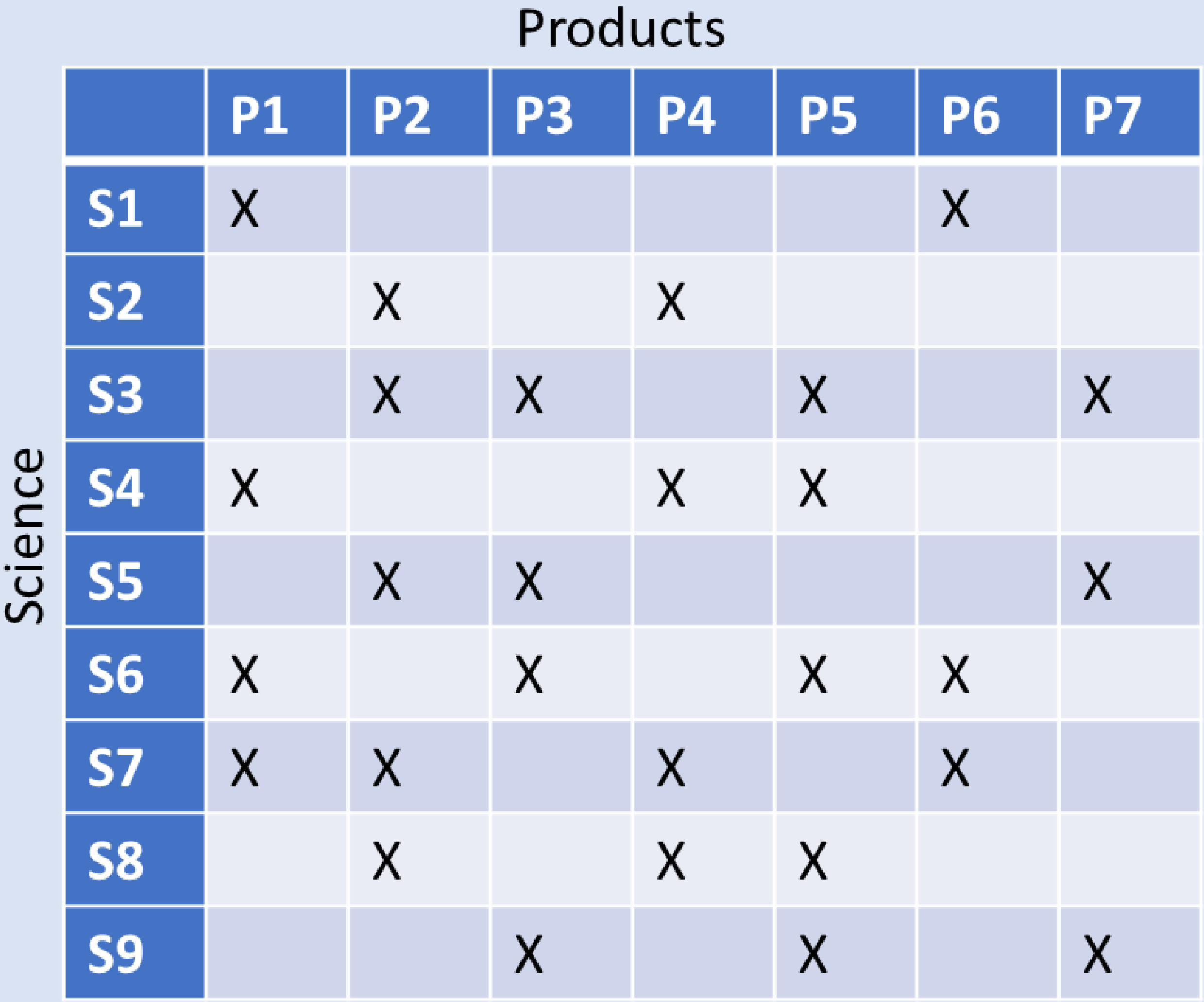
Formulation Scientist's Toolkit



Products

Science

	P1	P2	P3	P4	P5	P6	P7
S1	X					X	
S2		X		X			
S3		X	X		X		X
S4	X			X	X		
S5		X	X				X
S6	X		X		X	X	
S7	X	X		X		X	
S8		X		X	X		
S9			X		X		X

Prof Steven Abbott

The Formulation Scientist’s Toolkit

Prof Steven Abbott

Steven Abbott TCNF Ltd & University of Leeds

With thanks for inspiration and encouragement from:

- Prof Massimo Bonino University of Florence;
- Prof Gerardo Palazzo University of Bari;
- Dr Seishi Shimizu University of York.

Timeline:

- First demo: February 2023
- Alpha version: July 2023
- Beta release: September 2023
- v1.0: 16 October 2023
- v1.01: 30 October 2023 – minor tweaks
- v1.02: 29 January 2024 – a few updates/corrections

Introduction

When you have to fix something around the house, you often find that a “simple” job turns out to be hard because it requires a specific tool you don’t have in your toolkit. You go out and buy the right tool, finish the job … and maybe the tool then sits at the bottom of the toolkit, unused … till the day you have a similar job. It takes a few minutes to find it, but then the job is easily finished.

Formulation science isn’t so different. Many problems can be solved quickly if you have the right scientific tool. But formulation science is very broad and none of us can know every relevant bit of science or have the time to search in the literature for a tool that might help.

That’s why I’ve created this Toolkit. There are over a hundred scientific tools ready for you to use. They are described in the Science chapters. And there’s a set of representative Products that each use relevant Science chapters. But there’s a problem – how do you find what you need; what links everything up? Behind the scenes there’s the **FST Matrix** made from Product columns and Science rows, with an X showing intersections of interest – the cover gives you an idea. Using the concept of a “self-assembling book” the matrix ensures that the Science is linked to the Products and the Products are linked to the Science. Within the text, these inter-Science and inter-Product links also take you parts of the book you want to explore. All the (1000+) links are auto-checked so you can click with confidence. The text also has links to relevant external apps that open in your browser, so you can explore a wide formulation space with a few clicks.

Navigation

In a PDF there aren’t the familiar browser-style navigation tools. But Alt-Left-Arrow takes you back to where you were, and Alt-Right-Arrow takes you forward. However, unlike a browser, the PDF lets you open the bookmark or page views to help you navigate. And there is always Ctrl-F to search for what you are looking for.

Each time you go to a new Science chapter you find a set of hyperlinks to those Product chapters that use the same science. These might suggest alternative ways of thinking about your own issues. And each Product chapter has its own set of hyperlinks to Science chapters. This double hyperlinking is an important feature of the FST. Science is universal, so seeing your own types of problems in different contexts can often bring you fresh insights. Something which might be common knowledge in industry A might be almost unknown in industry B, even though the core science is the same.

By widening your knowledge of how your core science applies to other products, you know where to look for fresh knowledge. You also become more flexible and employable.

The FST Community

Welcome! You are part of the community. You will quickly find things you don't like, things missing, things you'd like changed. No problem. Let me know. Behind the scenes the self-assembling book goes from matrix to final PDF. So missing bits of science, missing links, new Science or Product topics, new matrix links, or just errors in the text - at the press of a few buttons a new version of the book can be created and uploaded to the FST site.

A Toolkit this big is bound to have repetition, omissions, biases and errors, so don't hesitate to point out any faults you find.

Send comments, suggestions, corrections to me at steven@stevenabbott.co.uk.

My thanks to Prof J-M Aubry who suggested numerous improvements from the beta release and helpful corrections/additions for v1.02.

AI and the future of the FST

ChatGPT, Bard and similar tools emerged while I was writing the FST. The question immediately arose: is the FST unnecessary, are the AIs our Toolkits? So I've put plenty of time into learning how to ask suitable questions via GPT4 + plugins (the most powerful combination of tools at the time of writing). The FST has been improved through this process because GPT4 is more methodological (and sometimes tedious) in answering interesting questions, so has spotted some details that I'd not thought of and was happy to add. I've also asked it some tricky detailed questions about issues that cross conventional boundaries, and got impressive replies. But I could only ask those questions because I knew the answers, and knew why they were important to formulators. I've not found a way for it to generate the sort of content that I think formulators need beyond their routine training.

Will the FST become irrelevant as the AIs get better? Of course! Till then, I'll carry on doing my best to improve the depth and breadth of the content. After all, the AIs will be reading the FST so in a small way I'm helping them get smarter.

Abbreviations

The list is not exhaustive.

- AI Artificial Intelligence

- APG Alkyl Polyglycoside
- API Active Pharmaceutical Ingredient
- BP Boiling Point
- CAPB Cocamidopropyl betaine
- Cc The “characteristic” value of a surfactant in HLD theory
- DoE Design of Experiments
- EACN The Equivalent Alkane Carbon number in HLD theory
- HLD Hydrophilic Lipophilic Difference
- HSP Hansen Solubility Parameters
- Mc Critical Entanglement Molecular Weight
- ML Machine Learning
- MVol Molar Volume
- MW Molecular Weight
- PCA Principal Component Analysis
- PE/PP Polyethylene/Polypropylene
- PEG/PPG Same as PEO/PPO
- PEO/PPO Polyethyleneoxide/Polypropyleneoxide
- PSA Pressure Sensitive Adhesive
- PVOH Polyvinylalcohol
- SLS Sodium Lauryl Sulfate; SLES Sodium Laureth Sulfate (extra EOs)
- Tg Glass transition temperature

Contents

- The Products -

You can explore the FST via the product categories. Your specific products might not feature but the idea is to take inspiration from those product classes that are close to yours.

Experience shows that exploring outside your usual product domain is a great way to find fresh inspiration. In too many formulation domains, everyone else is working with the same, tired ideas. If that applies to you, splash out and learn the ideas that are common in domains that might interest you.

It also helps with career advancement and employability to have knowledge of a broader set of products and the core science elements that connect them.

Adhesive

Links

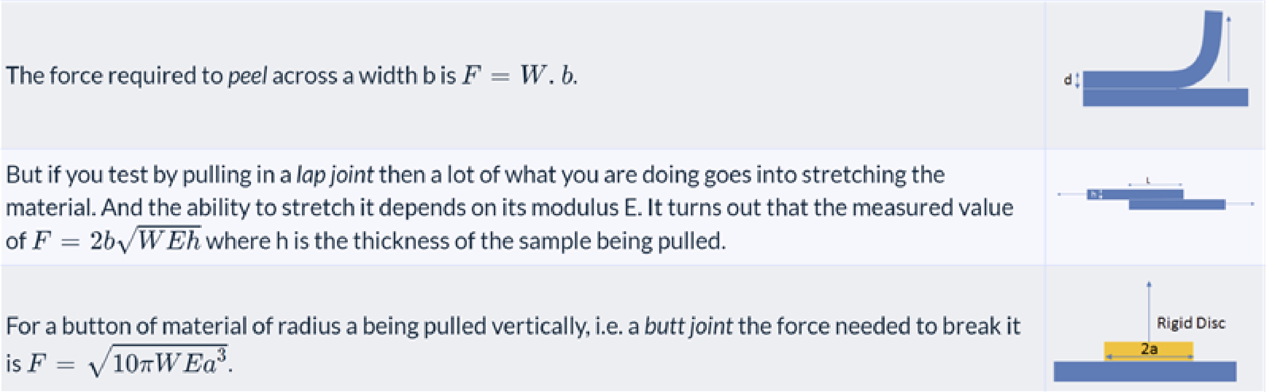
Adhesion_Adhesion promoters, Adhesion_Butt, Adhesion_Crack resistance, Adhesion_Crosslinking, Adhesion_Entanglement, Adhesion_Intermingling, Adhesion_JKR, Adhesion_Peel, Adhesion_Shear, Adhesion_Surface energy Interactions, Adhesion_Testing, Diffusion_Basic Diffusion, Dissolution_Flory-Huggins, Dissolution_Hansen Solubility Parameters, Evaporation_Basics, Evaporation_Diffusion limited, Flow_G' and G", Flow_Oscillatory rheology, Flow_Polymer Viscosity, Flow_Rotational rheology, Flow_Shear dependent, Flow_Stefans Squeeze, Flow_TTS-WLF, Flow_Thixotropy, Flow_Yield Stress, Mechanical_Surface Roughness

> Making a good adhesive is hard. Look at the list of things you need to know to be able to formulate well. But it is even harder if the formulation team aren't familiar with the core science behind a good adhesive. Those working on Water-based Adhesive and PSA have specific challenges discussed in those separate chapters.

Adhesion is a Property of the System

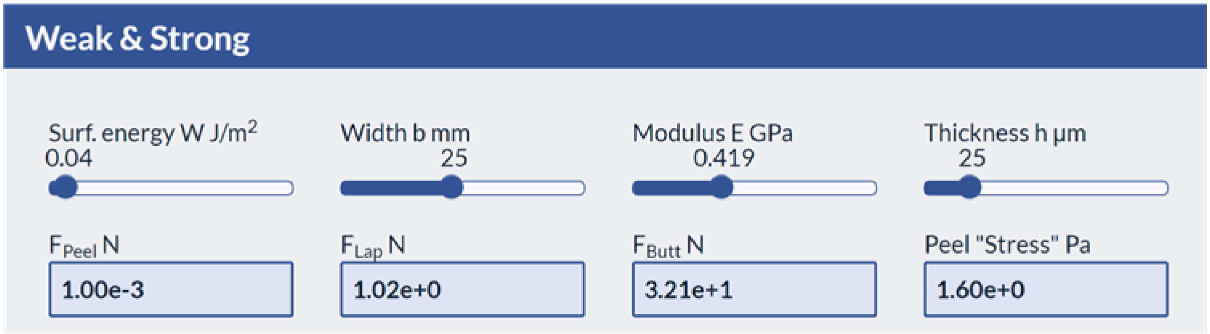
This phrase is key to understanding adhesion. The best adhesive in the wrong system is useless. A mediocre adhesive in the right system will be fine for the application.

Here are two examples where "the system" has a big influence. The first compares the adhesion (via pure surface energy) of two pieces of super-smooth rubber, tested in peel, shear and butt modes:



The force required to *peel* across a width b is $F = W.b$.	
But if you test by pulling in a *lap joint* then a lot of what you are doing goes into stretching the material. And the ability to stretch it depends on its modulus E. It turns out that the measured value of $F = 2b\sqrt{WEh}$ where h is the thickness of the sample being pulled.	
For a button of material of radius a being pulled vertically, i.e. a *butt joint* the force needed to break it is $F = \sqrt{10\pi W E a^3}$.	

The app does the calculations for you:



Weak & Strong

Surf. energy W J/m^2 0.04

Width b mm 25

Modulus E GPa 0.419

Thickness h µm 25

F_{Peel} N: 1.00e-3

F_{Lap} N: 1.02e+0

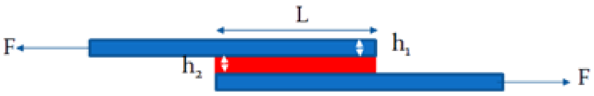
F_{Butt} N: 3.21e+1

Peel "Stress" Pa: 1.60e+0

https://www.stevenabbott.co.uk/practical-adhesion/weak-strong.php

In the case of peel, the force needed to break the bond is 1 mN. Pulling the same system apart under shear requires 1 N – 1000x times more even though the same surface energy is involved. Pulling apart as a butt joint requires 32 N, 32,000 times more.

More interesting is that although (of course) the peel force depends only on the surface energy, the shear and butt joints depend on the modulus and thickness of the polymer, neither of which is an “adhesive” property.

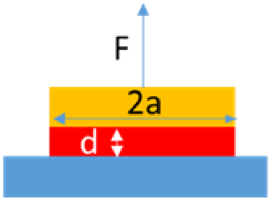


Another example is discussed more fully in Adhesion_Shear. The “lap shear joint” test has two adherends stuck together via an adhesive with a length L of overlap.

As you can see, a force F is applied to each end to induce shear in the joint. It is obvious that the joint will (eventually) fail in shear (it’s called a shear test!) and that the more adhesive, the longer L, the stronger the bond (it’s a *lap* shear test!).

Both intuitions are wrong. Adhesion is a property of the system, and in this system the joints generally fail in peel and after a minimum L, the overlap makes no difference to the strength of the joint.

Which brings us to Adhesion_Testing in general. Many standard adhesion tests give numbers that are irrelevant for the specific application. So when you test for adhesion you need to think of your system and find tests that are relevant to that system. If you put a carbon composite into a standard lap shear test it fails catastrophically. Yet carbon composites are wonderfully strong in shear. They fail because the “shear” test is, in fact, a peel test, which carbon composites are known to fail easily. For the true shear behaviour of a carbon composite you need to use the “double overlap test” which is less convenient and far more informative.



A final example of system thinking is connected to the notes on adhesive packages telling you to apply only a thin coating. Why would those selling adhesives want us to use less of it?

When you pull up on this butt joint (this time with adhesive, the one above was pure surface energy), the force, F, needed to break the joint of area A, for an adhesive with modulus E and Poisson ratio ν depends on the bulk modulus $K = \frac{3E}{1-2\nu}$, the (dissipative) work of adhesion W and adhesive thickness d, given by:

$$F = A\sqrt{\frac{KW}{d}}$$

The thinner the adhesive, the stronger the bond.

Why did I say "(dissipative) work of adhesion"? Think about it. Most adhesives have similar surface energies, so if W was a classic surface energy work of adhesion, we could use just about any adhesive. But a good adhesive, as we shall see, is able to dissipate crack energy, so it really matters that you have an adhesive suited to the types of crack forces that joint will receive during use.

Handling the adhesive

Whether the adhesive is a solution or a self-soluble mix of two components, the adhesive has to be handled from production to application and the drying/curing process itself. At the very least you need three sets of information:

- Flow_Rotational rheology,
- Flow_Shear dependent.
- Flow_Polymer Viscosity,

The basic rheology gives core information about the flow properties. The shear dependent data give the behaviour during different parts of the process. Pouring from a pot might have a shear rate of 1/s but squeezing a thin film between adherends might be 1000/s. And the sorts of polymers that are likely to give strong adhesion because of entanglement (see below) are automatically going to have a rapid increase in viscosity at concentrations above a critical value. Something as simple as a desirable reduction of solvent by 10% might lead to a catastrophic increase in viscosity.

Each property changes with temperature and with drying/curing in use. Without a core set of such data, it is easy for an adhesive to be unusable at some critical stage in its use.

Things that aren't important for adhesion

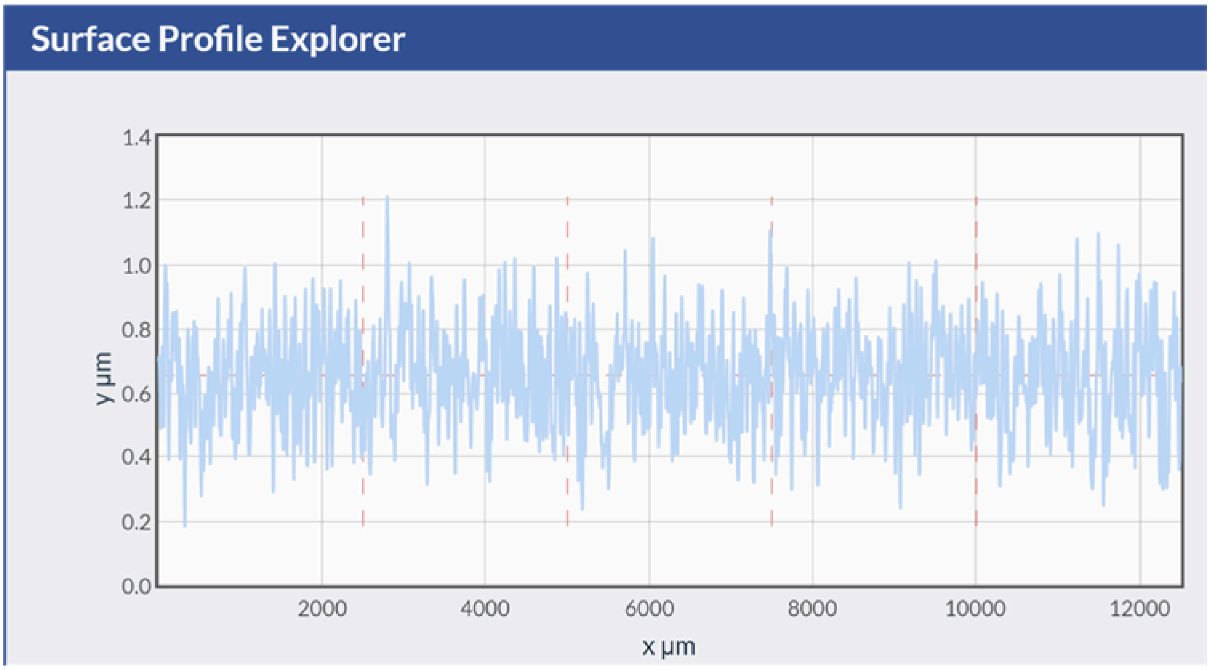
Surface energy is 10000x too small to be of relevance. A reasonable surface energy is 40 mJ/m² while a reasonable pressure sensitive adhesive is 400 J/m² - and that's not even a "strong" adhesive. It follows that measuring surface energies is a waste of time. The idea that "if it doesn't wet it can't stick" is also wrong. It is trivially shown that all liquids wet all surfaces (so water wets Teflon) when sandwiched between two surfaces. In the diagram the liquid has a 90° contact angle so "doesn't wet", yet it nicely wets between two surfaces of the same material.

> People ask why I get so angry about surface energies. It's because 50 years of adhesion science have been wasted by people obsessing about measuring them, rather than thinking about what's really going on. The day I realised, after many hours of surface energy measurements, that not once in my professional career had they ever been of use to me, was the day I started my search for ideas that really *do* help solve adhesion problems

The one good reason for (usually) quick measures of surface energy is a test for obvious surface contamination. Everything discussed in this chapter assumes that you have A and B when you are sticking A to B. Surface contamination would give not A-B but A-X-B. So if some drops of water+ethanol, or some dynes pens can tell you day after day that your surface is the same as before, that's fine. And if today it's lower then that's a signal that something has changed, which is usually a bad thing. But if the surface energy is *higher* you don't celebrate – you realise that something has changed, which is usually a bad thing.

If corona/plasma is your preferred way to remove contamination from the surface, and if surface energy is good at showing the removal, that's great. Adhesion is a property of the system and any system with junk that hasn't been removed is unlikely to give good adhesion. Focus on junk removal, not surface energy.

Surface roughness does not, and cannot, help adhesion via "extra surface area" and "mechanical interlocking". Although a surface profile might look like this:



https://www.stevenabbott.co.uk/practical-coatings/surface-profile-explorer.php

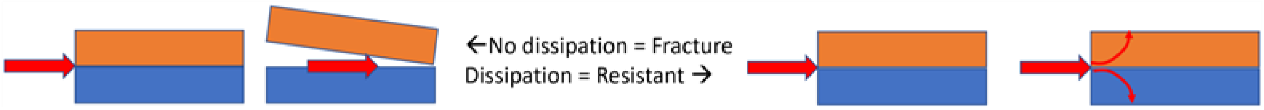
when you explore this diagram in Mechanical_Surface Roughness you find that the extra surface area (compared to a smooth surface) is 0 and that the chances of mechanical interlocking are similarly 0. Both ideas are an illusion based on mixed axis scales for x and y.

Old ideas of electrostatics have proven to be unhelpful.

Chemical bonds as a sole strategy are a failure. If you have complete chemical bonding you gain 1 J/m², tiny compared to our 400 J/m² earlier, see https://www.stevenabbott.co.uk/practical-adhesion/chemical.php. A *low* level of chemical bonds is a potent source of strong adhesion, for reasons explained shortly. Yes, too few chemical bonds might give poor adhesion, but too many (> ~ 1%) make things worse.

Dissipation is key

Adhesion fails when a crack propagates along the interface. Although you can easily *start* a crack, if it doesn't propagate, the adhesion is fine. If the energy in the crack is dissipated into the system then the crack won't propagate:

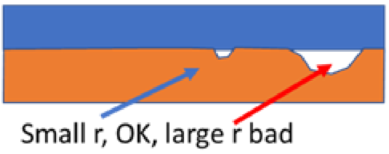


A pure surface energy bond has no way to dissipate crack energy. As found from hysteresis in Adhesion_JKR JKR experiments, intermingling of polymer across a boundary, Adhesion_Intermingling, can help dissipate energy via friction. A lot of strong adhesion comes from Adhesion_Entanglement where polymer chains are so entangled that if you try to pull them apart they cause motion over many nm which in turn dissipates energy. In these cases, the generic idea of viscous dissipation is helpful. This is particularly relevant to the crack resistance of PSA systems.

To get a full picture of your cured adhesive, you need to do some basic Flow_Oscillatory rheology, to measure Flow_G' and G". Dissipation is via G" (the loss modulus), but for a strong adhesive you need a large G', the elastic modulus. Getting the balance right is hard if you use rheology and impossible if you don't.

Thinking seriously about how your adhesive system might or might not dissipate crack energy is a productive process. Because of Flow_TTS-WLF a system that works well at, say, room temperature and "normal" test speeds might fail easily at low temperatures or high test speeds because dissipation can no longer take place.

Removing your own cracks

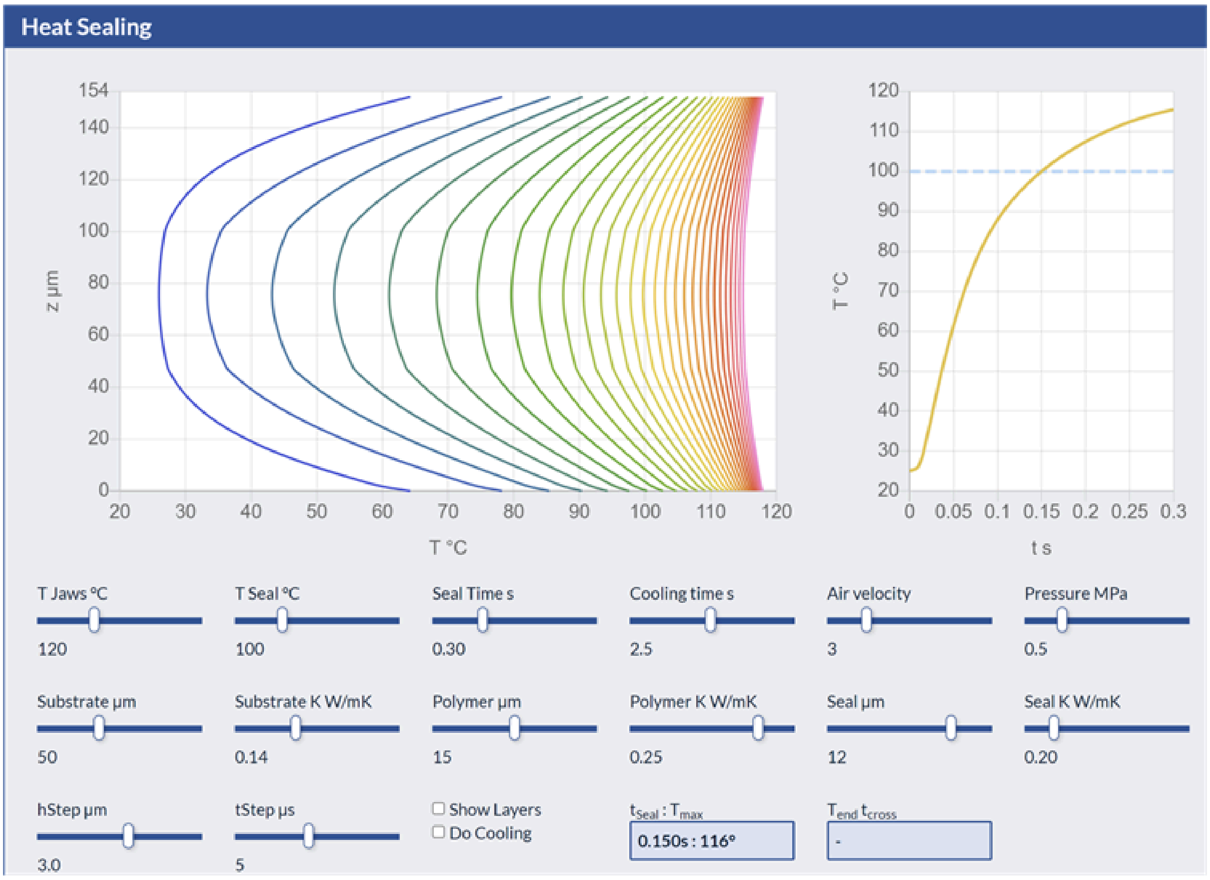


Another way to increase the strength of your adhesive system is to remove anything that provides an excuse for a crack to form and propagate. As explained in Adhesion_Crack resistance, the stress needed to open an existing crack of radius r is proportional to $\frac{1}{\sqrt{r}}$. If your joint contains super-small defects then maybe you will be OK, but a larger defect from, say, an air bubble can be devastating. Putting thought into removing air bubbles (or lumps of general dirt) can often give you more adhesion than trying to improve the formulation of the adhesive itself.

Reaching across the interface

In order to dissipate energy across an interface there must be ways for molecules of adhesive and adherend to cross the interface.

The simplest way to achieve this is a PolymerA-PolymerA interface created by heat sealing. Heat the interface above the MPt of the polymer and keep the surfaces in contact for 0.5s and on rapid cooling without accidentally pulling the two surface apart, you have effectively a solid A polymer. It is self-entangled, with the bond approximately as strong as the polymer itself. You can explore the thermal effects via the app:



T Jaws °C 120

T Seal °C 100

Seal Time s 0.30

Cooling time s 2.5

Air velocity 3

Pressure MPa 0.5

Substrate µm 50

Substrate K W/mK 0.14

Polymer µm 15

Polymer K W/mK 0.25

Seal µm 12

Seal K W/mK 0.20

hStep µm 3.0

tStep µs 5

☐ Show Layers

☐ Do Cooling

t_{Seal} : T_{max} 0.150s : 116°

T_{end} t_{cross} -

https://www.stevenabbott.co.uk/practical-mechanical/Heat-Seal.php

> "It is self-entangled, with the bond approximately as strong as the polymer itself." That's true for heat sealing. When it comes to 3D printing with hot, extruded, polymers there is an unfortunate trade-off. At lower temperatures, there isn't enough heat & time to allow the polymer chains to entangle, so adhesion is low. At higher temperatures, there's plenty of entanglement, but the whole structure sags and deforms under its own weight. It turns out that the temperatures used in practice give only 1 or 2 entanglement lengths at the interface, too few to give strong adhesion. These 3D printed structures easily fracture along the interlayer-planes – they can't dissipate enough crack energy.

If you can't use heat, you can use solvents. Typically the adhesive will be in a solvent chosen for good compatibility, so you can get a high concentration of adhesive. If the solvent cannot attack the adherend surface then no entanglement can take place. That is why many adhesives have a solvent optimized for a range of compatible polymers (typically excluding polyethylene and polypropylene), where there is sufficient "bite" into the adherend polymer surface without destroying it.

Choosing the right solvent for mutual compatibility can be done rationally via Dissolution_Hansen Solubility Parameters. Getting the right evaporation rate is described in Evaporation_Basics and minimizing the amount of trapped solvent is discussed in Evaporation_Diffusion limited.

If you can't get polymers across an interface (for example, when sticking to a metal) then you can resort to chemical bonds. As with everything in life, too much of a good thing is a bad thing. It turns out that if you have a *lot* of chemical bonds, adhesion is poor. Why? Because the interface is so brittle (think of glass) that a crack easily travels along it. The chemical bonding done correctly at a low level is there to deliver entanglement into the polymeric adhesive. No matter whether the bonding is between groups naturally on the surface or those added with Adhesion_Adhesion promoters, you just need enough bonding to create plenty of tangles. It's a common observation that when you try especially hard to increase the level of adhesion promoter, or heat the system to encourage more bonding, the adhesion gets worse.

Adhesive mechanical properties

Assuming the interface is good, what else is needed from the adhesive? That depends on the system.

For many *structural* adhesives, a key requirement is strength, i.e. a high Mechanical_Modulus. This will be especially useful in guarding against Flow_Creep and Relaxation, the slow distortion of the bond with prolonged loads. Often these adhesives will be required to keep those good mechanical properties over a large temperature range. If the temperature range is large then stresses from mismatch of thermal expansion coefficients of adhesive and adherend can be a problem. The more rigid the adhesive, the more critical it is to have a good match of expansion coefficients as there is no ability of the adhesive to relax under local thermal stresses.

So structural adhesives are the highly crosslinked systems such as epoxies or UV. Achieving sufficient Adhesion_Crosslinking is difficult because as the crosslink density increases it becomes harder for subsequent links to be formed. Shrinkage during crosslinking risks damaging the bond – the low shrinkage of epoxies makes them especially desirable.

Conversely, many adhesives are better when they are relatively low modulus and flexible. For example, wood expands and contracts drastically with temperature and humidity. So an adhesive that is able to stretch and shrink is often preferable. Similarly, you do not want a rigid adhesive for shoes, handbags and many autoparts where light weight and flexibility are part of the design.

Some years ago I was researching optimal wood adhesives. I kept finding papers using sophisticated structural adhesives, while observing that no one was using them. The PVA used in common wood glues is notable for being a relatively poor polymer, especially compared to the wonder polymers used in all those academic papers that everyone was ignoring. That's when I found the true explanation – as wood expands/shrinks with humidity, low-strength PVA can flow to accommodate the changes. The joints made from wizzo adhesives break under the stresses.

Once again, because of Flow_TTS-WLF, any single measure of mechanical properties is meaningless – you have to know the temperature and time-scale of any likely threat to the adhesive.

Adhesive chemical properties

For non-crosslinked adhesives, resistance to solvents can be straightforwardly estimated via Dissolution_Hansen Solubility Parameters. For crosslinked adhesives the problem isn't one of dissolution but rather of swelling. The Absorbency_Swelling guide introduces you to Flory-Rehner theory which depends on the Flory-Huggins χ parameter which links directly to Hansen Solubility Parameters.

Resistance to aggressive chemicals is a more specialist topic and you probably know a lot already for your specific system.

The real challenge for everyone is water and humidity. Water is a small molecule available in large quantities from a humid atmosphere so it can diffuse into the adhesive, where it might cause some weakening of structure (e.g. by plasticisation). See Diffusion_Permeation OTR and WVTR. But more serious is diffusion to the interface. Relatively small changes to the polymer(s) at the interface can have a significant effect on local stresses, relaxation times and viscous dissipation, with potentially significant reduction in adhesion. If you are relying on relatively few chemical bonds (too many are bad in other ways) and if they are susceptible to hydrolysis then you have big problems.

So make sure you have the minimum of spare -OH and other polar groups at the interface, and if you are using chemical bonds across the interface, try to make them multi-functional. Maleic acid and, especially, trialkoxysilanes, are favourite adhesion promoters because their multifunctionality gives them extra resilience against hydrolysis.

Stefan's Squeeze

An under-appreciated difficulty with adhesives is getting the right, even, thin layer of adhesive across the whole joint. It seems easy – add a central blob with a slight excess of adhesive, squeeze, and the job is done.

Unfortunately, Flow_Stefans Squeeze tells us that it is hard to get things right. If you feel the urge to jump off a bridge with a bungee cord attached to a lump of steel stuck to another lump of steel with 1 large drop of superglue – don't do it. You will die. But if you have the same volume of superglue supplied as 10 drops, Stefan is on your side and you will have a satisfying dive and rebound.

Finding the right pattern of drops of adhesive for the best uniformity with zero chance of trapping air bubbles is a scientifically unsolved puzzle. The closest I can find is a way to think about dissolving the air bubble during the pressure of the squeeze: https://www.stevenabbott.co.uk/practical-adhesion/Bubble-Gone.php.

Short summary

Now you know that Adhesion is a Property of the System and have seen examples of how our intuitions about the system can often be wrong, you have some core ideas for tackling your own adhesive issues. We've given you no direct advice about your specific adhesive polymers, crosslinkers, promoters, surface preparation. Each of these is super important, but this isn't a guide to formulation *specifics*, it's a guide to formulation *principles*. This is especially important for adhesives where historically there has been too much focus on the specifics, without stepping back and seeing the whole problem. If it turns out that you can fix an adhesion problem by changing, say, the modulus of the adherend or removing crack-inducing air bubbles, that saves a lot of time which might have been wasted trying to improve an already good formulation.

> I once helped fix an adhesion problem without changing the adhesive or the adherends. Instead, a polymer layer further away from the site of adhesion failure was to blame. When that remote layer was LDPE, the standard "tear test" deformed the whole structure and stresses got concentrated at the adhesive interface. When it was HDPE, the same test passed with no problem – stresses went where they were meant to, and the adhesive was fine. I have to say it again – Adhesion is a property of the system.

Deodorant Sticks

Links

Dissolution_Hansen Solubility Parameters, Dissolution_Solubilizers and Hydrotropes, Flow_Oscillatory rheology, Flow_Shear dependent, Flow_Yield Stress, Fragrance_Activity Coefficients, Fragrance_Vapour Pressure, Gelling_Networks and Percolation, Gelling_Syneresis, Optics_Scattering and Opacity

We have two formulation decisions:

- The type of deodorant package to use
- The matrix for the deodorant package

Because antiperspirants (stopping sweat from reaching the surface of the skin) raise too many regulatory complexities, they are not included in the discussions.

The Deodorant package

The simplest is a general-purpose pleasant fragrance, along with any marketing claims about the fragrance being extracted from or plausibly related to some suitably exotic and natural source.

More sophisticated is the recognition that many fine perfumes include malodorous molecules such as scatole. Somehow a hint of malodour can be converted into a desirable perfume with the right combination. The distinctive underarm odours are specific hexenoic, octanoic and hydroxyhexanoic acids plus sulfanylalkanols. Amazingly a specific combination of terpene-derived esters works with the malodours to create a pleasant fragrance.

An attractive semi-natural ingredient is zinc ricinolate, based on natural ricinoleic acid. Although other zinc ingredients work adequately, the ricinolate seems especially good at intercepting the carboxylic acids and the sulfurous molecules and removing the odours. The reasons are unclear, but the efficacy is impressive.

The sweat itself is odourless. The malodorous molecules are produced by enzymes breaking down some natural (odourless) glutamides into glutamine and the offending carboxylic acids. If your fragrance package includes glutamides that are broken down to *odourless* molecules, you can overwhelm the enzymes so they don’t touch the odour-producing one. Or you can add gluatamide mimics that bind to and block the enzymes.

So far there is no controversy. If your matrix is water based then you must add some preservatives. Although these are designed to inhibit microbes the rules of the game say that these are preservatives and not antimicrobial. But if you go one logical step to add antimicrobials to the deodorant to zap the skin microbes that produce the enzymes that cause the odours then you are in difficult territory. You are potentially offending two very different groups:

1. The regulatory authorities that require proof of efficacy, safety and absence of build-up of resistance to the antimicrobials
2. The users who are increasingly aware that their skin biome is there not just to create malodours –

it should be a healthy part of their overall defence system against threats to the outside of the skin.

Using potassium alum, “natural extracts” and maybe chelators such as a salt of EDTA or GLDA may merely upset some microbes (e.g. the chelators starve them of iron) so they aren’t really antimicrobials. And maybe they don’t upset the skin biome too much. And maybe the good microbes have already been eliminated by excess washing with powerful soaps, so zapping the bad microbes is reasonable. Sorting out these issues is an interesting balance of science and marketing claims.

It is scientifically crazy that we damage our skin biomes with too much hot, soapy washing then try to fix the odours which might well be caused by excess “bad microbes”, using an array of chemicals which might themselves further disrupt the biome. This leads to a formulation bifurcation. One way is to add pre-/pro-biotics to the soaps and deodorants, the other way is to formulate minimalist soap/deodorant packages.

The matrix

We have to choose between water-based and oil-based

Water-based

With water as a low-cost evaporative component along with propylene glycol or dipropylene glycol as the (effectively) non-volatile carriers of the relatively small % of chemicals in the deodorant package, our key need is to gel the system. With “sodium stearate” (actually a mix of different alkane chain lengths, giving a range of properties) we have, essentially, cheap soap holding everything together. Minor ingredients such as chelators, preservatives and antioxidants can easily be incorporated. The problem is that the fragrances are generally hydrophobic. So how do we compatibilise the fragrance with the hydrophilic portion?

First, why do we want to make them compatible?

- Because incompatibility leads to Mie scattering (Optics_Scattering and Opacity) and consumers like clear gels which suggest no possibility of staining of clothes
- Because incompatible fragrances have a high activity coefficient (Fragrance_Activity Coefficients) and will evaporate too quickly, losing claims of long-lasting action
- Because incompatibilities can lead to syneresis of the gel – with fragrances or water being eliminated from the gel during storage, producing an undesirable surface liquid and a poor consumer experience. Although the Gelling_Syneresis chapter concentrates on particle gels, the general background to syneresis is still valid.

It would be nice to say (given how they feature in many parts of the FST) that the problem of compatibility could be fixed by suitably matching the Dissolution_Hansen Solubility Parameters of the fragrances and the gel. Indeed, some of the relevant product literature invokes HSP. But that can’t be true. HSP and other solubility approaches are “mean field” theories, meaning an absence of long-range order. We can be certain that gelling agents such as sodium stearate are acting with plenty of long-range order. As mentioned above, they are merely soap and although there isn’t a clear water/oil distinction relevant to normal soap use, there has to be some aspect of head/tail ordering to produce the gelation. Although it is possible to claim that the popular long chain dimethylamide additives have an HSP that matches some fragrances, it seems more plausible that the long chains have interesting interactions with the various long chains in the “sodium stearate" creating, for example, local pockets of lower crystallinity that allow the fragrance molecules to escape the hostile aqueous environment.

Instead we have to invoke the ideas of "solubilizers" Dissolution_Solubilizers and Hydrotropes. Unfortunately, these have been subject to decades of fruitless hand waving rather than solid science. Although the newer Dissolution_Kirkwood-Buff statistical thermodynamic approaches can work well to describe solubilization, we have no predictive framework to invoke for solving problems like these.

How do we proceed under these circumstances? By being intelligent and inquisitive. We don't know how these solubilizers might work, but we can be certain that they will impact the gelation behaviour – this is certain because they are interfering with long-range order. So we need to be ready to look for patterns of desirable increases of solubility of the fragrances (greater optical clarity) while attending to changes in gelation characteristics. Which, as so often in the FST takes us to rheology, discussed below.

Oil based

Such formulations eliminate the need for preservatives and avoid fragrance incompatibility issues. We use any of a myriad of pleasant, light plant oils, some waxes for strength and cetearyl alcohol as a thickening and "glide" agent. Zinc ricinolate fits in easily along with antioxidants such as tocopherol that consumers love to see. If you don't like to add fragrances (which imply some artificiality) you can add essential oils instead. They are fragrances, but you don't have to call them that.

The first downside of these formulations is that the stick is opaque, which points towards the second downside that all this oily gunk can create a white stain your clothes, and that the stain is relatively hard to remove compared to the water-soluble aqueous formulation ingredients.

Feeling good

The different chain lengths in the different "sodium stearates" give different levels of crystallinity and different responses to temperature during storage and during application to the warm skin surface. The same applies to the different waxes (beeswax, carnauba, shea butter etc.) and oils for the non-aqueous products. The formulations must have a Flow_Yield Stress behaviour nicely tuned to feeling good during application – not too stiff, not too runny, not applying too much or too little.

Clearly this needs the help of considerable amounts of smart rheology. Equally clearly, there is no way to go directly from rheology to a successful formulation. You need to build up a smart map of how the different formulation ingredients and rheological measurements map on to the product performance. This needs measurements of good and bad products, internal and competitor materials.

How to do this is described in the Lipstick chapter. Although lipstick and deodorant sticks are different products, science is universal and the problems of cost, marketing claims, ingredient fads, rheology measurements, skin feel, mapping across complexity are so similar that it's a good idea, not mere authorial laziness, to send the reader to lipstick.

Dishwashing Liquids

Links

Cleaning_Boundary removal, Cleaning_Contact Angles, Cleaning_Surfactants, Cleaning_Temperature effects, Flow_Basic viscosity, Surfactancy_Anti-foaming, Surfactancy_CMC and Langmuir, Surfactancy_CPP and phases, Surfactancy_Dynamic Surface Tension, Surfactancy_Foam Drainage, Surfactancy_Foam Ostwald, Surfactancy_Foam Rheology, Surfactancy_Foaming, Surfactancy_HLD, Surfactancy_Roll-up and Eötvös Number

The official test for a dishwashing liquid is *not* how well it washes dishes, but how many greasy plates can be washed before the foam created at the start of the test finally disappears. Plenty of formulations superior at washing dishes are therefore rejected in favour of foamy formulations.

So although this chapter will explain how to get the best formulation for washing, to be relevant to the real world it has to focus on foaming.

Foaming

The science of Surfactancy_Foaming is remarkably simple. Take any reasonable surfactant solution, mix air with it, then put in enough energy to break down large bubbles to small ones … and you get an impressive, long-lasting foam. More than good enough to pass a standard dish washing test. The physics of Surfactancy_Foam Rheology, Surfactancy_Foam Drainage and Surfactancy_Foam Ostwald converge on the fact that small bubbles have a high yield strength (a “strong” foam), slow drainage and a greater resistance to Ostwald ripening – bigger bubbles getting bigger at the expense of smaller one.

So why are so many dishwashing liquids based around SLES and CAPB? Because they happen to give relatively large volumes of relatively stable foam when tested under standard dish washing conditions. Also, because they are cheap. But they are cheap because they are used in vast quantities – the argument is a bit circular.

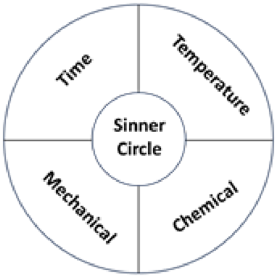
There are no satisfactory explanations for why a few surfactant combinations give foaming characteristics that are ideal for the standard dish washing market. There are plenty of ad hoc explanations, but a good explanation would let formulators predict newer, more interesting formulations … and we don’t have this.

Are the standard formulations good because they have good Surfactancy_Anti-foaming resistance? No. The science of anti-foaming remains relatively crude and there is no reason to assume that SLES/CAPB is particularly good, especially when it’s not clear how much the fats in the washing water present a classic anti-foam challenge. Anyone who has tried to add anti-foamers to reduce foaming will find (as explained in the anti-foam chapter) that a good anti-foam is hard to find.

The grease eventually wins because surfactant molecules end up wrapped around the fat globules. It is an interesting fact that SLES is not especially good at wrapping up fat globules (as we shall see, it is far too hydrophilic), so maybe the long-lasting nature of the foam is due to the fact that the surfactant isn’t doing a great job at cleaning.

It is very easy to be negative about marketing-driven formulations for consumer products. So it was great to have the chance to talk to an experienced European formulator making products for those large cut-price supermarkets who demand top performance at low cost. So everything unnecessary is eliminated – what matters is cost and performance. The feedback was clear. They could make a superior dishwashing product no problem, but because "performance" is defined by the foam test, they have no choice but to make technically inferior "foamy" products.

The Sinner circle



According to Dr Herbert Sinner of Henkel, you can clean with any combination of Temperature, Time, Mechanical Action and Chemicals. The **mechanical** action is necessary for Cleaning_Boundary removal which, thanks to the no-slip boundary condition is impossible without such action. Many dishes will clean themselves adequately if left to soak for a **time**. And **temperature** is good for dissolving soluble residue on the dishes as well as helping to melt off fats and grease. So what are the distinctive **chemical** aspects to the cleaning (rather than foaming) roles of the surfactants?

Reducing surface tension?

There are those who like measuring Cleaning_Contact Angles and are concerned about Surfactancy_CMC and Langmuir in order to get the lowest contact angle for the least amount of surfactant. They tell us that wetting and cleaning are strongly related. But they aren't. A dish in the washing bowl is fully wetted. When scrubbing with the brush, the power of the washing action overwhelms relatively weak surface tension forces.

The time we need a small-ish contact angle is during the final rinse when we would like the water to drain off evenly. It's not often stated, but this happens best if the hydrophobic tail of the surfactant is contaminating the dish (OK, it's only a chemical monolayer), leaving the hydrophilic head to let the rinsing water (generally fresh water, low in surfactants) drain as a nice sheet.

As has often been pointed out, if CMC was super important then we would all be using non-ionic surfactants which have CMCs far below those of the anionics that dominate the dish washing world.

What about Surfactancy_Dynamic Surface Tension. For dynamic processes, the static surface tension measured during CMC tests is irrelevant – fresh interface is created on the millisecond timescale and it can often take seconds before an equilibrium surface tension is reached. Interestingly, the secret to a good dynamic surface tension is a high CMC; a "bad" surfactant might actually be better. But as surface tension effects are of minimal importance for dish washing, we don't have to worry too much about dynamic surface tension effects.

Real cleaning

In a dish washer, or a washing machine, where foam is actively not wanted, the formulations can focus on real cleaning. Dish washers use plenty of caustic to attack fat and grease. Washing machines use an array of enzymes to chop up the lipids, proteins and starches on the clothes. We can't use caustic or enzymes in dish washing liquid. What *could* we use?

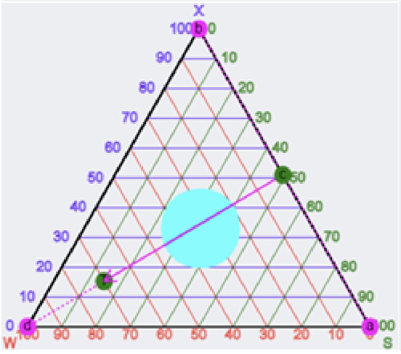
To get fats to “roll up” from a surface is straightforward if you formulate for a Eötvös number: Surfactancy_Roll-up and Eötvös Number. Translated, this means a low Surfactancy_Interfacial Tension and Rigidity which is trivially obtained by formulating so that the Hydrophilic Lipophilic Difference, Surfactancy_HLD is zero. When you explore HLD (which balances oiliness, surfactancy and temperature) you quickly find that a surfactant such as SLES is hopeless – it is far too hydrophilic. At best it can wrap around whatever fat blobs you can create, so does help to clean. But compared to the sorts of surfactants chosen for washing machines (where roll-up and therefore HLD balance is critical) dish washing formulations are sub-optimal.

You can buy *real* dish washing formulations that are effective rather than foamy, but they are not popular and cannot claim to be good, because the standard test is the foamability one.

Making the product

At first it sounds trivial. Add the different surfactants to water, mix a bit to create a concentrate and the product is ready to be delivered into the final containers either via dilution during the final step or by selling a concentrate to reduce the environmental footprint.

However, concentrated surfactants can form difficult phases, on their own or in mixtures. If these phases are hexagonal or cubic then they are semi-solid and difficult to disperse/dissolve. The problem can arise at any stage in dilution from concentrate to the final formula.



https://www.stevenabbott.co.uk/practical-surfactants/pde1.php

In this diagram, the concentrate (green dot at the right) is being diluted to the final form (green dot to the left) and passes through a cubic phase (cyan circle).

It’s important for this, and many other reasons, to be familiar with Surfactancy_Phase Diagrams.

I’d known that hexagonal or cubic phases were “viscous” but had imagined that claims for how difficult they were to handle, or the dangers of them appearing during dilution we being exaggerated … till on a lab visit I got a chance to compare an impressively solid hexagonal phase to the two liquid samples sitting next to it, one higher and one lower concentration. Mixing these two liquids produced the impossible-to-handle gel. There is a difference between “knowing” something and *knowing* something.

Doing something new

The pressures on you to produce low-cost, foamy products force you back to SLES-CAPB or whatever “sulfate-free” equivalent you can find to replace SLES. The pressures to go green require either genuinely greener surfactant types or clever schemes to offset the carbon footprint and environmental damage of your current type. But genuinely greener products don’t give the same foamy signals desired by the consumers and required to pass the irrational cleaning tests.

This forces you to adopt one of three novel strategies:

- Use honesty in marketing to explain to consumers that your formulation cleans better and greener, despite having few bubbles. Convert a bug (no foam) into a feature (Look! No Foam!!!)
- Find some novel dispensing device to get lots of bubbles from a formulation that doesn't foam well in the standard test.
- Find a magic additive that transforms your formulation from low foaming to high foaming

I once had to find an additive to transform a surfactant from no-foam to high-foam. A simple, high-throughput, small-volume test allowed us to try a wide variety of "co-foamers" – something impossible with slow, high-volume conventional tests. I tried 3 that seemed to me to be reasonable – with no effect. A colleague suggested a 4th. The test produced so much foam I thought I'd accidentally used SLES so carefully repeated the test, with the same amazing amount of foam. To our dismay we haven't found a good explanation for the effect. But it shows that such effects are possible – and that imaginative HT tests are a good idea.

Emulsion Paint

Links

Adhesion_Adhesion promoters, Adhesion_Crosslinking, Adhesion_Entanglement, Adhesion_Intermingling, Coating_Dewetting Theory, Coating_Levelling Theory, Coating_Pinholing Theory, Dispersions_DLVO, Dispersions_Zeta potential, Dissolution_Hansen Solubility Parameters, Evaporation_Basics, Flow_Polymer Viscosity, Humidity_Water mechanical isotherm, Humidity_Water vapor isotherm, Optics_Gloss, Optics_Scattering and Opacity, Surfactancy_Emulsion Inversion, Thickeners_Associative Thickeners, Thickeners_Polymeric Thickeners

The basics of paint have been discussed in Solvent-Based Paint. Here we focus on those special features that arise because the paint is water-based.

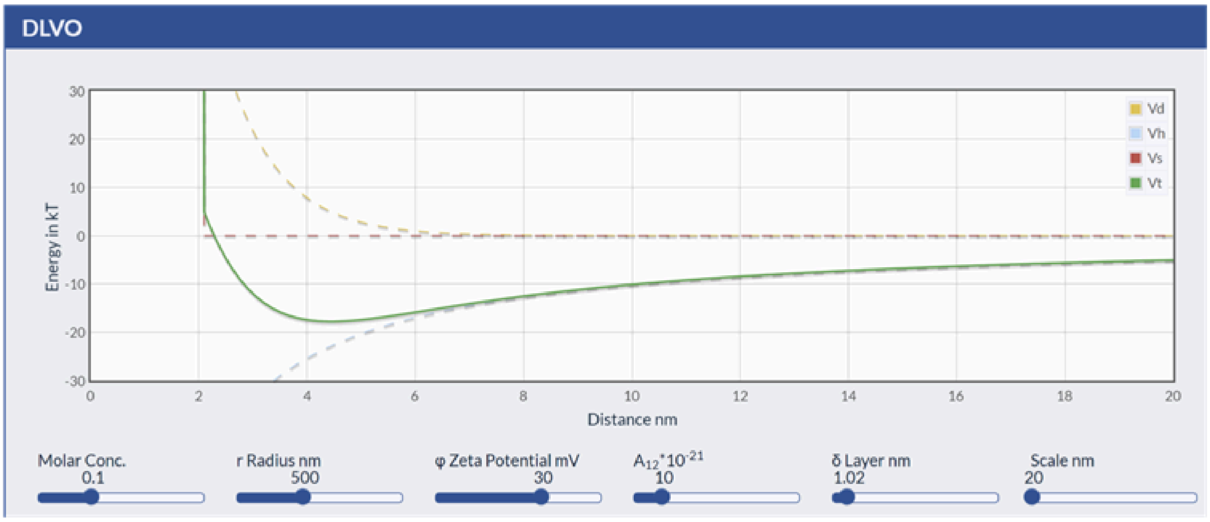
From dispersed to insoluble

For simplicity we will call the dispersed blobs of paint polymer emulsion drops, even those that aren’t strictly emulsions. By doing this we can use much of the world of emulsion science without bothering too much about the fine details.

The problem at the heart of these emulsion systems is that the individual drops must be infinitely stable when dilute, yet should form a continuous film of polymer once the water has (largely) disappeared. Each property on its own is not a problem – getting both right is the challenge.

A stable dispersion

For these aqueous systems, Dispersions_DLVO theory is adequate.



Molar Conc. 0.1
r Radius nm 500
φ Zeta Potential mV 30
$A_{12}*10^{-21}$ 10
δ Layer nm 1.02
Scale nm 20

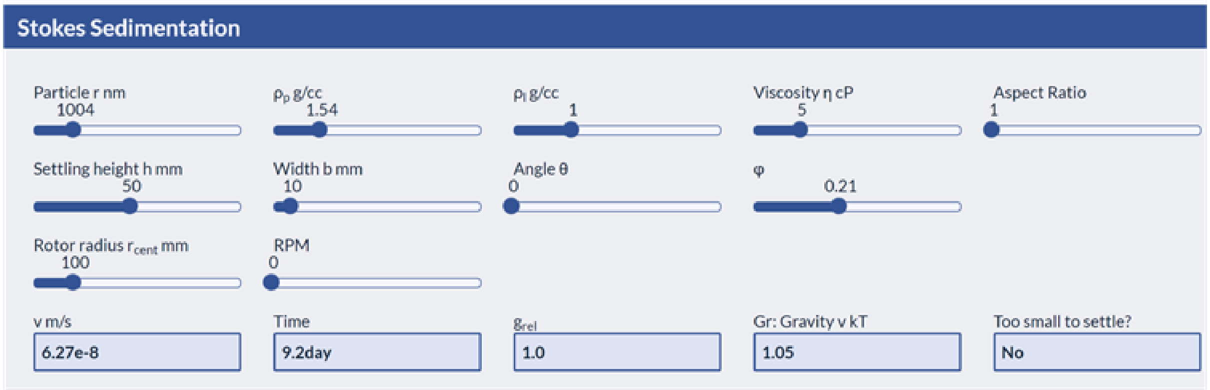
https://www.stevenabbott.co.uk/practical-surfactants/dlvo.php

The particles can be stabilised via:

- Charges, typically via an anionic surfactant. If the zeta potential (https://www.stevenabbott.co.uk/practical-solubility/zeta.php) is somewhere over 30 mV then the particle has a good chance of being stable.
- Steric effects, a long hydrophilic (typically ethoxylate) chain sticking out.

Arguably, charge-charge repulsion is the stronger stabilization method because steric repulsion can be changed into attraction (bridging or depletion flocculation) via extraneous polymers. The downside is that those water-loving ionic groups are not desirable in the final paint because they attract water from the environment.

The dispersion also has to be stable against settling or creaming:



Stokes Sedimentation

Particle r nm 1004
ρ_p g/cc 1.54
ρ_l g/cc 1
Viscosity η cP 5
Aspect Ratio 1
Settling height h mm 50
Width b mm 10
Angle θ 0
φ 0.21
Rotor radius r_{cent} mm 100
RPM 0
v m/s 6.27e-8
Time 9.2day
g_{rel} 1.0
Gr: Gravity v kT 1.05
Too small to settle? No

https://www.stevenabbott.co.uk/practical-solubility/stokes.php

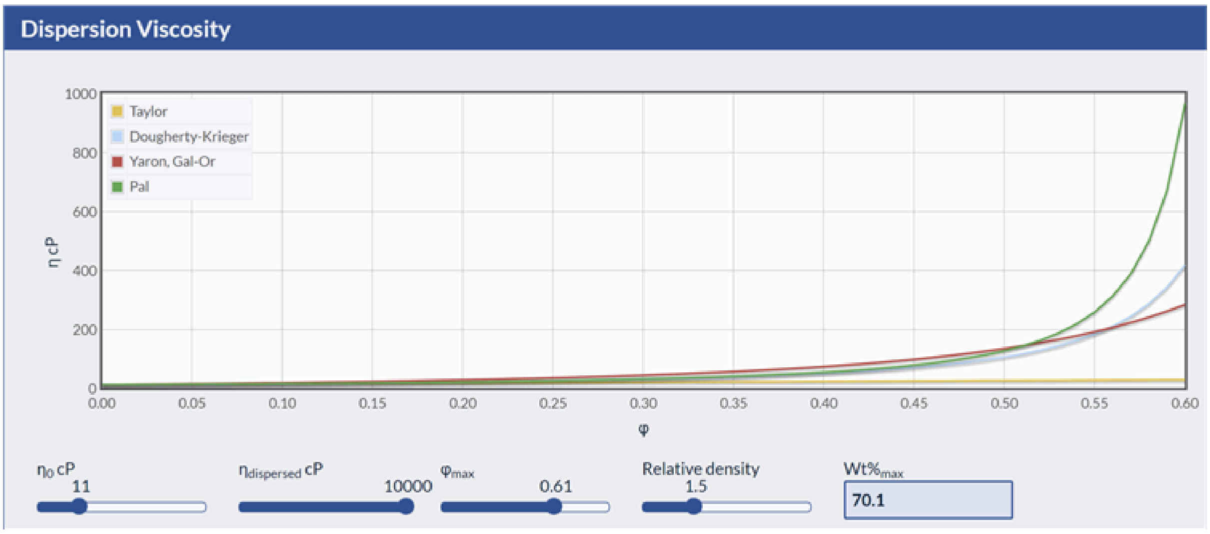
Stokes law tells us that the time for a particle (the pigment or emulsion drop) of radius r and density ρ_p to fall a distance h in water of viscosity η and density ρ_w (and therefore density different $\Delta\rho = \rho_p - \rho_w$) is given by:

$$t = \frac{h\eta}{2.18 g \Delta\rho r^2}$$

In the screen shot, a modest particle density is assumed. The volume fraction input, φ, allows the use of the Richardson-Zaki formula where the velocity is reduced by $(1-\varphi)^{5.65}$. A higher volume fraction also increases the viscosity, as discussed next.

To reduce settling, in addition to the necessity, for other reasons, of a high viscosity, smaller particles are helpful. Unfortunately with densities like 4.2 and 2.7 for TiO2 and CaCO3 respectively, pigments are trickier to keep in suspension.

Viscosity and rheology



η_0 cP 11

$\eta_{dispersed}$ cP 10000

φ_{max} 0.61

Relative density 1.5

Wt%$_{max}$ 70.1

https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

A complication of understanding the viscous behaviour of these systems is that we have two types of dispersed particles – polymer emulsion and pigments. If, for simplicity, we use a single value for the volume fraction, φ, of "particles", then over 60%, the viscosity increases, depending on the model and on the internal viscosity of the droplet, by about 40x. In this region the close packing limit for rigid spheres is being reached and the precise increase to higher φ values is too complex to model. In any case, these models are for smooth spheres with no attractive forces between them. As we try to go to a higher volume fraction required for a practical paint, we lose any good theory and have to rely on rheology to compare/contrast formulations.

Although in production and QC it is necessary to use simple measurements of "viscosity", no one in formulation should ever measure "the" viscosity of a paint. Viscosity depends on multiple factors so formulations should be, at a minimum, be tested via flow curves.

The standard flow curve Flow_Shear dependent is a single sweep from low to high shear rate. For paints the reverse sweep is required to get information about thixotropy Flow_Thixotropy.

Because low shear performance (e.g. avoidance of sagging if applied to a vertical surface) is important, the Levelling & Sagging app at https://www.stevenabbott.co.uk/practical-rheology/Levelling-&-Sag.php is useful. This emphasises the need for careful measurements in the 0.01/s to 5/s range.

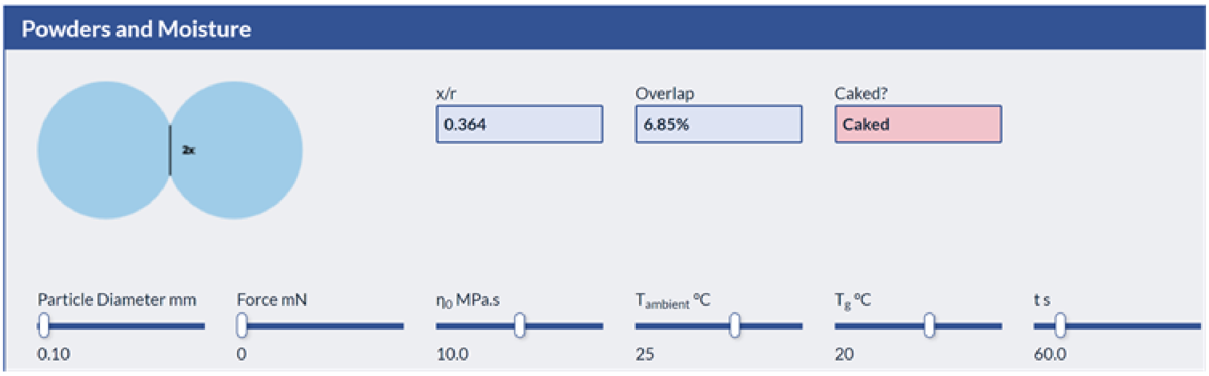
For the most information with the least effort, checking out Flow_G' and G" is effective. The balance of elastic (G') and lossy (G") shear properties offers insights not so much into any specific formulation, but into trends that occur with changes of concentration or ingredients. This means that you set up a few experiments with different concentrations or ingredients and see the G' & G" trends. From these you get a deeper understanding of how quickly/slowly paint properties change during, say, drying.

Thickeners

Controlling behaviour in the can, on the brush/roller, in the minutes after application and, for vertical surfaces, controlling sag, usually requires smart thickeners as part of the formulation package. There is a choice between Thickeners_Associative Thickeners and Thickeners_Polymeric Thickeners, with the associative type being generally preferred because their kinetics of breaking and re-forming are normally faster than those of classic polymers.

Film formation

It is helpful to introduce a neutral, and slightly unusual, word into the discussions: sintering. There are many terms used to describe the joining of blobs that are packed close together, but they lack a coherent theory to capture the key parameters that influence the phenomenon. The Rumpf equation for sintering describes the key elements behind the joining of particles, and allows us to identify how changes in our system can affect film formation for better or worse.



Powders and Moisture

x/r: 0.364

Overlap: 6.85%

Caked?: Caked

Particle Diameter mm: 0.10

Force mN: 0

η_0 MPa.s: 10.0

$T_{ambient}$ °C: 25

T_g °C: 20

t s: 60.0

https://www.stevenabbott.co.uk/practical-mechanical/Powders-W.php

The Rumpf equation tells you how overlap x (in the diagram you see 2x) changes over time t for particles with radius r and “viscosity” η and surface energy γ. The equation includes any applied external force, F, which we can assume to be 0.

$$\left(\frac{x}{r}\right)^2 = \frac{t}{\eta}\left(0.8\frac{\gamma}{r} + 0.4\frac{F}{\pi r^2}\right)$$

The use of x/r in the formula is deliberate because we’re generally interested in the fractional sintering – a value of x on its own doesn’t convey much.

There are a few interesting aspects to this equation:

- Contrary to instincts, a low surface energy, γ, *reduces* the rate of sintering. This isn’t surprising because the process is driven by surface energy. The surfactants necessary to disperse the blobs now become unhelpful. However, the surfactant effect on surface energy is likely to be no more than a factor of 2, so it’s not worth worrying about.
- Smaller particles fuse faster. Unfortunately they are also harder to keep separate in the original formulation, so there are trade-offs
- We need to attend to the viscosity of our emulsion particles (here we ignore the complexities of the pigment particles). In addition to this being a difficult thing to measure (I have found no helpful literature on how to do this), the variation of viscosity with water or concentration of “coalescing solvent” is an added complexity.

The apps associated with the app above are all related to food science. But physics is universal and the idea that the Tg changes dramatically with moisture content which in turn is governed by the water-vapour isotherm means that there is a logical chain for investigation. Your paint is designed to dry out, but what is the water-vapour isotherm – how dry is dry? And what is the effect of small amounts of water on the Tg? If film formation is complete long before the water has evaporated then these questions are irrelevant. The important thing is to know if/when film formation happens.

If the particle Tg is not affected by the water (an advantage for long term performance) then either the Tg has to be low enough for the film forming temperature to be close to room temperature. Or you have to use a coalescing solvent. The solvent must have a lower volatility than water so its concentration increases sufficiently to lower the Tg and decrease η. It must also have a good Dissolution_Hansen Solubility Parameters compatibility with the polymer.

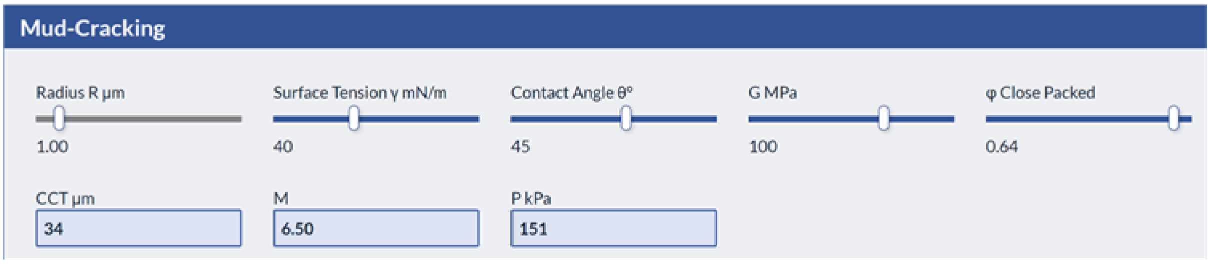
Good, but not too good. The simple story is that the coalescing agent reduces the viscosity of the whole particle. But that would require a large amount of agent, which then has to escape the paint over days, weeks, months. What is really wanted is for the agent to soften the outer portion of the particle, without going too deep inside. If, as seems likely, the outer portion has more hydrophilic components then its HSP will feature a higher δH than the bulk. So the HSP of the coalescing agent should be optimised for a Distance to that higher δH system. At the same time, the δH should not be so high that it partitions too readily into the remaining water which itself might have an HSP closer to that of the polymer thanks to the various soluble components that increase % concentration as the water evaporates.

Everyone wants “low VOC” paints. But they also want paints that form wonderfully tough, water-resistant films. Any film former that remains in the paint is a potential form of weakness, so having them volatile is optimal for paint performance. How can you have a volatile film former while legally claiming “low VOC”? Any molecule with a BP > 250°C or VP less than 0.01 kPa at 20°C [those values are confused in the literature so don’t quote them] is officially classed as zero VOC so you can add as much of it as you like to your paint.

Mud cracking

If the drying process happens too quickly, before the sintering takes place, then the paint can start to show mud crack patterns, such as the one shown here, taken from the app, https://www.stevenabbott.co.uk/practical-coatings/Cracks.php, which describes the phenomenology while giving little advice about how to solve the cracking issue. It’s interesting that the cracks join at right angles, though sometimes you get patterns that crack at 120°. Interesting, but not so helpful.

For some formulation insights we need to turn to the mud-cracking app.



Mud-Cracking

Radius R μm 1.00

Surface Tension γ mN/m 40

Contact Angle θ° 45

G MPa 100

φ Close Packed 0.64

CCT μm 34

M 6.50

P kPa 151

https://www.stevenabbott.co.uk/practical-coatings/Mud-Cracking.php

Here we calculate the CCT, the Critical Crack Thickness. Above the CCT, cracking is likely – telling us that one way to reduce cracking is to coat thinner.

The cracking is driven by a capillary pressure, P, calculated from the particle radius R, the surface tension γ, the contact angle θ and the close packed fraction φ.

$$P = \frac{2\gamma}{r}\frac{3\varphi cos\theta}{2(1-\varphi)}$$

Already we can see that cracking is reduced if we reduce γ and increase r.

The critical crack thickness depends on a on a "coordination number" M, which is related to φ and the "how well the particles are stuck together" parameter G. One well-known formula is from Tirumkudulu, and tells us that:

$$CCT = 0.64\left(\frac{GM\varphi R^3}{2\gamma}\right)^{½}\left(\frac{2\gamma}{PR}\right)^{½}$$

To increase the CCT (so we can get thicker coatings without cracking), we can either give up on otherwise desirable small radius particles or increase G. For entirely the uninteresting systems studied by academics, G is the "shear modulus" arising from pure particle-particle interactions. For the rest of us, G is more related to dissipative shear resistance (the "energy release rate") from polymers that are (the emulsions) or surround (additives) the particles. Unless we use polymers that are especially weak, G is likely to be high if we give the system time to sinter. If our coalescing solvents and/or surfactants also reduce γ (which ends up as a $\frac{1}{\sqrt{\gamma}}$ dependency) that can also help.

The advice, from the web page is not super helpful, but decades of mud-cracking research hasn't come up with anything more profound:

- Although the correlation with R is good in simple systems, real-world systems correlate better with SSA, Specific Surface Area. For spherical particles the conversion from SSA in m^2/cm^3 is 1/R when R is in μm so for your real particles you can try using 1/SSA instead of a nominal R. However, the SSA data has a dependency of CCT on $\sim SSA^{-2.2}$ rather than the expected 1 from $\frac{\left(R^3\right)^{½}}{R^{½}}$
- Shape, over a range of typical particles, makes relatively small difference.
- Soft particles have no CCT - they simply fuse.
- A mix of soft and hard particles can start to crack when hard particles start to exceed 50% and can percolate across the film.
- Lowering the (dynamic) surface tension isn't a straightforward win. That will also decrease the contact angle, which increases P.
- Experiments often show small effects of drying rates - but these are often super-slow idealised experiments or for slow-drying paints. For coatings that dry in seconds it's highly likely that "happy" formulations that keep moving particles for as long as possilbe, will crack less than "unhappy" ones. Putting it another way ...
- ... the longer the available "relaxation time", the more viscous flow that can reduce stress.
- A bi- or multi-modal particle distribution where small particles fill gaps between larger particles obviously help increase CCT.
- You can (at least in principle) measure stress build-up in a coating using the cantilever approach described in the Stress from Bending app, https://www.stevenabbott.co.uk/practical-coatings/

stress.php.

- Extra polymers can sometimes hinder (if they make G too small) but mostly help as they can link up otherwise weakly-connected particles. There's a subtle issue here described in the Ceramic Green Strength app, https://www.stevenabbott.co.uk/practical-adhesion/ceramic-green-strength.php. If the polymer congregates at the intersections of the particle it has a bigger effect on strength than if "wasted" by coating most of the particle.

Emulsion Polymers

Links

Dispersions_DLVO, Dispersions_Zeta potential, Flow_Polymer Viscosity, Surfactancy_Dynamic Surface Tension, Surfactancy_Emulsification, Surfactancy_Emulsion Inversion, Surfactancy_HLD

Typical uses of emulsion polymers are described in Water-based Adhesive and Emulsion Paint. The variety of polymerisation techniques and their complexities are beyond the scope of this chapter. Instead we focus on the “emulsion” part of making them and provide a toolkit within a toolkit.

What’s in this toolkit?

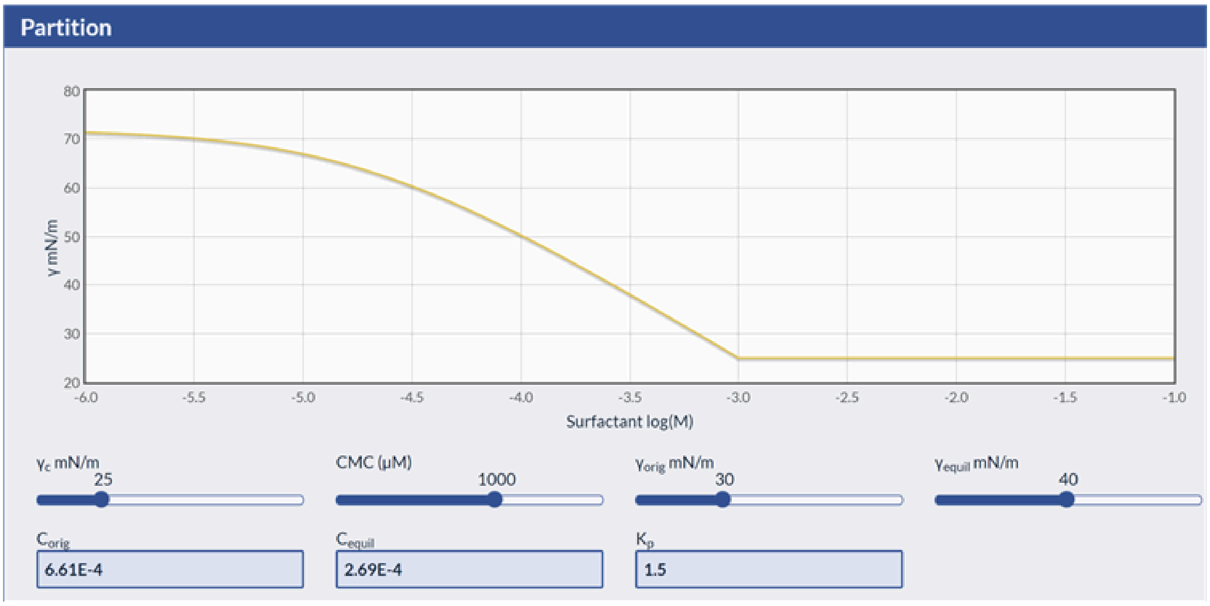
In addition to basics such as surfactant MW, Krafft Point and Cloud Point, you need:

- A knowledge of surfactant o/w partition coefficients
- Knowledge of dynamic surface tension effects
- A method & data to tell you how to respond to:
    - Changes in the oil (e.g. acrylates, methacrylates)
    - Changes in the temperature
    - Changes in salinity (ionic surfactants change their own salinity…)
    - Changes in surfactant type and surfactant mixtures
    - Changes in surfactant efficiency
    - Changes in polar oils (defined later) and other additives
    - A knowledge of interfacial tensions in delicate balanced emulsion systems
    - A knowledge of curvature effects
    - Thermodynamic AND Concentration inversions

Why a toolkit in the FST? There wasn’t enough usable science for a training session to a large corporation on emulsion polymerisation (which relies on “this is how we’ve always done it”). So the idea of a toolkit emerged, with a reverse guarantee: Success cannot be guaranteed if you adopt the toolkit’s science, but you can guarantee that life will be unnecessarily difficult without it.

Partition coefficients

Ideally, your surfactant should be at the o/w interface. But if it is too soluble in the water or oil, it will be mostly wasted. Measuring the partition coefficient should be routine, but it sounds difficult. In fact it is delightfully simple so should be routine:



γ_c mN/m 25

CMC (μM) 1000

γ_{orig} mN/m 30

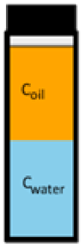
γ_{equil} mN/m 40

C_{orig} 6.61E-4

C_{equil} 2.69E-4

K_p 1.5

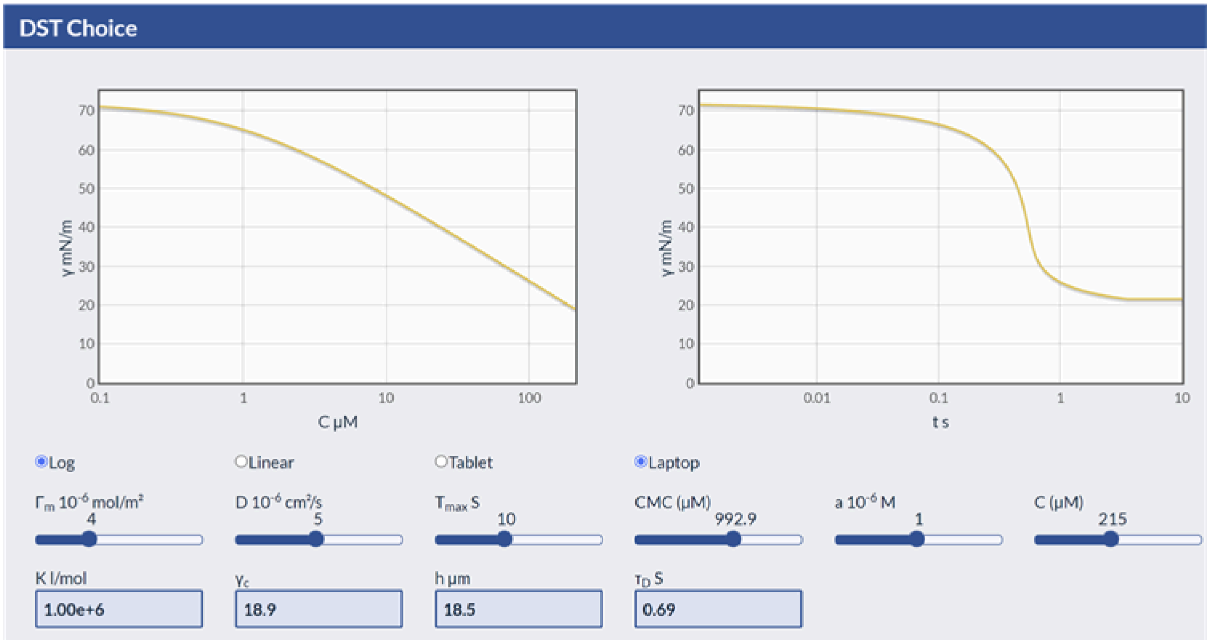
https://www.stevenabbott.co.uk/practical-surfactants/partition.php



Put a small amount of surfactant into a tube containing a known volume of water. Measure its surface tension, γ_{orig}, in the example above that's 30 mN/m. Now add the same volume of oil, shake and allow to equilibrate. Take out (most of) the water phase and measure its surface tension, γ_{equil}. Separately you will have measured the surface tension versus concentration for the surfactant (Langmuir isotherm), a typical experiment used to measure the CMC and the saturated surface tension γ_c. From a calculation based on the Langmuir isotherm you can work out the starting concentration (in this example 6.6 μM) and the equilibrium concentration 2.7 μM. From this the partition coefficient, $K_p = C_{oil}/C_{water}$, turns out to be 1.5.

Dynamic surface tension, DST

Your polymerisation is a dynamic process so a knowledge of how fast or slow the surfactant is to get to the interface will be helpful. Modern bubble pressure tensiometers can measure DST over a good range of time-scales. How do you interpret the data? Although it is common to analyse DST to get ad hoc parameters such as those of Rosen, https://www.stevenabbott.co.uk/practical-surfactants/dst.php, there is a more insightful technique:



● Log
○ Linear
○ Tablet
● Laptop

Γ_m 10^{-6} mol/m² 4
D 10^{-6} cm²/s 5
T_{max} S 10
CMC (μM) 992.9
a 10^{-6} M 1
C (μM) 215
K l/mol 1.00e+6
γ_c 18.9
h μm 18.5
τ_D S 0.69

https://www.stevenabbott.co.uk/practical-surfactants/dst-choice.php

The app text explains how you extract the core surfactant value Γ_m which is the surface excess concentration and via the diffusion coefficient D (which doesn't change much over a wide range of surfactants) you can derive the characteristic diffusion time τ_D. You then find that to get a low value of τ_D you need a surfactant with a relatively high CMC, a fact that is little appreciated.

Hydrophilic-Lipophilic Difference, HLD

To access most of the effects that are important during an emulsion polymerisation, the Surfactancy_HLD approach is the only one powerful enough to be of help. The HLB approach has long-since been discredited and should not be used.

Briefly, HLD says that the balance of a surfactant, oil and water system depends on salinity S, temperature T, a measure of oiliness EACN and a characteristic value for the surfactant Cc. That balance is the HLD given by:

$$HLD = F(S) - k.EACN - \alpha(T - 25) + Cc$$

There are 3 sets of parameters:

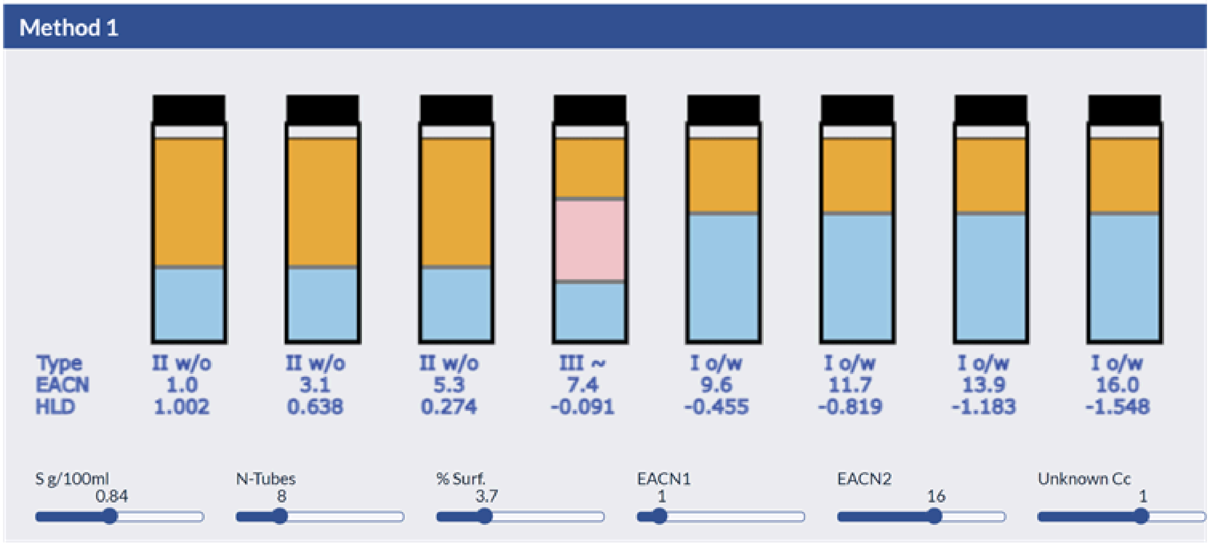
- For ethoxylates, F(S) = 0.13S, α = -0.06 and k~0.16
- For ionics, F(S) = ln(S+Surf), α = 0.01 and k~0.16. The extra term in F(S) is because the ionic surfactant creates its own contribution to salinity.
- For Spans, APGs, Polyglycerols, F(S) ~ 0, α = 0 (maybe -0.01 for polyglycerols), k~0.16

There is some debate about k values but given the many other uncertainties we don't have to worry about it here.

The EACN, Equivalent Alkane Carbon Number, gives us a measure of the oiliness of the oil. If your acrylate behaves (in surfactancy terms) like octane, then its EACN = 8. Going in the more hydrophilic direction we get EACNs of 3 or -1. That's because it is a scale of virtual hydrocarbons, we're not proposing that you can make an emulsion with a molecule containing -1 carbon atoms.

The Cc *should* be a value provided by your surfactant supplier. It has been a battle to get them to measure and provide these numbers but slowly users are getting their way. As described in the HLD chapter, users can set up simple *scans* (next paragraph) to check whether the "same" surfactant (different batch or different supplier) has the same Cc and, when it doesn't (commercial surfactants are mixtures) they can compensate using the HLD equation.

Scans are routinely used to measure EACNs and Ccs. Across a set of tubes you systematically change one component of the HLD equation (e.g. S) with two knowns kept constant (e.g. T and EACN) and from the tube where HLD ~ 0 (i.e. the Type III transition between Types I and II) you can calculate the unknown (in this case the Cc).



Method 1

Type	II w/o	II w/o	II w/o	III ~	I o/w	I o/w	I o/w	I o/w
EACN	1.0	3.1	5.3	7.4	9.6	11.7	13.9	16.0
HLD	1.002	0.638	0.274	-0.091	-0.455	-0.819	-1.183	-1.548

S g/100ml 0.84

N-Tubes 8

% Surf. 3.7

EACN1 1

EACN2 16

Unknown Cc 1

https://www.stevenabbott.co.uk/practical-surfactants/measure-cc.php

These scans can be used to answer many different questions, from the Cc values of new surfactants (typically broad scans) to changes in Cc between batches of the "same" surfactant (typically a high-precision scan) and, by looking at what happens when you add the same small amount of the ingredient to each tube, how that ingredient changes the HLD (the position of the transition tube) and the efficiency of that system (the size of the intermediate phase). A detailed examples covering polar oils is described below.

For your emulsification process, you first decide what HLD values you require during the different stages of your process. If, for example, you want a classic o/w emulsion then an HLD of -0.5 to -1 is optimal. If you are using microemulsions then you need HLD ~ 0.

Now we can go through each aspect of the toolkit, listed above.

Changing oil

If for the past year you have used an oil with EACN = 3 and you now want to polymerise an oil with EACN = -1, you can instantly work out that the Cc has to reduce by 4x0.16 = 0.66. The beauty of HLD is that Cc values are a molar-weighted average of individual Cc values. So if you previously had an X:Y ratio of two surfactants, you can shift the Cc by 0.66 by going to X+δ:Y-δ.

Changing temperature

If you *wanted* to have an o/w emulsion at room temperature then a microemulsion at the reaction temperature of 55°C then you can choose an ethoxylate system. The ethoxylate coefficient of -0.06 means that a 30°C change gives an HLD change of +1.8. This means that you need to start with an inefficient hydrophilic system of HLD = -1.8 to achieve your microemulsion at 55°C. If you happen to go to 65°C then you are encouraging a phase inversion to w/o which may not be what you want.

The point of the previous paragraph is that many ethoxylate-based emulsion polymerisations are undergoing very large shifts in HLD … and because people have fixated on HLB, which contains no explicit T-dependence, they have been unaware of this issue.

If, on the other hand, you wanted to move to a greener APG, if your emulsion system required a T-induced change in HLD during the process, your APG will fail because APGs have no T-effect on HLD.

For those who like to stay in control of their systems, the large T-dependency of ethoxylates is a good argument for *not* using them.

Changes in salinity

For non-ionic systems, addition of salts, for whatever reason, has little or no effect on the HLD. The Ln() dependence for ionic systems means that if you start with a low salt concentration (maybe just the surfactant's own salt contribution) of S ~ 0.001 g/100cc, an extra 0.001 g/100cc changes the HLD by Ln(0.002)-Ln(0.001) a change of 0.7! Hence it is a good idea to have a modest amount, say 0.01 g/100cc of "background salt" to stabilise the system.

When you have to choose, for other reasons, between ionics and non-ionics, those who want to keep in good control of their systems might want to take into account whether a background salt would be acceptable in the overall process and final product.

Changes in surfactants

Those (and that seems to be most formulators) who are forced to make changes to their surfactant package for cost, greenness, supply chain or other reasons can seriously reduce the pain of the transition by using HLD.

If you are swapping within the same class of surfactants then by knowing the Cc you can instantly know how much change there is from the previous one and what to do about it. If you are swapping class then you need to think deeper if you are relying on T effects or if you have the need for significant amounts of salts.

Changes in surfactant efficiency

The full theory, HLD-NAC, shows not only *where* you have to be in surfactant space, that's the HLD part, but *how efficient* the system will be when you're there, that's the NAC part. You can have two surfactants which produce identical HLD values but where one (there's a special class called "extended surfactants") is far more efficient than another. There is also a phenomenon that the efficiency for a given surfactant decreases as the EACN of the oil increases. You can add "linkers", or "cosurfactants" that increase the efficiency, but read this caution, then the next section on polar oils, before trying these.

There are few known rules for success in finding additives to improve an HLD system but there is a key rule for avoiding failure: Additives may change HLD, efficiency or both, so test additives within a scan.

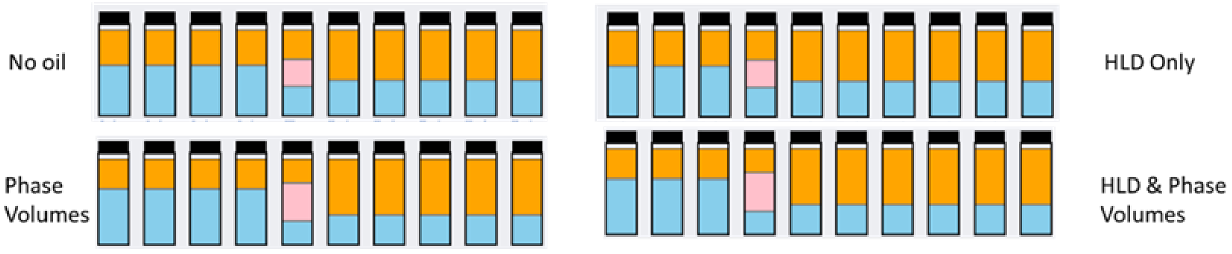
If you don't use a scan, then if you additive produces a wonderful improvement but changes the HLD, the result in your tube will be worse and you will have missed that a small change in your surfactant blend to restore the HLD would give you better results.

This brings us to the next issue.

Changes in polar oils and additives

Suppose you have to add some octanol for some reason. This is an oil, so might have an EACN, it's a "linker", it's even a surfactant. So how do you incorporate it into the HLD equation? The general answer is: you can't. There are no known rules for these sorts of molecules which are conveniently called "polar oils".

However, because HLD is based on scans, it is easy to set up a scan around your current system, with the central tube being where HLD = 0, then you add the same small amount of the polar oil to each tube and see how things change. You can repeat the additions to get a quantitative measure of how both the HLD and the efficiency change. Here we start with the reference scan at the top right and see 4 possible outcomes: No change, change to HLD only, change to phase volumes (efficiency) only, change to both.



Testing additives would be easy if they changed only the phase volumes. Given that it is likely the that they will change both HLD and phase volumes, the experiments are a bit slower, while the chances of success are much higher.

Curvature

As HLD transitions from negative through zero to positive, the curvature changes from negative, o/w to neutral, microemulsion, to positive, w/o. For control of your emulsion polymerisation, knowing where you are in that curvature space is clearly important.

The change also affects the interfacial tension. If a typical IFT between an oil and water is 5 mN/m at a large HLD , then as the HLD decreases the IFT might decrease to 1 but around HLD = 0 it might decrease to 0.01.

If you have a specific emulsion polymerisation method you are trying to control, it would be a good idea to know that you were safely in the > 1 mN/m zone for one kind and in the ~ 0.01 mN/m zone for another.

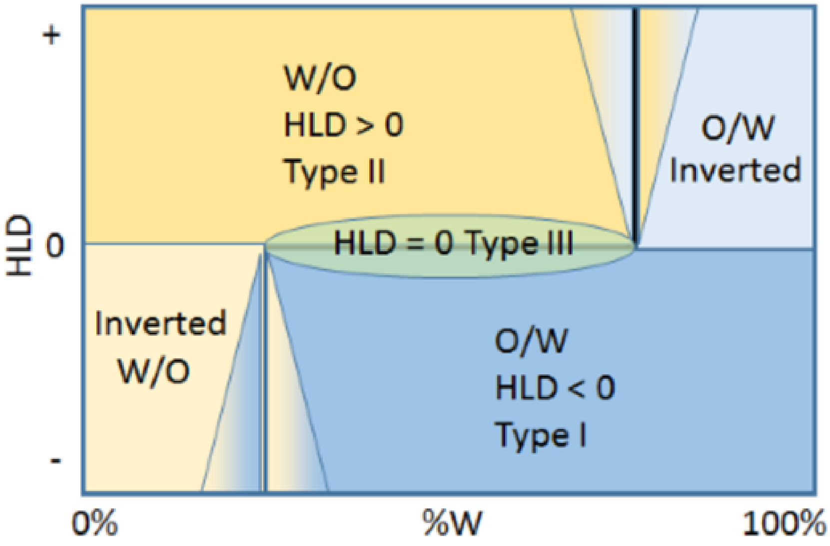
The other reason for avoiding the low IFT region is that if you stray a little further, you are in a zone of opposite curvature and you risk a nasty phase inversion.

Inversion

There are times when you want to carry out controlled emulsion phase inversions. HLD tells you how to do it via T, S, EACN or CC. This is a *thermodynamic* inversion. The other way to do it is via having too little of the bulk phase, so the emulsion is forced to switch to being the "wrong" way round. This is a *concentration* inversion. Each on its own is a useful process when done deliberately and a disaster if done accidentally.

The reason there is a risk of this happening with emulsion polymerisation is that we would generally like the least water and the most emulsion in the final product. Take this too far and the oil phase will take over as the bulk phase and you have an undesirable w/o batch in your reactor.

The famous diagram worked out by Salager helps us see where our formulation is within inversion space:

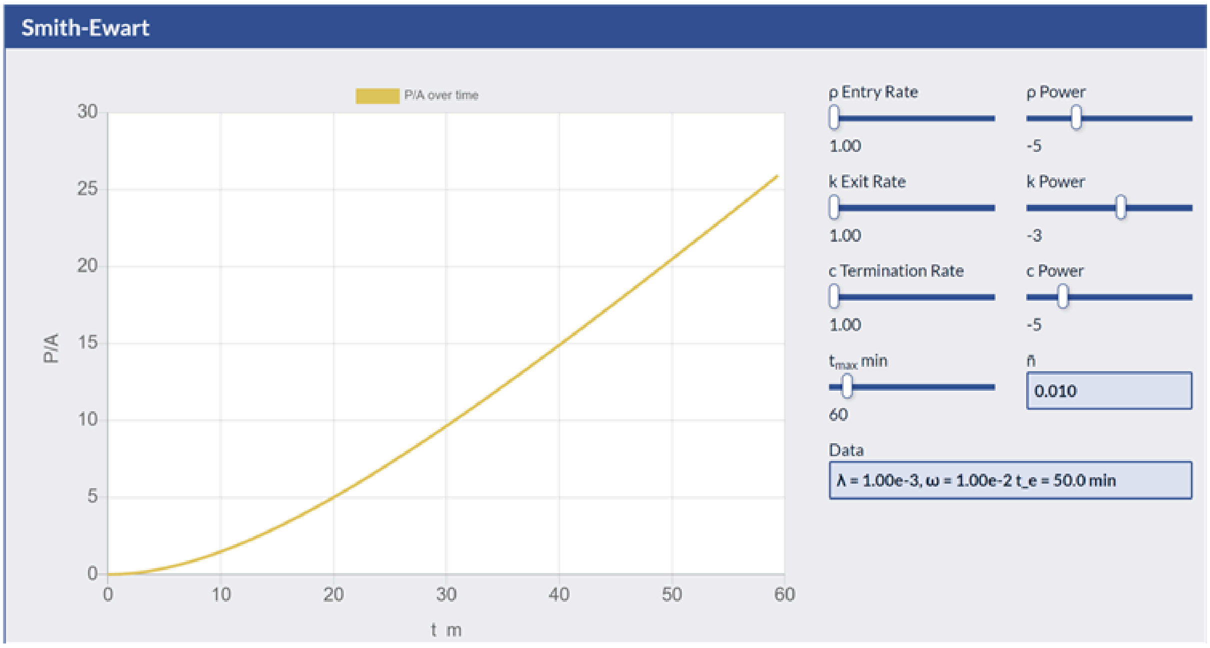


https://www.stevenabbott.co.uk/practical-surfactants/inversion.php

The uncertainties around the edges of the non-HLD transitions are due to the complex kinetics and hysteresis common with these transitions. The point of the diagram is that if, like many formulators, you are unaware of the HLD effects, you can get transitions when you don't expect them. It is a good idea to use the diagram to ensure you are safely in your desired zone or that you get the thermodynamic or concentration inversion you require.

Smith-Ewart

For those doing classic emulsion polymerisation, there is a popular idea that you can describe it using the Smith-Ewart formula. Although an app exists, it is not clear what value it adds. It is provided for those who might find it useful:



ρ Entry Rate 1.00

ρ Power -5

k Exit Rate 1.00

k Power -3

c Termination Rate 1.00

c Power -5

t_{max} min 60

n̄ 0.010

Data λ = 1.00e-3, ω = 1.00e-2 t_e = 50.0 min

https://www.stevenabbott.co.uk/practical-surfactants/Smith-Ewart.php

Water whitening

Because emulsion polymers are created and coated from water, they will have at least a modest attraction for water in their final form. If this is a general absorption that might create issues around long-term stability. It is also observed that atmospheric water builds up as little droplets. These then scatter light, leading to water whitening.

Because obvious explanations such as attraction to the surfactants used do not seem to be useful, the app can merely model the effect and suggest tricks to reduce it:

R drop radius nm 22

φ water % 5

d Thickness µm 50

RI Polymer 1.50

λ_{From}m nm 400

λ_{To} nm 650

m 0.89

C 10^9nm^4 49.5

%T Average 49.3

https://www.stevenabbott.co.uk/practical-coatings/Water-Whitening.php

The relevant equations and references are provided in the app.

Fragrances

Links

Diffusion_Barrier Properties, Evaporation_Basics, Evaporation_Enthalpy of Vaporization, Evaporation_Temperature and Antoine Coefficients, Fragrance_Activity Coefficients, Fragrance_Barrier Properties, Fragrance_Vapour Pressure

> A fragrance has been selected by a specialist team. Now it's your job to formulate it into the product. There are plenty of challenges ahead.

Controlled disappearance

The point of a fragrance is to evaporate. As the different fragrance molecules have different volatilities, the mixture that reaches the nose changes over time. Rather than fight against this inevitability, the fragrance is envisaged to have a certain top note (the instant hit of the most volatile), middle note ("the" fragrance) and bass note (the lingering memory of the fragrance).

As a formulator you might be asked to provide some control over this aromatic symphony. The starting point is to know what would happen to a simple mix of the aroma molecules. So you need to know their Antoine Coefficients (Evaporation_Temperature and Antoine Coefficients) in order to calculate their vapour pressures at the relevant temperature (e.g. 20°C for a typical cleaning product, 35°C for something on the skin). The three coefficients, AA, AB and AC are merely fitting constants that describe the vapour pressure of the pure liquid over a relevant temperature range, usually up to the boiling point. Databases of Antoine values tend to be in units of °C and mm/Hg so that's what we will use here. The vapour pressure, p, at temperature T is given by:

$$log_{10}(p) = AA - \frac{AB}{AC + T}$$

In the space above the fragrance, component 1 will have a pressure p_1 which is a fraction $p_1/\Sigma p$ of the total pressure. If the fragrance is measured in weight fractions, w_n, then the individual pressure values need to be corrected by the molecular weight, M_n, of each component:

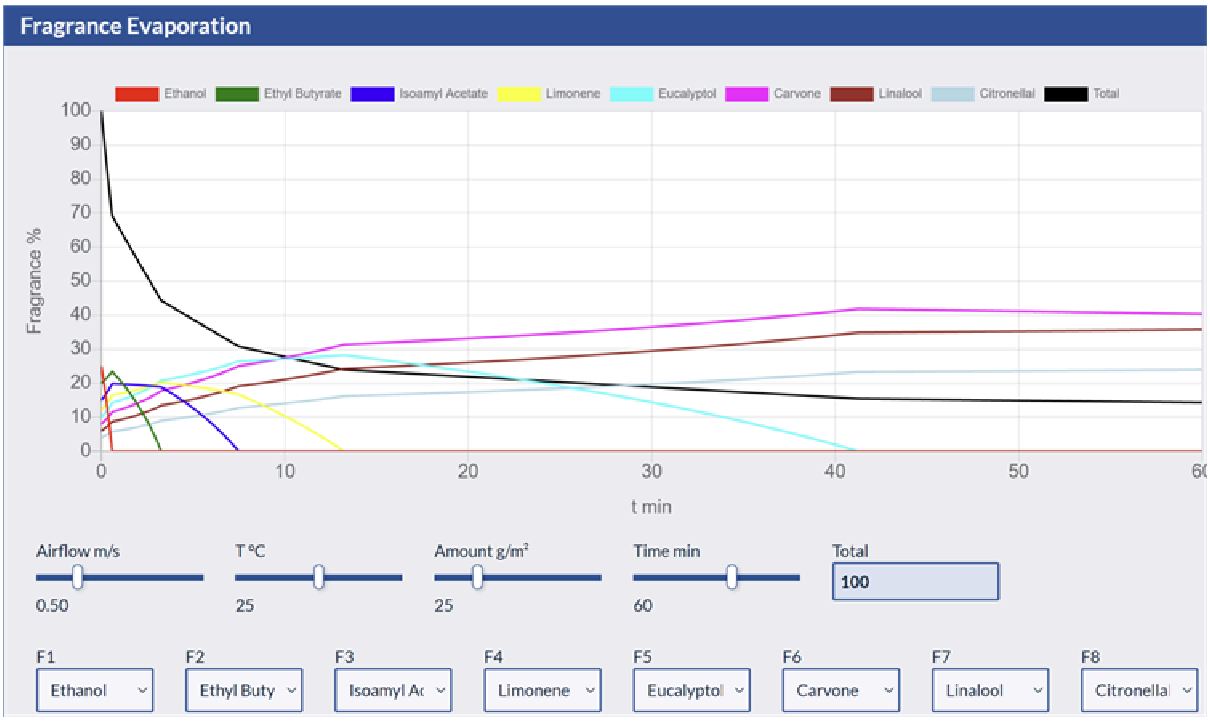
$$\sum p = \frac{w_n p_n}{M_n}$$

At a given pressure, the mass, m_n, of the n^{th} molecule in a volume V of the vapour is the number of moles per litre of vapour multiplied by the molecular weight, M, so, using 760 mm/Hg as atmospheric pressure and 24.7 as the number of moles per litre of an ideal gas at 25°C:

$$m_n = \frac{p_n M_n V}{760\ x\ 24.7}$$

The rate at which the solvents are removed depends on the air velocity. The dependence is complex and is described in Evaporation_Basics.

Putting these things together we can calculate how a complex fragrance mixture will change over time. The app, which, for simplicity, only has 8 ingredients, chosen from a list of 40+, shows the principle:



Airflow m/s 0.50

T °C 25

Amount g/m² 25

Time min 60

Total 100

F1 Ethanol

F2 Ethyl Buty

F3 Isoamyl Ac

F4 Limonene

F5 Eucalyptol

F6 Carvone

F7 Linalool

F8 Citronellal

https://www.stevenabbott.co.uk/practical-coatings/Fragrance-Evaporation.php

The ethanol used to create the fragrance has, in this example, disappeared in less than 1 minute. After 1hr, only 15% of the original formulation remains, and it's made up of the three low volatility components, carvone, linalool and citronellal.

It's not ideal

The calculations have assumed ideal solutions, i.e. where the activity coefficients (Fragrance_Activity Coefficients) of everything are 1. For a fragrance delivered from ethanol, once the ethanol has gone, the assumption is not unreasonable – because molecules with very different solubility properties and chemical functionalities tend not to be useful as fragrance molecules.

This changes once the fragrance is formulated into other products. A candle is a very different solubility environment from a shower gel. More hydrophilic molecules will have a higher vapour pressures than hydrophobic ones within a candle, and vice-versa in the shower gel. A difference in activity coefficients of 2 is not unreasonable, meaning that an ingredient might be twice as volatile in one formulation than another. This would mean a very different aroma.

This is such a significant effect that when a product range uses the "same" aroma for, say, candles and shower gels, the formulations have to be significantly different.

Indeed, a reader kindly sent me the following which shows that subtle effects occur even within a single type of formulation, such as shower gels:

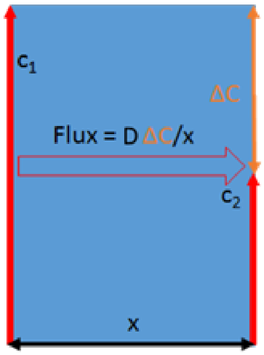
"An interesting observation to note is the phenomenon of partitioning of fragrance molecules within micelles in scented shower gels and shampoos. Hydrophobic molecules (such as limonene) position themselves in the core of the micelle, while more polar molecules (aldehydes, alcohols) prefer the micelle's

outer layer. Consequently, the olfactory sensation experienced by consumers using these hygiene products varies depending on the nature of the surfactants employed (e.g., SDS vs SLES) and the polarity of the fragrant molecules."

It is an ongoing debate whether these different variants can be designed via art & craft or whether a series of GC head space analyses allow the activity coefficients to be worked out and the formulae adjusted accordingly.

Fragrance migration

As discussed in the Diffusion_Basic Diffusion chapter, diffusion is simple to understand. It is driven by the concentration gradient, the difference in concentration, Δc, divided by the distance, x, over which the gradient exists. For a given gradient, the flux (amount per unit time) is simply a diffusion coefficient, D, times the gradient.



If we are concerned with fragrance/flavour molecules escaping, so-called flavour scalping, then we can assume that the external concentration is zero and that we have a packaging film from a single ingredient (we will modify this later) of thickness x.

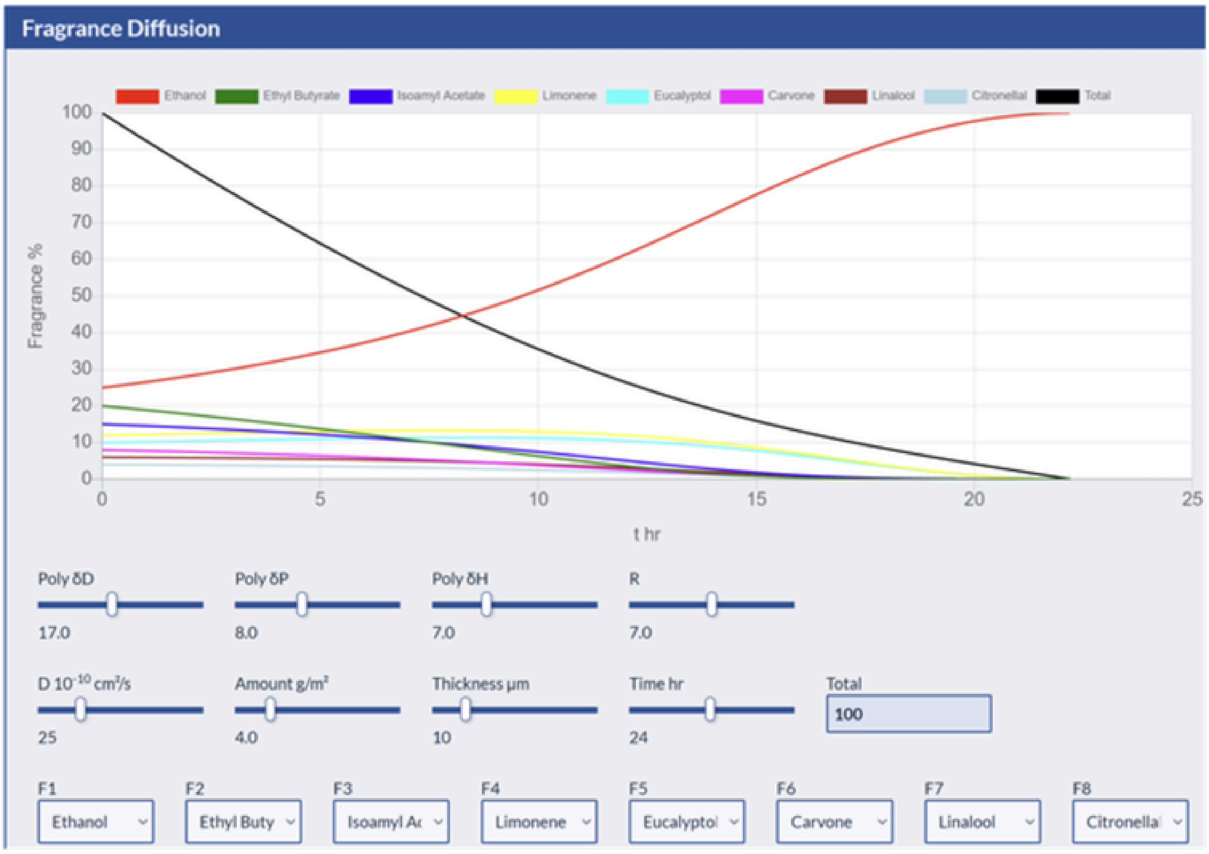
Different flavour molecules have different MWs and we can make a basic assumption that the diffusion coefficient is proportional to MW, so a molecule with MW = 200 will diffuse at half the speed of one with MW = 100. The precise dependence on MW is debatable, but because the concentration factor is much more important, we don't need to worry too much.

The concentration is really a partition coefficient between the internal environment and the packaging film. If the internal environment is dry then we are talking about Henry constant partitioning. If it is liquid then we have differential partition between solution and packaging.

To create a full model of any specific system is complex, so to illustrate the principles we use an app that makes the following (drastic) approximations:

- The fragrance has a maximum of 8 ingredients, chosen from a list of ~ 40
- The total amount of fragrance is assumed to be sitting on the packaging film as a liquid film with a g/m^2 equivalent to the total amount in your package.
- The partition coefficient depends only on the Dissolution_Hansen Solubility Parameters Distance between each fragrance molecule and the packaging material (you specify its 3 HSP)
- The maximum concentration for a Distance = 0 molecule is 1 vol%, with a fall off given by exp(-(Distance/Radius)^2) where the Radius is the distance over which solubility in the polymer goes from maximum to (effectively) 0. An "open" polymer has a large Radius, a "closed" one has a small Radius.
- Concentrations are too low to create concentration-dependent diffusion coefficients.
- The concentration on the outside remains as zero.
- The diffusion coefficient for a molecule of MW = 100 varies from 10^{-10} to 10^{-8} cm^2/s (a unit more common than m/s) via a slider from 1 to 100.
- Diffusion coefficients go as 100/MW.

Although you may object to some or all of those restrictions, remember that the app is meant as an illustration of broad effects – and is much better than the usual alternative which is no model at all.



Poly δD 17.0

Poly δP 8.0

Poly δH 7.0

R 7.0

D 10^{-10} cm²/s 25

Amount g/m² 4.0

Thickness μm 10

Time hr 24

Total 100

F1 Ethanol

F2 Ethyl Buty

F3 Isoamyl Ac

F4 Limonene

F5 Eucalyptol

F6 Carvone

F7 Linalool

F8 Citronellal

https://www.stevenabbott.co.uk/practical-solubility/Fragrance-Diffusion.php

Fragrances in water

Most fragrance molecules are, at best, only slightly soluble in water. As we increasingly move away from a solvent-based world, how do we formulate fragrances in water?

One obvious way is via a classic oil-in-water emulsion as described in Surfactancy_Emulsification. This is generally not appreciated for stand-alone fragrances because of the scattering of an emulsion, and the problems of long-term stability when you can't use high viscosities to stabilise the emulsion drops. For personal care products with their relatively large amounts of surfactant, there isn't much of a problem as the surfactant molecules are more than enough to keep the fragrance happy.

A way to solve both the visible scattering and longevity problem is to use a microemulsion. Using Surfactancy_HLD theory it's not hard to create a reasonable microemulsion. The difficult part is to find (via the NAC part of HLD-NAC theory) an *efficient* formulation using the minimum of surfactant. The tricks (rational cosurfactants, linkers or extended surfactants) described in the HLD chapter can make this possible. As fads of what are or are not acceptable surfactants come and go, reformulating with acceptable alternatives is relatively easy if you know the Cc values of the available alternatives, something that will become easier as suppliers are persuaded to provide these values routinely.

The idea of using the science of "solubilizers", "hydrotropes", "surfactantless emulsions" and so forth is generally hopeless because there is so much confusion about what these terms mean, and so little usable theory. That is changing, as described in Dissolution_Solubilizers and Hydrotropes where Dissolution_Kirkwood-Buff Kirkwood-Buff theory allows us to find out what is going on inside our formulation.

And many of these ideas are hopeless in practice. Who wouldn't want a "surfactantless emulsions"? It sounds such a great idea. You can do this with "solvents" but solvents are bad. So you do it with hydrotropes or co-surfactants or solubilizers because they are nice words. But they are still solvent-like molecules and, worse, you often have to use a lot of them to get the emulsion you like. Terms like "surfactantless emulsions" are great for academic papers and grant proposals, but usually not so brilliant for real-life formulation.

Why formulators dislike fragrances

Many formulations are smart balances of competing properties. Thickening by Thickeners_Wormlike Micelles, clever emulsification via Surfactancy_HLD and efficient Surfactancy_Anti-foaming are examples of such balances.

Things are fine till marketing come along with a new fragrance, carefully tuned to capture some essence of the zeitgeist. Taking out the old 0.1% and adding the new 0.1% often destroys the previously excellent formulation, and formulators have few tools for understanding why. It's worse than that. Many fragrance ingredients are common (low price is vital) so it might be that the new, destructive, fragrance isn't so different from the old one. But even the "same" fragrance molecule can be a mixture. So it may be that the problem comes from one sub-component of one of 20+ ingredients.

My view, based on the "polar oil" problem discuss in the HLD chapter, is that the fragrance effects are due to subtle partitioning effects not in the tail (these are boring and predictable for such low levels of molecules) but in the critical area between the head and tail. The scientifically relevant app, but one too hard to use in practice is based on Tchakalova's CIT technique: https://www.stevenabbott.co.uk/practical-surfactants/curvature.php.

It seems to me that the future is to take inspiration from the CIT approach and find ways to identify such effects swiftly via curvature-critical measurements.

Hair Conditioners

Links

Diffusion_Diffusion into Skin and Hair, Evaporation_Humectants, Flow_Basic viscosity, Surfactancy_Phase Diagrams, Thickeners_Wormlike Micelles

> If your hair has been damaged or its natural protective oils have been removed by aggressive Hair Shampoos then a fix exists via a conditioner. If the hair is suffering from excess conditioner then that problem can be fixed by a conditioner conditioner.

The problem to be solved

The hair follicle and the scalp's biome do a good job in keeping your hair in good condition. A combination of free and bound 18-methyl eicosonoic acid (18-MEA) makes the hair highly hydrophobic, with easy slip between individual hairs. Regular shampooing strips off the free 18-MEA and aggressive hair drying and combing helps to remove the bound 18-MEA. The cysteines in the hair keratin become oxidised into cysteic acids. The hair is, therefore, not in good shape.

A root-cause fix for this is to shampoo less frequently (a water wash will remove most "dirt" and a light shampooing will remove excess oil), and dry/comb more gently. The follicles and scalp biome can look after things rather well as they have had millions of years to work out how to do this.

The alternative fix is to:

- Add a convenient oil to coat the hair;
- Add a long-chain cationic molecule (quaternary ammonium, "quat") to lock onto the cysteic acids. As a bonus, the resulting salt is hygroscopic, adding some anti-static functionality for free;
- Ensuring that the formulation is low pH to encourage the cysteic acid form and to ensure that non-quat amines such as stearamidopropyl dimethyl amine are in their cationic form;
- Optionally (but they are becoming unacceptable), add some silicones for that wonderful silicone feel;
- Optionally (but they aren't as good as silicones) add some "natural oils" to smooth out rough hair surfaces;
- Optionally include some "strengthening" polymers.

As has been known for decades, this can be done with ~2.5% of cetyl alcohol and ~ 1% of cetrimonium chloride, plus <1 % of a thickening gum such as HEC, some fragrance and whatever preservative package fulfils the triple role of (a) stopping bad microbes from growing in the formulation, (b) without harming all the good microbes in, say, the scalp biome and (c) conforming to whatever is the latest trend in preservative claims.

If you want a "deep" conditioner, just add a lot more cetyl alcohol (etc.) and a lot more quat to emulsify it. As a bonus you later get to sell conditioner conditioner to restore the hair to some form of normality.

The silicones undeniably give a beautiful soft feel to the hair but too much of a good thing is a bad thing, especially because silicones (being insoluble) build up on the hair over time. They are also surprisingly sticky as the build up, so hair attracts more dust and dirt. In any case, regulations are taking us to silicone-free formulations as silicones are seen as allergenic and toxic for aquatic life.

With the need to replace silicones, formulators have to master the art of getting drops of "natural oils" from, say, rice bran or coconut to be delivered from emulsions, maybe trapped in coacervates, and then once on the damaged hair fibre, to spread spontaneously – something the silicones do and the oils tend not to do. Although relatively simple Cleaning_Contact Angles thinking can be used, the fact that the hair surface changes from relatively good quality near the root to highly damaged near the tip, and that spreading can be interrupted by surface roughness, the formulator will find this difficult.

Polymers (beyond the thickeners added because consumers have been conditioned to believe that viscous = luxurious) will, as claimed, "strengthen" the hair or "add body" but in general are just more gunk that sooner or later has to be removed via a conditioner conditioner.

Via clever coacervate science (Surfactancy_Coacervation) these ingredients can be added to the shampoo itself – allowing the shampoo to remove natural oils (18-MEA) and replace them with, say, cetyl alcohol which, while arguably "natural" is not natural to the hair system. For these systems the coacervate trick requires polymeric versions of the cationics to be used, polyquats (synthetic or natural guar gum) instead of simple quats. As above, these have the upside of providing "body" and the downside of being "gunk".

The formulator's problem

The simplicity and efficacy of a standard conditioner formulation is a problem for marketing. Consumers should be smart enough to buy the least amount of the lowest-price me-too product and that's not what marketing want.

Therefore the formulator has to add ingredients that allow marketing to claim, via plausible-sounding "logic", that *this* conditioner will condition especially well:

- Protein from whatever source sounds both sciencey and green, vegan etc;
- Anything with keratin in the title ("keratin polypeptides") because hair is made from keratin so more keratin polypeptides must be good for it;
- Stearyl alcohol and cetearyl alcohol (a mix of stearyl and cetyl alcohol) because more ingredients must be better;
- Aloe vera because everyone likes aloe vera because the polysaccharide acemannan is from aloe vera;
- Vitamins and antioxidants because who doesn't want extra vitamins in their keratin and maybe the antioxidants will help reduce the damage to keratin cysteines;
- Extract of some exotic plant because exotic plants must be good.

Although those ingredients can all be thrown in with little obvious problem, the more stuff that's added, the more unexpected problems can arise. Maybe the exotic plant extract forms a complex with the acemannan in the aloe vera, but only after 1 month of storage. That's a tough problem to identify. And all those extra nutrients may make your previously excellent preservative package incapable of passing microbial tests (the nutrient boost might overwhelm the preservative effect) so a difficult round of balancing antimicrobial efficacy against marketing trends has to take place.

As explained in the Shampoos chapter, well-controlled experiments on shampoos show that the average consumer forms their judgement based on the packaging rather than what's in the formulation – the same standard shampoo is “better” if provided in an expensive package and consumers score randomly when judging a standard or expensive shampoo delivered in the same standard package. We can be confident that many expensive conditioners, packaged accordingly, exert a similar placebo effect.

Conditioner-Conditioner

If there is too much build-up of conditioner residues then the consumer has to buy yet another conditioner to remove those residues. Removing a mix of silicones, quats and oils requires some powerful chemistry with risk of damage to the user’s hair and scalp … so you probably have to provide a post conditioner-conditioner conditioner.

Maybe consumers will work out that if they treated their hair more carefully, using fewer products, less frequently, they won’t need multiple levels of hair warfare.

Although management and marketing would object to a strategy of urging customers to purchase less, science is on the side of those companies who find a way to do more with less.

Hair Shampoos

Links

Cleaning_Boundary removal, Cleaning_Contact Angles, Cleaning_Surfactants, Cleaning_Temperature effects, Flow_Basic viscosity, Surfactancy_Anti-foaming, Surfactancy_CMC and Langmuir, Surfactancy_CPP and phases, Surfactancy_Coacervation, Surfactancy_Dynamic Surface Tension, Surfactancy_Foam Drainage, Surfactancy_Foam Ostwald, Surfactancy_Foam Rheology, Surfactancy_Foaming, Surfactancy_HLD, Surfactancy_Roll-up and Eötvös Number, Thickeners_Wormlike Micelles

A basic shampoo is very simple – just about any soap or surfactant will get rid of the small amounts of dirt and oils on our hair. For a mixture of marketing and functional reasons, a typical formulation is much more complex.

No-poos and dry shampoos

Many people use no shampoo other than when their hair gets super dirty after some tough job. Water alone does a good-enough job at keeping their hair in good condition. Others use dry shampoos – powders that absorb excess oil and, via combing, help remove other bits of dirt.

So for those who choose to use a shampoo, the absolute minimum of ingredients should do a good job of cleaning whatever needs to be cleaned. Using a minimum would be good for the pocket and good for the planet. It is worth noting that a good Laundry Liquids cleans at ~ 200ppm while a typical shampoo is used at 20,000ppm (i.e. 2%), an astounding over-use of a complex product for a simple task.

For those who want a compromise position using half the amount of shampoo, half as often cuts the carbon footprint 4x and saves money as well.

The essentials

Because of the high concentrations, applied in high doses, we can use baby shampoos, soap bars, simple SLS formulations and super-complex formulations, each with a different class of surfactants (sorbitan ethoxylates, fatty acid salts, SLS, complex mix) and get adequate cleaning of the generally small amounts of dirt and oils that need to be removed for a clean hair.

Users with hard water and those whose hair picks up various metal salts from the urban environment will benefit from small amounts of chelation from EDTA or a greener alternative.

All aqueous formulations need a preservative package, discussed below.

Used sparingly, the simplest formulation, surfactant, chelator, preservative, will achieve the nominal requirements for a shampoo.

Packed with features

For whatever reason, consumers have decided that their shampoo should be packed with extra features, none of which is relevant to the core task of cleaning excess dirt without destroying the hair's own protective biome and oils:

- Viscosity to give a luxury feel;
- Foaming partly to help disperse the viscous surfactant, partly because users associate "foaming" with "cleaning";
- Oils to replace natural hair oils removed by the shampoo;
- Silicones (or, as these are becoming unacceptable, "natural oils") to smooth out hair that is partially damaged by over-use of shampoos;
- Quats to avoid static from hair that is over-cleaned;
- Polymers to add "body" to hair that might have been damaged by over-cleaning;
- Fragrances for a feel-good factor;
- Probiotics and their equivalents;
- Ingredients with no functionality other than to provide marketing claims.

Let's discuss each in turn.

Viscosity

There is little consumer utility in making a shampoo viscous. Some say it stays better on the hand during application in a shower, but this isn't much of a justification. It is anti-functional because it makes it harder to spread evenly through the hair. But consumers have been conditioned to associate viscosity with luxury, so most shampoos are thickened.

Because most Thickeners_Polymeric Thickeners leave unpleasant sticky residues (though they might be used to create "body"), the most usual thickener is NaCl, via the magic of Thickeners_Wormlike Micelles. The salt can also provide some mild water softening effects so this cheap thickener is very attractive.

The downside of this trick is that many alternative surfactant formulations that might be considered greener, don't have the ability to form wormlike micelles. So formulators are trapped. They want to use newer, greener surfactants but have to preserve a functionless feature that consumers expect.

As is known to everyone who has formulated via salt thickeners, the other ingredients can have subtle effects on the thickening. If a new fragrance happens to provide optimal thickening at a lower salt concentration and Production add the standard amount of salt, the formulation might end up in the "past thickening" state and the entire batch is an expensive write-off. If the additive interferes with the thickening then maybe an extra bit of salt will save the batch – but maybe it simply won't thicken and the batch is a write-off.

Because the rheology of salt-thickened formulations contains time/speed/temperature complexities, it is especially tricky to check for subtle impacts of extra ingredients. A simple QC check with a single viscosity might say that a new additive is OK. Only a proper rheological scan, unlikely to be used in QC would reveal the subtle problems.

Some basic flow curve rheology, Flow_Rotational rheology, will allow judgement about flows through production (for example, calculating Flow_Poiseuille Flow through pipes) and for the end-user experience (pourability through the shampoo dispenser).

Foaming

Any surfactant molecule that is sitting in a foam wall is contributing nothing to cleaning the hair. Foam is, therefore, not just useless but has a negative impact on efficacy … except where the foam helps distribute the viscous surfactant more evenly through the hair.

But because decades of marketing has convinced consumers (see also Dishwashing Liquids where the standard test is foam lifetime rather than cleaning efficiency) that "*foam = clean*", formulators are forced to abandon otherwise excellent surfactants for those that create lots of stable foams in standard tests that may not apply to the real world.

As described in Surfactancy_Foaming, most surfactants will foam adequately if sheared efficiently. The shear of a viscous surfactant rubbed between the user's hands and the shear from fingers rubbing surfactant through the hair is very different from the standard tests such as a blender, a cylinder of solution being regularly inverted, or a stream of bubbles rising inside a cylinder.

This means that formulations are optimised for a non-core function via tests that don't replicate the real phenomenon. This problem is an opportunity. Finding imaginative ways to deliver foaming to consumers using less conventional surfactants might be an interesting way to use the laws of physics to your advantage. It's probably better than trying yet another combination to create an equivalent of SLES/CAPB to mimic what everyone else is doing.

Oils

Whether the added oils are "natural" or not, they aren't the oils that the consumer's scalp naturally produce. The chosen mix will be based partly on hard-to-make-scientific sensorial effects and partly on marketing claims for whatever oil happens to be fashionable.

In a formulation containing so much surfactant, the challenge of incorporating the oil into the formulation is not very hard. The real problem is if/when the oils interfere with the all-important salt thickening mechanism. Again, a marketing-driven non-necessity presents formidable formulation challenges for which we have few rational tools other than plenty of rheology.

Silicones

The highly desirable silky feel from silicones has to be balanced with the increasing need to claim "silicone-free" formulations. The non-volatile silicones suffer the problem of build-up over time. They even become sticky – attracting dust and dirt, making the hair worse. Like so many such problems, this is an opportunity to formulate "deep cleaning" shampoos to remove the problems created by the regular shampoos. Scientifically, the best surfactants for silicones are silicone-based surfactants. That's because the "hydrophobic tail" of a conventional surfactant is an alkane … and silicones prefer silicones to alkanes. If silicone surfactants are unacceptable then relatively high Cc (hydrophobic) surfactants are needed to be able to remove the silicones.

Some silicones can also be quats.

With the need to replace silicones, formulators have to master the art of getting drops of "natural oils", such as those from rice bran or coconut, to be delivered from emulsions, maybe trapped in coacervates (see below), and then once on the damaged hair fibre, to spread spontaneously – something the silicones do and the oils tend not to do. Although relatively simple Cleaning_Contact Angles thinking can be used, the fact that the hair surface changes from relatively good quality near the root to highly damaged near the tip, and that spreading can be interrupted by surface roughness, the formulator will find this difficult.

Quats

Quaternary amines polymers ("polyquats", synthetic and natural like guar gum) are salts, so attract water and provide a conductive pathway along the surface of the hair. So they provide both anti-frizz and anti-static properties. Although hair nearer the root is neutral as it hasn't been damaged, nearer the tip the hair contains more anions (e.g. sulfates from oxidation of cysteine groups in the keratin) so the quats are nicely attracted to these damaged areas, giving the chance to provide "body" to the weakened fibres as well as a degree of smoothing of the roughened surface.

They can be delivered via Surfactancy_Coacervation, the concentration-dependent solubility/insolubility of complexes of the polyquats with anionic surfactants such as SLES. Unfortunately, the science of coacervates remains rudimentary and there seem to be no useful guides as to how to formulate these effects rationally. They can clearly have a big effect on wormlike micelles – they can destroy them by removing surfactant, or create complex equivalents. Again, disentangling these effects scientifically seems to be beyond our knowledge.

Polymers for body

It seems irrational to want to clean junk from the hair and replace it with other stuff intended to stiffen the hair. But especially for those who use a lot of hair care products, the damage caused to the hair by the aggressive cycles of chemicals, hot water and hair dryers is enough to require repair work using a coating of polymer.

The polyquats can do some of this. Other uncharged polymers might also perform this function as long as they don't act as glue between strands. For those who can't use salt thickening for their surfactant formulation, a "body" polymer might also provide the required thickening.

Although not as strong as the interactions of polyquats with anionics, polymers can show significant interactions with surfactants so a well-intentioned addition of a polymer might have a bad effect on other subtle properties of the formulation such as thickening. The extended discussion in the Coacervates chapter covers some of these issues.

Fragrances

The general challenges of formulating a fragrance, including avoidance of flavour scalping through the packaging, are discussed in the Fragrances chapter. With so much surfactant in a shampoo, getting the relatively low quantities of fragrances into the formulation is not much of a problem.

As with the other ingredients, the effects on the liquid crystal (wormlike) phases created by salt thickeners can be subtle and difficult to diagnose and formulate around.

Probiotics

As with Soaps and Washing, after years of selling products which aggressively interfere with the skin and scalp biomes (to use a convenient shorthand term), industry wants to sell products to restore our systems to their optimal natural balance.

The “no-poos” who don’t used shampoos claim that it takes a few weeks for their systems to recover to a natural balance, with their biome then doing a good job of caring for their hair. There’s an obvious selection bias here as this doesn’t work for those who try, but fail, to go shampoo-free. Who knows whether their natural biome isn’t great for their hair, or whether the biome that evolves is out of balance because of the history of shampoo use.

Whether the addition of one or more “pre-biotics” (food for your microbes) or actual biotics (live microbes) can create a biome as natural and healthy as yours (presumably) would have been without lots of shampoos is a tricky question. Even harder for the manufacturer is making a solid claim. Use vague terms, with no proof of efficacy, and people might buy the product. Make precise claims, based on extensive safety and efficacy trials, and it’s probably a pharmaceutical product.

Preservatives

For any aqueous product that contains ingredients that are nutritious to microbes (and that very much includes the surfactants used in shampoos as well as the oils, vitamins etc. added for whatever purposes) there needs to be something in the formulation to stop the microbes from growing. By definition, these preservatives have a negative effect on living cells, which means that it’s highly likely that some of them could have some negative effect on some human cells. It makes no difference if the preservative is entirely synthetic or is found in nature – it has to have some bad effects on living cells, that’s its job.

Formulating a preservative into a shampoo is not hard. Proving its efficacy involves tedious standard tests, but they are routine. The hard thing is finding a way to say that your shampoo contains ingredients that are harmful to living cells without scaring the customers who often believe that all “chemicals” are harmful and only “natural” things are good.

The only way out of this dilemma is to have a formulation with insignificant amounts of water, as discussed soon.

Ingredients for marketing claims

If your product proudly claims that it contains extract of mountain blossom and has a picture of a mountain to prove it, and if, indeed, some mountain blossom has been extracted and added to the shampoo, then, assuming you know that mountain blossom contains no harmful chemicals, your product can be sold. If consumers think that extract of mountain blossom must be good (otherwise, the logic goes, you would not be proudly adding it), then you are probably OK. Only “probably”. Authorities are so irritated by marketing claims designed to fool the customers that they have become more strict even on things that merely imply, but don’t explicitly state, some benefit to the consumer that cannot, in fact, be substantiated.

As a formulator, your job is to add what marketing want you to add. As long as extract of mountain blossom is present only at the minimum that allows the marketing claim, it should pose no real formulation problem.

The right packaging

Sadly, well-controlled experiments using either bland, industry-standard shampoo formulations in different packaging, or bland versus expensive formulations in standard packaging, have shown that consumers overwhelmingly rate the effect of the shampoo according to their expectations from the packaging. It might be distressing to the formulator that all those extra ingredients added to their top-of-the-range product are not detectable to the average consumer, but it's what is expected from the science.

If you think that I'm making up these assertions, get hold of a copy of the Robbins *Chemical and Physical Behaviour of Human Hair*. The author has a lively sense of what is good science and what is mere marketing.

Dry shampoos

Dry shampoos mean you don't have to ship large amounts of useless water, with the packaging and transport costs the water requires. They can also be free of preservatives as microbes are unlikely to grow in an environment largely free of water.

There are two broad types.

- Absorbent powders such as starch that attract excess oils and can then be combed out of the hair, leaving it feeling OK. The tricky formulation problem is getting the right Particles_Size distribution. If the particles are, on balance, too small then the product is a dust hazard and hard to comb out of the hair. If, on balance, they are too large, then they feel "gritty" and the surface area to volume ratio is too small to be an effective remover of oil.
- Shampoo bars, which are generally old-fashioned soap with whatever ingredients (e.g. excess glycerol from the manufacture or excess oil) make it less alkaline and aggressive, plus the usual fragrances and colourants for a feel-good factor.

Neither presents a tough formulation challenge, though they are each a marketing challenge as is the challenge of a soap dish that is convenient and keeps the bar relatively dry without falling into the sink or shower.

Inkjet Inks

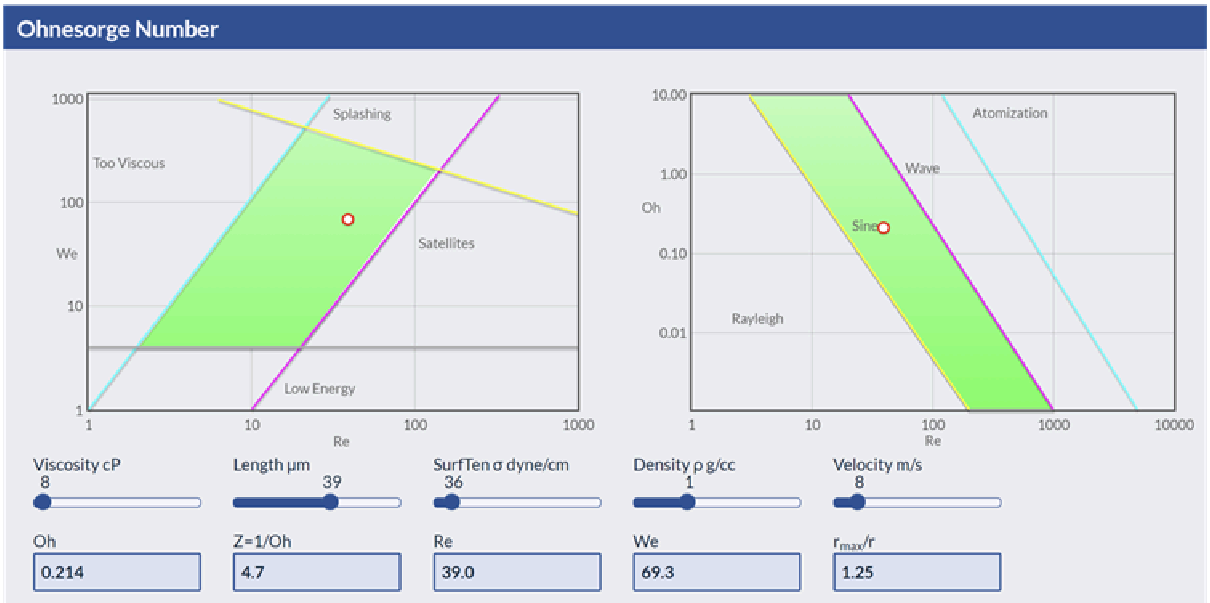
Links

Absorbency_Porosity and Darcy, Dispersions_DLVO, Dispersions_ODC, Dispersions_Rheology (Low shear), Dispersions_Zeta potential, Evaporation_Basics, Flow_Ohnesorge, Flow_Rotational rheology

> Although it is easy to specify and create a workable inkjet ink, the frustration is with the difficulty of getting perfect drops, infinite lifetime as a stored ink and zero problems within the head and on the nozzle. Following everything in this chapter will get you a workable ink, the rest is hard work.

The limits of Ohnesorge

As discussed in Flow_Ohnesorge, you can’t get a good inkjet drop if the combination of drop size, velocity, viscosity and surface tension takes it outside the core area defined by Ohnesorge, Reynolds and Weber numbers:



Viscosity cP 8
Length µm 39
SurfTen σ dyne/cm 36
Density ρ g/cc 1
Velocity m/s 8

Oh 0.214
Z=1/Oh 4.7
Re 39.0
We 69.3
r_{max}/r 1.25

https://www.stevenabbott.co.uk/practical-coatings/ohnesorge.php

For the velocities and (small) drop sizes from typical commercial inkjet heads, the limits of surface tension are modest, maybe 30-60 mN/m is fine, with lower tensions more problematical for interactions with the nozzle plate than for Ohnesorge. The limits on viscosity are tougher 5-25 mPa.s are often quoted, with something like 15 mPa.s being a common target. Outside these ranges the drops can’t form or satellite drops are formed excessively.

For typical UV inkjet formulations, at room temperature the viscosity limit is tough to achieve from the limited set of usable UV monomers. Heads routinely run at 40°C to take advantage of the significant drop in viscosity with temperature that is typical of acrylates.

For water-based formulations, the equilibrium surface tension is a worthless value, what matters is the dynamic surface tension (Surfactancy_Dynamic Surface Tension) measured, typically, via the bubble pressure tensiometer.

In principle, viscosity should be judged using a proper rotational viscometer Flow_Rotational rheology to check for properties over a range of shear rates. In practice most reasonable inkjet inks have happily dull rheologic properties at the rates and strains that can be measured. What's happening at the super-high extensional flows coming out of the nozzle is outside the capabilities of this book.

Basic knowledge of particle viscosities at low shear, Dispersions_Rheology (Low shear) can help understand the modest effect of particle volume fraction up to, say, 30%, with suitable caution that a change in viscosity of the *base* formulation, without particles, from, say, 5 to 7 cP, which is a trivial 2 cP change will take a 30% formulation from 13 to 19 cP (13 * 7/5) which is a significant jump for an inkjet ink.

The basic particle viscosity curve assumes no attractive particle-particle interactions. Getting the right dispersant is vital, using either the Dispersions_ODC approach for non-aqueous systems or standard Dispersions_DLVO for aqueous ones.

Another constraint on formulation comes from subtle internal issues of specific heads. Having read the specs of a specific head, we created a reactive ink that delivered exactly the desired effect … until the head suddenly stopped working. An enquiry to the head manufacturer led to them asking about the pH of the ink. "Ah, we failed to put into the spec that the pH must not be below 5.5". The chemistry of this specific ink required a pH of 5. A project that had so much promise had to be abandoned, thanks to a pH difference of 0.5.

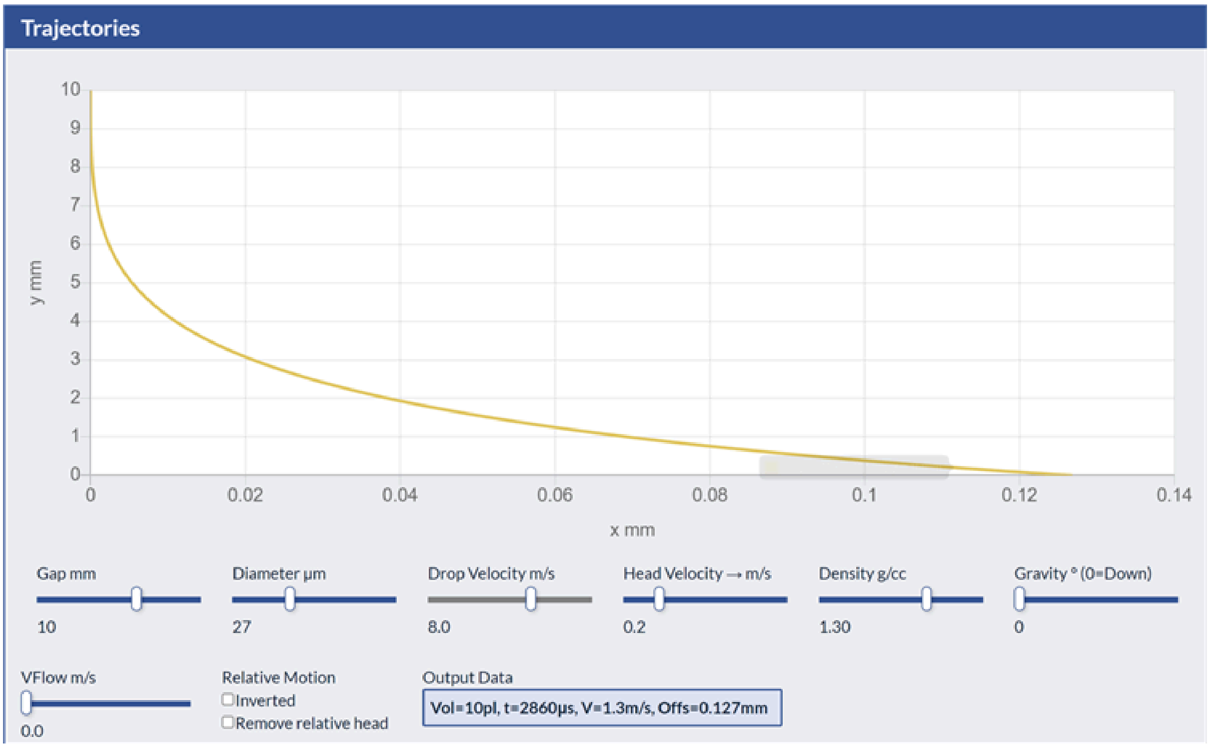
Early drop dynamics

Assuming you are in the right Reynolds/Weber/Ohnesorge range, success in drop behaviour is not guaranteed. Complex interactions while leaving the nozzle, and complex fluid dynamics as the stretched-out drop pulls in on itself can cause many wild things to happen. There seem to be many observations via high-speed drop watchers and plenty of computer simulations, but nothing appable that can help the formulator. So let's assume you have a nice drop with, at worst, a single satellite following harmlessly directly behind (some satellites shoot off to the side, which is most undesirable).

Flying through the air

There are three issues of interest about the drop's flight from the nozzle to the substrate:

1. By how much it gets slowed down by air resistance
2. By how much it is influenced by gravity (answer, the effect is insignificant)
3. By how much it is deflected from its straight path to the substrate by the fact that the head is moving in the cross-web and along-web directions.



Gap mm 10

Diameter µm 27

Drop Velocity m/s 8.0

Head Velocity → m/s 0.2

Density g/cc 1.30

Gravity ° (0=Down) 0

VFlow m/s 0.0

Relative Motion
☐Inverted
☐Remove relative head

Output Data
Vol=10pl, t=2860µs, V=1.3m/s, Offs=0.127mm

https://www.stevenabbott.co.uk/practical-coatings/Trajectories.php

The app answers all 3 questions, but the gravity and deflection issues are not relevant to this book. What is important is that a typical 10pl drop (you define drop diameter and its volume is calculated for you) coming out of the head at 8 m/s lands on the substrate after a 10mm gap at 1.3m/s.

Contrary to instinct, to a typical drop, hitting a surface at 1.3 m/s is a gentle landing, so there’s no dramatic splashing. Instead it quickly forms a hemispherical cap of $\sqrt[3]{2}$ the diameter of the full drop, in this case a 27 µm drop becomes 34 µm on the surface. What happens after that needs another app.

Drop spread

Drop● Wedge○ Cylinder○

Viscosity Pa.s	ST Dyne/cm	Radius or Wedge µm	θ_{Start}°	$\theta_{Equilibrium}$°	
0.01	50	17	90	0	
T_{Max} s	R_{Max} µm	Height µm	Width µm	Volume	End Rad. µm
1	50	15	25	10.3pl	66.8
Start Height µm	End Height µm	Start µm	Start Angle°	End Angle°	End µm
17.0	1.5	34.0	n/a	n/a	n/a

https://www.stevenabbott.co.uk/practical-coatings/drop-spread.php

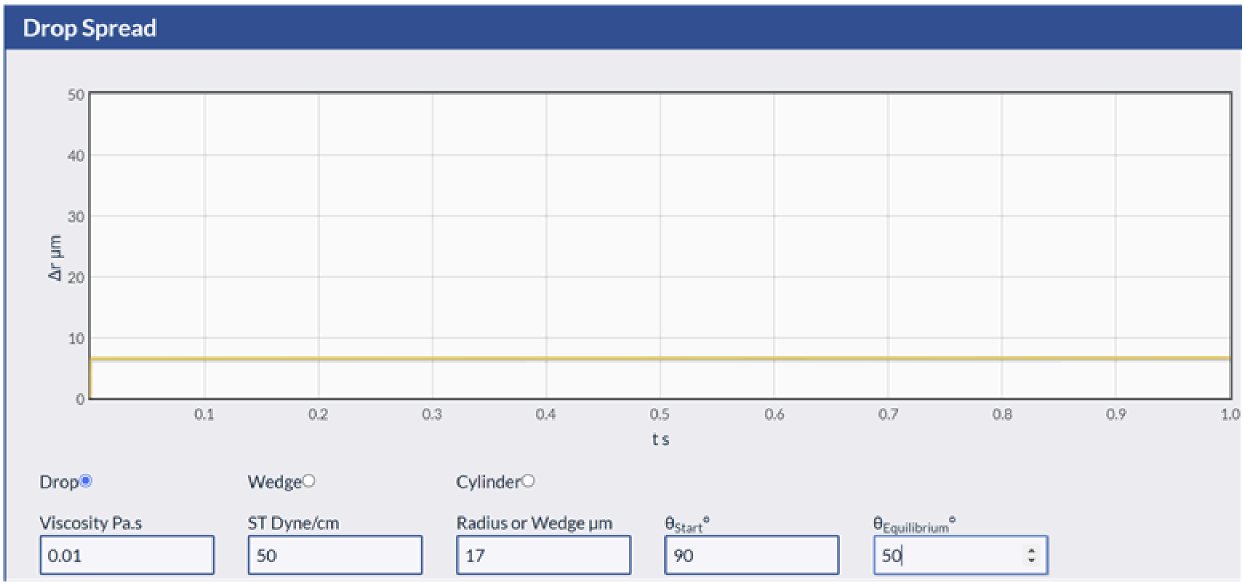
What is scary about drop spread is how fast it is. We assume a 90° contact angle as the drop hits the surface and we've set the starting radius to 17 µm. Within 50 ms (you can use your mouse to confirm this) the drop has spread to 92 µm diameter. After 1 s it is 134 µm. As explained in the app, the dramatic fast early speed is due to the fact that for a contact angle of θ, surface tension σ and viscosity η, the spreading velocity is:

$$v = \frac{\theta^3 \sigma}{\eta}$$

That θ^3 dependency is very harsh if you want to avoid drop spread.

Given the limits of surface tension and viscosity for an ink, drop spread cannot be seriously controlled via the ink. It's the substrate that matters most.

If you make sure that the equilibrium contact angle of the ink is, say, 50° then you find:



Drop ● Wedge ○ Cylinder ○

Viscosity Pa.s: 0.01

ST Dyne/cm: 50

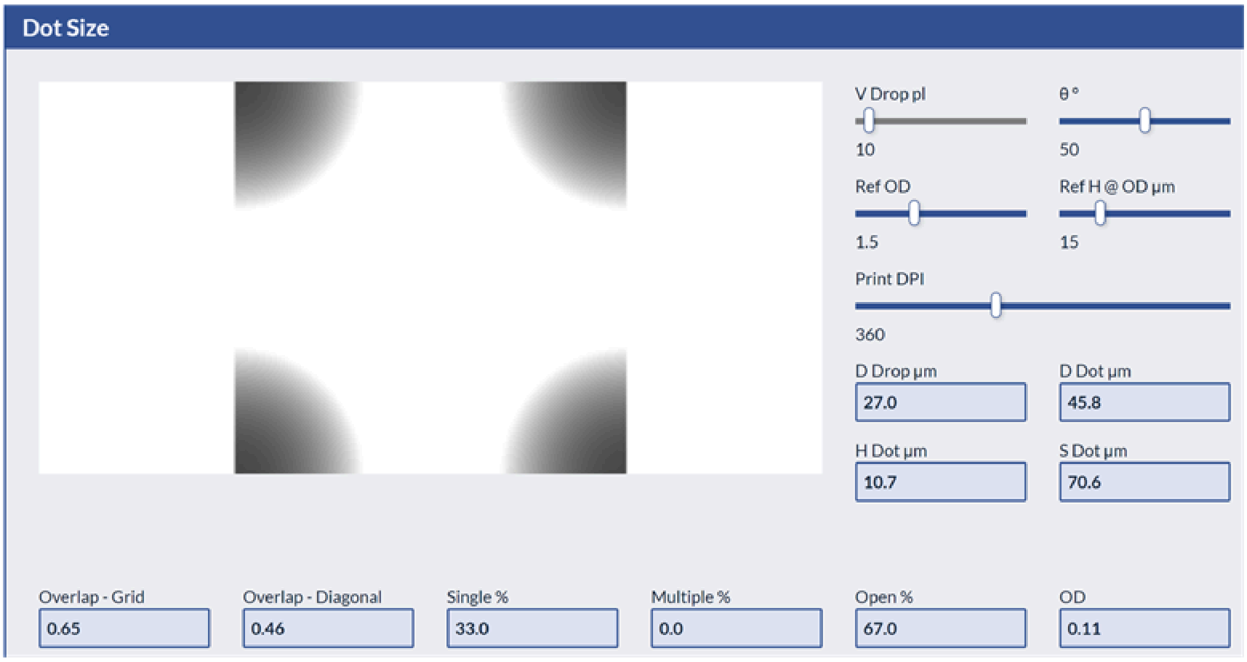
Radius or Wedge µm: 17

θ_{Start}°: 90

$\theta_{Equilibrium}$°: 50

that the drop goes near instantly to 48 µm, and stays there pinned at the 50° limit. Inkjet printing is one of the rare portions of formulation science where contact angle is of crucial importance.

You can explore the consequences of this via another app:



V Drop pl: 10

θ°: 50

Ref OD: 1.5

Ref H @ OD µm: 15

Print DPI: 360

D Drop µm: 27.0

D Dot µm: 45.8

H Dot µm: 10.7

S Dot µm: 70.6

Overlap - Grid: 0.65

Overlap - Diagonal: 0.46

Single %: 33.0

Multiple %: 0.0

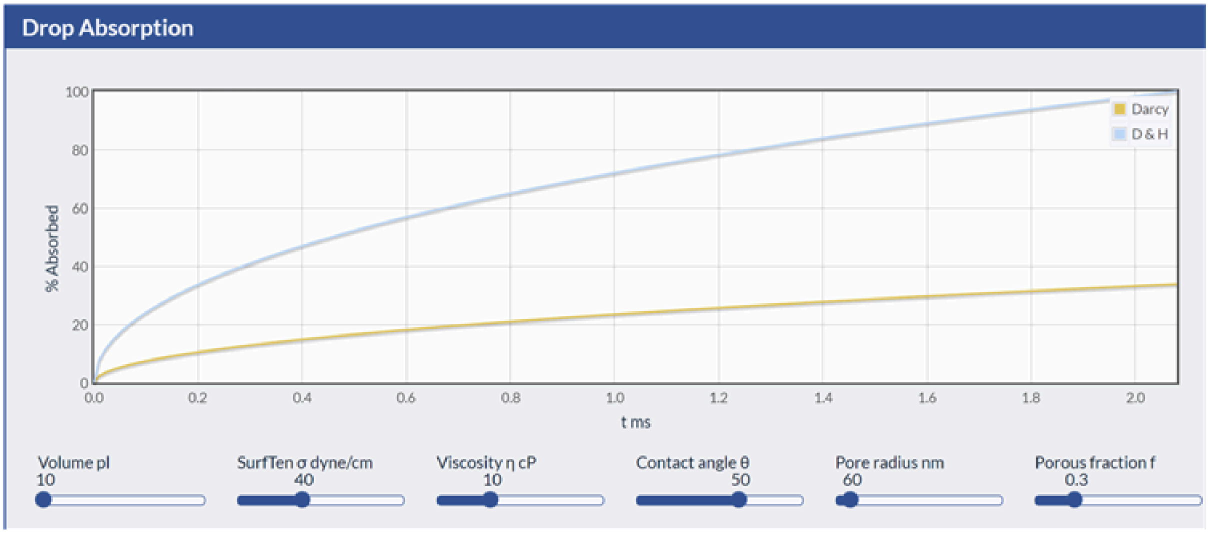
Open %: 67.0

OD: 0.11

https://www.stevenabbott.co.uk/practical-coatings/Dot-Size.php

Here you can explore the impact of final drop size on the optical density (OD) of the print. Too little dot gain (contact angle too high) and the dots don't join up. Too much and your print becomes a mess.

The other way to stop drop spread is to absorb the ink into a porous substrate:



Volume pl 10

SurfTen σ dyne/cm 40

Viscosity η cP 10

Contact angle θ 50

Pore radius nm 60

Porous fraction f 0.3

https://www.stevenabbott.co.uk/practical-coatings/drop-absorption.php

The standard Absorbency_Porosity and Darcy model seems to over-predict the sorption (70% in 1 ms). The Davis & Hocking version predicts only 23%, but that's still going to have an impact on drop spread as there may well be some pinning implications.

Both Darcy and D&H depend on the distance, d, travelled in time t into the porous medium with pore radius r.

$$d = \sqrt{\frac{r\sigma\cos(\theta)t}{2\eta}}$$

D&H tells us that the volume, V, absorbed depends on the starting radius of the drop, R via:

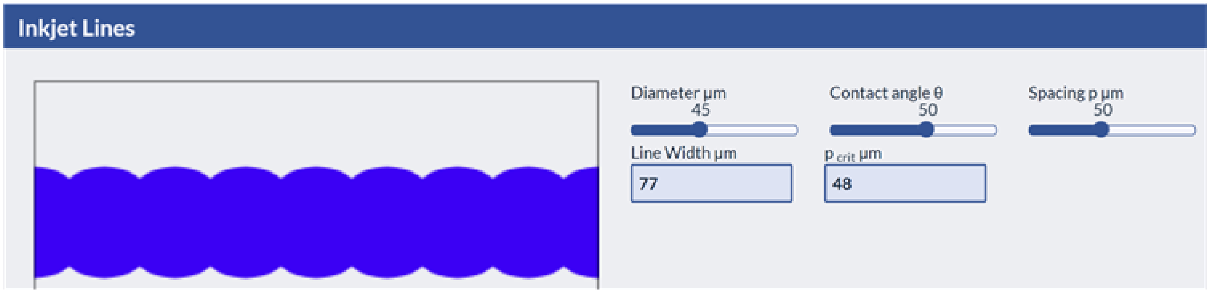
$$V = \frac{\pi R^2}{2}\left(2d - \frac{d^2}{D}\right)$$

It is worth noting that for a nanoparticle ink, the chances are that the particles will block the pores, so absorption stops very quickly.

There are so many subtle ways for drops to fail to arrive at the desired location that naïve dreams of "printed electronics" via inkjet have regularly failed. I have been astonished that so much funding went to so many companies that stood zero chance of success. The one success story is OLEDs, but the printers cost in the $10's of millions. Those with printers at home note the large amount of ink consumed during "necessary head cleaning cycles". It's not that the printer manufacturers are especially greedy, it's just that they are paranoid about losing a single nozzle which may lead to losing a customer.

Dot join-up

Given the interesting calculations of drop spread, the next logical question is about dot join-up. It turns out that the rule is trivial: if the drops touch then they join and it's essentially instant. That's frustratingly imprecise, but seems to be the case. One interesting calculation involves the task of writing a nice straight line:


Inkjet Lines

Diameter µm 45

Contact angle θ 50

Spacing p µm 50

Line Width µm 77

p_{crit} µm 48

https://www.stevenabbott.co.uk/practical-coatings/InkJetLines.php

The screenshot shows that these dots have joined to create an uneven pattern. With some variations in spacing, drop size and contact angle you can get separate dots or the desired uniform straight line.

Is that it? I’ve tried hard to find more examples of reliable inkjet science we can all use, but it doesn’t seem to exist.

Laundry Liquids

Links

Cleaning_Bleaching and Oxidation, Cleaning_Boundary removal, Cleaning_Contact Angles, Cleaning_Enzymes, Cleaning_Surfactants, Cleaning_Temperature effects, Dissolution_Solubilizers and Hydrotropes, Flow_Basic viscosity, Fragrance_Activity Coefficients, Fragrance_Barrier Properties, Fragrance_Vapour Pressure, Surfactancy_Anti-foaming, Surfactancy_CMC and Langmuir, Surfactancy_Dynamic Surface Tension, Surfactancy_Foaming, Surfactancy_HLD, Surfactancy_Interfacial Tension and Rigidity, Surfactancy_Phase Diagrams, Surfactancy_Roll-up and Eötvös Number

Modern laundry detergents are based on complex trade-offs of science, cost, environmental footprint and marketing claims. They are over-engineered for their typical requirements which are to freshen up hardly-soiled clothes, but must give satisfactory results with any tough soil that the user happens to provide. Much of what is discussed applies to solid laundry powders, but their formulation challenges are relatively easier, so the focus here is on the more challenging liquids.

A basic wash

Given some soft domestic water and a typical load of hardly-soiled clothes, just about any low-foaming surfactant in the fairly hydrophobic range (discussed with more precision shortly) will render the clothes clean by any rational standard. Any hydrophilic dirt will be removed by the water and typical hydrophobic dirt particles will attract the tail of most surfactants, leaving a sufficiently hydrophilic head sticking out to keep the particle suspended. The difficult challenge of removing oil on the surface is discussed below. The surface tension reached by the surfactant in an ideal lab test is irrelevant to what is found in a washing machine. It's nice to have a lower surface tension, but whether it's 36 or 41 mN/m is of no great significance.

The low-foaming need is paramount. Consumers have been conditioned to associate "foaming" with "cleaning", which is irrational. In a washing machine, there is no need to impress the consumer with foam and the formulation can deal with the obvious fact that foaming interferes with cleaning by tying up surfactant, impeding free flow of cleaning water, and overflowing the machine.

Of more importance than the surfactant is the fact that overcoming the no-slip boundary condition (Cleaning_Boundary removal) is necessary to get *any* cleaning. This all comes down to the amount of agitation in the machine which depends on drum design, orientation (vertical drums are far less efficient), on the relative volumes of fabrics and water, and on the relative absence of foam.

To cope with hard water, a typical chelator such as EDTA or an environmentally friendlier one will keep the calcium/magnesium salts from interfering and "builders" such as phosphates, silicates or polyacrylics are included to help with chelation and tie up ions such as iron or copper that can discolour the wash or the washing machine.

Such a product would have a minimum environmental footprint and would be low cost. It would also be unsellable.

What consumers want

The first requirement is to mask the usual “chemical” smell from traces of surfactants and any smell of damp clothes. So it needs a fragrance (Fragrances). A typical fragrance molecule will be nicely soluble in the concentrated detergent and will happily dissolve in the washing water … to disappear down the drain unless is strongly partitions onto or into the fibres of the clothes, something that is hard to guarantee. Via Microencapsulation most of the formulation problems disappeared as the capsules are readily trapped by just about any fabric and their slow release of fragrance via leakage or via rupture through handling of the fabric gives the desirable long-lasting properties. The transition from microcapsules that produce microplastics to ones that are fully biodegradable is happening as this book is being written.

The next requirement is that whatever soil is on any fabric will just disappear in the wash. To make this happen we need the following:

- A surfactant that can produce efficient roll-up of oils and grease, Surfactancy_Roll-up and Eötvös Number.
- Optionally a lipase enzyme to break up fatty acids and therefore reduce the need for the surfactant to do all the removal of fats. It’s optional because a formulation that doesn’t then remove the resulting fatty acids can have a slight “sweaty” odour.
- A protease enzyme to break up protein stains including blood and grass.
- An amylase to break up starch residues, usually from food stains.
- Optionally a cellulase to remove cellulose microfibres that make the fabric look “dull”. The removal also removes and dirt or stain that’s on those microfibres.
- A bleaching system (Cleaning_Bleaching and Oxidation) that magically destroys all coloured stains (including wine, blood, tea/coffee, grass) while never harming any of the fabrics and never causing fade to any of the colours of the fabrics themselves.
- A high alkalinity is great for destroying many of the soils but risks damaging the fabrics. A modest alkalinity (often provided via the package of builders) helps a little and is especially good for solubilizing the fatty acid odours from stale sweat and those released by lipases.

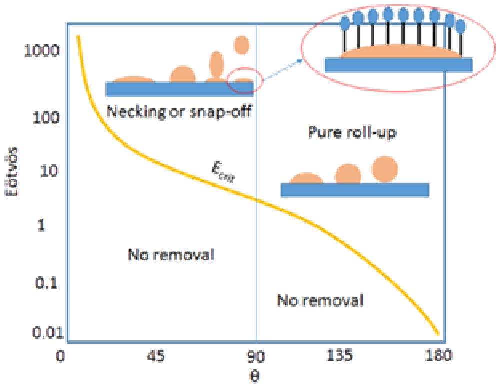
Consumers also want their detergents to be chemical free, natural and low price. And although the most natural parts of most formulations are the enzymes, many consumers want “bio-free” formulations, recognising that just because something is natural doesn’t mean that it’s harmless. A mix of lipase, protease and amylase sitting on your skin could, indeed, result in plenty of damage.

The obvious fact that all residues of surfactants, enzymes and bleaches can be removed via thorough rinsing meets the requirement to reduce water usage during a wash. As is hopefully well known, 3 rinses with 1/3 the amount of water is better than 2 rinses with 1/2 the amount, which is more efficient than one rinse with the full amount – hence the trend to more rinse steps with a good spin after each rinse. Smart washing machines can adjust the volume of water and number of rinses based on various sensors (such as turbidity), but formulators need to formulate for dumb machines too.

The right surfactant system for oil removal

As we know from general cleaning, the choice of surfactant isn’t of great importance – you can clean most things adequately with SLS. But for removing oils from fabrics you need to be in the right part of surfactant space, as defined by HLD (Surfactancy_HLD) to achieve the necessary super-low interfacial tension γ (Surfactancy_Interfacial Tension and Rigidity) between water and oil which gives you a large Eötvös number:

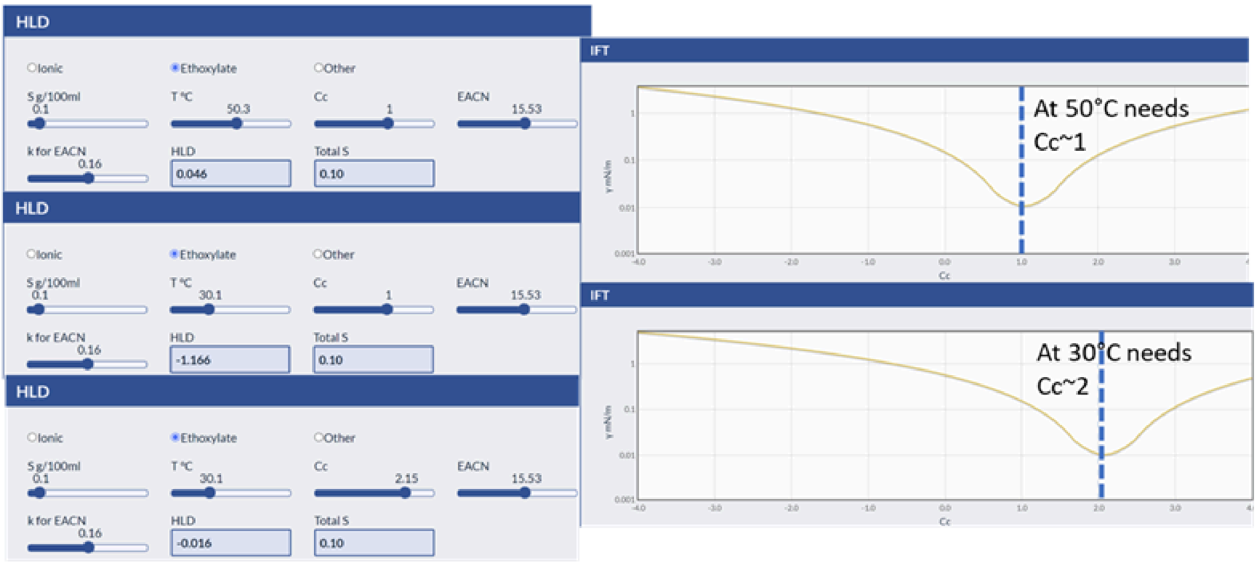
$$E = \frac{\Delta\rho g R^2}{2\gamma}$$



The interfacial tension becomes super-low (100x lower so Eötvös is 100x larger) only when the Hydrophilic-Lipophilic Difference is close to zero, and because HLD depends on the oiliness of the oil (its EACN), the salinity of the formulation (all those chelators/builders that are added), the temperature of the wash and the characteristic value, Cc, of the surfactant blend, the focus for the formulator is to get the Cc right.

Under any typical set of values for oil, salinity and temperature, you need relatively high Cc (hydrophobic) surfactants which, happily, tend to be poor at stabilizing the water-air interface so give the low foaming necessary for a washing machine.

The temperature dependence of the much-used non-ionic ethoxylates was a key aspect of making them efficient at typical 40-60°C washes that were common in the past. In the example shown here, an ethoxylate with a Cc of 1 gives excellent washing at 50°C because HLD ~0 while at 30°C it is useless because HLD is ~-1. To work at 30°C needs a surfactant with Cc ~ 2.


HLD
○ Ionic ● Ethoxylate ○ Other
S g/100ml 0.1
T °C 50.3
Cc 1
EACN 15.53
k for EACN 0.16
HLD 0.046
Total S 0.10

HLD
○ Ionic ● Ethoxylate ○ Other
S g/100ml 0.1
T °C 30.1
Cc 1
EACN 15.53
k for EACN 0.16
HLD -1.166
Total S 0.10

HLD
○ Ionic ● Ethoxylate ○ Other
S g/100ml 0.1
T °C 30.1
Cc 2.15
EACN 15.53
k for EACN 0.16
HLD -0.016
Total S 0.10

https://www.stevenabbott.co.uk/practical-surfactants/hld.php & https://www.stevenabbott.co.uk/practical-surfactants/ift.php

The problem now is that consumers might use a low temperature wash at a high temperature and the interfacial tension at 50°C will now be ~1 mN/m instead of 0.01.

There are two broad approaches to solving this issue.

1. Create a temperature-insensitive combination of ethoxylate and ionic surfactants. Because the T effect of an anionic is 6 times smaller, in the opposite direction, a suitable anionic-ethoxylate ratio might give adequate temperature independence. Unfortunately, simple HLD cannot be used for this because anionic/ethoxylate mixes show non-ideal behaviour. Instead, a careful experimental program of HLD scans is required.
2. Use something like an APG or polyglycerol formulation because these have very low and low T

dependence respectively. Finding effective packages of these that are in the right HLD space and have the required efficiency is a challenge.

Anti-foams for detergents

Even though efficient, lower foaming surfactants can be used, there is still a general requirement for some Surfactancy_Anti-foaming. Interestingly, the challenge isn't so much to find a really good anti-foam, but rather one that does only a modest job – and that is a difficult challenge. Scientifically a good anti-foam is perfect for washing. Unfortunately, consumers might see an absence of foam through the window of their machine and think that the detergent must be useless. Had the marketing departments of megacorps not wasted decades persuading consumers that foam = clean, we could have better formulations for most of our consumer products. Given the expectation of *some* foam being visible, finding a reliable not-quite-good-enough anti-foam is a frustration for the detergent formulators and suppliers of anti-foams. In the end, it comes down to trial-and-error and "experience", both of which can be upended when, say, a "bad" sulfate-based surfactant is replaced by a "good" sulfate-free equivalent. Worse than this is when marketing want a new fragrance for the product. As happens so often in other parts of personal care, 0.1% of a new fragrance can wreck an anti-foam formulation that worked wonderfully for 0.1% of a different fragrance.

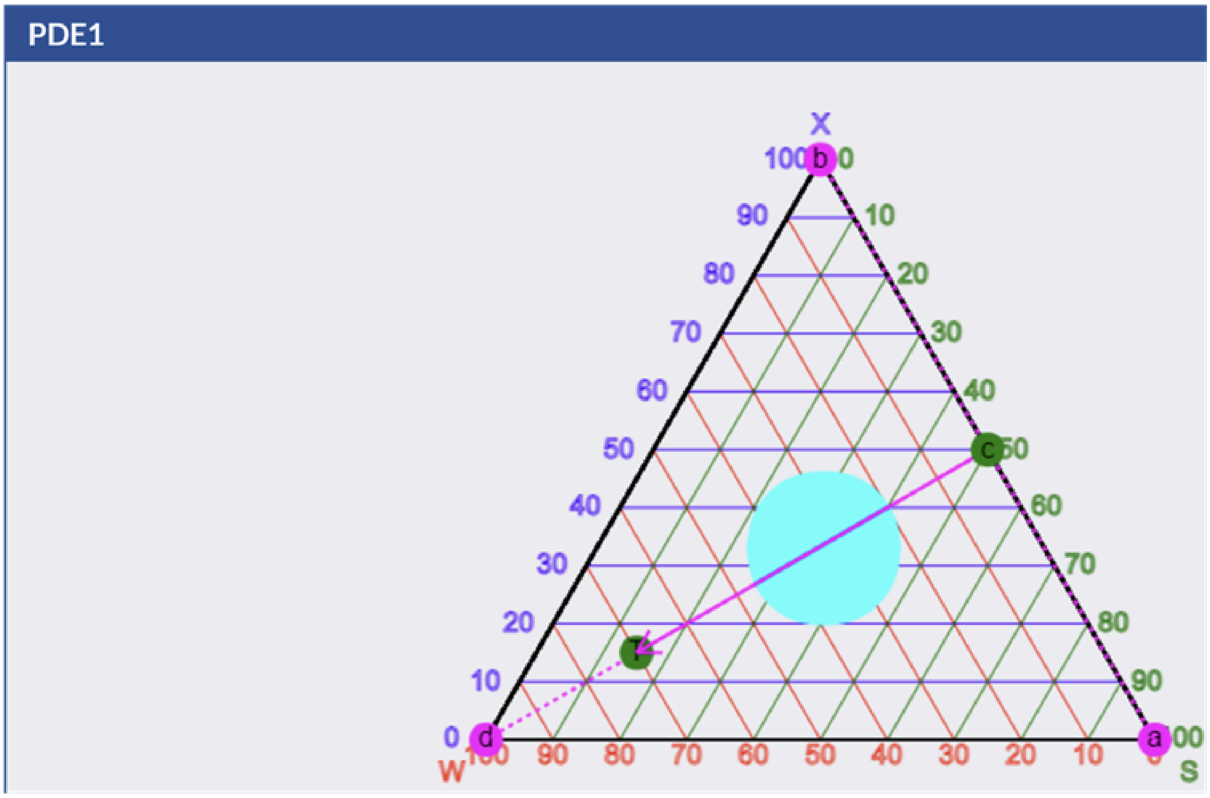
Why can't we design (or purchase) anti-foam packages that are right first time? Read the anti-foam chapter.

The right surfactant for a concentrated product

Concentrates have become the norm, meaning that less water and packaging are required per wash. There are two problems with concentrates – of which the first is obvious and the second is hidden, requiring more work:

1. Some surfactants get too viscous at the desired concentration so they cannot be used without whatever ad hoc tricks destroy whatever strong self-association is causing the high viscosity.
2. Some surfactants are fine at high viscosity, fine (of course) at low viscosity, but catastrophic at the intermediate concentrations that arise when the concentrate is diluted in the washing machine. The problem is formation of cubic or hexagonal phases which are semi-solid. As there is no way to predict these (ideas around Critical Packing Parameter (Surfactancy_CPP and phases) are far too naïve for real-world detergent formulations), you have to do the work to explore the phase diagram (Surfactancy_Phase Diagrams). As the formulations tend to be complex, a ternary phase diagram is often needed.

You might find something like this:



https://www.stevenabbott.co.uk/practical-surfactants/pde1.php

The concentrate on the X:S axis to the right is OK, the diluted formulation at the end of the arrow is OK, but the cyan blob represents an hexagonal phase which will result in blobs of surfactant floating in the wash and ending up on the clothes at the end.

Obtaining a phase diagram is often seen as difficult, but the https://www.stevenabbott.co.uk/practical-surfactants/QPD.php page, courtesy of Seth Lindberg at P&G shows that you can obtain the relevant information relatively easily.

The trade-offs

Hot water makes enzymes and bleaches work faster and the kinetics, if not the thermodynamics, of surfactancy are improved. Soluble dirts are more soluble, tough greases start to melt so the only downside, which is overwhelming, is the combination of direct energy costs for the user and the impact on the planet.

With enough effort, enzymes can be tuned to be efficient at cold water temperatures and catalysts can encourage the peroxide bleaches to work faster … or you just add more of them, with an obvious cost penalty. If the temperature isn't high enough to melt some grease, maybe the lipase will be able to break it up.

A longer wash cycle will be especially useful for the enzymes and bleaches, but unappreciated by the busy end users.

If the surfactants, bleaches, enzymes, chelators/builders and fragrance package interact in some unfavourable ways then one solution is to package the individual components inside a "pod". The pods are made from a water-soluble polymer such as PVOH which will hopefully dissolve quickly. Unfortunately, pod residues can sometimes get caught in the fabrics and, because they are denied the flow of water required to dissolve them completely, end up as gunk on the clothes. This can happen if the PVOH dries out during storage, making it less responsive to the first seconds of contact with the water. One potential fix is to add Evaporation_Humectants that keep a fairly stable amount of water even if the local atmosphere is of low RH, though the app throws some doubt on the standard explanation for the efficacy of humectants.

Lipstick

Links

Adhesion_Adhesion promoters, Adhesion_Surface energy Interactions, Cleaning_Boundary removal, Cleaning_Solvents, Cleaning_Surfactants, Cleaning_Temperature effects, Diffusion_Barrier Properties, Diffusion_Basic Diffusion, Diffusion_Diffusion into Skin and Hair, Diffusion_Permeation OTR and WVTR, Dispersions_ODC, Dispersions_PVC and CPVC, Dispersions_Rheology (Low shear), Dispersions_Settling, Flow_Basic viscosity, Flow_G' and G", Flow_Oscillatory rheology, Flow_Rotational rheology, Flow_TTS-WLF, Flow_Yield Stress, Optics_Colour

> Lipstick contains its own contradiction: the colour should be indestructible on the lips, never transfer to an object such as a cup or glass, yet be fully and easily removable when required, leaving no stain/residue on the lips.

The basic formulation

For this chapter we will assume that the pigment is a given and that it is the optimal size for hiding power (too small and there’s insufficient scattering, Optics_Scattering and Opacity) and colour intensity (too large and the interior of the particles adds no extra absorption) and has an excellent dispersant (Dispersions_ODC) optimized for the oil/wax environment. The pigment might be ground using one or more of the oils as a convenient medium.

A possible exception to this is that the common mica pigments may be providing global properties such as slip, barrier or strength properties beyond their colourant, gloss and sparkle effects. If these extra effects are important, you can include them in the formulation space mapping discussed at the end of this chapter.

The wax component provides the necessary semi-solid form at room and skin temperature (conventionally said to be 32°C) while being meltable at a reasonable temperature (conventionally 55-75°C) to be liquid for combining the ingredients and then pouring into a mould to produce the desired shape.

The oils provide the desired softness, emollience and general feel-good effects.

Melting together some beeswax, coconut oil, shea butter and just about any hydrophobic colourant gives an adequate lipstick when poured into a tube and allowed to cool.

A suitable fragrance and an antioxidant such as BHT or vitamin E (tocopherol) to stop the oils being oxidised completes the basic formulation.

Even without extra functional ingredients, discussed below, the formulator is faced with a daunting formulation challenge. There are many waxes and oils, and they interact with each other in multiple ways, especially in terms of the crystallinity of the waxes as the lipstick cools in the mould. Inevitably this means we need to use Differential Scanning Calorimetry (DSC) to get core information and have at least some idea of crystallization theory.

The subjective feel of the lipstick during and after application is too subtle to be properly analysed in the lab, but we have to find measurements that at least correlate, so that means rheology.

We address each of those in turn, after first thinking about what should be tested.

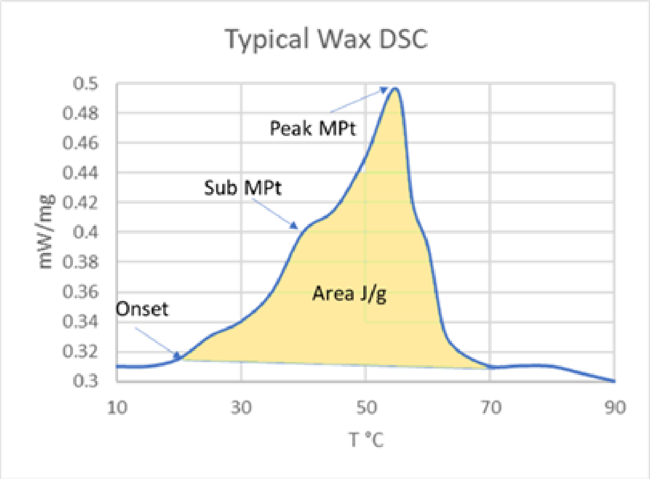
> This was a difficult chapter to write. It is very easy to go blah blah with truisms, but that doesn't help. There are no easy, relevant formulae to provide an app-route through formulation space. Instead we identify those measurements that *must* be relevant to the formulation, then work out how to map those data against real-world behaviour of products that are known to be good or bad, your own and competitors. This is an unfamiliar approach, but a powerful one.

Core data

It is tempting to throw together some typical formulations and measure the key data. Although, as discussed below, it's important to measure some real formulations, especially the low, medium and high-end competitors' formulations, to get out of the habit of adding too many ingredients too soon, it is necessary to get some core data on simpler sub-components.

So choose a few waxes (ones favoured by marketing for their claims) and a few oils (similarly chosen by marketing) and gather pair-wise data to see what is affecting what in what ways. For the future digitalisation of your product development, you can build on core data from key ingredients. It is hard to build on the data from complex mixtures selected for historical reasons such as "we've always done it this way".

DSC, MPts etc.



The DSC plots the mW/mg of sample needed to raise the temperature by 1° at the given heating rate, °/s. Without melting, the value (when multiplied by the timestep) is simply the heat capacity of the material. The latent heat of fusion is responsible for the large increase in power needed to maintain the heating rate. Even a "pure" wax melts over a significant temperature range and a typical mix of waxes for a lipstick melts, as here, over wide range. This example shows two peaks, nominally for two types of waxes. In reality, the number of peaks found depends strongly on heating rate – the slower the rate the more individual peaks can be found. The integration gives us the total latent heat of fusion, in J/g. A fully amorphous wax would show a small peak, the fully crystalline equivalent would show a large peak, so the area is used as some indication of the degree of crystallinity.

Any single graph should be treated with caution. This "heating" curve will be different from the "cooling" curve because not only is there supercooling (so the MPt peaks will be at a lower T), there is also a different degree of crystallization. Both of these will be cooling-rate dependent. Scientifically we might measure "isokinetic" cooling where the DSC cools the sample at a fixed rate. In production the cooling is not only non-isokinetic (cooling rate is faster at higher temperature) but also faster at the outside of the sample than in the core that remains hot for longer.

What do we do with the data? In principle we can relate the experience of the lipstick in use to:

- The onset temperature (first sign of something beyond standard heat capacity);
- What's happening at 32°C
- The number of MPts and the maximum MPt
- The "degree of crystallization", with a higher degree making it harder to apply the lipstick while making it more resistant to accidental transfer.

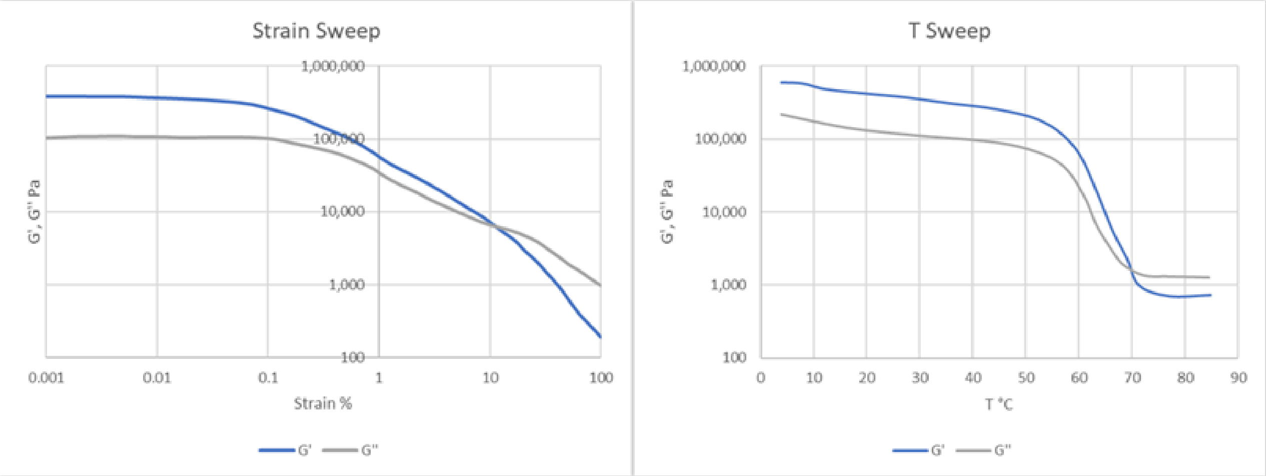
We return to these ideas later.

Rheology

We know that lipstick must be a combination of elastic and plastic properties – it must be relatively rigid (elastic) yet flow when applied to the lips (plastic). The standard way to examine the relative values is via oscillatory rheology, measuring G' (elastic) and G" (plastic), Flow_G' and G". Most uses of G'/G" data rely on "small strain" data, where the oscillation is in the linear region of, say, 0.02% strain, SAOS (Small Angle Oscillatory Shear). Small strain data are scientifically simple and we know how to analyse them. For lipsticks we must look at data up to 100% strain. This takes us into LAOS (Large Angle Oscillatory Shear) techniques with their Lissajous figures and Chebyshev coefficients. With modern software, we can extract "simple" G'/G" data (as shown in the plots) or any other feature such as strain stiffening and shear thinning and find ways to use the parameters to help characterise the complex trade-offs of lipstick.

Typically a G'/G" curve would be measured at a "reasonable" oscillation rate of, say, 1/s that matches the application process, though in an ideal world it would be measured at a range of frequencies representative of typical challenges to the lipstick. Measurement at 32°C makes sense as a reference point.

To complement the DSC data, a standard low-strain, 1/s measurement should be made across a relevant temperature range to get some idea of the intrinsic physical properties (elastic/plastic) during processing. Here are some typical scans showing what to expect. On the left is a less-routine LAOS sweep up to 100% strain, on the right is a routine SAOS sweep over a convenient range of temperatures:



In both cases, at low strains and low temperatures, the lipstick is mostly an elastic solid, with G' 3x larger than G". A key property of lipstick is that it should flow under strain. The strain sweep shows that above 1% strain it starts to become much weaker and above 10% strain G" dominates. The T sweep shows that not a lot happens to G' and G" in the DSC region below ~55°C, above which where there is a lot of melting behaviour and then, at these low strains, the lipstick finally starts to soften significantly.

What can you do with these data? Again, on their own, not much. We discuss below how best to use them.

Adhesion and abhesion

Although we could readily add adhesion promoters to the formulation that would provide strong bonding to the lips, these would result in a stained lip as the bulk is removed. That is unacceptable.

The only source of adhesion, therefore, is dissipation, just as is the case for pressure sensitive adhesives, PSA. We measure the potential for dissipation via the balance between G' and G", so our rheology will provide insights into the adhesive properties. As is the case for PSAs there is no simple formula relating G' & G" to strong adhesion, just the notion that too much or too little of either is a bad thing. That unhelpful remark points, again, to the discussions below.

Abhesion, the tendency to *not* stick is easily attained by having a high G' and low G", a brittle, high-wax formulation. This means that we either have to compromise or use the opposite effect by providing something that migrates to the surface to provide a thin (sub-µm is good enough) layer with low G' and G" and which, ideally, doesn't include much pigment. Although this can be achieved with silicones, they are becoming less acceptable, so some oil partition is desirable – without being excessive and showing up as syneresis or "weeping" from the lipstick (Gelling_Syneresis). All upsides have downsides – a thin layer on the surface of the bulk lipstick might make it too slippery to apply to the lips or might give a disconcerting "slip" when first brought into contact.

It is common to measure variants of adhesion and abhesion: tack and friction coefficient.

- A probe tack test pushes a flat-ended cylinder (don't use a spherical end!) onto a block of lipstick with a known force then measures the force needed to separate it. The peak force and the integrated force from any "stringing" of the lipstick give two measures of "tack". Too little tack and adhesion to the lip is likely to be low, too much and the lipstick is obviously "sticky" and unacceptable.
- The friction coefficient (Mechanical_Friction) can be measured with various degrees of sophistication, but placing a block of lipstick onto a surface of interest, tilting the surface and finding the angle at which the block starts to slide is simple and sufficiently accurate. Tests on a simulated skin surface will give an idea about applying the lipstick, and tests on a glass surface will give an idea of whether the lipstick will slip away when taking a drink from a glass.

Finally, adhesion of the lipstick to itself (cohesion) can be measured via a standard "breaking load" test – the weight needed to snap a sample.

Ingredient compatibility

Although we *could* use a tool such as Dissolution_Hansen Solubility Parameters to make sure that like is compatible with like in the formulation, we don't need to bother. All the ingredients to be used in significant quantities have to be waxy or oily, which means an absence of any functionalities except the esters holding fatty acids and alcohols together. As is widely acknowledged, trying to incorporate castor oil with the extra -OH groups from the ricinoleate chains comes with its dangers.

Those adding polymers as film formers or adhesion promoters might add polyacrylates, but they do so ensuring that a good fraction of the acrylate groups are based on long-chain molecules.

User data

Via a panel of trained or (giving a different perspective) untrained testers it is possible to score a lipstick for properties such as:

- Ease of application
- Covering of lip imperfections
- Degree of “bleed”
- Feel on the lips
- Resistance to transfer
- Longevity on the lips
- Removability from the lips

These are the properties that matter, though different consumers have different preferences for the balances of these features. It is not obvious how to connect the individual properties to the core science, nor how to tune the ensemble of properties for the right part of consumer space. That is the challenge in the final section.

Two other properties:

- Uniformity of colour
- Shine

are relatively stand-alone, though if shine is reduced for a deliberate matt look, this might help with reducing transfer.

The uniformity of colour requires a good dispersant package for the pigments which, as mentioned at the start, is out of the scope of this chapter, though is relatively straightforward to arrange via the right dispersion science.

For simplicity these two factors are not discussed further.

Navigating through lipstick formulation space

Ignoring the pigment package, antioxidants and fragrances, the challenge is to link our core science values to individual properties and to enable the trade-offs of consumer preferences.

The need for this arises because formulation space is vast. If a typical “lean” formulation has “only” 3 waxes and 3 oils they have to be chosen from a palette of, say, 10 of each with multiple ratios of each component to be explored. Some reduction in that palette can be made if marketing insist on “chemical free”, “all natural” or “low cost for high margin”, but that still leaves an impossible space to explore, especially as the user tests are slow and expensive and, maybe, too late in the development process.

Let’s assume we have set ourselves up to do the minimax set of scientific tests on a wide range of samples. Minimax is the minimum number of different experiments giving us the maximum number of parameters relevant to optimising in formulation space. The experiments themselves need to be routine and reliable, with data analyses (e.g. from LAOS sweeps) automated to produce the parameters we want to use. A proposed minimax is:

- DSC to give you a map of the melting points and crystallinity
- SAOS and LAOS experiments. If you can't manage these then a rotational viscometer (Flow_Rotational rheology) able to cope with these semi-solids, with their yield stresses (Flow_Yield Stress) and shear dependence (Flow_Shear dependent) might work.
- Friction coefficient, either simple tilt or a more sophisticated speed/load-dependent device must surely feature.
- The tack test, with due thought about initial loads and probe speeds, provides nice data for relatively little effort.

In the early days, don't bother about your own formulations. Buy in, and test, a big range of lipsticks of well-known brands with plenty of user information about them: Soft, Long Lasting, Budget, Luxury, Natural, … which you can approximately map onto the user data classifications defined for your own formulations.

Now you have the data analytics challenge of mapping scientific values against user perceptions. At the start this will be "dumb mapping" as you don't know what you are doing. You *can* do this via some mindless algorithm that produces a PCA or ML or AI fit to every measured parameter. You gain precision but lose insight. Instead, the first task is to reduce internal duplication. If analytical parameter A correlates strongly with analytical parameter B, and if some reflection tells you that they are different ways of looking at the same physics, then choose whichever is easiest to get and remove the other from the fitting. And try to do fits to sets of parameters your scientific intuition says *should* work well. If the fits are OK then it's usually better to go with scientific meaning rather than overfitting to higher accuracy. With these scientific fits, outlier data points are a valuable challenge to the theory. If they force you to modify your theory, that's called good science.

Now you have a validated "smart mapping" process, you can use it in reverse with your new formulations. As you alter ingredients and their ratios, you see how they are affecting key scientific parameters and can quickly sort out a response surface that should get you to a desired point in formulation space. Because you are using minimax analytics, you can justify the resources needed to do this mapping – you already know that you are not doing worthless or duplicate tests (they have been eliminated by the original analysis) and that what you are measuring provides a direct route to whatever optimal balance marketing have requested.

The point of smart mapping is that it can make itself smarter. At first you might have no idea why waxes ABC work together better than DEF in the context of oils UVW yet it's the other way round with oils XYZ. As the data accumulate you can start sub-mapping of the wax and oil domains. Each individual component has its own melting point, viscosity etc. so now these parameters feed into the overall map, improving both the map and the scientific understanding of which features interact and why.

Why not let the AI do all this? Because unless you have set up to do this sort of mapping and assembled a minimax set of reliable data points, you don't have datasets worth plugging into the AI. In the end, the smart thing to do might be to ask the AI not only to provide the formulation predictions but to explain why they work. But you first have to do the hard work

Mascara

Links

Adhesion_Adhesion promoters, Adhesion_Surface energy Interactions, Cleaning_Boundary removal, Cleaning_Solvents, Cleaning_Surfactants, Cleaning_Temperature effects, Diffusion_Barrier Properties, Diffusion_Basic Diffusion, Diffusion_Diffusion into Skin and Hair, Diffusion_Permeation OTR and WVTR, Dispersions_ODC, Dispersions_PVC and CPVC, Dispersions_Rheology (Low shear), Dispersions_Settling, Flow_Basic viscosity, Flow_G' and G", Flow_Oscillatory rheology, Flow_Rotational rheology, Flow_TTS-WLF, Flow_Yield Stress, Optics_Colour

How do you rationally formulate for a suite of properties such as curling, volume, shine, smudge/crack resistance, water resistance, along with smooth/easy application, long pot-life and “natural”? By taking a step back from formulations with too many ingredients and thinking it through.

If 6 ingredients are good, 12 must be better

You can make a workable mascara at home with some carbon black, coconut oil, aloe vera gel and beeswax, just 4 ingredients. If you want to go wild you can get extra volume from bentonite clay and maybe some extra smoothness from shea butter which takes us up to 6.

But that shea butter is mostly oleic and stearic acid, so it’s really 7 ingredients, and fatty acids can go rancid, so you’d better add some antioxidant, maybe Vitamin E as that sounds good in the ingredients list, taking us to 8. And that aloe vera gel is more than 90% water so you’d better add a preservative package which might be two ingredients such as phenoxyethanol and caprylyl glycol. So we are up to 10 ingredients.

Finally, to be really natural, add some essential oil which takes us to 11 ingredients if we are allowed to count the oil as 1 rather than 10+ ingredients.

Looking at any commercial mascara we easily reach 12+ ingredients and when film forming polymers are added for anti-smudge and water resistance we can head to 16+.

For example, instead of just beeswax we might find a blend of C18-36 acid triglyceride, White beeswax, Stearic acid, Copernicia cerifera and Euphorbia cerifera. Why that specific combination of 5 waxy materials was chosen is often unknown, but the assumption has to be that someone with magical formulation fingers found that it gave just the right balance of final properties.

This was a difficult chapter to write. It is very easy to go blah blah with truisms, but that doesn’t help. There are no easy, relevant formulae to provide an app-route through formulation space. Instead we identify those measurements that *must* be relevant to the formulation, then work out how to map those data against real-world behaviour of products that are known to be good or bad, your own and competitors. This is an unfamiliar approach, but a powerful one.

Maybe as a formulation scientist it is a better idea to take a step back from the assumptions of complex mixtures and find ways to map out the formulation domain in a way that builds a foundation for all future formulations.

If we look at the desired properties we can plan for a minimax formulation – minimum (affordable) ingredients for the maximum performance. At each step we can identify potential contradictions and resolve them rationally rather than exploring an ever-wider range of ingredient tweaks that might get us there eventually.

Taking a step back – the Water phase

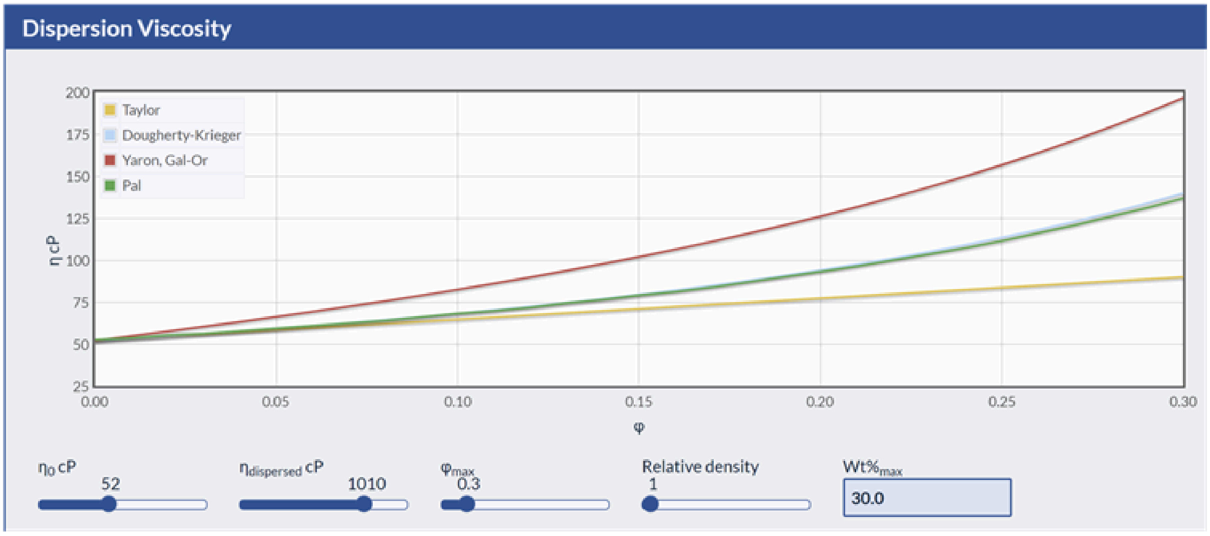
We can start with the most dangerous ingredient – water. We have to make it harmless by choosing a preservative package via whatever objective or marketing-led process makes sense. We know that preservatives will interact with other parts of the formulation, so let's find out those interactions sooner rather than later.

Assuming the water needs to be thickened to make a stable emulsion and to stop the mascara dripping during application and drying, the starting point is a marketing-led decision on the type of thickener. Xanthan gum, gum Arabic and similar might be good for "naturals", hydroxyethylcellulose is artificial but some consider it natural, and something like polyvinylalcohol is purely synthetic.

Given a choice of one or two thickeners, some standard rheology, Flow_Rotational rheology, Flow_Shear dependent, Flow_Thixotropy, Flow_Yield Stress will quickly capture the performance landscape of the key properties we need to understand:

- Low-shear viscosity and yield stress
- Shear-dependent viscosity & thixotropy
- Concentration-dependence of low shear viscosity

You can find out the sort of values required for the first two by measuring whatever commercially successful products are seen as market leaders. You will need a viscosity somewhat lower than the real product because the emulsion from your other ingredients increases the viscosity by a factor easily estimated from the desired volume fraction using the app described in Flow_Particle Viscosity.



η_0 cP 52

$\eta_{dispersed}$ cP 1010

φ_{max} 0.3

Relative density 1

Wt%$_{max}$ 30.0

https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

If we assume that the Yaron, Gal-Or formula applies, that our wax+oils has a 1000 cP viscosity and that we have 20% of it, then an initial viscosity of ~50 cP rises to ~125 cP. If you started with 500 cP then it would rise to 1250 cP – the effect of emulsion particles is a proportionality factor, in this case just over 2.

The shear-thinning behaviour is controlled by the degree of entanglement of the polymer which, in turn, depends how far above its entanglement molecular weight (see Adhesion_Entanglement) you are. You want some shear-thinning to allow the thick formulation to flow easily when being applied to the lashes. You *don't* want thixotropy (slow recovery back to full viscosity) because that would allow the formulation to sag on the lashes.

Sorting out concentrations and MW to give you the desired properties is tedious, but this is giving you permanent data on relatively pure systems that can be applied to other formulations. This is very different from trying to optimise these things within a complex formulation.

The final pieces of data are the concentration-dependences of the low-shear viscosity when you systematically vary concentrations of key ingredients, below and above what you expect to be reasonable. You need this for two conflicting purposes:

1. Desirable: A high concentration-dependence gives a rapid increase in viscosity when drying on the lashes so the final mascara properties are attained before drips or sags can happen;
2. Undesirable: That same high concentration-dependence gives a rapid increase in viscosity in the mascara container over days or weeks of use.

Building up this concentration-dependent and MW-dependent information is tedious, but it is timeless data for current and future formulations.

Taking a step back – the Oil phase

Again the starting point is marketing requirements for oils, fatty acids and waxes that are acceptable for the desired products. Paraffin or mineral oil might be preferred for inertness and lack of impact on food crops; or coconut oil might be preferred because it's natural.

The fatty acids might be dual use – surfactants/emulsifiers as triethanolamine salts in the formulation, or simply present as individual acids or blends such as shea butter. Because the triethanolamine salts revert to the acids on drying, for the moment we consider only their fatty acid properties in the final oil phase.

For the moment we can focus on the issues of what individual waxes do in the context of the chosen oil phases. Whether we need to create a blend is something to be decided if there is a problem that only a blend can solve. Again, the choice of natural versus synthetic waxes is mostly a marketing issue.

Across a reasonable range of oil/fat/wax ratios there are 3 key questions to be answered:

1. Over a user-relevant temperature range (say, 10-35°C) are the key mechanical properties stable enough to provide protection against drooping (high T) and flaking (low T)?
2. How do the wax crystals compare to your desired size (not too small to be irrelevant, not too large to be obviously visible), number (not too few, not too many) and stability against disappearance or growth during storage?
3. Given that cooling rates for a batch might vary, how sensitive is the crystal distribution to fast or slow cooling

Crystallization in the real formulation might be affected by the pigments (discussed later), but it’s still important to get to know the core oil/wax interactions.

The “key mechanical properties” can be most conveniently measured across a relevant temperature range via oscillatory rheometry Flow_Oscillatory rheology to get the G' & G" values Flow_G' and G". An over-brittle, poor adhering mascara (maybe showing up only at lower T) will have high G' (elastic) and small G" (plastic) values and an over-soft one (maybe showing up only at higher T) will be low G' with a lot of G".

The trick is to find a “strong” formulation that can retain “body” and “curl” and be anti-smudge, while being soft (“rich”) and flexible. By gathering G' & G" (probably at a convenient 1/s frequency) over the, say, 10-35°C range, any alarming changes in properties can be picked up quickly.

What are the ideal G' and G" values? Doing the same cycles of measurements on dried commercial mascaras of known low, medium and high quality will give some indications.

With a few examples you can have a look at whether low-level additives in the oil phase, such as tocopherol antioxidants, have unexpected effects. It is better to understand unexpected interactions at this relatively simple phase of building knowledge than to discover them late in the formulation cycle.

Taking a step back – Pigment effects

If you have put in the effort to ensure that your pigment will be nicely dispersed in your hot oil phase then, at the 5-10% addition range, they should make little difference to the oil. This can be seen from the Dispersion Viscosity app where increases of, say, 25% in viscosity might be expected. If, at least for generally spherical particles, you see a doubling of viscosity this is a signal that the dispersion is poor, producing strong pigment-pigment interactions that you don’t want. A well-dispersed pigment (see Optimal Dispersant Concentration, Dispersions_ODC) in the few 100nm size range gives optimal optical performance (colour, hiding power, see Optics_Scattering and Opacity) and it is likely that it will offer minimal interference with crystallization when typical crystals are in the multi-µm range.

Your formulation is complicated enough without having to take into account particle-particle, particle-wax, particle-oil and particle-water interactions. Obtaining pigments with good oil dispersants, or grinding them in the presence of some of your oil along with a good dispersant, should not pose great difficulties.

An exception is the micas. As described in Flow_Particle Viscosity, particles, such a mica, with a high aspect ratio (length/thickness) rapidly form a percolated network of high viscosity. Even with good dispersion, there will be significant viscosity increase, with a poor dispersant the formulation will be unusable. Because a typical bentonite clay is hydrophilic, micas and colour-coated micas need specific dispersants with a mica-friendly head and a long alkane tail. Tallow quats were a good example, though tallow is out of favour in many formulations.

After the investment in all those G' & G" measurements of basic oil/wax mixes, some follow-up experiments with pigmented versions will either provide confirmation that the pigments are mostly harmless to your formulation, or alert you to specific issues, such as particle seeding producing different crystallinities, or micas needing (probably) more oil and less wax to achieve the same balance of properties.

Taking a step back – the Emulsifiers

Although the emulsifier ends up in the dried-on mascara as (probably) an extra bit of oil/wax, choosing it for its anticipated contribution to the final product is not yet a priority. We want an emulsifier that gets us to a great o/w emulsion with the least effort. Given that emulsification takes place at something like 80-90°C in order to melt the waxes (unless our previous work encourages the use of a lower MPt wax) we need to use a theory that understands the "oiliness" of the oil/wax solution as well as the temperature effect on the characteristics of the surfactant (which is the word we will now use instead of the vague word "emulsifier"). This automatically means HLD: Surfactancy_HLD.

Experience with HLD tells us that the Equivalent Alkane Carbon Number (EACN) that defines the oiliness of these types of formulations is likely to be in the range of 15 – depending on the ratio of simple alkanes (maybe we're adding isododecane), esters, triglycerides and waxes. However, we know that relatively small quantities of "polar oils", which include fragrances and some preservatives, can change (usually reduce) the EACN, but we can't say by how much.

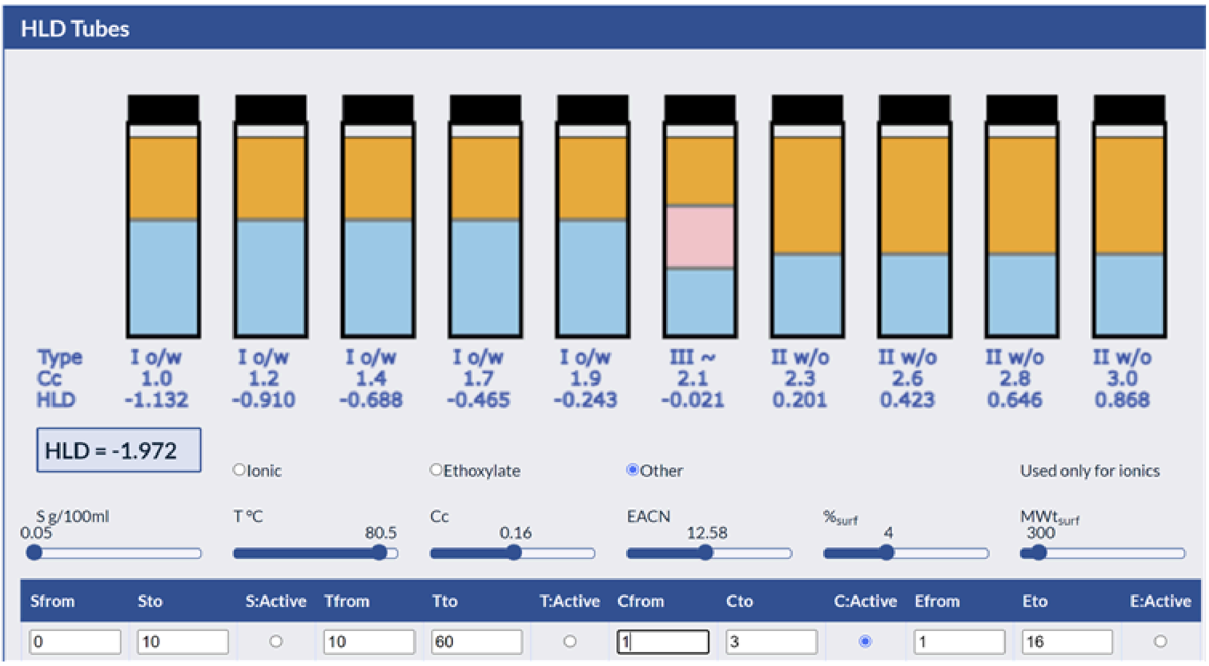
We also know that we have to think carefully about the classes of surfactants we intend to use. The ethoxylates become *much* more hydrophobic at these high temperatures, with the danger of creating w/o emulsions which are not good when you want a rapidly drying o/w emulsion. The anionics (including the fatty-acid salts) become *somewhat* more hydrophilic, which is not so good for these oily oils. And APGs and polyglycerols are temperature insensitive. Whatever the class, we need its characteristic (Cc) defined at 25°C.

To get efficient o/w emulsification we need an HLD ~ -0.25 at the high temperature. For *normal* o/w emulsions we want to ensure HLD ~ -1 to give stability against coalescence and flocculation. An HLD near -0.25 gives rapid emulsification but also relatively rapid de-emulsification, so we need to get from -0.25 to -1 for a stable room-temperature emulsion. But given that our emulsion will be semi-solid particles, as long as we cool quickly, we can benefit from the low Surfactancy_Interfacial Tension and Rigidity with a small, negative HLD to give us a fine emulsion with little need for high energy dispersion.

Once again, marketing have to rule on which classes of surfactants are acceptable. Once we know the classes we do a quick HLD calculation to tell us what starting Cc we need to give us a good emulsification then we find the 2-surfactant blend (usually no single surfactant is a perfect match) that gets us to that point. This *should* be done via manufacturers' lists of Cc values, but industry is only slowly giving us the data we need.

Looking at the list of Cc values on https://www.stevenabbott.co.uk/practical-surfactants/cc.php we find that the fatty acid salts are probably of little use (though the table only lists the Na salts, not the amine salts) as they are too hydrophilic, even more so at high T. The Spans, some APGs and polyglycerols look in the right sort of range and their temperature insensitivity is a bonus. There are plenty of ethoxylates that can do a good job at 80°C but they will be much too hydrophilic at low temperatures, which might be a problem.

Given the double uncertainty of EACN and Cc values, how can we formulate rationally? By doing intelligent scans:



HLD Tubes

Type	I o/w	I o/w	I o/w	I o/w	I o/w	III ~	II w/o	II w/o	II w/o	II w/o
Cc	1.0	1.2	1.4	1.7	1.9	2.1	2.3	2.6	2.8	3.0
HLD	-1.132	-0.910	-0.688	-0.465	-0.243	-0.021	0.201	0.423	0.646	0.868

HLD = -1.972

○Ionic ○Ethoxylate ●Other Used only for ionics

S g/100ml 0.05

T °C 80.5

Cc 0.16

EACN 12.58

$\%_{surf}$ 4

MWt_{surf} 300

Sfrom	Sto	S:Active	Tfrom	Tto	T:Active	Cfrom	Cto	C:Active	Efrom	Eto	E:Active
0	10	○	10	60	○	1	3	●	1	16	○

https://www.stevenabbott.co.uk/practical-surfactants/HLD-Tubes.php

Here we have 10 tubes containing a 50:50 mix of our oil and water+package phases. We know we have some nice hydrophobic sugar- or glycerol-based surfactants, maybe in the Cc range of 1 to 3. So in each tube we place 4% of consecutive mixes of the surfactants to take us from a nominal Cc = 1 to Cc = 3. We shake up the tubes in their 80°C water bath and allow them to phase separate. When we see that (in this case) tube 6 contains the 3 phases characteristic of HLD = 0, we have a good idea of where we need to be and, specifically, the mix in tube 5 (HLD ~ -0.25) should be a good starting point for emulsifying our mascara.

This is highly idealised and reality doesn't give you a great scan right first time. The payoff of working to get good scans is that it is then easy to, say, change preservative packages, or make gradual changes to your oils/waxes. If the optimum tube changes slightly then you have expanded your range of easy-to-swap ingredients if some sudden new requirement emerges. If the optimum tube changes significantly then you now know that ingredient swapping will require more care, with adjustments to the surfactant package to accommodate the change.

Bringing it all together

It has been a lot of work. But now you have a palette of scientific information that you can start to combine in real formulations. And you have the basis for a rational digitalisation process. When the individual components come together with only modest new interactions, you know in advance how to tune some ratios to quickly get you to the desired end point. When they come together and give very different properties than those you expect, that can be welcomed by optimists as a sign of an interesting effect – interestingly bad or good, it's all valuable information.

Had you tested that formulation without the preliminary data, there would be no way of knowing that some synergy or antagonistic effect had taken place – it would be just another confusing datapoint in mascara formulation space.

Bells and whistles

You might still be required to add some exotic plant extract for marketing reasons. With the efficient science-based methodology, you can quickly test its effect on whichever phase it is going to be added in (water or oil) and with luck you will see, at worse, the need only for a minor tweak.

The film-forming polymers for greater water resistance are more of a problem. Some are water-based (acrylics), with a strong pH dependence, e.g. requiring the triethanolamine salt to be water soluble, reverting to insolubility when the triethanolamine evaporates, leaving the insoluble acid form. Others are oil based and some formulations even use mixes of water-based and oil-based polymers to make life even more complicated.

Getting these polymers to be right, not too rigid, not too glossy/plastic-looking, not too hard to remove is a whole topic beyond the scope of this chapter. But maybe the rational approach discussed here can get you to the desired properties without those extra ingredients. A glossy, water-resistant, well-adhering oil/wax combination is a real possibility and marketing can put “no nasty polymers” onto the packaging for the many customers who like pure simplicity.

Microencapsulation

Links

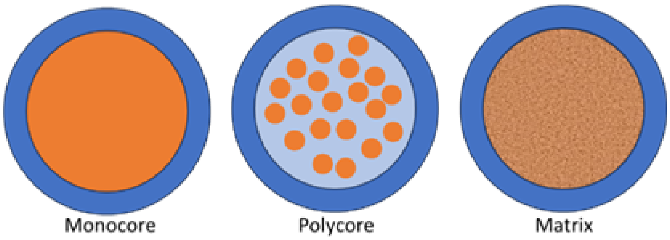
Absorbency_Porosity and Darcy, Diffusion_Basic Diffusion, Diffusion_Concentration-Dependent Diffusion, Diffusion_Diffusion Coefficients, Dispersions_Settling, Dissolution_Flory-Huggins, Dissolution_Hansen Solubility Parameters, Dissolution_Ideal solubility, Dissolution_Solubilizers and Hydrotropes, Flow_Basic viscosity, Flow_Particle Viscosity, Flow_Polymer Viscosity, Fragrance_Activity Coefficients, Fragrance_Barrier Properties, Fragrance_Vapour Pressure

We have 3 broad ways to make microcapsules:

1. Covering solids with a spray
2. Making a solution or emulsion and spray drying it
3. Making an emulsion and adding a shell

Here we are interested in not so much in microencapsulation itself but the core science that a formulator can use to influence the process during making and then using the capsules.

Release from the capsule



There is no point in making a microcapsule if its release properties are inappropriate for the desired application. So let's make sure we understand release mechanisms before working out the best way to make the capsules. Although what follows sounds as if it is specific to monocore, there is little difference in principle if we have a polycore or matrix form.

Mechanical fracture

A typical microcapsule will need to be robust enough to withstand mechanical stresses during manufacture and general use, yet weak enough to rupture under the desired stress conditions, such as when scratching a film or when a fabric is being worn. So it should be standard practice to measure the relevant parameters and compare them to the expected stresses.

However, the science is murky and indirect. For example, the force, F, needed to deform a capsule of radius R by amount δ depends on the Poisson ratio ν which is either 0.3 or 0.5 (in reality it makes little difference which you choose) and the modulus E of the capsule:

$$F = \frac{4E\sqrt{R}\,\delta^{1.5}}{3\left(1-\nu^2\right)}$$

This Hertz model seems to say that larger capsules are stronger – bigger R means bigger F – but because we are interested in the stress (force per unit area), we have to divide by πR^2, meaning that there is a $\frac{1}{R^{1.5}}$ dependence on effective strength – bigger capsules are easier to break.

What about the modulus, E? This is some combination of the modulus of the contents and the modulus of the shell, with the combination depending on the ratio of shell thickness to capsule radius and the ratio of the shell modulus to that of the contents. Details are again sketchy but we can probably assume that if the contents are purely liquid then doubling the shell thickness will double the strength. And maybe doubling the modulus of the shell material will double the strength … for a constant δ.

This leaves us with δ itself. If you add an extra shell of silica to a typical shell of melamine-formaldehyde then you increase both the shell thickness and its modulus. But as the shell is now more brittle, you will reduce δ. It's the classic dilemma – "stronger" can mean "more brittle" which can mean, on balance, weaker.

This leaves the formulator with the general idea that thicker, stronger shells and smaller capsules will be either more robust (good) or hard to break for the user (bad). If you aim for thicker shells you automatically decrease the loading of your ingredient, which is not good. If you aim for stronger shells, these might be brittle, leading to a smaller δ and a weaker capsule. All of which, sadly, says that you have to resort to trial and error. That is not a sentence we want to include in the FST, but we can only work with whatever the literature provides.

Porous escape

Approximately speaking, if you have a porous capsule, escape will be "fast". Although we could think of applying Absorbency_Porosity and Darcy, with the typical thin walls and small capsule sizes, release will be effectively instantaneous.

Diffusion

The laws of diffusion (Diffusion_Basic Diffusion) tell us that the steady-state rate of diffusion or flux, F, out of a capsule with shell thickness h, diffusion coefficient D and concentration gradient C is:

$$F = \frac{DC}{h}$$

Because, for practical/economic reasons, you can't change h by more than a factor of ~2 in the real world, you can't make much of a difference with shell thickness.

The next factor is the diffusion coefficient, D. Although it would be nice to aim for a superb, highly crystalline polymer with $D < 10^{-10}$ cm²/s, most shells are created very rapidly and are too disordered for such a low D. Although crosslinking can lead to low D values, again a typical crosslinked shell does not have an impressively low D. Larger molecules have a lower D value. If D for a molecule with MW = 100 is 4.10^9 then, assuming a MW² dependency, a molecule of 200 Daltons will have $D = 1.10^9$. Clearly D is an important factor, but if the shell's other desirable properties don't come with a low value of D, what else can you do?

The final factor is the solubility of the molecule in the shell. The rate of diffusion is driven by the concentration gradient, C, from inside to outside, and we can define this as the saturated concentration of the molecule within the inner few nm of the shell if we assume that the concentration outside is effectively zero. Because of the small effects (factor of ~2) of shell thickness and MW and with only a modest ability to lower D for a given type of capsule, your flux is largely driven by C, which can be a factor of 1000 because the range of solubilities can be anything from, say, 10% down to 0.01%.

How do you control the solubility? As so often in this book, the simplest way is via Hansen Solubility Parameters, HSP: Dissolution_Hansen Solubility Parameters. By knowing the HSP of your shell polymer and that of the molecule(s) inside the shell, it is easy to calculate the HSP Distance which tells you whether the solubility is low, medium or high. An important consideration here is that if you have a mixture of molecules you can calculate the average Distance and, therefore, get a general idea of solubility. But you can also look at the individual Distance values and estimate (along with their individual MWs) which molecules will diffuse quickly or slowly. If you have a fragrance formulation, big differences in HSP and MW may show up as a fragrance that changes over time – with the smaller, closer molecules escaping sooner than the larger, more distant ones.

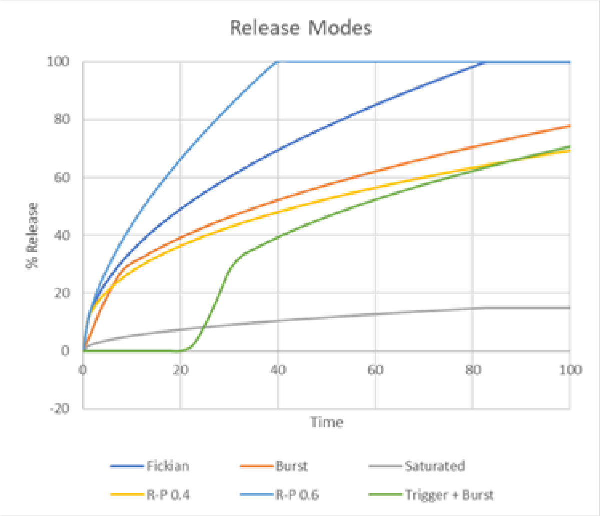
Mixed modes

Typically, the amount released at time t, m_t is related to the total mass, m_0, by:

$$\frac{m_t}{m_0} = kt^n$$

For Fickian release, n = 0.5 (classic square root of time) and for the more general Ritger-Peppas (R-P) model, n can be anything from 0.35 (somehow constrained) to 0.6 (maybe swelling helping diffusion) to 1 (the much-desired zeroth order release).

If we release into an infinite sink, so that the molecular concentration outside the capsule is always (effectively) zero, we can obtain graphs that go from $m_t = 0$ to $m_t = m_0$ as in these graphs:



Our porous release might give just burst kinetics. Our mechanical restraints might give us triggered release. We might get pure Fickian (or R-P if $k \neq 0.5$) with a steady curve, or we might get a mix of burst (maybe molecule caught in or on the shell) and Fickian or R-P.

The release coefficient, k, is basically $F = \frac{DC}{h}$ so the different graphs (ignoring the bursts) represent either different diffusion coefficients, D, saturated concentrations C or wall thicknesses h.

If k is purely a function of the shell, then the size of the capsule makes no difference. If k depends on the core then h will be some complex mix of core and shell sizes, so larger capsules will show slower release. Although we *could* produce some fancy equations for this, in reality we are unlikely to know all the parameters so we can treat k as just another constant.

If our encapsulated molecule is a solid then k is a complex mixture of solvent (usually water) coming through the shell, dissolving the solid and the concentration gradient then driving out the molecule. Once again we can subsume that into some generic k. This is arguably more Higuchi kinetics than Fickian kinetics, but as they each have a $t^{0.5}$ dependency, we won't really notice the difference.

If our aim is to keep the molecules in the capsule even though the capsule is in some environment where release is possible (e.g. a fragrance capsule in a liquid laundry detergent) then we must reduce C. Although we say that this is a *concentration* gradient, in reality it is a gradient against *activity*. So if the saturated concentration in the external environment is very low (e.g. a fragrance oil in water), after a small m_t the concentration has reached saturation and the activity gradient is now 0, so F is 0 and k is 0.

Strong barriers

If diffusion through our shell is too fast for the intended application then we can try a few more tricks:

- Evaporate a shell of Al
- Coat with some sort of sol-gel to give something like a silica coating
- Incorporate nanoclays that provide a tortuosity barrier, as explained in Diffusion_Barrier Properties

Each of these has obvious downsides in terms of cost and complexity. The tortuosity barrier is especially a problem for spray-dried capsules, as discussed below.

Polycore and matrix capsules

Whether we have monocore, polycore or matrix capsules makes no obvious difference to the science. The bursting behaviour is just as inscrutable, and the release mechanisms can follow any of the curves shown above. This is not to say that these different capsule types are of no consequence. For a given shell, matrix and molecule, the performance might depend strongly on the capsule type. The point here is that each capsule type is able to produce each of the different release modes. It is likely that a monocore capsule will tend to be Fickian and that a matrix capsule might show interesting R-P behaviour, but when you add in the effects of capsule size distribution, damage during processing, variabilities in shell thickness and integrity, difference between capsule types become blurred.

Making the microcapsules

We will look at the 3 most typical ways to make microcapsules:

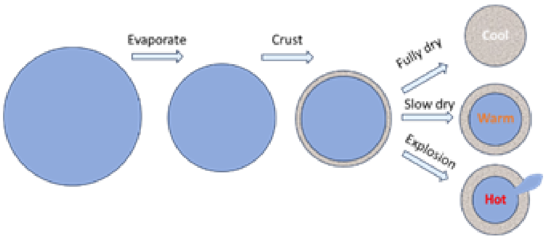
1. Coating onto a particle
2. Spray drying from solution or dispersion/emulsion
3. Adding a core/shell to an emulsion

Coating onto a particle

There is surprisingly little that can usefully be said. Assuming, for the moment, that you are using a fluidized bed, then your particles can't be too large (can't be fluidized at reasonable air velocities) or too small (uncontrollable) and, therefore, you need control over the size distribution so all your particles are in the viable range, see Particles_Size distribution. Then you need to spray your coating using some sort of atomizer where, again, the bad news is that there is little useful formulation science (Flow_Atomization) other than the obvious advice to have low viscosity and an absence of viscoelasticity.

Whether you use top spray or bottom spray or a Wurster device is usually a given so you have to formulate with what you have.

Spray drying



We have the same problem as with coatings. Spray drying, mostly from water, is mostly about getting a good atomizer and good airflows, neither of which is under the formulator's control. In this section we don't worry if we have a solution or an emulsion. The emulsion aspects are described in the next section.

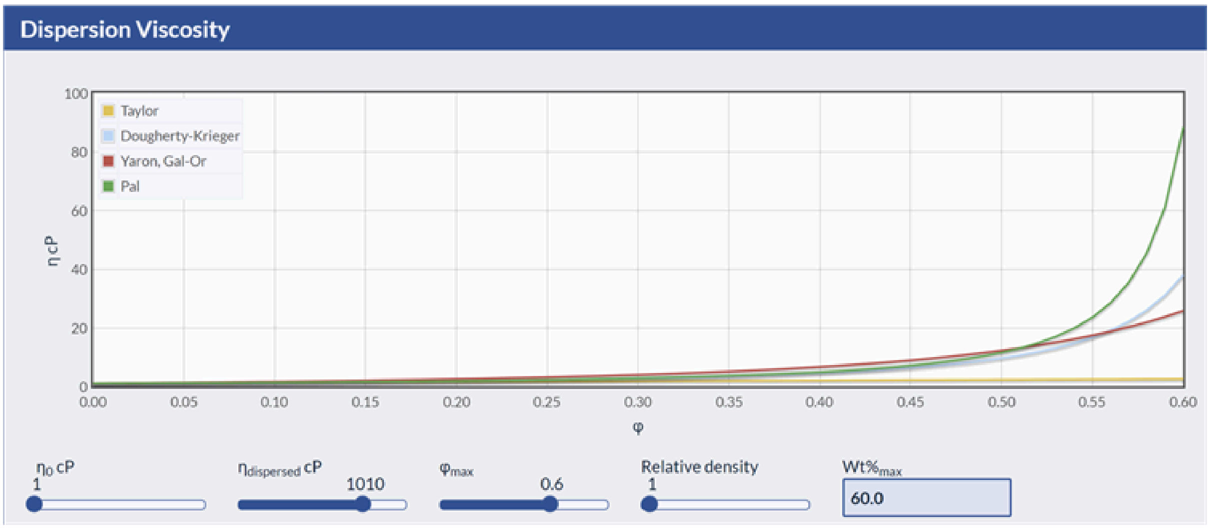
There are difficult tensions over which we *can* have some control.

- Production will always demand higher concentrations (less water to evaporate) and lower viscosity (easier atomization) which are usually contradictory.
- They also want a rapid transformation on the outside (crust formation) so that particles that touch each other or the wall don't stick, without that crust inhibiting rapid evaporation of the water. Such inhibition can slow down the process or, at worst, lead the water to explode as the internal temperature exceeds 100°C.
- We want maximum hot air temperature for rapid drying, but we don't want to destroy the contents of the particle, such as enzymes or friendly biome microbes.

Let's try to do something about those contradictions.

Maximum concentration, minimum viscosity

If our water contains particles or (it's the same thing) emulsion drops, then the same app shows the viscosity dependence on volume fraction, φ and on the viscosity of the dispersed particle (a few cP for a typical oil and for a solid particle 10,000 cP is good enough:



η_0 cP 1

$\eta_{dispersed}$ cP 1010

φ_{max} 0.6

Relative density 1

Wt%$_{max}$ 60.0

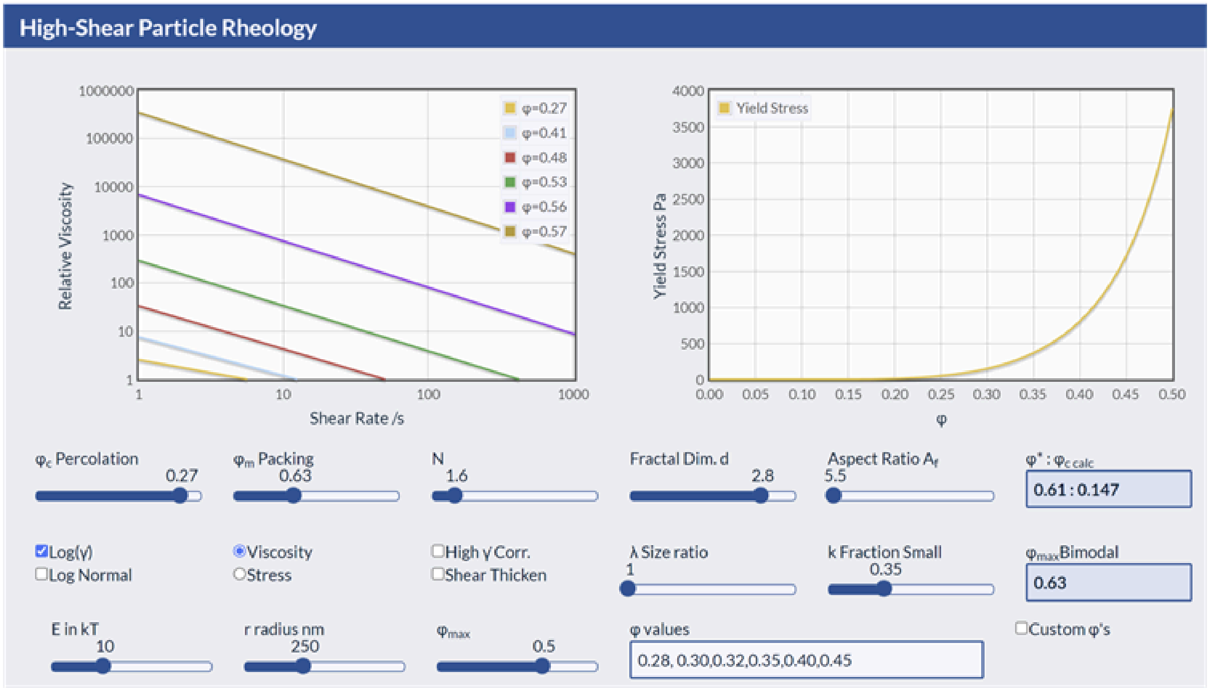
https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

This tells us that without particle-particle interactions, just them getting in each other's way, the viscosity doesn't change much from water's 1cP till about 50% volume fraction when it reaches 10cP and is starting to grow rapidly.

But be careful. If your "water" starts at 10cP because of some soluble additive, then at 50% volume fraction the viscosity is now 100cP – the crowding effect is multiplicative, not additive. If you really need a few % more particles, do whatever it takes (and it's often relatively easy) to reduce your "water" viscosity by a few cP. This is an important and little-known trick.

If, however, your particles show significant interactions because, for example, you have a poor dispersant on solid particles, or because (and here is one warning about nanoclays for extra barrier properties) they have a high aspect ratio (length/thickness) then viscosities can shoot up.

We can explore these ideas in Flow_Particle Viscosity which includes this app:



φ_c Percolation 0.27

φ_m Packing 0.63

N 1.6

Fractal Dim. d 2.8

Aspect Ratio A_f 5.5

φ^* : $\varphi_{c\,calc}$ 0.61 : 0.147

☑ Log(γ)

☐ Log Normal

● Viscosity

○ Stress

☐ High γ Corr.

☐ Shear Thicken

λ Size ratio 1

k Fraction Small 0.35

φ_{max} Bimodal 0.63

E in kT 10

r radius nm 250

φ_{max} 0.5

φ values 0.28, 0.30,0.32,0.35,0.40,0.45

☐ Custom φ's

https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php

As you play with the sliders (the N and Fractal Dimension sliders describe particle association, there is also the Aspect Ratio slider) you can see two things:

- It is scarily easy to shoot up to super-high viscosities *at low shear.*
- But if your pumps are good enough to create high shear in the atomizer nozzle, then maybe you will be OK.

If you are going to rely on the very strong shear thinning of particle formulations, make sure you have tested many likely variations in production batch formulations. It is very easy to go from a 1000x low shear increase in viscosity to 10000x with a small change in dispersant or aspect ratio.

If you are dealing with polymer solutions then you need to be familiar with Flow_Polymer Viscosity as shown in the app:

R_g @0.5+40K nm 10
M_{mono} 100
Polymer MWt 28053
Polymer M_c 9759
Flory-Huggin χ 0.2
C_{max} g/dl 17.6
R_g @χ+MWt nm 11.4
Ne 32.5
C* g/dl 3.2
Ce g/dl 4.6
|η| dl/g 1.58
N.bonds 400
M_{bond} 100
Bond length nm 1.4
MWt 40000
R_g nm 11.4

https://www.stevenabbott.co.uk/practical-solubility/polymer-viscosity.php

The core idea is not so much that higher polymer concentrations lead to rapid increase in viscosity (we all know that), nor that high MW polymers are worse than low MW (again we know that). The key is MW compared to Mc, the "critical entanglement MW". You can have two similar polymers, with similar solubilities and similar MWs, but one of them will be much more viscous. This is because the number of tangles depends on MW/Mc. If the two polymers have Mc values differing by a factor of 2, then the one with the lower Mc is twice as tangled so will be more viscous.

In many applications (Adhesion_Entanglement) you *want* lots of entanglement, so you have a real fight when increasing concentration and MW. For spray drying you may not need so much entanglement so can choose either a lower MW or a polymer with a large Mc, so a high MW polymer (desirable for other reasons) becomes acceptable.

However, if you want a low diffusion coefficient for your encapsulated molecule in the polymer, all other things being equal, an entangled system leads to lower diffusion. Everything in formulation is a trade-off.

Another way to decrease viscosity of a polymer solution is to decrease its maximum solubility. Compared to a "happy" polymer with chains reaching out and tangling, the "unhappy" polymer is more coiled up on itself so isn't so tangled. The obvious flaw in this approach is that you need a high solubility to get enough polymer into solution.

Fast drying, all the way to dry

Everyone who dries coatings wants maximum drying speed with minimum formation of a crust that inhibits further drying. See Evaporation_Basics and Evaporation_Diffusion limited for a more detailed discussion.

The key idea is that in "constant mode", which is evaporation-limited drying, you are relying on a combination of heat and mass transfer, readily arranged in a modern drying system. As soon as you become "diffusion limited", you can't dry faster than the solvent gets through the crust, which becomes slower as the crust thickens, then there is nothing you can do other than increase temperature because diffusion goes faster at higher temperatures. High temperatures of a coating or a particle are generally not desirable.

So how do we keep our crust "open" enough to keep the drying evaporation limited? There are a few standard tricks:

- Add a "high boiler" (e.g. propylene glycol) in which your formulation remains at least partially soluble. This keeps the crust "open" for much longer.
- Use the formulation tricks for highest concentration with minimum viscosity to keep your formulation "happy" for as long as possible. You don't care that the shell is now very high viscosity, you just want to make sure that there is still enough water for diffusion to be rapid and not rate limiting.
- Add some junk (e.g. rough particles) that stops the crust getting too compact. There's an obvious downside if this makes the capsules too leaky, but with some thought you can get a reasonable compromise. If those particles are sticking out they will also reduce capsule-capsule and capsule-wall adhesion while the drying is incomplete.

This problem of the crust becoming too good a barrier is especially severe for those trying to use nanoclays as tortuosity barriers. There are no obviously good workarounds for this.

Keeping cool

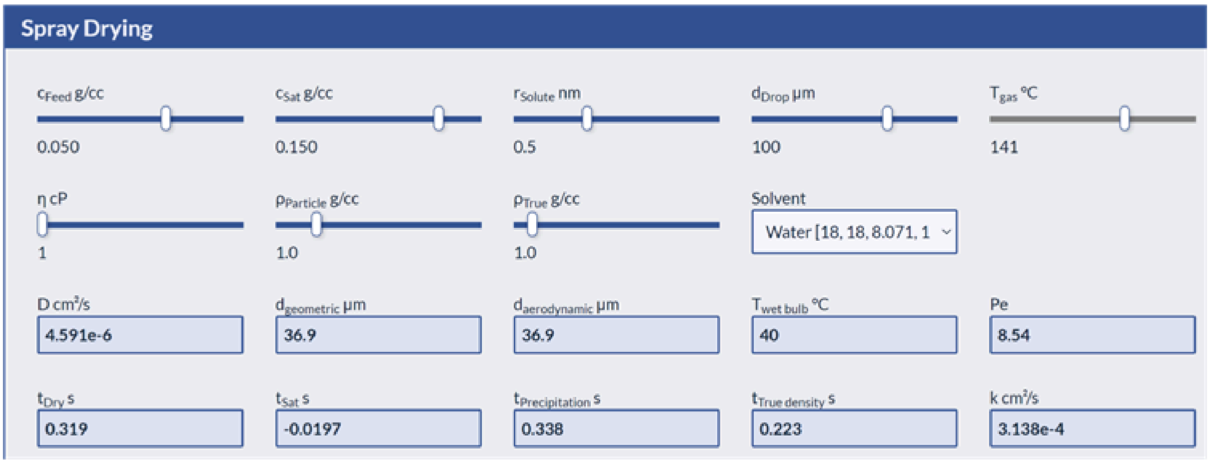
How do we spray dry heat-sensitive enzymes or microbes in 140°C (413°K) air, T_{gas}, without destroying them? A handy approximate formula (found in the app below) gives us the answer for water with a BP, T_b = 373°K. It tells us that the temperature of the microcapsule, T_{mc} is given by:

$$T_{mc} = 137\left(\frac{T_b}{373}\right)^{0.68}\log_{10}\left(T_{air}\right) - 45$$

For T_b=373 then we have 137 $\log_{10}$(413)-45 = 313°K = 40°C. Yes, our rapidly evaporating water is only at 40°C, so our enzyme or microbe inside the capsule is OK. If, however, we create a barrier crust then the capsule rapidly approaches 140°C and we have failed. Avoiding diffusion-limited drying really is important for spray drying.

A spray drying app

For those keen to explore some of the complexities of a pharma spray-drying system, this app gives more than enough to think about:



Spray Drying

c_{Feed} g/cc 0.050
c_{Sat} g/cc 0.150
r_{Solute} nm 0.5
d_{Drop} µm 100
T_{gas} °C 141
η cP 1
$\rho_{Particle}$ g/cc 1.0
ρ_{True} g/cc 1.0
Solvent Water [18, 18, 8.071, 1
D cm²/s 4.591e-6
$d_{geometric}$ µm 36.9
$d_{aerodynamic}$ µm 36.9
$T_{wet\ bulb}$ °C 40
Pe 8.54
t_{Dry} s 0.319
t_{Sat} s -0.0197
$t_{Precipitation}$ s 0.338
$t_{True\ density}$ s 0.223
k cm²/s 3.138e-4

https://www.stevenabbott.co.uk/practical-coatings/Spray-Drying.php

The meaning of the multiple inputs and outputs are described on the app page.

Why only this one app? Because the spray drying community has not provided any other usable models.

From emulsions to capsules: Getting the emulsions

One route to an emulsion-based microcapsule is to make an emulsion and spray dry it. The discussions here are tangled up with the more general principle of using the emulsion as the basis for the whole capsule. If you are oriented towards simple spray drying, you can ignore the other aspects.

In other parts of the FST we treat emulsification as a relatively simple process of putting an oil into water or water into oil. Here we focus on “oil” in water, but now our oil is much more complex; it might be the content of the capsules, or it might be an oil containing the real content dissolved within it. In these cases, a shell has to be provided around the emulsion. Or it might be a hydrophobic pre-polymer emulsion containing the content, with the aim being to create a “matrix” capsule consisting of content dissolved in the crosslinked polymer, with no need for a shell. Or it might be a hybrid – a matrix *plus* a shell.

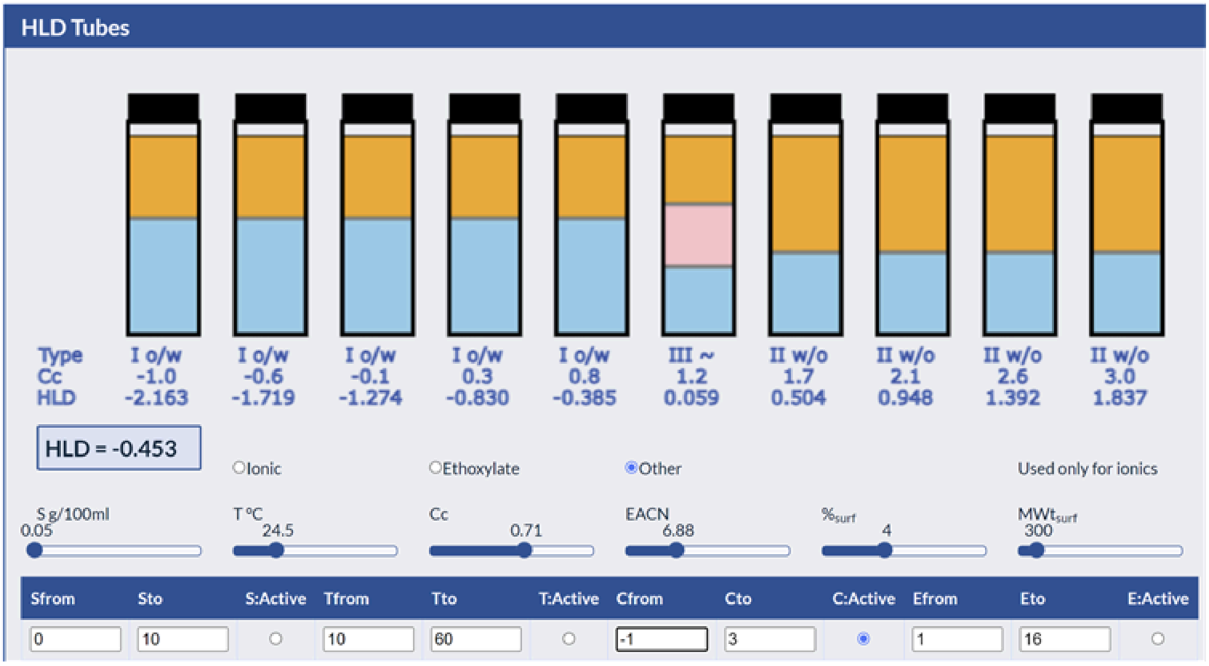
The FST approach to emulsification is to use HLD theory, Surfactancy_HLD, as the backdrop to understanding Surfactancy_Emulsification itself.

The reason for using HLD is that we can characterise the oiliness of the oil (its EACN, Equivalent Alkane Carbon Number), the characteristic properties of the surfactant (blend), Cc, and any salinity, S and temperature effects, T. We can tune the system to get HLD ~ 0 to allow easy emulsification with low energy, then move to HDL ~ -1 to give a stable o/w emulsion. This rational approach avoids the trial and error necessitated by classic surfactant science which has no coherent way to balance EACN, Cc, S and T effects.

The obvious problem applying this to microencapsulation is that other than for the first type of capsule delivering a simple oil, we don’t have classic oils that are likely to follow nice EACN rules. What is the EACN of a simple oil containing a 20-component fragrance and some crosslinkable pre-polymer? No one knows.

Here is where some patience in the early formulation steps saves a lot of time later on.

The assumption in what follows is that you already know the types of surfactant you prefer to use: anionics, cationics, ethoxylates, sugar-based, polyglycerols, and their level of greenness. This means that you have some idea of how temperature and/or salinity will affect your chosen surfactant class. It is also assumed that you have some choice, within your range, of hydrophilic and hydrophobic variants. And although you might not know the Cc values of all your chosen class of surfactants, you have at least some idea of the Cc values of at least two of them. If this paragraph is mostly meaningless, then invest some time in the HLD chapter before proceeding.



Type	I o/w	I o/w	I o/w	I o/w	I o/w	III ~	II w/o	II w/o	II w/o	II w/o
Cc	-1.0	-0.6	-0.1	0.3	0.8	1.2	1.7	2.1	2.6	3.0
HLD	-2.163	-1.719	-1.274	-0.830	-0.385	0.059	0.504	0.948	1.392	1.837

HLD = -0.453

○Ionic ○Ethoxylate ●Other Used only for ionics

S g/100ml 0.05

T °C 24.5

Cc 0.71

EACN 6.88

$\%_{surf}$ 4

MWt_{surf} 300

Sfrom	Sto	S:Active	Tfrom	Tto	T:Active	Cfrom	Cto	C:Active	Efrom	Eto	E:Active
0	10	○	10	60	○	-1	3	●	1	16	○

https://www.stevenabbott.co.uk/practical-surfactants/HLD-Tubes.php

The idea is that we start with our base, carrier, oil. Using two surfactants, (in this example, one hydrophilic (Cc = -1) and one hydrophobic (Cc = 3)) we do a scan of 50:50 o:w of stepwise mixtures from left to right and find in which tube we change from HLD < 0 to HLD > 0. We are not interested in 50:50 o:w nor are we especially interested in the base oil (though it plays an important role so if we change it, everything will change along with it). Instead we are pleased to have one nice point in the formulation map relevant to our spray drying. Although we say "water" you might choose to start with whatever aqueous phase you will be using (containing a range of ingredients) or you might prefer to use pure water then (not discussed further) do a similar process to find the effects of the extra ingredients.

Now we repeat the scan with a small % of, say, the oil-soluble molecule we wish to deliver. Maybe the scan shifts one tube to the left. We then double the % and the scan shifts 2 tubes. We now know that our molecule is making things effectively more hydrophilic, and we have some approximate rule for the shift in HLD versus concentration.

We now repeat this but adding our pre-polymer. Maybe this shifts things to the right.

Assuming you are doing this via some sort of robotic system (the scans are simple but tedious), you quickly build up a map of what's going on. You can now see if the effects of the molecule and the pre-polymer are a linear mix (which is nice) or have some non-ideal mixture interactions.

Why is it worthwhile to have done all this? There are three reasons:

1. If by some rational trick you can set HLD ~ 0 during the emulsification stage you can get a fine emulsion easily, with little expenditure of time and energy. You then need to complete the trick by shifting -1 < HLD < -0.5 to get a stable emulsion for further processing (discussed shortly). This is easy (but costs energy) to do via raising the temperature in ethoxylate systems (so-called PIT method) and requires other tricks for ionics and the sugars/glycerols.
2. If you choose to emulsify with -1 < HLD < -0.5 (without the HLD ~ 0 trick) then you get the most emulsion for least cost/effort by using the correct surfactant blend
3. Any time you change your formulation, you can readily adjust your surfactant blend to stay in the

optimum zone for emulsification.

By bringing control to the emulsification step and by understanding why well-intentioned changes in other parts of the system (a different oil, prepolymer etc.) can disrupt the emulsification, the whole process of product development and improvement is easier. Suppose, for example, that marketing want you to abandon, say, ethoxylates because of some internet scare. Without the understanding from HLD it would be a nightmare to find a replacement using, say, APGs. With the understanding, you can readily find an APG blend that gives an approximation to the equivalent ethoxylate blend and if you have relied on PIT you don't have to waste time with fruitless temperature cycles – you have to get to HLD ~ 0 by Cc blending.

Why do we want -1 < HLD < -0.5? What we really want is high concentrations of stable emulsion particles that don't give Surfactancy_Emulsion Coalescence or Surfactancy_Emulsion Creaming and Flocculation. We know that when HLD ~ 0 the interfacial tension is wonderfully low for creating emulsions but that also means that the emulsions coalesce easily. When HLD > 0 we are in the w/o region that we must avoid, And when HLD < -1 the surfactant is likely to be very happy in the water rather than at the interface, so is less efficient for stabilizing the natural o/w curvature we have obtained.

From emulsions to capsules: Getting the capsules

There seems to be little incisive information for the formulator. If you are creating the shell via an interfacial polymerization, say an isocyanate in the emulsion core and a polyamine in the continuous phase then you have complicated kinetics and thermodynamics with few reliable tools to inform your choices. If you are creating a water-insoluble polymer via water-soluble pre-polymers then it's the same thing – you have complicated kinetics and thermodynamics leading, one hopes, to the polymer forming around the emulsion drops rather than as gunk within the reactor.

And if you are relying on complex coacervation, Surfactancy_Coacervation, you are entering a world with plenty of general ideas, lots of handwaving and little usable theory.

Let me know of any FST-style science that can provide some usable ideas so this chapter doesn't just end like this.

Pharma Formulation

Links

Diffusion_Concentration-Dependent Diffusion, Diffusion_Diffusion into Skin and Hair, Dissolution_Crystallization, Dissolution_Dissolution Kinetics, Dissolution_Dissolution Rheology, Dissolution_Flory-Huggins, Dissolution_Hansen Solubility Parameters, Dissolution_Ideal solubility, Dissolution_Kirkwood-Buff, Dissolution_Solubilizers and Hydrotropes, Flow_Basic viscosity, Flow_Couette Flow, Flow_Oscillatory rheology, Flow_Poiseuille Flow, Flow_Rotational rheology, Flow_Shear dependent, Flow_Yield Stress, Surfactancy_Coacervation, Surfactancy_Emulsification, Surfactancy_Emulsion Coalescence, Surfactancy_Emulsion Creaming and Flocculation, Surfactancy_Emulsion Ostwald, Surfactancy_HLD, Surfactancy_Phase Diagrams

To deliver a dose of an Active Pharmaceutical Ingredient, API, other than via direct injection (ignored in this chapter) requires it to be dissolved or dispersed in some delivery medium, either as a solid or liquid formulation which dissolves at the desired rate once ingested, or as an external cream or patch for delivery through the skin. Here we look at a range of challenges to getting a viable delivery vehicle for your chosen API

Getting a pure API

Assuming you want to purify by crystallization, the first task is to go to the Dissolution_Crystallization chapter to discover why something that seems so simple is so tricky. It is made all the harder because the standard theory is provably wrong, and there hasn’t been an adjustment to the newer ways of thinking about how to obtain the right crystals at the right speed from the right solvents in the right polymorphic form.

Given that controlling solubility either via temperature or evaporation is key, having a good theory of solubility and temperature dependence is vital. The starting point is the idea of ideal solubility, Dissolution_Ideal solubility, where, approximately, the maximum possible solubility is given by a combination of MPt and Heat of Fusion. If you have molecule with strong crystal packing then both the MPt and heat of fusion will be large so solubility, whatever you do, will be low. Some (not very good tricks) for dealing with this are discussed below.

Assuming you have a sufficiently high ideal solubility, finding the right solvent requires a good solubility tool, which means COSMO-RS. The starting point for COSMO-RS calculations is either the ideal solubility or the measured solubility in one solvent. What you need to find is a solvent with a rather low solubility at the lowest temperature of your crystallization process (it might be near 0°C) and a high solubility at the highest reasonable temperature. The constraint is the relatively small number of solvents generally acceptable to pharma. It would be useful if there were understandable rules for why different solutes have different temperature dependencies of solubility, but it mostly comes down to subtle entropic effects that are not captured in any intuitive way.

Getting the right polymorph

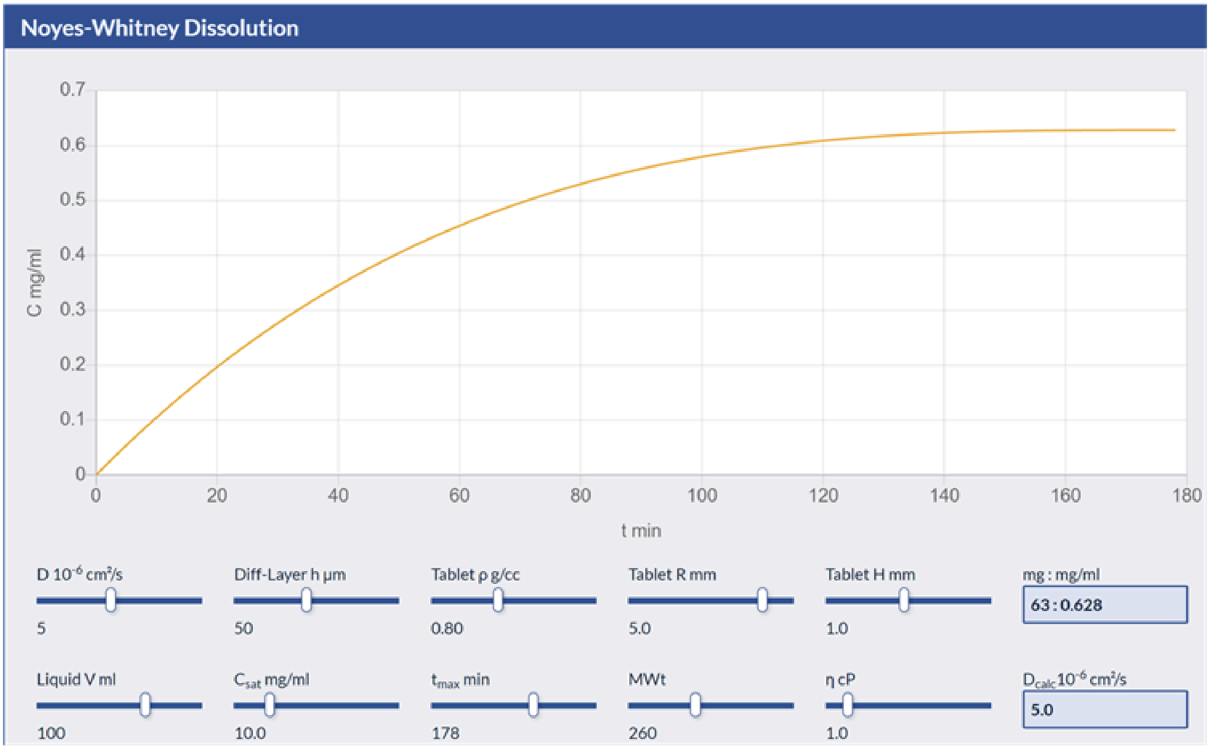
As discussed in the Crystallization chapter, your API crystals can grow in different crystalline forms, polymorphs. The different crystal packing of the different forms produces different shapes, MPts and enthalpies of fusion and, automatically, different solubilities. Given that pharma is all about certainty of dose, it is a requirement to produce the same polymorph every time. It is a catastrophe when something changes in a plant and the polymorph that has been grown for the past 10 years is replaced by a different one caused by, it is assumed, minute traces of seeds of the new one brought in to the plant as a contaminant (yes, such stories are true).

The problem isn't so much that the different solubility makes the new polymorph unacceptable – changes in solubility between polymorphs are usually minor. The problem is the *kinetics* of dissolution. If your drug has been designed around a release profile of, say, 1hr, then a factor of 2 in the kinetics is unacceptable.

The kinetics are, in turn, affected by two effects:

1. The way the crystals interact with the excipient (e.g. lactose) used in creating a tablet can have a profound effect on the tablet's rate of dissolution, affecting porous diffusion, mechanical strength and available surface areas of contact.
2. Crystal dissolution is limited by the rate at the slowest face, and if the crystal packing along this face is especially strong, dissolution is slowed down.

These effects end up affecting the "diffusion layer" in the standard Noyes-Whitney dissolution model:



D 10^{-6} cm^2/s 5
Diff-Layer h μm 50
Tablet ρ g/cc 0.80
Tablet R mm 5.0
Tablet H mm 1.0
mg : mg/ml 63 : 0.628
Liquid V ml 100
C_{sat} mg/ml 10.0
t_{max} min 178
MWt 260
η cP 1.0
D_{calc} 10^{-6} cm^2/s 5.0

https://www.stevenabbott.co.uk/practical-solubility/Noyes-Whitney.php

That diffusion layer is an imaginary distance between saturated concentration at the surface of the crystal and the current bulk concentration – the greater it is, the slower the dissolution.

Trying to increase aqueous solubility

There are two types of problem of getting your API delivered in an aqueous environment.

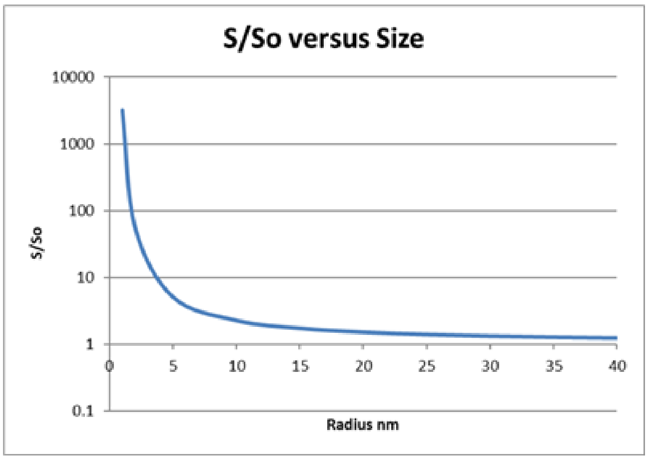
1. The API is relatively happy in water, but the ideal solubility is less than the required dose;

2. The ideal solubility is fine in some solvents, but the API isn't happy in water.

We discuss each in turn.

Ideal solubility too low

If the API is a salt, then changing the counterion to give a worse crystal packing and a higher solubility is a root-cause cure, but we now have a different molecule so we go back to an earlier stage in development.

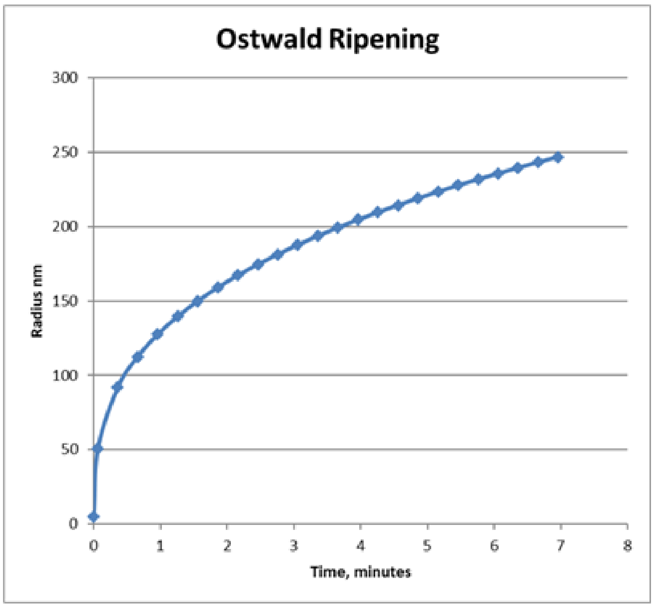


Assuming we can't change the molecule, then we have two tricks, neither of which works well in practice.

The first uses the Ostwald-Freundlich relationship which shows that for super-small particles with a large surface area to volume ratio, the surface energy effects overwhelm the crystalline packing effects and solubility increases. Where S_0 is the original (low) solubility, the solubility S for a radius r, surface energy γ, molar volume V_m and the usual gas constant/temperature term RT then

$$\ln\left(\frac{S}{S_0}\right) = \frac{2\gamma V_m}{rRT}$$

This sounds a great idea till you see, from the above graph, how fine you need to grind your particles.



Even if you succeed at getting your 5nm particles, Ostwald intervenes, this time in terms of Ostwald Ripening where larger particles grow at the expense of smaller ones – driven by the same surface energy forces that enabled the higher solubility. The theory tells us that the average radius at time t, $\langle r_t \rangle$ depends on the starting radius $\langle r_0 \rangle$, the surface energy γ, the diffusion coefficient of a typical molecule through the solvent (~10^{-6} cm²/s), the mole fraction solubility of the solute (obviously a low value, but still finite), and, as above, V_m and RT:

$$\frac{\langle r_t \rangle^3}{\langle r_0 \rangle^3} = \frac{8\gamma Dct}{9RT}$$

With c = 0.0001 mole fraction, a low solubility, then our 5nm particle grows to 250nm in 7 minutes.

The obvious fact that Ostwald ripening essentially destroys the prospects of using Ostwald-Freundlich to enhance solubility has not deterred countless researchers attempting to fight the laws of thermodynamics … and losing.

The second trick is to crash-precipitate the API so that it solidifies as amorphous particles. In the lab you can inject a solution in a water-miscible solvent, maybe with some ultrasonication, into water and remove the solvent via, say, dialysis.

With no crystal packing energy, their solubility is greatly increased (easily 10x). By making them nanosized, but in the 100's of nm range, their kinetic solubility is also increased. But we hit the same problem of Ostwald ripening, the particles, over minutes or hours steadily grow not only to a slower-dissolving size but also risk the chance of crystallizing over time without the constraints of the nano-sized particles.

Fortunately, as described in more detail in Surfactancy_Emulsion Ostwald there is a neat trick. If you can co-dissolve a very water-insoluble molecule into the API then as the API starts to migrate from smaller to larger particles, the concentration of the co-solute increases and the thermodynamics (call it osmotic pressure) brings the ripening to a halt.

Assuming that the amorphous particles are indefinitely stable against crystallization, it might seem that everyone would be using amorphous formulations. For lab and development experiments where having high concentrations of an API is useful, these formulations are of proven utility. It is worth the effort to find a pharma-acceptable Ostwald inhibitor. For end-user formulations there are a few problems:

- Regulators are happier with crystalline purity compared to amorphous materials
- Everyone is nervous about shelf-life. What might trigger the amorphous material to crystallize, maybe sitting on a shelf in a warm pharmacy?
- If there is a trace of crystalline material then a sort of Ostwald effect takes the API from the amorphous particles and the crystals grow rapidly, with crystal packing forces overriding the small osmotic pressure effect of the inhibitor.

Solubilizers

If the ideal solubility is OK but the API doesn't like water, then one obvious approach is to add some co-solvent, such as DMSO. This *might* have been acceptable for liquid formulations, but there are obvious issues of acceptability. And this doesn't help for a solid tablet where we need to increase solubility in the stomach.

So we add a "solubilizer" – a vague term that has caused much confusion. The root cause of the confusion is the well-known fact that some surfactants can increase the solubility of an API when it associates with the surfactant micelles. Tweens are the classical type of surfactant used for this effect.

S_0 mM 0.6

MSR 0.24

CMC μM 85

Max % 1

Increase 4.03

Surfactant MWt 1307

Micellar Aggr. No. 61

Surfactant Chain # 18

Solute Volume Å³ 350

Solute Area Å² 150

Solutes/Micelle 14.6

%Core Volume 10.0

%Surface Area 33.0

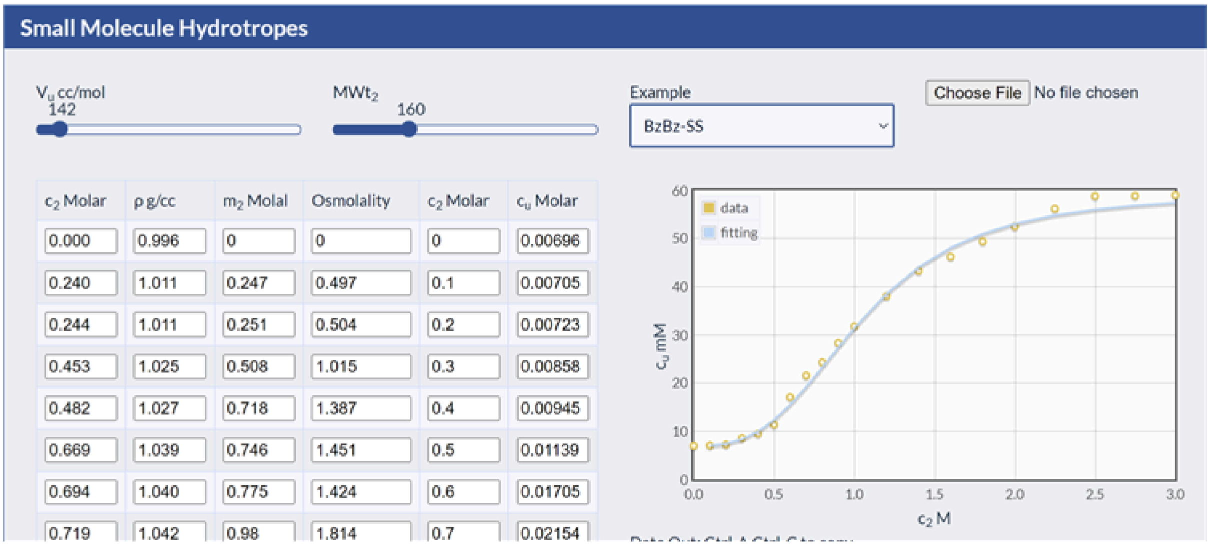
Log(K_m) 4.25

https://www.stevenabbott.co.uk/practical-solubility/surfactants.php

In this example we have Tween 80. The API has a starting solubility of 0.6mM and with 1% Tween we have a 4x increase. The hardly-visible kink in the curve near the start is where concentration rises above the CMC – below the CMC there is no solubilization effect.

The usual explanation for the solubilization is that the API is hydrophobic (which is why it's insoluble in water) and the Tween's C18 tail is hydrophobic so "therefore" the API is dissolved into the tail. This is obvious nonsense. The tail is something like hexadecane and the API is probably less soluble in hexadecane than in water. So where is the API? That's a good question. With the PEO head of Tween 80 there's a good chance that at least some of the API is happier in that environment than in the water. In the app, if you know the solubility enhancement at a given % surfactant you can slide the MSR slider (Molar Solubility Ratio) till you get the correct increase. Then you can look down at the number of solutes per micelle and, from the solute's size, what the chances are that they are filling the hydrophobic core or occupying the surfactant's head surface area. In reality it is likely to be a bit of both.

When we move away from real surfactants we enter a confusing world. Small molecule solubilizers are often called "hydrotropes" though the word has multiple fuzzy meanings. A typical hydrotrope experiment shows an S-shaped curve, as in the app, with an MHC, minimum hydrotrope concentration that reminds people of the CMC.



Small Molecule Hydrotropes

V_u cc/mol 142

MWt_2 160

Example BzBz-SS

c_2 Molar	ρ g/cc	m_2 Molal	Osmolality	c_2 Molar	c_u Molar
0.000	0.996	0	0	0	0.00696
0.240	1.011	0.247	0.497	0.1	0.00705
0.244	1.011	0.251	0.504	0.2	0.00723
0.453	1.025	0.508	1.015	0.3	0.00858
0.482	1.027	0.718	1.387	0.4	0.00945
0.669	1.039	0.746	1.451	0.5	0.01139
0.694	1.040	0.775	1.424	0.6	0.01705
0.719	1.042	0.98	1.814	0.7	0.02154

https://www.stevenabbott.co.uk/practical-solubility/kb-hydrotropes.php

Instantly we then get theories saying that the hydrotropes produce pseudo-micelles and that's a pseudo-explanation of the hydrotrope effect. As explained in Dissolution_Solubilizers and Hydrotropes this explanation is not only wrong (urea is a hydrotrope but shows no self-clustering) but backwards – those hydrotropes that self-cluster are less efficient than they otherwise would be. The real explanation, via Kirkwood-Buff theory is "solute-induced hydrotrope clustering". Although it is great to have a correct theory, sadly we have no predictive model based on it. At least by removing the bad theories there is hope that someone will put the effort into finding ways to predict the optimum hydrotrope.

Another class of solubilizer is represented by polymers ranging from the PPO-PEO-PPO triblocks to the tri-polymer, Soluplus, with its combination of PEO, polyvinylcaprolactam and polyvinylacetate. In its solid form it is friendly to many APIs as the HSP average of those three polymers places it nicely in the relatively high δD, δP and δH range of many APIs – so they have good solubility for controlled release. But Soluplus isn't fully soluble in water so provides a complex environment for APIs to partition into in some, as yet, unspecified manner. Hopefully someone will carry out a Kirkwood-Buff analysis of the effects of Soluplus, which would help disentangle the solubilization mechanism.

Despite decades of research, finding a solubilizer is still mostly trial-and-error. To sort things out via big data would require an approach that looks at different systems in a unified manner, and that can only mean a Kirkwood-Buff analysis of the key interactions. With high-throughput and some thought, it shouldn't be so hard.

Tablets

Grind the API with an industry-standard excipient such as lactose, along with any sub-ingredients, add the right amount of the mix to a tablet press, squeeze, eject the tablet and repeat. It sounds very easy. However, squeeze a bit too hard and the tablet won't dissolve at the required speed. Squeeze a bit too lightly and the tablet lacks the strength to traverse the manufacturing process without falling apart.

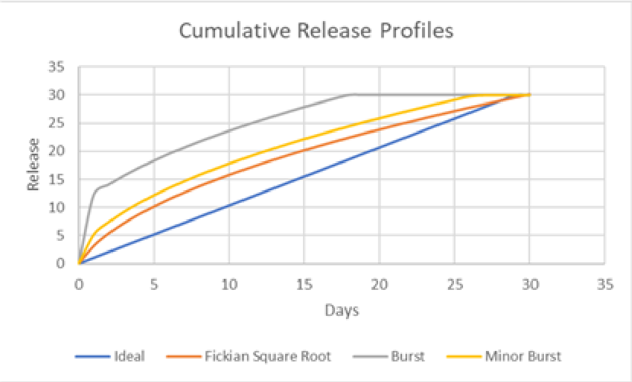
Sadly, the state of the art in understanding how particles adhere remains rudimentary, Particles_Basic behaviour. Different batches of the "same" lactose, without the API, can show very different compaction and dissolution behaviours for reasons that can't be pinned down via our standard ways of looking at the particles. It is interesting that Inverse Gas Chromatography (IGC) can reliably distinguish between two batches of the "same" lactose which show different tablet compaction properties. Frustratingly, no causal chain of logic has been identified to explain this, or to provide a root-cause fix.

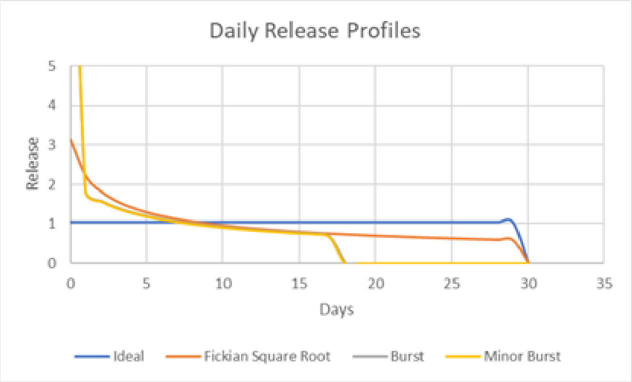
When you then add the API and small batch-to-batch variations, the chances of unexpected tablet behaviour increase.

The only reason for this section is to point out to the reader that their frustrations with tablet formation are not unique to them, and the reason they can't find the science to allow them to predict, optimize and troubleshoot tablet properties is that the science doesn't exist.

Controlled release

Dissolving an API in a polymer is a proven way to provide a controlled release from whatever polymeric form is chosen. The ideal is usually a continuous, steady release. The worst case is a burst, and a typical case is a minor burst followed by a relatively high rate falling off with time. Mostly we see a Fickian square root of time dependency on the cumulative amount, i.e. a steady fall-off of rate. We can see these in a pair of graphs – cumulative and daily release.





Ideal release can only be obtained by some sort of core-shell system where the release rate is controlled by the shell. Ignoring that, to overcome burst kinetics we need the API and the polymer to be reasonably compatible so that whatever method we use to get the API into the polymer results in a fairly homogeneous distribution. For this, knowing the HSP Dissolution_Hansen Solubility Parameters of the API and the polymer and calculating the HSP Distance gives us a good idea of (in)compatibility.

The square root of time behaviour can be obtained by pure Fickian diffusion from the API fully dissolved in the polymer, Diffusion_Basic Diffusion. But it can also be found from Higuchi kinetics where the solubility is much lower than the loading, so solid API particles are evenly distributed within the polymer. The key assumption behind Higuchi is that the rate of diffusion of the low level of API dissolved in the polymer is slower than the speed at which the polymer particles can re-stock the API in the polymer.

Skin delivery

The science discussed in SkinCare is generic, applying as much to getting an API through the skin and into the bloodstream as it does to delivering the desired ingredients in a skin cream. The key is knowing relative solubilities and diffusion coefficients of key ingredients. If you can't get enough API into your delivery system, then you can't deliver it effectively. If the API is compatible with elements in the delivery system, such as a Tween, that happen to be too large to diffuse through the skin, then the API will be trapped on the surface. If the different components of the formulation diffuse through the skin at different rates, then the composition on the surface will change; in the worst case this can cause a slow-to-diffuse (large MW) API to be stranded on the surface.

I learned this the hard way. A test formulation applied to my face used a “fast” solvent, DMI, which disappeared relatively quickly. The large API reached its solubility limit on the skin surface and crystallized … in the form of *very* painful sharp needles. I later learned of a real product recall in countries with high humidity. The cream worked fine everywhere else, but the high humidity caused the API to crystallize out … again into painful needles.

This is all very simple when stated, but for historical reasons the pharma skin world remained obsessed with LogP (Octanol/Water partition coefficient), ignoring straightforward science.

Another obsession was with surfactant-assisted delivery and the supposed wonderful attributes of liposomes and the many other “ome” variants. The surfactants aren’t surfactants after 20min when the water has gone, and there never was a reason why a large liposome would transport an active through the skin.

A special problem was that testing focussed on delivering from infinite doses of water-based formulations, measured with Franz cells, when something applied to the skin is finite dose and the water has all gone after ~20min. The more obviously sensible approach of examining what is on the skin surface over time was never adopted, so they missed the fact that water disappears, that APIs can get trapped, that fast-moving cosolvents might strand the API on the surface.

It is no coincidence that despite decades of research following the old methodology, the number of APIs delivered through the skin remains very small.

Science bandwagons are an enormous waste of precious resources. It takes only a few minutes to realise that surfactants on the skin are no longer surfactants when the water has gone in a timescale far shorter than skin delivery. But this didn’t stop decades of naïve papers comparing skin delivery between o/w with w/o emulsions. It equally took decades for the “ome” bubble to burst despite it being obvious that they couldn’t do what paper after paper claimed that they must do.

Franz cells continued to be used long after it was clear that they provided the minimum relevant information for the maximum work. Confocal Raman is a welcome but expensive way to get rather more information. But imaginative ways to directly sample what’s on the skin surface over time, which should be the gold-standard approach, are still rudimentary.

PSA

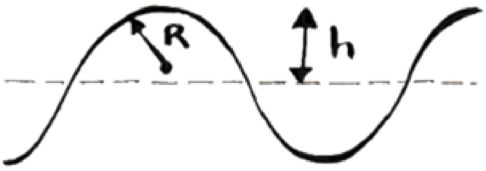
Links

Adhesion_Peel, Adhesion_Shear, Adhesion_Testing, Flow_G' and G", Flow_Oscillatory rheology, Flow_TTS-WLF, Mechanical_Hardness, Mechanical_Modulus

> Pressure Sensitive Adhesives rely on a tricky set of compromises. They must flow easily to create good surface contact, yet must put up a resistance to being stretched and broken. They can only be optimized via good rheology, but you can't (yet) create a great PSA via pure rheology because it's a big multi-physics challenge. They offer a reverse guarantee: the right science cannot guarantee success, but the wrong science guarantees failure. So it's good to get the science right.

PSAs aren't pressure sensitive

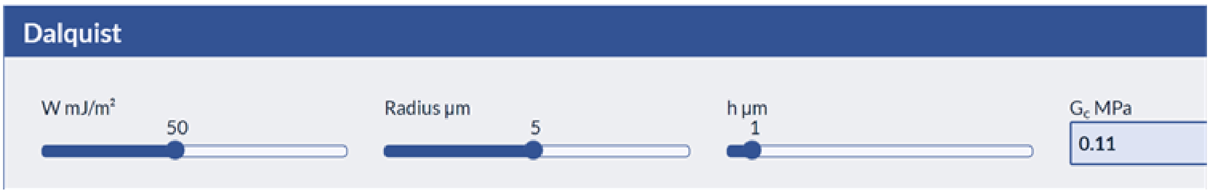
The adhesion of a PSA doesn't depend (much) on the pressure you apply to them. They get most of their adhesion via spontaneous flow into perfect contact with the substrate. At a minimum, therefore, a PSA must meet the Dahlquist criterion of a 1Hz G' (see below) no greater than 0.1 MPa. This automatically means that PSAs are weak polymers.



Dahlquist tells us that if we have a surface like this with an average radius of curvature R and height h then the PSA will flow into it if its modulus is less than the critical value G_C given by:

$$G_c = W\sqrt{\frac{R}{h^3}}$$

You can do the calculation here (an alternative formula is provided, linked to other roughness measures as described in Mechanical_Surface Roughness):


Dalquist

W mJ/m² 50

Radius µm 5

h µm 1

G_c MPa 0.11

https://www.stevenabbott.co.uk/practical-adhesion/dahlquist.php

There are four key take-aways from this:

1. if you have a very rough surface, most PSAs will fail because their G' will be in the 0.1 MPa region, too high to flow into a rougher surface.
2. If you want a "stronger" PSA and increase its modulus, it will fail because it won't flow into contact on a typical surface.
3. To get both "strength" and Dahlquist you need tricks like using a heat-applied PSA so that G' meets Dahlquist during application and reaches a higher value during use.

4. You need to understand G' and G" behaviour via rheology

How does a PSA stick?

Given that the PSA must be weak, how can it provide strong adhesion? As explained in the Adhesive chapter, the trick is to use viscous dissipation to stop the crack energy from running along the interface. In the language of standard PSA rheology (to be discussed shortly), we get viscous dissipation via G", the loss component. But although you need viscous behaviour, you can't make a good PSA from a purely viscous material – you need plenty of elastic strength, G'. You now see an instant contradiction – you can't have lots of G' because you will fail Dahlquist.

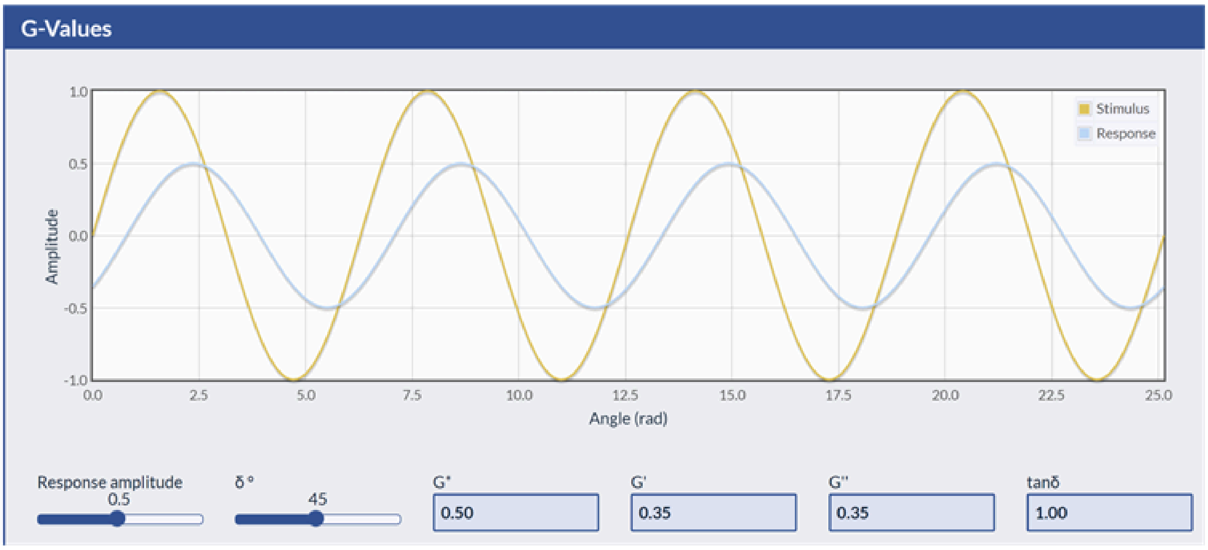
All this can only be disentangled via rheology. Here we have the reverse guarantee in the introduction. If you don't understand and control your rheology, failure is guaranteed. So although, because of the complexities of the system, you can't guarantee success via rheology, you are at least reducing the chances of failure.

Reverse guarantees? Smart mapping? Why this strange idea of reverse guarantees and the later one of smart mapping? Because there are times when we know that a direct formula/app-based approach can't work. These phrases deliberately alert you that "business as usual" cannot work and you have to adopt a different mindset to efficiently navigate the complexities of these formulations.

Rheology for PSAs

Although for other aspects of handling (e.g. during coating) you need Flow_Basic viscosity and Flow_Rotational rheology, here we focus on Flow_Oscillatory rheology to measure the elastic (G') and loss (G") moduli which you need to juggle, Flow_G' and G".

Basically you apply an oscillatory stimulus (e.g. a stress) and measure the response (in this case a strain). You might find something like this:



Response amplitude 0.5

δ ° 45

G* 0.50

G' 0.35

G" 0.35

tanδ 1.00

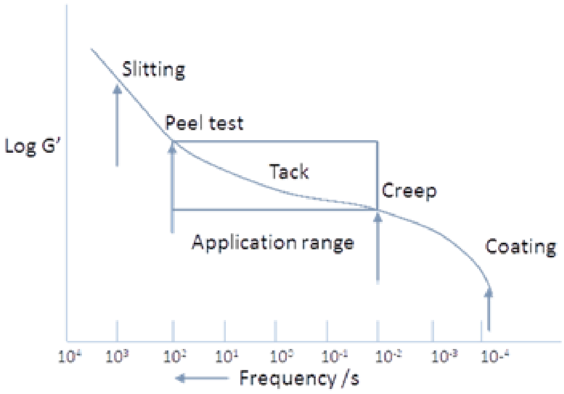
https://www.stevenabbott.co.uk/practical-adhesion/g-values.php

The response is lagging behind the stimulus in this case. If the material was purely elastic (all G') it would respond in phase. If it were purely viscous, lossy, (all G") it would be 90° out of phase. Here it is 45° out of phase which means it is an equal mixture of G' and G", which is the sort of zone you need to be in for a typical PSA. The ratio of G"/G' is called tanδ. When tanδ = 0, you have pure elastic, when it is very large you have mostly lossy behaviour.

Temperature, Time, WLF

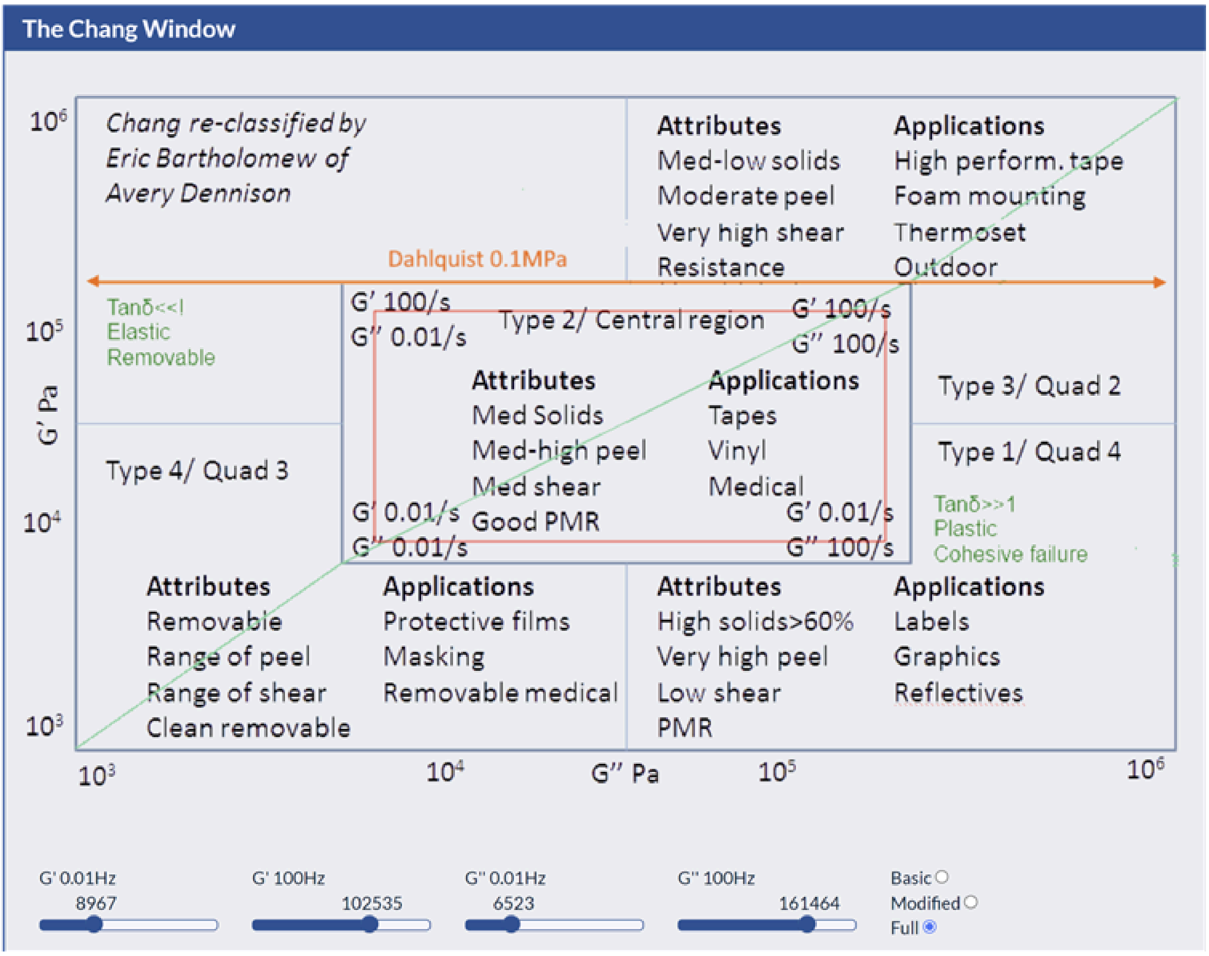
Although the measurement is valid, in terms of a PSA it is meaningless without further information. A PSA at low temperature isn't (unless specifically designed for this) a PSA. The polymer is too rigid, G' is too high. A PSA at high temperature isn't a PSA because the polymer is too soft, both G' and G" are too low.

Just as important is that if, at room temperature, you test the PSA at very low speed, it behaves as if it was at high temperature – it's soft and useless. If you test at very high speed, it behaves as if it was at low temperatures – it's brittle and useless.



Although your PSA may be designed for a modest temperature range, it will experience various timescales in use. If it is perfect when tested slowly yet fails at high speed, then that might make it useless for a specific application. Here are some typical timescales and an indication of how G' changes over those timescales. We shall shortly see what those G' numbers are, but you already know from Dahlquist that the Tack G' at 10^0/s (i.e. 1/s) must be ~0.1 MPa.

We can take this graph and look at it through the Chang Window, which says that in use you care about G' and G" in the region between 100/s (Peel test) and 0.01/s (Creep).



G' 0.01Hz 8967

G' 100Hz 102535

G" 0.01Hz 6523

G" 100Hz 161464

Basic ○
Modified ○
Full ●

https://www.stevenabbott.co.uk/practical-adhesion/chang.php

A standard PSA has G'/G" criteria broadly in the middle – in this example the central brown box represents the rheology of the product, with a range slightly smaller than Chang, but maybe that's a positive choice to optimize performance near the centre. There are 3 views of the same information, this is the “Full” version with a lot of other information added via Bartholomew.

Experts in PSA formulation are clear that Chang is another reverse guarantee. If you are in your desired window you still might not achieve success, but if you are outside, failure is likely.

A deliberate feature of Chang is that the values are defined at room temperature. So a high temperature PSA must have G'/G" values way too high when measured at room temperature (top right corner) which will shift to the centre when in use at those higher temperatures.

The right level of crosslinks & tangles

If you have a highly crosslinked or entangled formulation (in many ways the two things are the same) it will have a high G', spoiling the chance of meeting Dahlquist. A low level of crosslinks or tangles will be far too weak to be an effective PSA. So the formulator has to find ways to tune the level of crosslinks and tangles. With rubber formulations it's convenient to use styrene triblocks because the crystalline styrene regions act as crosslinks – and you can tune properties by the styrene-to-rubber ratio and (because the rubber itself can tangle) the MW of the blocks between the styrenes. For acrylics you can add low levels of di-acrylates – or rely on entanglement of high MW components. For UV PSAs you can create the required low level by tuning the levels of mono, di-, tri- functional acrylates.

)

An important aspect of these crosslinked systems is *strain hardening* discussed below.

As a beautiful proof that too much of a good thing is a bad thing, there is one type of UV PSA used in large quantities with a life between use and throwing away measured in minutes. It solves a problem in silicon chip manufacture where you need to flip 100s or 1000s of chips in order to carry out a process on the reverse side. You roll the PSA over the chips and can easily pick them up. After the required process, you probably need to transfer them to another bit of PSA for a different process. You apply the new PSA but how do you unstick the previous one? You give it a blast of UV that fully cures it. Now it is a brittle film with near-zero adhesion. It sounds unlikely but I once spent a couple of days on such a production line marvelling at the process while also helping to solve some technical issues with the process.

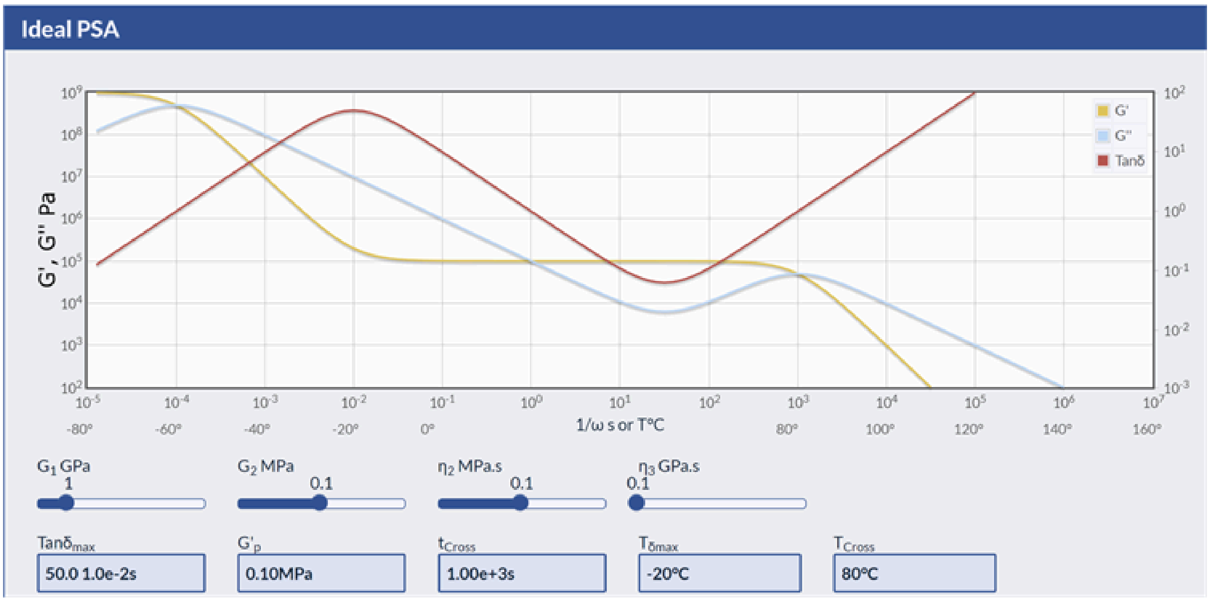
The Ideal PSA

When you get off a train near Amsterdam and the stranger who is there to meet you says “I’m a rheologist, would you like a beer?” the obvious thing is to forget the rheologist part and say yes to the beer. However, the beer was accompanied by some heavy rheology talk which was only about 5% comprehensible. But that 5% was enough to realise that something important was being taught … and that led to the app below. The wonderful rheologist is acknowledged in the app.

There is an amazing law of physics that says that “temperature is equivalent to time”. This is captured as TTE (Temperature Time Equivalence), TTS (Temperature Time Superposition) or WLF (Williams, Landel and Ferry who produced a convenient formula we all use), Flow_TTS-WLF.

If you do a set of rheology measurements at one temperature but different speeds, or different temperatures and one speed (or any other combination) you can re-plot them as if they were your chosen speed over a large temperature range, or a chosen temperature over a large speed range. So a single rheology plot can capture what happens to your PSA over a set of conditions way beyond anything you can test in the lab. If you are not doing WLF analysis of your PSA rheology, you are making formulation life very difficult for yourself.

For the classic rubber-based PSAs, the rheology can be captured in one “Ideal PSA” plot:



G_1 GPa 1

G_2 MPa 0.1

η_2 MPa.s 0.1

η_3 GPa.s 0.1

Tanδ$_{max}$ 50.0 1.0e-2s

G'$_p$ 0.10MPa

t$_{Cross}$ 1.00e+3s

T$_{\delta max}$ -20°C

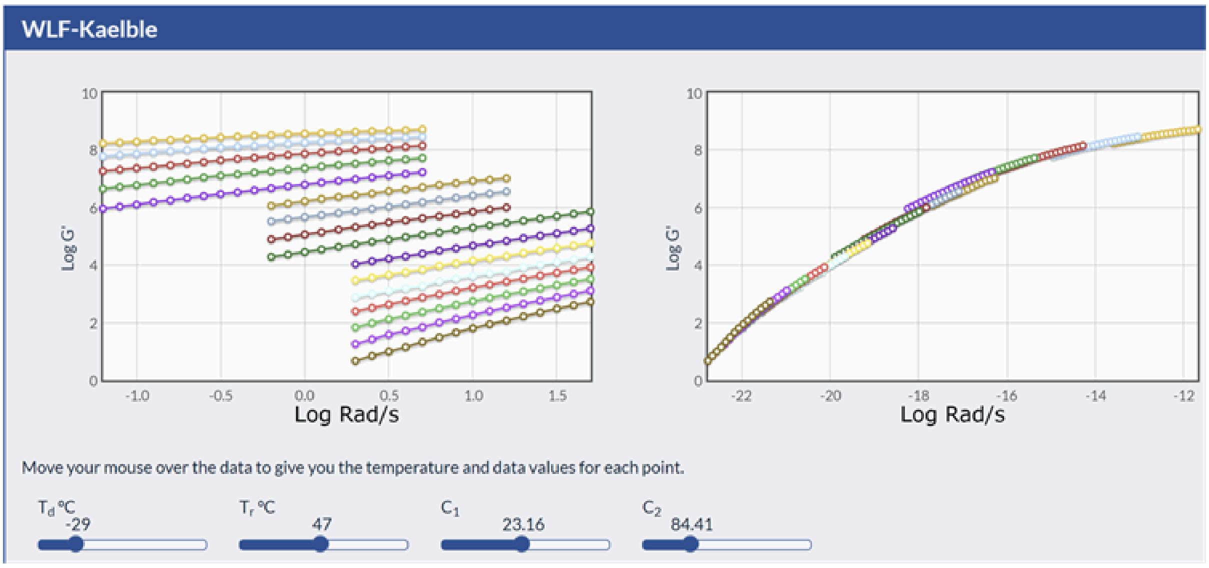
T$_{Cross}$ 80°C

https://www.stevenabbott.co.uk/practical-adhesion/ideal-psa.php

Here you are going from -80°C to 140°C and from 10^5 /s to 10^{-5} /s. If you explore the app you can learn a lot about plasticizers and tackifiers. In short, a plasticizer is some junk that takes your polymer's G' down into Dahlquist, but without adding anything else desirable to the formulation. A tackifier brings your G' down, but also raises the Tg, giving you a desirable higher temperature performance.

This is an idealised dataset and in the real world you will not obtain high quality data over such a large range, but the point is that a standard rheometer with a reasonable temperature controller can give a lot of data for little work. Although any single plot is hard to interpret, as you compare the data from formulations with different sets of product trade-offs you will learn how to take your current formulation to a different desired location in PSA space.

Although your rheometer software can do the WLF transformations for you, the WLF apps give you a good feel for what is going on:



Move your mouse over the data to give you the temperature and data values for each point.

T$_d$ °C -29

T$_r$ °C 47

C_1 23.16

C_2 84.41

https://www.stevenabbott.co.uk/practical-adhesion/WLF-K.php

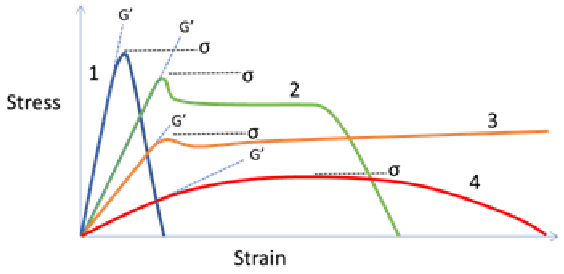
On the left we see the raw data from the rheometer at a range of frequencies and temperatures, on the right we see them all combined into a single curve described by WLF. While doing so we go from data gathered over 3 orders of magnitude frequency to the single curve over 11 orders of magnitude.

During some work on an unfamiliar rheometer we couldn’t find the WLF package in the rheometer’s software. In my hotel room that evening I had to work out how to do the transform in Excel. Although it’s trivial, it seemed very hard. In the morning I presented the fit – and they also had an overnight email explaining where to find WLF in their software. Fortunately the results were the same.

But that spreadsheet has been downloaded many times from my WLF page because everyone else finds their first WLF to be as hard as I did.

Rheology is not enough

You can’t formulate a PSA without rheology, but there is a fundamental issue with the rheological data which limits its applicability: it works over small strains (% stretching of the PSA), say 0.1%, while a PSA in action can easily see 100% strains.



To understand the behaviour at high strains, the ideal test is on a tensile tester. This is relatively unpopular, but gives the most direct information. More popular is the probe tack test which has a tricky relationship to PSA performance (described below) but is quick and convenient. In either case you get curves something like this:

The portion of the curve in the lower left is the elastic domain where you can get G'. After that you get various curves that show the yield stress σ then various elongations and failures. We certainly don’t want curve #1 because that’s just an elastic solid. We probably don’t want #4 because that is a weak solid. Something like #2 or #3 is more suitable for a PSA. The slight strain hardening in #3, discussed shortly, may or may not be desirable, it depends on the specific application.

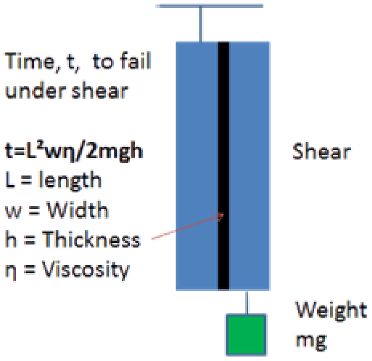
As with rheology, you can’t tell a lot from a single curve from on PSA, but as you compare and contrast good and bad PSAs you start to see the features you require.

Curve #3 shows a slight increase in stress with strain. This is *strain hardening*. Because the system is crosslinked/entangled, as you stretch it, it puts up some resistance. There is little agreement about how much strain hardening is optimal, but if (because you don’t do tensile tests) you have no idea if your system strain hardens or not, it’s hard to know how to optimise the level of crosslinks and entanglements.

Unfortunately, the WLF work in the hotel room showed that WLF G'/G" curves for their product and a competitor’s were identical, even thought the competitor’s product was superior. Some impromptu tests on an Instron-style machine with a light load cell showed obvious differences between the tensile curves: one showed strain hardening; the other didn’t. These differences could be translated into formulation insights which resulted in an improved, now market-leading, product.

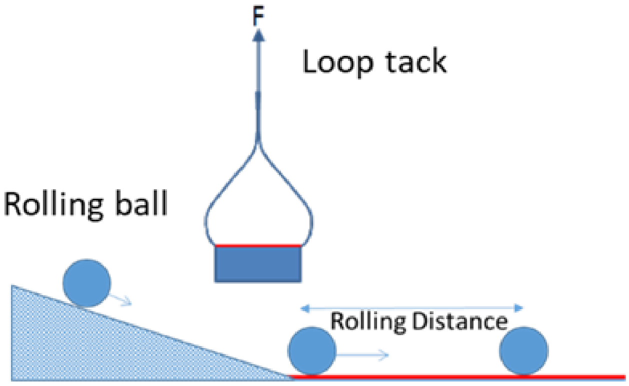
Testing PSAs

As the Adhesion_Testing topic shows, most adhesion tests provide little of the information you need to know on how to improve a formulation. This is true for PSA testing.



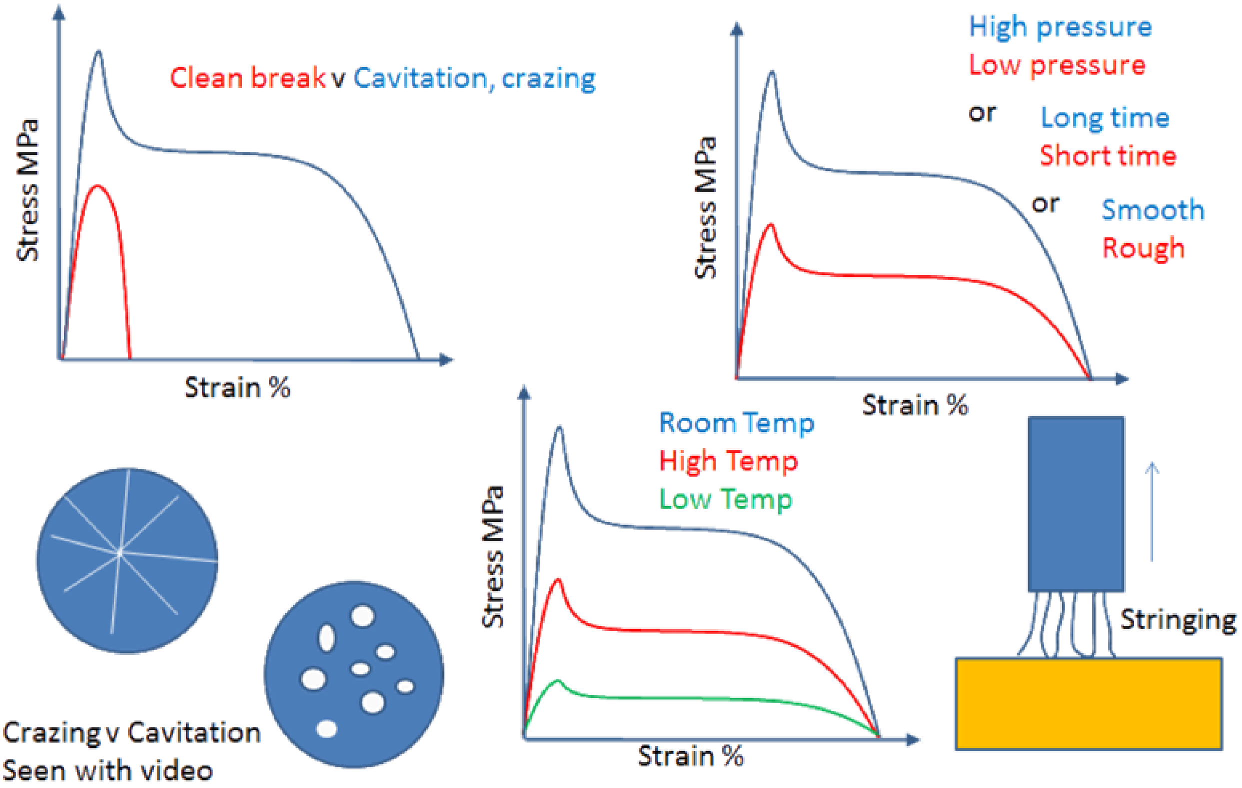
For example, the classic PSA shear test shown in the diagram is only a test for passing the shear test. It has very little relationship with the shear properties of the PSA for a good reason – it often fails in peel, a fact common in shear testing. So the formula shown in the image is not applicable, much to the puzzlement of people who think that the shear test provides information about shear behaviour.

If you really want shear behaviour, there are at least two tests (see https://www.stevenabbott.co.uk/practical-adhesion/psa-shear.php) that can do a better job at revealing the true shear behaviour. Shear properties *are* important, that is why, when you are optimizing the science, you should do the real shear tests rather than the industry-standard one. Of course your customers demand the industry-standard one, but first you have to get the real tests right.

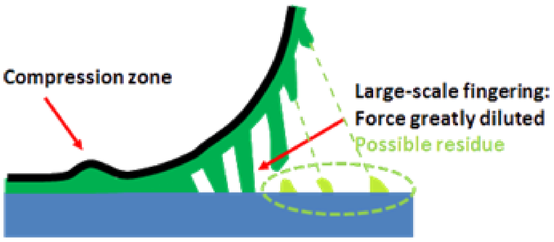


There are two tests that are scientifically valueless yet often used for QC testing: Rolling Ball and Loop Tack Tests. Although simple and convenient, there is a risk that their lack of scientific meaning allows you to miss crucial changes in the PSA properties.

Then there is the probe tack test:

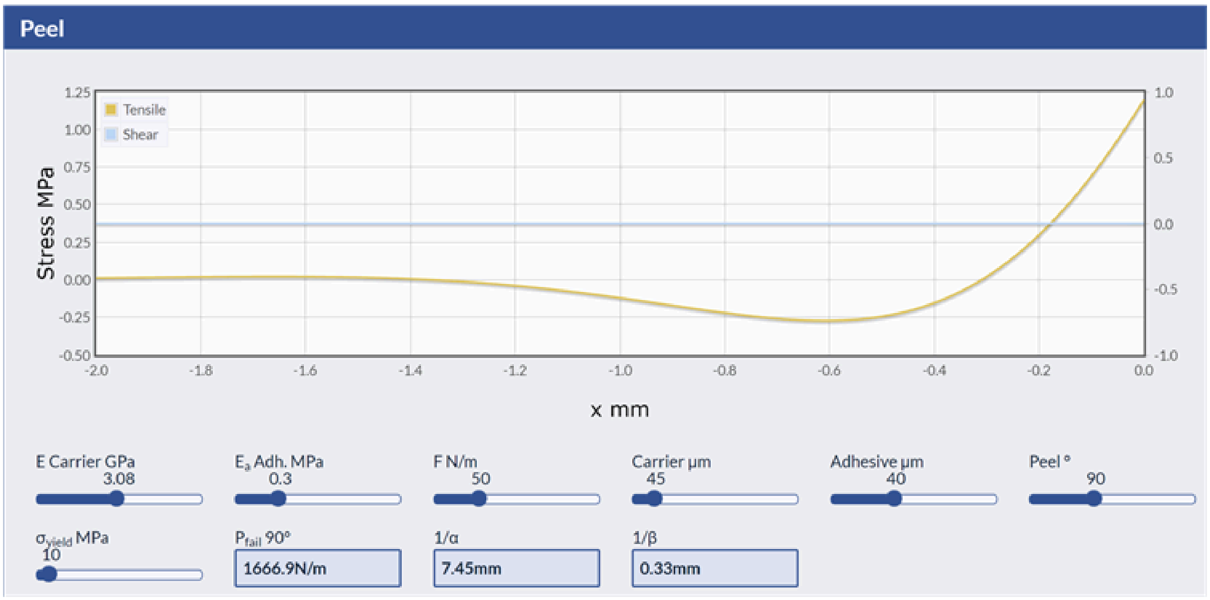


The probe tack test has lots of things happening at once, and because it is essentially a butt joint test, it is not directly relevant to the most important aspect of a PSA which is its peel. Using the probe test as an indirect tensile test has its merits, and the Simple Model described in https://www.stevenabbott.co.uk/practical-adhesion/Probe-Tack.php while anything but simple, is interesting. A recently developed direct link between probe and peel tests is described in this app, https://www.stevenabbott.co.uk/practical-adhesion/Peel-Tack.php, though it is still really for experts.



Which brings us to the peel test. This is at last a direct measure of what most of our customers (think they) want. Because adhesion is a property of the system, the peel test is not as simple as it might seem:

When we are peeling we can see the dramatic fingering part. But the compression zone is generally unknown despite, as we shall see, its importance. Although you can explore more in Adhesion_Peel, the core ideas can be found in the peel app:



E Carrier GPa 3.08

E_a Adh. MPa 0.3

F N/m 50

Carrier µm 45

Adhesive µm 40

Peel ° 90

σ_{yield} MPa 10

P_{fail} 90° 1666.9N/m

1/α 7.45mm

1/β 0.33mm

https://www.stevenabbott.co.uk/practical-adhesion/peel.php

The right-hand edge of the app is just before the first finger appears in the previous image. For simplicity we’re looking at a 90° peel – you can change the angle at any time.

The tensile stress is shown on the left-hand Y axis. You see that 2 mm ahead of the peel the stress is 0, as you would expect. But 0.6 mm ahead, the stress is -0.27 MPa, i.e. the adhesive is in compression. This is non-intuitive but very real. At the right-hand edge you reach the maximum stress, pulling the adhesive away.

The important fact to note here is that if you change the modulus or thickness of the carrier film, or the thickness of the adhesive, you change everything. In other words the peel adhesion properties of your adhesive depend on 3 parameters that are not part of the rheology of the adhesive itself. Yes, you must get the rheology right, but you should also attend to other aspects of your tape. For example, reduce the modulus of the tape and the curve “tightens up” – the compression gets deeper and closer to the peel edge, and the maximum tensile stress increases, reducing overall adhesion.

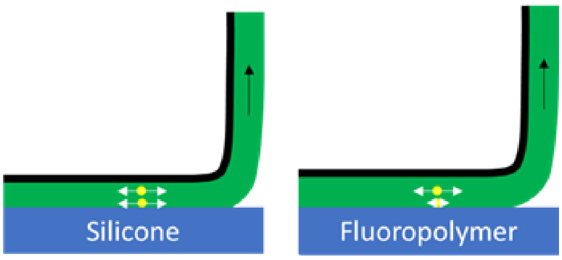
So a well-intentioned cost reduction via a cheaper, weaker, thinner backing tape can result in an expensive upgrade to the PSA formulation to restore peel strength to its original value.

I was asked about a customer complaint that the peel strength of a batch was lower than specified. In-house tests on QC samples at 25°C showed that the batch was in spec. It turned out that the customer was in a hot country and their lab was at 30°C. Using the additional calculator near the end of the WLF app, https://www.stevenabbott.co.uk/practical-adhesion/wlf.php, it was easy to show that if the customer performed the test 2.5x faster, that would be equivalent to them performing the test at standard speed at the specified 25°C lab conditions. The customer was astonished when the test at that speed proved the product to be in spec.

Release liners

Because silicone release liners have a low surface energy, it is usually stated that they are release liners *because* of the low surface energy. This is wrong. Fluoro-surfaces with the same surface energy show poor release. The real explanation has been known since the 1990s but the myth of surface energy still dominates.

The way that silicone release works is related to the compression zone ahead of the peel front!



If you put some fluorescent beads into the PSA and watch what happens as the compression zone arrives, you see that beads next to the surface and in the middle of the adhesive move no problem. This doesn't sound surprising, but it contravenes the "no slip boundary condition" which says that there is no fluid flow at a surface. With the fluoropolymer you see the normal behaviour, the beads at the surface are stationary.

Because in the fluoropolymer case there is differential motion as the compression zone arrives and leaves, there is viscous loss of crack energy, so the adhesive remains in contact. But on the silicone, there is no differential motion, no viscous loss, so the crack is able to travel easily along the interface.

The way the silicone overcomes the no-slip condition is by being relatively mobile itself. It's not the surface energy but the surface mobility that matters.

If you are trying to find a replacement for silicone release liners, finding ways to get surface mobility via some clever science will be more productive than trying to find alternative chemistries with low surface energies.

At the end of a talk on adhesion science in front of a large audience, someone asked a question to which I could only answer that I didn't even know what the question meant. Catching up over coffee with the (it turned out) distinguished professor who'd asked the question, led to a 15 min lecture on the science of PSAs. There is a direct link from that admission of total ignorance, through that 15 min lecture, to this chapter. The distinguished professor is Mikhail Feldstein and those trying to formulate PSAs without the magnificent 3-volume Handbook of PSAs and Products by Benedek and Feldstein are making life unnecessarily hard for themselves.

There are two types of ignorance. The first is that a topic has never come up in one's journey through life; so admitting ignorance is a chance to learn exciting new stuff. The second type is refusal to engage with correct ideas because they go against what you think you know. Although most of us are tempted by this second type, it's generally not a good way to make progress.

SkinCare

Links

Diffusion_Basic Diffusion, Diffusion_Diffusion Coefficients, Diffusion_Diffusion into Skin and Hair, Dissolution_Flory-Huggins, Dissolution_Hansen Solubility Parameters, Dissolution_Ideal solubility, Dissolution_Solubilizers and Hydrotropes, Evaporation_Basics, Evaporation_Humectants, Evaporation_Temperature and Antoine Coefficients, Flow_Basic viscosity, Flow_G' and G", Flow_Oscillatory rheology, Flow_Poiseuille Flow, Flow_Rotational rheology, Flow_Shear dependent, Flow_TTS-WLF, Flow_Thixotropy, Flow_Yield Stress, Fragrance_Barrier Properties, Fragrance_Vapour Pressure, Optics_Colour, Optics_Scattering and Opacity, Surfactancy_Dynamic Surface Tension, Surfactancy_Emulsification, Surfactancy_HLD, Surfactancy_Interfacial Tension and Rigidity

Although skin care involves cleaning the skin, that is covered in Soaps and Washing. Here we look at formulations designed to make the skin be, or look, better.

What problem are we trying to solve?

Skin has evolved to be an adequate solution to the twin problems of keeping good things in and bad things out of the body. Although it is routinely said that skin is an amazing barrier, it is relatively poor. It easily lets water in and out (by design) and many liquid chemicals such as solvents can easily get through. It is also readily damaged by a range of common household chemicals such as caustic, bleach, enzymes in laundry detergents.

But because the outer part of the outer stratum corneum is designed to be shed continuously it’s a smart system that can repair itself with little help from ourselves. The skin provides the surface with a range of nutrients via the hair follicles, which add to the nutrients provided to the skin biome. This array of interacting microbes, each looking after their own interests, provides mutual assistance against threats to the biome and the skin.

The skin will remain in a good state if we don’t carelessly expose the skin to excessive threats from sunlight, chemicals and abrasion, and if we remove dirt without devastating the skin and its biome via excessive (even if well-intentioned) cleansing.

So the key to a healthy skin is to do as little as possible. Washing with water cleans away most dirt and most sweat (which is mostly salts dissolved in water), and the lightest dose of soap does the rest, unless we’ve undertaken a really dirty job without good skin protection.

As is well known, if the skin needs repair from temporary exposure to an aggressive environment (e.g. too much sun), there are 3 ingredients proven to be of use:

1. Water
2. Glycerol
3. Mineral oil (or some modern equivalent for those who don’t like mineral oil)

The water does rather little because it mostly evaporates after 20-30 minutes from any reasonable formulation – though the glycerol is a good humectant (Evaporation_Humectants) and keeps some of the water as well as interact itself inside the stratum corneum to make it more fluid, less dry. Glycerol is not quite perfect – it ends up feeling sticky on the skin. The mineral oil provides some occlusion to stop water loss from within the skin … and lets you have plenty of glycerol without the sticky effect.

Excellent water/glycerol/oil emulsions with, perhaps, a little added fragrance have been around for decades and pose no new formulation challenges. The choice of surfactant for the emulsion will change with fashions (and is discussed below) and maybe some Thickeners_Associative Thickeners or Thickeners_Polymeric Thickeners need to be added to keep the emulsion from separating and to give the required "luxury" Flow_Yield Stress behaviour. But, again, such formulations have been known for decades.

Because all aqueous formulations containing nutrients (oils, surfactants, fragrances …) are attractive to microbes, the product must contain a proven preservative package. Whether the preservative ingredients are entirely synthetic, entirely natural or a mix cannot alter the fact that they are designed to stop living organisms from proliferating. They therefore, by definition, pose a potential threat to the living cells in and on our skin including, especially, our microbiome. The formulation problem is trivial – proven packages of preservatives are well known. The problem is a marketing one – finding a package that is provably great at inhibiting bad microbes whilst being entirely friendly towards the good cells and microbes and sounding good to the end users.

So this chapter could/should end at this point if we define the problem to be solved in the rational manner above. Instead, the problem to be solved is "How to offer hope to consumers that their skin will look like the images seen in advertisements?" Now we enter a much more complex world.

Cosmetic or Pharmaceutical?

Although the ~15µm top layer of the skin, the stratum corneum, SC, is far from dead (only the top few µm of skin cells are dead, but the skin biome is happy in that environment) it is a convenient fiction that if our formulation only interacts with the SC then we're probably not doing much damage (though biome research increasingly questions that assertion). If we want to add an oil, and if that oil is a large molecule that mostly stays on the surface, then that's fine. We can also add water, glycerol and the ingredients of NMF, natural moisturizing factor, already in the SC, such as amino acids, urea and pyrrolidones.

A large surfactant molecule such as a Tween is too bulky to migrate through the SC so is safely on top.

Vitamins might seem OK, but normally our bodies receive vitamins at low concentrations absorbed from the gut, not in relatively high concentrations delivered to a part of the body not used to them (apart from Vitamin D which is manufactured in the skin).

The problem arises when you want to do something significant to boost the skin's function by sending chemicals into the viable epidermis or deeper into the dermis. These chemicals are now in your body, with few barriers to getting into the numerous blood vessels near the surface of the skin. If they have a specific biological function then they are pharmaceuticals. There is a strong tradition of trying to deliver APIs (Active Pharmaceutical Ingredients) through the skin so aspiring to improve the skin in this manner is laudable … provided you then follow the regulations surrounding pharmaceuticals. As these regulations are (rightly) difficult, the skin care industry somehow has to promise that the wonder chemicals they add are doing great things to the skin biology without being pharmaceuticals.

This used to be easy because regulations were lax. You could claim to be doing wonderful things to, say, “boost DNA”. Who wouldn’t want a skin cream that boosted their DNA? Well, anyone with common sense. Cancer cells “boost DNA”. So you’d want to be sure that the boost from the cosmetic formulation was 100% upsides and 0% downsides. Which means pharma-grade proofs of efficacy and safety.

So claims are now carefully crafted by lawyers and marketing experts familiar with just how far they can make something sound sciencey and biologically effective without making a specific pharmaceutical claim that would need a decade of hard work to validate – in an industry that changes fashion in months.

Cosmetic claim validation

I once was asked to step in to deliver a 1hr training course on cosmetic claim validation. Being entirely ignorant of the topic I had to quickly educate myself. The task was made easier by the existence of a UK database of challenges to cosmetics claims, along with the decisions, and reasons, delivered by the authorities. The large companies frequently pushed claims to the limits and beyond. But “beyond” only meant a slap on the wrist after weeks or months of a campaign that was probably ending anyway.

One type of claim is “Contains X proven to do Y”. Although the claim is factual, it is now generally regarded as unacceptable if, as is usually the case, X does Y only at concentrations far higher than allowed in the formulation. Retinol, for example, is proven to do wonderful things to the skin at high levels that, inevitably, come with the risk of pharma-scale side effects. So retinol can only be added at levels that have an insignificant effect.

“Contains extract of [some exotic plant]” is a claim that leads the consumer to conclude that it has been added because this exotic extract must do something wonderful. This is really an ingredient rather than functional claim so is probably OK. “Tap into the healing power of [some exotic plant]” isn’t claiming that the product provides any healing power, but because it’s obviously designed to give the impression of an (unproven and, at the given concentrations, unlikely) healing power it’s probably unacceptable.

What about serious claims?

By using only ingredients approved by the cosmetics standards, the issue of cosmetic safety can be largely bypassed. So now the claims of, say, “proven to reduce wrinkles” need to be validated.

In some cases, you don’t even need statistics to validate the claim. Any reasonable “filler” formulation (by definition) fills the spaces between wrinkles, and everyone can agree that wrinkles have vanished for as long as the filler stays in place.

For most other claims, the industry demands, and gets, reasonable statistics, taking into account sample size and statistical differences. Those lucky enough to have a wonder formulation will need only a small sample size. The more common formulations need a larger sample size (expensive) from which to extract a statistically significant effect.

One reason for this rigour is industry standards. But a key force driving those standards has been competitors who have been taken aback by a claim from a competitor that sounds too good to be true and who then come together to show (there are famous historical precedents) that the claim is indeed invalid.

Claims can be subjective measures such “X% of users found that Y improved the Z aspect of their skin”.

The challenge here is to find a relevant control. For typical skin creams, a water/glycerol/mineral oil control performs distressingly well. Tricks like packaging the control in a way that makes it look cheap (or using no fragrance), and the real product in a luxury package (with a luxury fragrance) benefit strongly from the placebo effect so their objectivity has to be questioned.

Or they can be objective measures such as:

- Changes in TEWL (Trans Epidermal Water Loss)
- Objective reductions in wrinkles
- Changes in elasticity
- Changes in colour (e.g. for age spot reduction)

Here the key issue is that the claim “90% users showed a reduction in wrinkles” can be factually correct but irrelevant – the scientific measure is correct but the consumer cannot see the subtle differences detected by the equipment.

The history of cosmetics claims shows that either they are great results based on levels of chemicals that take them into the pharmaceutical zone (as with retinol) or are marginal results of no significance to the user. This claim is rather obvious – if any company had a truly superior product and the patents to protect it, they would dominate the field.

Your job as a formulator is to make sure that at least one molecule of the wonder additive is present (“Contains X” must be factually correct, and presumably one molecule makes it correct) … and more importantly that the additive doesn’t mess up your carefully balanced formulation, discussed below.

Getting your molecule through the skin

For decades, the standard technique for measuring the penetration of a desired molecule, let’s call it an “active”, through the skin used an “infinite dose Franz cell”, i.e. on one side of the skin sample was a container with a large amount of the (aqueous) formulation and on the other side was some buffer that could be sampled over, say, 24 hours to measure how much of the active made it through the skin.

This has been maximum work for minimum relevant information. For any skin formulation that is water-based, the water evaporates after 20-30min so the interactions with the skin are nothing like those in the Franz cell. And much of the formulation that was on the skin disappears after a few hours through the contacts of daily life. So decades of research has produced data of essentially zero value to the problem. One confusion from this work was that LogP (water-octanol partition coefficient) was a key determinant of skin permeation. It simply isn’t. Another confusion (helped by LogP) was that molecules permeated via the “lipid only” route through a “brick and mortar” skin. Again, this is simply untrue, not least because the “lipid” portion is full of molecules such as cholesterol containing plenty of polar functionalities.

Another confusion was the role of surfactants. When the water doesn’t evaporate in the Franz cell, the surfactant is a surfactant. But in a real formulation, when the water goes, the surfactant is no longer a surface active agent as there is no longer a meaningful water:oil interface. The surfactant therefore becomes another molecule in the ingredient mix. If that surfactant is a Tween, a large molecule that does not penetrate the SC, then on the surface you have plenty of alkyl chain, plus a lot of ethoxylate oligomers that can act as an attractive environment for some of the active ingredients. They are then stuck in this environment and cannot deliver any of the desired benefits to the skin. This is an obvious fact, but one that was not obvious to decades of researchers who kept insisting that surfactants acted as surfactants on the skin.

As we know from Diffusion_Basic Diffusion, diffusion depends on the concentration gradient across the barrier plus the diffusion coefficient of the molecule in the barrier. The diffusion coefficient can be approximated by something like a MW^2 effect, leading to the rule of thumb that anything above 500 Daltons won't permeate much. Obvious exceptions to this rule are discussed below.

To get a high concentration gradient needs three things:

1. A high concentration of molecule in the original formulation (once the water has evaporated);
2. A good partition coefficient into the barrier;
3. A semi-liquid environment at the surface to deliver a continual supply of the ingredient to the skin surface.

To address the first part, if we consider the formulation after the water has evaporated to be a mix of oils, emollients, hydrotropes, surfactants (such as the Tween example above), humectants (and more) then we need to estimate the solubility environment and how friendly it is to the active ingredient. The only viable tool to do this (and hence, much used in the skin world) is Dissolution_Hansen Solubility Parameters. We can create a volume % weighted average of the formulation to get its HSP then compare it to the HSP of the desired ingredient. If the HSP Distance is small, they will be compatible, if far, incompatible.

To address the second part, we need to estimate the solubility environment of the skin. Far from being "lipid" (which would imply HSP parameters in the range of [16, 2, 2] typical of ingredients such as IPM (isopropyl myristate), it turns out that the skin is something like [17, 8, 8]. For those who believe in the lipid-only route, this non-lipid value emerges quite naturally when you include the HSP of cholesterol (present at ~20%) into a mix with the typical lipid value. Knowing 1 and 2 and the HSP of the desired ingredient we can see how likely it is that the ingredient will stay above the skin or partition into the skin.

The third part can be tricky. Suppose you have a lot of smaller molecules that together give an HSP not far from [17, 8, 8]. These will partition into the skin and start their journey diffusing into the SC. With a low MW their diffusion will be relatively fast. If the active molecule has a large MW it will diffuse slowly … and can get left behind, stranded on the surface and/or locked into any other large molecules left on the surface.

Beyond Franz cells

Because of the multiple interactions of the multiple components in a formulation, it makes no sense to use Franz cells to analyse a single component coming through the skin. We need to know which parts of the formulation (including our active) go into the SC. Those with confocal Raman spectrometers can get a good idea of this, but the signal of the skin itself is complex enough, and changes when the formulation enters, so it's optimum only for looking at a few components with strong, distinctive absorption bands. Far better is to take samples off the surface of the skin at appropriate times and use advanced mass spectrometry to characterise the composition. You know the quantities at the start, so can normalise the signals. As time goes on you find out what is staying and what is leaving.

If the component is not on the surface then it's in the SC. Does that mean it penetrates the SC? Once inside, what else can it do? It's not coming out again, so unless it is actively destroyed by the biology in the skin, it's going through – maybe slowly, but inevitably.

Although such a technique is not simple in practice, it is so significantly better than the alternatives that it should become the norm.

What happens next?

The active and maybe some other ingredients have gone into the SC and will inevitably get through to the other side. The SC is designed to be a barrier. The epidermis is not a barrier – it is a relatively open, relatively aqueous environment. So unless the active is amazingly hydrophobic, it will continue through the epidermis and reach the numerous blood vessels in the area. There is no good reason for most actives to stop once they've made it through the SC. If this is OK with the consumers and the regulatory authorities, that's fine. Otherwise it's a problem.

Biology

So far we've just used simple solubility and diffusion physics to analyse what's happening. In this biological system there are other considerations.

The rule of thumb around 500 Daltons has some obvious exceptions, especially with peptides that can easily be in the 1000+ range yet disappear rapidly into the skin. This can only be via an active process. The existence of "transporter systems" in all cells is both well-known and under-appreciated. Cells can't rely on molecules passively wandering in and out – they are actively included and excluded via transporter systems. Whether cosmetics could/should take advantage of transporter systems takes us to the cosmetics/pharma boundary once more.

The other way for biology to intervene is if the ingredient of interest is actively trapped by a biological sub-system. If the ingredient is a "key" in a "lock and key" system inside the SC or epidermis then it delivers its function and goes no further. This is good news in terms of ingredients not penetrating further into the body, but maybe bad news if this is an effect that could be termed "pharmacological" as, again, we are at the cosmetics/pharma boundary.

The follicular route

The skin contains plenty of hydrophobic channels, hair follicles, crossing the SC, providing a hydrophobic flow from inside to out. A suitably hydrophobic formulation might be sufficiently soluble in that hydrophobic environment to have a chance to diffuse into the skin. But diffusion is generally slower than the flux through the follicles so without some pumping action in the reverse direction, nothing much happens. Because each follicle contains a hair, a massaging of the skin causes the hair to move which, in turn, provides a pumping action. So follicular delivery *can* be achieved if aided by some modest massaging of the skin.

Again, this is taking the active to a very-much alive system (the base of the hair follicle) so we have another cosmetics/pharma boundary question.

Formulating the emulsion

Once you have decided on the *type* of emulsion (o/w, w/o, macroemulsion, nanoemulsion, microemulsion) and decided (mostly for cost/marketing reasons) which class of surfactants to use, the first task is to make sure you are in the right part of surfactant space to have a good chance of creating the right, stable emulsion with relatively little effort and the minimum amount of surfactant. Remember that the type of emulsion can make no difference to anything after the ~20min it takes for the water to evaporate, meaning that there is no longer an emulsion. The type has an effect on the initial look and feel and for 10-15 minutes, so base your decisions on those timescales.

Given that HLB is useless, the rational approach is to use HLD: Surfactancy_HLD. We start with (https://www.stevenabbott.co.uk/practical-surfactants/hld.php):

$$HLD = F(S) - k.EACN - \alpha(T - 25) + Cc$$

Assuming a low content of salts, the dependence on salinity, F(S) can be ignored and, for the moment, assuming that the temperature T = 25, to calculate HLD we need only to worry about the oiliness of the oil, its EACN, Equivalent Alkane Carbon Number and the “characteristic” of the surfactant, Cc, which goes from hydrophilic (large negative values) to hydrophobic (large positive values).

For an o/w emulsion we need HLD < 0, for w/o we need HLD > 0 and for microemulsions and efficient manufacture of nanoemulsions we need (temporarily for the nanoemulsions) to be in the HLD = 0 zone.

If we are using commercial surfactants, their Cc values should be available from the supplier. Failing that you can look them up on the Cc web page: https://www.stevenabbott.co.uk/practical-surfactants/cc.php. Given that the Cc of a surfactant blend is the mole-fraction ratio of the individual Cc values, it’s easy to generate your desired Cc if you start with surfactants in the right range. Don’t try to combine Tweens with Spans. Although this works in theory, each has a severe tendency to partition into, respectively, the aqueous and oil phases, rendering the assumptions behind HLD invalid.

The problem is the EACN of your oils. It is not hard to work out the EACN of any reasonable oil, using values from the EACN web page: https://www.stevenabbott.co.uk/practical-surfactants/eacn.php. The volume-weighted average of your oils is the EACN. The problem is that real-world oils are not pure. For example, octyl myristate might contain small levels of octanol and myristic acid. A triglyceride might contain small amounts of di- and mono-glycerides. In addition the oils will contain various fragrances with functional groups such as -OH or -C=O. It turns out that “polar oils” (ones with groups like -OH) exert a big effect on EACN values. There is no known way to predict the EACN of these real-world formulations.

Fortunately HLD allows us to easily measure the EACN of any mix … and how it might change if we tweak the fragrance or the oils or if there are batch-to-batch variations.

Knowing the EACN we can now choose the surfactant blend to get us into the desired part of HLD space. For an o/w emulsion we need -1 < HLD < -0.5 and for w/o we need 0.5 < HLD < 1. Sufficient emulsification energy will produce an adequate emulsion. For those concerned with taking the product into production, there are some subtleties about water-to-oil ratios as the batch is created so there is a need to look out for emulsion inversions: Surfactancy_Emulsion Inversion.

For effortless *creation* of a fine emulsion, the trick is to take the formulation to HLD = 0. Now the interfacial tension Surfactancy_Interfacial Tension and Rigidity is super low and emulsion drops form easily. For a microemulsion that is the end of the story – they are thermodynamically stable and can be created by simple mixing. For nanoemulsions and stable o/w or w/o emulsions you create the emulsion at HLD = 0 then change one of the parameters in the HLD equation to place HLD safely into the optimal o/w or w/o zone.

The classic way to do this, which allows wax ingredients to be melted into the blend, is to heat ethoxylate formulations to a temperature where the net +0.06(T-25) term takes a negative HLD to 0. At that “phase inversion temperature” (PIT) the emulsion is created with just standard mixing. Rapid cooling to room temperature keeps the fine emulsion. For a standard emulsion “rapid” can be slower than for a nanoemulsion where Ostwald ripening (Surfactancy_Emulsion Ostwald) is a serious problem.

With some consumers wanting "ethoxylate free" formulations (worries about 1,4-dioxane) and/or the use of greener surfactants such as APGs or polyglycerols where the α term is ~0 and ~-0.01 respectively, the trick is to start the formulation with the higher Cc component so that HLD is ~0 then quickly add a more negative Cc component to take the formulation to the desired zone. Similar tricks could be adopted using part of the oil to start at a lower EACN then raise the EACN to take the HLD to the desired negative value. These ideas are unfamiliar because of the historical preference for HLB. As HLD is becoming more accepted, these ideas will become increasingly normal.

The rules of HLD for formulations with plenty of glycerol or other humectants are unknown – the effects on oils and surfactants will vary. Again, the standard scanning method used to measure EACN effects can be used for humectants. With a set of test tubes, some standard formulations and a convenient robot, it doesn't take long to map the new surfactant space and to formulate accordingly.

What about alternative to surfactants such as "hydrotropes" and "co-surfactants"? This takes us to a world of marketing speak. "Surfactants" are bad while "hydrotropes" are good. Maybe hydrotropes are not-very-good surfactants and may be added to "boost" some property of the main surfactant. In that case it's a co-surfactant. Other co-surfactants are things like "linkers" that go to the head/tail interface and potentially increase efficiency by making the surfactant head and tail longer, with an extended influence. Sometimes these hydrotropes and co-surfactants are simple alcohols like ethanol or octanol.

In general these extra ingredients with ill-defined names and functions have been added by those lost in surfactant space and in need of a tweak via a relatively simple ingredient. The problem, as discussed below, is that every extra ingredient added to solve some specific problem ends up as a permanent part of future variants of that formulation – because future formulators don't want to risk removing something they assume has some vital, but unknown-to-them function.

One of the beauties of the HLD approach is that with an approach based on core scans for the main ingredients and sub-scans for the effects of extra ingredients, it is possible to record that, for example, the addition of X shifted HLD too far in one direction so Y had to be added to bring it back. A future formulator, not needing X, can revert to the original HLD and work out what happens when they add ingredient A and whether they need B to fix the problem that A creates.

Formulating for surfactant efficiency

The ideal surfactant would be low cost, green, kind to users' skin, tuned to the desired part of HLD space … and be highly efficient in producing emulsions. At its simplest, efficiency, assuming 100% partitioning to the emulsion interface, is the ratio of head area to MW, as you can explore in the emulsion surface area app: https://www.stevenabbott.co.uk/practical-surfactants/emusa.php. If you have a large head area for a small MW, a small amount will cover a large area of emulsion.

Once you get into microemulsions, efficiency is more related to the de Gennes ξ parameter used in the more sophisticated HLD-NAC theory: https://www.stevenabbott.co.uk/practical-surfactants/nacmore.php. For this edition of FST HLD-NAC is not discussed, but the apps and the description in Practical Surfactants https://www.stevenabbott.co.uk/practical-surfactants/the-book.php are readily available.

Formulating for feel

We don't know how to formulate for a perfect "feel" of the product – it is a subtle balance of effects best judged by trained tactile experts.

We can eliminate a popular factor – surface tension. This isn't just because surface tensions of most cosmetic oils (excluding silicones) cover such a narrow range – at most a factor of 2. And it isn't just because surface energy forces are factors of 100s or 1000s too small compared to the shear stresses during rubbing. It's not even because the formula for calculating the stresses in Flow_Couette Flow does not contain surface tension or surface energy (it assumes the no-slip boundary condition). It's because when you rub the lotion, the value you've carefully measured (oil-air interfacial energy) is irrelevant because both surfaces of the lotion are in contact with skin.

Silicone oils are different because, for subtle reasons due to the siloxane bonds, the silicones at the skin surface act as a pseudo-liquid, overcoming the no-slip boundary condition and giving much lower shear stresses.

If you find a blend of (non-silicone) oils that give a special feel, you can be certain that the effect isn't due to surface tension and is unlikely to be due to the viscosity effects discussed next. It probably means some sort of phase separation, maybe including water that might be in the formulation, leading to some sort of liquid-liquid boundary that, again, overcomes the no-slip condition that leads to shear stresses in Couette flow.

But we *can* formulate to avoid obviously bad feel.

The starting point is rheology, which gives you more understanding than just a few viscosity measurements.

Yield stress. Without the correct yield stress behaviour, Flow_Yield Stress, the product is an instant failure. Some consumers want instant flow, with no yield stress, others want a "rich" blob that retains its shape until gently rubbed. Because yield stress is hard to measure objectively (there are at least 6 different ways to measure it!) the trick is to find a set of samples (your own, or competitor products) that have been judged by your tactile experts. Now measure their yield stress behaviour as best you can on your specific equipment. If there is a good correlation between measured values and the experts' judgement of flowability then you have a good basis for testing new formulations. If there's a lack of correlation, have a go at measuring yield stress via an alternative route.

Flow curve. Next is a flow curve showing the shear-dependent behaviour, Flow_Shear dependent. With some tweaks, the flow curves can also be used to estimate thixotropy, Flow_Thixotropy, i.e. the time-dependence of viscosity. There's a chance that neither of them will correlate much with feel, but will be relevant to manufacture, pourability, fillability and maybe to the user's experience of squeezing the product from a tube or dipping their fingers into a jar. Something as simple as Poiseuille flow calculations for pumping product through pipes (Flow_Poiseuille Flow) benefits from estimates of the shear rate (flow velocity/ pipe diameter) and, therefore, the relevant viscosity during pumping. The consumer-relevant aspects are probably best checked by correlation between flow curve and thixotropy measurements and in-use evaluations by hands-on experts.

Rubbing in. The rheology of a formulation being rubbed into the skin is too complex for most of us. A 10 µm layer being rubbed at 1 m/s has a shear rate of $1/10^{-5} = 10^5$/s which is impossible for most practical devices. This is unfortunate because glitches in the formulation can show up as an unpleasant lack of smoothness during rubbing. One trick is to place a thin layer between two glass slides then apply a one-off rapid shear. Under the microscope, formulation failure modes such as forming temporary fibrils from polymers that could not respond sufficiently quickly (relatively slow relaxation kinetics) can show up, perhaps under crossed polarizers. It's not a reliable technique, but it's simple, and great when it shows up something unexpected from an otherwise well-behaved formulation.

I once got lucky with this technique in the lab of a famous Italian cosmetics formulator. He couldn't understand why an otherwise great formulation didn't "feel" right. We tried this technique and could instantly see that one key polymer was stringing under "rubbing in" shear forces. He was most impressed to have learned so much for so little effort.

Stickiness is obvious to a human and hard to identify with objective measurements. A "probe tack tester", a flat-ended cylinder pushed gently into the formulation then withdrawn at controlled speed with measured force gives some idea of the instantaneous resistance (close to an elastic response) followed by some degree of resistance via any tendency to form strings. Because any polymers in the system (e.g. thickeners) have natural relaxation times, the probe test needs to be done at different speeds to see if a formulation that offers little resistance at low speeds becomes much stringier at (relevant) high speeds, faster than the polymer relaxation rate.

For "thin" formulations, probe tests will be at a limit of measurability. A more general technique is to measure G'/G" (Flow_G' and G") over a wide range of temperatures and speeds in your rheometer and to use time temperature superposition/WLF (Flow_TTS-WLF) to give you the elastic (G') and plastic (G") responses over a super-large speed range (outside the real measurement speed of the machine itself). As above, there is little chance of you directly going from measured values to formulation insights. Instead the idea is to compare the curves of "good" and "bad" formulations and gain insights into, say, speed onset of high G" behaviour and (by implication or from your rheology software) insights into relaxation rates of components in your formulation.

Time-dependent feel. All the previous measures take place on the full formulation. Although aqueous formulations lose their water within 20-30 min, the instant feel dominates perceptions. But consumers take note of how the formulation changes feel at least for a few minutes, so it's a good idea to check for stickiness measured on formulations with known water loss. The G'/G" rheology is probably the most suitable technique with, as always, no obvious prior knowledge of what measured behaviour is desirable, so a correlation with subjective feel data are the best to be hoped for.

Formulating for being microbe free

This is an area of scientific and regulatory complexity. But some principles are clear. If your formulation contains water then microbes can grow. So by law you need to stop this growth via a package of preservatives. By definition this means molecules that discourage or kill living organisms. So by definition, this means a risk that they are harmful to skin cells or skin biome. So we have the nightmare scenario for marketing – they must assure users that everything in the product is wonderful, chemical free and nourishing (indeed, microbes might find these to be super-nutritious), whilst admitting that it contains chemicals that kill or inhibit natural organisms.

Given that "natural" does not necessarily mean "good" or "safe", marketing can't hide behind the use of natural preservatives. Many parabens are natural, benzoic acid and benzoates are natural but they are demonised as "chemicals". Some "essential oils" such as thymol are otherwise known as phenols, and phenols are "chemicals" too.

The *scientific* advice is to go with a minimum package of approved, effective preservatives. But because in these areas the science is of little importance to marketing and consumers, the formulator has to juggle whatever are the latest fads from the latest internet scare.

Formulating for other claims

The chemicals for removing dark spots, exfoliating the skin and other specialised effects are beyond the scope of the FST.

For those who wish to claim that the formulation contains the soothing extracts of an exotic plant, contains vitamins, antioxidants, UV blockers (see Sun Screens) it is likely that levels are small enough to do little other than modify the EACN in a manner that is easily fixable by adjusting the surfactant ratio to obtain the desired Cc and, therefore, HLD.

There are three real problems:

1. Permeation through the SC into the living skin. If the formulation happens to encourage soothing extract to get through the SC, while safety tests on the soothing extract did not lead to permeation, then maybe something will go wrong – for example a small percentage of irritated customers.
2. Unforeseen interactions. If the soothing extract happens to be a wonderful stimulant for the growth of some microbe then the previously effective preservative package no longer works. If the antioxidant reacts with the UV blocker in some odd way, maybe there’s a malodorous side product that your end user wouldn’t appreciate.
3. Each extra ingredient might require some extra extra ingredient to avoid a long-term problem. That’s OK. But when the product needs a marketing upgrade to provide yet another claim, it is unlikely that anyone can remember which part of the formulation was required to solve a problem of a no-longer-needed ingredient. So formulations become ever more complex.

There comes a point in this over-complication of formulations where the formulator has to fight back: return to a basic formulation with the essentials then carefully add the marketing claim ingredients looking out early for unwanted interactions. Marketing will object that you are too slow, but ultimately this back-to-basics approach saves time and money. With luck, marketing can see that the new formulation, with fewer ingredients can be marketed as “pure beauty” or whatever wording they care to use. It’s win-win.

Soaps and Washing

Links

Cleaning_Boundary removal, Cleaning_Contact Angles, Cleaning_Surfactants, Cleaning_Temperature effects, Flow_Basic viscosity, Flow_Yield Stress, Fragrance_Activity Coefficients, Fragrance_Vapour Pressure, Surfactancy_Anti-foaming, Surfactancy_CMC and Langmuir, Surfactancy_Coacervation, Surfactancy_Dynamic Surface Tension, Surfactancy_Foam Drainage, Surfactancy_Foam Ostwald, Surfactancy_Foam Rheology, Surfactancy_Foaming, Surfactancy_HLD, Thickeners_Wormlike Micelles

The formulator has two challenges. The first is to create a soap that is effective for washing (note that we are not discussing clothes washing, covered in Laundry Liquids). That is trivially easy. The second is to create a soap that meets the ever-changing requirements of marketing. That is mostly easy, but has its challenges.

How to clean

Most dirt on our skin is easily removed with water and some effective rubbing with our hands, a fine cloth or, if necessary, a brush. It is easy to remove because most of the time we aren't "dirty" – we merely have some bits of dirt, sweat (which is mostly non-greasy), and maybe the dust and pollutants from a day walking around a city.

The water drastically reduces adhesion between particles and the surface (Particles_Basic behaviour) and the rubbing motion overcomes the no-slip boundary condition (Cleaning_Boundary removal).

A small amount of any surface active agent will wrap around generally hydrophobic particles and help them to wash away more easily. Here we meet the ambiguity of the word "soap". It has two meanings:

1. The sodium or potassium salt of a long-chain fatty acid mix, e.g. sodium "palmitate", the solid soaps;
2. A surfactant solution based on classics such as SLS or SLES, the liquid soaps.

Solid soaps

Whether you want to call a bar of the first type "natural" or not, it is some mixture of excess alkali and the glycerol produced as a by-product of the process of heating the starting triglyceride oil with the caustic. One key advantage of the generally high pH soap is that it helps to destroy the dead skin cells on the surface, leaving a shiny newer skin and taking any strongly-adhered dirt with it on the dead cells. Adding some phenol to the soap ("carbolic soap") helped with this – providing an effective, anti-bacterial but obviously harsh cleaning product.

The form factor of this type of soap – a solid bar – used to be perceived as a disadvantage, hence the creation of liquid soaps. Now it is seen as a low-waste, "green" form factor.

It turns out to be relatively difficult to add marketing ingredients to solid soaps – they tend to ooze out over time. So soaps are relatively unexciting products, requiring claims such as being "French milled" to encourage us to buy them.

The downside that the calcium & magnesium salts of the fatty acids are insoluble, giving rise to soap scum helped swing the market towards the second type of soap.

Bars are relatively inconvenient in use – where do they go when you are washing (soap trays are of varying utility), and how do they stay dry when not being used or transported (wrapped in beeswax?) in a wash bag? The fact that a bar of soap lasts a long time shows that we can get adequate cleaning from small amounts of surfactant. With a perception of a lower carbon footprint and of being "natural", there is some fight back against the liquid soaps that dominate the market.

Liquid soaps

Whether or not liquid soaps are more convenient and more effective, there is one overwhelming advantage if you are in marketing: you can add all sorts of ingredients to impress the users. Fragrances, "natural oils" (an odd ingredient when the soap is intended to remove natural oils from the skin), vitamins, antioxidants can all be added (as we shall see, these can impose some formulation challenges) to give the impression that the cleaning will be, somehow, better. Add an impressive volume of foam, and the consumer is convinced that the product must be good.

The main surfactant ingredients could easily be selected for cleaning efficacy, especially for oil removal. Applying HLD theory Surfactancy_HLD to optimise for Eötvös number (Surfactancy_Roll-up and Eötvös Number) could give excellent cleaning. But this is precisely *not* what is required. A really good cleaner would remove all the natural oils from the skin and destroy the skin biome. Instead the need is for a guaranteed poor-quality surfactant, in the hydrophilic part of HLD space and for this SLS and SLES are ideal. As these happen to give good foaming (Surfactancy_Foaming) under normal washing circumstances, and are low cost because they are used so much, it is not surprising that they remain the core ingredients.

Scientifically, SLS is proven to have a strong, and still surprising, destructive effect on the skin if left in contact at high concentrations for long times. That is, even more destructive than the already destructive effects of leaving high concentrations of surfactants on the skin for a long time. This bit of science of SLS has led to the demand for "no sulfate" surfactants which are conveniently provided by the isethionates and taurates which are scientifically *not* sulfates (they lack one oxygen) but which behave more-or-less the same so are convenient drop-in replacements.

The alternative ethylene oxide non-ionic class don't have the reputation of being harsh, but now have the problem that hyper-low levels of 1,4-dioxane need to be removed which is inconvenient and expensive. In any case, consumers often choose "ethoxylate-free" surfactants along with "sulfate-free".

The alkyl polygycosides (APGs) and the polyglycerols are, presumably, making steady in-roads, though those used to the "feel" of the previous standard formulations are less happy with these … along with the problem that these formulations foam rather differently (arguable "different" rather than "worse", but consumers again note differences) and are not easily thickened with sodium chloride.

Real cleaning

In the rare cases that we are really dirty, given that soaps are designed more for effect than for efficacy, we have two choices.

1. Carefully use a scientifically proven cleaning system – a laundry detergent, Laundry Liquids.

These are especially good for oils as they use the right types of relatively hydrophobic surfactant. It is interesting to note that these work very well at levels around 200ppm, compared to a typical 20,000ppm of a typical liquid soap.

2. Use lots of scrubbing with particles, brushes, sponges, pads

In both cases, the skin will be temporarily damaged, so make sure it's fully rinsed if you've used a good detergent and take care of yourself if you've used the rigorous scrubbing.

Luxury feel

It must have been a wonderful day for marketing when the formulators showed that they could produce exactly the same formulation at the same cost, but with a higher viscosity that made it look luxurious. The trick was to add a few % of NaCl which created Thickeners_Wormlike Micelles which provided the magical combination of being "rich" at low shear (they have a significant Flow_Yield Stress) but under shear handle more normally so aren't too difficult to apply.

This simple advance has come at a huge cost to the formulation community. Wormlike micelles are complex, and their structure/behaviour is easily changed. As is well-known, addition of slightly too much NaCl tips their behaviour away from having a high yield stress. But if for the past year you have had a great formulation with a fragrance package X, when marketing require fragrance package Y, the wormlike magic fails … and there is not a serious body of science that allows formulators to tweak things to restore the viscosity. The same applies to the oils, vitamins, antioxidants that marketing wish to be added.

At a conference, a spontaneous sub-group formed during a coffee break to swap stories about the huge waste of time, resources and energy caused by small changes in a formulation or the manufacturing process that tipped the wormlike micelles into an unhappy state. Sadly, between us there were no solutions to the problem – even though it is huge for the whole industry. The root cause solution, selling soaps without this unnecessary viscosity effect, was not even discussed.

It also means that sulfate-free and ethoxylate-free formulations, demanded by consumers, are hard to produce because APGs, for example, are unaffected by NaCl. Vast amounts of formulation effort have gone into producing products with an effect, viscosity & yield stress, that is only a marketing requirement, not a genuine consumer benefit.

Of course there are opportunities to use Thickeners_Associative Thickeners. But it is sad that we have to add extra chemicals to an already over-complex formulation merely to add a marketing feature.

Luxury foam

Foam has no role in skin cleaning. Any surfactant tied up in a bubble is not at the skin surface doing something. So foam is a bug, not a feature. But because it is hard to make a surfactant formulation that doesn't foam, the standard marketing approach is to convert a bug into a feature. Most consumers would be offended by a non-foaming soap.

In standard lab foaming tests, which bear little relationship to real foam generation during washing, combinations of CAPB with SLES give somewhat more foam that SLES alone. Hence we see SLES/CAPB formulations as a common base formulation to which the other marketing attributes are added.

The science of foaming (Surfactancy_Foaming) shows that true foamability is not strongly dependent on the surfactant – you can get great foam from just about anything if you put in the right amount of shear energy. So a not-very-good foamer can produce large volumes of foam if, for example, you use a Japanese shaving foam net. Why anyone wants a lot of shaving foam (which provides no actual shaving benefit) is a different question, the point is that you can get lots of foam even from something that happens to fail "standard" foam tests. The common shower net creates lots of foam by the same high-shear mechanism.

It is said that foam offers two advantages:

1. It provides a signal that surfactant has been applied to all relevant parts of the body.
2. It helps to spread the otherwise hard-to-spread surfactant formulation

Given that most of the time we are not trying to clean off lots of tough dirt, and that the benefits of the surfactant are modest at best, it's probably more important that all relevant parts of the body are wetted and wiped (the true cleaning process) than guaranteeing that surfactant has reached everywhere.

The "spreading" argument is necessary because the surfactant formulations are high viscosity and, therefore, more difficult to spread. The marketing feature of luxury viscosity becomes a bug which has to be fixed with the other marketing feature.

Destroying the skin biome

The human skin has evolved to look after itself rather well without the need for highly viscous, highly foaming, vitamin-containing body washes. The array of microbes on the surface are actively taking care that the skin barrier remains nicely balanced and that harmful microbes don't get a chance to build up their own rich biome community. By regularly attacking our own skin biome, we are making it hard for the biome to look after us.

Recognising the damage caused by aggressive soap formulations, marketing require the formulators to add extra oils to replace those removed from the skin. Even if these oils are "natural" they aren't the complex set our skin provided after millions of years of natural selection.

So now the next demand on formulators is to add pre-biotics, pro-biotics or maybe biotics. There is a nice irony here because one super-strict demand on the soap formulations is that they are essentially microbe free.

Destroying the microbes

A solid soap bar is a hostile environment to microbes. Indeed, as mentioned above, a high pH soap bar can be nicely destructive to the skin surface, enabling it to clean more effectively, with obvious collateral damage to the biome and the risk of harming ("drying out") the skin itself. But at least your soap bar doesn't come with any extra pathogens.

The surfactants, oils, vitamins, pre-biotics etc. added to a liquid soap are a feast for microbes. Even if each bottle of soap was fully sterilized during manufacture, once it is opened, microbes can enter and over weeks or months the soap would be smelly and dangerous as the microbes took over.

So the liquid soaps need preservatives. By definition this means molecules that discourage or kill living organisms. So by definition, this means a risk that they are harmful to some organisms on (our biome) or in humans. So we have the nightmare scenario for marketing – they must assure users that everything in the soap is wonderful, chemical free, nourishing and beautiful, whilst admitting that it contains chemicals that kill or inhibit natural organisms. It is a legal requirement that soaps demonstrate the ability to stop microbes from proliferating.

Given that "natural" does not necessarily mean "good" or "safe", marketing can't hide behind the use of natural preservatives. Many parabens are natural, benzoic acid and benzoates are natural but they are demonised as "chemicals".

The *scientific* advice is to go with a minimum package of approved, effective preservatives. But because in these areas the science is of little importance to marketing and consumers, the formulator has to juggle whatever are the latest fads from the latest internet scare.

Soaps to save the planet

The science of a "save the planet" soap is super easy:

- Use half as much
- Half the time
- Of the simplest-possible formulation

This is a 4x reduction in your carbon footprint, a 4x reduction in cost, with far fewer unnecessary chemicals washed down the drain, with your skin biome less under attack, and with no downside to your general cleanliness. Sweat itself is odourless and contains very little oil – so just a warm shower and gentle rubbing removes it. The smell of sweat comes from the skin biome (especially a disturbed one) converting odourless sweat molecules into odorous ones, such as hexenoic acids which are easily washed away with water.

This is *not* what your management and marketing teams want to hear. But sooner or later, consumers will catch up with greenwash from those selling large amounts of unnecessary chemicals. The laundry detergent industry cut product volumes at least 20x along with the carbon footprint of the washing process thanks to cold water formulations. Interestingly, further reduction of their carbon footprint requires dealing with the relatively high footprint of the unnecessary fragrances that consumers demand. The footprint of natural fragrances happens to be larger than their synthetic equivalents, so detergents with less fragrance and fewer natural fragrance molecules will be better for the planet. These sorts of considerations will eventually reach the soap community.

If you are going to choose to aim for a 10% reduction in carbon footprint, still encouraging consumers to use too much of your product, then think hard. Life cycle analysis of bio-based or bio-sourced surfactants rarely shows them to be significantly greener. For those keen to use biosurfactants, the triple challenge is to make them with desirable properties, with a low carbon footprint and to make them in the million ton quantities required to make a meaningful impact on the soap industry. Anything in the 10, 100 or even 1000 ton range is greenwash in terms of saving the planet.

Some smarter formulators are recognising that changing to bio-based surfactants "because they are green" is not a good way to make profitable products. Instead, they identify some unique property of their favourite molecule and create a novel product to deliver a genuine consumer benefit.

And if you are going to use a, say, C12-C14 chain, is it best for it to be bio-based or synthetic? Technically, branched long carbon chains can be better surfactants because they don't form such compact, liquid crystalline phases. Because nature does not provide these, if you can reduce surfactant use by 10% through using a branched chain, is that better or worse for the environment?

If, as is clearly a strong trend, much of petrochemical feedstock comes from advanced chemical recycling (converting to something like pyrolysis oil then putting that into a standard cracker) maybe it’s better to do this than to debate whether palm oil or coconut oil is more sustainable.

The formulation scientist’s dilemma

The science of soap and washing is so simple that we don’t really need formulation scientists. We could all agree on an effective general-purpose formulation that gives adequate cleaning, doesn’t mess up the skin biome, uses sustainable ingredients and effective/safe preservatives, and which is used sparingly by consumers.

That’s what the science says.

Management and marketing want the opposite – complex formulations with multiple ingredients offering marketing claims that can out-compete other products, but all based on the lowest-cost ingredients because price is a premium. The requirement for lowest-cost means that everyone needs to use the same ingredients, produced at large scale. They want the impossible: to be the same but different.

At some point, reality always wins. Given that the science is clear, something has to change – either industry first, leading the change, or consumer first with industry scrabbling to catch up.

Solvent-Based Paint

Links

Adhesion_Adhesion promoters, Adhesion_Crosslinking, Adhesion_Entanglement, Adhesion_Intermingling, Coating_Dewetting Theory, Coating_Levelling Theory, Coating_Pinholing Theory, Dispersions_ODC, Dispersions_PVC and CPVC, Dissolution_Hansen Solubility Parameters, Evaporation_Basics, Evaporation_Diffusion limited, Evaporation_Marangoni, Evaporation_Temperature and Antoine Coefficients, Flow_G' and G", Flow_Oscillatory rheology, Flow_Polymer Viscosity, Flow_Rotational rheology, Flow_Shear dependent, Flow_Thixotropy, Flow_Yield Stress, Optics_Gloss, Optics_Scattering and Opacity

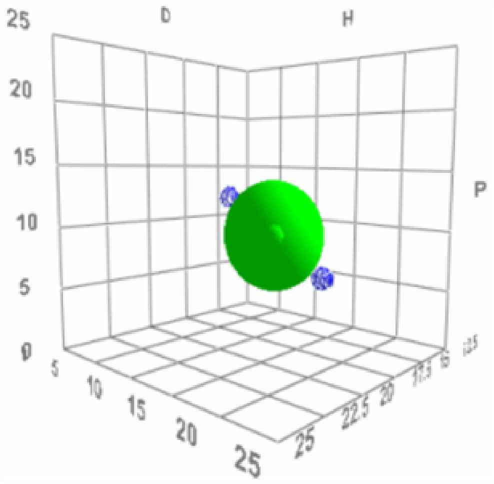
You are the specialists in the alkyds, acrylics, urethanes etc. so we are not going to comment on the chemistry. Instead we focus on the general principles behind all solvent-based paints.

Polymer solubility

Using standard techniques you can readily measure the Dissolution_Hansen Solubility Parameters of your polymer. The three parameters, δD, δP and δH describe the Dispersion, Polar and Hydrogen bonding characteristics of your polymer.

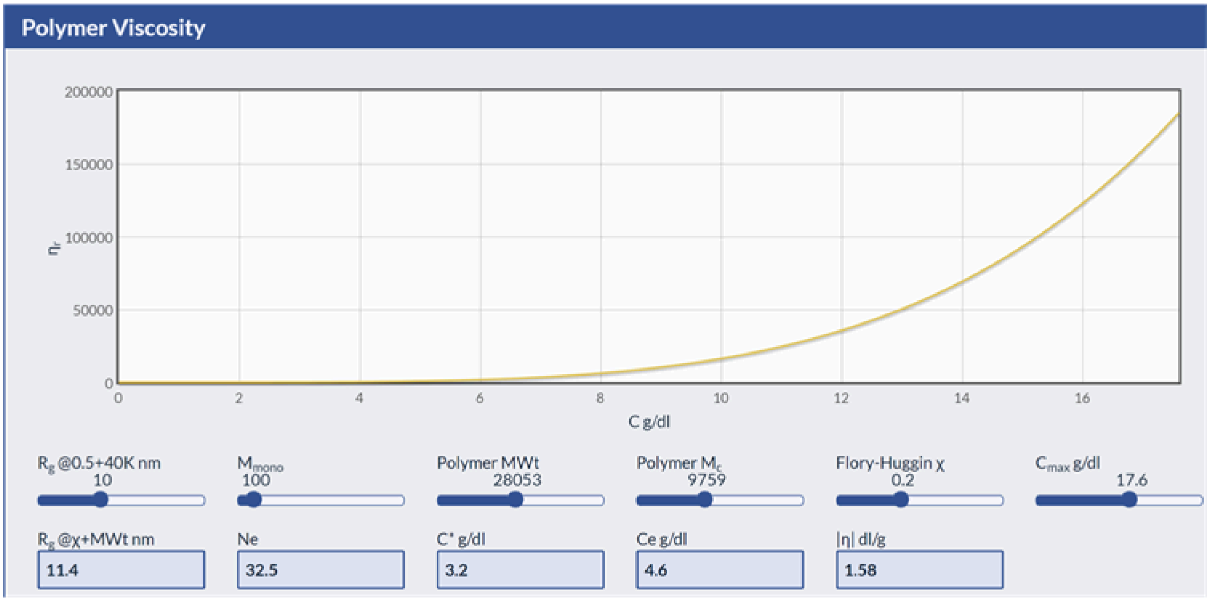
You can then choose a solvent with closely matching HSP values, using the 3D Distance metric to give a number to the quality of the match. The closer the match the “happier” your polymer is in the solvent and this generally makes your formulation life much easier. Your problem is that a solvent that is good in terms of HSP match might be bad in terms of cost, safety, odour, evaporation rate etc. If you are restricted to a single solvent, it can be hard to find an optimum.

Fortunately, the HSP of a blend of solvents is the volume weighted average of the two. This means that it is possible to take two solvents that have desirable properties, yet a poor HSP match and create a blend where the match is perfect. Surprisingly this even extends to creating a good solvent from two bad solvents.



The green dot in the centre of the green sphere is at the HSP of the polymer. Any solvent inside the larger sphere is OK, outside it is bad. The two blue dots show solvents that are each bad. But you can see that a 50:50 blend would be exactly in the centre. The point is *not* that we recommend making good solvents out of bad ones, but to emphasise the potential of creating great solvent blends from individual solvents that, by themselves, are not as good as desired.

The Flow_Polymer Viscosity behaviour can be captured in the app:



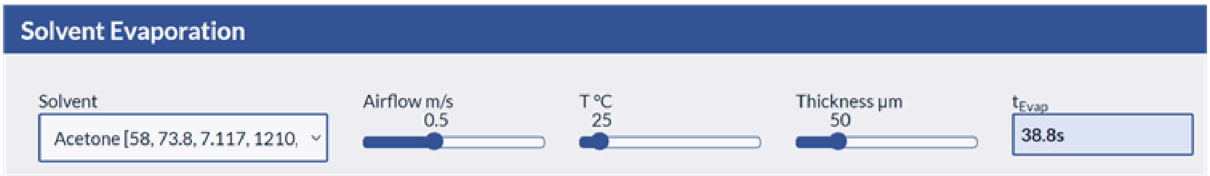
R_g @0.5+40K nm 10
M_{mono} 100
Polymer MWt 28053
Polymer M_c 9759
Flory-Huggin χ 0.2
C_{max} g/dl 17.6
R_g @χ+MWt nm 11.4
Ne 32.5
C* g/dl 3.2
Ce g/dl 4.6
|η| dl/g 1.58

https://www.stevenabbott.co.uk/practical-solubility/polymer-viscosity.php

We want the maximum possible polymer for the minimum solvent, but as the app shows, the viscosity increases rapidly with concentration. As discussed in the Polymer Viscosity chapter, the rules for this depend on M_C the critical entanglement molecular weight. This is a complex phenomenon which means that relatively small changes to your polymer can lead to relatively large changes in viscosity at higher concentrations. This *should* mean that polymer suppliers (including internal suppliers) should be confident about formulating in terms of M_C – choosing polymer MW to match the desired entanglement phenomena. Sadly, although polymer physicists have known for decades that M_C is super-important, manufacturers generally don't know, or choose not to reveal, the M_C values of their polymers. If you are a small-volume customer, there's nothing you can do. If you are from a megacorp, demand M_C values from suppliers – maybe they'll listen to you.

Solvent evaporation

As you find in Evaporation_Basics up to a point, the rate of evaporation of the solvent is governed by the vapour pressure of the solvent and the airflow. You can see this in a simple app:



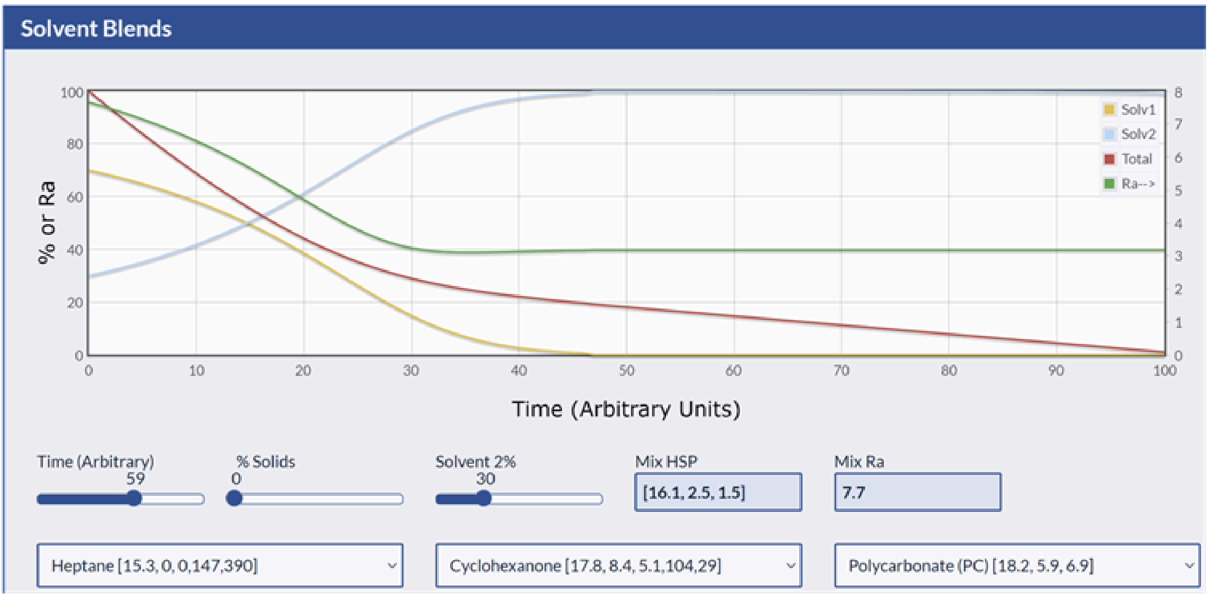
Solvent Acetone [58, 73.8, 7.117, 1210,
Airflow m/s 0.5
T °C 25
Thickness µm 50
t_{Evap} 38.8s

https://www.stevenabbott.co.uk/practical-coatings/evaporate.php

The numbers you see after Acetone are its MW, MVol and the first two of the three Antoine coefficients, Evaporation_Temperature and Antoine Coefficients, provided with each solvent. The rate of evaporation is proportional to the vapour pressure (calculated from the Antoine coefficients and temperature) and to the square root of air velocity – paint dries *very* slowly in a room with no air flow!

Solvent blend evaporation

There is a key scientific advantage to using a solvent blend, and a significant disadvantage. First, the advantage (in addition to the cost, safety, green advantages opened up by solvent blends):



Time (Arbitrary) 59

% Solids 0

Solvent 2% 30

Mix HSP [16.1, 2.5, 1.5]

Mix Ra 7.7

Heptane [15.3, 0, 0,147,390]

Cyclohexanone [17.8, 8.4, 5.1,104,29]

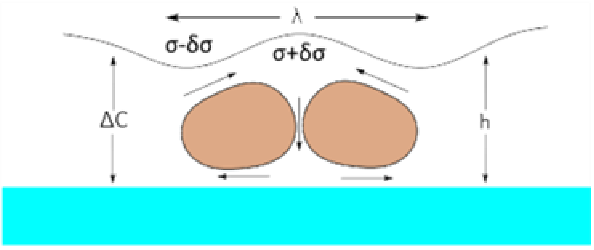
Polycarbonate (PC) [18.2, 5.9, 6.9]

https://www.stevenabbott.co.uk/practical-solubility/solvent-blends.php

Here we have chosen to dissolve polycarbonate using a blend of 70% heptane and 30% cyclohexanone – [not because they are green but because we get a nice graph to tell the story]. At the start, the solvent blend has an HSP Distance (Ra) of 7.7, making the mix a relatively poor solvent, but OK for the start of the process where the concentration is relatively low. Because heptane evaporates faster than cyclohexanone (their Relative Evaporation Rates are 390:29) by the time we are down to 20% of the original total amount of solvent, we have mostly cyclohexanone which has an Ra of just over 3. This means that in the difficult diffusion-limited part of the drying process (discussed soon), the chances are that the polymer is “happier” and more open. If this coating is a gloss paint then the surface will be smooth because imperfections have a longer time to flow out.

If the relative evaporation rates had been reversed then the good solvent would have left early and the polymer would have crashed out early.

But now the problem of these solvent blends. The surface tension, σ, of heptane is ~20 mN/m while for cyclohexanone it is ~25.5 mN/m. If there are any local variations in surface composition then there will be local surface tension gradients, δς, … and these can rapidly result in Evaporation_Marangoni effects.



The local variations drive flows which spontaneously form into patterns from subtle through to full hexagons. A lot of “orange peel” in paints and coatings is due to these Marangoni patterns.

If you have Marangoni problems you can either fix the relative evaporation rates of your solvents, or choose solvents with closely matched surface tensions. Or, if you look at the relevant equations, a thin coating and a large viscosity is enough to stop the effect.

In terms of additives, any component that moves rapidly to the surface and swamps the surface tension effects of the solvents will do the job, because there will be no surface tension gradient to drive surface flows. People say that you need “low surface tension” to level out defects. This is simply untrue; levelling is *driven* by surface tension so the higher the better. Instead molecules that partition to the surface to swamp Marangoni effects have to be low surface energy, otherwise they don’t partition. The end result is the same, but the science is very different. Silicones are especially good at “levelling” (in reality, destroying Marangoni) because they partition so well and have a low, temperature-independent, surface tension that swamps other effects. The downsides of silicones are equally well known, so these “levelling agents” have to be used with caution.

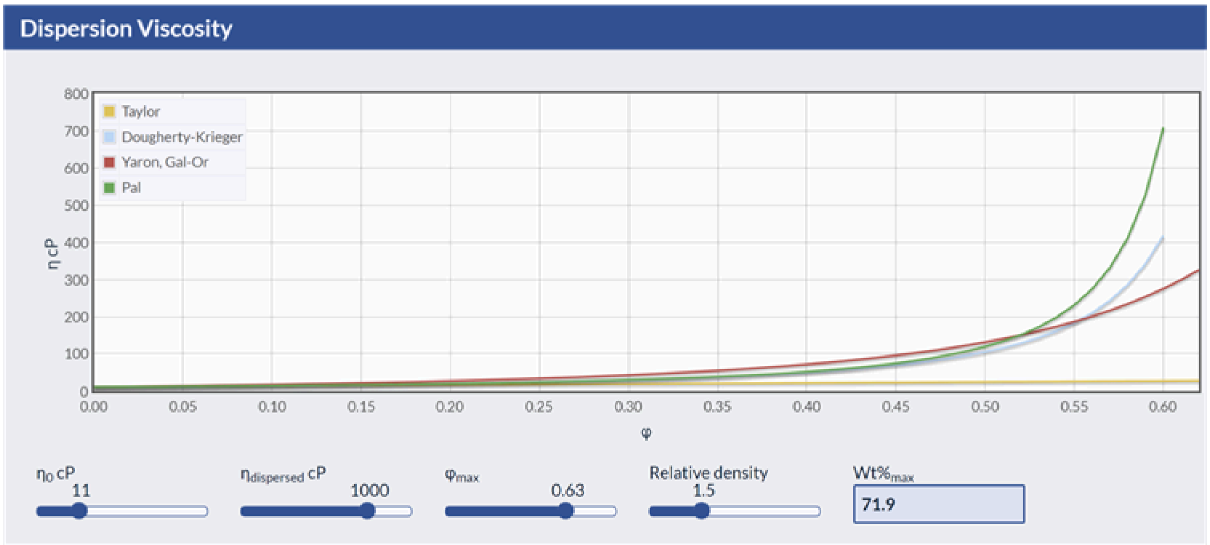
Diffusion-limited drying

At some point, all drying becomes limited not by the evaporation but by the rate at which solvent molecules can reach the surface. As discussed in full, Evaporation_Diffusion limited, there are only three things you can do:

1. Use a solvent that keeps the coating open for as long as possible, as discussed above.
2. Heat the system, because diffusion coefficients increase exponentially with temperature.
3. Delay the curing of your polymer system till the solvent has mostly gone, because the crosslinking will decrease the diffusion coefficient.

Controlling the pigments

Your specific pigments are chosen to give hiding power (Optics_Scattering and Opacity), to provide cheap bulk that also forms the core of a tough paint, and to provide colour. Although we naturally work with weight %, what really matters is volume fraction, φ. If your pigments are well-dispersed then simple viscosity theory shows Dispersions_Rheology (Low shear) that you can add a lot of pigment with only a modest effect on viscosity:



η_0 cP 11

$\eta_{dispersed}$ cP 1000

φ_{max} 0.63

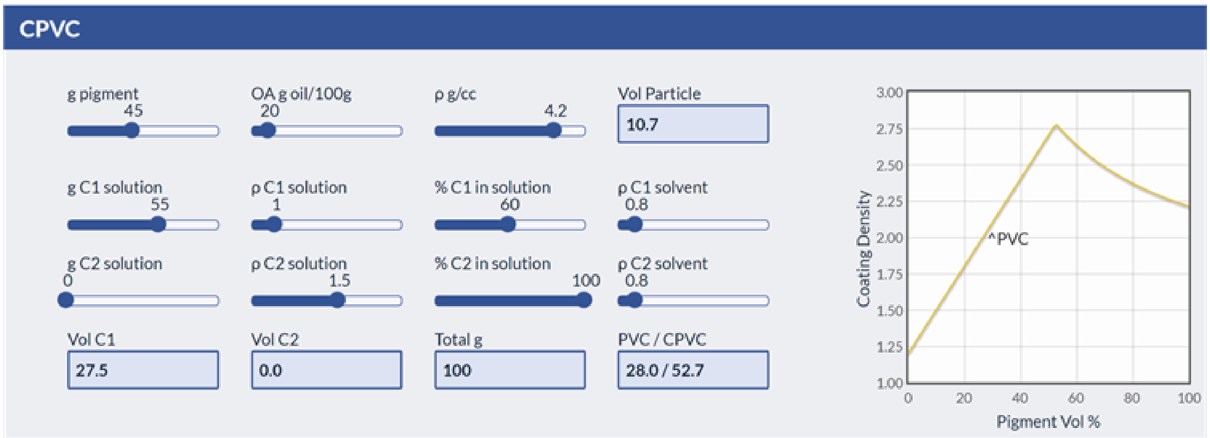
Relative density 1.5

Wt%$_{max}$ 71.9

https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

For spherical particles, up to about 45% nothing much happens. The problems then take off, with the rate of increase depending on the point at which the particles have no ability to move past each other, the close packing limit. Things get more complicated when particles are elliptical and when your dispersant isn’t ideal. The Dispersions_Rheology (High shear) discussion takes you through those issues.

More pragmatically, you need to calculate the Pigment Volume Concentration (PVC) and its critical value, the CPVC, Dispersions_PVC and CPVC:


CPVC

g pigment 45

OA g oil/100g 20

ρ g/cc 4.2

Vol Particle 10.7

g C1 solution 55

ρ C1 solution 1

% C1 in solution 60

ρ C1 solvent 0.8

g C2 solution 0

ρ C2 solution 1.5

% C2 in solution 100

ρ C2 solvent 0.8

Vol C1 27.5

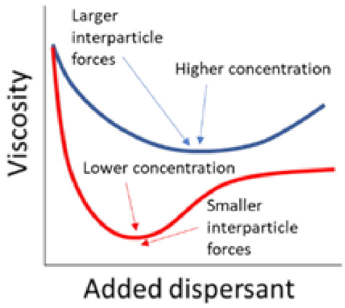
Vol C2 0.0

Total g 100

PVC / CPVC 28.0 / 52.7

https://www.stevenabbott.co.uk/practical-coatings/CPVC.php

As you get close to the CPVC, paint properties such as coating density, gloss or water resistance become susceptible to small changes in the formulation so you have to work harder on your pigment and its dispersant to be able to push the limits.



A good dispersant makes life easier, so finding the Dispersions_ODC Optimal Dispersant Concentration is necessary in both senses: a good dispersant has the lowest optimal concentration and the smallest interparticle forces (lowest viscosity) at that concentration. It is a good strategy to get the dispersant optimized on the pigment before it gets into the paint. Adding the dispersant to the paint is likely to lead to disappointing results because a bad dispersant already on a pigment is surprisingly hard to displace with a better one.

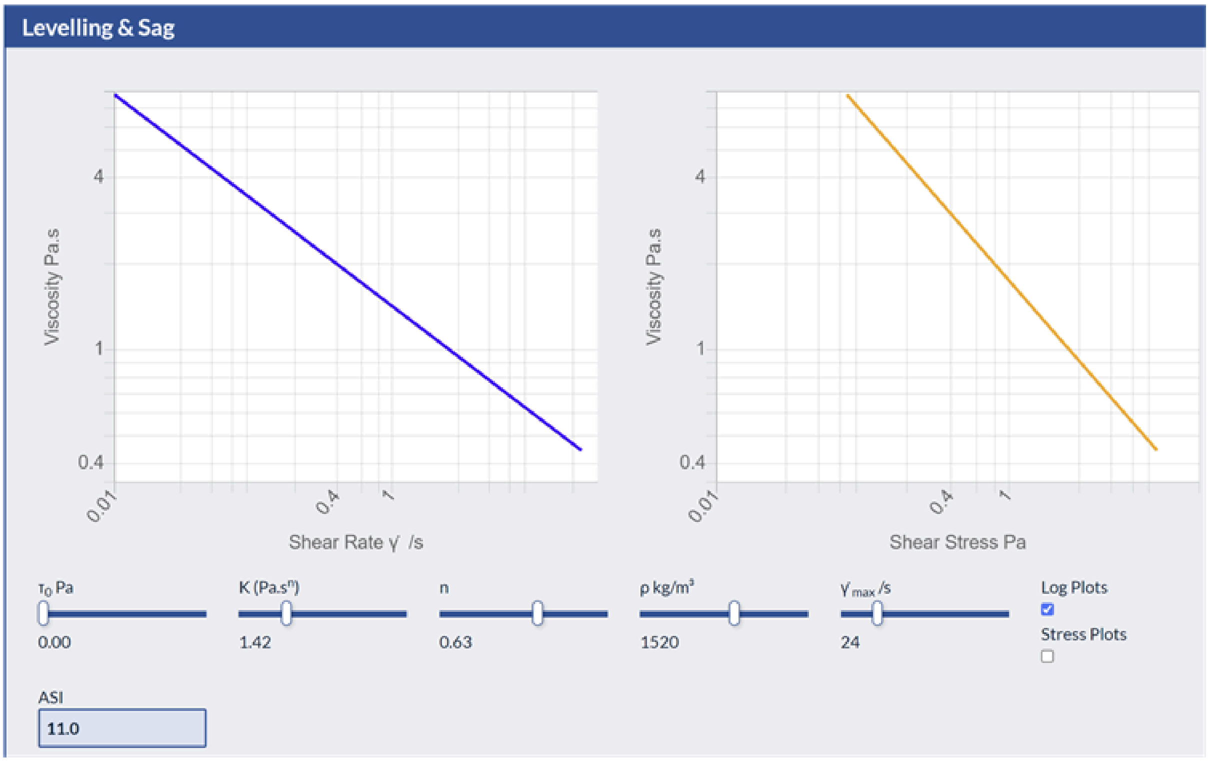
Controlling the rheology

The formulation team need to embrace rheology. There seems to be a lot of it, Flow_G' and G", Flow_Oscillatory rheology, Flow_Polymer Viscosity, Flow_Rotational rheology, Flow_Shear dependent, Flow_Thixotropy, Flow_Yield Stress but it really comes down to flow curves in rotational rheology (looking at viscosity as a function of shear rate – both with increasing and decreasing rates) and, less popularly but more fundamentally, via oscillatory rheology and G' and G".

> The Abbott Guide to Rheology at https://www.stevenabbott.co.uk/practical-rheology/the-book.php is an attempt to demystify rheology. It really is unfortunate that something so essential is often seen as some deep magic knowable only to rheologists. Knowing how to get good data from a rheometer is indeed tricky, which is why we need rheologists. But knowing *why* we want the data and knowing *what to do with it*, is not at all hard.

A good specific example is in controlling levelling and sag. You want easy flow to allow Coating_Levelling Theory yet you want a very high low shear viscosity to resist sag – flowing down a vertical surface.

An app lets you think through these issues:



τ_0 Pa 0.00
K (Pa.s^n) 1.42
n 0.63
ρ kg/m³ 1520
$\dot{\gamma}_{max}$ /s 24
Log Plots ☑
Stress Plots ☐
ASI 11.0

https://www.stevenabbott.co.uk/practical-rheology/Levelling-&-Sag.php

We examine the flow curve over a very small range of shear rates, $\dot{\gamma}$, e.g. up to 5/s. The data can be fitted to a Herschel-Bulkley law:

$$\eta = \frac{\tau_0 + K\dot{\gamma}^n}{\dot{\gamma}}$$

Here τ_0 is the yield stress. As it happens, given day-to-day accuracy of measurements of real paints, there's little advantage trying to extract a yield stress value and you can simply fit the data to a power law:

$$\eta = K\dot{\gamma}^{n-1}$$

Because shear stress is simply viscosity times shear rate, we can plot the same data in terms of shear stress, the right-hand curve. The reason is to see that typical stresses are in the 1-5 Pa range which, if you do a calculation based on a typical paint thickness and density is the sort of driving force inducing sag. From this you can even estimate the ASI (Anti-Sag Index) used in the paint industry; it is the largest thickness (alas, in US mils) of a line which doesn't sag down to the next thinnest line below.

What is interesting about the paper (quoted on the app page) that links the power law fit to the ASI is that they can get just as good a predictive value from measuring G' via oscillatory rheology. But it's not just "the" G'. Instead the paint is sheared for a time then the oscillatory measurement is started. Over time the paint structures itself (discussed next) and achieves a relatively large G'. It is the short-time G' (not surprisingly) that links to the tendency to sag.

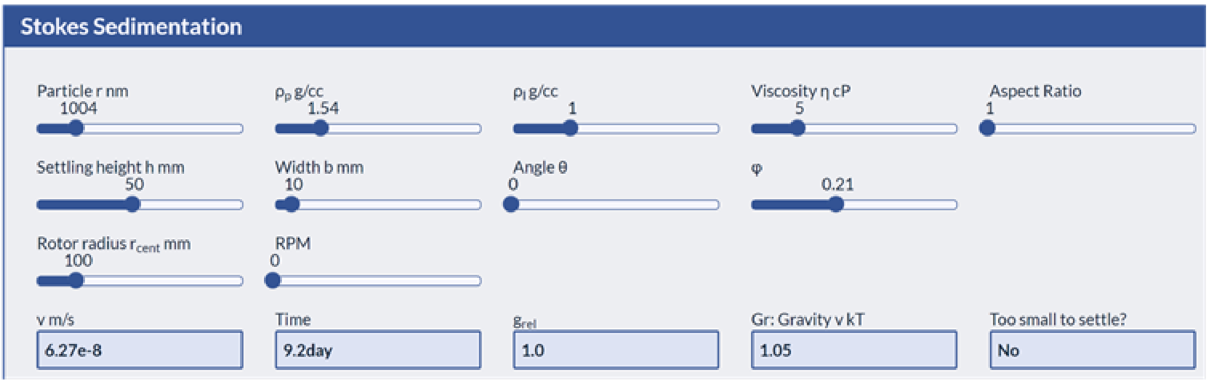
A theme running through FST is that G'/G" measurements are underutilized because of their perceived difficulties. With modern rheometers being so user-friendly, this oscillatory approach to understanding the paint structure seems more insightful than more conventional and difficult low shear rate rotational viscosity measurements.

Paint structuring

The rheological effects of the polymers and well-dispersed particles are a core necessity of a paint. To get the subtle details right, such as the correct power law behaviour for anti-sag with Emulsion Paint is relatively easy via Thickeners_Associative Thickeners. There are fewer fancy options in the solvent-based formulation world, so attention to the subtleties of polymers and pigments is a necessity.

One key trick is to perform rheology analyses not just of your standard paint formulation but also on formulations with less solvent. This lets you see how quickly (usually good) or slowly (usually bad) the paint gains structure as the solvent evaporates. A paint with a higher starting structure but a less-steep rise in structure on drying may well be inferior to one with a lower structure but a steeper rise.

Avoiding particle settling


Stokes Sedimentation

Particle r nm 1004
ρ_p g/cc 1.54
ρ_l g/cc 1
Viscosity η cP 5
Aspect Ratio 1
Settling height h mm 50
Width b mm 10
Angle θ 0
φ 0.21
Rotor radius r_{cent} mm 100
RPM 0

v m/s: 6.27e-8
Time: 9.2day
g_{rel}: 1.0
Gr: Gravity v kT: 1.05
Too small to settle?: No

https://www.stevenabbott.co.uk/practical-solubility/stokes.php

Stokes law tells us that the time for a particle of radius r and density ρ_p to fall a distance h in a medium of viscosity η and density ρ_l (and therefore density different $\Delta\rho = \rho_p - \rho_l$) is given by:

$$t = \frac{h\eta}{2.18 g \Delta\rho r^2}$$

So less dense particles or a higher-density solution, higher viscosities and smaller particles help keep particles in suspension. An extra factor is the volume fraction φ of particles. The higher φ the more the particles self-crowd each other, slowing the speed. The Richardson and Zaki equation tells us that velocity is reduced by $(1-\varphi)^{5.65}$.

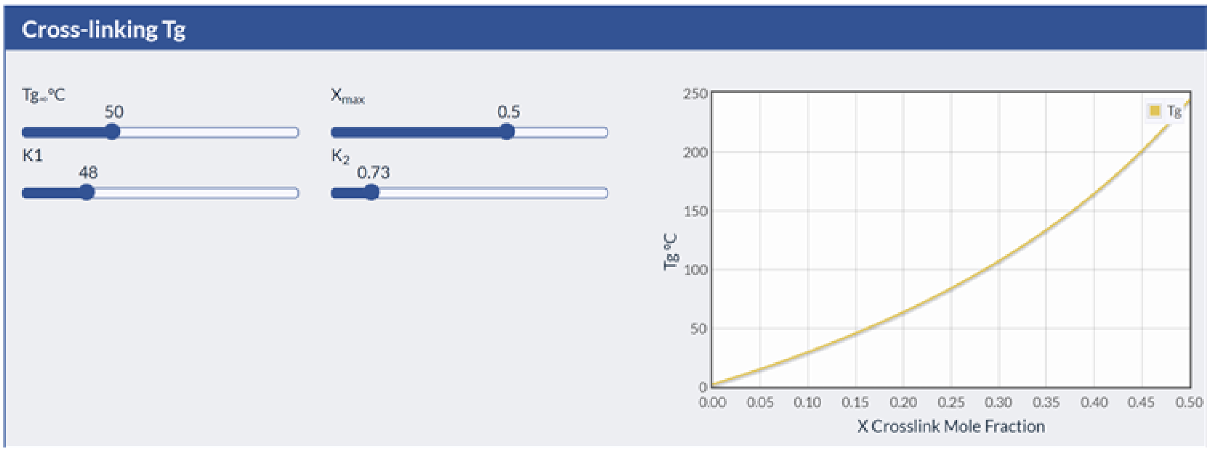
Elliptical particles fall a bit slower, via an equation shown in the app.

Crosslinking and adhesion

The build-up of a crosslinked network is difficult to simulate numerically, but, as described in Adhesion_Crosslinking the chances of converting a large fraction of the reactive groups in low-functionality systems is much higher than the same number of groups in a high-functionality system. The difficulty is that highly crosslinked systems restrict the motion of the remaining reactive groups so they are unable to reach another reactive site. Unreacted crosslinking groups risk the danger of undesirable reactions over time (e.g. with water and oxygen), which in turn lead to deterioration of the paint.

Putting it another way, the Tg of a highly crosslinked system is higher, so main-chain segmental motion is restricted as soon as the system reaches this Tg. As is obvious, raising the temperature allows more reaction … till the Tg rises to this new temperature.

An idealised view of this is found at:



Tg_∞ °C 50

X_{max} 0.5

K1 48

K_2 0.73

https://www.stevenabbott.co.uk/practical-coatings/X-Link-Tg.php

A paint, therefore, is a pragmatic compromise between crosslink density for general strength/resistance and the sort of temperatures or times allowed to reach the desired density.

If the chemistry crosslinking the paint also allows a modest amount of reaction onto functional groups on the surface of the coating (e.g. reacting onto surface -OH groups) then there is the necessary Adhesion_Entanglement across the interface to give a robust bond to the surface.

Sun Screens

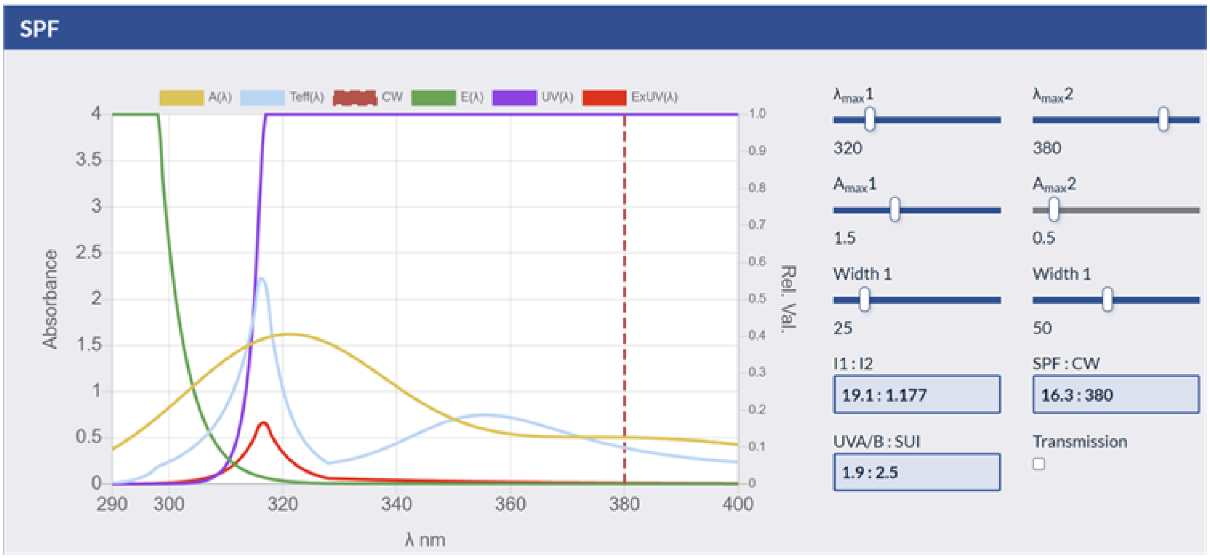
Links

Diffusion_Diffusion into Skin and Hair, Dispersions_Rheology (High shear), Dispersions_Rheology (Low shear), Dissolution_Hansen Solubility Parameters, Dissolution_Ideal solubility, Evaporation_Basics, Evaporation_Humectants, Flow_Basic viscosity, Flow_Yield Stress, Optics_Scattering and Opacity, Surfactancy_Emulsification, Thickeners_Associative Thickeners, Thickeners_Polymeric Thickeners

> A good sunscreen is hard to achieve. It is easy to have lots of UV-blocking chemicals or particles, it is easy to make a water-and-sweat-resistant product, it is easy to make a super-safe product, it is easy to make an aesthetically pleasing product. But those desirable properties provide many conflicts and trade-offs.

SPF

Whatever your formulation, the product you sell has to have a known, validated Sun Protection Factor, SPF. The app makes it easy to explore the options for creating whatever SPF you are aiming for. In principle you can use any number of UV absorbers; the app lets you use 2 as a good-enough basis for exploring.



λ_{max}1: 320

λ_{max}2: 380

A_{max}1: 1.5

A_{max}2: 0.5

Width 1: 25

Width 1: 50

I1 : I2: 19.1 : 1.177

SPF : CW: 16.3 : 380

UVA/B : SUI: 1.9 : 2.5

Transmission ☐

https://www.stevenabbott.co.uk/practical-coatings/SPF.php

The SPF is the ratio of the integrated amount of potentially damaging UV (this turns out to be the red line) divided by the amount that gets through your filter, the blue Teff(λ) line which has been greatly magnified so you can see what it's doing. So the SPF is the ratio of integrals, I1 and I2 which are provided in the output. The CW is the Critical Wavelength which helps describe the range of filtering capability, you also get the ratio of UV-A and UV-B absorption and, finally, the SUI, the Spectral Uniformity Index. The equations and deeper explanations are provided in the app.

Choice of absorbers/scatterers

There are plenty of small molecules with strong absorption in the appropriate wavelengths. Dissolving them in some pleasant cream used to be easy, and you would have an adequate sun block. The number of such molecules that you can use is diminishing rapidly, and the number of cosmetically-approved oils is also diminishing so it's hard to get the required concentrations. But even if you can dissolve them and the molecules themselves are entirely safe sitting on the surface, there are multiple problems:

- When they absorb UV they can react and create problematic chemicals that, at the very least, are irritants.
- By reacting they will (generally) lose their UV absorbance so their efficacy wears off over time in the sunshine.
- In water they can dissolve/disperse and potentially damage the pool or marine environment
- Being small molecules in a skin-friendly cream, they can migrate through the stratum corneum of the skin (see SkinCare) and once through that, there's nothing much to stop them getting into the rest of the skin and into your bloodstream.

The problem of UV reactivity and loss of absorption can only be solved via a clever choice of absorbing molecule or by adding co-molecules that take away the UV energy from the absorbent, before it reacts.

For a fuller understanding of diffusion through the skin, see Diffusion_Diffusion into Skin and Hair. It is a polite fiction in cosmetics that chemicals only go in to the dead stratum corneum and magically stop there, but the reality is very different.

Given that the number of viable small-molecule UV absorbers is diminishing rapidly, and given the problem of marine pollution, there are two logical alternatives. First, make the small molecules rather large, such as Bemotrizinol, which is also water insoluble. Second, make them into a polymeric form. They certainly won't migrate through the skin and their relative lack of mobility decreases the chance of UV-created molecules directly irritating the skin surface. With the right polymer, they won't wash off into the water.

The other approach is to use particles as scatterers and/or absorbers. For optimal scattering, Optics_Scattering and Opacity, you need relatively large particles with a high refractive index difference from your cream. Particles like ZnO (n~2) and TiO2 (n~2.5) do this very well. They are in themselves safe (the TiO2 surface is treated to ensure that it doesn't generate reactive oxygen species) and because they are large (in order to scatter) there are no issues with nanoparticles. Formulating the cream so that it doesn't easily wash off means that these formulations are ideal sun screens in every respect other than aesthetics. People generally don't like going around smothered in white gunk.

The same particles formulated to nanoparticle dimensions no longer scatter but have plenty of UV absorption. Formulating a stable nanodispersion is tricky, but good performance from such strong absorbers, that don't degrade in the sunlight, is easy to achieve. But now we swap scares of chemicals for scares of nanoparticles. The Australian approach to nano-sunscreen safety has a lot to be recommended: "Sure, there are theoretical small risks from nanoparticles, but compared to the high risk of skin cancer, it's a no brainer."

Formulating the base cream

Some formulations are basically the UV blockers themselves dissolved in a volatile solvent such as ethanol.

A typical (old-fashioned) mix would be Homosalate, Octocrylene, Ethylhexyl Salicylate and Avobenzone (or Oxybenzone). Add a bit of polymer, maybe a siloxane, some pleasant fragrance and a few marketing-friendly natural extracts and you have a classic, effective, sunscreen. Because absorption is a combination of extinction coefficient and path-length, the concentrations of active ingredients have to be high because the final layer is thin.

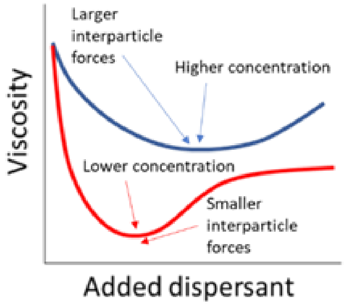
You need to understand the solubility of any solid component in any of the liquid components. For this Dissolution_Hansen Solubility Parameters are especially helpful because you can readily calculate the parameters of the initial solvent blend (a volume-weighted average of the individual components) and can see what happens when volatile components (such as ethanol) leave. What you don't want is sudden crystallisation of a solid component (becoming "gritty" on the skin) because it finds itself in an alien solubility environment.

The alternative is to apply a typical 10µm coating of some pleasant base cream with a lower concentration of the actives or, similarly, a cream-based dispersion of particles. Because there is plenty of warm sunshine to evaporate the water, the cream is likely to be an o/w emulsion. So choose your oil by whatever marketing, cost, safety, sustainability or sensory criteria you like and formulate the emulsion rationally. Although there are many ways to perform Surfactancy_Emulsification, doing it rationally via Surfactancy_HLD theory will let you use minimum surfactant for maximum efficiency and will allow you to explore alternatives to the standard ethoxylates and ionics that arouse passions around health, safety and natural/sustainability issues.

The rheological requirements are the big challenge. First you need to get the Flow_Basic viscosity right. Then a luxurious cream should have a significant Flow_Yield Stress so that a blob on the hand just sits there. But it needs to be highly Flow_Shear dependent so that it flows easily under the shear of the fingers rubbing it onto the skin. If the final coating is thin, it probably doesn't need to show Flow_Thixotropy, but it's probably worth doing a thixotropy scan to see if this is a feature or a bug of your formulation.

You have to decide whether to use Thickeners_Associative Thickeners or Thickeners_Polymeric Thickeners. Polymeric thickeners have relatively modest shear-thinning properties so their properties change less during application and use. The polymers can be chosen for their other properties such as water resistance. The associative thickeners lose their viscosity very easily with shear, and can restore their viscosity and yield stress very quickly, so they are good during the moments of application.

Particle Dispersions



You need to find a great dispersant for your particles, ideal using the Optimal Dispersant Concentration method: Dispersions_ODC

The all-important rheological behaviour will be a combination of simple Dispersions_Rheology (Low shear) behaviour then, beyond a typical volume fraction (usually 0.28) you need the Dispersions_Rheology (High shear) behaviour. Whether you have desirable or undesirable yield stress and flow under shear depends partly on ideal particle rheology (which is what the apps mostly describe) and partly effects (fractal dimensions, N-clusters) owing to poor dispersions. The ODC approach should minimize those and make life easier.

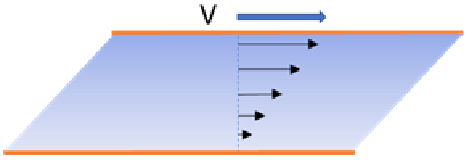
Surface Cleaning

Links

Cleaning_Bleaching and Oxidation, Cleaning_Boundary removal, Cleaning_Contact Angles, Diffusion_Basic Diffusion, Diffusion_Diffusion Coefficients, Dissolution_Dissolution Kinetics, Dissolution_Dissolution Rheology, Evaporation_Basics, Flow_Basic viscosity, Fragrance_Activity Coefficients, Fragrance_Vapour Pressure, Surfactancy_CMC and Langmuir, Surfactancy_Coacervation, Surfactancy_Dynamic Surface Tension

Designing an effective surface cleaner is easy. Designing an “easy wipe” surface cleaner is somewhat harder. Designing an effective, easy wipe surface cleaner is fighting the laws of physics, so is hard.

An effective surface cleaner



To remove dirt from a surface you need to get a flow of the cleaning fluid (that high velocity V in the diagram) to reach the dirt along with sufficient mechanical energy to remove the dirt into the bulk fluid, without damaging the surface. There are two barriers to this, the first one of which is shown in the diagram – zero flow at the surface:

1. The No Slip Boundary Condition (Cleaning_Boundary removal) tells us that the fluid velocity at the surface is zero, so there is no flow of fresh cleaning agent.
2. We need mechanical energy at the sub-µm level, yet cloths and brushes are, of necessity, on the sub-mm to mm level.

Partial answers to the second problem are:

- Microfibre cloths which have the necessary strength of large fibres, along with many sub-fibres that greatly increase the mechanical energy across the surface.
- Particulate systems where the particles are given sufficient energy (inertia and particle-particle contact) to hit the surface and provide cleaning action.

If you aren’t in a hurry and can allow time for cleaning chemicals to diffuse to the surface and for the dissolved dirt to diffuse away then you can use a powerful cleaning formulation such as a laundry detergent (Laundry Liquids) which contains the optimised surfactants plus lipases to attack the oils, proteases to attack protein stains, lipases for any starches and bleaches for hard-to-remove stains. It will contain chelators such as EDTA or a greener alternative to remove stubborn calcium and magnesium salt residues. This chemical combination is sufficiently potent to be effective, while not likely to endanger any typical surface to be cleaned.

There are two big problems with any effective cleaning formulations:

1. Human skin contains many of the elements we wish to remove from a dirty surface. So a good

cleaning formulation requires the user to wear gloves – a big practical barrier for an every-day product.

2. If the formulation dries out, all those chemicals leave a nasty-looking residue. So the cleaning formulation has to be wiped away with clean water. This is more work, and hardness in the water can end up as unsightly residue on the surface.

There is one obvious fact linked to the second problem, yet seldom addressed by users or formulators. Any formulation that's more than, say, 1 µm away from the surface is of no use – it's just excess wasted chemicals that require a lot of effort to remove. An ideal cleaning system would deposit just 1 µm of formulation onto the surface with enough physical motion to enable adequate mixing to avoid long diffusion times. The Easy-wipe formulations described next achieve the thin layer by being delivered as a fine spray. But as this is a dilute formulation with little intention of providing cleaning power, it is just thin and insufficient for removing tough dirts. A spray that can handle the higher viscosity of a potent cleaning solution would enable a formidable type of cleaning system.

I was seriously unimpressed by a new spray cleaning product I was asked to help improve. They showed me their standard cleaning test, which their product passed with no problem. I said that plain water & a microfibre cloth would do the same thing. "OK, prove it!". They set up the test, gave me a water spray and a microfibre cloth. I realised that the cloth would be swamped by the large amount of gunk, so I did a quick wipe with a paper towel, then with the microfibre cloth. They accused me of cheating by using the paper towel, but I'd proved the physics. Happily, work on the project was soon abandoned, freeing up resources to work on better projects.

Easy-wipe formulations

Light dirt is easily removed with (deionized) water (see the Adhesion part of Particles_Basic behaviour), helped by a low level of surfactant to wrap around oily particles to give them a hydrophilic surface, and not so much to reduce surface tension and decrease the standard measured contact angle Cleaning_Contact Angles but to provide a low receding contact angle so the film of water does not bead up.

As has been commonly observed, using a microfibre cloth significantly increases the cleaning efficacy of these formulations, though you don't use easy-wipe cleaners for removing significant dirt.

As we are not requiring the surfactant to do anything very interesting, we can choose the lowest-cost surfactant which also has whatever marketing claims are desired: "natural", "biobased", "sulfate free" etc. However, as discussed in the next section, surfactants that "dry out" are likely to leave an undesirable visible residue.

As these are likely to be spray formulations to give a broad, light coverage, and because the light loading of surfactant isn't going to change the spray characteristics very much, the focus should be on the design of the spray system itself. As the Flow_Atomization chapter shows, there is distressingly little good science to help with this.

My wife and I once used a large mirror to test out a few different easy-wipe sprays. The winner was obvious – the design of their spray nozzle was superior, delivering a fine, even mist that wiped away easily. [Later discussions with an industry expert confirmed that the spray device is at least as important as the formulation.] One of the "good" formulations showed a problem some days later. A visible residue appeared. Puzzled, we re-tested at a different area and a few days later the same thing happened. One explanation is provided below.

Easy-wipe, effective formulations

Although "effective" cannot mean powerful cleaning, it can mean a surprising level of efficacy over the longer term. The trick is to include some sort of coacervate system (Surfactancy_Coacervation) which involves cationic polymers along with anionic surfactants. The coacervate can't help with the cleaning itself (how could it?). Instead it precipitates as a nano-thick layer on the surface. As is well known (see Hair Conditioners) the cationics are relatively hygroscopic so the surface remains hydrated. This has two benefits:

1. As the formulation stays semi-liquid, it remains invisible, solving the previously-mentioned drying-out problem. [In the failed mirror test, I assume that the formulation held on to the water for a day or two, but eventually dried out. Later discussions with an expert confirmed that this is the case]
2. Any oil/dirt reaching the surface now sits on a hydrated layer with low adhesion. So the next treatment with the easy-wipe formulation renders the surface beautifully clean once more.

How do you know that there is a nm layer of a coacervate on the surface? While ellipsometry can tell you there's *something* super-thin on the surface, how do you know what it is? Amazingly, although ATR-FTIR has a natural depth resolution of a few µm, with some standard interfacial tricks, you can get a good spectrum of the coacervate sitting at the interface. Those who are really keen can use AFM-FTIR. Those interested in such methods can read more in Analytical Techniques.

Toothpastes

Links

Cleaning_Boundary removal, Flow_Oscillatory rheology, Flow_Particle Viscosity, Flow_Polymer Viscosity, Flow_Rotational rheology, Flow_Shear dependent, Flow_Yield Stress, Particles_Size distribution, Surfactancy_Foaming

A toothpaste is a necessary compromise – too good at cleaning and it damages the teeth, too poor and the teeth suffer. The key properties of a toothpaste come from the particles – they control the cleaning function as well as the way the toothpaste handles as a squeezable paste. Apart from the proven efficacy of the fluoride, other ingredients play a minor role in dental health, though are vital for consumer perception and marketing claims.

To a large extent, cleaning is about overcoming the problem of Cleaning_Boundary removal. The No-Slip Boundary condition says that any cleaning fluid has zero velocity at the surface … which is the place where you most want the cleaning fluid to be flowing.

A toothbrush overcomes this problem so there is effective cleaning power even without a toothpaste. Unfortunately, the total contact area between brush bristles and the teeth is very small, so cleaning is inefficient.

A paste of particles allows the mechanical energy of the brush to impact a wider area, so cleaning is more effective.

A stiff brush and a paste of small, sharp particles provides outstanding cleaning … but also erodes the enamel of the teeth. Given that enamel does not replace itself, this is catastrophic. A soft brush and a paste of soft, round particles leaves the enamel undamaged, but is also poor at removing residue on the teeth.

So cleaning teeth is about a medium hard brush (the overwhelming majority of brushes sold) and a paste with a balance of particle sizes and hardnesses.

Optimal particle size distribution

Typically a toothpaste mixes together different silicas, carbonates/bicarbonates and calcium phosphates on the basis of fashion, cost, marketing claims and technical performance. Focussing only on the latter, what is the optimal Particles_Size distribution? As the link explains, knowing what your distribution *is* and extracting numbers to capture it is a hard combination to achieve. For our purposes, smaller is generally better as you get more surface area to do the cleaning, and oversized particles feel “gritty” and unpleasant when brushing. Because of fears of nanoparticles, life is easier if the distribution has nothing below 100nm.

In terms of formulating, you want the maximum amount of particle to pad out your toothpaste. If all particles were the same size then you could not get much beyond 60% volume fraction because of the close packing limit. As discussed in Flow_Particle Viscosity, one way around this is to have a multi-modal size distribution where smaller particles fill the voids between larger one. We also see that in terms of having a tractable viscosity, spherical particles are better than elliptical or rod-shaped ones.

Particle dispersion

Any supplier of particles to the toothpaste industry knows that the particles should be readily dispersed in water. Any of us could buy such particles, mix them with a little water into a thick paste and have a viable toothpaste. The rising popularity of toothpaste “tablets” made from particles compacted together show that cleaning teeth is not a super-sophisticated activity

The extra ingredients such as a bit of surfactant to remove oils and create foam, a humectant (Evaporation_Humectants) such as glycerol to stop the toothpaste drying out underneath the cap, a bit of rheology modifier polymer, the fluoride and extras such as nitrates and the minty flavour we generally seem to like, make little difference to creating an adequate paste.

So where is the challenge?

A blob of toothpaste sitting on a brush should just stay there. So it needs a high Flow_Yield Stress. Yet when we squeeze the tube it should flow out easily, so it needs Flow_Shear dependent behaviour which, happily, comes naturally from particle-filled systems. But not *too* easily. There are strict tests that hold a tube of toothpaste with opening down, timing how long it takes for the contents to flow out – too fast or too slow and the batch is rejected.

The challenge is that relatively small changes to particle sizes, shapes, dispersion, surfactant, humectant, concentration can make subtle changes to any one of the critical rheological properties.

How, then, does the formulator formulate? A common theme in FST is the need for “smart mapping”. You cannot know precisely how a change in particle size. or change of surfactant, or changes to your production process, or change to rheology modifier will affect the several competing rheological needs. But by regularly measuring the core rheological properties, especially via Flow_Oscillatory rheology, you gain a map of how the competing requirements change. With a reasonable map, understanding how to fix one problem (e.g. a change in yield strength) without causing another (e.g. difficulty to squeeze out of the tube) becomes possible.

It was once famously said that the only significant ingredient in toothpaste is the fluoride. So how can manufacturers claim that their special formulation is objectively better thanks to another key ingredient? The trick is to do the measurements under idealised test conditions which might show a 1% improvement in some desirable tooth property. Although this 1% improvement would not be detectable by any customer and, in any case, toothpaste isn’t used in the same idealised way, the claim is valid, and consumers are persuaded to buy the “improved formulation”.

Foaming

Consumers instinctively feel that foaming in cleaners is a “good thing”. But given that teeth cleaning is about delivering adequate cleaning force to the tooth surface via brush and particles, foaming plays no part. If a surfactant is required to help remove oily residues from the teeth, then, again, foaming does not help – it just ties up surfactant molecules uselessly around air bubbles. If, as seems likely, foaming is there as a signal to the user that they’ve used an adequate amount of toothpaste, and, additionally, as a feel-good marketing symbol, the question is how to get the right amount of Surfactancy_Foaming. When you delve into that link you will find that there are far too many myths about foamability and that standard tests are near useless for your specific problem. By all means formulate for foam, but do it rationally.

Why are there so many different toothpastes?

Marketing. Given that most of the cleaning comes from the brush and the particles, other than extreme whitening (lots of sharp particles, so lots of enamel erosion) there’s little practical difference across the confusing array of “different” toothpastes with remarkably similar ingredients on the box.

If any of them was especially better, we’d all be buying it. And we’re not.

How do you make a rational choice? The ones at eye level on the supermarket shelf are the unnecessarily expensive ones. Just look on the bottom shelf for any “normal” fluoride toothpaste at a more sensible price and it will work fine.

Water-based Adhesive

Links

Adhesion_Adhesion promoters, Adhesion_Butt, Adhesion_Crack resistance, Adhesion_Crosslinking, Adhesion_Entanglement, Adhesion_Intermingling, Adhesion_JKR, Adhesion_Peel, Adhesion_Shear, Adhesion_Surface energy Interactions, Adhesion_Testing, Diffusion_Basic Diffusion, Diffusion_Concentration-Dependent Diffusion, Diffusion_Diffusion Coefficients, Dispersions_DLVO, Dispersions_Rheology (High shear), Dispersions_Rheology (Low shear), Dispersions_Settling, Dispersions_Zeta potential, Dissolution_Hansen Solubility Parameters, Evaporation_Basics, Evaporation_Diffusion limited, Flow_G' and G", Flow_Oscillatory rheology, Flow_Polymer Viscosity, Flow_Rotational rheology, Flow_Shear dependent, Flow_Stefans Squeeze, Flow_TTS-WLF, Flow_Thixotropy, Flow_Yield Stress, Humidity_Water mechanical isotherm, Humidity_Water vapor isotherm, Particles_Sintering, Surfactancy_Dynamic Surface Tension, Surfactancy_Emulsification, Surfactancy_Emulsion Coalescence, Surfactancy_Emulsion Creaming and Flocculation, Surfactancy_Emulsion Ostwald, Surfactancy_HLD

Here we ignore all the Adhesive aspects of an adhesive and focus on those special features that arise because the adhesive is water-based.

Soluble adhesives

There is little to discuss – the physics are the same as standard solvent-based adhesives. Their obvious downside, susceptibility to water from the atmosphere or from direct contact make them of little interest to the FST community, except for those working with young children where the ability to clean up with water after some adhesive use is a real advantage.

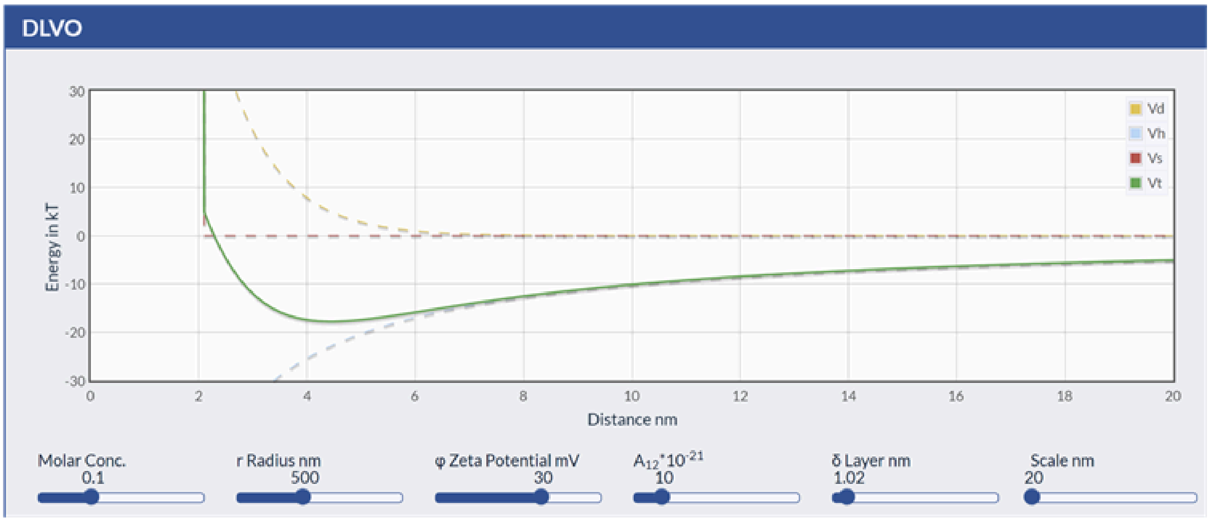
From dispersed to insoluble

For simplicity we will call the dispersed blobs of adhesive polymer emulsion drops, even those that aren’t strictly emulsions. By doing this we can use much of the world of emulsion science without bothering too much about the fine details.

The problem at the heart of these emulsion systems is that the individual drops much be infinitely stable when dilute, yet should form a continuous film of polymer once the water has (largely) disappeared. Each property on its own is not a problem – getting both right is the challenge.

A stable dispersion

For these aqueous systems, Dispersions_DLVO theory is adequate.



DLVO

Molar Conc. 0.1

r Radius nm 500

φ Zeta Potential mV 30

A_{12}*10^{-21} 10

δ Layer nm 1.02

Scale nm 20

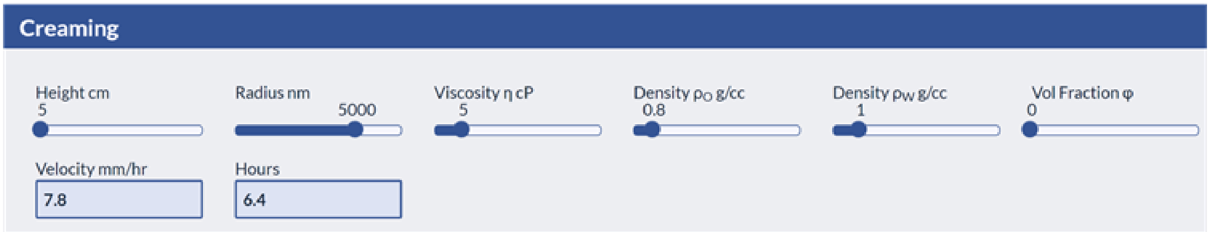
https://www.stevenabbott.co.uk/practical-surfactants/dlvo.php

The particles can be stabilised via:

- Charges, typically via an anionic surfactant. If the zeta potential (https://www.stevenabbott.co.uk/practical-solubility/zeta.php) is somewhere over 30 mV then the particle has a good chance of being stable.
- Steric effects, a long hydrophilic (typically ethoxylate) chain sticking out.

Arguably, charge-charge repulsion is the stronger stabilization method because steric repulsion can be changed into attraction (bridging or depletion flocculation) via extraneous polymers. The downside is that those water-loving ionic groups are not desirable in the final adhesive.

The dispersion also has to be stable against settling or creaming:



Creaming

Height cm 5

Radius nm 5000

Viscosity η cP 5

Density ρ_O g/cc 0.8

Density ρ_W g/cc 1

Vol Fraction φ 0

Velocity mm/hr 7.8

Hours 6.4

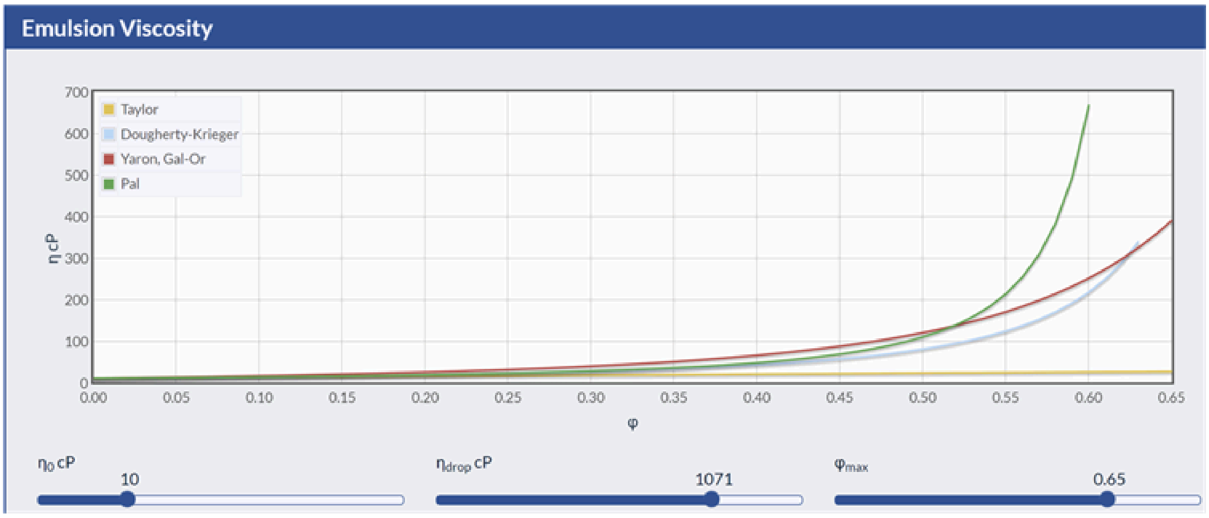
https://www.stevenabbott.co.uk/practical-surfactants/creaming.php

Stokes law tells us that the time for an emulsion drop of radius r and density ρ_O to rise a distance h in water of viscosity η and density ρ_W (and therefore density different $\Delta\rho = \rho_W - \rho_O$) is given by:

$$t = \frac{h\eta}{2.18 g \Delta\rho r^2}$$

In the screen shot, a classic oil emulsion is assumed, hence the low density. Tuning an adhesive emulsion drop to have a density of 1 would eliminate issues of creaming or settling. The volume fraction input, φ, allows the use of the Richardson-Zaki formula where the velocity is reduced by $(1-\varphi)^{5.65}$. A higher volume fraction also increases the viscosity, as discussed next.

Viscosity and rheology



https://www.stevenabbott.co.uk/practical-surfactants/emuvisc.php

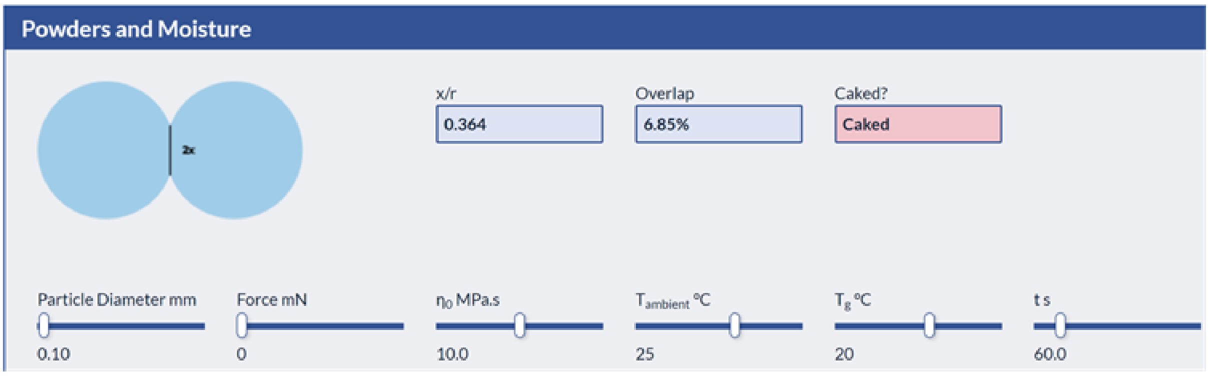
Up to a volume fraction, φ, of emulsion over 60%, the viscosity increases, depending on the model and on the internal viscosity of the droplet, by about 40x. In this region the close packing limit for rigid spheres is being reached and the precise increase to higher φ values is too complex to model. In any case, these models are for smooth spheres with no attractive forces between them. As you try to go to a higher volume fraction to make the adhesive viable, we lose any good theory and have to rely on rheology to compare/contrast formulations.

For the most information with the least effort, checking out Flow_G' and G" is effective. The balance of elastic (G') and lossy (G") shear offers insights not so much into any specific formulation, but into trends occurring with changes of concentration or ingredients.

Straightforward flow curves Flow_Shear dependent and maybe up/down curves for thixotropy Flow_Thixotropy will provide information relevant to handling the adhesive during production, filling and application. If the low shear performance (e.g. avoidance of sagging if applied to a vertical surface) is important then the Levelling & Sagging app at https://www.stevenabbott.co.uk/practical-rheology/Levelling-&-Sag.php is useful.

Film formation

It is helpful to introduce a neutral, and slightly unusual, word into the discussions: sintering. There are many terms used to describe the joining of blobs that are packed close together, but they lack a coherent theory to capture the key parameters that influence the phenomenon. The Rumpf equation for sintering describes the key elements behind the joining of particles, and allows us to identify how changes in our system can affect film formation for better or worse.



Powders and Moisture

x/r: 0.364
Overlap: 6.85%
Caked?: Caked

Particle Diameter mm: 0.10
Force mN: 0
η_0 MPa.s: 10.0
$T_{ambient}$ °C: 25
T_g °C: 20
t s: 60.0

https://www.stevenabbott.co.uk/practical-mechanical/Powders-W.php

The Rumpf equation tells you how overlap x (in the diagram you see 2x) changes over time t for particles with radius r and “viscosity” η and surface energy γ. The equation includes any applied external force, F, which we can assume to be 0.

$$\left(\frac{x}{r}\right)^2 = \frac{t}{\eta}\left(0.8\frac{\gamma}{r} + 0.4\frac{F}{\pi r^2}\right)$$

The use of x/r in the formula is deliberate because we’re generally interested in the fractional sintering – a value of x on its own doesn’t convey much.

There are a few interesting aspects to this equation:

- Contrary to instincts, a low surface energy, γ, *reduces* the rate of sintering. This isn’t surprising because the process is driven by surface energy. The surfactants necessary to disperse the blobs now become unhelpful. However, the surfactant effect on surface energy is likely to be no more than a factor of 2, so it’s not worth worrying about.
- Smaller particles fuse faster. Unfortunately they are also harder to keep separate in the original formulation, so there are trade-offs
- We need to attend to the viscosity of our particles. In addition to this being a difficult thing to measure, the variation of viscosity with water or concentration of “coalescing solvent” is an added complexity.

The apps associated with the app above are all related to food science. But physics is universal and the idea that the Tg changes dramatically with moisture content which in turn is governed by the water-vapour isotherm means that there is a logical chain for investigation. Your adhesive is designed to dry out, but what is the water-vapour isotherm – how dry is dry? And what is the effect of small amounts of water on the Tg? If film formation is complete long before the water has evaporated then these questions are irrelevant. The important thing is to know if/when film formation happens.

If the particle Tg is not affected by the water (an advantage for long term performance) then either the Tg has to be low enough for the film forming temperature to be close to room temperature. Or you have to use a coalescing solvent. The solvent must have a lower volatility than water so its concentration increases sufficiently to lower the Tg and decrease η. It must also have a good Dissolution_Hansen Solubility Parameters compatibility with the polymer.

Thickeners

In theory, the rheology of an aqueous adhesive can be controlled by additional Thickeners_Associative

Thickeners and Thickeners_Polymeric Thickeners. There is a trade-off between the tunability of these systems and their interference with the intrinsic properties of the adhesive.

Where does the water go?

If you do some calculations using the Concentration Dependent Diffusion app at https://www.stevenabbott.co.uk/practical-solubility/diff-cdepend.php and make generous assumptions that the diffusion coefficient of the wet adhesive is 10^{-7} cm²/s and that of the dried adhesive is 10^{-9} cm²/s then in 1000 min you can remove 90% of the water from a blob of adhesive 200µm wide used to stick together two impermeable adherends. If we wanted to stick something that was 2cm square that's a 100x increase in length so a 10^4 increase in time, taking us to 10^7 min, nearly 20 years.

So we have a choice:

1. Pre-dry the adhesive on each surface and hope they act as a *cold-seal* adhesive, like a PSA but somewhat tougher because the surfaces don't need to (and probably don't) stick to anything other than the other surface. Clearly this isn't super-strong, but because Adhesion is a Property of the System, if you think of the system clearly, this might be strong enough.
2. Only stick onto permeable surfaces such as wood, leather or paper. Although we say that in such joints we need to allow the glue to "dry", really the water is diffusing/absorbing through/into the material hopefully finding ways out into the atmosphere eventually.

- The Science -

You can explore the FST via the science categories. We have chosen to classify them one way. You might have chosen a different classification, but because you create your own path through the book, the exact choice of science topics is not so critical. What is important is that you actively explore to find the scientific ideas and resources that help you formulate more effectively.

One of the certainties in formulation life is that you can find help and inspiration in unexpected places. Like all explorers, sometimes it’s necessary to be outside your comfort zone.

Absorbency Capillarity

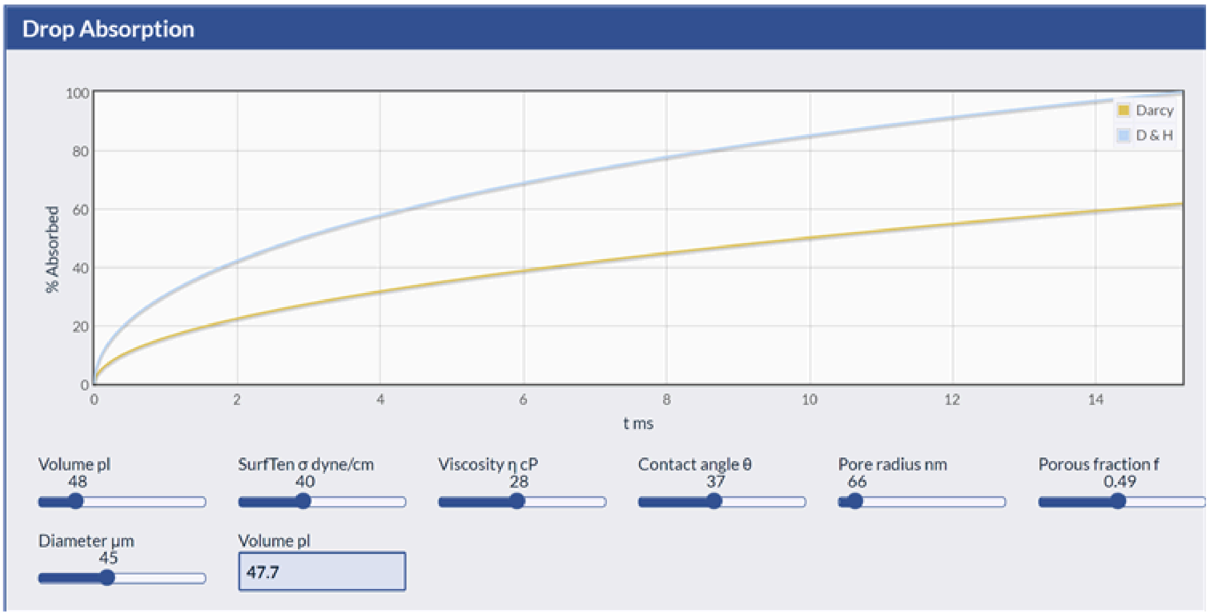
Links

The distance travelled by a liquid into the pores of a coating depends on:

- The radius of the pores, r: larger is faster
- The surface tension of the liquid, σ: larger is faster
- The contact angle of the liquid with the solid, θ: smaller is faster
- The viscosity of the liquid, η: smaller is faster

The key equation for calculating this is the Washburn equation

The current app most relevant is:



Drop Absorption

Volume pl 48

SurfTen σ dyne/cm 40

Viscosity η cP 28

Contact angle θ 37

Pore radius nm 66

Porous fraction f 0.49

Diameter μm 45

Volume pl 47.7

https://www.stevenabbott.co.uk/practical-coatings/drop-absorption.php

This provides the Washburn equation which is the distance d travelled in time t of a fluid into pores of radius r for a fluid with surface tension σ, contact angle θ and viscosity η:

$$d = \sqrt{\frac{r\sigma cos\theta t}{2\eta}}$$

Instead of d, the app describes the volume absorbed if a drop is applied.

Absorbency Porosity and Darcy

Links

Inkjet Inks, Microencapsulation

If we have a powder or fibrous material (non-wovens, paper) then it can absorb liquid via capillary forces. The standard method for calculating the flow is Darcy’s law. The standard law is not so useful, so here we calculate the progress of the liquid into the medium with a more practical version.

Basic Darcy’s Law

We have a porous medium and apply a pressure ΔP (maybe a real pressure, or maybe a capillary pressure) across a distance L to help a liquid of viscosity η flow into the pores. The flow per unit area, q, is given by:

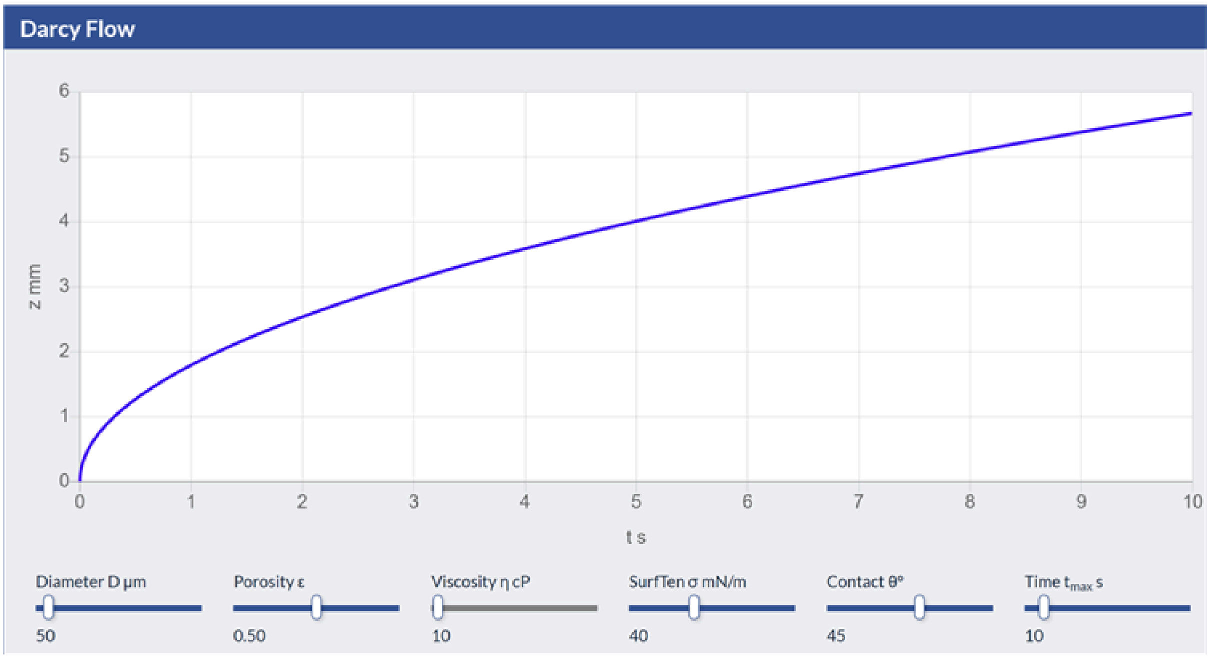
$$q = \frac{k\Delta P}{\mu L}$$

The parameter, k, is the permeability. If you have a flow meter and can apply different pressures across some known lengths then you can determine k.

But the chances are that you aren’t interested in that sort of setup.

Absorbent Darcy flow

We have a porous medium and want to see how the liquid flows into it over time, drawn in by capillary forces. The app shows an example:



https://www.stevenabbott.co.uk/practical-rheology/Darcy-Flow.php

The particles or fibres making up the medium have a diameter D. The porosity (fraction of open space) is ε, the liquid viscosity is η and surface tension is σ, and there is a contact angle of θ. Over your chosen t_{max} you can see how the flow advances.

To calculate the flow we create two porosity-dependent parameters S and T

$$S = -4(1-\varepsilon) - (1-\varepsilon)^2 - 3 - 2\ln\left(1-\varepsilon\right)$$

$$T = \ln\left(1-\varepsilon\right) + \frac{1-(1-\varepsilon)^2}{1+(1-\varepsilon)^2}$$

The permeability k is calculated as:

$$k = -\frac{D^2}{16(1-\varepsilon)}\frac{ST}{(S+T)}$$

The rate of flow in the z direction into the medium $\frac{\delta z}{\delta t}$ is calculated via Darcy's law and k, but it's convenient to eliminate k and just use the input parameters:

$$\frac{\delta z}{\delta t} = -\frac{D\sigma\cos\left(\theta\right)}{8\eta z}\frac{ST}{(S+T)}$$

Because of the $\frac{1}{z}$ dependence, the flow slows down over time (the infinite flow when z=0 is accommodated by a fudge in the code). Other than that, the general trend is obvious: faster flows with larger particles/fibres, larger surface tension, smaller contact angle and smaller viscosity. The porosity effect is a bit complicated – larger porosities allow faster flow, but the flow has more porosity to fill so that gives a slower advancing front.

Washburn Tubes

The fact that the rate of flow in a powder depends on the contact angle gets people excited about the possibility of measuring powder surface energy by monitoring the take-up rate of different solvents into the powder packed within a tube. The Darcy law flow under pure capillarity can be seen as classic Washburn capillarity, Absorbency_Capillarity, so the tubes used in these surface energy measurements are called Washburn tubes. Via a chain of logic, it can be argued that the porosity effects can cancel out between the different solvents and with some arithmetic the surface energy can be calculated.

It is important, therefore, to pass on the advice from a distinguished expert in powder surface energies: "The best thing to do with Washburn tubes is to find an empty drawer in a remote corner of the lab, place the tubes at the back of the drawer, lock the drawer and throw away the key". He deeply regrets the time spent trying to get meaningful results from a technique full of artefacts.

Absorbency Swelling

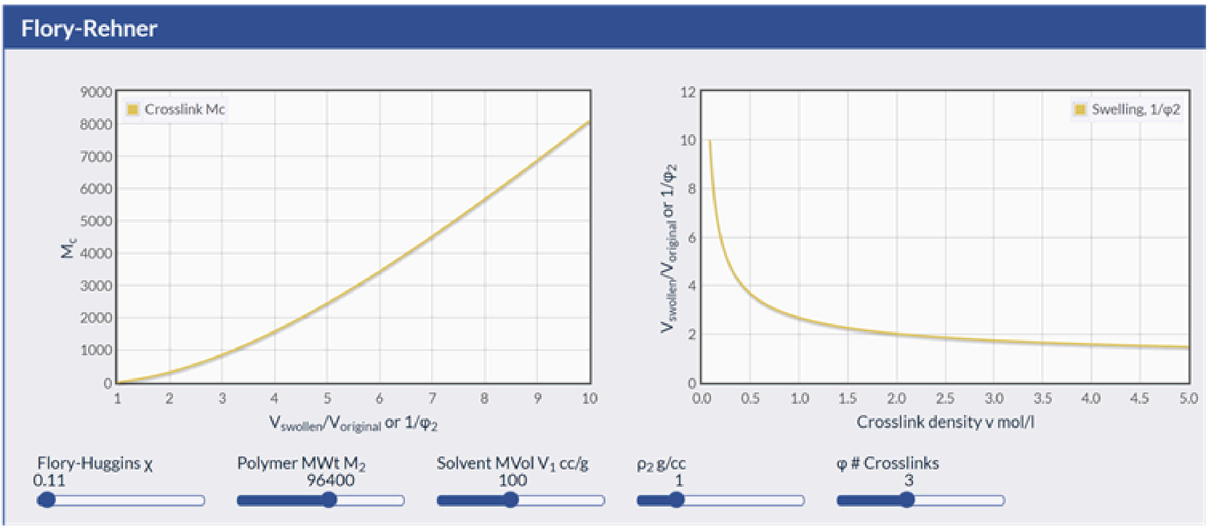
Links

With absorbency of a liquid into a swelling medium we usually want a high speed of absorption and a large amount absorbed. This requires good solubility compatibility between liquid and sorbent.

But we also need to handle the medium (e.g. for disposal) once all the liquid is absorbed. This involves a key trade-off between a high number of crosslinks to provide mechanical strength with a low number to provide sorption capacity.

Flory-Rehner is the standard model for thinking through these issues.

Although apps are supposed to make things easy, this balance of solubility, swelling and crosslinks is tricky to grasp. Let's see the app then try to make sense of what's going on.



Flory-Huggins χ 0.11

Polymer MWt M_2 96400

Solvent MVol V_1 cc/g 100

ρ_2 g/cc 1

φ # Crosslinks 3

https://www.stevenabbott.co.uk/practical-solubility/polymer-swelling.php

The X-axis of the left-hand graph is the swelling ratio, $V_{Swollen}/V_{Original}$. The Y-axis shows M_c, the molecular weight of the average chain between crosslinks. For a high swelling ratio, obviously you must have a large M_c value.

In the right-hand graph we have the swelling ratio on the Y-axis and see the crosslink density, v, that will allow a given amount of swelling.

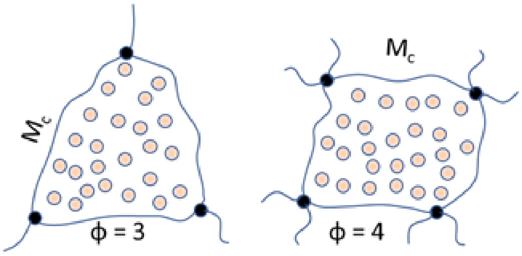
Why two plots? Because the system is complicated and we often come to the problem with different sets of data. We might know the crosslink density from the amount of crosslinker in the system. Or we might have some swelling data and want to determine the crosslink density or M_c. By using the mouse on either of the graphs you can generally find what you need.

The swelling behaviour depends on the MW of the unlinked polymer chain and the Flory-Huggins χ parameter with the solvent and the solvent's molar volume, as described in Dissolution_Flory-Huggins. The χ parameter controls the amount of swelling of the polymer itself – a smaller χ giving more swelling.

The swelling behaviour is analysed by the Flory-Rehner equation:

$$-\left[ln\left[1-\varphi_2\right]+\varphi_2+\chi{\varphi_2}^2\right]=\rho_2\frac{V_1}{M_c}\left(1-2\frac{M_c}{M_2}\right)\left({\varphi_2}^{0.333}-\frac{\varphi_2}{2}\right)$$

This isn't an intuitive equation, but it is what it is. V_1 is the molar volume of the solvent, φ_2 is the volume fraction of polymer in the swollen gel or, to put it another way, $\frac{V_{swollen}}{V_{original}}=\frac{1}{\varphi_2}$.



To relate to the crosslink density, v, we have the formula:

$$v=\frac{2\rho_2}{\varphi M_c}$$

Here φ is the average # crosslinks per crosslink chain. My interpretation of this is shown in the diagram.

How do we use this?

For some communities, such as the rubber industry, Flory-Rehner is built in to their ASTM standards for measuring crosslink density. For some academics it's a nice way to formalise their careful experiments. Can you use it in your own formulation work? Well, let's ask a different question – is there a better way to capture the tradeoffs between solubility (χ), MW, crosslink density and solvent type/size? It's a genuinely difficult topic and maybe Flory-Rehner is better than the non-alternatives.

Adhesion Adhesion promoters

Links

Lipstick, Mascara, Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint

Adhesion promoters exists to provide Adhesion_Entanglement across the interface. Some help to create physical entanglements between polymers. Others provide chemical entanglements. We don't much care which it is; instead we care that it's not too little (of course) and not too much.

Promoting Physical Entanglement

Solvent-based adhesives create physical entanglement by mutual solubility across the interface, encouraged by mutual Dissolution_Hansen Solubility Parameters compatibility of polymers and solvents.

Adhesion to polymers such as polyethylene (PE) or polypropylene (PP) or polyethylene terephthalate (PET) is difficult because they are each sufficiently crystalline that solvents at room temperature don't open up the polymer, so entanglement can't take place.

For PE and PP, heating above their melting point allows excellent entanglement, that's heat sealing (see https://www.stevenabbott.co.uk/practical-mechanical/Heat-Seal.php). The melting point of PET is impractically high, so heat sealing has to take place via co-extruded amorphous polyesters on the surface.

There *are* solvents that can open up PE and PP – for years chlorinated alkanes were used as adhesion promoters. But these are now unpopular. Some phenols are good at opening up PET but again are not practical for most uses.

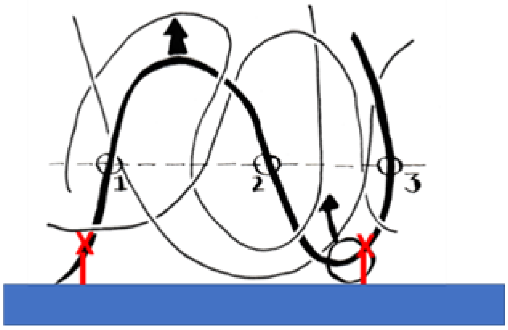
A very nice way to create an amorphous PET surface is to hit it with an excimer laser or a xenon flash (see https://www.stevenabbott.co.uk/practical-coatings/flash-cure.php). This raises the temperature of the top few nm by 100s of degrees, rendering it amorphous. It then cools rapidly and remains in an amorphous state. Solvent-based adhesion and heat sealing are both straightforward on such surfaces.

Flame, corona or plasma treatment of PE, PP and PET does three things:

1. It opens up the crystalline structure, allowing physical entanglement.
2. It adds some functionality that might allow chemical entanglement
3. It increases the surface energy. But this is irrelevant. A 10-20% increase in surface energy cannot be the cause of a 100x increase in adhesion.

Pure PP cannot be usefully corona treated – the surface is totally destroyed. That's why "PP" films contain copolymers at their surface. These can readily be treated.

Chemical Entanglement



Because entanglement is three crossings, it doesn't take much to get one surface entangled with a polymer which itself is nicely entangled.

In the diagram, some functional group is on the surface (e.g. an -OH or -NH2 group) and this reacts into the polymer, for example via a maleic acid co-polymer group or into a crosslinking systems such as epoxy or polyurethane.

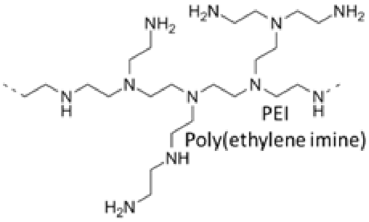
Because so many surfaces contain usable -OH groups, a generic class of adhesion promoters are the X-TES, the triethoxysilanes (or trimethoxy silanes). The silane portion reacts onto the -OH groups which might be on bricks/cements, metal (oxide) surfaces, wood, paper... The X can be -NH2 (in APTES), an acrylate, an epoxy or more.

The trick with adhesion promoters is to not try too hard. We've all added more molecules, provided more heat, increased the UV power to get stronger adhesion … only to see adhesion fall dramatically. When we try too hard, the interface is both too brittle and not sufficiently entangled into the rest of the system:

Great promotion from a bad polymer

Corona treat some PE or PP. A few % of -OH, -C=O and -CO2H groups are produced. In this case, the -OH groups are irrelevant.

Coat the surface with a highly branched PEI, polyethylene imine, a polymer with no desirable mechanical properties. It's coated from aqueous solution so there is zero chance of any physical entanglement:



A few of the -NH2 groups react with the -C=O and -CO2H groups to form an entangled network with the PE/PP. Now coat this surface with an epoxy, a polyurethane or a UV acrylate. Some of the remaining -NH2 groups react into the coating, so the whole system is entangled and adhesion is strong.

This works wonderfully subject to one strict condition: the PEI coating must be thin, preferably 10-20nm, and no more than 50nm. Why? Because it is such a weak polymer that a thicker layer will lead to cohesive failure within the PEI … and because it's highly water soluble so will attract humidity to do bad things at the interface.

Yet again in adhesion, less is more.

Your own chemical entanglement

Life is easy if you can use one of these standard adhesion promotion tricks. But you can't simply add "an adhesion promoter” – they aren’t magic molecules. You have to be specific about what you are attaching to what. For your own system you might have to use some thought. To stick to steel, starting with a phosphate group is always a good idea, as is the case when sticking to teeth. To stick to gold, sulfur is a great choice. What’s at the other end depends on the specifics of your system. If it’s a classic urethane, epoxy or UV coating that subsequently cures then an amine is a good idea. Similarly, if it’s a curing unsaturated system then some sort of double bond will be fine. If there really is no chemical functionality, then the chain sticking out from your surface will need to be long enough to entangle with the other surface. We don’t care how we get entanglement, we just need to use our knowledge of the system to ensure enough entanglement across the interface.

If you can make your promoter polydentate or a chelate, that helps with long-term adhesion. The three silane sites on X-TES cannot react to give a tripod, but having three sites allows for backup if one fails. Although you can get good adhesion with an acrylic acid, maleic acid is preferred because there is a backup acid in situ.

Just be intelligent with your choice and remember that too much of a good thing is bad.

Adhesion Butt

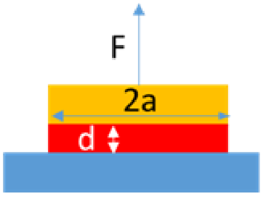
Links

Water-based Adhesive, Adhesive

A pure butt joint provides strong adhesion. The problem is that a small shift in the applied angle changes it from a butt joint to a Adhesion_Shear joint with a large Archimedes lever effect that will destroy the bond.

Less is more

When an adhesive package tells you to use small amounts of the glue that's a big clue about the science of butt joints. Manufacturers make money by selling *more* not less, but they want happy customers and a thick layer of adhesive makes for a weak butt joint.



With our adhesive of modulus E we get off to a good start because the modulus relevant to adhesion in butt mode is the *bulk* modulus K which depends on the Poisson ratio ν which is by how much the adhesive shrinks in the horizontal direction when pulled in the vertical direction:

$$K = \frac{3E}{1 - 2\nu}$$

For a typical adhesive, ν ~ 0.33 so K is ~ 8x larger. For a pure rubber adhesive, ν = 0.5 so K is infinite!

The force needed to break the bond for a disk of radius a with an adhesive thickness of d is:

$$F = \pi a^2 \sqrt{\frac{KG}{d}}$$

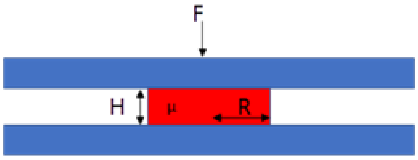
G might be the naïve surface energy W or, more realistically, some plastic deformation value 100s of times larger.

There are three key points:

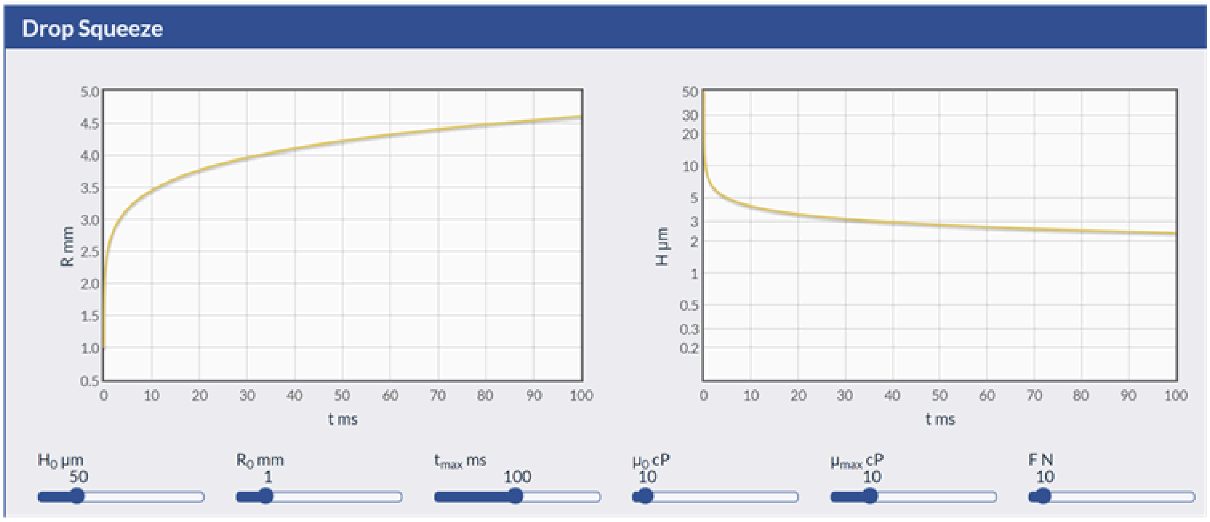
1. You want a strong adhesive (of course) but not too strong/brittle otherwise G might be small through lack of dissipation.
2. You need the thinnest possible layer of adhesive
3. This only works when F is perfectly vertical

The reason for (3) is that adhesion is a property of the system. This idealised large F comes because the forces are distributed equally around the bond. As soon as F is at an angle, you get large concentrations of forces at certain points, and once these concentrated forces are large enough (with Archimedean lever arms), a crack will start and the whole bond will break.

Stefan’s squeeze law



Because a super-thin adhesive layer is so advantageous, we need to know how to obtain one. It seems obvious that you just put a drop of adhesive in the centre of the joint and squeeze hard. But Flow_Stefans Squeeze law gets in the way. For a given force, the speed at which the thickness, H, decreases goes as H^3 and as the radius, R, of the drop expands, speed decreases by R^4. This combination means that the force to obtain a thin layer is impossible to achieve in practice.



H_0 μm 50

R_0 mm 1

t_{max} ms 100

μ_0 cP 10

μ_{max} cP 10

F N 10

https://www.stevenabbott.co.uk/practical-adhesion/drop-squeeze.php

Adverts for superglue show butt joints created between two super-smooth steel cylinders via application of 10 small drops instead of 1 large drop. If we say, naively, that this reduces the Stefan’s law R by a factor of, say, 5, then the force needed to create a thin drop decreases by 5^4, 625x. That is a big extra safety margin if, as per the adverts, you want to bungee jump off a bridge, supported by those two glued cylinders.

While writing my book *Sticking Together* I found that I needed to refer to Stefan at some point … then at another, then another… I’d not realised that it’s such an important part of adhesion. It subsequently turned out that the question of how many drops you use, in which geometry, is a huge question for the adhesives industry. There seems to be no appable algorithm about the big question which is whether a bubble will get trapped. The only app I could manage is https://www.stevenabbott.co.uk/practical-adhesion/Bubble-Gone.php which has plenty of formulae and discussion for those who are interested.

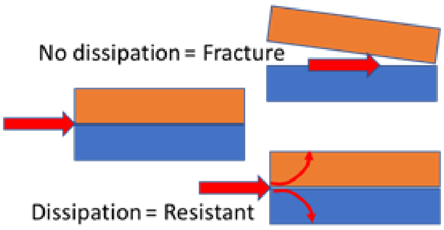
Adhesion Crack resistance

Links

Water-based Adhesive, Adhesive

Adhesion is about stopping a crack growing under an applied force. So let's see what we can do to resist crack growth.

Dissipation



It's quite simple. If the force you are applying to the interface merely has to break a few bonds straight ahead of it, it will carry on doing so. If you have a piece of "strong" glass you can easily break it by introducing a small crack onto the surface. As you bend around the crack, the glass fractures where the crack energy is focussed … and continues to crack.

If the force is sent away into the bulk, then there is no longer enough energy to keep the crack going. Chewing gum is hard to remove not because it is strong but because it is weak enough to absorb the crack energy, dissipating it as viscous heat as it flows. Car windshields include a soft polymer, polyvinyl butyral, because the soft, weak, polymer dissipates the stresses away from the "strong" glass.

Because adhesion is a property of the system, we can easily make glass more resistant to a crack – simply heat it close to its melting point. Similarly, chewing gum is brittle and easy to remove if it is cooled to liquid nitrogen temperatures.

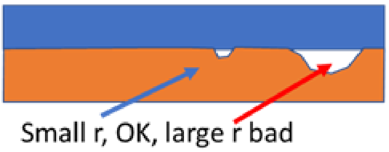
And because of Flow_TTS-WLF Temperature-Time Superposition we can say that room temperature glass under a very slow stress will resist the crack and that a super-fast force will remove the chewing gum as its polymer molecules have no time to move and absorb the crack energy.

In between these two extremes of glass and chewing gum (or, more scientifically, Pressure Sensitive Adhesives, PSA) the best way to ensure dissipation across an interface is to create Adhesion_Entanglement, so that attempts to move one bit of a polymer chain will be dissipated across many nm of chains, just as trying to pull hard on a tangled piece of string makes things worse by spreading the load across many pieces.

Relaxation times

In order for the adhesive to dissipate the crack energy, the polymer chains need to move. Another way of thinking about temperature and time is via the relaxation times of the polymer. When we do rheology we tend to focus on Flow_G' and G". One of the wonders of rheology is that you can interconvert between different measurements to see the system in a different way. When you start to think of your system in terms of relaxation times, new ways to build in crack resistance suggest themselves. See the https://www.stevenabbott.co.uk/practical-rheology/Interconversions.php Interconversions app for an introduction to this fascinating way to expand your knowledge of your system.

Griffith Cracks



The other way to reduce cracking is by not providing your own sources of cracks through voids or junk. Instead of over-focussing on the adhesive, paying attention to removal of defects can be a great strategy. The basic idea is simple but powerful. The Griffith crack law says that the stress σ at which a crack yields depends on a modulus E (stronger is better), a factor G which we'll discuss in a moment and the size of the defect r:

$$\sigma = \sqrt{\frac{EG}{r}}$$

If you reduce the average size of your defects by a factor of 4 then you double the strength of your joint. It can be hard to double the strength of an adhesive, so it can often be a good idea to find ways to, say, remove bubbles from your adhesive system.

In naïve adhesion theory, G is the surface energy. In the real world (the Irwin modification of Griffith) G can be 1000s of times larger because it represents plastic deformation around the crack, in other words dissipation of the crack energy over a nm scale.

You tell people about Griffith Cracks and they say "Yeah, yeah" and ignore you – it's far too boring compared to focussing on fancy adhesion formulation science. But I've had several cases where it turned out that some pesky dirt or bubbles have, indeed, been the cause of unreliable adhesion, fixed by relatively simple filtration and de-aeration steps.

Adhesion Crosslinking

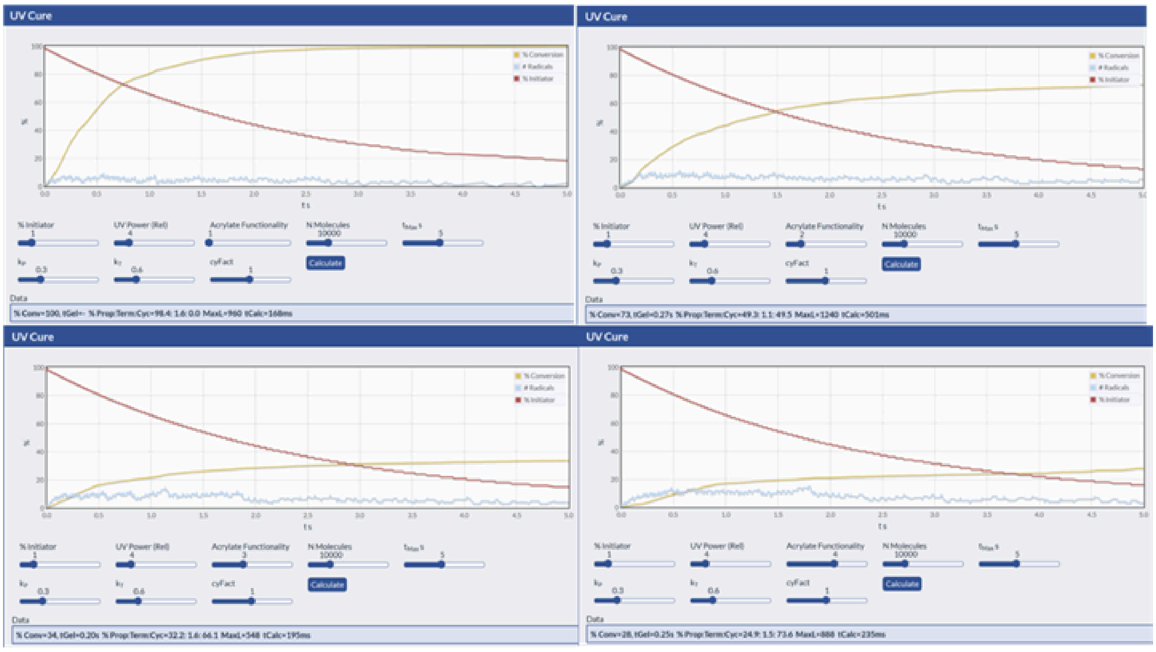
Links

Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint

> If, on average, we have X-functional components in our system with a Y% fraction of reactive bonds, it would be good to know, after a given time, how much of the Y% has been converted into permanent bonds and how much of the system is comprised of 1-, 2- or more functionalities.
>
> Although we can usually answer the Y% question, and although the questions are inter-related, the functionality question is mostly obscure.

Degree of cure

A UV curing app provides a good example of how complex crosslinking can be.



% Initiator 1, UV Power (Rel) 4, Acrylate Functionality 1, N Molecules 10000, t_{Max} s 5, k_P 0.3, k_T 0.6, cyFact 1
Data: % Conv=100, tGel=- % Prop:Term:Cyc=98.4: 1.6: 0.0 MaxL=960 tCalc=168ms

% Initiator 1, UV Power (Rel) 4, Acrylate Functionality 2, N Molecules 10000, t_{Max} s 5, k_P 0.3, k_T 0.6, cyFact 1
Data: % Conv=73, tGel=0.27s % Prop:Term:Cyc=49.3: 1.1: 49.5 MaxL=1240 tCalc=501ms

% Initiator 1, UV Power (Rel) 4, Acrylate Functionality 3, N Molecules 10000, t_{Max} s 5, k_P 0.3, k_T 0.6, cyFact 1
Data: % Conv=34, tGel=0.20s % Prop:Term:Cyc=32.2: 1.6: 66.1 MaxL=548 tCalc=195ms

% Initiator 1, UV Power (Rel) 4, Acrylate Functionality 4, N Molecules 10000, t_{Max} s 5, k_P 0.3, k_T 0.6, cyFact 1
Data: % Conv=28, tGel=0.25s % Prop:Term:Cyc=24.9: 1.5: 73.6 MaxL=888 tCalc=235ms

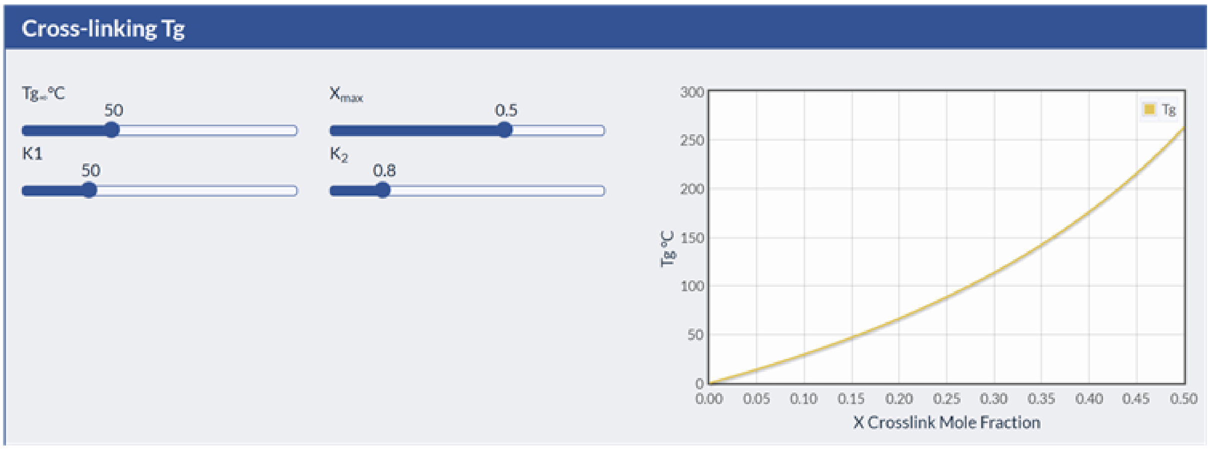
https://www.stevenabbott.co.uk/practical-coatings/UV-cure.php

We have a UV curing system with a UV power, a level of photoinitiator and propagation and termination rates. In the 4 screen shots we just change the acrylate from mono- up to tetra-functional. The yellow % conversion line shows that each increase in functionality decreases the conversion over time – from 100% for the mono-functional down to 28% for the tetra-functional. So although our extra functionality can deliver benefits in strong crosslinking, they are undermined by ending up with ~70% of the groups being unreacted.

It's always important to monitor the degree of cure via FTIR, confocal Raman, ultrasound reflectometry (which, interestingly, gives you the rheology, Flow_G' and G") or calorimetry. But why is the degree of cure so limited?

Rising Tg

If your system is currently above the glass transition temperature, Tg, then monomers, oligomers and polymer chains are free to move. But at a certain degree of cure, the Tg will have risen so you are now just below it – and free motion stops. We can see this in an app:



Cross-linking Tg

Tg_∞°C 50

X_{max} 0.5

K1 50

K_2 0.8

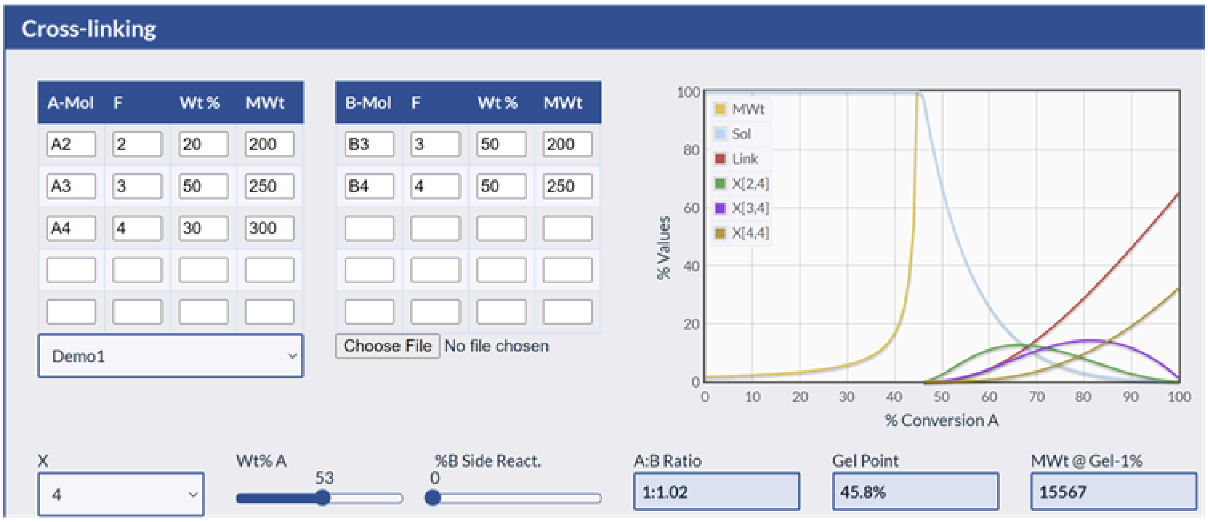
https://www.stevenabbott.co.uk/practical-coatings/X-Link-Tg.php

In this particular example, an 8% mole fraction of crosslinks is enough to raise the Tg to above room temperature. To get to 20% you need to raise the temperature to 65°C.

And that is the answer to the crosslinking crisis described in the previous section. Carrying out the curing at a higher temperature would give you more conversion in a given time because:

1. A 10°C rise might typically double the speed of cure
2. You remain above the Tg so the reactive species can continue to move to find a reactive site.

Distribution of functionalities



Cross-linking

A-Mol	F	Wt %	MWt
A2	2	20	200
A3	3	50	250
A4	4	30	300

B-Mol	F	Wt %	MWt
B3	3	50	200
B4	4	50	250

Demo1

Choose File No file chosen

X 4

Wt% A 53

%B Side React. 0

A:B Ratio 1:1.02

Gel Point 45.8%

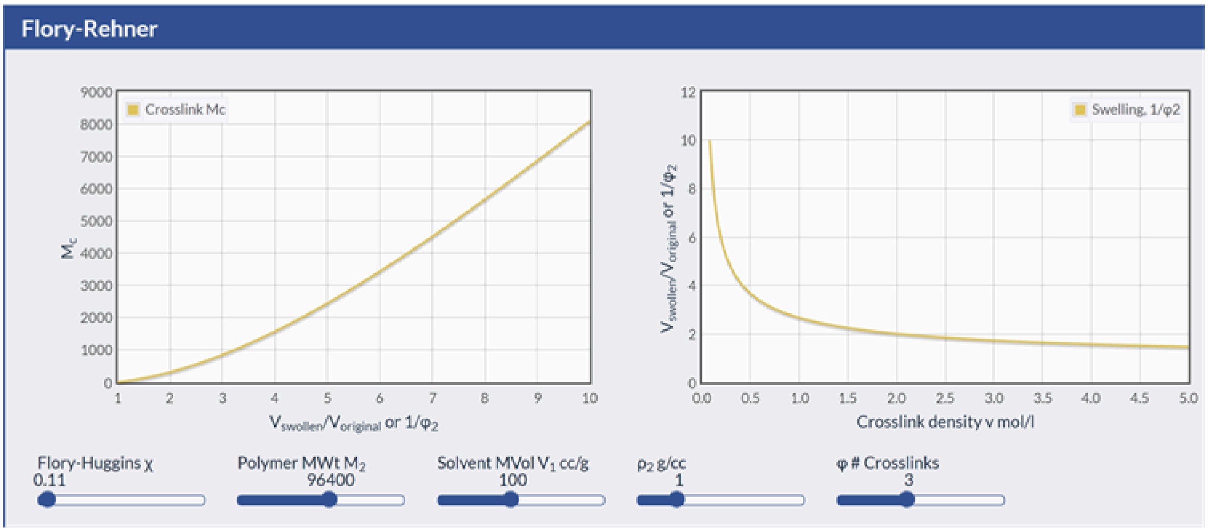
MWt @ Gel-1% 15567

https://www.stevenabbott.co.uk/practical-coatings/X-Link.php and the similar https://www.stevenabbott.co.uk/practical-coatings/X-Link-UV.php for UV crosslinking

Miller-Macosko theory is not for the faint-hearted. But if you have ever wanted to know about how the overall MW changes (in the example up to infinity at the gel point of 45.8%) and how the di-, tri- and tetra-functional group concentrations change with conversion, then the app shows you how to do it. At 70% conversion there are 12, 9 and 8% respectively of 2, 3 and 4 functional links.

Swellability

One reason for creating a crosslinked system is to reduce the potential for swelling by solvents or other small molecules (e.g. plasticisers) in the system. The shorter the distance between crosslinks, the lower the potential for swelling. By measuring the degree of swelling if you stop the curing reaction at different stages, you can get a good idea of this distance between crosslinks. The standard approach is to use Flory-Rehner theory, which is also discussed in Absorbency_Swelling:



https://www.stevenabbott.co.uk/practical-solubility/polymer-swelling.php

The equation, which is not especially obvious or intuitive, is shown in the app. The graphs represent two ways of plotting the data from the calculations. Suppose you find at some point that there is 4x swelling. From the graph on the left you see that there is a molecular weight of ~1500 between links or, from the right-hand graph that the crosslink density, v = 0.4 mol/l.

Although any single measurement may not be too insightful (is 1500 large or small, is 0.4 mol/l good or bad?), measurements during the curing process as well as measurements comparing formulations will lead to some deep insights into your crosslinking system.

Mechanical properties

By following the cure via rheology, Flow_G' and G", using the full capabilities of Time Temperature Superposition, Flow_TTS-WLF, you will get to see how successful you are at establishing a balance between raw strength and resistance to brittle fracture. Even if the mechanical properties look good at room temperature, you need to know if you are close to becoming too weak if used under higher temperatures or too brittle at low temperatures or, equivalently, prone to creep over long timescales or prone to brittle failure from high-speed impacts.

It is common to get super-high crosslink densities with epoxies by curing them above 100°C. They can then become desirably strong and undesirably brittle. The fix for the brittleness is to put core-shell rubber particles into the adhesive. The rubber core absorbs the crack energy, while the shell ensures that the rubber is linked into the epoxy matrix.

Entanglement with the adherends

If the system is not entangled across the interface, the adhesion will be poor. Unless the adherends have functional groups at the right (low) concentration and of suitable reactivity into your system, you need to select appropriate Adhesion_Adhesion promoters that will give the necessary cross-interface entanglement.

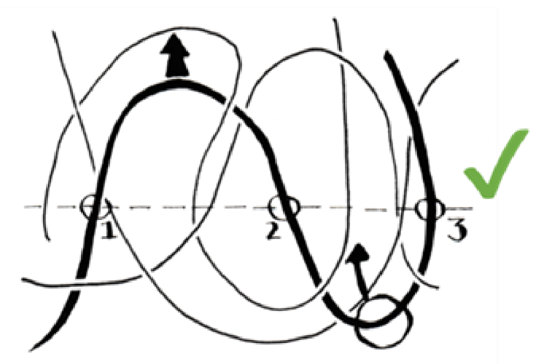
Adhesion Entanglement

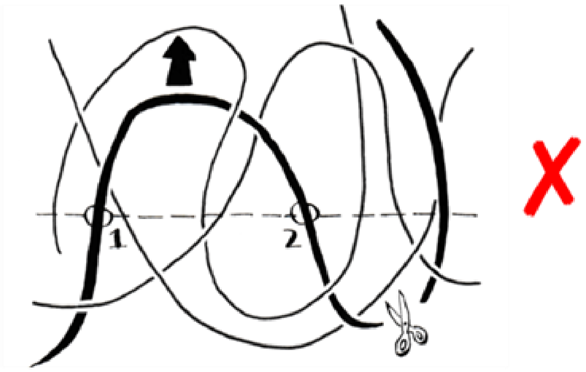
Links

Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint

> If you try to disentangle a mess of string you find that the harder you try, the more it resists. Your energy gets dissipated across many sub-strands and they each help stop the main strand from being pulled out. That’s exactly how strong adhesion is achieved … by dissipating crack energy across many individual polymer loops.

Entanglement = 3 Crossings



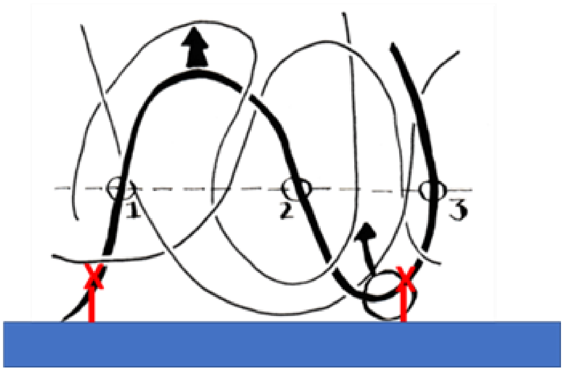


The dotted line is a real or virtual interface. In the left-hand image, the polymer chain of interest crosses it 3 times. This means that it is entangled. If you pull up on the big arrow at the top left, the lower loop gets tangled with the other loops.

In the right-hand image, the chain has been cut so it now only crosses 2 times. When you pull up, the loop is merely intermingled and can be pulled out.

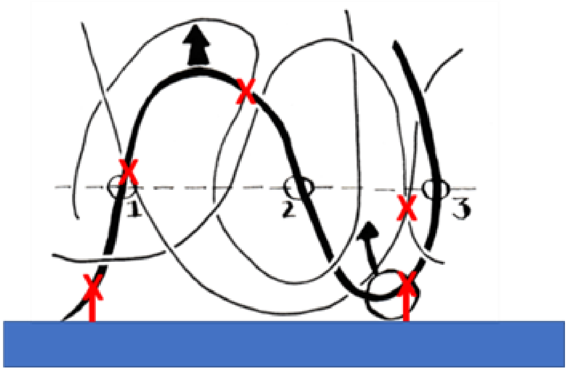
For polymer-polymer adhesion, you just need to get sufficient lengths of each polymer tangled up with the other, lots of three crossings, and that’s it.

But how do you entangle a polymer with, say, an aluminium surface?



You make sure that the polymer chain has a low level of reactive groups that can link on to some reactive sites on the metal surface. Now the two systems are entangled. An example is maleic acid co-polymers of polyethylene.

You can do the same thing with a crosslinking polymer.

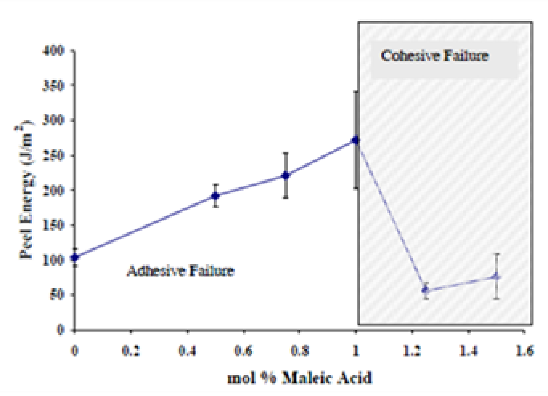


The groups on the metal surface might be something like APTES (aminopropyl triethoxy silane) which allow the amine group to participate in a classic epoxy or urethane crosslinking. Now the tangles are via chemical links (the red Xs) rather than the physical crossings in the previous image. In terms of entanglement *there is no difference* between these cases. To break the tangle you have to break a main chain bond or a crosslink bond, basically the same degree of difficulty. Of course a crosslinked system is less resistant to long-term creep, but at the core, a tangle is a tangle, whether it is physical or chemical.

Too much of a good thing …

Many of us imagine that if we add adhesion promoters or functionalise the polymers to react more strongly with a surface that we would get stronger and stronger adhesion. But as is universally found, as you increase the functionality adhesion goes up, up, up … then crashes down. Too much of a good thing is bad.

Too few bonds are, indeed, not good enough. But with too many bonds you have a brittle interface which is also not properly integrated into the bulk of the structure.



The famous experiments from Prof Richard Wool showed this effect. The image (reproduced with his kind permission) shows that 0.5% adhesion promoter doubled the adhesion, 1% tripled it, but 1.25% halved it. The interface was now brittle and the crack ran along the thin layer of adhesive at the interface, allowing him to call it “cohesive failure”, though I would argue that it should just be called “brittle failure”. This paper was an especially clear example of, and explanation for, a phenomenon many of us had observed but had never explained.

Like many formulators I had often noticed that trying too hard to get adhesion resulted in failure. This made no sense. While struggling to understand adhesion science I was encouraged to read the papers of Prof Richard Wool. I scan-read one that seemed of no interest and closed down the pdf. But something made me re-open and re-read it. It was the one with the image above. A decade of confusion disappeared by looking at one image. It is a powerful demonstration that Adhesion is a Property of the System.

Adhesion Intermingling

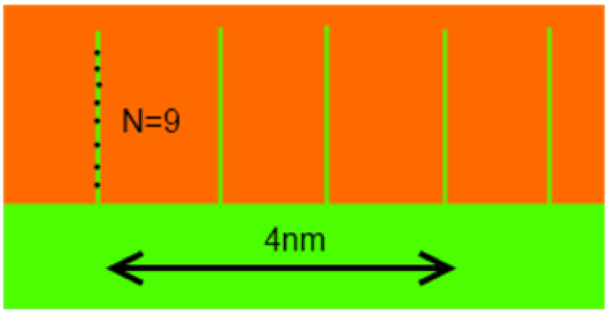
Links

Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint

If you can get some polymer chains across an interface, you automatically get at least 10x the level of adhesion obtained by Adhesion_Surface energy Interactions. This small step towards strong adhesion is called intermingling.

The adhesion, G, from polymer chains of an interfacial density Σ (number per unit area), with chain length N monomer units and a friction coefficient U is given by a formula from the great de Gennes:

$$G = \Sigma N U$$

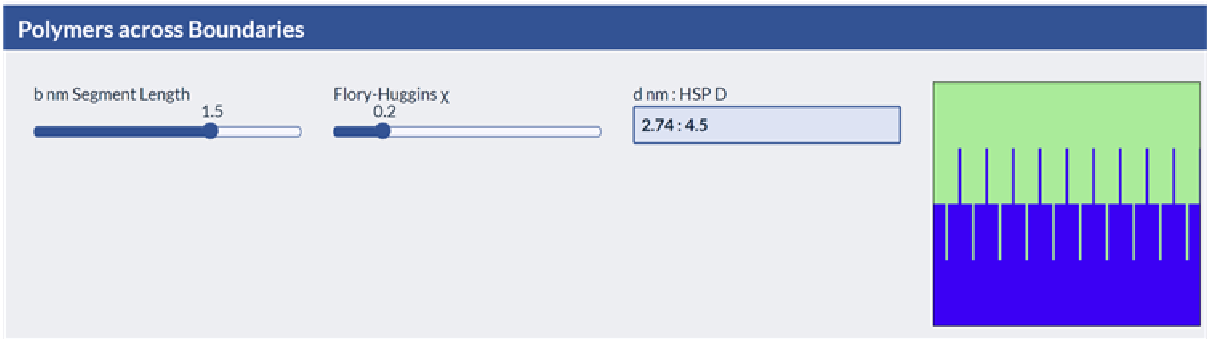


In the image, N is 9, we have 16 chains in 16 nm², so Σ = 1/nm² and U is typically 3.5 J/m². Adding in Avogadro's number and changing to m² from nm² we find that G ~ 400 mJ/m², a factor of 10 greater than a typical surface energy of 40 mJ/m². If a typical monomer length is 0.2 nm, then this 10x improvement comes from only 1.8nm of cross-interface polymer. How hard is it to get 1.8nm of intermingling? The Helfand formula tells us, as discussed next.

The famous Prof Richard Wool investigated intermingled or, as he called it, "nail" adhesion using planks of wood and nails. Enlisting the help of the university carpenter he compared the de Gennes theory for polymers with his data for planks held together with different lengths (n), nails/length (Σ) and nail friction (U). The beauty of science is that the de Genne formula worked well at both the macro- and nano-scale.

But it's impossible!

You sometimes see claims that this sort of beneficial cross-interface intermingling is impossible "because most polymers are immiscible". While it is surprisingly true that most polymers are immiscible, it is an irrelevant fact. As was shown a long time ago by Helfand, you can get plenty of intermingling if the polymers aren't too incompatible:



Polymers across Boundaries

b nm Segment Length 1.5

Flory-Huggins χ 0.2

d nm : HSP D 2.74 : 4.5

https://www.stevenabbott.co.uk/practical-solubility/polymers-across-boundaries.php

The Helfand formula is trivial. For a Kuhn segment length (something like 5 monomer units) of b and a polymer-polymer χ parameter, the distance d of intermingling is given by:

$$d = \frac{b}{\sqrt{6\chi}}$$

Getting a few nm is no problem, so intermingling boosts to adhesion are common. This almost certainly happens even with some PSA systems. An acrylic PSA has no problem intermingling (maybe after a bit of time) with an acrylic surface to give a significant boost to the overall adhesion.

Estimating the χ parameter is simple via the Distance between the Dissolution_Hansen Solubility Parameters of the two polymers.

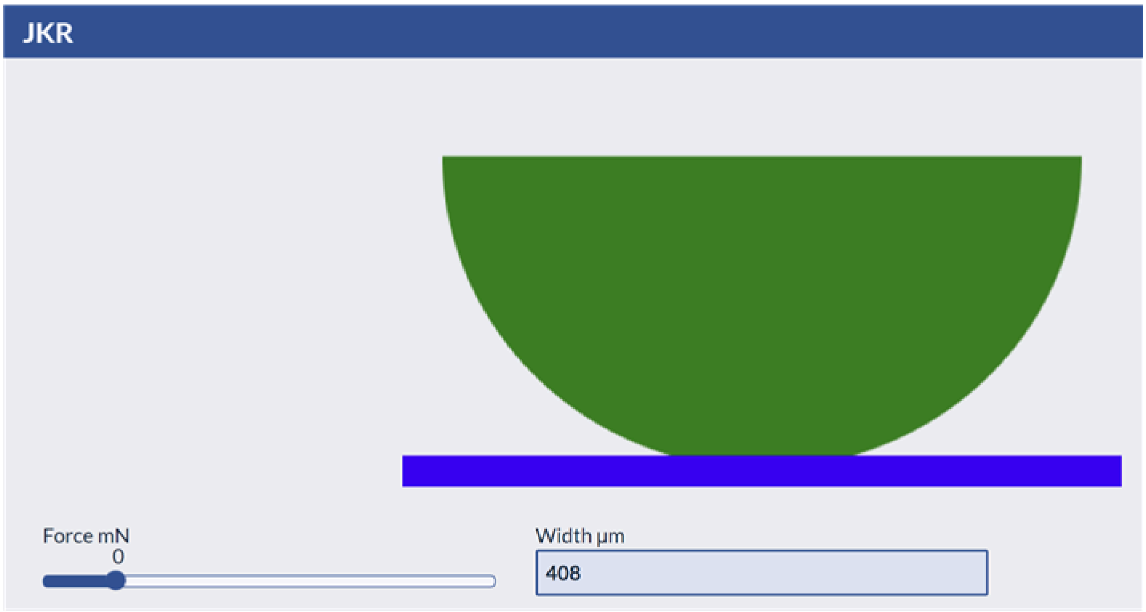
How do you get the boost? Helfand is thermodynamics, getting the boost involves kinetics. Heat, solvent and time help increase intermingling, crystallinity makes it almost impossible. The inability to adhere to polyethylene is not due to any difficulty in finding Helfand-enabled polymers. It’s due to the crystallinity of the surface, with HDPE being harder to stick to than LDPE.

Adhesion JKR

Links

Water-based Adhesive, Adhesive

JKR theory is one of the cornerstones of adhesion science. Because it focusses on pure surface energy then on its own it's not all that useful (see Adhesion_Surface energy Interactions) but once you start using it to measure hysteresis, a world of understanding opens up.

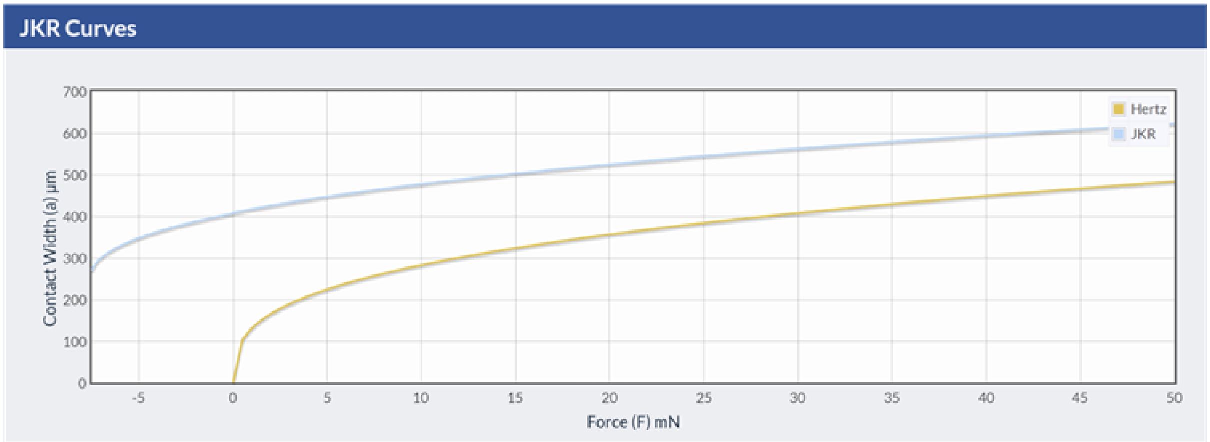


Force mN 0

Width μm 408

https://www.stevenabbott.co.uk/practical-adhesion/jkr.php

The app image shows a JKR experiment that has just started. A smooth rubber ball of radius 20mm and modulus of 5 MPa is placed next to a smooth surface, with zero force applied. You'll notice that the ball is squashed, i.e. some extra force has pulled it into contact. That force is surface energy. Via some means the contact length, *a*, is measured. You then apply some modest extra forces and measure how *a* changes. By plotting *a* versus force you get a curve like the one in blue (the Herz curve is a curve with no surface energy)



https://www.stevenabbott.co.uk/practical-adhesion/jkrcurves.php

Notice that you have to go to a negative force (pulling) to get the ball to unstick itself.

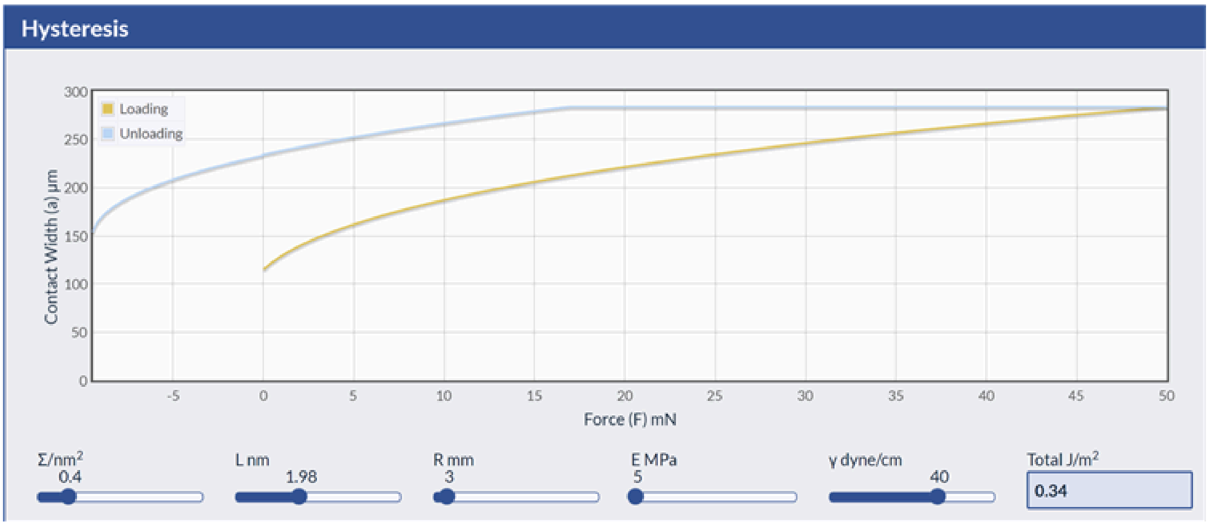
The JKR (Johnson, Kendall, Roberts) formula for *a*'s dependence on radius R, modulus E*, surface energy γ and force F is:

$$a^3 = \frac{3R}{4E^*}\left(F + 6\gamma\pi R + \sqrt{12\gamma\pi RF + (6\gamma\pi R)^2}\right)$$

> This famous formula was created in a pub in the city of Derby where Kendall was working on adhesion of train wheels to rail tracks. Johnson was a professor at Cambridge and Roberts was in Johnson's lab researching windscreen wiper rubbers on glass. It's a lovely example of apparently mundane industrial problems producing one of the great formulae of adhesion science.

Unlike classic methods (Zissman plots, Owens-Wendt…) the JKR approach gives a real measure of surface energy. But as surface energies aren't so interesting for adhesion, why do we care?

Now look at this Loading curve (increasing force) and Unloading curve (after some time at a constant high force.



Σ/nm² 0.4
L nm 1.98
R mm 3
E MPa 5
γ dyne/cm 40
Total J/m² 0.34

There is a large hysteresis, i.e. the unloading curve corresponds to a higher "surface energy" than the loading curve. What’s going on?

This is the first sign of real adhesion. Something is happening across the interface. For example, polymer chains might start migrating (in this case by 2nm) across the interface. As you try to separate the surfaces, these chains have to slide past each other and that extra friction force creates a higher net adhesion force.

Much of real adhesion is due to this cross-interface movement of molecular chains. Small migrations give “intermingling” which adds frictional forces (Adhesion_Intermingling). Large migrations lead to “entanglement” which gives strong adhesion (Adhesion_Entanglement).

Adhesion Peel

Links

Water-based Adhesive, Adhesive, PSA

> Peel seems simple and obvious – but because adhesion is a property of the system, it is surprisingly complex.

It's complicated

The app, although designed for peel of PSA, captures the essence of a typical peel test.

E Carrier GPa 1.5

E_a Adh. MPa 0.3

F N/m 50

Carrier μm 45

Adhesive μm 40

Peel ° 90

σ_{yield} MPa 10

P_{fail} 90° 1392.5N/m

1/α 5.20mm

1/β 0.28mm

https://www.stevenabbott.co.uk/practical-adhesion/peel.php

The adhesive has lifted off at the right-hand edge and we are looking at stresses ahead of the peel front, hence the -ve x values. As this is a 90° peel there are no shear stresses, so the blue line is flat.

At the right-hand edge, the tensile stress is maximum, as expected. At the left-hand edge, the stress is 0, again as expected. What is surprising is that at 0.25mm the stress passes through 0 and there is a compressive peak at 0.5mm. In the Adhesion_Shear chapter you see that not only does a "shear" test produce (and fail because of) peel forces, it shows the same compressive zone.

It's non-intuitive

When thinking about adhesion we tend to focus on the adhesive. But if you change the modulus of the "carrier" (the backing tape for a PSA, this would be the adherend in a conventional peel joint), you find that everything changes. An increased modulus spreads the whole peel zone so the maximum tensile and compressive stresses decrease – the bond is stronger.

Now increase the modulus of the adhesive layer. This should surely make the bond stronger. Yes and no. Yes, the material is stronger, but the tensile stresses are now more concentrated so there's an increased probability of passing a critical strain.

Unlike the Adhesion_Butt joint where a thinner adhesive is mandatory, a thicker adhesive reduces the concentration of forces, which might improve adhesion.

Those who do 3D printing have a classic peel problem. They need light adhesion of the object to the base plate so things stay stable during printing. They then need to remove the object. Pulling straight up is butt-joint mode, with strong adhesion. Peel *would* be easy if the peel could start, but the object can't be distorted into a proto-peel. That's why some objects have a thin rim around the base which *can* be pulled up to start peel.

Why so vague?

Missing from the app is any reference to the strength of the interface, whatever that means. For PSAs there is only 40 mJ/m² of surface energy, but massive dissipation within the adhesive itself (not captured in the app) provides the typical 400 J/m² adhesion found in practice. For conventional adhesives, dissipation through entanglement across the interface is not something that can be specified.

So we have some precise calculations about tensile stress distributions that depend on the bulk aspects of the joint, and handwaving about dissipation and entanglement for the interfacial portion.

If you happen to know the yield stress σ_{yield} of your adhesive then you can at least get an idea of what peel force would lead to yield failure. What you don't know is whether the interface will fail before then.

It's not just peel. The app for shear has many parameters, but none of them are to do with the interfacial capabilities of the adhesive.

And it's not just these apps. Standard finite element analysis is merely a more precise way to calculate the stresses. Working out whether dissipation across the interface will keep the bond together is not part of standard FEA packages.

One famous example in the literature measured the peel adhesion of a polyimide (PI) film to an aluminium (Al) strip. Holding down the PI and pulling the Al gave one value for the peel strength, holding down the Al and pulling the PI gave another. After some more sophisticated analyses they concluded from their high-quality analysis that the "true" peel strength was either 5, 300, 625 or 850 J/m².

There are no simple answers to adhesion problems. The key is to think through those aspects of the system for which there are good theories and use the theory to optimize those parts of the adhesive system that are under your intelligent control.

If the backing stretches

The peel test is supposed to measure the work going into disrupting the adhesive. But if some of the work goes into stretching the backing tape then the peel measurement is, well, what is it? It's a measure of the work needed to peel that material, and if that includes stretching the backing, maybe that's fine. But if you want to know something about the adhesive itself then having the backing effect isn't helpful. So "therefore" you need a stronger backing tape in order to measure the "true" peel.

The scare quote are there to remind you that, as we've seen above, the backing, even when it doesn't stretch, has a significant effect on the measured peel value. And because "adhesion is a property of the system", the "true" peel is an uncertain value.

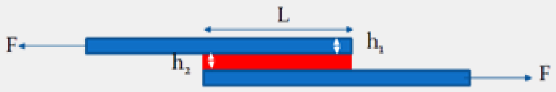
All the same, using a stretchy backing film, unless it's for some positive reason, is not a good idea. To get an idea if your backing is strong enough, this app, https://www.stevenabbott.co.uk/practical-adhesion/peel-value.php, gives you an estimate of whether your current backing (modulus, thickness) is strong enough.

I was asked to help on a tricky adhesive problem and we had discussed their test results that made no sense. So I asked to see the test setup. We walked into the lab and I could see their test machine across the room … and before we'd made another step into the room I said "There's your problem". Even from that distance I could see that their backing tape was massively distorted. No wonder their results made no sense. As this weak backing was a feature of the product, to be able to measure something like the real adhesion the fix was easy – stick some standard adhesive tape to the back, so the backing was now strong enough.

Adhesion Shear

Links

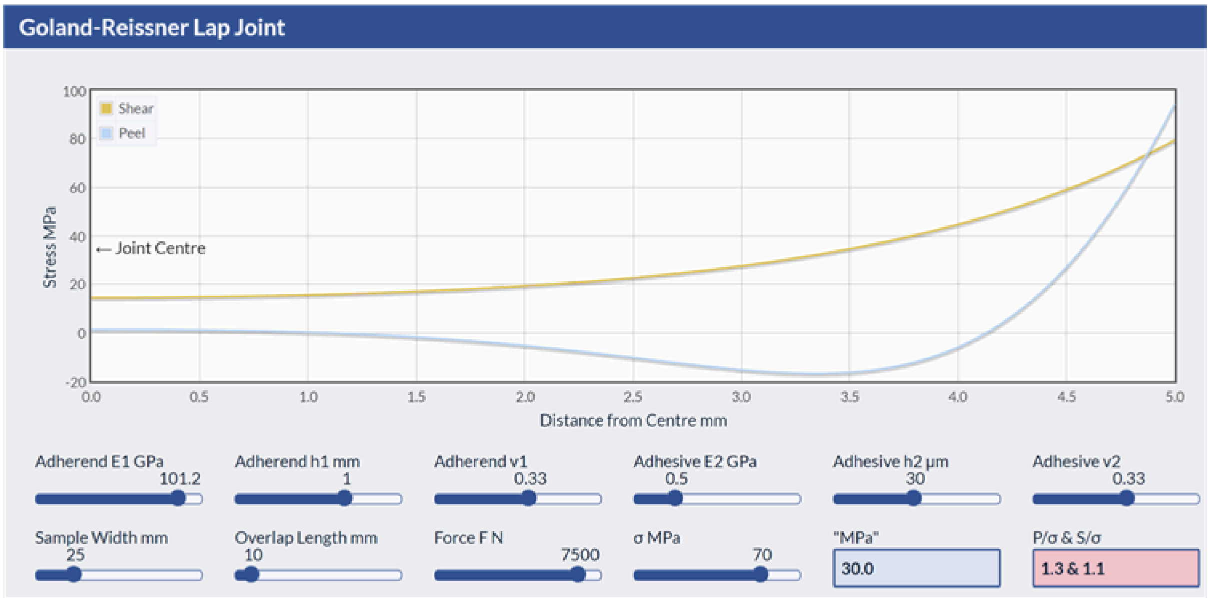
Water-based Adhesive, Adhesive, PSA



The Lap Shear Joint is a beautiful example of why adhesion is a property of the system. Contrary to intuition, the amount of lap, L, (above a certain minimum) makes no difference, and the joints normally fail in *peel*. And the thickness of the adherend, h_1, has a big effect, even though it's nothing to do with the adhesive.

Below we shall see that the standard "shear test" for PSAs is another example of a test that doesn't do what it claims.

A very bad test



Goland-Reissner Lap Joint

Adherend E1 GPa 101.2
Adherend h1 mm 1
Adherend v1 0.33
Adhesive E2 GPa 0.5
Adhesive h2 µm 30
Adhesive v2 0.33
Sample Width mm 25
Overlap Length mm 10
Force F N 7500
σ MPa 70
"MPa" 30.0
P/σ & S/σ 1.3 & 1.1

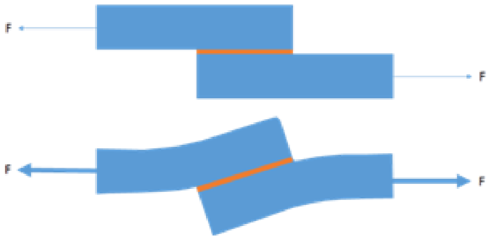
https://www.stevenabbott.co.uk/practical-adhesion/g-rlap.php

If you don't have access to finite element analysis, then the Goland-Reisner equation for a lap shear joint is a rather good approximation. The key lesson from this is that the lap-shear test is so fundamentally flawed that it should seldom be used. Unfortunately, it remains widely used so we have to see why it is so bad and why the quoted values from the test are misleading.

The app shows the stresses from the right-hand edge to the centre of the bond. So with L set to 10mm we see 5mm, with the other 5mm being identical. Now let's see why the test is so bad:

1. We see is that the shear stresses are close to 0 for 2 of the 5 mm. So most of the joint isn't doing anything helpful. Academic tests where they deliberately remove adhesive from a joint show that you can take away 60% of the adhesive with no impact on the force needed to break the bond.

2. The blue peel stresses go negative (in compression) for a few mm inside the bond (something discussed in the Adhesion_Peel chapter), but at the edge they rise sharply *and exceed the shear stresses*. In other words, these shear tests usually fail in peel. A beautiful example of this is with carbon composites. They fail "standard" lap shear tests because although their shear strength (along the fibres) is huge, the peel strength (fibre-to-fibre) is low. For these you have to do the less convenient but more scientific double-lap shear test.
3. Those who have a lap L, a width b and a force to failure of F claim that the bond has a strength, $\sigma = \frac{F}{Lb}$ MPa. This is nonsense. If (above a minimum) they doubled L, F would remain almost unchanged, but they would have to quote a value that is halved.
4. If you decrease the thickness or modulus of the adherend, the peel and shear stresses increase, so the bond will fail earlier. You are not just testing the adhesive, you are testing the system.
5. A well-intentioned increase in the modulus of the adhesive will make the adhesive inherently stronger, but it also concentrates the peel and shear stresses. It's not necessarily the case that "stronger = better".
6. Whether you are using Goland-Reisner or full finite element analysis, there is no way to include the dissipation of crack energy across the interface, no way to work out the effect of entanglement. This means that if you are trying to understand the "real" adhesion at the interface, the lap-shear test is of no help other than distinguishing between cohesive failure (the adhesive was weaker than the interface) and adhesive failure (the interface was weaker).



Why are there peel forces in a lap-shear joint?

The necessary offset of the two forces across the adhesive produce a peel bending moment.

If your real-world system works in lap-shear mode, then by all means use the test. If you are trying to understand the behaviour of an adhesive in shear mode, the test is not fit for purpose.

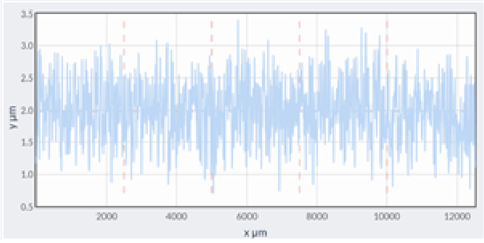
The cockroach of adhesion tests

The lap shear test is used extensively. Because it tells you surprisingly little about real adhesion, why is it still around? Because everyone uses it. A famous centre of adhesion science calls it the cockroach of adhesion tests: "It's everywhere, it's completely useless and no matter how hard you try you can't get rid of it."

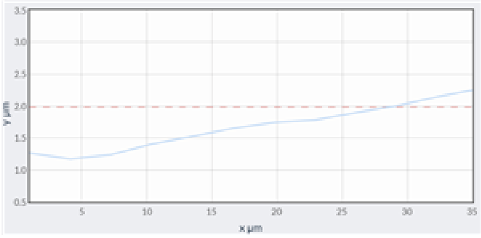
Roughening the surface

See Mechanical_Surface Roughness for a fuller discussion of this topic.

The literature of the lap-shear test is especially full of tests on the effect of roughening the surface. "Everyone" knows that roughening the surface increases the surface area and therefore enhances adhesion via surface energy, while also providing mechanical interlocking between peaks and valleys. If you go to the Surface Profile Explorer app, https://www.stevenabbott.co.uk/practical-coatings/surface-profile-explorer.php and check out the extra surface area, and the peaks and valleys from a rough surface such as this:



you find that there is no significant increase in surface area, and in any case, surface energy is 1000x too small to be relevant. And there is zero chance of peaks and valleys interlocking. It is all an illusion from the mismatch of scales. This same surface with the x-axis scale expanded to be similar to the y-axis (but still a factor of 10 different) is a gentle landscape:

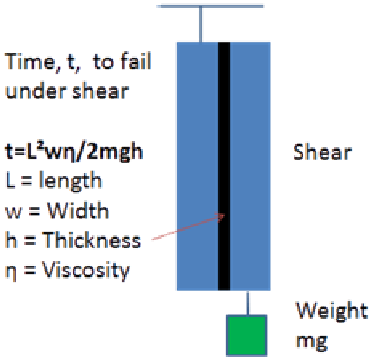


We can all agree that roughening the surface can remove junk and maybe provide some fresh chemical functionality for reaction into the adhesive. But you can also get a clean surface, and the same enhancements, without roughening.

If you shot blast a steel surface you *can* get re-entrant surfaces so that the liquid adhesive is trapped inside. There is good evidence that you can get ~2x more adhesion compared to the same (cleaned) steel without the surface deformations. But this is a rare example of cause-and-effect backed up by good data. You cannot take this special case and generalise to other surfaces.

By all means, if roughening is the best way to get a clean surface, go ahead and use it. Just don't claim that the extra adhesion is due to extra surface area and/or mechanical interlocking.

The PSA Shear test



The classic, official PSA shear test shown in the diagram is only a test for passing the shear test. It has very little relationship with the shear properties of the PSA for a good reason – it often fails in peel, as we saw with the lap joint. So the formula shown in the image is not applicable, much to the puzzlement of people who think that the shear test provides information about shear behaviour.

Just as carbon composites need a non-standard test to get a better idea of shear failure, if you want shear behaviour for PSAs, there are at least two tests (see https://www.stevenabbott.co.uk/practical-adhesion/psa-shear.php) that can do a better job at revealing the true shear behaviour.

Adhesion Surface Energy Interactions

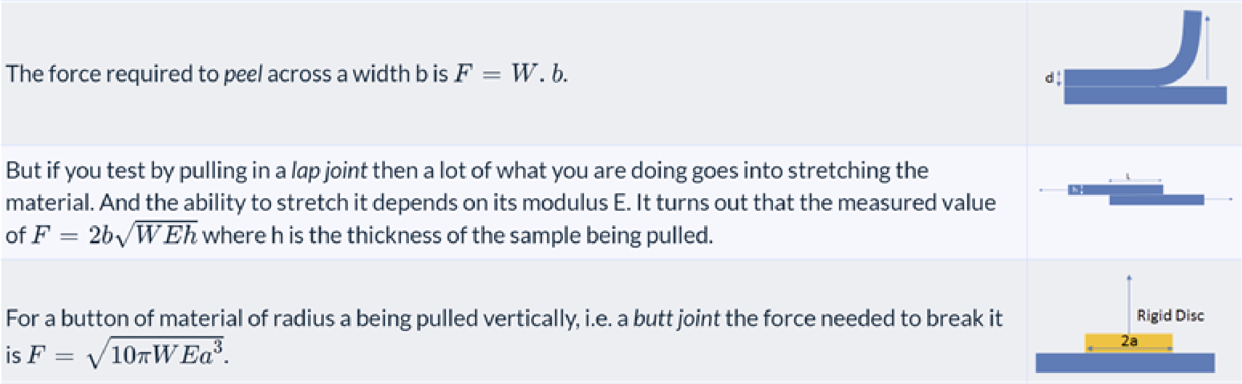
Links

Lipstick, Mascara, Water-based Adhesive, Adhesive

> Most people think that surface energy is important for adhesion. In fact it’s 3-4 orders of magnitude too small to be of any significance. A typical strong adhesive bond might be 400 N/m while a typical surface energy is 40 mN/m.
>
> If you are a gecko then surface energy is all you’ve got. But remember, a gecko does *not* want strong adhesion – it has to easily break adhesion each time it moves its feet.

Adhesion is a property of the system

We can show that this is true in a setup where the *only* adhesion is surface energy between two super-smooth pieces of rubber:

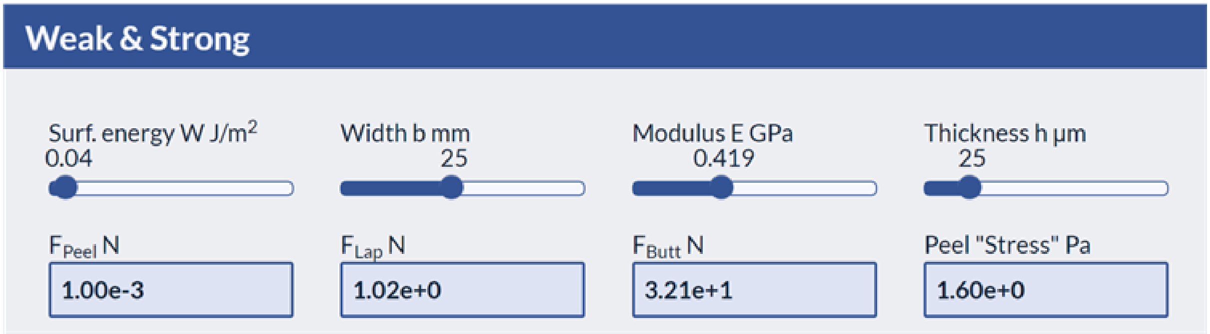


The force required to *peel* across a width b is $F = W.b$.

But if you test by pulling in a *lap joint* then a lot of what you are doing goes into stretching the material. And the ability to stretch it depends on its modulus E. It turns out that the measured value of $F = 2b\sqrt{WEh}$ where h is the thickness of the sample being pulled.

For a button of material of radius a being pulled vertically, i.e. a *butt joint* the force needed to break it is $F = \sqrt{10\pi W E a^3}$.

The app does the calculations for you:



Weak & Strong

Surf. energy W J/m^2 0.04

Width b mm 25

Modulus E GPa 0.419

Thickness h µm 25

F_{Peel} N 1.00e-3

F_{Lap} N 1.02e+0

F_{Butt} N 3.21e+1

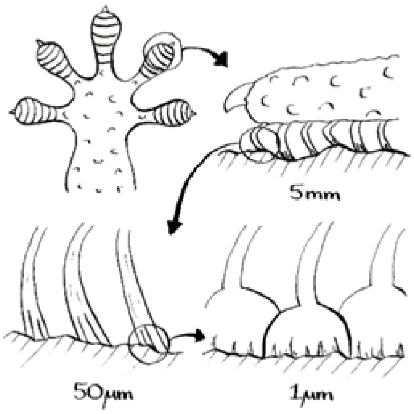
Peel "Stress" Pa 1.60e+0

https://www.stevenabbott.co.uk/practical-adhesion/weak-strong.php

In the case of peel, the force needed to break the bond is 1 mN. Pulling the same system apart under shear requires 1 N – 1000x times more even though the same surface energy is involved. Pulling apart as a butt joint requires 32 N, 32,000 times more.

More interesting is that although (of course) the peel force depends only on the surface energy, the shear and butt joints depend on the modulus and thickness of the polymer, neither of which is an "adhesive" property.

Standard 3D printing using "fused filaments", i.e. squirting hot polymer, requires enough adhesion to the substrate to keep the object from moving around under the forces of the extruded filaments (the results of too little adhesion are not pretty) but small-enough adhesion that the object is easy to remove. Pulling the object straight up requires overcoming impossibly large butt joint adhesion. Peeling it allows easy release, but it's hard to start the peel in the first place. So it is common to print "peelable" structures on the base of the object to allow the peeling process to get under way.



Remember that the above example is for super-smooth rubbers. As soon as you have even micro-roughness, the adhesion falls much further. The gecko gets around this problem via layers of compliance. The toes can bend to accommodate large (5mm) changes in the surface, then the next level is flexible at the 50 μm scale, then the final level is flexible in the nm level. But, as mentioned above, gecko adhesion *must* be weak, otherwise the gecko would not be able to move. There *are* times when you might want a gecko-style adhesion, and there *are* ways to approximate the highly compliant gecko interactions, but generally we want much stronger, more resilient adhesion.

For PSA (pressure sensitive adhesives) the adhesive is deliberately weak so that it can spontaneously flow into full surface contact. For PSA the contact area is pure surface energy, yet they are 1000x stronger for other reasons. Adhesion is a property of the system.

But if the surface doesn't wet, you'll get no adhesion!

We all know that a drop of water does not wet polyethylene. Yet if you put another piece of polyethylene on top (see the Sticking Together video: https://www.youtube.com/watch?v=w1dsiMOvdRc&list=PLLnAFJxOjzZuXlacXKsuBSNk3AH-NuuIr&index=15) the water spontaneously wets both surfaces. The equations for this are well known, and in fact you get full wetting of water even between two PTFE surfaces. So all adhesives wet between all surfaces. Of course many adhesives don't stick to polyethylene or PTFE, but this is nothing to do with wetting.

If surface energy is so obviously irrelevant to adhesion, why do so many people teach that surface energy is *the* important attribute? When the key influencers in any field assert that X is important, it is very hard for non-experts to question the orthodoxy. If surface energy doesn't seem to work for *your* problem, it's obviously your fault, not the fault of the orthodoxy. Fortunately the newer generation of formulators have learned that surface energy is a few orders of magnitude too small to be relevant, so the old story is fading fast.

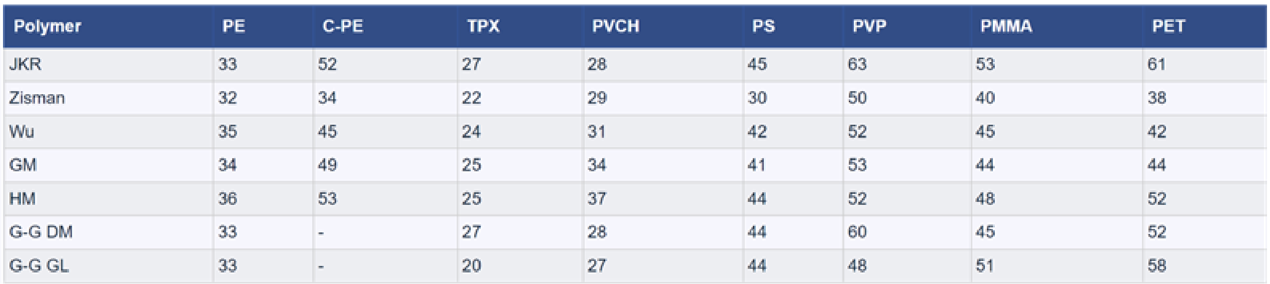
Measuring surface energy

Although everyone obsesses about surface energy, measuring it is surprisingly difficult. The convenient fiction is that we use the Young equation which tells us that θ depends on the surface tension of the liquid σ_L (easily measured), the solid surface energy σ_S which is what interests us and σ_{SL} the solid-liquid surface interaction about which we are very vague:

$$\sigma_S = \sigma_{SL} + \sigma_L \cos\left(\theta\right)$$

This leads to lots of work measuring contact angles with water, diiodomethane and formamide and calculations via Owens-Wendt, Wu or Lewis Acid/Base algorithms (see the app: https://www.stevenabbott.co.uk/practical-coatings/Surface-Energy.php). Or it involves Zismann plots where the 0° intercept of a line of contact angles versus surface tensions is supposed to give you a surface energy.

The scientifically correct method, which requires a more sophisticated apparatus, is via the JKR method (Adhesion_JKR). It is amusing to compare real JKR values with the various attempts using contact angles. The table from the Real-SE app shows that not only do none of the methods give reliable values, you can't even find an objective "least bad" method.



Polymer	PE	C-PE	TPX	PVCH	PS	PVP	PMMA	PET
JKR	33	52	27	28	45	63	53	61
Zisman	32	34	22	29	30	50	40	38
Wu	35	45	24	31	42	52	45	42
GM	34	49	25	34	41	53	44	44
HM	36	53	25	37	44	52	48	52
G-G DM	33	-	27	28	44	60	45	52
G-G GL	33	-	20	27	44	48	51	58

https://www.stevenabbott.co.uk/practical-adhesion/Real-SE.php

The acronyms for the polymers and techniques are described in the app.

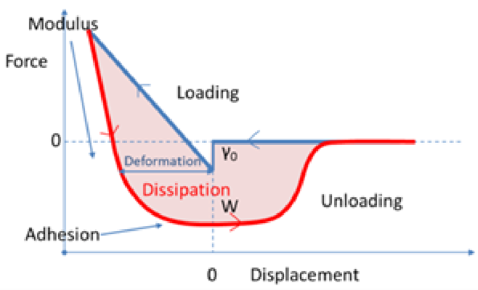
Those who insist that corona or plasma treatment works "because of higher surface energy" should note the columns PE and C-PE, where C means corona-treated. The C-PE column is doubly interesting.

- The standard methods for measuring the surface energy are wildly wrong compared to the real JKR value, so people are drawing conclusions from erroneous values.
- PET usually has a *higher* surface energy than corona-treated PE and yet nothing sticks to untreated PET.

If corona/plasma is your preferred way to remove contamination from the surface, and if surface energy is good at showing the removal, that's great. Adhesion is a property of the system and any system with junk that hasn't been removed is unlikely to give good adhesion. Focus on junk removal, not surface energy.

For those tempted to measure the surface energies of powders under the mistaken impression that (a) the measurements are meaningful and (b) the measurements will help with adhesion, the caution on the use of Washburn Tubes in Absorbency_Porosity and Darcy is worth reading.

Adhesibility



The more productive idea of "adhesibility" is described in an app, https://www.stevenabbott.co.uk/practical-adhesion/adhesibility.php, that invokes a nanoindenter or JKR-style machine for measuring the hysteresis effects that illustrate the potential for adhesion. The initial jump in the loading curve gives the pure surface energy, and the subsequent cycle during loading and unloading gives you a lot of information about modulus and the dissipation that is a key part of adhesion.

A phone call from an irritated reader of attacks on surface energy came up with the question: "OK, so if surface energy isn't important, how else can we measure *adhesibility* of a surface?". What an interesting word, and what an interesting challenge. That's how the adhesibility app was created. I found that although the loading/unloading diagram was "well-known", it was not used in the adhesion science world. Hopefully this will change.

At the time of writing, an "adhesibility" device is under development by a major equipment supplier. More will be added to this chapter once the device is being widely used.

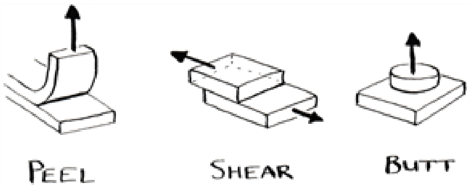
Adhesion Testing

Links

Water-based Adhesive, Adhesive, PSA

Everyone wants a test for the “real” adhesion in their system so they can optimise the adhesive. They won’t find one. Let’s see why.

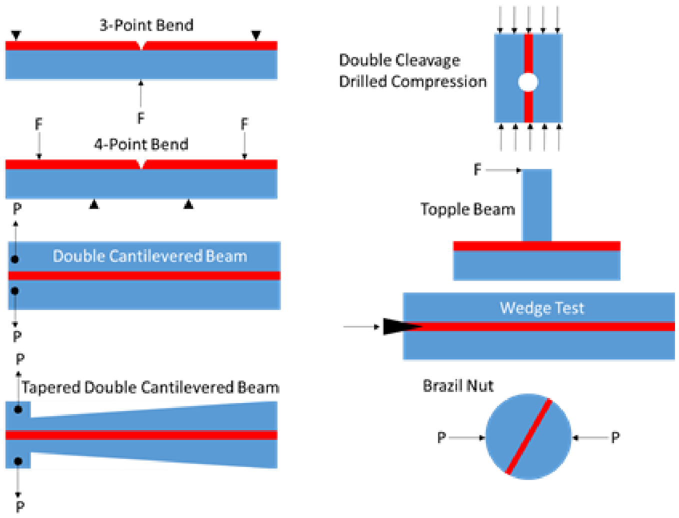
Adhesion is a property of the system



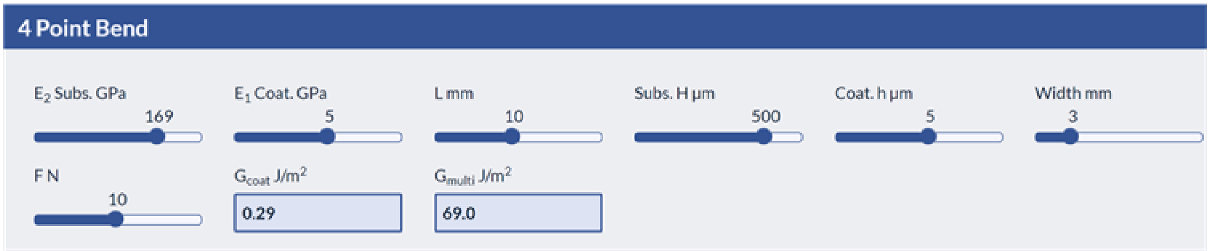
The three core types of test are good examples of why it is impossible to test for the “true” adhesion of a system. Let’s see in turn why standard adhesion tests in each of these modes do not, and cannot, give the “true” adhesion:

- Adhesion_Peel. The stresses ahead of the peel front can extend by several mm and even go into compression. The shape and magnitude of these stresses depend not just on the adhesive but on the adherend, its modulus and thickness. A weak adhesive, a PSA, can show great strength because the peel forces are dissipated over many µm of adhesive. For a strong adhesive, success depends on the dissipation of crack energy across a few nm, something that we hardly know how to analyse and test for.
- Adhesion_Shear. Famously the lap-shear joint does not depend (much) on the overlap, and mostly fails in peel, not shear. Again, the distribution of forces depends as much on thickness and modulus of the adherend as on the adhesive. And, as with the peel joint, the dissipation of crack energy across the (hopefully) entangled interface is something about which we have little information.
- Adhesion_Butt. Now the strength of the joint depends on the bulk modulus of the adhesive, which in turn depends on the Poisson ratio, something we generally know little about. It also depends inversely on the square root of the adhesive thickness, so the strength depends a lot on our ability to create a super-thin, even layer without (say) bubble defects that would be the source of Griffith crack failure (see Adhesion_Crack resistance). It is surprisingly hard to squeeze a drop of adhesive to produce a thin layer, see Flow_Stefans Squeeze, so the strength of a butt joint can depend more on some basic laws of fluid flow than on the adhesive itself. And as soon as the butt force is not perfectly vertical we start to get failure in shear or peel.

What about fancier tests? Professionals tend to like crack tests. At least, unlike the tests above, they start with a known crack in a known position. But look at how many crack tests there are:



Which should you choose? Is 3-Point bend worse than 4-Point. Is a Topple beam better than a Wedge? The answer is that they are all so complex, and so indirectly connected to the "true" adhesion that it's hard to like any of them. If, for example, you look at the app for the 4-point bend,



4 Point Bend

E_2 Subs. GPa 169

E_1 Coat. GPa 5

L mm 10

Subs. H μm 500

Coat. h μm 5

Width mm 3

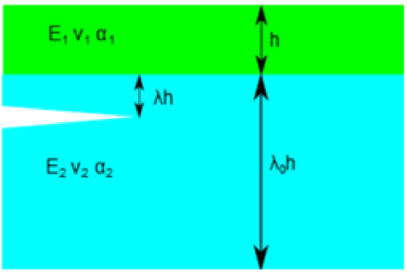
F N 10

G_{coat} J/m^2 0.29

G_{multi} J/m^2 69.0

https://www.stevenabbott.co.uk/practical-adhesion/4point.php

you find from the equations that the journey from data to "adhesion" is complex. The test is used in the world of multi-layer electronic devices for which it may be suitable. For the rest of us it is not a good investment of time and energy.



And cracks are complicated. Even if you start with a known setup:

Depending on a number of parameters, including Dundurs α (see https://www.stevenabbott.co.uk/practical-coatings/Interlayer.php for an explanation of Dundurs α and how it affects other cracking situations), the crack might continue where it is, move higher or lower in the current layer, head to the interface or even jump across the interface and propagate in the top layer. It's complicated.

ASTM-style tests

Many of us carry out tests according to ASTM (or other standards such as ISO) because our customers demand that we pass them. That's fine, life's like that so you have to go along. But just because it is an ASTM test doesn't mean that it's of any use for your system. Carbon composite aircraft structures fail an ASTM standard test for aircraft structures. This isn't because carbon composites are bad for aircraft (they're great) or because ASTM created a stupid test. It's because the ASTM experts at the time made the pragmatic (and correct) decision that for existing aluminium aircraft structures, their lap-shear test was a good pass/fail criterion where lives really matter. A footnote in the standard points out that it is really a peel test and, as it turns out, carbon composites easily fail in peel. ASTM (and the aircraft industry) know that the correct shear test for carbon composites is the double-lap test.

The problem isn't with ASTM. It is with the general adhesion industry assuming that a test designed for one purpose could provide insights into totally different types of adhesion issues.

Note, too, that as the story below shows, you can't mindlessly apply an ASTM test. Subtle details of your product might cause problems unforeseen even by the ASTM experts.

The advice, therefore, is to spend the *minimum* time doing the ASTM-style tests required for your industry and diverting your resources to tests that are more insightful for your specific adhesion issue. Passing those tests will *automatically* help you to pass the ASTM tests.

> Despite the company gathering lots of ASTM test data on a PSA, their formulation efforts were getting nowhere. In such circumstances, my instinct is to blame the test so I asked to see it being performed. They took me to the lab and left me with the technician, who was very good and performed the test exactly as per the standard. As the test proceeded it was clear that their product was doing something odd which meant the results were meaningless. I pointed this out to the technician who agreed: "I tried to tell them, but I'm only a technician". I asked for the test to be re-set and when it was ready I got the senior managers to look at the test. They, too, instantly saw the problem and without too much effort an alternative method was worked out.

Your own tests

You are the expert in the adhesion challenges in your system. So what *really* affects adhesion?

- Is it chemical reactivity across and interface? Well, test for chemical functionality and reactivity.
- Is it performance in cold conditions? Test for Flow_G' and G" and, to make sure you can cope with high-speed shocks (if relevant) extend your rheology via Flow_TTS-WLF.
- Is the problem ageing performance? It's often something to do with water, so do tests that look for, say, water sorption then tweak the formulation to reduce the sorption at high humidities.
- Is it a challenge of getting enough dissipation while resisting long-term creep? This needs some more challenging rheology. But far better doing rheology than meaningless adhesion tests.

Then there are some other considerations:

- Do you need to explore a wide formulation space before optimizing? Find a quick, good-enough test (maybe one that your robot can do) that helps you decide between hopeful and hopeless formulations. You often don't have the time to do detailed ASTM-style tests on a broad range of samples, so you do too few tests within a narrow formulation zone and can miss superior options.
- Do the tests require live judgement of what's going on? This is the 21st century so set up a camera to record every test. Most of the time you don't need to re-run the video, but when something interesting (good or bad) happens, having the video is vital.
- If you are using any form of camera/video, remember to put a lot of thought into good lighting and good lens positioning. It is easy to lose key information which could easily have been recorded with a little more thought.
- Find tests that most closely relate to the most likely end-use failure modes. Sometimes this needs imagination.

> One customer complained of poor adhesion, while all other customers were happy. It turned out that adhesion failure occurred during a slitting operation and their very bad set-up was causing

extraordinary stresses that caused adhesion failure. In this case the solution to an “adhesion” problem was to help them set up their slitting machine correctly. But if this had been an industry-wide problem, we would have had to devise a test to mimic these extraordinary fast and extreme stresses, something that no ASTM test could mimic.

Those specific examples may not be relevant to you. The idea is to encourage you to think through the issues then use the tools within the FST to find solutions.

Cleaning Bleaching and Oxidation

Links

Laundry Liquids, Surface Cleaning

> It's easy to clean and protect using bleach. The problem is how to ensure that there is not too little activity and not too much, doing it for long enough but not too long.

Chlorine or peroxide

Although some specialist *reducing* bleaches such as thiourea dioxide are available, here we discuss the classic oxidizers, chlorine, delivered in liquid form as hypochlorite, and peroxide delivered in liquid form as hydrogen peroxide or in solid form as a "per" salt or (surprisingly perhaps) via a combination of enzymes and glucose.

In terms of killing microbes or reaching difficult stains, the chlorine bleaches win out because chlorine has a higher redox potential, and the gas is a small molecule that can diffuse into hard-to-reach and/or hydrophobic places, which hydrogen peroxide as a hydrophilic solute can't so readily reach. That advantage comes with the side-effect of chlorine readily reaching the nose of the user which is a useful signal that some powerful cleaning is going on, while also being a warning of the presence of a highly toxic gas.

The extra power of chlorine comes with the extra downside of it being more likely to lead to damage to the item being cleaned and (from accidental splashes) to the clothes of those using it.

For general purpose stain removal, if we assume that the stain is only on the surface (even if that surface is within a fabric via capillary flow) then the surface-only action of the peroxide is an advantage as it can still reach the stain while not being able to penetrate deeper into, say, a coloured hydrophobic fabric to bleach out the colour or react with and weaken the fibres. This advantage is less clear for cotton with its more hydrophilic structure.

Sodium/calcium hypochlorite, sodium dichloroisocyanurate

Reacting chlorine with sodium hydroxide solutions gives liquid sodium hypochlorite, the classic bleach. It has to be kept at high pH as hypochloric acid breaks down rapidly to chlorine.

Reacting chlorine with slightly wet calcium hydroxide (lime) gives "bleaching powder" as a mix of calcium chloride, remaining lime and calcium hypochlorite. It is a convenient solid form of bleach.

Sodium dichloroisocyanurate is another convenient solid form of chlorine bleach.

Sodium perborate and percarbonate

It is a convenient fiction that sodium perborate tetrahydrate or the more efficient monohydrate are what they say they are. Instead they are a composed of the 1,4-diboratetroxane di-anion, a fact we can conveniently ignore. It is also a convenient fiction that sodium percarbonate is a percarbonate. Instead it's sodium carbonate with hydrogen peroxide locked into the crystal lattice, again a fact we can ignore.

Stabilising the unstable

We use hypochlorite and peroxide because they are thermodynamically unstable molecules able to react with relatively hard-to-remove microbes and stains. How, then, can we keep them stable up to the moment of use?

Sodium hypochlorite is mostly stable enough in a sealed container to keep the chlorine from escaping, as long as the pH is high enough and the extra ingredients (see below) are chosen wisely. The literature shows no problem with hypochlorite solutions remaining unchanged for at least 1 year.

The per-salts are highly sensitive to transition metal ions such as iron, so it is important to use low transition metal ingredients and avoid contamination in manufacture and to add chelators such as EDTA or a modern greener equivalent to the formulation.

The dry per-salts are relatively stable, though percarbonate needs extra protection on the outside of the particle via sodium silicate or sulfate. Liquid formulations are tougher to stabilise, and need to be relatively acidic and shielded from UV light and heat. For detergents that work best at alkaline pH, there is a difficult contradiction. Again, the liquid formulations need a wise choice of the extra ingredients around them.

Making them more reactive

The per-salts bleach well at temperatures above 40°C but require a boost to be effective at ambient temperatures. Although it is possible to provide smart transition metal (manganese) catalysts, the standard technique is to add an activator that converts the per-salt into a peracid. In Europe the default activator is tetraacetylethylenediamine (TAED) which reacts with the per-salt to generate 2 equivalents of peroxyacetic acid. In the US the default is sodium nonanoyloxybenzene sulfonate (NOBS) which reacts to create peroxynananoic acid. The peracids are, respectively, hydrophilic and hydrophobic, so TAED tends to be better for hydrophilic stains and NOBS for hydrophobic ones.

Famously, one manufacturer introduced a manganese catalyst that boosted the cleaning power of their laundry products. This gave a strong competitive advantage, so the competitors were quick to do tests which proved that this formulation was *too* powerful and that fabrics were getting damaged by the bleach. Whatever the truth of the situation, that specific product was withdrawn and a consensus level of catalysts was achieved throughout the industry. The manufacturers of the manganese catalyst were left with large stocks, so we tried to see if it could help in our specific technical bleaching application. Unfortunately it didn't. We tried a more imaginative formulation with a different catalyst system. This gave astonishing cleaning but was so dangerously reactive that it was unusable. Getting the right balance is very hard.

A reader kindly informed me that *" [There are] other applications for the same catalyst and new ones based on iron, including products for dishwashing, bleaching of paper pulp, and siccatives for alkyd paints."*

Creating peroxide with enzymes

By adding glucose to a formulation containing glucose oxidase, hydrogen peroxide is produced. It is an attractive option in terms of using only natural ingredients, but it has obvious formulation difficulties in practice.

Bleach-stable surfactants

A thickened chlorine bleach is useful for its ability to cling to surfaces such as toilet bowls to provide longer-lasting action. Surprisingly (as they say in patents) many standard surfactants have no problem surviving the high pH, highly oxidative environment of a bleach and a bleach can be thickened simply with classic alkyl ether sulfates (such as SLES), though more complex blends of anionics with amine-oxide, quaternary ammonium or betaine surfactants are also common. The fact that such a wide range of surfactant types can be used shows that bleaches aren't all-powerful destroyers of complex chemicals. Using relatively low-cost surfactants as thickeners also allows the claim of extra cleaning power plus the production of foam which consumers associate (wrongly) with cleaning – though clever foam bleach formulations that cling to the walls of facilities such as dairies are a smart way to provide the longer-term activity to remove tough microbes.

For those who want to thicken with polymers, the sodium salts of polyacrylic acids can usually survive long enough, though some radical stabilizers such as benzoic acid derivatives might be required.

Partitioning

Many of the contradictions between the various ingredients in a cleaning formulation for laundry or dishwashers can be solved by placing ingredients within different parts of a "pod". So a strong peroxide bleach can be kept at acid pH, while other compartments in the pod can contain enough base to produce an overall basic environment for the washing, along with any transition metal catalysts or activators required to increase the bleaching power, especially for cold washes.

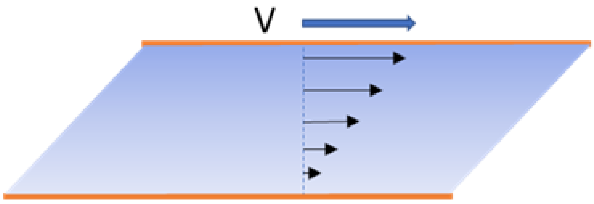
Cleaning Boundary removal

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, Surface Cleaning, Toothpastes, Lipstick, Mascara

> The exact place where we want our cleaning fluid to interact with the dirt is the place where the "no-slip boundary condition" tells us that the fluid velocity is zero. Unless we do something about this, we really can't do much cleaning.

No-slip



In Flow_Couette Flow we see that as we drag some fluid along (using a cloth, brush or just some pressure flow) at a high velocity v, the velocity of the flow is zero at the stationary surface. This is the "no-slip boundary condition" which is a fundamental law of physics verified down to the nm scale.

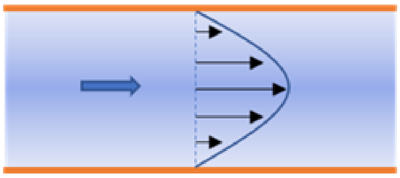
This automatically means that dirt is not being swept from the surface and that fresh cleaning fluid is not being provided to help attack the dirt.

There are three ways to overcome the problem:

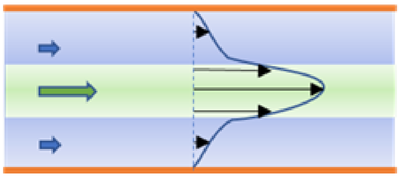
- Create lots of turbulent flow, applying lots of energy with your hands, or by using high-powered jets like those in commercial dish washers
- Add plenty of particles that can transmit energy right down to the surface. For spherical particles, there is a "percolation threshold" (see https://www.stevenabbott.co.uk/practical-coatings/percolation.php) which for spheres is 28% volume fraction, so you need a lot of particles to guarantee that energy applied above reaches down to the surface
- Use a microfibre cloth. The main structure of the cloth transmits forces down to the surface, while the microfibres coming from each main fibre have a bigger chance of contacting the surface and inducing cleaning.

> At a time when the use of microfibre cloths was new and they were being marketed as good at cleaning, my wife, a non-scientist, suggested I should use them for a product development challenge. A proper lab experiment on a known-to-be-difficult real-world scenario was set up to compare a standard cloth with the microfibre one. The control cloth, along with some smart chemical formulation, cleaned fairly well so then it was time to try the microfibre one. We were stunned. It was no contest – the microfibre cloth was instantly and obviously better – we'd never seen cleaning like it. The hunt for an explanation (overcoming the no-slip boundary condition) took years while the decision to change to microfibre cloths was instant.

Cleaning pipes



There are two problems with trying to clean out a pipe full of one fluid by pumping another fluid through it. The first, this time showing Flow_Poiseuille Flow is that the velocity of the fluid at the surface is still zero:



A more severe, and less well-known problem arises if you try to pump a low viscosity cleaning fluid (green) through a pipe with a higher viscosity contaminant. Thanks to the Rayleigh-Taylor instability, the low viscosity fluid simply punches through the centre of the pipe, providing little cleaning towards the edge.

The problem is reduced if the interfacial tension between the phases is super-low, (see Surfactancy_Critical Capillary Number for an analogous situation in oil recovery). If you can't match the surface tensions of the two phases then the trick is to add something to the cleaning fluid which increases its viscosity but which is also an acceptable “contaminant” once the pipe has been cleaned.

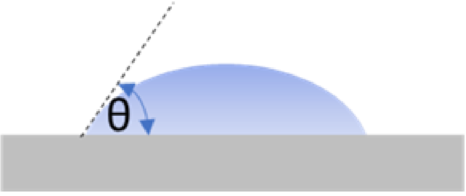
Cleaning Contact Angles

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, Surface Cleaning

There is an unhealthy obsession with contact angles as being important for cleaning. Often their role is minor or even irrelevant. Let's see why.

Contact Angle measurement



The story goes that we put a drop of cleaning liquid onto a surface and measure the contact angle, θ. Pure water on a typical surface will have a contact angle in the range of 40-60°. A water containing Cleaning_Surfactants might have a contact angle in the range of 20-40° and "therefore" cleans better because it wets more.

The relevant theory, the Young equation tells us that θ depends on the surface tension of the liquid σ_L (easily measured), the solid surface energy σ_S which is what interests us and σ_{SL} the solid-liquid surface interaction about which we are very vague:

$$\sigma_S = \sigma_{SL} + \sigma_L \cos\left(\theta\right)$$

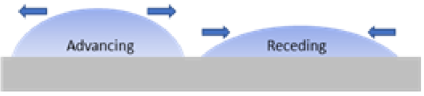
As discussed in Adhesion_Surface energy Interactions there are many problems getting anything useful from this apparently simple equation because of the uncertainties of σ_S and σ_{SL}.

There are good reasons why it is too simplistic to think that we can go from measured contact angle to cleaning efficiency.

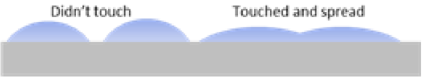
Surface energy forces are very weak. We can easily cover a surface with liquid by adding some extra energy such as a spray or applying a cloth. The capillary number, $Ca = \frac{U\eta}{\sigma}$, represents the balance of inertial forces, velocity times viscosity, compared to surface tension. We can vary U and η by orders of magnitude, while σ varies at most by a factor of 2. It is very easy to overwhelm surface tension effects.

Any liquid between any two surfaces (even water between Teflon) will fully wet those surfaces, because although the liquid (in this example) has a 90° contact angle so doesn't wet the surface, when there is a second surface, for example a cleaning cloth, it is happier to contact the surface than be in air. You can see this effect on my Sticking Together YouTube channel, using coloured water and some polyethylene: https://www.youtube.com/watch?v=w1dsiMOvdRc&list=PLLnAFJxOjzZuXlacXKsuBSNk3AH-NuuIr&index=15.

The measured contact angle is very susceptible to slight contaminants. If you *really* want to know the contact angle of a specific surface, you first have to clean it thoroughly with, say, acetone, alcohol and DI water. If a slight contaminant (1 molecule thick) can significantly change a contact angle, maybe we should focus on contamination rather than contact angle.



Our problem often isn’t one of *wetting* the surface but of *dewetting*. If our cleaning fluid contains any “contaminant” such as a surfactant where the hydrophobic portion sticks to the surface and the hydrophilic head pokes up into the liquid, the *receding* contact angle might be much lower than the equilibrium contact angle that is typically measured. Because, as noted in the previous paragraph, contact angles are highly dependent on contamination, just about anything in a liquid might alter the surface enough to make a measurement of “the” contact angle meaningless.



A lot of the things we care about, such as what happens as drops of water dry on a dish or whether we get the coffee-ring stains involve effects and timescales not covered by classic contact angle measurements. For example, if two drops happened to be a bit too far apart, they will dry as two separate drops. But if the same drops were a bit closer, they would coalesce and dry as a single drop:

> Why these reminders of why contact angle is generally useless? I had, for many years, used fancy contact angle measurements each time there was a problem of bad adhesion in a product. After one such session where again there was no correlation between contact angle and product performance, I realised that *not once in years of measuring contact angles had they helped me solve a problem.* We had been measuring them because everyone else measured them. By banning future use of the technique (unless there was a *very* good reason) we stopped wasting time on useless hypotheses, forcing us to find the real causes much faster.

What is contact angle good for?

For most purposes, the human eye looking at a drop of liquid on a surface will tell you whether a surface will easily be wet by that liquid or not. Such a test is also useful if you need a quick check as to whether one liquid is similar to another or whether one surface is similar to another.

If your real concern is about the ability to clean the surface then Cleaning_Bleaching and Oxidation, Cleaning_Boundary removal, Cleaning_Enzymes, Cleaning_Solvents, Cleaning_Surfactants, and Cleaning_Temperature effects are likely to be much more important.

Getting into tight spaces

Flow of cleaning fluid into small spaces, pores, fibres, fabrics is governed by capillary forces, Absorbency_Porosity and Darcy. When we look at the core equation:

$$\frac{\delta z}{\delta t} = -\frac{D\sigma\cos(\theta)}{8\eta z}\frac{ST}{S+T}$$

the two parameters of relevance are σ, the surface tension and θ the contact angle. Adding surfactants to reduce the contact angle has the downside of reducing σ. And because the contact angle works via cosine term, the effect of changing contact angle is relatively small: cos(30) = 0.87 while cos(60) = 0.5, so a halving of the contact angle increases flow by ~70%. But if to achieve the 30° contact angle, the surface tension had to be reduced from 60 mN/m to 30 mN/m that’s a halving of the speed so the overall effect is a 13% reduction in flow.

The trade-off becomes positive for difficult-to wet surfaces with contact angles greater than 65° and by 80° the net benefit is a factor of 2.5.

If you are interested in predicting contact angles of liquids on polymers, using Dissolution_Hansen Solubility Parameters an app can be found on the Hansen-Solubility site: https://www.hansen-solubility.com/HSP-science/Surface_Energy.php.

Cleaning Enzymes

Links

Laundry Liquids

Enzymes are now so tough and reliable that they are regularly used in cleaning applications.

An enzyme for every occasion

Although we can use powerful general-purpose cleaners such as surfactants, caustics and bleaches, there are times where precision is required. To clean protein stains we can use proteases, fat stains via lipases, starches via amylases. Others, such as cellulases aren't for cleaning clothes but for keeping cotton fabrics looking nice by removing cellulose microfibres. We can even create hydrogen peroxide for stain removal using glucose oxidases if we add glucose to the cleaning fluid.

Getting enzymes tough enough to be used in real-world cleaning used to require finding extremophiles in places like the hot springs of Yellowstone. Now we can engineer them by synthesising promising DNA sequences that combine efficient active sites with 3D structures (checked using AlphaFold) that confer stability to the cleaning environment.

Many years ago my then PhD supervisor, an expert on enzymes, recounted a phone call he'd just had, to see if he could help enzymes work in washing machines at the normal (at that time) high temperatures. We laughed at the idea that enzymes would ever be used in laundry detergents. Yes, we were wrong. Prediction is hard!

Or we can add specific crosslinks to stop them unfolding, or immobilise them onto a surface to keep them away from harm.

Naturally aggressive

Because humans are made of proteins, fats and starches, the enzymes used in cleaning products are equally capable of causing damage to our skin or (e.g. from powder detergents) lungs. So a key requirement for these enzymes is reliable removal from the clothes during the rinse cycles. Those consumers who are nervous of biological detergents might like to reflect that rather than miss out on the proven efficacy of enzymes, they should at the same time decrease the amount of detergent per wash (most of us use far too much) and increase the number of rinse cycles.

Formulators working regularly with these enzyme formulations need check-ups at agreed intervals for enzyme-related problems.

Cleaning Solvents

Links

Lipstick, Mascara

> Although solvents are thought of as “bad”, sometimes water, even with surfactants and bleach, isn’t good enough. So how to select the least bad solvent?

The right combination

The solvent has to simultaneously meet a set of requirements:

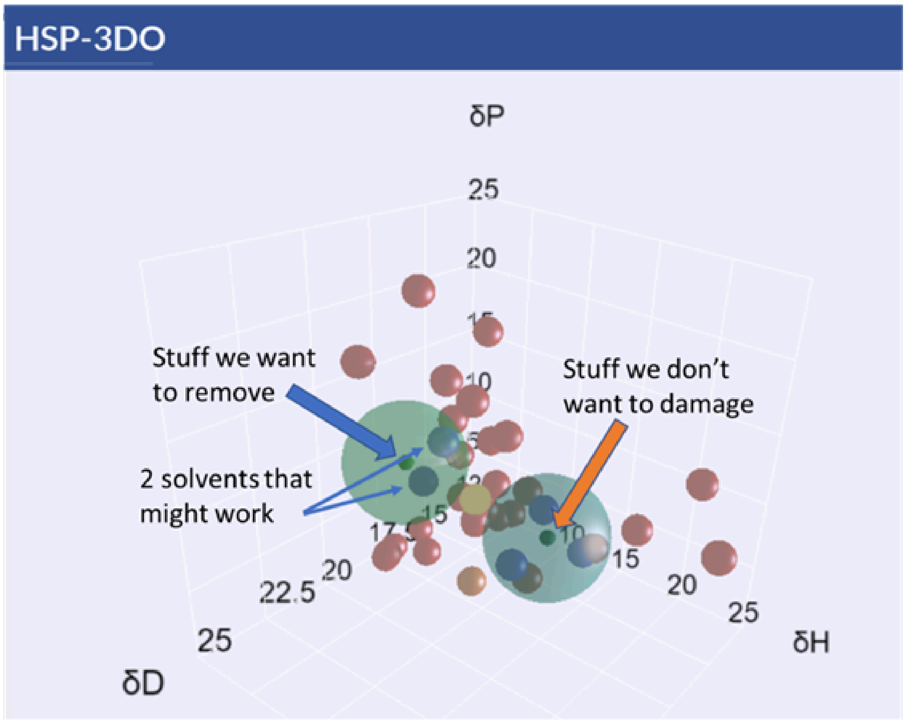
1. Efficacy as a solvent;
2. Correct evaporation (not too fast, not too slow);
3. Acceptable environmental impact.

Let’s look at each in turn.

Solubility match

You need to know where the thing you are trying to clean is in solubility space. You then need to find solvents that are sufficiently close in solubility space to be able to dissolve the problem material, while being sufficiently far from other things in the system (e.g. plastics) that you don’t want to be touched by the solvent.

The standard way to know all this is via Dissolution_Hansen Solubility Parameters. Because they show solubility space in three dimensions, δD (Dispersion), δP (Polar) and δH (Hydrogen bonding) we need a 3D viewer:



https://www.stevenabbott.co.uk/practical-solubility/HSP-3DO.php

As the image shows, we can quickly see where we want to be, i.e. close to the "bad" stuff, and where we don't want to be, close to the "good" stuff. Using the mouse to see all the right solvents (those near the bad stuff) we can get a list of solvents that will work well. We can then check for their other properties before selecting one.

Right evaporation

Solvents that evaporate too fast have two problems:

1. They disappear before they've cleaned;
2. They will have a low flashpoint that makes them a fire hazard.

Solvents that evaporate too slowly will have plenty of time to act but:

1. Might have time to attack stuff you don't want to damage;
2. They are difficult to remove, their odour lingers too long.

By knowing about Evaporation_Temperature and Antoine Coefficients we get a good idea for their vapour pressure at the desired cleaning temperature, and by combining that with an estimate of the local air velocity we can use Evaporation_Basics to estimate the evaporation time.

Can we use our solvent?

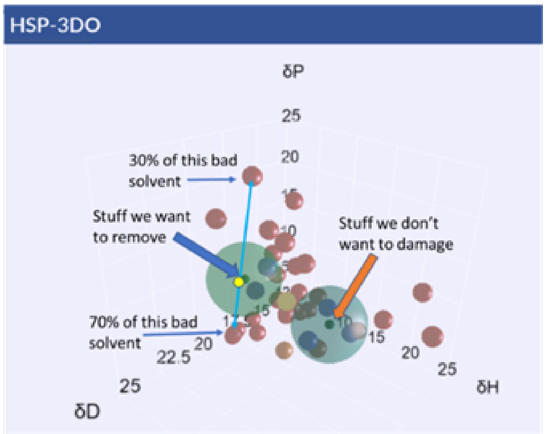
From the Safety Data Sheet we can find if our prospective solvents are harmful to us and to the environment and can then decide if the risks are manageable with proper PPE and containment.

Very often the 3-step analysis proves discouraging. There is no single solvent which can achieve all our goals. There are two things we can do about this.

1. Smart solvent blends
2. Microemulsions

Let's look at each in turn.

Smart solvent blends



If you take an X:Y blend of two solvents, the HSP of the blend is the weighted average of each of the 3 parameters. When you look at the diagram, you see that a 70:30 blend of two bad solvents creates a new solvent very near our solubility target:

In other words, two bad solvents can create a good solvent. This is a very powerful idea. Now we can choose solvents for their green qualities, for their odour, safety, evaporation rates … and combine them to achieve the desired solubility. The ability to work through issues like this in 3D space is a powerful feature of the HSP approach.

Microemulsions

One obvious way to deliver a solvent in a gentler fashion is as an emulsion in water. But there are at least two things wrong with this idea:

1. If the surfactant system isn't good enough, the emulsion can separate (Surfactancy_Emulsion Coalescence and Surfactancy_Emulsion Creaming and Flocculation which makes the system inconvenient and ineffective.
2. If the surfactant system is very good, the solvent is locked inside the surfactant shell, so gets little chance to clean.

The smart thing to do is to use Surfactancy_HLD to produce a microemulsion with a super-low Surfactancy_Interfacial Tension and Rigidity. A microemulsion is *thermodynamically* stable, so it is easy to mix (no high-shear required) and never separates. It is also characterised by a super-low interfacial tension which means that when the solvent/water hit the surface to be cleaned, the solvent is fully accessible.

A refinement of these microemulsion cleaners is to optimise the formulation to make them fully dilutable. This requires some understanding of microemulsion Surfactancy_Phase Diagrams, but if you do it right then you can ship a concentrated cleaner which can be diluted with water with no risk of phase separation.

Cleaning Surfactants

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, Lipstick, Mascara

> Lowering the surface tension of the water is the least important aspect of cleaning with surfactants. Providing feel-good foam is key to selling surfactant systems but is irrelevant to cleaning. Let’s sort out what is and isn’t important.

Removing oils and fats

Most forms of dirt don’t get significant cleaning benefits from surfactants – you can’t speed up dissolving jam, removing a stain or loosening some dried-on starch via surface active agents. As explained in Cleaning_Contact Angles, which in turn refers to Absorbency_Porosity and Darcy, changing the contact angle on the surface is rarely directly significant in terms of wetting, because we can force wetting through the action of a spray or cloth.

For getting into porous surfaces which have a contact angle greater than 65° with water, a surfactant will speed things up – by a factor of 2.5 for an 80° contact angle.

Most of the time, however, we need the surfactant to remove oils and fats. And now we have a real contradiction between consumer perception and scientific cleaning.

Standard dishwashing and personal care formulations are full of highly foaming, hydrophilic surfactants like SLES and CAPB. They give a good impression of the surfactants doing something.

But as we can find from Surfactancy_Roll-up and Eötvös Number we really need relatively hydrophobic surfactants that produce a low interfacial tension with the oils and fats, allowing them to roll up away from the surface and become engulfed as a microemulsion drop. Using Surfactancy_HLD, such surfactant packages are easy to formulate and perform their cleaning function much better. They are ideal for the industrial market where real cleaning is a requirement. But no one will buy them for the consumer market because they don’t behave as users have been trained to expect.

Surface cleaners

Surface cleaners are interesting because the users don’t want showy foam, they want good, instant cleaning (so the more hydrophobic surfactants should be used) plus a low surface tension and low advancing contact angle that helps spray drops to spread and coalesce plus a low receding contact angle to avoid de-wetting and the coffee-ring effects that might appear if isolated drops dry out.

This low receding contact angle trick provides a further benefit that’s real but not so obvious to the user. By adding some cationic polymers to complement the anionic surfactants, it is possible to create nm-scale thin films of “coacervates” – complex polymer/surfactant phases, see Surfactancy_Coacervation. These stick well to the surface and are humectants (Evaporation_Humectants) which attract atmospheric water and keep the surface hydrophilic. This means that oily dirt cannot adhere so well to the surface and is easily wiped away. To the user, the cleaner seems to provide “stay clean” or “EZ Clean” properties.

Specifying a surfactant (blend)

For the purposes of digitalisation of surfactant formulation we might expect to require some key parameters to allow us to make a rational choice of the surfactants to add.

The key parameters are “price” and “perception by the internet”. However green or scientifically excellent, your Sales team won’t let you use an expensive ingredient. And Marketing will be insistent that you are, say, “sulfate free” or “zero ethylene oxides” or “palm oil friendly”.

Whether a surfactant is bio-based or not should be irrelevant. What matters is its PCF (Product Carbon Footprint) or some similar LCA (Life Cycle Analysis) to tell us the real impact on the environment. It is far from obvious that bio-based surfactants are superior.

Scientific factors like CMC (Surfactancy_CMC and Langmuir) or Krafft temperature or cloud point (https://www.stevenabbott.co.uk/practical-surfactants/cloud-krafft.php) play almost no role. CMCs are ill-defined for industrial surfactants and play little role (if they did, we would not be using ionics with CMCs 1000x higher than non-ioincs). Although Krafft point *is* important (you can’t use the surfactant if it’s not soluble in the formulation), all the affordable surfactants easily meet the requirement – otherwise they wouldn’t be produced in the large quantities that make them affordable.

But everyone teaches CMC in surfactant classes

Lots of formulators are sent on surfactant science classes and come back with knowledge of worthless HLB and a detailed examination of CMC. The reason is that CMC is “sciencey” – it can be measured, and it has phrases like “Langmuir Isotherm”, so if you’re a teacher, what’s not to like?

Yet everyone who comes back from such classes finds that CMC is irrelevant to *their* formulation challenges so assume that they are just unlucky. It’s not unlucky, it’s bad teaching to focus on things that are “sciencey” rather than things that are usable.

Similarly, although there is a tenuous link between Critical Packing Parameter and surfactant phases (Surfactancy_CPP and phases) for pure surfactants, CPP is of little help for a real formulation.

If you *are* aiming for efficacy over Marketing (and this is the case for laundry detergents) then you need the Cc values of your surfactants to do the relevant calculations for Surfactancy_HLD, leading, as above, to efficient roll-up cleaning of oils.

Viscosity

Marketing have convinced users that viscous cleaning products are “rich” and “luxurious”. Apart from times when a clingy cleaning fluid is required, viscosity is unnecessary for cleaning. It costs nothing to create within standard formulations, using sodium chloride to create Thickeners_Wormlike Micelles. But those wanting to use alternative surfactants hit the problem that they don’t thicken in the same way, and alternative thickeners such as polymers have obvious downsides.

Foamability

There are few occasions (foamed bleaches for disinfecting places like dairies are an example) where foam adds anything to cleaning. All the surfactant molecules in the foam are, by definition, not on the surface doing something useful.

If you really want foaming for Marketing reasons then there are two broad choices.

1. Use classic, standard foaming tests which tend to confirm that standard systems like SLES/CAPB are the best. This is unfortunate if you have a surfactant system that is superior in other ways but foams poorly in the standard test.
2. Use the relatively new understanding that nearly all surfactants foam well if you create the foam properly: the right amount of trapped air, broken down into small bubbles via shear. A simple device such as a Japanese shaving net can produce awesome foam in formulations that fail the standard tests. A CAF (Compressed Air Foam, like a traditional shaving foam canister) also ends up with naturally small bubbles with a good, stable foam. With some imagination perhaps you can find a method that is simple, green, cost effective and appreciated by your customers.

Smart dosing

There is a simple way to reduce the carbon footprint of surfactant use: get the user to use half the amount, half the time. That's a 4x reduction in carbon footprint. "Half the time" can mean "Half the damage to the skin or scalp biome" (see SkinCare and Soaps and Washing). "Half the dose" means mostly education of the user that surfactants can only do so much to remove (usually) minor amounts of oils and fats, so throwing most of it down the drain is not a good use of their money or good for the environment. Helping them to dose the correct, small, amount via some smart dosing system would establish good habits.

The obvious downside that this means selling less product is partly true. But laundry detergents have seen a dramatic reduction in the amounts used (and the temperature required for them to work efficiently) and the industry is still alive.

The comparison with laundry is sobering. A modern, effective, laundry wash uses ~200ppm of surfactant. A typical shampoo or shower gel dose, to clean far less, uses 20000 ppm, i.e. 100x more. There is much room for improvement.

Cleaning Temperature effects

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, Lipstick, Mascara

> We all use Arrhenius as a rule of thumb for temperature effects. It's usually good enough.

The famous factor of 2 for 10°C

To get molecules to react (e.g. Cleaning_Bleaching and Oxidation) or to interact, requires thermal energy. Where R is the universal gas constant of 8.31 J/Kmole and T is the temperature in Kelvin then the Arrhenius equation tells us that a typical reaction proceeds at a rate, k, given by a constant, A, and an activation energy E_a:

$$k = Ae^{-\frac{E_a}{RT}}$$

At T=300K, RT ~2.5 kJ/mole. A typical value for E_a is 50 kJ/mole. Doing the arithmetic shows that k changes by a factor of 2 for each 10°C rise in temperature.

A factor of 2 doesn't sound very exciting. But a 30min wash at 40°C becomes a 1hr wash at 30°C and a 2hr wash at 20°C. Users won't adopt lower temperature washes with such time penalties. So formulations have to be smarter to win back those factors of 2. You can't change Arrhenius, so explore the FST for smarter ways to formulate around the factor of 2 issue.

Coating Dewetting Theory

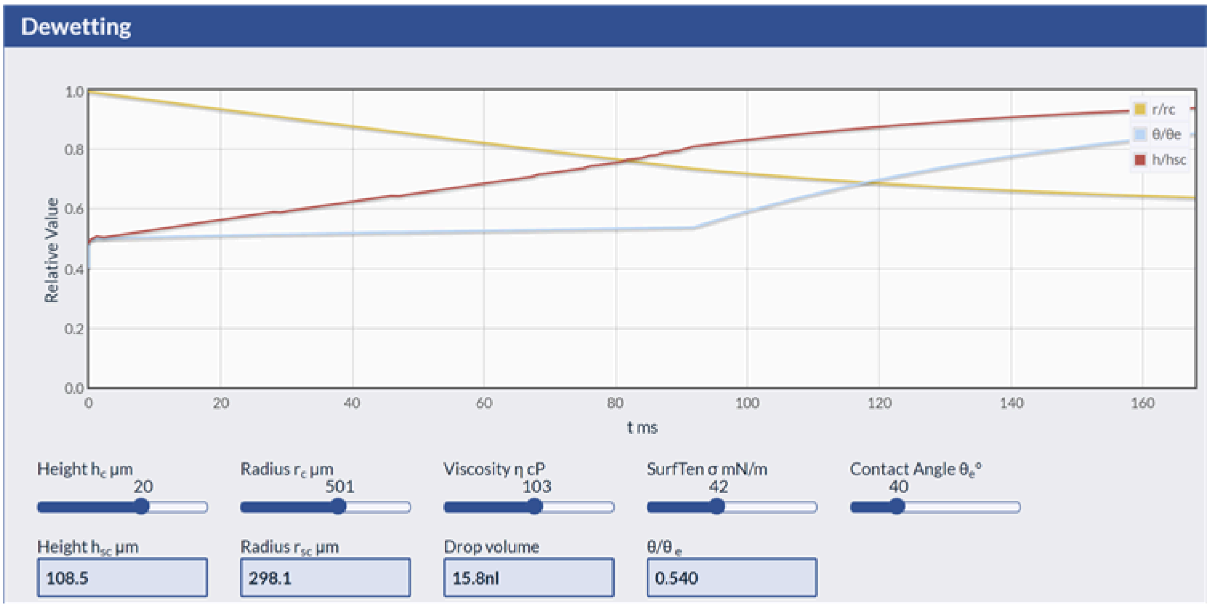
Links

[Solvent-Based Paint](), [Emulsion Paint]()

Everyone adds “wetting agents” to help coatings to, well, wet. But for everything other than inkjet, the processes in coating/printing machines fully wet the surface with forces orders of magnitude more powerful than surface tension.

The additives, therefore, aren’t helping with wetting. Instead they are helping to avoid dewetting.

Rapid dewetting should be normal



Height h_c µm 20

Radius r_c µm 501

Viscosity η cP 103

SurfTen σ mN/m 42

Contact Angle θ_e° 40

Height h_{sc} µm 108.5

Radius r_{sc} µm 298.1

Drop volume 15.8nl

θ/θ_e 0.540

https://www.stevenabbott.co.uk/practical-coatings/Dewetting.php

We’ve just printed a 20µm high dot, 500µm radius onto a surface with a 40° contact angle. The viscosity is a respectable 100cP. According to the app, within about 160 ms, the dot has dewetted to form a 100µm high drop with radius 300µm.

If you play with the settings you find that no practical printing technique, with any reasonable contact angle of ink, would ever print a dot that was a close approximation to the original size. Dewetting is fast and unavoidable.

The elegant paper (referenced in the app) describing and experimentally validating dewetting theory was based on large drops of silicone fluids on controlled surfaces. Having validated the app with the large drops, it was expected that because small printed drops of more normal fluids with smaller contact angles don’t dewet, it must be a feature of their size and contact angle. So it was a considerable surprise to find the theory predicting that printing was essentially impossible. Such a clear negative result tells you either that the theory is wrong (it's not) or that something else is going on. That led to the following section.

Measuring contact angles

Everyone measures contact angles (Cleaning_Contact Angles), because that's what everyone does. Typically you syringe a tiny drop onto the surface and follow the advancing contact angle with a camera and automatically fit the final shape to a curve that gives the contact angle. As everyone knows, the key to getting a reliable value is to carefully clean the surface. Any contamination would invalidate the reading.

But this is entirely pointless. When we coat/print we, by definition, "contaminate" the surface. As mentioned above, we do this under stresses far larger than surface tensions so there's no problem getting the shape/size that we want. Our equilibrium contact angle is meaningless because any retraction of the drop is over a surface that has been contaminated by the ink/coating. If that contamination gives a fully wetting substrate, e.g. because we have added a wetting agent that changes the surface, then the drop won't dewet back to whatever the equilibrium contact angle is supposed to be.

If we routinely measured *receding* contact angles then we'd choose a wetting agent that rapidly contaminated the surface under the drop and gave a small, maybe near 0 receding contact angle. This would mean that the drop is effectively pinned at its current position – it doesn't want to advance because its *advancing* contact angle would be larger, and it can't contract because its *receding* angle is small, meaning that it would return to the pinning spot.

The reason we measure the irrelevant equilibrium contact angle is because everyone else does, and it's easy to do. Receding contact angles are notoriously tricky and unreliable because, well, the surface has been contaminated.

Note that the wetting agent is *not* achieving its desirable effects by changing the surface tension of the coating fluid, i.e. it's interface with the air. It works by migrating to the fluid/solid surface, changing the solid/air and solid/fluid surface energy components that each figure in the Young contact angle equation.

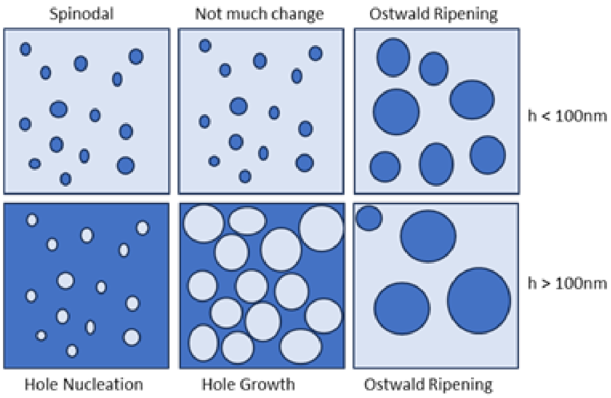
So how *should* you solve dewetting problems? The wrong approach is to choose an additive because marketing literature tells you that XYZ is a superb dewetting additive. The right approach is to imagine a molecule with a head that will instantly lock onto your solid surface and a tail that will strongly attract your solvent – and which will partition rapidly to the solid/liquid interface. You are the expert on your surface and solution. Once you know what *your* surface/solvent combination needs, finding something with a good chance of success becomes a rational process.

Pinholes

We often add wetting agents to avoid pinholes. The theory of avoiding pinholes is described in Coating_Pinholing Theory, but key to this is that the thickness of the coating and the diameter of the original pinhole (very often from a bubble) are usually more important than the contact angle used in the theory.

There's a fudge in that theory. For the pinhole to close, the advancing (or, as an approximation, equilibrium) contact angle is important. For it to open wider the receding contact angle is the one to use. As the advancing contact angle is the more pessimistic value, and because the theory emphasises that thin coatings need greater attention to cleanliness and bubble reduction, it's good to leave the focus on the advancing angle.

Thin film instabilities



If you have a layer in the region of 100nm you can have spontaneous dewetting. Below 100nm (top 3 images) you tend to get "spinodal" dewetting where you get sudden, isolated droplets. Above 100nm (lower 3 images) you tend to get "nucleation and growth" where holes appear at some defects on the surface. The holes then grow into some sort of patterned structure. If you can provide some specific defects in a specific pattern then you can make a patterned nanostructure. In both cases, the individual blobs can undergo Ostwald ripening, so that smaller drops get smaller, larger get larger, till you have just a few large drops.

The thermodynamic stability/instability depends on surface energies, Hamaker constants and charges; equations are available but we tend not to know the values to put into apps. The kinetic stability depends on the mobility of the material at the given temperature, humidity or solvent environment. So whether you get your desired stability or desired nanostructure is not easy to predict.

Those two paragraphs give a taste of a large, controversial area that I don't properly understand. Obviously the effects are somehow linked to Coating_Pinholing Theory but the core theory of dewetting says that any film of thickness less than the critical dewetting thickness h_c *(depending on surface tension* γ*, density* ρ *and gravity,* g*) will spontaneously break into drops. The problem is that this thickness is typically in the mm range, not the nm range shown here and the μm range typical of pinholes in coatings. I've not been able to make sense of this.*

$$h_c = 2\sqrt{\left(\frac{\gamma}{\rho g}\right)}\sin\left(\frac{\theta}{2}\right)$$

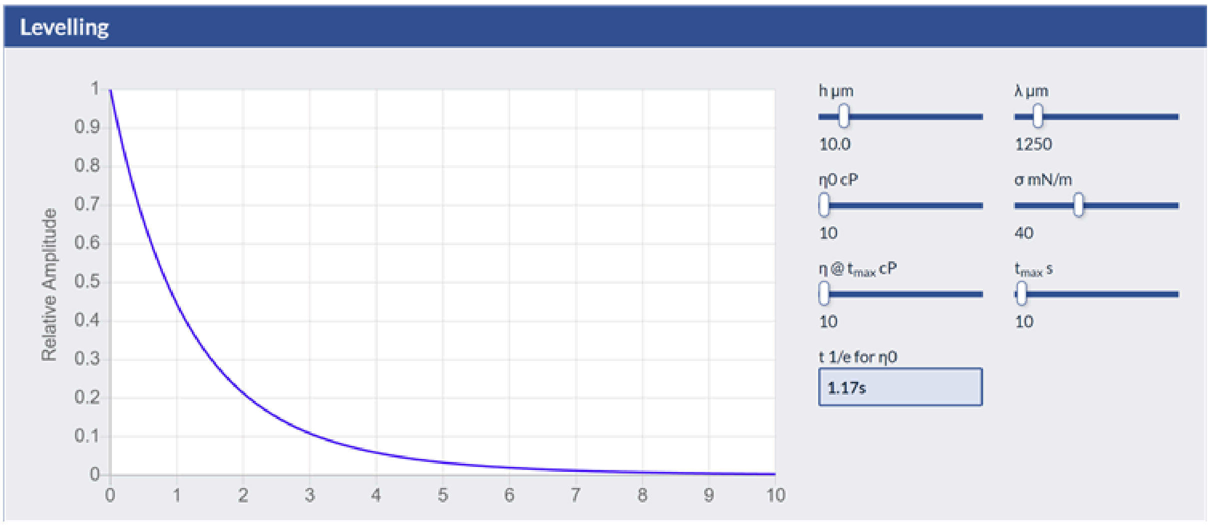
Coating Levelling Theory

Links

Solvent-Based Paint, Emulsion Paint

If you have some coating lines from a brush or roller, you can measure that they have an average spacing of λ. The coating has a thickness h, a viscosity η and a surface tension σ.

You can easily work out the time for the lines to disappear. Although people focus on the viscosity and surface tension, they only have a *linear* effect on levelling. Much more important are the h^3 and λ^4 dependencies.



h µm: 10.0

λ µm: 1250

η0 cP: 10

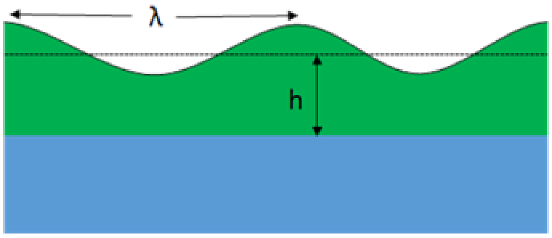
σ mN/m: 40

η @ t_{max} cP: 10

t_{max} s: 10

t 1/e for η0: 1.17s

https://www.stevenabbott.co.uk/practical-coatings/levelling.php



A coating will often contain some sort of line defect visible during creation (coating head, brush, roller) but invisible once dry. As the wet coatweight is reduced the same wet defect now shows up in the dry coating. Why? Because the defect has not had time to level out. The reasons for this are not obvious until Orchard levelling theory is invoked.

The levelling time for a defect of wavelength λ in a coating of thickness h with viscosity η and surface tension σ to level by 1/e, t_e, is given by:

$$t_e = \frac{3\eta\left(\frac{\lambda}{2\pi}\right)^4}{\sigma h^3}$$

Or in terms of height at time t:

$$h_t = h_0 \exp\left(\frac{t}{t_e}\right)$$

It immediately becomes clear that relatively small decreases in thickness can give a large increase in levelling time because of the h^3 term. On the positive side, increasing the sharpness (decreasing λ) of the defect greatly decreases levelling time. That's why those who buy cheap paint brushes with coarse bristles (large λ) have to put on more paint (larger h) than those who buy better brushes with finer bristles.

The app includes corrections for the fact that viscosity of a drying formulation increases with time.

Useless "levelling agents"

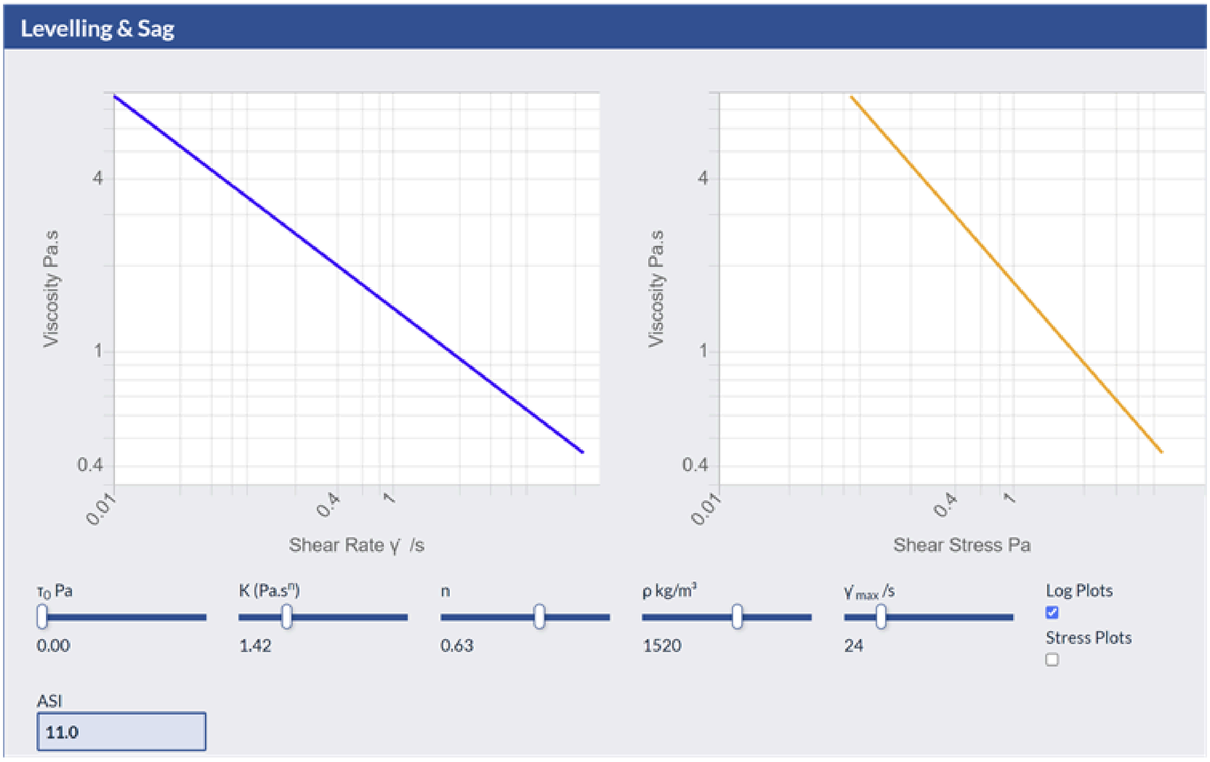
Because some "levelling agents" work for some types of defects, it is assumed that you can solve these linear defects with such agents. At best they lower the surface tension which, from the equation, *increases* the levelling time. The confusion arises because these agents address defects such as orange peel not by lowering the surface tension but by swamping the surface with a uniform surface tension that obliterates the real problem which is due to Evaporation_Marangoni effects caused by self-ordering surface tension gradients.

The origins of lines

Although the brush has *some* effect on line spacing, many of the lines with brushes and with rollers are due to the **ribbing** defect that has a complex dependence on speed, viscosity and gap. Maybe we'll have an app about it sometime.

Rheology of levelling & sag

It's well known that if you increase the viscosity of a paint, the tendency to sag goes down (you have a higher ASI, Anti-Sag Index) but the levelling performance gets worse. There is no simple "viscosity" metric than can resolve this contradiction. It needs more sophisticated measures such a Power Law fluid fit as discussed in the Levelling & Sag app:



τ_0 Pa 0.00

K (Pa.s^n) 1.42

n 0.63

ρ kg/m³ 1520

$\dot{\gamma}_{max}$ /s 24

Log Plots ☑

Stress Plots ☐

ASI 11.0

https://www.stevenabbott.co.uk/practical-rheology/Levelling-&-Sag.php

You can also use G' and G" values as discussed in Flow_G' and G".

Coating Pinholing Theory

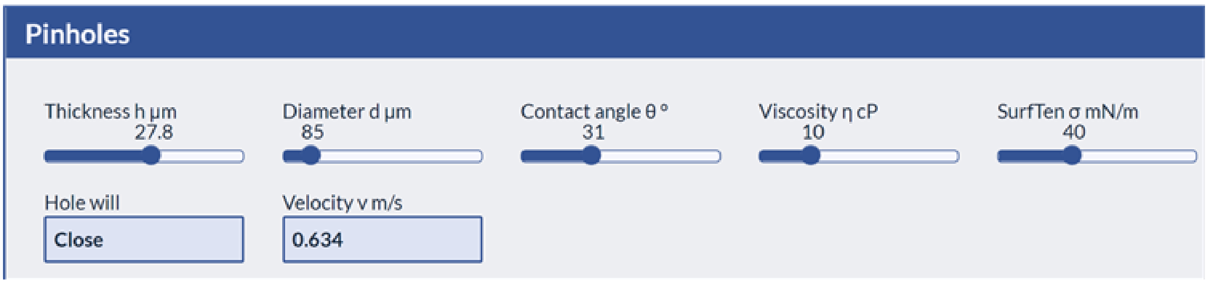
Links

Solvent-Based Paint, Emulsion Paint

Everyone worries about surface tension σ and contact angle θ when they see pinholes. But there are two key factors that are less talked about:

1. The diameter of the initial pinhole, *d*, and, therefore, what caused it
2. The thickness, *h*, of the coating

You will see from the app that large pinholes are inevitable if *d* is large and *h* is small. So, focus on small *d*'s and thicker *h*'s if you want to reduce pinholes. Yes, a lower θ helps, but not as much as controlling *d* and *h*.



Pinholes

Thickness h μm 27.8

Diameter d μm 85

Contact angle θ ° 31

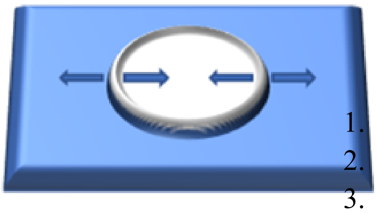
Viscosity η cP 10

SurfTen σ mN/m 40

Hole will: Close

Velocity v m/s: 0.634

https://www.stevenabbott.co.uk/practical-coatings/pinholes.php



A coating will often contain a little pinhole. Whether it spontaneously grows (outer arrows) or self-heals (inner arrows) depends on three factors:

1. contact angle
2. coating thickness
3. hole diameter.

While everyone focusses on contact angle, there is little understanding that you can transform a pinhole-free coating to one full of holes merely by reducing the coating thickness.

The theory shows why you must keep all 3 factors in mind.

A pinhole of diameter d in a coating of thickness h and contact angle θ will grow if:

$$\frac{h}{d} < 2(1 - cos\theta)$$

The speed at which it grows depends also on the viscosity η and surface tension σ is given by:

$$v = \frac{\sigma\theta^3}{\eta}$$

We can now see why thinner coatings are more likely to show pinholes. If a pinhole-free coating is halved in thickness then $\frac{h}{d}$ can fall below the critical value. The only cure is to have a lower contact angle (better wetting) or a cleaner, bubble-free coating with less initial pinholes. As the (true) story below shows, eliminating the source of defects fixes the pinhole problem. In that case, the problem was an unexpected source of dirt, in many other cases (a fact confirmed by talking to experts around the world) the cause is bubbles, often from an unexpected source – see Dissolution_Gases.

I was working on a coating line where senior managers were paranoid about the use of WD40 near the machine as this was guaranteed to create “repellencies”, defects caused by oil droplets During an outbreak of these defects, doubts about them being caused by oil droplets were sufficient for me to get permission for a tiny “psst” of WD40 to see how long such repellencies would last. Unfortunately my tiny “psst” was a big blast covering about 1 m^2 of coating. But instead of hours waiting for the repellencies to go away, the coating fixed itself in a few metres, and the oil repellency hypothesis was killed. It turned out that the continuing outbreak of defects was due to people like me going into the coating environment to find the cause of the defects. By banning entry for a short time, the problem went away.

The original source of dirt? I’d insisted on cleaning the machine before the run. This kicked up a lot of dirt which hadn’t had a chance to settle.

We all make mistakes. The key is to learn from them.

The right contact angle

There’s a fudge in the theory. For the pinhole to close, the advancing (or, as an approximation, equilibrium) contact angle is important. For it to open wider the receding contact angle is the one to use. As the advancing contact angle is the more pessimistic value, and because the theory emphasises that thin coatings need greater attention to cleanliness and bubble reduction, it’s good to leave the focus on the advancing angle.

Diffusion Barrier Properties

Links

Fragrances, Lipstick, Mascara

Everyone wants barrier films for, say, oxygen and water. They must be low cost, flexible, impermeable. The ideal would be a single polymer, or, at least, something that can be made in a single pass. And in general they should be transparent, colourless and fully recyclable. Getting a good barrier is easy, getting one that meets the contradictory demands is hard.

Barrier properties

If you find yourself confused by WVTR (MVTR) and OTR (Water Vapour/Moisture and Oxygen Transmission Rates) then you are like everyone else. The WVTR/OTR world is a disaster zone of bad units and bad measurements. This is all discussed separately: Diffusion_Permeation OTR and WVTR.

For the rest of this chapter we assume that you want low values for both, plus low values for, say Fragrances and flavour molecules.

Aluminium

We aren't going to run out of aluminium, it's highly recyclable and as a thin foil it is a perfect barrier. As a vacuum coated film on polymer its barrier properties depend mostly on the cleanliness of the coating line. This is because pinholes have an outsized effect on the barrier properties. Take some time to explore the subtleties of the 3 models for pinholes, https://www.stevenabbott.co.uk/practical-coatings/3models.php, but if we take the Coverage model (which is a bit optimistic) then the sad fact is that 1% pinholes means that you improve the barrier only by a factor of 100 compared to the many orders of magnitude improvement from a perfect film.

One of the quirks of vacuum coating is that you can't fix pinholes by just delivering more metal. More metal mostly increases cost, with a surprisingly small effect on % pinholes – it's all to do with how metal atoms flow on the fresh surface.

A clean roll of polymer, an in-situ cleaner to remove the last remnants of dust, and a good protocol for cleaning the vacuum chamber between runs are all necessary to get good barriers.

The core problems with Al coated barriers is that you can't see the contents and these "metal" packages have been demonized by the public as not being recyclable.

AlOx/SiOx

In theory, and in lab samples, these are ideal barriers. A few nm of some oxide (it doesn't have to be perfect, hence the "x" in the name) of Al or Si vacuum coated on clean film and in a clean machine (see the previous section) provide a transparent, excellent barrier to both water and oxygen. Because the layer is thin, it is sufficiently flexible to be wound into a roll and processed into a final product.

The problems are obvious – these oxide barriers are brittle and easily scratched. So the barrier properties of commercial rolls of film are surprisingly modest and some glitch during a production run of the end product can introduce barrier-breaking stresses that would not be detectable; so products with poor barriers can be (and have been) shipped.

A factor that makes these barriers especially prone to cracking under modest stresses is that the strains "focus" on the layer via Dundurs α parameter in fracture mechanics. An app that shows the significance of this effect is at https://www.stevenabbott.co.uk/practical-coatings/Interlayer.php. It shows that the problems arise due to a big mismatch between the modulus of the substrate (say 1 GPa) and that of the oxide coating (say 50 GPa). If you can do something to increase the substrate modulus (adding a thin layer of 2 GPa) that does a lot to decrease the severity of the Dundurs α effect.

Biopolymers

If you take almost any natural polymer, starch or protein, you can produce a film with surprisingly good oxygen barrier properties, for reasons discussed below.

The obvious problem is that the water barrier properties are generally poor, and highly %RH dependent. At high %RH even the oxygen barrier properties get worse.

Various chemical tricks can reduce the %RH dependence, but then the polymer is becoming less natural.

> It isn't good enough to have good intentions and a friendly biopolymer. Being green means not wasting precious resources, but the predictable failure of most biopolymer barriers has meant much wasted resource.

PE, PP

The standard food packaging films based on polyethylene or polypropylene (PE/PP, for our purposes indistinguishable) are wonderful water and oxygen barriers. But that's only because they are multi-layered with one or more thin layers of EVOH created by co-extrusion.

The fact that they are not pure polymers is very annoying to those who demand that food packaging should be easily recycled. So why do evil packaging film suppliers insist on contaminating the recycling stream? Because pure PE/PP films are useless oxygen barriers. [Some new grades are more oriented/crystalline with somewhat better barrier properties, but it remains a challenge.]

The reason is simple. HSP Dissolution_Hansen Solubility Parameters are used throughout the FST because they deliver so much power with so little effort. The key is that if the Distance between the HSP of, say, a polymer and a solute is large then they are incompatible, a small Distance means they are compatible.

It is obvious that the HSP Distance between water and PE/PP is large. So water does not partition into the polymer. And, because (Diffusion_Basic Diffusion) the rate of diffusion depends on the concentration gradient, a low level of water in the surface of the polymer means a low rate of diffusion.

It happens that the HSP of O2 is relatively close to the HSP of PE/PP, so O2 partitions readily into the polymer, there is a large concentration gradient, so diffusion is rapid.

Now let's put a thin layer of EVOH into the middle of the PE/PP film. EVOH is a poor water barrier (the HSP Distance to water is obviously much lower) but its Distance from O2 is large, so there is only a low concentration that can partition into the EVOH, so a small diffusion. It is remarkable that even 1µm of EVOH provides a near perfect barrier to O2 inside a co-extruded film. In practice the thickness is larger and there is more than one layer. This is because it is very hard to extrude a pinhole-free 1µm EVOH barrier, so going to 2µm and having multiple layers is a cost penalty worth paying in order to obtain the astonishingly good barrier properties from apparently cheap and simple plastic films.

The reason the biopolymers have good O2 barriers is the same – their HSP Distance is large. And, of course, their distance to H2O is smaller, so they are poor water barriers.

It's well known that PE/PP film is a poor barrier for common fragrance molecules like limonene. Again, the HSP Distance is small. The EVOH multilayer packaging films provide, not surprisingly, excellent barrier properties for a wide range of fragrances and flavours.

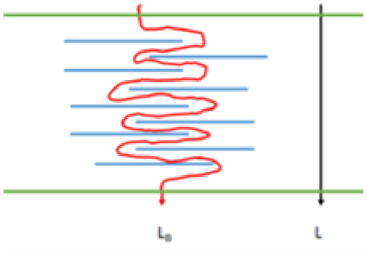
PET

The discussion of PE/PP implies that any polymer will be permeable to something if the HSP Distance is small. The exception would be a polymer that has a very low value for the diffusion coefficient, D. Even if there is reasonable partition to set up a concentration gradient, if D is small you won't get too much permeation.

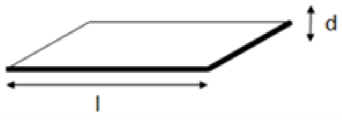
PET is used because, when oriented, (both as PET film and as PET bottles) the polymer is relatively dense and crystalline while remaining crystal clear. This high density means that there is little "free volume" for molecules to diffuse through. While PE/PP D values are in the 10^{-9} cm^2/s range, for PET we are in the 10^{-13} range. Four orders of magnitude is a lot. And that's why we use a lot of PET.

The thickness of barrier is only linear in its effect on permeation. Usually you cannot solve a diffusion problem by increasing thickness. However, when (for strength, not barrier, reasons) we go to a PET bottle with a thickness of 500µm compared to a packaging film of 50µm that's an extra order of magnitude. So even a relatively small, mobile molecule like CO2 has a five orders of magnitude barrier, allowing us to store fizzy drinks for a long time.

Clay barriers



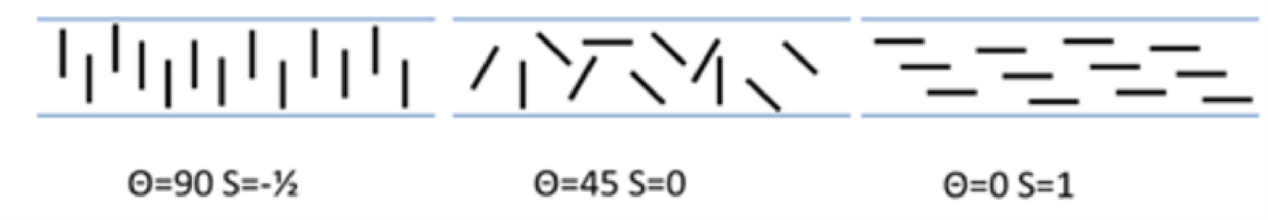
Everyone loves the idea of taking a cheap, natural clay and creating a tortuous barrier layer. For a thickness L, a molecule has to travel a distance L_B to cross the barrier. The more clay we can pack in, the more tortuous the path and the better the barrier. The problem is that clays, when well-dispersed, add considerable viscosity to the dispersion so there is always a trade-off between the amount you would *like* to add and the amount you *can* add.



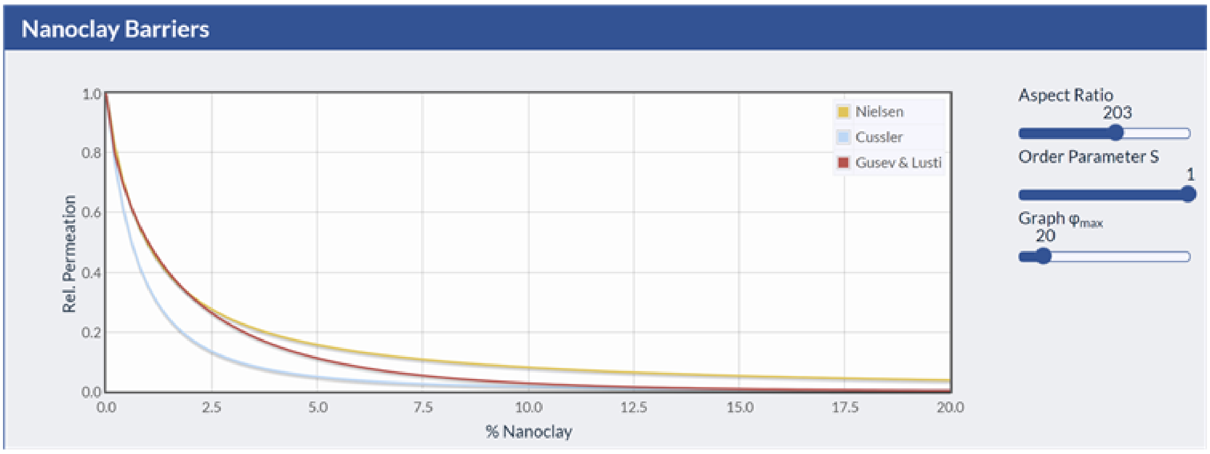
To create such a barrier we must have a high aspect ratio, l/d, for the particle, and clays are ideal for that *provided you can separate them into individual layers*. That is harder than it sounds and there are many failed attempts to provide good clay barriers.

While writing a chapter on the use of HSP for understanding so-called green polymer PLA, it was heartbreaking to see how many PhDs had been wasted on trying to add clays to the polymer. A few seconds thought showed that the dispersants they were using were guaranteed to fail – which they did. The one commercially-available dispersant that, from HSP, would give great dispersion happened to react with the PLA during processing. A few minutes thought could have saved years of failed research.

Finally, you need the order parameter, S to be close to 1 to maximize the barrier effect. It is not intuitive that the order parameter go from -½ to 1, but that is the scale that everyone uses:



Now we can calculate the relative permeation compared to no nanoclay:



Aspect Ratio 203

Order Parameter S 1

Graph φ_{max} 20

https://www.stevenabbott.co.uk/practical-coatings/nanoclays.php

There are some disagreements between models, but you can see that by, say, 15% high aspect-ratio clay, nicely oriented, you have a great barrier. That 15% is volume fraction – depending on relative densities its weight fraction is higher.

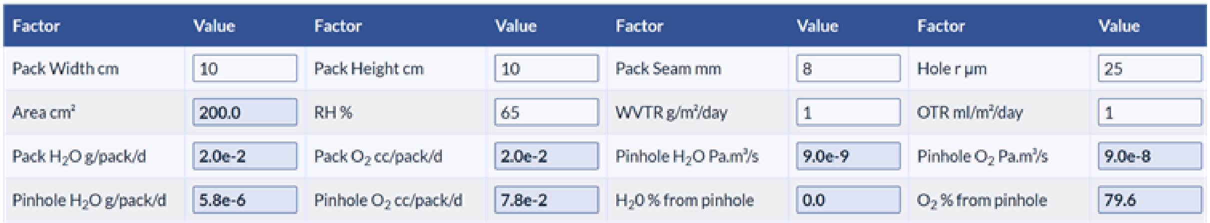
If these barriers are so good, why don't we see more of them.

- Because they are based on nanoclays, there are issues of nanosafety. Many would argue that these are non-issues, but the regulatory infrastructure defines them as nano, so you need to apply nano rules.
- Because clays (and therefore the barriers) are often yellow.
- Because dispersing the clays down to individual plates, required for a high aspect-ratio, is very hard. And such dispersions can be highly viscous.
- Orientation during production might not be perfect.
- Because these tortuous barriers can be deceptive. For short-term measurements, the barriers seem amazing. But once the water or oxygen has traced the tortuous path, you "only" have the increase

in path length. This might be a factor of 100, which is great, but isn't the factor of 1000s that they appear to be when you measure them before breakthrough via the tortuosity.

Pinholes in seams

If, see Adhesive, we use a heat seal to create a seam in a package, then much of our barrier efforts will be wasted if, through carelessness or dirt, a pinhole goes through the seam, allowing oxygen and water a simpler way to get inside. The Leakage app shows that even a 25 µm pinhole can make your OTR irrelevant:



Factor	Value	Factor	Value	Factor	Value	Factor	Value
Pack Width cm	10	Pack Height cm	10	Pack Seam mm	8	Hole r µm	25
Area cm²	200.0	RH %	65	WVTR g/m²/day	1	OTR ml/m²/day	1
Pack H_2O g/pack/d	2.0e-2	Pack O_2 cc/pack/d	2.0e-2	Pinhole H_2O Pa.m³/s	9.0e-9	Pinhole O_2 Pa.m³/s	9.0e-8
Pinhole H_2O g/pack/d	5.8e-6	Pinhole O_2 cc/pack/d	7.8e-2	H_20 % from pinhole	0.0	O_2 % from pinhole	79.6

https://www.stevenabbott.co.uk/practical-coatings/Leakage.php

The theory is simple, the flux, Q, through the pinhole depends on the diffusion coefficient in air, D, the pressure difference ΔP across the pinhole (0.21 atm for oxygen), the length of the seam, l, through which the air passes and the radius of the pinhole, r.

$$Q = D\Delta P\frac{\pi r^2}{l}$$

Q is in units of Pa.m²/s and, as so often with barriers, the difficulty is making sure you get all the units right in order to compare the various fluxes. In this case, the 25 µm pinhole has no effect on the WVTR but is almost 4x larger than the OTR through the packaging film.

Diffusion Basic Diffusion

Links

SkinCare, Surface Cleaning, Lipstick, Mascara, Water-based Adhesive, Adhesive, Microencapsulation

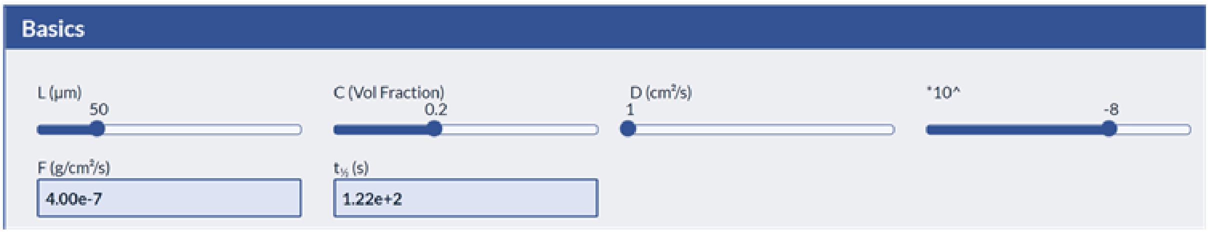
You can calculate how much of a molecule diffuses into (or out of) a coating by knowing three things:

1. The thickness of the coating;
2. The maximum solubility of the molecule in the coating;
3. The diffusion coefficient of the molecule inside the coating.

#1 is easy to measure, #2 needs a bit of experimentation (or can be predicted via Dissolution_Hansen Solubility Parameters) and #3 can be estimated via its molecular weight if there is a known value for a reference molecule.

With a bit of training it becomes easy to use the relevant apps to understand and calculate diffusion.

The basics of diffusion are easy to understand.



L (µm) 50

C (Vol Fraction) 0.2

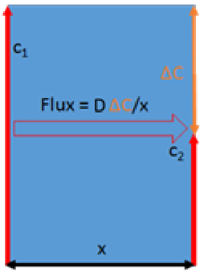
D (cm²/s) 1

*10^ -8

F (g/cm²/s) 4.00e-7

$t_{½}$ (s) 1.22e+2

https://www.stevenabbott.co.uk/practical-solubility/diff-basics.php



The formula to calculate diffusion is simplicity itself and is called Fick's First Law. It says that the amount of material per unit area that will flow in a given time (the Flux, F) over a given distance x depends on a constant, the Diffusion Coefficient D and the change of concentration ΔC over that distance (the concentration gradient), i.e. $\frac{\Delta C}{x}$, or $\frac{\delta C}{\delta x}$ if you want to be fancy. Therefore:

$$F = D\frac{\Delta C}{x} = D\frac{\delta C}{\delta x}$$

Note that purists add a negative sign on the right-hand side. Fick's Second Law of Diffusion is not quite so intuitive. It tells us that the change of concentration with time is the change of flux with distance. In other words:

$$\frac{\delta C}{\delta t} = D\frac{\delta F}{\delta x} = D\frac{\delta^2 C}{\delta^2 x}$$

Although these two laws are simple, they are differential equations (for example, the concentration gradient changes with the flux which depends on the concentration gradient) and the solutions in most real cases (where D also depends on concentration) are not straightforward. Therefore numerical simulations are required to capture what really happens. Fortunately, we can explore some simpler cases which happen to be very relevant in practical terms. Once we're happy with the basics we can then look at more general cases, with the app doing all the hard work of carrying out the numerical simulations.

Steady state diffusion across a barrier

We already know the answer to how much chemical will diffuse across a barrier of thickness L if the concentration is C at one side and 0 at the other. We have a concentration gradient (C-0)/L, a diffusion coefficient D, so the flux F is:

$$F = \frac{DC}{L}$$

The real problem is not the formula, but getting inputs and outputs in the right units. D is usually expressed in terms of cm²/s and has values of 10^{-5} for self diffusion in liquids, thru 10^{-7} for diffusion through rubbery polymers up to 10^{-12} for typical polymers and 10^{-15} for very close-packed polymers. *If you have a D value in m²/s multiply it by 10^4 or, the same thing, subtract 4 from the exponent*. 10^{-13} m²/s is the same as 10^{-9} cm²/s.

> From time to time I try to convince myself to switch to m²/s, but we live in the real world and cm²/s are so common and comfortable that they are the default in the relevant apps.

Concentrations can then be expressed in the familiar terms of g/cm³. Thicknesses are typically in µm but have to be converted to cm for the calculation. Fluxes are, therefore, in g/cm²/s. A complication is that it's often easier to think of solubility in terms of volume fraction. Assuming, for simplicity, that all densities are 1g/cc, a volume fraction of 1 equals a concentration of 1g/cc. This is convenient. It is intuitive to think of a polymer containing a 0.1 fraction of a solvent (i.e. 10%) and this simply becomes 0.1g/cc for the calculations. Adjusting for relative densities is a trivial task.

Half Time

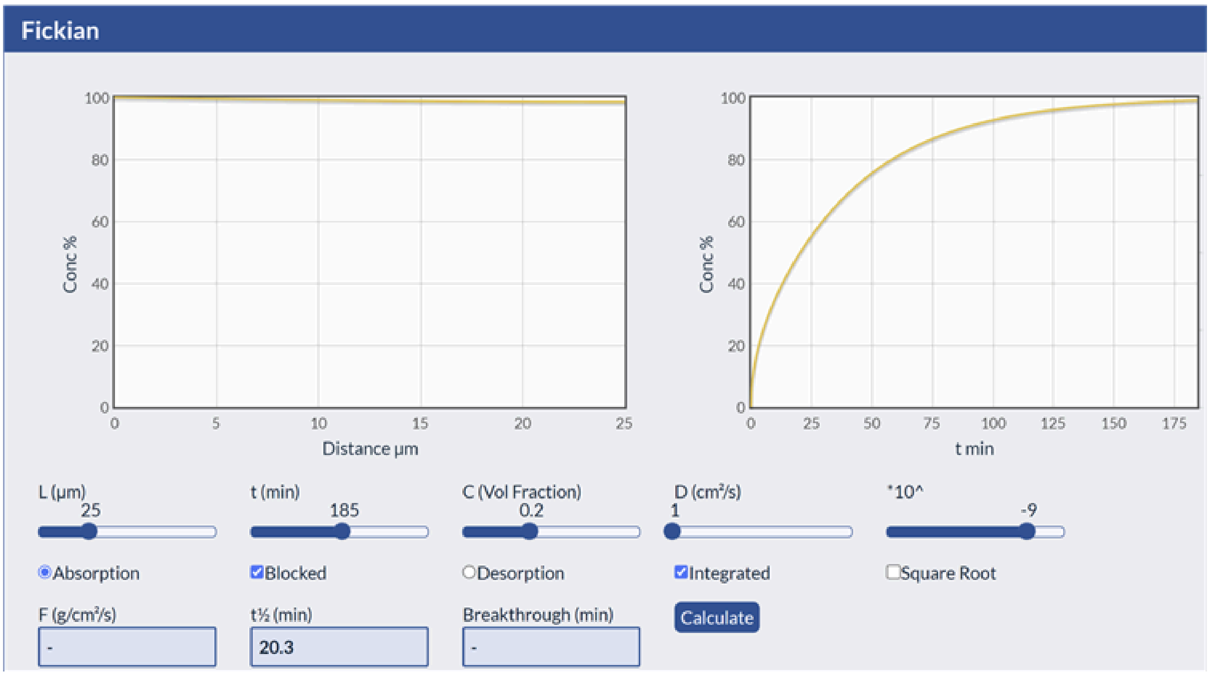
In addition to the Flux, the time, $t_{½}$, to reach half the equilibrium amount (either by absorption from empty or desorption from equilibrium) is given by $t_{½} = 0.049\frac{L^2}{D}$. This equation follows from the fact that absorption or desorption goes as the square root of time. It is often used to calculate the diffusion coefficient from experimental half-time measurements.

Now we can see what happens as we change our key parameters. *Note that because of the limitations of app interfaces the diffusion coefficient is entered as a value between 1 and 9.9 and an exponent from 10^{-5} to 10^{-15}.*

Because of the plane of symmetry in the middle of a free film, $L = \frac{Thickness}{2}$ when modeling absorption or desorption where mass is transported through both surfaces. When modelling permeation, which is one-sided exposure, then *L=Thickness*.

Real data

Although it *can* be difficult to get diffusion data, there are many cases where it's both easy and useful. A typical example is where you need to know the equilibrium concentration (which controls ΔC) *and* the diffusion coefficient. Let's look at a typical absorption experiment using the Fickian diffusion app:



L (µm) 25

t (min) 185

C (Vol Fraction) 0.2

D (cm²/s) 1

*10^ -9

● Absorption

☑ Blocked

○ Desorption

☑ Integrated

☐ Square Root

F (g/cm²/s) -

t½ (min) 20.3

Breakthrough (min) -

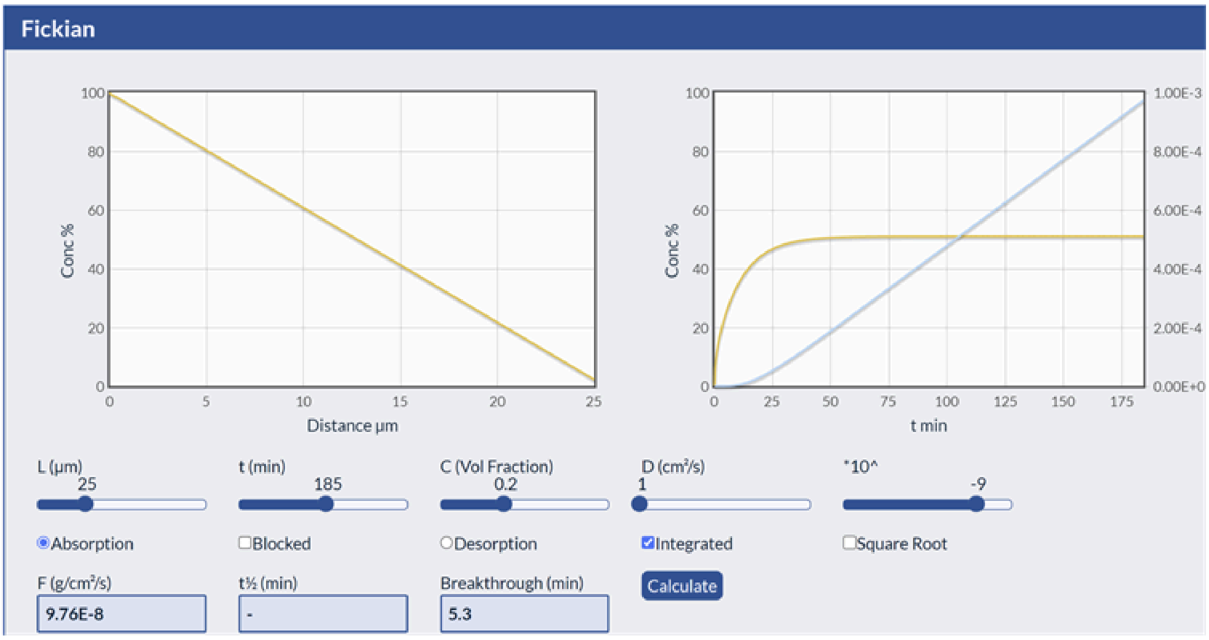
https://www.stevenabbott.co.uk/practical-solubility/diff-fickian.php

A 50 µm film is weighed then dipped into a solvent. After, say, 5 min, it is pulled out, quickly dried and weighed, before being put in and removed after 10 min (total), then 20, 40 up to 180min. From the increases in weight we can work out a time versus % max concentration curve. In the app we enter L as 25 µm (Thickness/2 because solvent comes from both sides) and by sliding sliders we find a good match to the experimental curve if D is 1.10^{-9} cm²/s. From the asymptotic weight gain we also know the equilibrium solubility. With one experiment we have fully characterised the system.

Of course, it's not as easy as that. What sized sample is used, what thickness chosen, how do you reliably wipe dry for each measurement, how do you weigh accurately? Because I've done this many times, in different labs with different abilities to weigh accurately, and with different polymers covering a range of D values, I can say that it's not as difficult as it sounds. Within 1hr you can easily work out many of the issues and if you are lucky you can have some good values in less than a day. If you have a difficult polymer, it might need some time to get the right sample thickness and some overnight measurements to get the equilibrium value.

If you cannot fit your diffusion data to the standard Fickian curve, the chances are that you have the interesting case of concentration-dependent diffusion. This is discussed in its own chapter, Diffusion_Concentration-Dependent Diffusion.

Alternative measurement techniques are available. For example you can attach your sample to an FTIR ATR crystal and wait to see how long it takes for solvent on the other side to become detected and link that time to the "breakthrough" time calculated in the non-"blocked" mode in the app, in this case 5.3 min:



L (µm) 25

t (min) 185

C (Vol Fraction) 0.2

D (cm²/s) 1

*10^ -9

● Absorption

☐ Blocked

○ Desorption

☑ Integrated

☐ Square Root

F (g/cm²/s) 9.76E-8

t½ (min) -

Breakthrough (min) 5.3

If you want a good enough estimate of D for a molecule through a common polymer then you can use the Diffusion Coefficients app at https://www.stevenabbott.co.uk/practical-solubility/diff-d-values.php. The equations used have been developed for the food and pharma industries where they realise it is unrealistic for everyone to measure D values for all their molecules.

Diffusion Concentration-Dependent Diffusion

Links

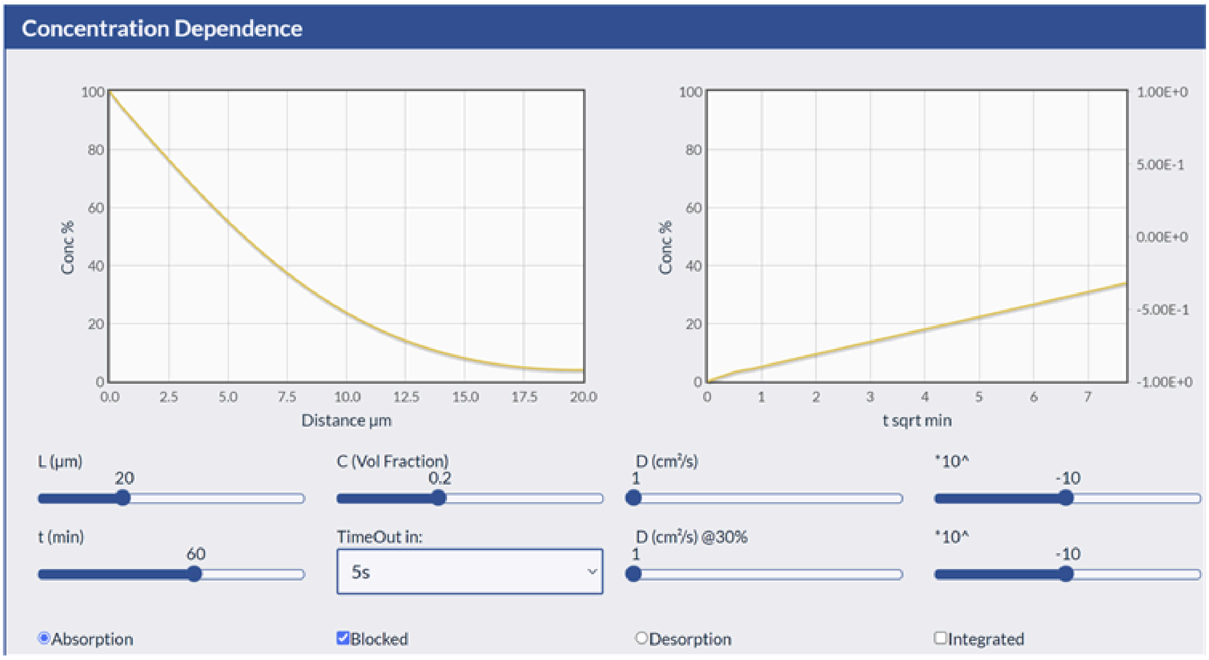
[Water-based Adhesive](), [Pharma Formulation](), [Microencapsulation]()

> Standard Fickian diffusion theory works fine for many cases – but if the liquid is significantly soluble in the polymer, diffusion coefficients get much larger, with some expected and unexpected results.

2 orders of magnitude

The diffusion coefficients described in Diffusion_Basic Diffusion and in Diffusion_Diffusion Coefficients are constant. Some polymers have large constants, say, 10^{-8} cm²/s and others might be lower, say 10^{-10}. The difference is due to the amount of "free volume" in the polymer. To diffuse, a molecule must have somewhere to move to, which means some free space that opens up when, say, the polymer chain wiggles. Compact polymers and those below their Tg have little free volume.

Now let some solvent diffuse into the polymer. Because the concentration is a bit higher, the polymer is slightly swollen and there is more free volume – so the solvent can diffuse in faster, the concentration gets higher, the diffusion coefficient gets larger. It can easily increase by 2 orders of magnitude.



L (µm) 20

C (Vol Fraction) 0.2

D (cm²/s) 1

*10^ -10

t (min) 60

TimeOut in: 5s

D (cm²/s) @30% 1

*10^ -10

● Absorption

☑ Blocked

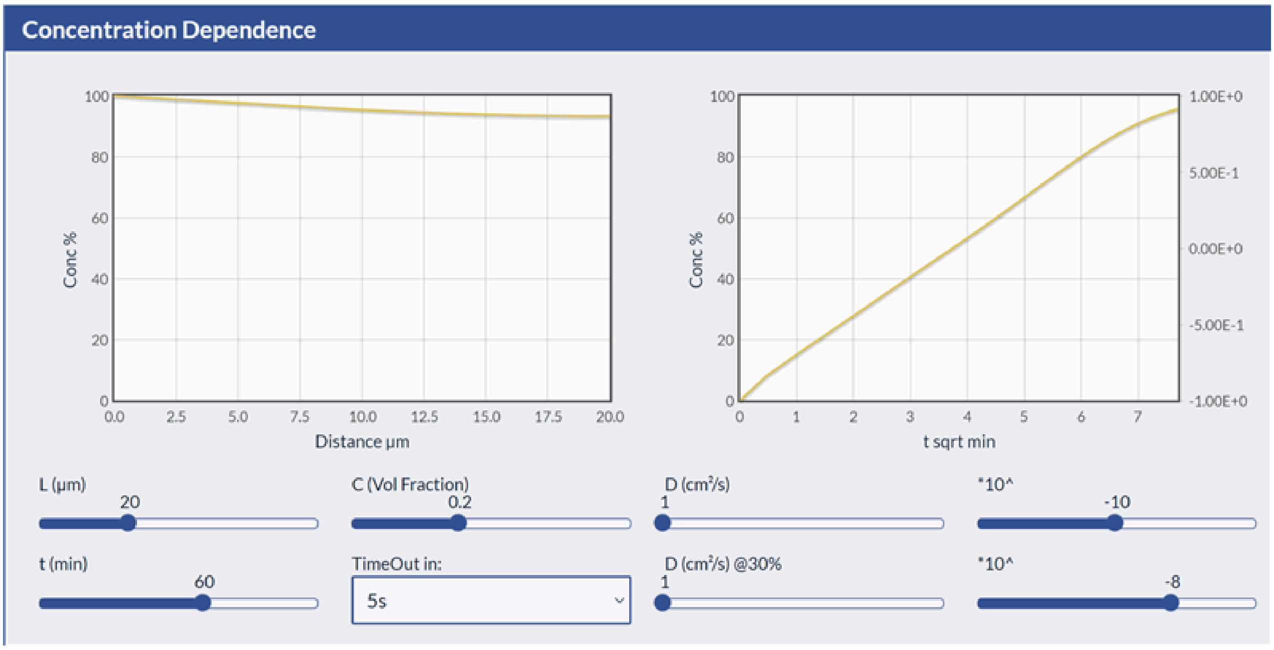
○ Desorption

☐ Integrated

https://www.stevenabbott.co.uk/practical-solubility/diff-cdepend.php

Here we have a 40µm film of polymer. Because the "Blocked" option is checked, solvent coming in doesn't exit from the other side. Exposing a 40µm sample on both sides is equivalent to 20µm blocked. After 1hr the solvent has just reached the middle. The diffusion coefficient, D, is concentration independent at 10^{-10}.

Now let’s make it concentration dependent so at higher concentration D = 10^{-8}.



L (µm) 20

C (Vol Fraction) 0.2

D (cm²/s) 1

*10^ -10

t (min) 60

TimeOut in: 5s

D (cm²/s) @30% 1

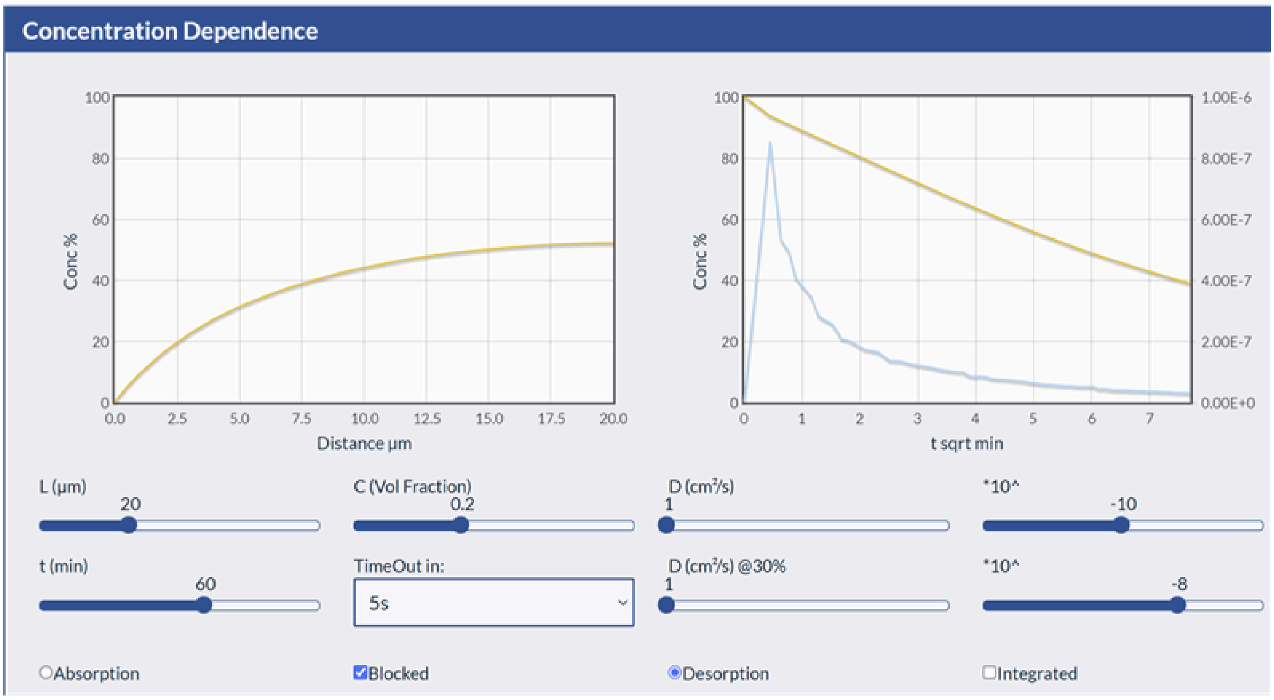
*10^ -8

Now after 1hr the solvent has almost saturated the polymer. This is what you might expect when you think about it.

So what is the unexpected consequence?

Asymmetry

In 1 hour, dipped in solvent, the film gets full of solvent. If we take it out and the solvent is free to evaporate from the surface (so the concentration on the outside is 0) we might expect the solvent to empty in 1hr. But that’s not what happens:



L (µm) 20

C (Vol Fraction) 0.2

D (cm²/s) 1

*10^ -10

t (min) 60

TimeOut in: 5s

D (cm²/s) @30% 1

*10^ -8

○Absorption ☑Blocked ●Desorption ☐Integrated

We are now in Desorption mode. After 1 hour, the concentration near the middle is still 50% and the total remaining (right hand graph) is 40%.

How can this asymmetry arise. Why can it fill to 100% in 1 hour but only lose 60% via desorption?

The answer is that the solvent has to escape via the surface with its low concentration of solvent and, therefore, a low diffusion coefficient. Absorbing the solvent is like having a tap wide open on the outside, desorption is via the same tap allowing only a drip of fluid to escape.

Many of us have applied coatings to polymers using solvents that can diffuse into the substrate. We dry our coatings and think that we have no solvent left. But it might have rapidly diffused in by a few µm and will take a long time to escape. We might be lucky and never notice – or it might cause a problem in subsequent processes for us or our customers.

I learned about this desorption phenomenon the hard way. We had coated a product using cyclohexanone, CHX, as the solvent. Obviously we dried it well; CHX has a pungent odour. The product was cut into packs of sheets, sealed in polyethylene bags, and shipped abroad. When opened after some weeks there was an overpowering odour of CHX. Our solvent had rapidly diffused in and we had been unaware that it would slowly diffuse out during storage and get trapped inside the polyethylene which, for HSP reasons, is a good barrier.

Those who print multilayers of ink sometimes find that 4 layers adhere well, while 5 layers show poor adhesion. The instinct is to blame one of the colours and to change the print order to try to fix the problem. The real reason for failure is a slow build-up of solvent in the substrate; each printed/dried colour leaves a little extra solvent in the top µm of the substrate. At, say, the 5th layer, the ink/substrate interface is sufficiently weakened by the solvent and adhesion fails.

Diffusion Diffusion Coefficients

Links

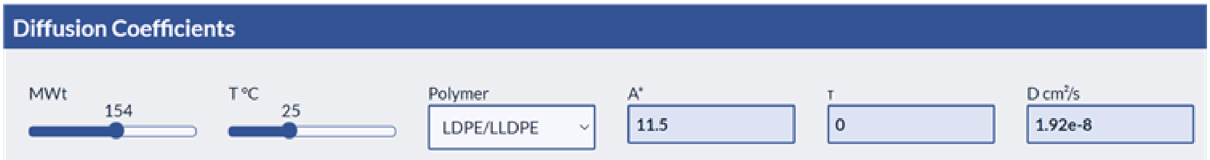
SkinCare, Surface Cleaning, Water-based Adhesive, Microencapsulation

To calculate diffusion we need to know the diffusion coefficients of our materials. Where can we find them, and how do they depend on temperature and on the MW of the molecule diffusion?

And if we can’t find them, how can we measure them?

A good-enough guide

Because lots of us need diffusion coefficients, D, to calculate diffusion for purposes such as regulatory compliance, and because regulators know it is unreasonable to expect everyone to know D values for every molecule in every polymer at every temperature, there are some semi-official formulae available which, if you use them, allow you to show that you are putting in reasonable effort.



Diffusion Coefficients

MWt 154

T°C 25

Polymer LDPE/LLDPE

A* 11.5

τ 0

D cm²/s 1.92e-8

https://www.stevenabbott.co.uk/practical-solubility/diff-d-values.php

You enter the MW of your molecule, the temperature of interest and your chosen polymer. D is calculated for you.

The formula used is:

$$D = 10^4 e^{A - C_1 MW^{0.667} + C_2 MW - C_3 / T}$$

Where $C_1 = 0.135$, $C_2 = 0.003$ and $C_3 = 10454$ and $A = A^* - \tau / T$ and T is in °K. Those two constants, A* and τ are polymer specific.

The app has some alternative equations should you be interested. The exponential dependence on $MW^{0.667}$ is one common form but the way it is used here gives an effective $\left(\frac{100}{MW}\right)^2$ dependence which is a commonly-used approximation, using 100 as a convenient value for typical common solvents such as toluene.

Measure your own

Although there are many fancy academic ways to measure diffusion coefficients, it’s not so hard to get a good-enough value via some simple experiments.

The easiest method is to get a thin sample, say 50µm, weighing a few grams, and to place it in some solvent that you know is compatible, but not too compatible, with the polymer. At increasing intervals take it out, wipe it dry, weigh and return to the solvent. Finally, leave it overnight and remove/wipe/weigh to get the equilibrium absorption value. Using the https://www.stevenabbott.co.uk/practical-solubility/diff-fickian.php Fickian diffusion app in “blocked” mode, slide the various sliders (remember to set the thickness to half your sample thickness as solvent is coming in both sides) till your data match the experimental values. If the solvent is too good then you will get a poor match. Swap over to the https://www.stevenabbott.co.uk/practical-solubility/diff-cdepend.php Concentration Dependent app and play around with values till you get an OK fit.

This sounds crude, but it is remarkably effective. You generally don’t need D values to high accuracy, and the advantage of the simple method is that you can do it with various solvents to check on dependence on molecular size.

The trick is to get a sample thin enough to allow you to reach equilibrium in a reasonable time, with a weight matched to whatever balance you have so that weight increases can be measured accurately. Before carrying out the detailed work you need two preliminary experiments:

1. The simple-sounding “wipe it dry” needs a bit of practice, so before committing to a proper experiment, try dipping the weighed sample into the solvent, removing it immediately and testing that your wipe/dry procedure leads to no significant weight increase.
2. The “compatible but not too compatible” solvent can readily be found by setting up, say, 4 different solvents, placing a sample in each and leaving overnight. A solvent that has seriously damaged the sample is obviously unsuitable, and any solvent with too small a weight increase is equally unsuitable.

This approach of doing “scouting” experiments is highly recommended; it is a small investment of time to avoid large wastes of time when attempting unfamiliar experiments.

A $multimillion contract depended on proving that a certain hydrophobic material would not absorb more than a certain % water over a long time. Doing the test would have required 3 months of dunking the sample in water – time they didn’t have. Nor did they have the diffusion coefficients to calculate the effect. I suggested they made a thin sample (their real object was thick) and doing the dunking, removing, drying, weighing and re-dunking technique to measure D. To their astonishment after a couple of days they had a reasonable value. Refining their technique, a few days later they had a reliable value, could prove to themselves and their client that their material met the spec, and they won the contract. Not bad for a few days of simple weighing of thin samples of their material.

If you have a convenient ATR-FTIR spectrometer then you can place your ATR crystal onto the polymer which has the solvent trapped on the other side. Keep scanning till you start to see the peaks from your solvent. This gives you a breakthrough time which you can find in the Fickian app. There is some uncertainty about how much solvent constitutes “breakthrough” and how sensitive your spectrometer is to that solvent, but again you can get surprisingly good values for relatively little effort.

If you want to test for breakthrough using some sort of solvent detector (e.g. a simple mass spec “sniffer”) then the biggest difficulty is getting a good seal between the sample and whatever is separating the solvent from your sniffer. There is a long history of false high D values due to inadequate seals, especially for polymers with very small D values.

Diffusion Diffusion into Skin and Hair

Links

Hair Conditioners, Sun Screens, SkinCare, Lipstick, Mascara, Pharma Formulation

The dual (and sometimes conflicting) demands of beauty and safety require an understanding of how chemicals diffuse into skin and hair. The combination of solubility and diffusion science in the FST allows us to make more informed formulation decisions.

Skin diffusion

A paper that happened to be technically correct but deeply deceptive convinced the skin community that permeation was controlled by LogP, the water/octanol partition coefficient. This is provably wrong. The real correlation is with molecular size which, in turn, is the single most useful predictor for LogP. And the correlation was with a partition ratio from water *not* with what matters which is the flux from whatever medium is being applied.

The true flux dependency is simple and obvious:

- Solubility of the ingredient in the medium being applied to the skin.
- Partition between the medium and the skin.
- Molecular size correlation with diffusion coefficient.
- Skin thickness (highly variable for different parts of the body).
- Skin diffusion coefficient (probably ~ 10^{-9} cm²/s but depends on quality and moisture level of the skin.

With the use of Dissolution_Hansen Solubility Parameters you can readily work out correlations with the first two and with a diffusion modeller Diffusion_Basic Diffusion and Diffusion_Diffusion Coefficients you can work out the final piece.

Because skin diffusion is relatively slow, the one ingredient you can ignore is water. There is a vast amount of work (using, especially, Franz cells) measuring permeation from aqueous solutions over long periods such as 24hr. This is all a complete waste of effort. The water in a typical skin formulation is lost within ~ 20 minutes. This means that all experiments that look at other aspects of aqueous delivery, such as the type of emulsion (w/o, o/w, microemulsion) are meaningless because after 20 minutes there is no emulsion because the water has evaporated.

When you do the modelling described above it is, therefore, important to model the % of ingredients on the skin. A lovely example is a skin formulation that delivered none of the desired API (active pharmaceutical ingredient) through the skin. It was delivered as 70% water emulsion containing 3% of a surfactant such as Tween. When the water disappeared, the 3% surfactant was now 10% of the formulation and because there was no water, it was no longer surface active. It was merely a chemical. In this case, a chemical with a lot of polyethylene oxide(PEO) chains, and one so large (above the typical 500 Dalton MW limit) that it does not permeate through the skin. The API happened to have a great HSP match for PEO, so it remained happily on the surface, trapped in the surfactant.

Another aspect of HSP relevant to skin permeation is the fact that two bad solvents can make a good one – their HSP end up as an average which might be a good match for a chemical or for skin. A classic example was the observation made some years ago that although propylene glycol, PG and "essential oils" like limonene were relatively ineffective at aiding skin permeation, a mix was quite potent. When you do the calculation, the HSP is a good match for the HSP of skin.

HSP of skin?

If we think of the skin as a polymer, and use standard HSP techniques for measuring its compatibility with various solvents, you can work out that it behaves like a polymer with an HSP of something like [17, 8, 8]. People get angry at this value. After all, they "know" that skin is permeated through the lipid layer and lipids are hydrophobic. But the lipid layer contains 20+% of cholesterol which easily takes the HSP up into the [17.7, 8, 8] range. In any case, the idea that chemicals permeate only via the lipid layer has become less and less tenable.

If you look up the HSP of dimethyl isosorbide (DMI), you find that it's [17.7, 7.1, 7.5]. If you take the naïve HSP prediction as true, you would expect DMI to be great at penetrating skin. This is indeed the case. In fact it can be too good. If you happen to use DMI as a solvent for an API with a large MW and therefore a slow rate of diffusion, the DMI can penetrate the skin faster than the API. At some point the API on the surface has too little DMI and crystallises out. When the crystal form happens to be needle-like, the result can be very painful.

> I first came across DMI in a non-skin context. I worked out its HSP values and saw that it would be superb as a skin permeation enhancer. I mentioned this to someone in the skin care industry who, to my delight, confirmed that it had been found to be green, safe and an excellent permeation enhancer.

Hair

Because healthy hair has a strong armour of high-quality keratin, diffusion into the hair itself is minimal. It takes heat, chemicals and aggressive solvents to open up the keratin to allow molecules (such as dyes) to migrate inside.

The outer layer of high-quality fatty acids (either free or bonded onto the keratin) such as 18-MEA (18-methyleicosanoic acid) provides a truly hydrophobic layer which ideally would be mostly left alone.

As hair grows, the younger portions near the roots are in good shape while the older portions, near the tips, have sustained more damage so are more open to diffusion.

Add to this the assaults from surfactants, conditioners, silicones, quats, hair dryers, bleaches and dyes, it becomes less and less clear what "hair" is.

Until we know more, we can basically say that under normal circumstances diffusion into hair is, and should be, difficult and that after extensive human interventions, diffusion is probably easy and undesirable.

Diffusion Permeation OTR and WVTR

Links

Lipstick, Mascara

Everyone gets OTR (Oxygen Transmission Rate) and WVTR (Water Vapour Transmission Rate, or M for Moisture) values wrong. Literature values are wrong, internal publications are wrong, values provided by testing facilities are wrong. It's a mess. This chapter does its best to sort out the mess, with a life-saving app being at its heart.

We're all wrong and confused

When a leading expert in permeation admits in a paper that she and her colleagues have spotted errors in their own values, and in books, journals and databases, you know that the situation is bad. The wonderful paper from Prof Kay Cooksey and colleagues at Clemson University inspired the apps at https://www.stevenabbott.co.uk/practical-coatings/permeability.php which have been used 10's of thousands of times by people as confused as everyone else.

It is so easy to make an error, and even the app admits that a couple of errors had crept it which were spotted by smart users so they could be fixed.

Even when we're not wrong, we are often confused. So, take heart. You are not alone in finding this topic hard.

Permeability versus Transmission Rate

The first confusion is between Permeability, P and Transmission Rate, TR. Another layer of confusion is that people talk of "permeation" which might mean either P or TR.

Here's the difference:

- P is the amount that will go through a unit area of a unit thickness of a barrier in unit time and is usually defined for a single material.
- TR is the amount that will go through a unit area in unit time of your specific piece of barrier film. It can be the value for a given thickness of a single material or for multiple layers of different materials.

Calculating TR values

Let's say that we have 3 layers of materials. We know the TR values of each of the 3 materials (i.e. at the specific thicknesses used in our 3-layer barrier), TR_1, TR_2 and TR_3 so we can calculate the overall TR via:

$$TR = \frac{1}{\frac{1}{TR_1} + \frac{1}{TR_2} + \frac{1}{TR_3}}$$

If we know the thicknesses L_x of the individual layers and we know their P values then TR is given by:

$$TR = \frac{1}{\frac{L_1}{P_1} + \frac{L_2}{P_2} + \frac{L_3}{P_3}}$$

The app does the calculations for you, with other variants that are needed from time to time.

But it’s complicated

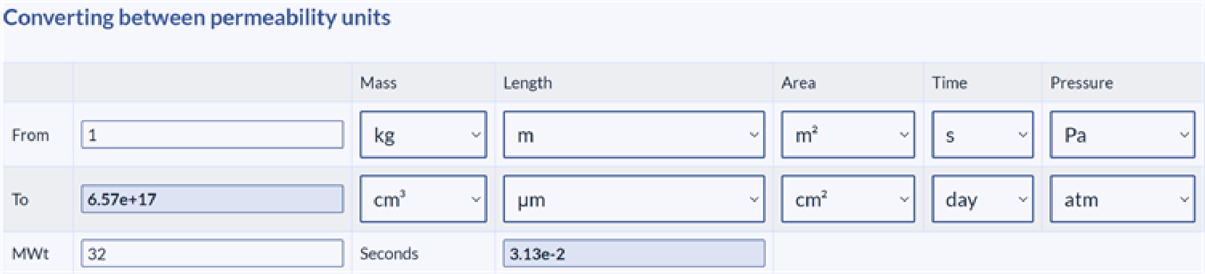
If (and that’s a big if) you know that the OTR was measured with air on one side of the barrier, then you have only 20% O2. If you know that P was measured and calculated with respect to 100% O2 then you have a factor of 5 difference if you are mixing and matching TR and P values.

Similarly, if WVTR was measured at 50% RH you might (or might not) have a factor of 2 correction for P. Why “might not”? Because if someone has a barrier that is very sensitive to water, they might want to quote P “at 50% RH” and, rightly, not including the factor of 2 correction because at 100% RH it might be a factor of 10 higher.

There is nothing the app can do to help with this mess. You just have to read the relevant literature or report carefully.

The units of P

This is another nightmare.



Converting between permeability units

		Mass	Length	Area	Time	Pressure
From	1	kg	m	m^2	s	Pa
To	6.57e+17	cm^3	µm	cm^2	day	atm
MWt	32	Seconds	3.13e-2			

https://www.stevenabbott.co.uk/practical-coatings/permeability.php

If a paper quotes O2 permeability in kg.m/m^2.s.Pa (that’s 5 units – mass, length, area, time, pressure) and you need it in cm^3.µm/cm^2.day.atm then you might have a problem knowing how to convert them. The app does it for you. If the original value was 1 then the converted value is 6.6.10^{17}.

Note that if you convert from a mass unit (here kg) to a volume unit (here cm^3) you need the MW of the gas, here 32 because it’s O2.

If everyone used clear 5-unit units then conversion is tedious but do-able via the app. But someone looking at cm^3.µm/cm^2.day.atm might say “Oh, we can cancel the cm” and give the units as cm.µm/day.atm. If their length unit had been cm then instead of cm^3.cm/cm^2.day.atm they might give you cm^2/day.atm, which is even more confusing.

The sad thing is that the “real” units for P are, surprisingly, seconds. In the above example P was 0.0312s. No one quotes P in s, but if they did, our nightmare of conversions would go away.

However, this wouldn't fix the problem of users recording P based on 1 atm of air rather than 1 atm of O2 (the factor of 5 error), and confusing 100% RH with 1 atm of water which, at 25°C is an error of 760/25 ~ 30. It's surprisingly common to find values that are out by a factor of 30.

> I enjoy helping people who email me with formulation issues. But when I see "permeation" in the Subject I open the email with dread. It means opening my own app, reminding myself of the complexities, then tracking down the chain of errors (often in literature values) that has caused the problem in the email. What encourages me to provide an answer is that the knowledge of the problems and their solutions, via the app, is steadily spreading. I now get fewer requests for help because the wider community is better able to help itself.

Different T and %RH

WVTR values measured at a given T and %RH can, in principle, be converted to other T and %RH values:

WVTR/MVTR at different T and %RH

WVTR	1	T°C	38	%RH	90
In Atm	1.70e+1	In mm/Hg	2.24e-2	In MPa	1.69e+2
T°C	23	%RH	50	$WVTR_{calc}$	2.35e-1

It's complicated and works only for those barriers where P doesn't change with %RH. But if someone quotes a WVTR at 38°C/90% RH and you need to know what it would be at 25°C/50% RH, then with a bit of thought and effort you can get there.

It's a mess

Wouldn't it be nice if a magic app could sort out all these problems. It would, but the problems are so deep and the data "out there" so poor and confusing that there is no substitute for sitting down and going step-by-step through whichever problem you are facing.

You *will* find values that are out by factors of 30 or 10^6, you *will* find wrong units or the right units cancelled out (as in the cm examples above), or just about any mixture of errors you can imagine. If you go slowly, if you imagine the sorts of errors that you might make, if you think through the traps of unit conversion, atmospheres, volumes etc. then you will find why the values you find in a paper make no sense and then reach a reasonable conclusion about the values their experiments have really measured.

That's the best you can do. Take inspiration from the experts who admitted their own failings and do your best to work around errors from others and to check that your own work isn't confusing TR with P, isn't cancelling out units, and isn't wrong by factors of 30 or 10^6.

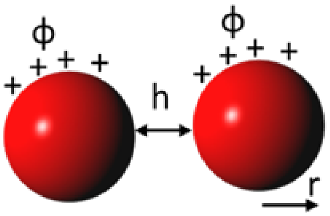
Dispersions DLVO

Links

Water-based Adhesive, Emulsion Paint, Emulsion Polymers, Inkjet Inks

> DLVO theory is often seen as *the* theory of dispersion. But it's surprisingly limited other than for charged species in water where it *is* useful. All relevant formulae can be found when you use the DLVO app.

Attraction and Repulsion



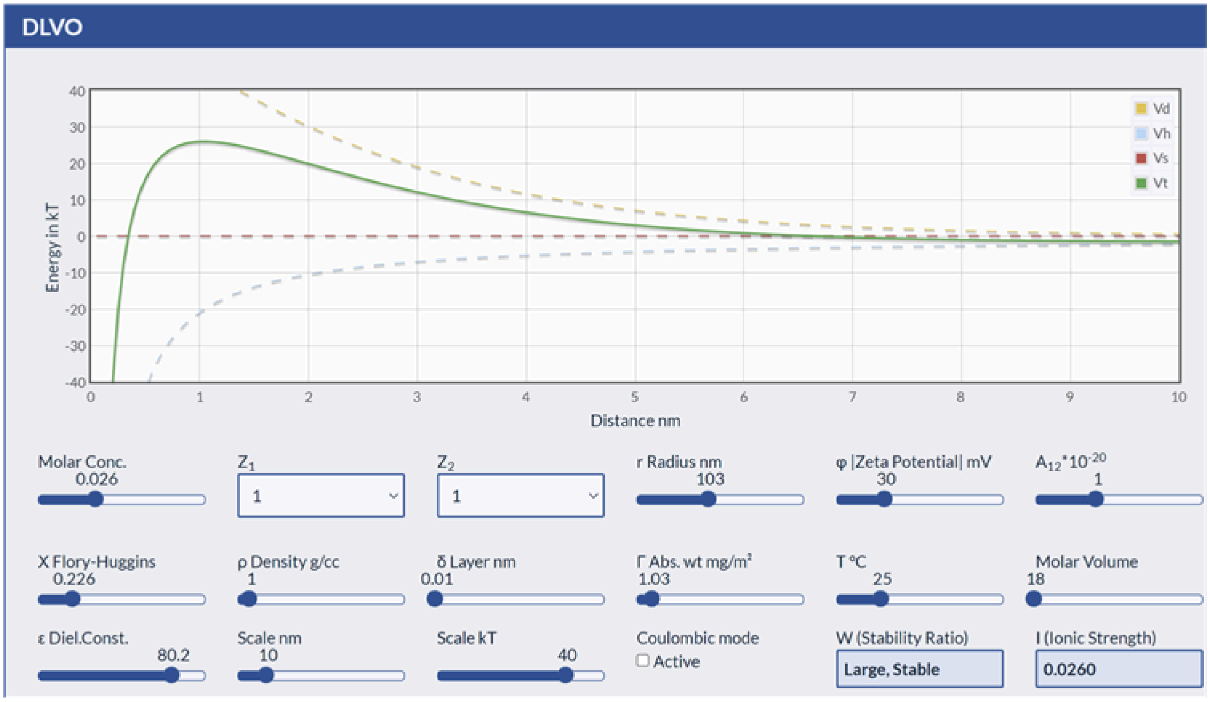
We have two particles of radius r separated by a distance h and having a surface of charge φ:

They have an attraction due to van der Waals forces which depend on the Hamaker constant, A_{12}, between the particles and, strongly, on r.

They also have a repulsion ("like charges repel") that depends on the charge and the dielectric constant and concentration of other ions in the water.

The charge is surprisingly ill-defined, for reasons discussed in Dispersions_Zeta potential.

DLVO (from Derjaguin and Landau, Verwey and Overbeek) theory lets us calculate the relative attraction and repulsion forces:



Molar Conc. 0.026

Z_1 1

Z_2 1

r Radius nm 103

φ |Zeta Potential| mV 30

$A_{12}*10^{-20}$ 1

X Flory-Huggins 0.226

ρ Density g/cc 1

δ Layer nm 0.01

Γ Abs. wt mg/m² 1.03

T °C 25

Molar Volume 18

ε Diel.Const. 80.2

Scale nm 10

Scale kT 40

Coulombic mode ☐ Active

W (Stability Ratio) Large, Stable

I (Ionic Strength) 0.0260

https://www.stevenabbott.co.uk/practical-solubility/dlvo.php

The y-axis is the interparticle energy in units of kT: positive is repulsive, negative attractive. The light-blue dotted line is the Vh curve, (H for Hamaker) which shows the rapidly increasing van der Waals attractions as the particles get very close. The light-yellow dotted line is the Vd curve (D for Debye, the theory used for the electrostatic repulsion) which shows increasing repulsion.

The solid green line is the net result. The repulsion peaks at ~28 kT then van der Waals takes over. But as long as the particles don't have enough energy to overcome the 28 kT, they will stay apart.

Controlling the barrier

Because the attraction depends on the Hamaker constant, which is typically 10^{-20} J, you might think it a good idea to reduce it. Unfortunately, most of us most of the time have no idea what the real value is and if we did there's little we can do to change it. There *is* however one trick. If you have, say, polystyrene particles in, say, a mix of decalin and tetralin, it is possible to adjust the refractive index of the solvent blend to exactly match that of the polystyrene … and A_{12} goes to zero. *[Strictly, you also have to match the dielectric constants, but the effect is usually negligible].* Some elegant experiments looking into particle behaviour unperturbed by attractions have done just that. But then we're not in water, and getting an exact match to the low RI of water is rather a challenge.

For charge repulsion if the barrier is over 20kT (some say 30kT) then the dispersion is likely to be stable. Unfortunately, the barrier size *decreases* as particle radius decreases, so smaller particles are more difficult to stabilize. The barrier depends on the Debye length which is confusingly called k^{-1}. This is because the effective voltage near a particle falls off with distance h as e^{-kh}. Because we want the exponential to be without units, and h is a distance we talk of k^{-1} as the Debye length, the inverse of the Debye exponential factor.

For a given particle you can generally change its charge by changing the pH. Particles might be -ve at high pH and +ve at low pH, passing through zero charge at the isoelectric point. So the first trick for stability is to control the pH to remain in a high charge region.

If you have to add extra salts to the formulation for other reasons, the repulsive forces decrease. The ions shield the mutual repulsive charges. The effect is controlled by the ionic strength which means that cations like Ca^{2+} or anions like SO_4^{2-} are especially troublesome because ionic strength depends on charge number squared, so if you are OK with a certain concentration of NaCl, you might have problems with the same molar concentration of $CaCl_2$ or Na_2SO_4. The app calculates the ionic strength as you change the charge numbers, and you see the effect on the charge barrier.

The app shows the simple theory. There are some subtleties not captured. For example, if you have two types of particles, one with a zeta potential of, say, -5mV and one with -50mV, there can be an *attraction* at low ionic strength, repulsion at medium concentration and attraction at high ionic strengths.

It is possible to accidentally quench the particle charge. For example, a small amount of cationic surfactant can be attracted to a negative particle, leading to a near-zero net charge and near-instant clumping of the particles.

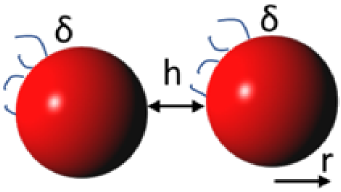
Finally, and importantly, charge stabilization is effective only in water and a few other solvents with a high dielectric constant. As you move to more typical dielectric constants, the repulsive force decreases rapidly, giving little effective stability. In these cases you need steric stabilization.

DLVO as a religion

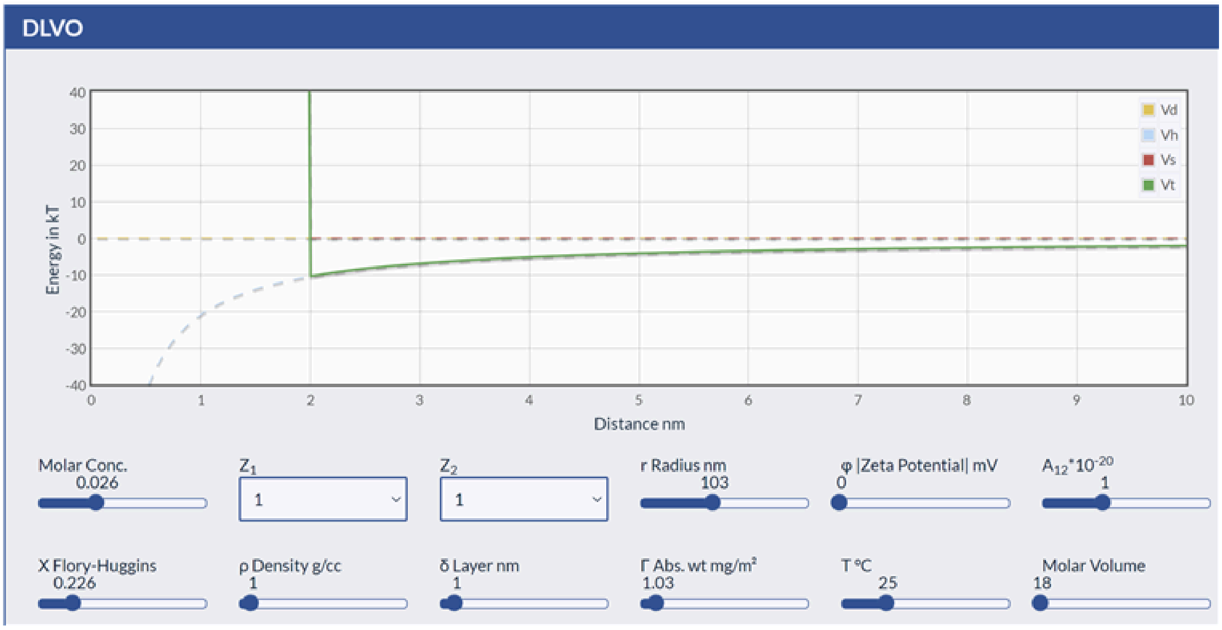
Famous colloid scientist Prof Barry Ninham has referred to DLVO has a “religion”. As he was friends with both D and O this was not a personal attack on the authors. Instead it was an attack on a lazy community that had taken a rough-and-ready, impressive-for-its-time theory and elevated it to *the* theory of colloid science. It is a recurrent theme of FST that once-useful theories take on a life beyond their (small) zone of utility to become unarguable orthodoxy. As a community we should always take the best from any theory but be actively alert to its faults. This especially applies to the FST. Please send in your critiques of its faults.

Steric stabilization

Although steric stabilization is not a part of DLVO, it is usually included under the DLVO heading. It works in aqueous systems but, unlike charge stabilization, also works well in non-aqueous systems, so is a more general technique.



We have polymers sticking out from the surface by a distance δ. Turning off the charge in the app and adding δ = 1nm layer of polymer results in a very strong particle-particle repulsion at a distance of 2nm.



Molar Conc. 0.026

Z_1 1

Z_2 1

r Radius nm 103

φ |Zeta Potential| mV 0

A_{12}*10^{-20} 1

Χ Flory-Huggins 0.226

ρ Density g/cc 1

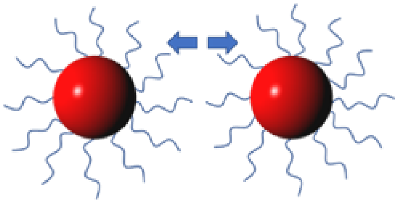
δ Layer nm 1

Γ Abs. wt mg/m² 1.03

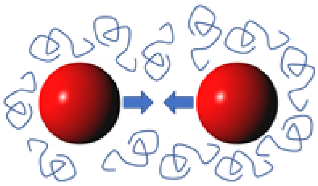
T °C 25

Molar Volume 18

This is steric stabilization and it works very well … until it suddenly doesn’t.

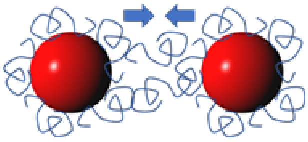


The key factor is the Dissolution_Flory-Huggins χ parameter, a measure of how much/little a polymer prefers itself to its solvent environment. When this is less than 0.5, steric stabilization is effective. As you change the solvent environment and χ approaches 0.5 you see no obvious problem, so you change the solvent a little more … until suddenly you get flocculation.



With χ > 0.5 the relevant term changes from repulsive to attractive and the formulation is ruined. This sudden switch from stable to unstable has surprised many formulators who have only made a "small change" to fix some other issue. For those who know about how Dissolution_Hansen Solubility Parameters relates to the Flory-Huggins χ value, any formulation changes can be checked to see if they are likely to tip things into the wrong part of solubility space.

Although the basic idea works well, such a simplistic approach cannot cope with the issue of extra polymer in the solution – something we frequently need to provide. This can give rise to extra stabilization or to destabilization via depletion or bridging flocculation. Although there are theories for depletion flocculation (https://www.stevenabbott.co.uk/practical-solubility/depletion.php) they aren't really of any use.



Here we have Bridging Flocculation, which is surprisingly hard to explain compared to depletion flocculation, especially because beyond a certain polymer concentration, stability is re-established. However, the powerful Dispersions_Scheutjens-Fleer technique copes with stabilization and de-stabilization phenomena and should be part of the toolkit of anyone working with particles and polymers.

The fact that a Scheutjens-Fleer app exists is an example of academia at its best. When I "discovered" it, I immediately saw its power and usefulness so wanted to write an app … then found that the theory was way beyond my capabilities. An email correspondence with Prof Frans Leermakers resulted in his kind permission to use his powerful SF Box engine on my server, with my front end simply sending parameters and plotting the results.

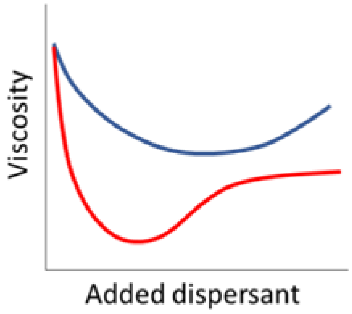
Dispersions ODC

Links

Lipstick, Mascara, Solvent-Based Paint, Inkjet Inks

Many formulators try a few dispersants in their full formulation and go with whichever works. A better approach is to find an Optimal Dispersant Concentration, which means both an optimal dispersant and an optimally low concentration required.

Optimal Dispersant Concentration



We add our pigment/particles to our solvent/water and (ideally in real time) measure the high viscosity due to the pigments being ill dispersed. We then slowly add a dispersant, measuring the fall in viscosity. After a while, the viscosity increases as the excess dispersant is probably a polymer, adding its own viscosity.

For the blue dispersant, the fall in viscosity wasn’t large, and it needed a lot of dispersant to reach a minimum viscosity with the maximum of dispersion. That its ODC.

For the red dispersant, the fall in viscosity is large and you need rather little to get a great dispersion. Its ODC is smaller and the results are better, so it’s the dispersant of choice.

Of course these curves are idealised and measuring the viscosity in real time is tricky, but by finding the closest possible equivalent in one’s own lab, it’s an effective way to screen multiple dispersants and find the ODC values (concentration and viscosity reduction) that meet your needs, along with cost, greenness etc.

The idea of an ODC was new to a major corporation working on a challenging pigment-based product using a range of pigments. They were using a general dispersant and had not realised that a dispersant optimal for each pigment would give them better performance without the compromises they were enduring. Having reviewed the amount of work needed to find the ODC and deciding that they didn’t have the time/resources to do it internally or to contract outside, they decided to see if they could purchase their pigments already optimised for their specific system. They found a supplier and not only did the formulations work better, it was cheaper to buy in easy-dispersing pigments rather than spending time and energy forcing their older ones into the formulation.

Optimal by calculation

With some basic thinking, we can work out an ODC for a rational choice of dispersant. Rational means that it has a head that likes to be strongly attached to the particle and a tail that will stick out into the solvent and provide the desired steric stabilization.

Optimal Dispersant Concentration

Particle Radius r nm 1000
Particle Weight g 1
Density ρ g/cc 1
Disp. Weight mg 2.5
Disp. MWt 450
Disp. Area Å^2; 90

Particle Area m^2: 3.00
Particle Area/g m^2/g: 3.00
Disp. Area m^2: 3.01
Disp. Area/mg m^2/mg: 1.20
Disp. mg/Area mg/m^2: 0.831
% Cover: 100%

https://www.stevenabbott.co.uk/practical-solubility/ODC.php

Here we are adding a dispersant with a head area of 90 Å² and MW of 450 to a 1000 nm particle. We've chosen 2.5 mg of the dispersant per 1 g of particle. We can calculate the surface area of the particle ($4\pi r^2$) and the total head area of the dispersant (g/Mwt to get moles, then x Avogadro x Area) so we find that we have a 100% coverage and a stable dispersion.

Of course we generally don't know the exact MW of our dispersant (suppliers provide less information than we might like) nor do we know the effective head area. But we can either choose to rely totally on the suppliers' promises that their dispersants are wonderful, or try to change the debate by using calculations such as this.

Why don't suppliers give us the information we need? Suppliers often claim that customers don't ask for the values, so there's no need to supply them. Customers often claim that it's not worth asking because suppliers never provide them. My view is that the core values could easily be provided, to mutual scientific benefit. The problem is that corporations prefer to sell magic rather than science. I'm not making this up. I've had heated arguments with marketing departments of megacorps who are terrified of their customers finding out that their wonder product is no different from everyone else's.

What is not included in the calculation is the chances of the dispersant wanting to stay on the particle. This requires some basic understanding of solubility and χ parameters, both captured in the Dispersions_Scheutjens-Fleer approach to improving our understanding of dispersants.

If you've never heard of Scheutjens-Fleer theory, that's not a surprise. Nor had I. The colloid science community has focussed on relatively useless theories while largely ignoring the one theory that is both powerful and usable. They had the excuse that the theory was hard to use – but that excuse has vanished now there's an app version plus a more powerful variant within the HSPiP software package.

Dispersions PVC and CPVC

Links

Lipstick, Mascara, Solvent-Based Paint

The Pigment Volume Concentration, PVC, and the Critical PVC are important formulation factors. If you exceed the CPVC the coating quality gets worse. Because they are *volume* dependent but pigments tend to get measured by *weight* you need to know the densities ρ of the pigments and the medium.

You also need to know the Oil Absorption, OA, which is the amount of a standard oil that the dry pigment can absorb. This is often a value given by the supplier.

CPVC

g pigment 45
OA g oil/100g pigment 20
ρ g/cc 4.2
Vol Particle 10.7
g C1 solution 55
ρ C1 solution 1
% C1 in solution 60
ρ C1 solvent 0.8
Vol C1 27.5
Total g 100
g C2 solution 0
ρ C2 solution 1.5
% C2 in solution 100
ρ C2 solvent 0.8
Vol C2 0.0
CPVC 52.7
PVC 28.0

https://www.stevenabbott.co.uk/practical-coatings/CPVC.php

We often need a lot of pigment in our coating or ink once all the solvent used to coat/print has disappeared. The problem is that pigment particles, when close packed, have lots of empty space between them. If we do not fill that space with binder then the properties are compromised: gloss, rub resistance, adhesion etc. will be lower than desired.

For each pigment there is a Critical Pigment Volume Concentration, CPVC, above which there is not enough binder present to fill all the voids (assuming that our binder is able to flow into and fill all available spaces). Pigment suppliers give us the density, ρ of the pigment and a value which tells us the amount of void space, expressed as the Oil Absorption, OA, which is the amount of a standard oil that the dry pigment can absorb. Because this oil has a density of 0.935 the CPVC is calculated as:

$$CPVC = \frac{1}{1 + \rho \frac{OA}{93.5}}$$

Now that we know the CPVC, all we need to know is our actual pigment volume concentration, the PVC. Typically we have a weight of dry pigment, then one or two extra "components", C1 and C2 (binders, fillers etc.) which are typically supplied as solutions. To get the PVC we need to convert our weight of pigment into volume (that's easy, it's g/ρ), then we have to take our g of component solution, the density of the solution, the % solids of the solution and the density of the solvent, from which the volume of the component can be calculated. Once we know all the volumes, the PVC is easy to calculate. If it exceeds the CPVC then it is shown up with a red background.

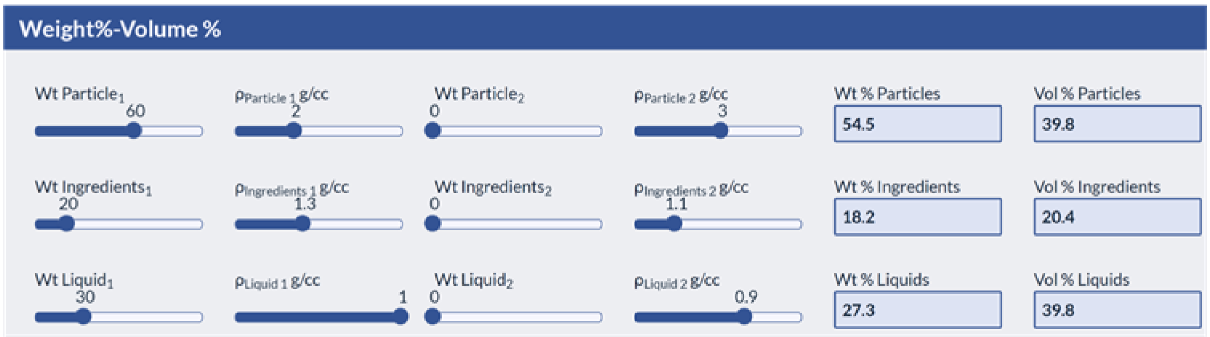
The complication for the app is that with 3 possible components, you have to adjust them individually so that the Total is 100g. If you accidentally miss getting close to 100%, the Total is shown up with a red background.

Various properties change with PVC and you can determine the CPVC from seeing where there is a sudden change. In the app the density is used to illustrate how you might go about this. As the CPVC is exceeded, the density decreases because of the air.

> My acquaintance with CPVC is rather recent. It seemed laughable to me that such an important aspect of formulation depends on “oil absorption” values measured by ancient techniques. However, I couldn’t find anything better, and everyone seems to be happy using them. If you (a) find CPVC calculations laughable and (b) have a better alternative, please let me know.

A handy Wt%, Vol% app

Although it’s trivial to convert weights of different particles, ingredients and solvents into wt% and vol%, users asked for a handy app to do just that, not focussing on CPVC. So here it is:



Weight%-Volume %

Wt Particle$_1$ 60
$\rho_{Particle\,1}$ g/cc 2
Wt Particle$_2$ 0
$\rho_{Particle\,2}$ g/cc 3
Wt % Particles 54.5
Vol % Particles 39.8

Wt Ingredients$_1$ 20
$\rho_{Ingredients\,1}$ g/cc 1.3
Wt Ingredients$_2$ 0
$\rho_{Ingredients\,2}$ g/cc 1.1
Wt % Ingredients 18.2
Vol % Ingredients 20.4

Wt Liquid$_1$ 30
$\rho_{Liquid\,1}$ g/cc 1
Wt Liquid$_2$ 0
$\rho_{Liquid\,2}$ g/cc 0.9
Wt % Liquids 27.3
Vol % Liquids 39.8

https://www.stevenabbott.co.uk/practical-coatings/Weight-Volume.php

For user convenience, weights are in arbitrary units going from 0 to 100. As long as all components are using the same units and weight scale, the app calculates the appropriate % values from whatever the total weight happens to be (so it doesn’t have to be 100). The densities of some well-known standard pigments/ particles/fillers are provided for convenience.

Dispersions Rheology (High shear)

Links

[Sun Screens](), [Lipstick](), [Mascara](),[Toothpastes](), [Water-based Adhesive]()

> With serious concentrations of particles we see a lot of complex rheological behaviours, not captured in the idealised [Dispersions_Rheology (Low shear)]() chapter. Unfortunately there are lots of ways that particles influence rheology so the app, and the chapter, are relatively complex.

Basic high shear properties

φ_c Percolation 0.28
φ_m Packing 0.63
N 1
Fractal Dim. d 3
Aspect Ratio A_f 1
φ^* : $\varphi_{c\ calc}$ 0.63 : 0.285
☑Log(γ)
☐Log Normal
●Viscosity
○Stress
☐High γ Corr.
☐Shear Thicken
λ Size ratio 1
k Fraction Small 0.33
φ_{max}Bimodal 0.63
E in kT 10
r radius nm 250
φ_{max} 0.5
φ values 0.28, 0.37,0.44,0.49,0.52,0.54
☑Custom φ's

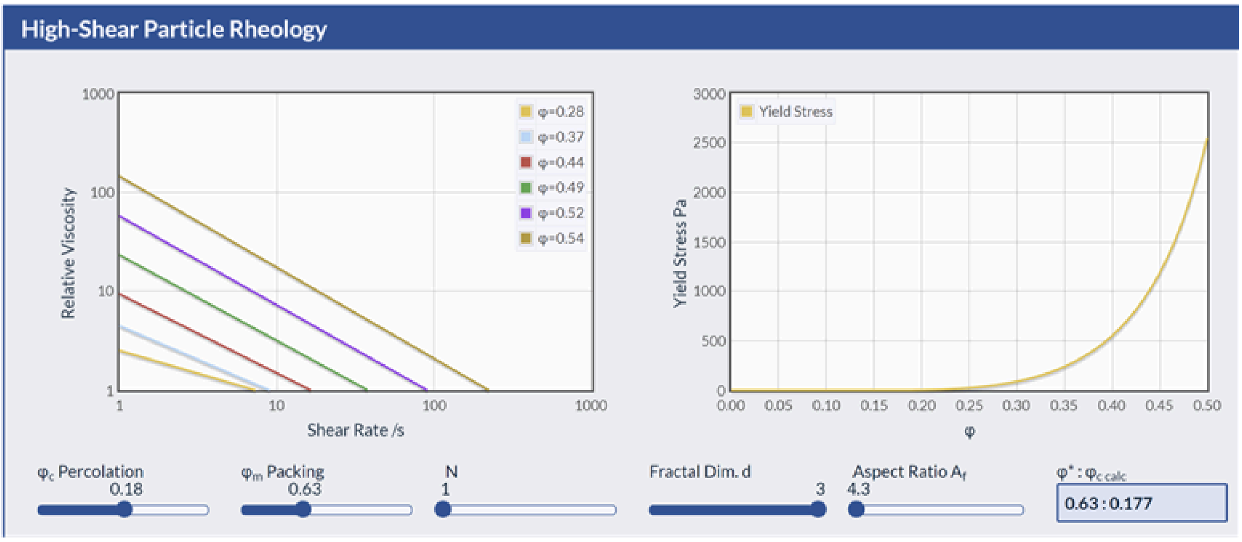
https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php

Ignoring all other settings, we see that the relative viscosity for a volume fraction φ = 0.28 is constant over the wide shear rate. For higher volume fractions, the viscosity gets much higher at low shear then decreases rapidly with shear. The formulae behind these curves are complex and you can read them in the app. But the take-home message is that although it is hard to handle high φ formulations at low shear, they aren't a problem at higher shear. The particles tend to line up like cars on a motorway and can flow along with little problem.

The viscosity approaches infinite at a critical φ* which we can set via φ_m to 0.63 in these screen shots. Why do we use φ* instead of φ_m? That will become clear later.

Aspect ratio

Now let's make our particles somewhat elliptical, with an aspect ratio (length/thickness) of 4.3.



φ_c Percolation 0.18

φ_m Packing 0.63

N 1

Fractal Dim. d 3

Aspect Ratio A_f 4.3

$\varphi^*:\varphi_{c\,calc}$ 0.63 : 0.177

The low-shear viscosity (you can use your mouse to read off the values) almost doubles. That’s because the percolation fraction, φ_c has decreased. What does this mean? As shown in https://www.stevenabbott.co.uk/practical-coatings/percolation.php, the chances of there being a continuous random network from one side of a container to another is near zero till φ reaches the percolation threshold where there’s now a high chance of such a chain being formed. Above the percolation threshold, viscosity is higher because the stresses at one side can be transmitted across to the other side via the percolation chain. Shear breaks up these chains, so viscosity is now shear rate dependent.

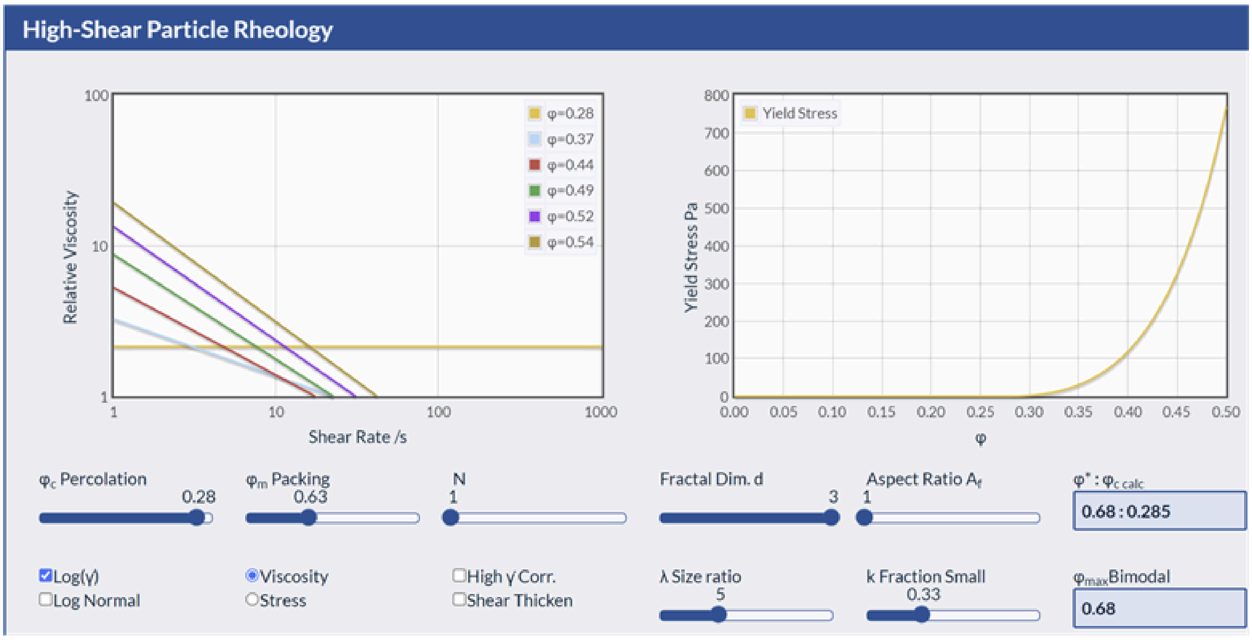
The percolation threshold depends strongly on the aspect ratio. For a sphere it’s 28%, for our 4.3 aspect ratio it’s 18%. So now the 28% formulation is no longer constant viscosity because percolation effects are strong.

The key lesson is that relatively small changes in the shape of your particles can have a significant effect on the rheology.

I was once asked to solve a weird particle formulation problem. I was using the app live in the discussions … and nothing was making sense. In desperation I changed the Aspect Ratio slider and the problem appeared in the graph data. There were some moments of silence. “Ah, we forgot to tell you that this specific particle is elliptical while all the others are round, we’d not thought it worth mentioning”. There are plenty of problems that can’t be solved by playing with sliders in an app, but it’s remarkable how often an app acts as a catalyst for finding a good solution.

Bimodal distribution

As you reach a high φ you are approaching the critical packing fraction where you can’t fit any more particles. That’s why viscosities take off strongly above 0.5. The close-packed spheres have lots of space between them which can be filled by smaller particles with no real impact on viscosity. So we often choose bimodal distributions (you can explore these via https://www.stevenabbott.co.uk/practical-rheology/distribution.php) to be able to pack in more particles:



High-Shear Particle Rheology

φ_c Percolation 0.28

φ_m Packing 0.63

N 1

Fractal Dim. d 3

Aspect Ratio A_f 1

$\varphi^* : \varphi_{c\,calc}$ 0.68 : 0.285

☑ Log(γ̇)

☐ Log Normal

● Viscosity

○ Stress

☐ High γ̇ Corr.

☐ Shear Thicken

λ Size ratio 5

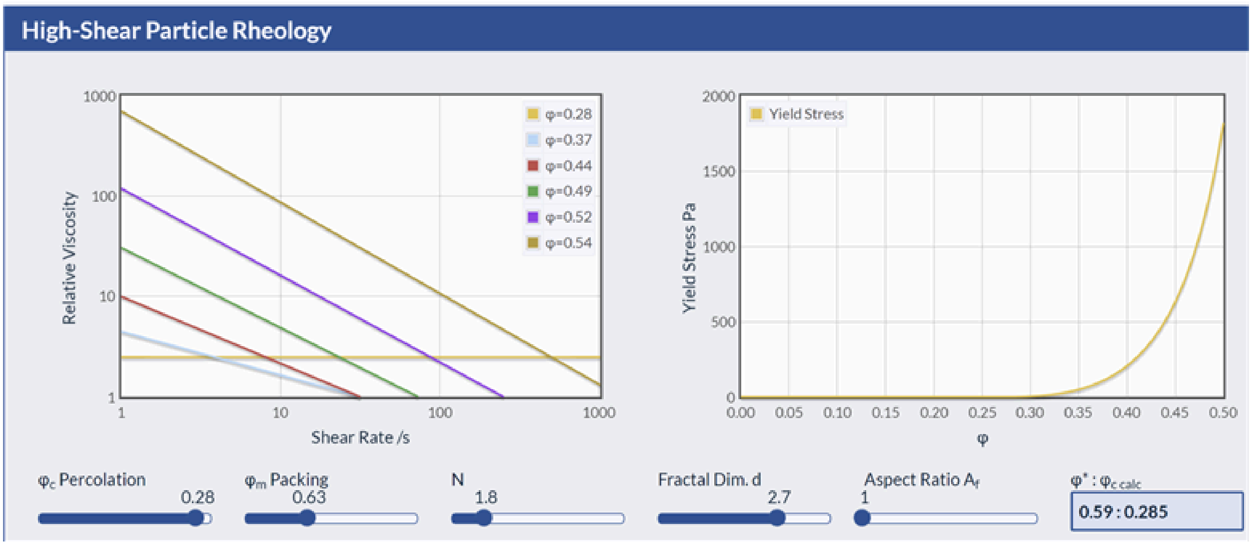
k Fraction Small 0.33

φ_{max} Bimodal 0.68

If we have a 5:1 size ratio and a fraction, k = 0.33, of the smaller particles then the critical packing fraction φ* rises from 0.63 to 0.68. The low shear viscosity of the 0.54 sample is now 1/3 of the original.

Real particles

Against our intuitions, the above calculations do not involve particle size. We know that in the real world, smaller particles tend to give more problems of self-association. There seems to be no app-able formula for this, so we have to use the two sliders N and Fractal Dimension d to get some idea of what might be causing higher viscosities.



High-Shear Particle Rheology

φ_c Percolation 0.28

φ_m Packing 0.63

N 1.8

Fractal Dim. d 2.7

Aspect Ratio A_f 1

$\varphi^* : \varphi_{c\,calc}$ 0.59 : 0.285

We have the same spherical particles, but now φ* has changed from φ_m = 0.63 to 0.59. This means that at, say, φ = 0.54 the particles are much closer to the critical parameter so the viscosity is more than 10x larger.

This has happened because on average there are 1.8 particles in some sort of association. That is a vague term because it also requires the Fractal Dimension. Perfectly random particles in 3D space have a fractal dimension of 3. If somehow aligned and confined to a plane, their dimension would be 2. Some modest alignment takes the dimension down somewhat from 3, in this case to 2.7. There's clearly some clustering to reduce the dimensionality; not a lot, but enough to change φ_m to this new φ* given by:

$$\varphi^* = \varphi^m N^{1-\frac{3}{d}}$$

It is frustrating that this formula is likely to contain a lot of truth while being mostly unusable as we know neither N nor d. Maybe someone will come up with something more usable. Until then, we just have to make impressionistic use of it and/or use any nugget of extra information available from our formulation to better understand what is going on.

Other effects

There are other options discussed in the app. They at least indicate other possibilities for observed behaviour and if your own formulation shows some peculiarities you can return to the app to see if the relevant phenomena have some explanation.

Yield stress

By definition, yields stress (see Flow_Yield Stress) is nothing to do with high shear. But it was convenient to include it in this app because the YODEL model from Flatt and Bowen on which it is based uses mostly the same parameters. The yield stress τ is given by

$$\tau = k\left(\frac{E}{r}\right)\frac{\varphi\left(\varphi - \varphi_c\right)^2}{\varphi_m\left(\varphi_m - \varphi\right)}$$

E is the attractive energy between particles, hopefully small, and r is the radius of the particle, showing, as intuition suggests, that smaller particles will give a higher yield stress.

Again it is unlikely that you will know E, but at least you get some sense of what influences it. Because φ_c is part of the equation, clearly higher aspect ratios (lower φ_c) will give higher yield stresses.

It is important in academia to have a good acronym for your fancy theory. Is it a coincidence that the acronym YODEL comes from Swiss researchers? I had a chance to ask them directly and, no, it’s not a coincidence…

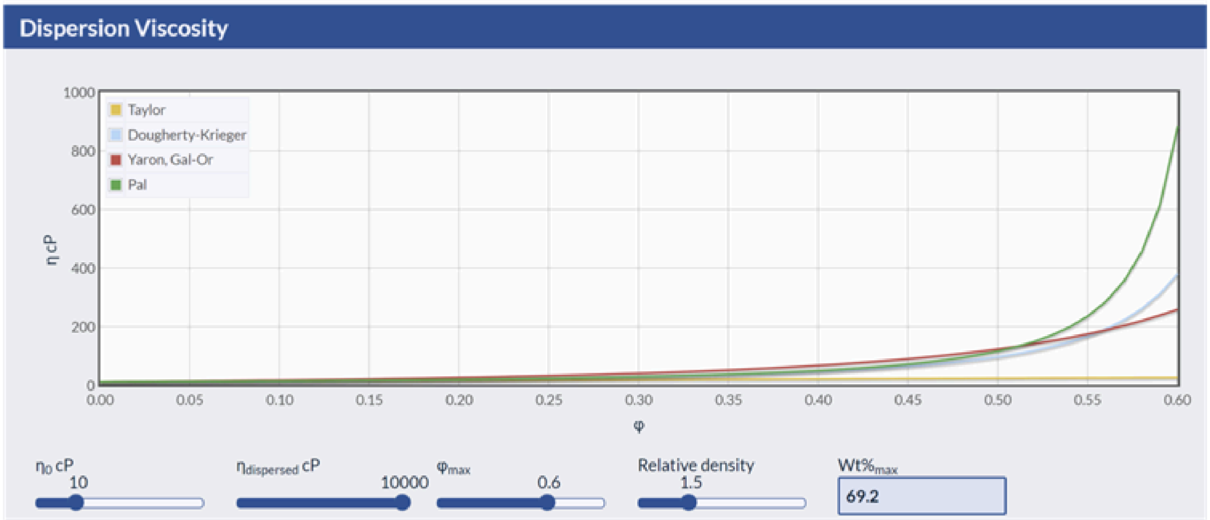
Dispersions Rheology (Low shear)

Links

Sun Screens, Lipstick, Mascara, Water-based Adhesive, Inkjet Inks

The low shear viscosity behaviour of particle formulations is described by a number of theories which give similar results up to a volume fraction φ of ~ 0.5. As the particles get very crowded, the theories diverge, but at that stage it is likely that your particles will be showing other behaviours, described in the Dispersions_Rheology (High shear) chapter.

It’s relative



η_0 cP 10

$\eta_{dispersed}$ cP 10000

φ_{max} 0.6

Relative density 1.5

Wt%$_{max}$ 69.2

https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

In this example, the solution without particles had a viscosity of 10 cP. At a volume fraction of 0.3 the viscosity has increased to ~ 35 cP because the particles bashing into each other dissipate energy. At 0.5 the viscosity is over 100 cP, a 10x increase.

At first this seems to indicate that if our particles are spherical and well-behaved, we don’t have to worry about viscous effects till we get closer to the close packing limits around 0.6. For most of us, the difference between 10 cP and 100 cP is not so worrying.

But now increase the solution viscosity to 20 cP. That’s only an extra 10 cP so you might think that the viscosity at φ = 0.5 will increase from 100 to 110 cP. But that’s not what happens. Try it yourself. You see that it goes from 100 to 200 cP. The effects of the particles are *relative*, so although 10 to 20 cP is only a 10 cP increase, it’s a doubling of that low viscosity. So the particle viscosity is also doubled.

The fact that an insignificant change to a base formulation viscosity can lead to a large change in the final viscosity is both bad and good news. It’s bad because some necessary new additive to the starting formulation can give an insignificant increase to the base formulation which is dramatic for the final formulation. It’s good because if you are hitting problems with your formulation and can’t go to a lower φ, then some modest consideration of your starting viscosity might solve the problem.

I once got asked at short notice to solve a formulation crisis involving a particulate coating. After much effort it became clear that the formulation was failing because its viscosity was slightly too high. Unfortunately, simple fixes such as altering the particle concentration were unacceptable. I'd only recently programmed the app and had never properly believed that a 1 cP shift from 10 cP to 9 cP in the base viscosity could give a 10% reduction in the final formulation. It turned out to be easy to make this reduction, keeping everything else constant … and the problem was fixed.

The formulae

The calculated viscosity is always relative to η_0, the viscosity of the bulk phase.

The app covers particles from oil emulsions through soft particles to hard particles. The ratio of viscosity of the particle to the medium is k which for solid particles is "very large" which means > 10000 which is all that's required for the formulae.

You can read about the monstrous Yaron, Gal-Or equation and the complexities of the Pal formula in the app. The very low φ formula from Taylor is included not because it's of use but because it's often quoted. It is basically the Einstein formula with a correction for k which, for large k comes out as 2.5φ:

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

The most common formula is Dougherty-Krieger:

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

The value of φ_m, the critical packing fraction is assumed to be 0.68 in the app. Others say it should be 0.58 which gives a much more dramatic take-off at higher φ values. The point is to not worry too much about the formula but to get the general idea that bad things will start to happen above φ = 0.5, which can be purely due to the formula but more likely due to particles wanting to flocculate at these high values, unless you apply high-enough shear to pull them apart. That's why we have the Dispersions_Rheology (High shear) chapter.

The number quoted as a critical packing fraction changes in different FST chapters. That's not simple carelessness. Different domains confidently quote different φ_m values and I use values common to that domain. The different values represent close packing, random packing, random close packing etc.

Particle size independence

It would make sense that smaller particles, at the same volume fraction, would show higher viscosities because they interact more strongly. The absence of a radius input therefore seems shocking. But these theories are all neutral about particle size.

To put it another way, if you see a particle size dependence in your formulations (and almost certainly you will) you know that other things are going on. The point of the app is to set a baseline expectation. If your formulation is close to this then you have the good news that your particles are well-dispersed. Congratulations. If the formulation shows a big divergence with φ and/or particle size then you automatically know that particle-particle interactions are significant – so you can focus your formulation energies on improving the dispersant system.

Dispersions Scheutjens-Fleer

Links

The most powerful and useful dispersion theory describing how particles and polymers interact was developed in the 1970s by Scheutjens and Fleer. But outside academia it has hardly been used. Now it's available in an app, there's no excuse not to use it.

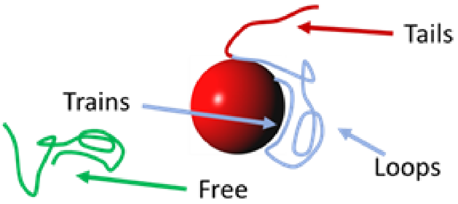
Deep solubility theory

It is unfashionable to apply solubility theory to particles (see Dispersions_Solubility) but the powerful capabilities of Scheutjens-Fleer theory show that fashion is wrong.

The theory starts with the well-known fact (Dissolution_Flory-Huggins) that the Flory-Huggins χ parameter based on simple lattice theory (https://www.stevenabbott.co.uk/practical-solubility/lattice.php) tells us that a polymer in a good solvent ($\chi < 0.5$) is extended while in a bad solvent ($\chi > 0.5$) is curled up on itself because the polymer-polymer segment interactions and solvent-solvent interactons are stronger than polymer-solvent interactions.

Using exactly the same logic you can get a χ parameter for particle-solvent interactions and for particle-polymer-solvent interactions.

Now put your particles, solvent and polymer into a virtual box and make some guess about where they might be. For example, the guess might have the polymer as mostly free, mostly as trains, mostly as loops or mostly as tails:



Via lattice theory, all the interactions can be calculated via the various χ parameters. This gives an energy value for this specific configuration. Now make some guesses about how to move some of the polymer chains. Re-do the calculation and if the energy is lower, carry on with those sorts of moves … and so on till you get a Self-consistent Field result (SF for self-consistent field theory).

Now you know what the polymer is doing. Let's see the results:

Type Polymer A
NA 100
NB 100
φ Particle 0.1
φ Polymer A 0.1
φ Polymer B 0.1
χA-->Solvent 0.25
χsA-->Particle -1.5
χB-->Solvent 0.25
χsB-->Particle -1.5
χA-B 0.4
NCombs 5
φ Grafted 0.015
NLayers 20
θ 1.80
Rg (# / nm) 4.1 / 1.2
φ Bulk 3.58e-2
Z-Max 7

https://www.stevenabbott.co.uk/practical-solubility/SF-Theory.php

What we see is that most of the polymer (it is 100 “lattice units” long, which is about 30nm) is concentrated in the first 2 lattice units (0.6nm) near the surface. The blue line shows that these are mostly trains and loops. The tails, the things we usually draw as dominating the situation, just about poke out to 6 units, 1.8nm. Finally, towards the right-hand side of the graph we see the free polymer.

Now replace this polymer with a simple di-block, 50 units each of A and B, so the same length as the previous system:

Type Di-Block (A-B)
NA 50
NB 51
φ Particle 0.1
φ Polymer A 0.1
φ Polymer B 0.1
χA-->Solvent 0.55
χsA-->Particle -1.5
χB-->Solvent 0.25
χsB-->Particle 0
χA-B 0.55
NCombs 5
φ Grafted 0.015
NLayers 30
θ 2.70
Rg (# / nm) 6.1 / 1.8
φ Bulk 2.51e-2
Z-Max 7

The A block has a high χ with respect to the solvent, so, not surprisingly, we find it (blue curve) near the particle. The B block (red curve) extends into the solvent. There is basically no free polymer, which means this will be an efficient dispersing agent.

Unlike the other apps, the engine powering these calculations is sitting on the server. It is called SFBox, developed over the years at U Wageningen and generously supplied by Prof F Leermakers.

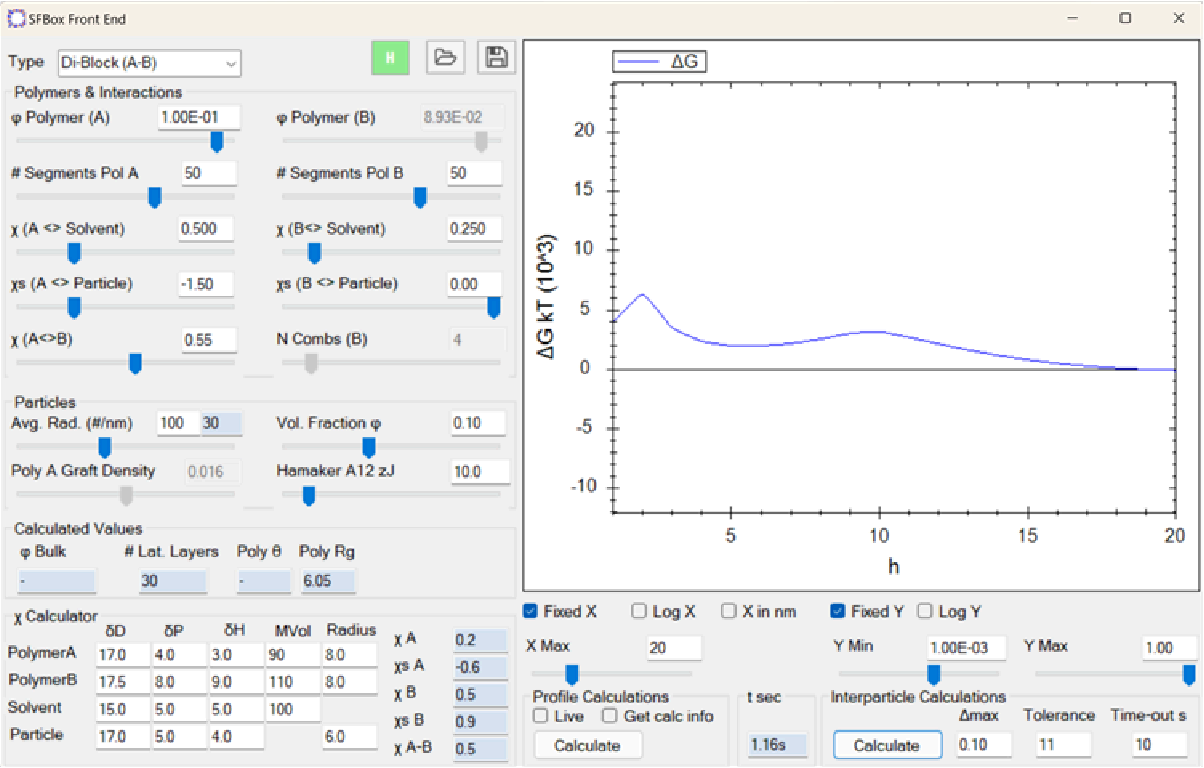
> Some top scientists choose not to waste their time talking to people like me who can't just read their published papers and translate them into glorious apps. There are, however, a rare few who have great theories combined with the patience/understanding to realise that not everyone knows as much about their area as they do. A number of the apps used in FST exist only because of such exceptional academics. The SF app is one of them.

You can explore other options, play with χ values and build your intuition about particles and polymers.

Particle stability

The full power of SF theory comes from calculation of particle-particle interactions involving polymers. You can set up scenarios such as the above, calculate, first, the configuration of polymers around the particles, then calculate whether those configurations are stable against flocculation.

Unfortunately those calculations are too complex for a web app so are only available conveniently within the HSPiP software package.



Type Di-Block (A-B)

Polymers & Interactions
φ Polymer (A) 1.00E-01
φ Polymer (B) 8.93E-02
Segments Pol A 50
Segments Pol B 50
χ (A <> Solvent) 0.500
χ (B<> Solvent) 0.250
χs (A <> Particle) -1.50
χs (B <> Particle) 0.00
χ (A<>B) 0.55
N Combs (B) 4

Particles
Avg. Rad. (#/nm) 100 30
Vol. Fraction φ 0.10
Poly A Graft Density 0.016
Hamaker A12 zJ 10.0

Calculated Values
φ Bulk -
Lat. Layers 30
Poly θ -
Poly Rg 6.05

χ Calculator

	δD	δP	δH	MVol	Radius
PolymerA	17.0	4.0	3.0	90	8.0
PolymerB	17.5	8.0	9.0	110	8.0
Solvent	15.0	5.0	5.0	100	
Particle	17.0	5.0	4.0		6.0

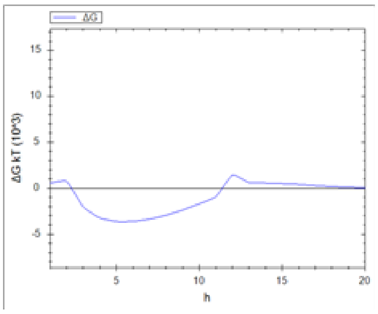
χ A 0.2
χs A -0.6
χ B 0.5
χs B 0.9
χ A-B 0.5

☑ Fixed X ☐ Log X ☐ X in nm ☑ Fixed Y ☐ Log Y
X Max 20
Y Min 1.00E-03
Y Max 1.00

Profile Calculations
☐ Live ☐ Get calc info
t sec 1.16s

Interparticle Calculations
Δmax 0.10
Tolerance 11
Time-out s 10

Using the same Di-block conditions, it is obvious from the interparticle calculation shown here that there is a very high barrier extending out to 15 lattice units. This would be an excellent dispersing agent.



Change the χ value for polymer B to greater than 0.5 and we see that the particles are going to strongly flocculate. The B polymer chains now like each other more than they like the solvent, so they give the bridging flocculation mentioned in the Dispersions_DLVO chapter. This is a more powerful way to look at these problems than the simplistic theories we have previously been offered. You can explore not just the χ parameter but the effects of relative sizes of the A and B blocks, the effects of free polymers, compare simple di-blocks to brush polymer equivalents, and much more.

The future for SF theory

Now that SF theory is available both as an app and as a powerful feature of HSPiP there are few excuses for staying with handwaving "dispersion" or "colloid" science arguments. SF theory isn't easy, but the basic ideas of interactions via sets of χ parameters are not hard to grasp. The more of us who use it routinely, the more we will demand relevant parameters from our suppliers, the academic community and ourselves. It is very much a theory that should be in the toolkit of many formulators.

> I had been entirely unaware of SF theory and found it by accident. It is tragic that everyone knows DLVO, which is so limited, and so few know SF theory which is so powerful. An obvious theme of the FST is that much of what we do in formulation work is based on theories that are totally inadequate, and which should be kicked out to be replaced by better ones.

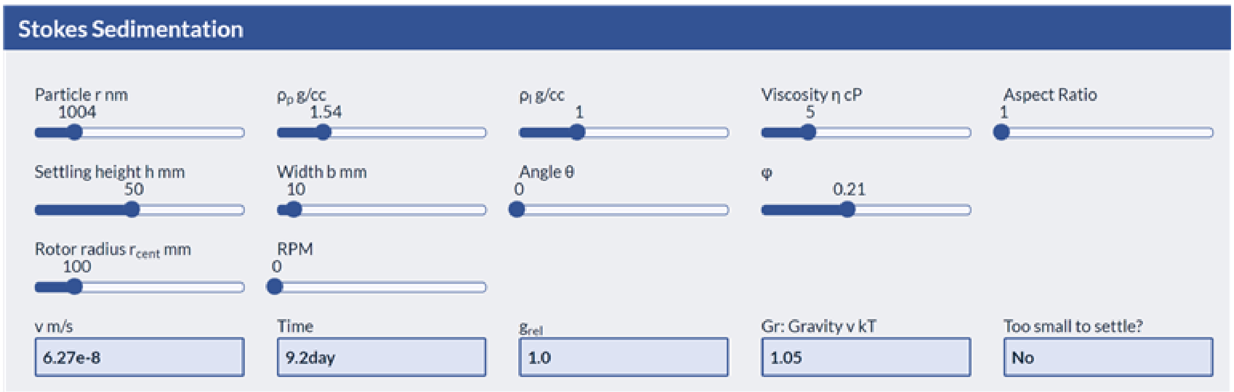
Dispersions Settling

Links

[Lipstick](), [Mascara](), [Water-based Adhesive](), [Microencapsulation]()

> If your particles are too large (perhaps from agglomeration) and are too dense for a liquid of too low a viscosity, they will settle, with a velocity governed by Stokes law, with some suitable modifications.
>
> Once they settle they might be easily re-dispersed or might, because they are close together, cohere permanently. Either way, avoiding settling is a good idea.

Stokes law



Stokes Sedimentation

Particle r nm 1004
ρ_p g/cc 1.54
ρ_l g/cc 1
Viscosity η cP 5
Aspect Ratio 1
Settling height h mm 50
Width b mm 10
Angle θ 0
φ 0.21
Rotor radius r_{cent} mm 100
RPM 0
v m/s 6.27e-8
Time 9.2day
g_{rel} 1.0
Gr: Gravity v kT 1.05
Too small to settle? No

https://www.stevenabbott.co.uk/practical-solubility/stokes.php

The velocity v at which particles settle depends on the radius r, the viscosity η and the density difference between particle and water, Δρ. In addition, the Richardson & Zaki dependence on volume fraction φ is included:

$$v = \frac{2gr^2\Delta\rho}{9\eta}(1-\varphi)^{5.65}$$

The time taken to settle a height h is given by:

$$t = \frac{h}{v}$$

If you are using a centrifuge of radius R spinning at an RPM of ω then you substitute g in the equation by g_{rel}:

$$g_{rel} = g(1 + 0.00118R\omega)$$

Although it is incorrect to say that some particles are too small to settle because of Brownian motion around kT, there is a useful heuristic which says that the particles won’t be observed to settle if the Gravity Number, Gr < 1:

$$Gr = \frac{0.768\pi r4^4 \Delta\rho g}{3kT}$$

The reason it is incorrect to talk of Brownian inhibition of settling is covered in the next section.

One final surprise about settling. If liquid of height H is tilted in a tube of width b by an angle θ then the Boycott velocity (he discovered the effect in the 1920’s) is given by:

$$v_b = v\left(1 + \frac{H}{b} sin\theta\right)$$

This significantly speeds up settling, as Boycott discovered when attempting to settle out red blood cells in small test tubes. For example, a 5 cm liquid in a tube 1 cm wide tilted at 30° increases settling by 3.5x, as you can confirm with the sliders in the app.

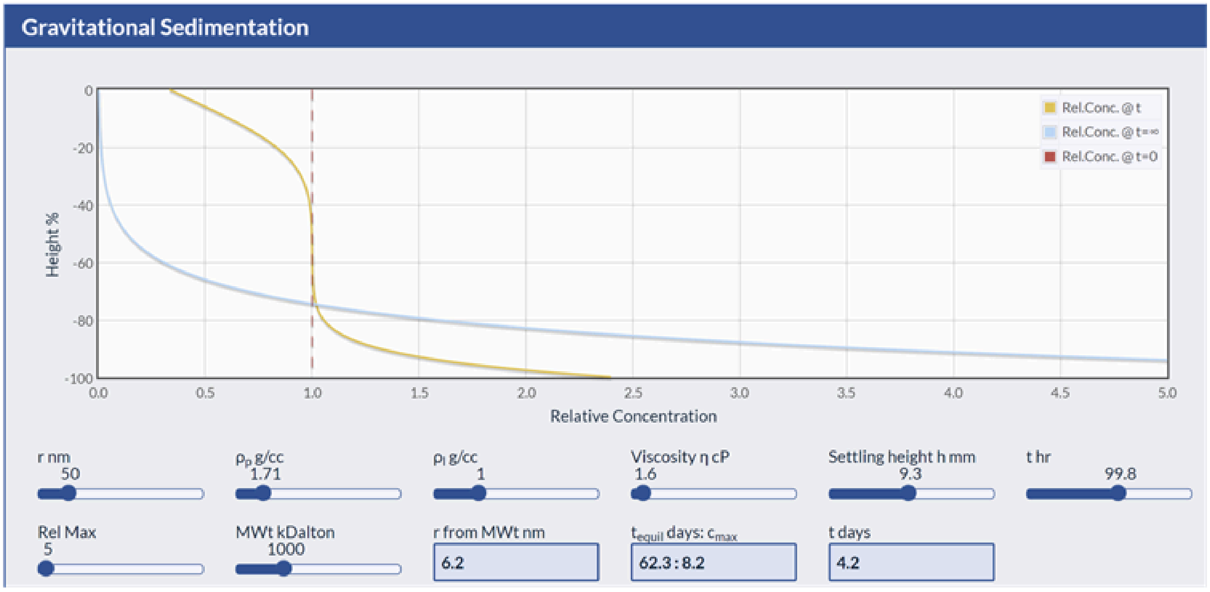
Like most people, I’d never heard of the Boycott effect. After a post about an improvement to a settling app, someone commented that it would be good to add the Boycott effect. It was great fun to learn of something completely new, to find the theory and to share it with many others for whom it was also new.

Gravitational Sedimentation

Everything, not just particles but molecules and polymers, can settle over time if there is a density difference. We know this from the common practice of putting biological samples into a centrifuge (large DNA or proteins) or ultracentrifuge (e.g. smaller proteins).

Instead of aiming for “settled” or “not settled” which is the typical focus of Stokes sedimentation, we can look for a concentration gradient of settled material.

If we just use gravity, then there is an algorithm from the 1920’s by Mason and Weaver which allows us to calculate the concentration profile from top to bottom of the tube. Solving the algorithm turns out to be very hard, but the Werts group at Ecole normale supérieure de Rennes found a way to implement the algorithm across all cases of densities and sizes.



r nm 50

ρ_p g/cc 1.71

ρ_l g/cc 1

Viscosity η cP 1.6

Settling height h mm 9.3

t hr 99.8

Rel Max 5

MWt kDalton 1000

r from MWt nm 6.2

t_{equil} days: c_{max} 62.3 : 8.2

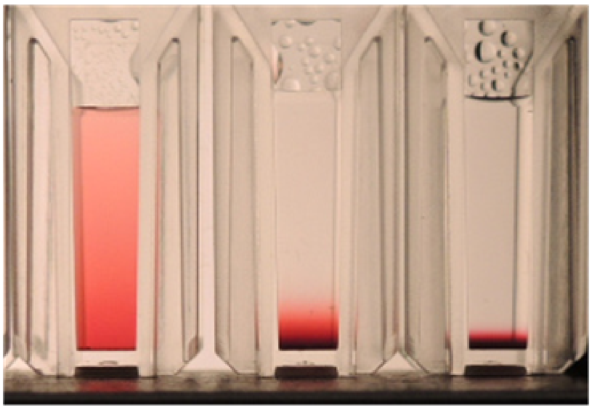
t days 4.2

https://www.stevenabbott.co.uk/practical-solubility/Gravitational-Sedimentation.php

The image shows the relative concentration profile (yellow) going from small (~0.35x) near the top to high (~2.5x) near the bottom with a middle section that retains the original concentration. This happens to be after 4.2 days and this sample would take 62 days to reach the equilibrium profile shown in blue.

There are a few reasons to be interested in this app:

- The myth that “dispersions settle but solutions don’t settle” is firmly busted.
- There’s a difference between particles settling slowly like this and particles flocculating and then (because they are much bigger) crashing out.
- The “Too small to settle” heuristic from the previous section is wrong in theory but OK in practice. These settling profiles over days require samples to be kept undisturbed and under good temperature control as convection currents could easily undo the settling.
- You can visualise the Boltzmann distribution!



Yes, the Wert group, who kindly provided this image, showed that the concentration gradient of these gold nanoparticles at equilibrium matches the Boltzmann distribution and because the particles are strongly coloured, you can easily see the distribution.

Dispersions Solubility

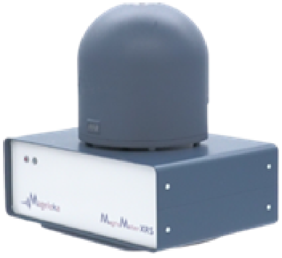
Links

Dispersions can't be "soluble" – they're dispersions! But because there is no agreed definition of "dispersion" we can't use "dispersion theory". However, there is an agreed statistical thermodynamics definition of "soluble" and nanoparticles happily fit that definition. And it turns out that if you use a pragmatic solubility theory, you can do all sorts of smart things with particle formulations.

Particles can be soluble

Sure, many particles are insoluble in many solvents, but the same can be said of many crystalline organic chemicals, of polymers and biomolecules. Just as you need the right solvent for those molecules we think of as soluble, you need the right solvent for particles.

Take samples containing the same small amount of particles and put them into 10 different solvents in centrifuge tubes. After a bit of dispersion, centrifuge the tubes. In some solvents, the particles will rapidly go to the bottom. In others, they are slow to sediment. Correcting for viscosity and relative density, you can produce a set of relative sedimentation times, RST. Those with a long RST are "happy" in the solvent. Those with a short RST are "unhappy".



Do the same thing with the particles in NMR tubes and place them not into a fancy NMR spectrometer, but a table-top device (image courtesy of Mageleka) to measure T1/T2 relaxation rates.

Those solvents that interact strongly with the surface, and which are "good" solvents show a big reduction in relaxation times compared to "bad" solvents which only weakly interact. These are not subtle effects requiring complex analyses – you get measurements in seconds.

Or just shake up your formulation with 10 different solvents and use your own judgement to see if the particles are "happy" or "unhappy".

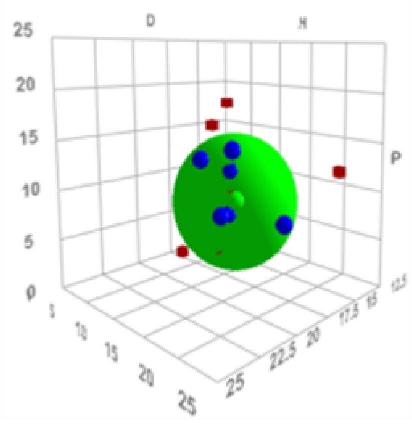
Whichever technique you use, the results are a set of good and bad solvents. Each technique has its strengths and weaknesses (and equipment costs), but they all lead to the same conclusion: *that you can apply solubility theory to the particles and, therefore, formulate intelligently*.

Independently of these pragmatic efforts, fundamental work by Shimizu and coworkers demonstrated, via assumption-free statistical thermodynamics, that "solubility" is indeed applicable to nanoparticles. Yes, for obvious reasons they are less soluble than small molecules or medium-sized polymers, but they are still soluble. There is no other scientifically coherent way to talk about their behaviour in solvents.

HSP of particles

The chapter on Dissolution_Hansen Solubility Parameters shows that you can characterise a polymer or particle via the three HSP values, δD, dispersion, δP, polar and δH, H-bonding. You do this by measuring a sphere that contains the good solvents and excludes the bad solvents. The best solvent for the polymer or particle is at the centre of the sphere. The further from the centre, the worse the solvent. This Distance is a measure of compatibility. You can calculate the Distance from any other solvent, polymer or particle and therefore have a good idea of the compatibility or incompatibility.

Here, for example, is a sphere of a specific carbon black, CB. This used values of RST from a LUMiSizer but similar results could have been obtained via NMR relaxation:



Its HSP are [17.2, 8.9, 13.1], meaning that it has quite high Dispersion (typical for CBs) is medium Polar and quite high H-bonding. If you happen to know that this batch of CB is great for making your Li batteries and the next batch came in at [17.1, 9.3, 12.9] you could be confident to use the next batch. But if it came in at [17.9, 3.2, 4.8] (measured values for a "hydrophobic" CB) you would know not to use it.

The correct HSP values will match nicely to the HSP of your chosen solvents and the polymers and other ingredients you are using. Sometimes you want to encourage phase separation during coating, so you *must* ensure significant differences in the HSP of the phases you need to separate.

Without these solubility-based measurements, it is extraordinarily hard to characterise a batch of CB (or any other particle) in terms that interest you. How else can you know its compatibility with your solvents and other ingredients?

You cannot get these insights via common analytical techniques such as AFM or SEM. Sorption techniques such as IGC give information about particles surrounded by low concentrations of solvent in the inert chromatographic gas; they cannot provide information about bulk solubility. Of course these techniques provide valuable other information, but in terms of formulation, having direct solubility measurements is important.

Not long after creating HSP for polymers, Dr Hansen was working in a paint company. He thought it worth trying to apply his solubility theory to paint pigments. If it worked, he'd have a powerful formulation tool, if it failed he'd learn something interesting about pigments. It worked. That was in the 1970s. It is only in the past decade that formulators started to realise that "dispersion science" offered little of practical value while solubility theory via HSP was a powerful methodology that delivers results.

Surface dependence

The theory is entirely neutral about the surface of the particle. Are the HSP values dominated by the general particle surface, specific "high energy sites" on the surface, polymers, contaminants, dispersants? The theory doesn't care. It pragmatically finds that this sample behaves with these solubility parameters, so if you want to formulate successfully, use those parameters. If you think that a new batch of the particles is different, just re-measure the good/bad solvents and calculate a new sphere.

When we add dispersants to the particles, what happens to solubility theory?

Again, the theory is neutral. If you deliberately change the particles via, say, a new dispersant, re-measuring the HSP tells you a lot about the success or failure of your change to get you into the solubility/stability zone you require.

The most powerful theory that takes into account the dispersants we tend to add to our particles is Dispersions_Scheutjens-Fleer which has been around since the 1970s, but little used because it was hard to implement. It is based on lattice theory, Flory-Huggins χ values and, therefore, solubility theory.

But particles are dispersed, not soluble!

Some people, and this might include you, refuse on principle to accept that particles can be analyzed via solubility theory. Particles are dispersed, therefore one should use dispersion theories. But what theories are those? Apart from Dispersions_DLVO which is really only useful in water (where HSP is weak), there is very little usable dispersion theory out there. Plenty of handwaving and good intentions. But very little usable theory. The conventional extension of DLVO to steric effects (described in the DLVO app) is OK, but cannot cope with effects of free or added polymer. These effects, plus effects such as depletion flocculation are better analyzed via Scheutjens-Fleer theory than via the relatively unhelpful Asakura & Oosawa model which is often presented as a key element of dispersion or colloidal science: https://www.stevenabbott.co.uk/practical-solubility/depletion.php.

Happily, the use of HSP for nanoparticles has reached critical mass. There are now committees working on ASTM-style standards for measuring the HSP of CB for battery manufacture, and “round robin” groups exist comparing HSP values obtained by different techniques (e.g. visual tests, centrifuge tests, NMR relaxation). These efforts to standardise tests and work out the advantages/disadvantages of the various techniques are arriving just as nanoparticle suppliers are receiving demands from their customers for the HSP of the particles.

So, yes, nanoparticles are soluble, and solubility theory is a powerful tool for nanoparticle formulators.

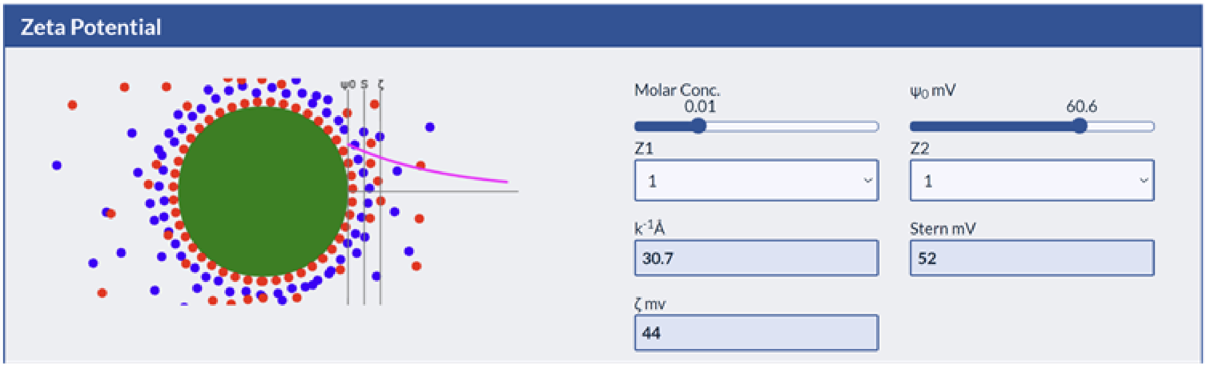
Dispersions Zeta potential

Links

Water-based Adhesive, Emulsion Paint, Emulsion Polymers, Inkjet Inks

Classic DLVO theory Dispersions_DLVO tells us that in water, particles can be stabilized by charge-charge repulsion. To know what these charges might be, we measure the zeta potential. As we shall see, it's a slippery concept and a wise rule is to never measure *a* zeta potential but to measure a *set* of them.

Real charge and measured charge



Zeta Potential

Molar Conc. 0.01

ψ0 mV 60.6

Z1: 1

Z2: 1

k^{-1}Å: 30.7

Stern mV: 52

ζ mv: 44

https://www.stevenabbott.co.uk/practical-solubility/zeta.php

We have a particle, green, with a "real" charge of 61 mV, shown by the red ring. The first thing a charge does is attract its counterion, blue, so the net charge must be less than 61mV. But that loose, blue ring attracts red ions and so forth till we reach ζ, the boundary between charges fixed to the particle and the general background solution. The measured voltage at this point is "the zeta potential", in this case 44 mV Along the way we pass through S, the Stern layer (52 mV) which some classify as the end of the original double layer, though there is much confusion about the relation between S and ζ.

If you increase the molar concentration of background salts then the Debye length, k^{-1}, decreases and the potential away from the particle (including the zeta potential) decreases.

If you put these particles and their background salts into a zeta measuring device (one such device is called a Zetasizer) then from the light scattering behaviour under an applied voltage gradient, "the" zeta potential is measured. Exactly what it is and what it means is less important than the fact that its sign and its general magnitude tells us a lot about the particle. The sign can typically switch from negative at high pH to positive at low pH, passing through the isoelectric point of zero charge and, therefore, the most perilous state in terms of charge stabilization. The magnitude can be roughly classed as "small" (less than 10 mV), unlikely to show any charge stability, "medium" (up to 25 mV), with borderline stability and large (40 mV is a comfortable value) for a "stable" system.

Never measure *a* zeta potential

Because of the messy nature of “the” zeta potential, surprisingly small changes in the formulation can result in surprisingly large changes in the potential. The “real” surface can change dramatically with small changes to the surface chemistry. Different counterions might have very different interactions within the Stern layer. Small levels of contaminants (such as cationic surfactants interacting with anionic surfaces) can have devastating effects. And, of course, pH changes might take you close to the dangerous isoelectric point.

That’s why measuring *a* zeta potential is so dangerous. The trick is to measure the potential over a range of likely conditions and possible contaminants. You will find a safe formulation zone and also know what things must be strongly avoided and therefore checked for before adding to the formulation.

> We were using a dispersion stabilised by a large (negative) zeta potential; it had always worked well. But we encountered some other problem and a special surfactant was suggested to fix the problem. Like all surfactants it just had a name, say, XY95. Only small amounts were required, so there were few risks of side-effects. On adding this tiny amount of XY95, the formulation crashed out. Digging into the surfactant datasheet we found that it was cationic. Even the small amounts that locked on to the anionic shell of the dispersion were enough to change the effective zeta potential to near zero, and the formulation failed.

And if you can’t find a stable zeta potential zone, then swap your effort over to steric stabilization as described in the Dispersions_DLVO chapter or even better in the chapter on Dispersions_Scheutjens-Fleer.

Dissolution Crystallization

Links

[Pharma Formulation]

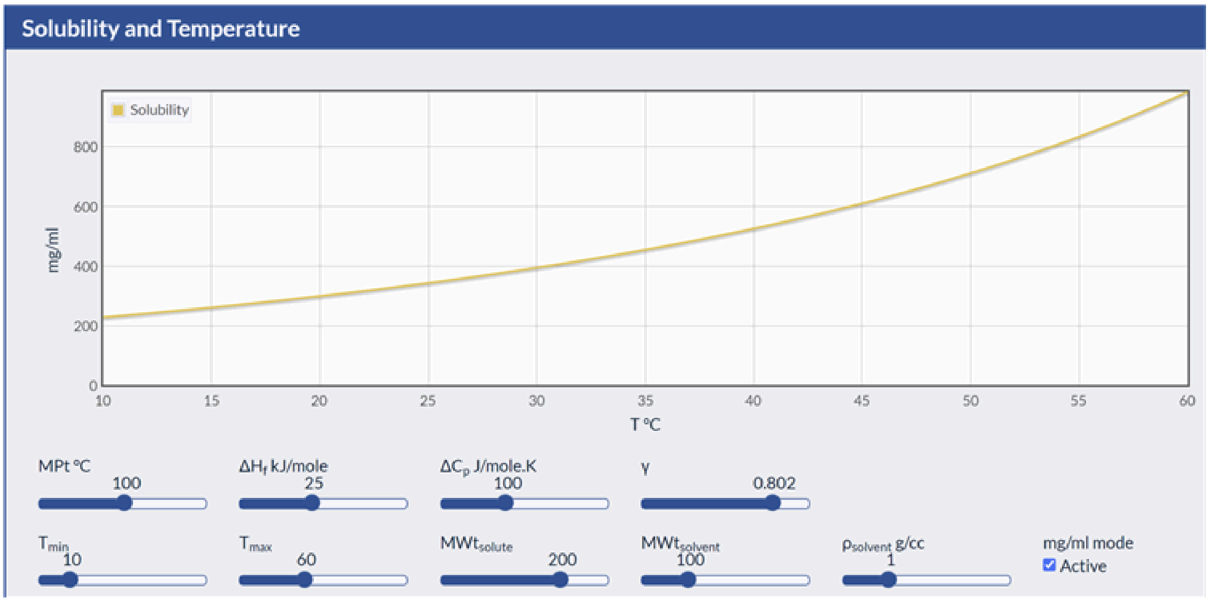
Crystallization is both vital and frustrating. It’s vital for obtaining pure compounds and frustrating because things don’t crystallize the way we want, when we want.

It hasn’t helped that for decades people have tried to use CNT (Crystal Nucleation Theory) despite everyone agreeing that it’s unfit for purpose. Some brief suggestions for a better approach are included here.

Solubility versus Temperature

For easy, high efficiency, crystallization you want a nice solubility curve – a large solubility at a temperature OK for handling your chosen solvent, and a low solubility at whatever temperature (say 20°C or 0°C) you can conveniently crystallize and filter.

Via a Differential Scanning Calorimeter (DSC) it’s not hard to measure the MPt and ΔHf, the melting point and enthalpy of fusion. If you use Dissolution_Ideal solubility theory you can then find how solubility in a perfect solvent changes over your temperature range. If, via COSMO-RS or Dissolution_Hansen Solubility Parameters you have an idea of the activity coefficient, γ, of the solute in any real solvent, then you can get a basic estimate of the sort of curve you can expect:



MPt °C 100

ΔH_f kJ/mole 25

ΔC_p J/mole.K 100

γ 0.802

T_{min} 10

T_{max} 60

MWt_{solute} 200

$MWt_{solvent}$ 100

$\rho_{solvent}$ g/cc 1

mg/ml mode ☑ Active

https://www.stevenabbott.co.uk/practical-solubility/Solubility-T.php

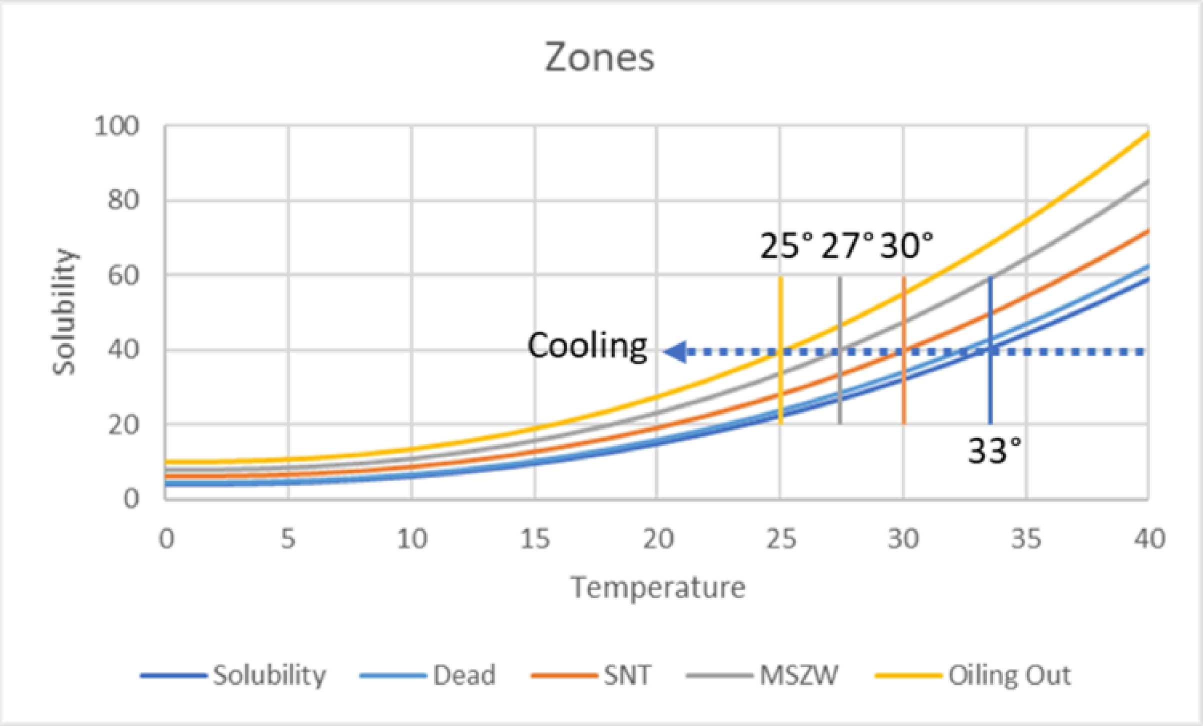
In this example, with a relatively low MPt = 100°C, then over the range 10 - 60°C there’s a 5:1 solubility ratio, but with a relatively large quantity of 200 mg/ml remaining in solution.

If you change the MPt to 200°C then you get only a 2:1 ratio, with 150 mg/ml remaining.

This information doesn't tell you if you can get good crystals – instead it's giving you an idea of the viability of the crystallization process given your molecule and the sorts of solvents (with their respective γ values) and temperature ranges that you might be prepared to use.

Real solubilities depend on more than the ideal solubility effect. The causes are subtle and it needs powerful tools such as COSMO-RS to model them. If you don't have access to such a package, use the ideal solubility calculations as a starting point.

Zones



Forgetting about units and focussing on principles, in this graphic we see that as T decreases from 40° to 33° you reach the solubility limit of this molecule. If you stayed there all day, nothing would happen. Indeed if you went to slightly lower you'd still be in the "Dead" zone and nothing would happen.

If, for whatever reason, some seeds (junk, old crystals, scratches on the walls …) are present then at 30° you might start to see some crystals – this is the Secondary Nucleation Threshold, SNT.

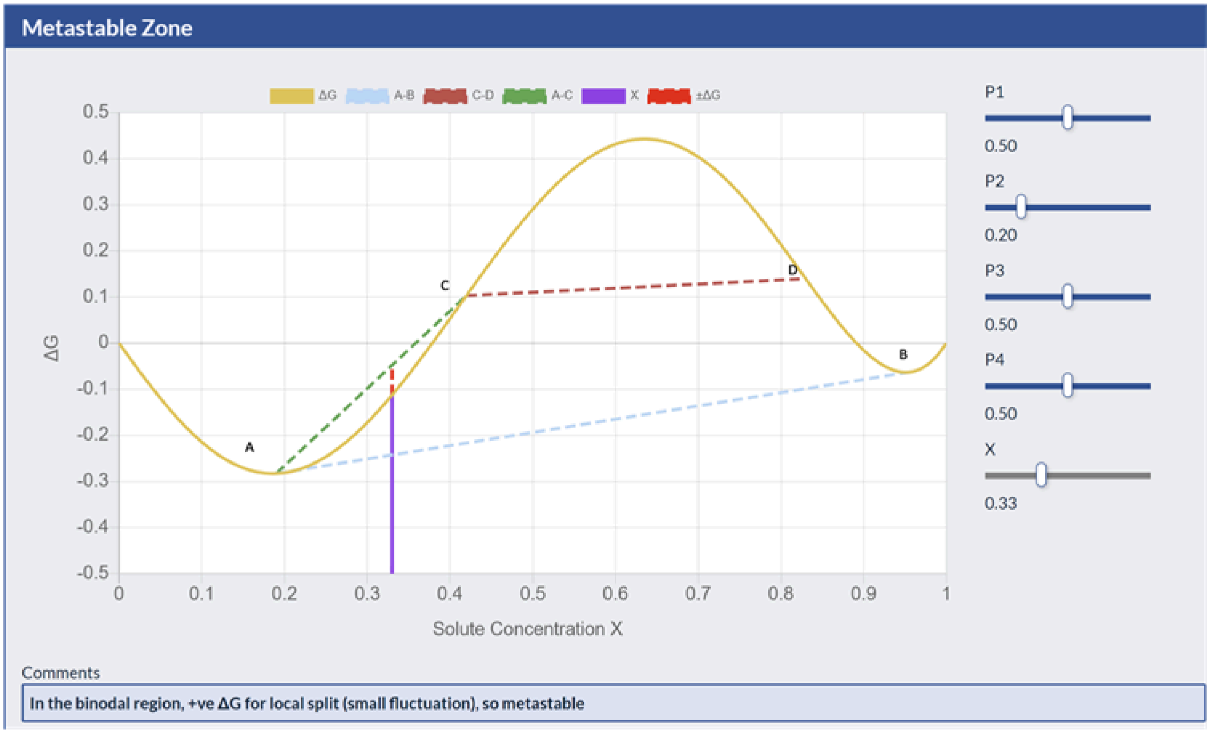
If everything were perfectly clean you might have to go to 27° to get anything to happen – you'd been through the Metastable Zone Width of 6°.

If you were impatient and cooled rapidly to 25° you might just get some oil to appear.

This story is highly idealized and contains some confused and confusing ideas, but it's the standard story so we have to tell it.

Metastability

Metastability leading to sudden (Spinodal) oiling out can be described via a straightforward theory:



P1 0.50

P2 0.20

P3 0.50

P4 0.50

X 0.33

Comments

In the binodal region, +ve ΔG for local split (small fluctuation), so metastable

https://www.stevenabbott.co.uk/practical-solubility/Metastable-Zone.php

You define a free energy landscape with 4 sliders (P1…P4) and you slide X to learn what the free energy landscape means. The app is of no practical use – it’s there for whoever wants to understand that metastability and oiling out are rational, even though highly undesirable for those who just want nice crystals.

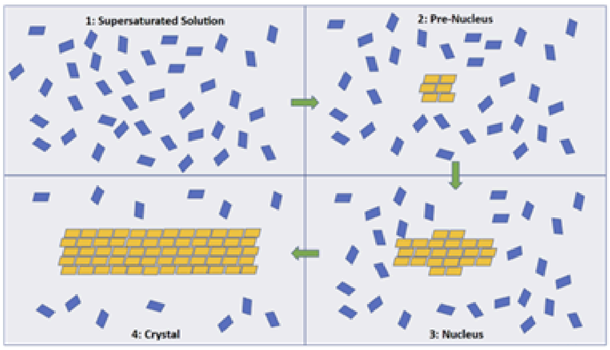
What the app stresses, however, is that when something oils out, it’s not a puzzlingly liquid version of the solute. What oils out is a fixed (low) concentration of solvent in solute, in a bath of fixed (low) concentration of solute in solvent.

CNT

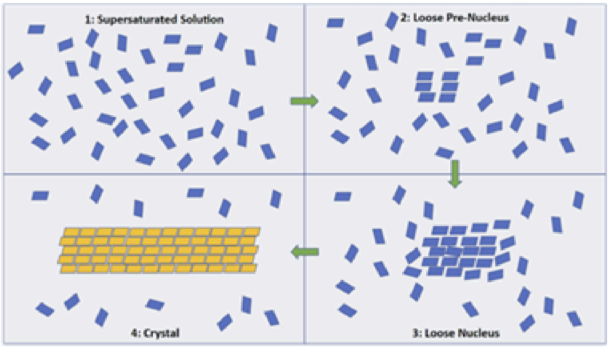
For those who want to learn about Crystal Nucleation Theory, the app https://www.stevenabbott.co.uk/practical-solubility/CNT.php describes it. Because (as it turns out) most practical crystallizations take place via seeding, CNT is completely irrelevant. The parameters that are endlessly measured and cited help no one. It’s much more important to understand and control the seeding phenomenon.

Crystallization Map

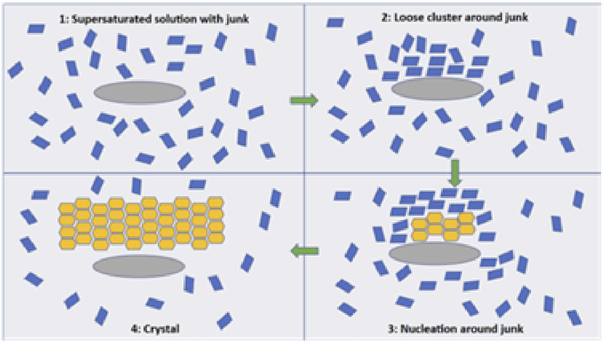
Far better, instead, to think in terms of a crystallization map where CNT is a relatively rare (it works for sodium chloride!) phenomenon. The following 4 images are from https://www.stevenabbott.co.uk/practical-solubility/Crystallization-Map.php which contains a set of 12 diagrams that cover a wider range of crystallization options



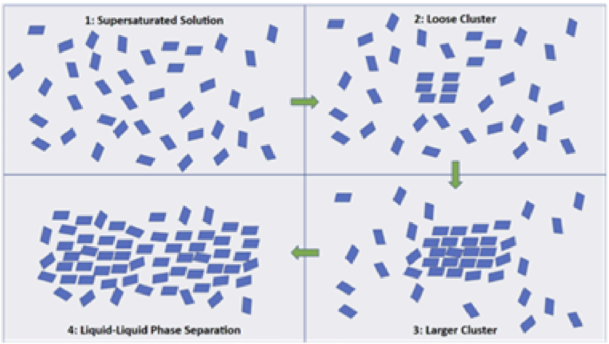
CNT



Loose CNT



Heteronucleation



Oiling Out

CNT itself says that a minimum viable crystal must first form, then everything happens after that. Loose CNT says that you can get loose clusters of molecules that can tip over into being crystals.

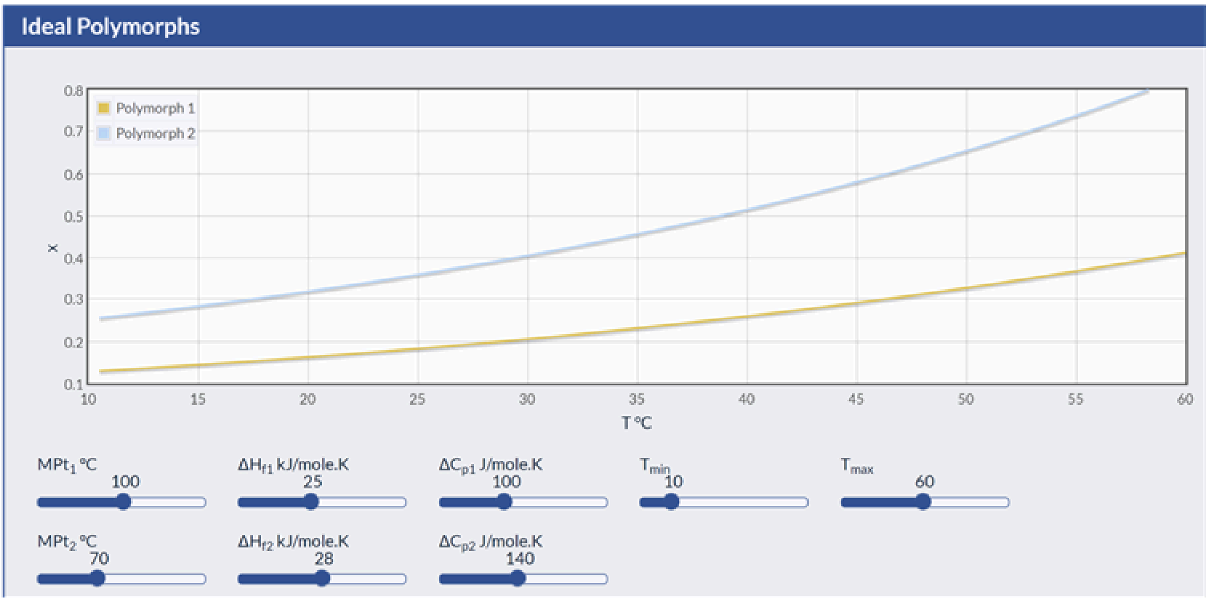
Heteronucleation encourages a loose cluster to form which in turn can change to a smaller crystal then grow to a larger one.

Finally, those loose clusters can become bigger and bigger till you get the liquid-liquid phase separation of a solvent-in-solute cluster in a sea of solute-in-solvent.

There are many reasons for preferring the cluster-based view over CNT.

- The assumptions behind CNT almost never apply
- The experiments to measure CNT values are tedious and unhelpful
- Clusters explain many otherwise weird phenomena such as sudden crystallization, selective enantiomorphs, memory phenomena around the solubility limit
- Clusters are a natural part of Dissolution_Kirkwood-Buff which naturally explains multiple solubility and solubilization phenomena
- Using cluster-sensitive techniques such as scattering will provide more molecular insight than CNT experiments.

Polymorphs



MPt_1 °C 100

ΔH_{f1} kJ/mole.K 25

ΔC_{p1} J/mole.K 100

T_{min} 10

T_{max} 60

MPt_2 °C 70

ΔH_{f2} kJ/mole.K 28

ΔC_{p2} J/mole.K 140

https://www.stevenabbott.co.uk/practical-solubility/Ideal-Polymorphs.php

It is annoyingly the case that many molecules can crystallize out as different crystalline shapes/sizes called polymorphs. Each polymorph has its own MPt and ΔHf so they have different solubility curves, as shown in the app.

Which form crystallizes out is a complex mix of solubility curves and metastable zone widths. Disentangling these effects can be challenging, so you need an app to make sense of what can happen:

Solubility Shift II 4.0

Skew 6.0

Metastable Width 2.0

Start Concentration 8.0

Cooled To 0.80

Crystallization outcome

Lots of I

https://www.stevenabbott.co.uk/practical-solubility/Polymorphs.php

In this example, if you cool to the relative temperature shown in red, you are likely to get polymorph I. But depending on the relative “skew” of the curves and the widths of the metastable zones, different outcomes are possible. Use the app to explore the other possibilities.

The statement above that CNT is usually irrelevant because crystallization usually happens via seeds is very much supported by the (true) stories of rogue polymorphs. A pharma company which has patented and certified one specific polymorph which reliably crystallizes in their facilities world wide … finds one day that a different polymorph appears at one site which can then never produce the previous one. Through inter-site travel, the new polymorph spreads and the company is forced to re-register the new polymorph for safety and efficacy – at vast costs. If a competitor happens to have patented this newer polymorph then life is very difficult. It is hard to know how polymorph seeds can travel and disrupt old processes, but it is impossible to explain by those who insist on using CNT.

Dissolution Dissolution Kinetics

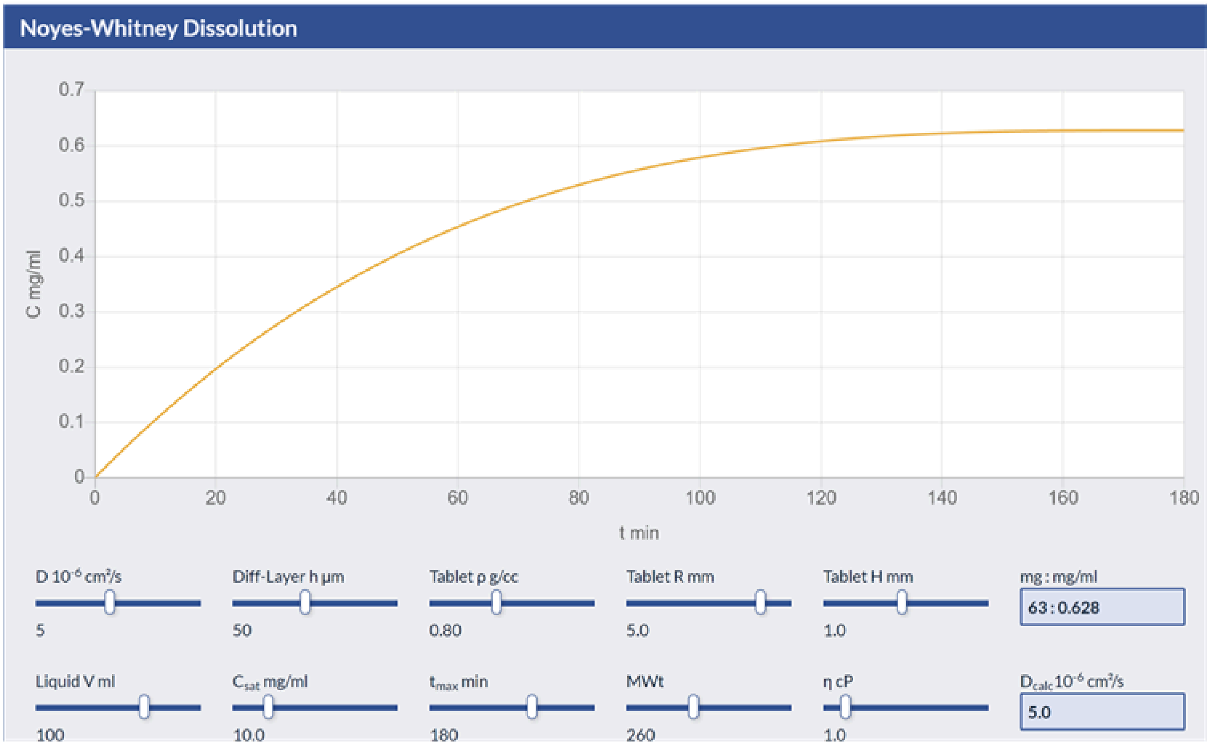
Links

Surface Cleaning, Pharma Formulation

Getting stuff to dissolve when you want, as fast as you want, can be frustrating. Here we cover some scenarios for which reasonable models exist. Take what you can from them to help with your own dissolution challenges.

Tablets and crystals

Dissolution of solids is conveniently modelled using Noyes-Whitney theory, that has been around for over a century. The app is for the specific case of a tablet, but the general principles are the same for a crystal.



$D\ 10^{-6}$ cm^2/s: 5

Diff-Layer h µm: 50

Tablet ρ g/cc: 0.80

Tablet R mm: 5.0

Tablet H mm: 1.0

mg : mg/ml: 63 : 0.628

Liquid V ml: 100

C_{sat} mg/ml: 10.0

t_{max} min: 180

MWt: 260

η cP: 1.0

$D_{calc}\ 10^{-6}$ cm^2/s: 5.0

https://www.stevenabbott.co.uk/practical-solubility/Noyes-Whitney.php

We have a tablet of radius 5mm, thickness 1mm. At a density of 0.8 g/cc that's 63mg. We are dissolving it in 100ml of water so its maximum concentration is 0.63 mg/ml, well below the saturated limit, C_{sat}, of 10 mg/ml. The solute has a diffusion coefficient of 5e^{-6} cm^2/s (you can estimate it from the MW and viscosity of the solution) and we follow the process for 180 min where the concentration levels out at the 0.63 mg/ml limit – the tablet has fully dissolved.

Everything is known in advance except for the thickness of the Diffusion Layer between the saturated solution at the surface of the tablet and the bulk solution. Something like 50µm is a reasonable starting point for an estimate. In reality you find h by fitting the experimental dissolution curve.

The theory is easy. For a surface area S, the rate of change of concentration, C, with time, t, depends on D, S, h and C_{sat}:

$$\frac{\delta C}{\delta t} = \frac{DS}{h}\left(C_{sat} - C\right)$$

This makes intuitive sense. Diffusion (per unit area) is simply (see Diffusion_Basic Diffusion) the diffusion coefficient times the concentration gradient which is the difference between the concentration at the surface and the bulk, divided by the thickness of the diffusion layer. The larger S, the greater the area for diffusions and the faster the rate.

But, of course, S decreases as the tablet dissolves and C increases; so for both reasons the rate of dissolution decreases over time. The app takes care of the details of S changing and C increasing.

Although in this example the dissolution ends long before the saturated concentration, the solubility of the solute makes a big difference to the rate; try changing C_{sat} to see the effect. This is because of the assumption that the concentration at the surface is C_{sat} so a larger value means a larger concentration gradient.

Conventionally, h is controlled by stirring. As a rule of thumb, if you increase the stirring rate by a factor of 2, h decreases by a factor of √2.

Dissolving a powder

We all know that if you add a powder (especially in the kitchen) to a solvent (water) with reasonable stirring, it can dissolve nicely. But if you add the same amount of solvent to the powder you can end up with a sticky mess that refuses to dissolve. In the world of technical polymers such as PVdF it is well-known that if you try to dissolve the PVdF powder in a good solvent (such as NMP) it is almost impossible – you get a sticky mess. If you disperse it first in a poor solvent (such as a ketone) and then add the good solvent, it's easy to dissolve.

If your powder, of radius R, doesn't form lumps, if the polymer is nicely soluble, if diffusion into the polymer is fast and if the polymer is not too entangled, then life is easy. Speed of dissolution is proportional to 1/R and to RPM, the rate of stirring.

From those simple rules, it seems obvious that finer powders will dissolve faster. But there's a problem. The gap between powder particles decreases as R decreases. Suppose that on contact with the solvent the surface of the powder absorbs the solvent and expands rapidly to form a gel. If the particles are closer than the swelling expansion then they almost instantly create a lump held together by the gelled state. If we assume that the lump has a dimension on the mm scale rather than the 0.1mm scale then we instantly have a 10x longer dissolution time. Only if dissolution of the gel is faster than its expansion can we avoid this problem.

We can explore this in the app:

Powder R µm: 150
Gel Speed: 5.0
RPM: 100
MWt/M$_c$: 2.0
t_{max} min: 120
f_{Powder}: 0.33
k: 6.00

https://www.stevenabbott.co.uk/practical-solubility/Powder-Dissolution.php

Here we have a 150µm powder which is a polymer that's somewhat entangled as MW/Mc = 2, where Mc is the Critical Entanglement MW, i.e. the MW above which the polymer chains are guaranteed to be tangled and therefore, for example, increase viscosity and decrease dissolution rates (it's hard to pull away from the tangles) (see Adhesion_Entanglement). You see a fairly rapid dissolution in the first 10 min but then things slow down. That's because we have a medium Gel Speed, meaning that the polymer swells quite quickly and you end up with a fraction of free powder, f_{Powder} = 0.33 which dissolves quickly and the remaining 0.67 fraction are lumps > 1mm so they dissolve slowly. The factor k is calculated from the RPM (higher k for higher value) and MW/Mc (lower k for higher value) and confirms what you see in the graph, the higher the value of k the faster the dissolution.

What is interesting is that although smaller powder particles dissolve more quickly (more relative surface area), if they can form rapid gels then grinding them to a smaller size can *increase* dissolution times because the gels form faster than the dissolution rate.

How do you know your "Gel Speed" and MW/Mc? The papers quoted in the app discuss various ways to calculated Gel Speed and they assume that you know both your MW and Mc. Sadly, most of the time most of us don't known Mc, which is largely a fault of the suppliers of our polymers who could easily measure the values of the products they sell.

Those familiar with entanglement theory will know of "reptation", the snake-like movement of polymer chains trapped in the tangles. We might imagine that the rate of pulling away the polymer from the powder surface would depend on the reptation timescale but interestingly it's more about the timescales of polymer chains coming out of tubes, which can be very long.

There is one well-known trick for solving the gelling problem. If you can add a non-polymeric co-powder in sufficient quantity and small size to cover a significant fraction of the polymer particles, then they interfere with the instant gelation, dissolve easily and allow the polymer to dissolve without problems. Those who like sweet cocoa drinks know this – the cocoa powder is covered with fine sugar crystals. Without the sugar, the cocoa becomes a gelled mess.

Polymer dissolution

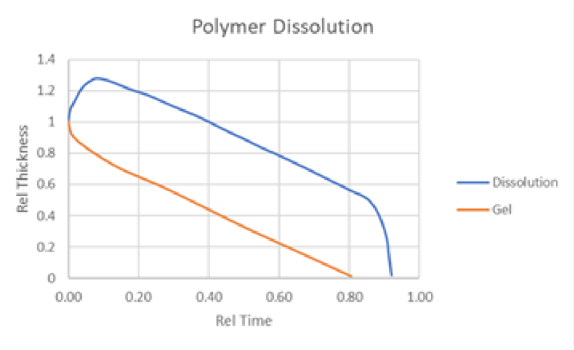
Now we have a slab or film of polymer. To get it to dissolve we must first ensure that it’s soluble in the solvent! A good HSP match Dissolution_Hansen Solubility Parameters ensures this.

Next the solvent has to diffuse into the polymer. At first this will depend on “the” diffusion coefficient, D, Diffusion_Basic Diffusion which varies strongly from polymer to polymer depending on the amount of available “free volume” through which the solvent can diffuse. To be above Tg and to have low crystallinity gives a large D with the opposite also being true.

But as solvent starts to penetrate, the amount of free volume increases so D increases. This is the phenomenon of Diffusion_Concentration-Dependent Diffusion. Once we get a solvent concentration (everything is expressed in volume fraction, φ) of something like φ=0.2 we are in a gel state with rapid diffusion. The polymer swells strongly which slightly slows down diffusion (the solvent has longer to travel) and greatly increases the difficulty of writing an app.

But even at a high φ we have generally tangled polymers in a viscous state so we haven’t swept it away – it’s not yet “dissolved”. So the issues of MW and Mc, discussed above, are relevant here too. A high MW polymer with a low critical entanglement MW will resist dissolution even though it’s well on its way to being dissolved. So the speed of flow from our mixing system will have a significant impact on the dissolution rate.

So far, a meaningful app has been difficult to provide. Instead, the general view of the process is something like the following:



We start at (relative) time 0 at (relative) thickness 1. As the solvent enters, the gel line, the point at which the polymer transitions from glassy to gel-like, moves down into the film. At the same time, the overall polymer swells so the surface rises, in this case by 30%, before dissolution erodes the swollen polymer. Once the gel phase has gone, the final dissolution is fast.

The exact shapes of these curves depend on, for example, entanglement. A more entangled polymer takes much longer to move from the surface into the solution, so the relative swelling is larger. The velocity of the solvent above the polymer has an effect, with a higher velocity removing swollen polymer sooner. The solvent quality also makes a difference as a poorer solvent will struggle to swell the polymer quickly so dissolution will be slower.

Dissolution Dissolution Rheology

Links

Surface Cleaning, Pharma Formulation

> Rheology affects dissolution, and dissolving stuff affects rheology. For the effects on dissolution, see Dissolution_Dissolution Kinetics. Here we look at how dissolving stuff affects the rheology of the final product.

Polymers and tangles

R_g @0.5+40K nm 10

M_{mono} 100

Polymer MWt 40000

Polymer M_c 10000

Flory-Huggin χ 0.25

C_{max} g/dl 15

R_g @χ+MWt nm 13.0

Ne 33.3

C* g/dl 3.0

Ce g/dl 4.1

|η| dl/g 1.67

N.bonds 400

M_{bond} 100

Bond length nm 1.4

MWt 40000

R_g nm 11.4

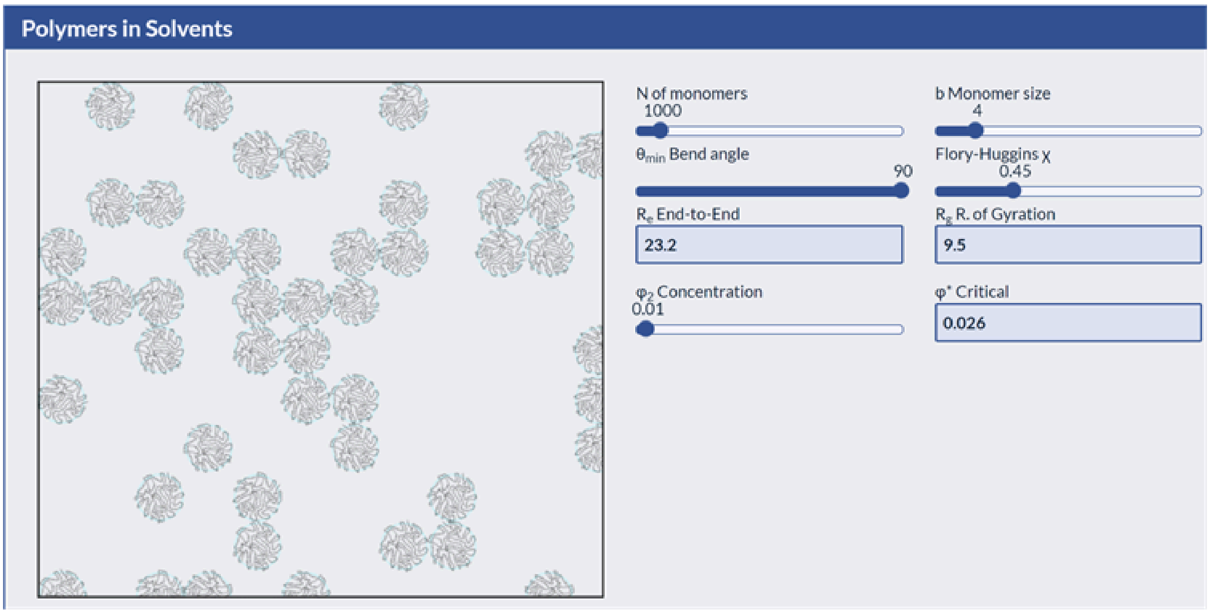
https://www.stevenabbott.co.uk/practical-solubility/polymer-viscosity.php

In general we need high concentrations of our chosen polymer for our formulation and we invariably find that the viscosity at our required concentration is just a bit too high. Or, more often, if we have a great product with the right viscosity, someone wants to use less water or solvent and our viscosity *becomes* too high.

A great way to reduce viscosity while keeping a high % solids is to use a lower MW. Some quick checks might show that the performance is unchanged, but usually a higher MW was chosen for a good reason. Polymers have an M_c, critical entanglement MW. Below M_c the polymer chains are too short to tangle. This means that their viscosity in solution is lower (good) but that their mechanical performance (e.g. as an adhesive) is worse. If your MW is, say, 10x M_c you have great mechanical properties but now the solution is too viscous because of all the tangles.

Traditional polymer science considers super-low concentrations of polymers. They are nice theories but irrelevant to our needs. For example, they like to work at concentrations, C, below C* which is when average polymer chains start touching each other. They certainly don't like to work above C_e which is when the polymer starts to be entangled and where we are likely to be working. This means that we have lots of good, but useless theory, and, above C_e we have very little. The app, therefore, is an attempt to capture some of the key themes above C_e.

The shape of a polymer



Polymers in Solvents

N of monomers 1000

b Monomer size 4

θ_{min} Bend angle 90

Flory-Huggins χ 0.45

R_e End-to-End 23.2

R_g R. of Gyration 9.5

φ_2 Concentration 0.01

φ* Critical 0.026

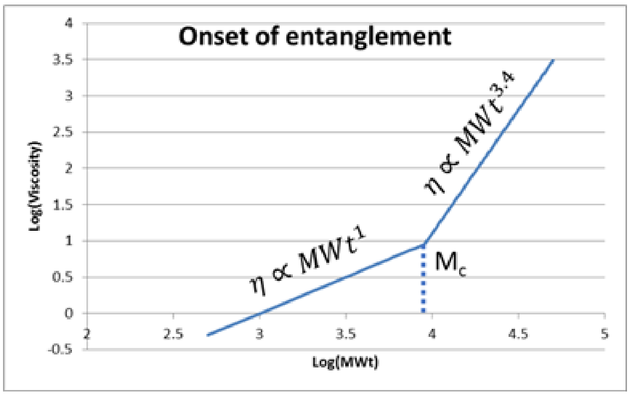
https://www.stevenabbott.co.uk/practical-solubility/polymers-in-solvents.php

As you change the Flory-Huggins χ parameter from 0 (completely compatible) to 0.5 (borderline, the θ solvent) the average polymer chain changes from nicely expanded to curled up in a ball. The app lets you look at low concentrations so you can explore the shapes with any χ value. But because we need high concentrations that are impossible when χ is large we need only confirm the obvious, that a happy polymer is nicely spread out, which means that its viscosity will be higher. There's nothing we can do about that, given that we need the high concentrations for our application.

However, we need a dilute-polymer value at $\chi = 0.5$, which is the radius of gyration, R_g, for a nominal 40K polymer. We use it to estimate the effective size for our actual MW and at our actual χ value. R_g is routinely measured by laser scattering. To make the estimate we need to know how many monomer units we have in our 40K polymer.

Knowing M_c

It should be a law that all suppliers of polymers tell us the M_c value. They aren't too hard to measure if you have (as a supplier should have) a reasonable polymer lab and can measure, say, melt viscosities at different MW.



Below M_C the viscosity increases linearly with MW, above it goes to the power of 3.4, a strange number with a good theoretical justification:

Other methods, such as https://www.stevenabbott.co.uk/practical-rheology/L-M_Mc.php are available.

The Polymer Viscosity app gives a list of M_C values for many typical polymers.

The polymer suppliers need to be persuaded to provide M_C values. It's up to you, the formulation community to demand that they provide you with the right science, so you can choose the right product.

Putting it all together

As you increase concentration beyond C_e viscosity increases rapidly. If you decrease M_C and/or increase MW, the viscosity increases faster as you have more tangles. As you increase polymer/solvent compatibility by reducing χ, viscosity also increases as the polymer chains expand into the solvent.

Because you might not know R_g, an estimator is provided using input values you either know or can estimate easily.

Shear dependence

Obviously, As discussed in Flow_Shear dependent, the viscosity will tend to decrease with shear rate:

η_0 Pa.s 10

η_{inf} Pa.s 0.1

α 0.052

n 1.3

n (Power Law) 0.5

γ /s 100

η Pa.s 1.1

τ Pa 113.9

Power Law

☐ Show Power Law

● Log(γ) ○ γ ○ Log(σ) ● Viscosity η ○ Stress σ

https://www.stevenabbott.co.uk/practical-rheology/Shear-Viscosity.php

It would be good if we could link the science of low-shear viscosity to the shear-dependent behaviour, but given that no good theories exist for the low-shear viscosity of polymers, we have to make guesstimates. The more tangles, the greater the fall in viscosity when they untangle, and the higher the shear rate at which that will happen. That’s little help, but there seems to be nothing better available.

Dissolution Flory-Huggins

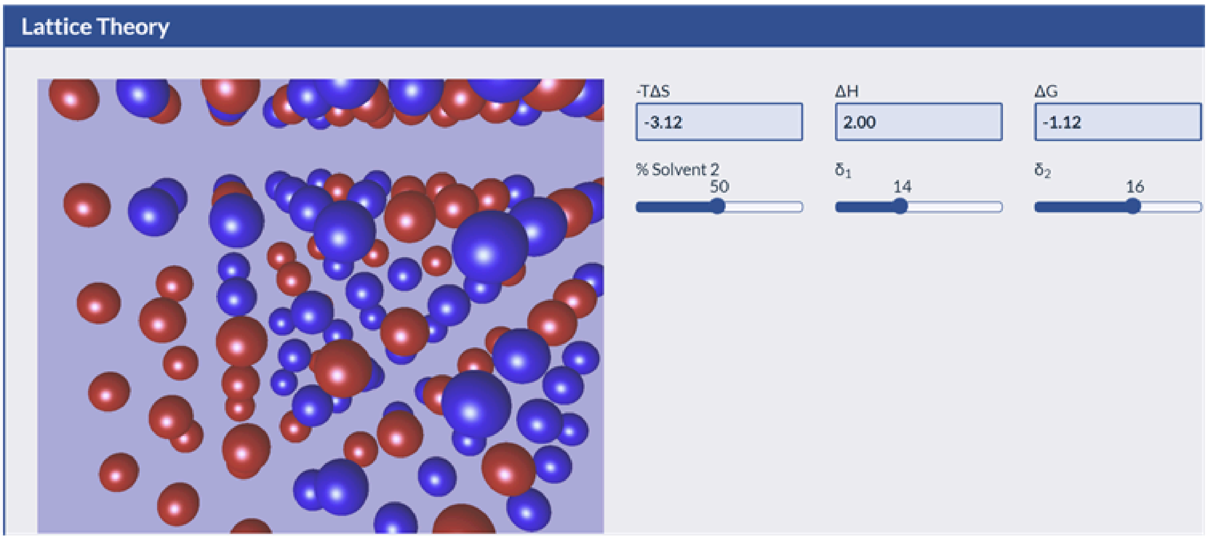
Links

SkinCare, Adhesive, Pharma Formulation, Microencapsulation

The solubility of polymers is usually described via the surprisingly simple assumptions behind Flory-Huggins theory. The χ parameter at the heart of the theory is then used across a wide range of formulation issues. Why do we use such a simple theory? Because it works well-enough, and we don't have any practical alternative.

Lattice theory

If we start with the absurd idea that molecules in solution can be described as being points on a lattice, we can use rather simple ideas to work out the entropy, ΔS, and enthalpy, ΔH, of the system from which we get ΔG, the free energy. A toy system shows us how:


Lattice Theory

-TΔS: -3.12

ΔH: 2.00

ΔG: -1.12

% Solvent 2: 50

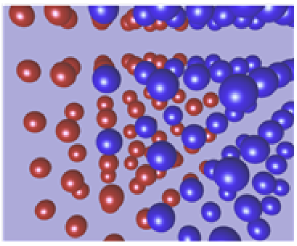
δ_1: 14

δ_2: 16

https://www.stevenabbott.co.uk/practical-solubility/lattice.php

In the screenshot we have a 50:50 mix of red and blue spheres. The energy of interaction of 1-1 is $\delta_1 = 14$ and for 2-2 it's $\delta_2 = 16$. We aren't worrying about units but these might be kJ/mole. Because 1 prefers 1 and 2 prefers 2, the interaction 1-2 is $\sqrt{(\delta_1 \delta_2)} \sim 15$ is de-stabilizing.

Because the spheres are randomly mixed there is a significant entropy gain of -3.12. The enthalpic balance of likes and dislikes is +2, so ΔG =TΔS + ΔH is negative, so the system is happy.



Now make the spheres more unlike each other. Reduce δ_1 to 13 and increase δ_2 to 17.1. The entropic gain does not change but the enthalpic loss is larger, so ΔG becomes sufficiently positive (this is inexact, as mentioned below what matters is the 2nd derivative…) that the system phase separates.

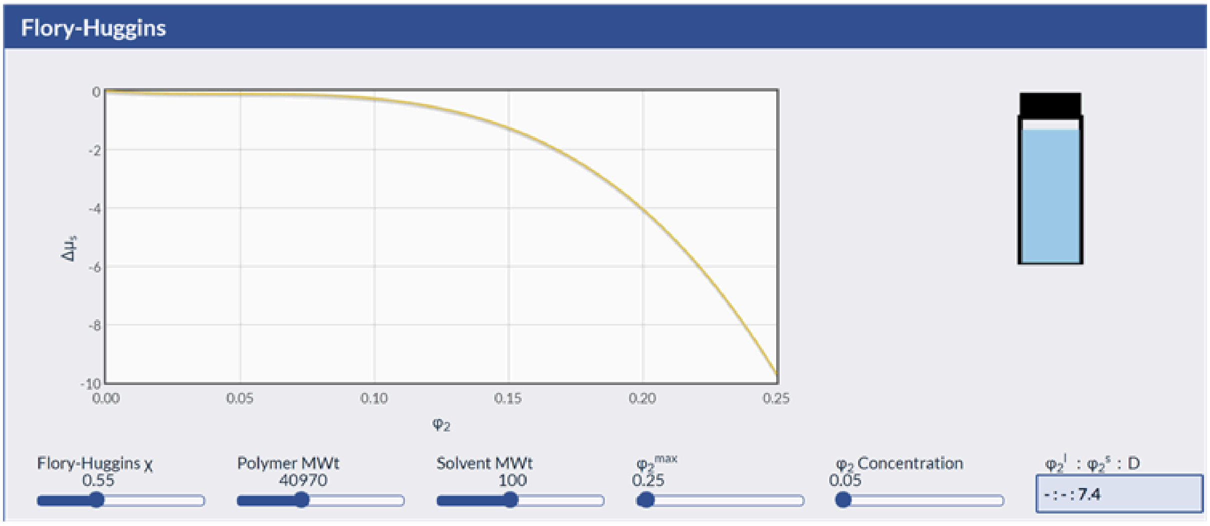
Moving away from this toy system to a polymer where the spheres on the lattice are joined together, and where we use a value χ which goes from 0 (polymer and solvent like each other as much as they like themselves) to 0.5 (bordering on instability) to 1 (they very much dislike each other) we can calculate the entropic and enthalpic terms via the Flory-Huggins equation where Δμ is the chemical potential φ_1 is the volume fraction of solvent, φ_2 is the volume fraction of polymer and x is $MW_{Polymer}/MW_{Solvent}$ which is a large number:

$$\Delta\mu = RT\left[ln[\varphi_1] + \varphi_2\left[1 - \frac{1}{x}\right] + \chi\varphi_2^{\,2}\right]$$

The first two terms are the entropy ones, and the last term, with the χ is the enthalpic one. It's not at all easy to see how it plays out. There are two reasons for this:

1. There is only a small range of conditions where the outcomes hangs in the balance.
2. It's not the sign of Δμ that's important but, rather, its 2nd derivative.

That's why we need an app:



Flory-Huggins χ 0.55

Polymer MWt 40970

Solvent MWt 100

φ_2^{max} 0.25

φ_2 Concentration 0.05

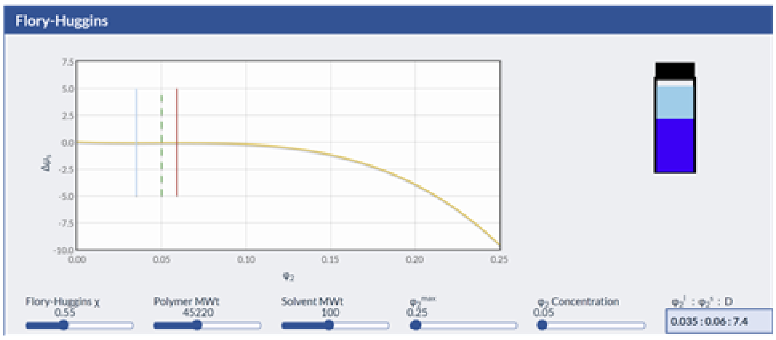
$\varphi_2^l : \varphi_2^s$: D

-:-: 7.4

https://www.stevenabbott.co.uk/practical-solubility/polymer-solubility.php

We have a polymer ~40K MW and a solvent of 100. The Flory-Huggins χ parameter is greater than 0.5 yet the test tube shows a single phase. The chemical potential is flat and just barely negative before, driven by entropy, it becomes strongly negative.

Now increase the polymer MW to 45K:



Flory-Huggins χ 0.55

Polymer MWt 45220

Solvent MWt 100

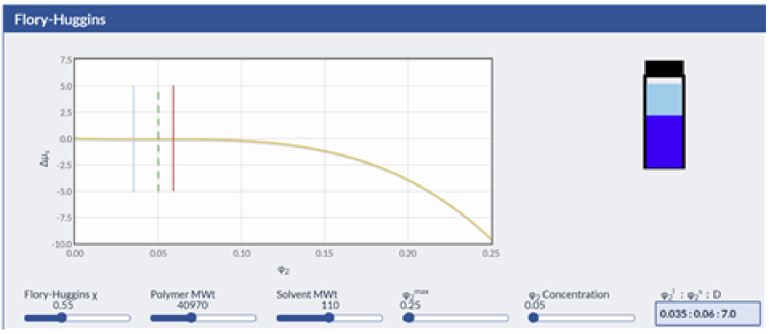
φ_2^{max} 0.25

φ_2 Concentration 0.05

$\varphi_2^l : \varphi_2^s$: D

0.035 : 0.06 : 7.4

It's not clear what's happening to the chemical potential, but its 2nd derivative flips and we get a spinodal decomposition. This means that there is a sudden phase separation and you get two phases. One is a dilute (0.035) solution of polymer in solvent, the other is a phase of polymer swollen with solvent. So although we say that the polymer is insoluble in the solvent, the solvent is happily soluble in the polymer.



Flory-Huggins χ 0.55

Polymer MWt 40970

Solvent MWt 110

φ_2^{max} 0.25

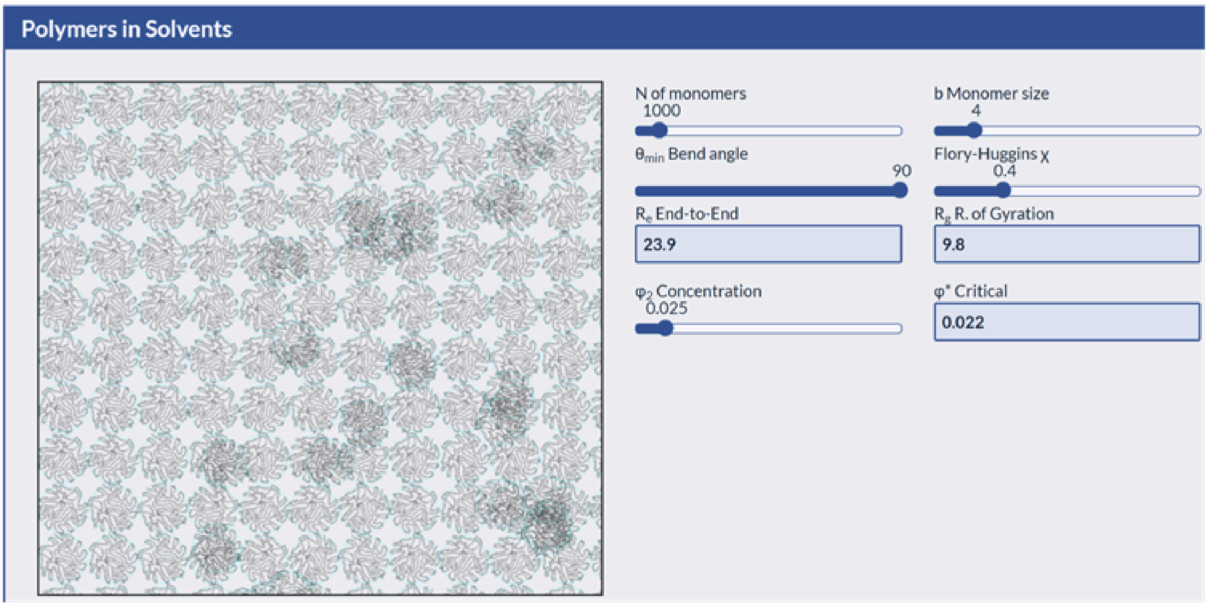
φ_2 Concentration 0.05

$\varphi_2^{\,l} : \varphi_2^{\,s} : D$ 0.035 : 0.06 : 7.0

We can revert to our original 40K polymer and get a spinodal decomposition in another way. This time we've changed the MW of the solvent. It's a universal that (on simple entropic grounds) larger solvents are less good than smaller equivalents with the same (enthalpic) χ value.

It's a common observation that many green solvents aren't as effective as we might like. This is often because they are larger molecules than the ones they are intended to replace.

This hovering state around χ = 0.5 is the "theta solvent" state, where polymers are in a neutral Gaussian coil rather than expanded in a good solvent or tightly coiled as they phase separate.



N of monomers 1000

b Monomer size 4

θ_{min} Bend angle 90

Flory-Huggins χ 0.4

R_e End-to-End 23.9

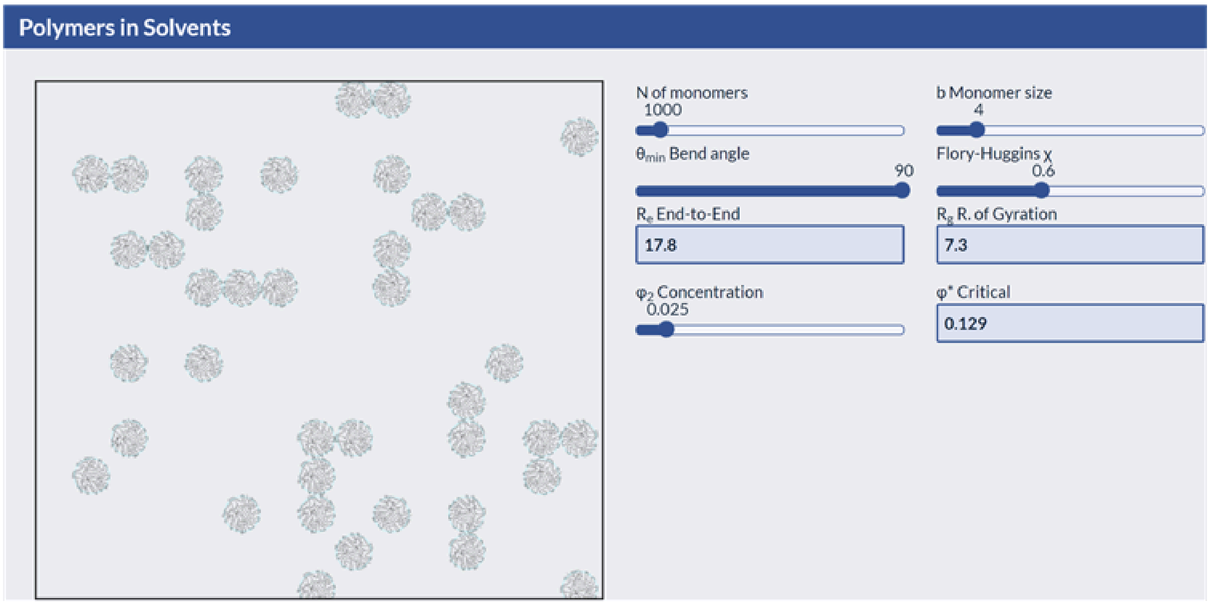
R_g R. of Gyration 9.8

φ_2 Concentration 0.025

φ* Critical 0.022

https://www.stevenabbott.co.uk/practical-solubility/polymers-in-solvents.php

At a relatively small χ, the polymers are expanded and even overlapping. At a relatively high χ their solubility is low and they are much smaller:



N of monomers: 1000

b Monomer size: 4

θ_{min} Bend angle: 90

Flory-Huggins χ: 0.6

R_e End-to-End: 17.8

R_g R. of Gyration: 7.3

φ_2 Concentration: 0.025

φ^* Critical: 0.129

Explanations for the End-to-End radius and Radius of Gyration are provided in the app.

Getting your χ values

A full academic group with access to lots of techniques and time can get a few χ values for a few polymers with respect to a few solvents. The rest of us have no real choice other than using Dissolution_Hansen Solubility Parameters. By knowing the HSP of the polymer (measuring it if necessary by testing yes/no solubility in a range of solvents) and the HSP of any given solvent, you can calculate an HSP Distance, D. The χ parameter for that value of D and a solvent of molar volume MVol is given by:

$$\chi = \frac{MVol.D^2}{4RT}$$

This is fine for the intended theoretical situation – an amorphous polymer of high MW. But we often need a χ value for polymers below their Tg, micro-crystalline (or highly-oriented) polymers, cross-linked (swellable) polymers such as rubbers or, going well outside this regime, pigments and nanoparticles. Then we use the radius of the sphere that encompasses the polymer in the practical solubility measurement.

For example, normal racemic polylactic acid, PLA, has a radius of 8 which relates to a normal χ value ~0.5. A high MW version of l-PLA that has been strongly extrusion oriented might have a radius of 4 because more distant solvents can't touch it. So now we define χ ~0.5 at a radius of 4.

I learned this the hard way. Dr Hansen and I were asked to design a green solvent formulation that was just inside the radius of PLA. We found a great solvent blend which was quickly tested … and shown to be useless. Only then did we learn that the application used highly-oriented PLA. When we got it measured, the radius was half the standard value. A reformulation to this value then worked as required.

Such an approach does not get the approval of purists. But the purists haven't given us anything we can use, so as practical formulators we use the tools that are proven to do an adequate job, and use them with due awareness that they are not as precise as we would like.

Dissolution Gases

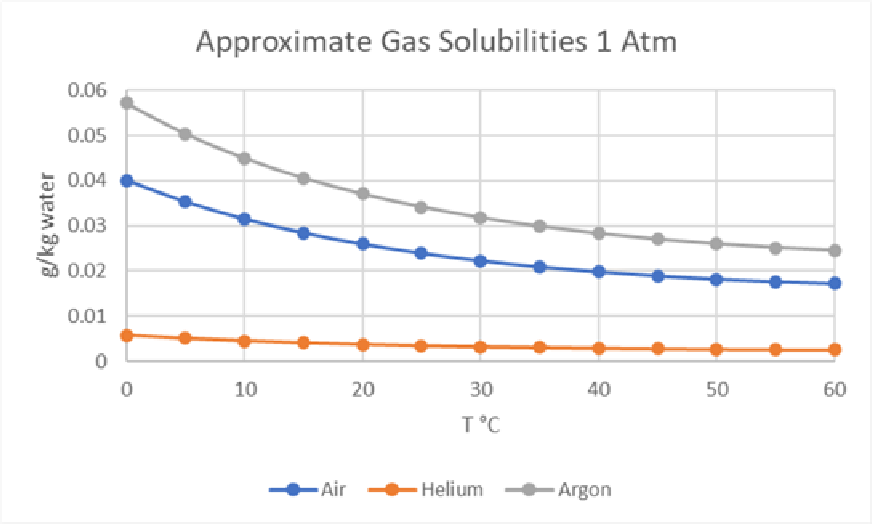
Links

The solubility of gases is something most of us can ignore most of the time – then you encounter an unexpected problem caused by an issue you didn't suspect would be important.

Little to know

The typical solubility, S, of a gas in a solvent in g/kg depends on pressure, P (bar), temperature T (°C) and MW via a simple formula valid (within a factor of 2) from 0°C to 80°C:

$$S = P\left(0.04 - 0.0008T + 0.000007T^2\right)\frac{MW}{28}$$



There are obvious exceptions when you have "reactive" gases. For CO2 the formation of H_2CO_3 gives a solubility 50x higher than expected. Ammonia forms ammonium hydroxide, hydrogen chloride splits into the acid form etc.

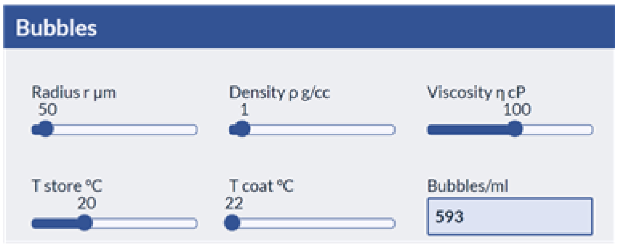
Pressure effects for normal gases (solubility is proportional to pressure) look equally dull until you go scuba diving, where the solubility contrast between high and low pressures can be life threatening.

The reason for this chapter is that there is a real issue with the modest temperature dependence, one that is little known and yet has caused problems for many. And the chapter gives us an opportunity to talk about champagne…

Coating defects

When people see a pinhole defect (Coating_Pinholing Theory) in their coating they are tempted to call it a "repellency". And because names imply causes, they look around for the oil contamination that has caused the repellency. They generally can't find any such contamination.

What else could cause it? The answer is a gas bubble. Of course people are aware that mixing and some aspects of a coating head can create bubbles in a formulation and work hard to eliminate them. But still the pinholes are there. The app at https://www.stevenabbott.co.uk/practical-coatings/bubbles.php explains why:



Radius r µm 50

Density ρ g/cc 1

Viscosity η cP 100

T store °C 20

T coat °C 22

Bubbles/ml 593

There are other inputs and outputs in the app (it also calculates time for bubbles to rise) but here we focus on a cause of coating defects. In this example, the aqueous coating was stored at 20°C where it had time to equilibrate with air so is a saturated solution. On moving to the coating machine, its temperature rises to 22°C. Amazingly, the reduced solubility of air, just in that 2° difference, is enough (if it all came out and all produced the same-sized bubbles) to create ~600 bubbles of 100µm diameter in each ml of solution.

Of course in real life we don't get all the dissolved air coming out. But the point is that even 1 bubble can ruin your day and the super-saturated solvent can readily produce that 1 bubble. If, however, the solution had been stored at 22°C and coated at 20°C, not only would there be no extra bubbles, there's even a chance that a rogue bubble from, say, mixing, might dissolve back into the water.

Champagne

A bottle of champagne might contain CO2 at 5 bar, which translates to 11g/kg at 10°C, so ~ 8g/bottle. When we open it, if 4/5 of it (6g) came out at once, that would be ~3l of CO2, something you wouldn't want to happen if you preferred drinking the champagne.

A simple formula shows us why it can't all come out easily. Suppose the CO2 coming out of the champagne produces a bubble of radius r and that the surface tension, γ, of champagne is ~ 50mN/m (the ingredients act as poor surfactants). The pressure, P, inside the bubble is given by:

$$P = \frac{2\gamma}{r}$$

If r = 1µm then P=0.1/10^{-6} = 100 kPa ~ 1 bar. So a bubble of r~0.2µm has an internal pressure of ~ 5 bar, and we know the CO2 is dissolved at that pressure.

So we have the problem that as a small amount of CO2 tries to come out of solution, producing a bubble < 0.2 µm, it can't. And because no small bubble can form, no bigger bubble can occur via diffusion of the gas into the bubble. So fizzy champagne is impossible because nucleation, formation of the proto-bubble, is impossible.

We know that the puzzle is solved via seeds, e.g. slight defects in the champagne glass.

Although, for fun, we've discussed champagne, this nucleation problem is universal. Many manufacturers of polyurethane foams, knowing that smaller foam cells are better insulating, have tried to seed their formulations in order to create many smaller bubbles rather than fewer larger bubbles. Although it can work in the lab, it is hard to scale up. Closer to champagne, some foams are made by dissolving scCO2 into a polymer and then releasing the pressure suddenly. Because the pressure difference is so large, even super-small bubbles are able to form so the seeding problem goes away.

Dissolution Hansen Solubility Parameters

Links

Deodorant Sticks, Sun Screens, SkinCare, Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint, Pharma Formulation, Microencapsulation

Many solubility issues, e.g. replacing one solvent blend with another, can be done efficiently using HSP – Hansen Solubility Parameters. These are 3 parameters for any solvent, polymer or pigment/particle that capture the essentials of their interactions:

1. δD is the Dispersion component – related to RI, polarizability, van der Waals forces
2. δP is the Polar component – related to dipole moment
3. δH is the Hydrogen-bonding component – usually obvious from the molecular structure.

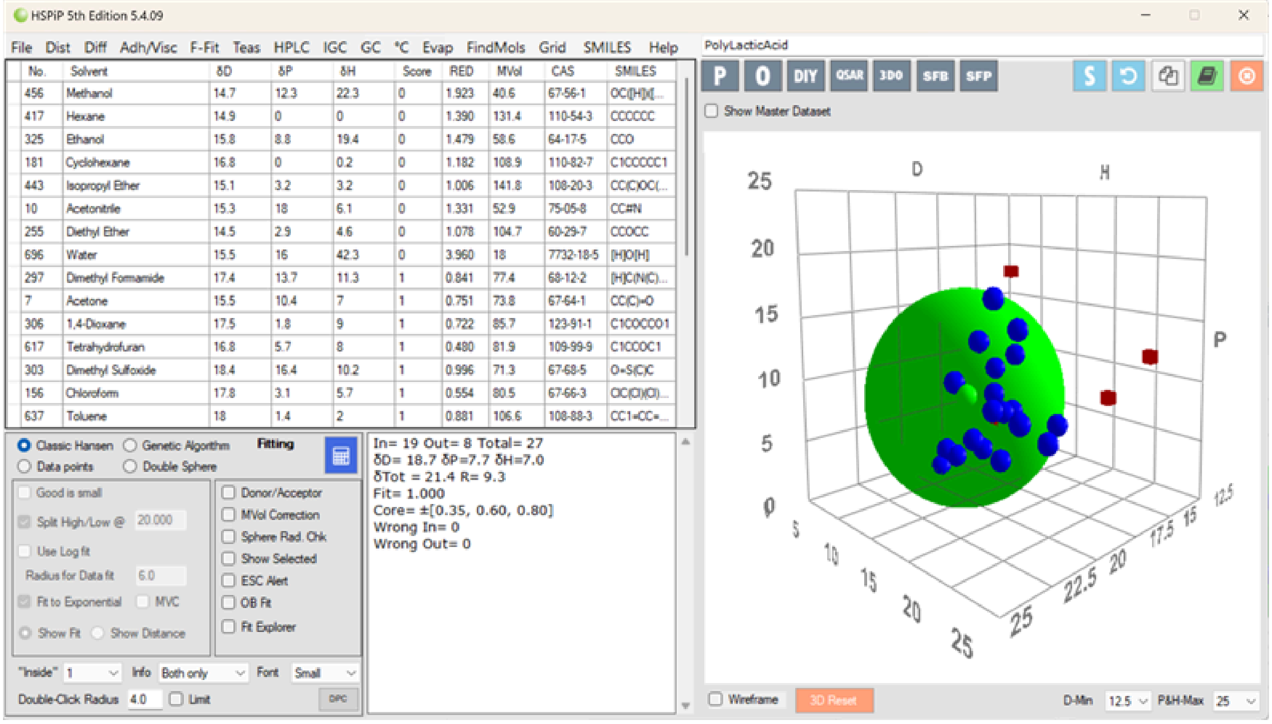
The HSP of a solvent blend is the volume-weighted average of the individual components

The compatibility of, say, a solvent (blend) with a polymer depends on the Distance, D, between the HSP values:

$$D^2 = 4(\delta D_1 - \delta D_2)^2 + (\delta P_1 - \delta P_2)^2 + (\delta H_1 - \delta H_2)^2$$

Because we know the HSP of most solvents and because we can know or measure the HSP of our other components, we can easily formulate by, for example, ensuring the smallest Distance between components we want to remain compatible.

Although there are plenty of HSP-related apps at https://www.stevenabbott.co.uk/practical-solubility/hspintro.php, the science is best explored using the HSPiP (Hansen Solubility Parameters in Practice) package of software, datasets and eBook created by Prof Steven Abbott, Dr Charles Hansen and Dr YAMAMOTO Hiroshi.



No.	Solvent	δD	δP	δH	Score	RED	MVol	CAS	SMILES
456	Methanol	14.7	12.3	22.3	0	1.923	40.6	67-56-1	OC([H])([...
417	Hexane	14.9	0	0	0	1.390	131.4	110-54-3	CCCCCC
325	Ethanol	15.8	8.8	19.4	0	1.479	58.6	64-17-5	CCO
181	Cyclohexane	16.8	0	0.2	0	1.182	108.9	110-82-7	C1CCCCC1
443	Isopropyl Ether	15.1	3.2	3.2	0	1.006	141.8	108-20-3	CC(C)OC(...
10	Acetonitrile	15.3	18	6.1	0	1.331	52.9	75-05-8	CC#N
255	Diethyl Ether	14.5	2.9	4.6	0	1.078	104.7	60-29-7	CCOCC
696	Water	15.5	16	42.3	0	3.960	18	7732-18-5	[H]O[H]
297	Dimethyl Formamide	17.4	13.7	11.3	1	0.841	77.4	68-12-2	[H]C(N(C)...
7	Acetone	15.5	10.4	7	1	0.751	73.8	67-64-1	CC(C)=O
306	1,4-Dioxane	17.5	1.8	9	1	0.722	85.7	123-91-1	C1COCCO1
617	Tetrahydrofuran	16.8	5.7	8	1	0.480	81.9	109-99-9	C1CCOC1
303	Dimethyl Sulfoxide	18.4	16.4	10.2	1	0.996	71.3	67-68-5	O=S(C)C
156	Chloroform	17.8	3.1	5.7	1	0.554	80.5	67-66-3	ClC(Cl)(Cl)...
637	Toluene	18	1.4	2	1	0.881	106.6	108-88-3	CC1=CC=...

● Classic Hansen ○ Genetic Algorithm
○ Data points ○ Double Sphere
☐ Good is small
☑ Split High/Low @ 20.000
☐ Use Log fit
Radius for Data fit 6.0
☑ Fit to Exponential ☐ MVC
○ Show Fit ○ Show Distance
☐ Donor/Acceptor
☐ MVol Correction
☐ Sphere Rad. Chk
☐ Show Selected
☐ ESC Alert
☐ OB Fit
☐ Fit Explorer
"Inside" 1 Info Both only Font Small
Double-Click Radius 4.0 ☐ Limit
D-Min 12.5 P&H-Max 25

```
In= 19 Out= 8 Total= 27
δD= 18.7 δP=7.7 δH=7.0
δTot = 21.4 R= 9.3
Fit= 1.000
Core= ±[0.35, 0.60, 0.80]
Wrong In= 0
Wrong Out= 0
```

The image is included because it captures so much useful information which we can explore.

The Table

With the table at the top left you get an idea of the δD, δP and δH parameters for some familiar solvents. Hexane has (of course) no polar or H-bonding and is held together as a liquid via dispersion forces of 14.9 (the units are the obscure MPa$^{1/2}$). We often neglect these because they seem dull, but they are what holds most things together, and they should never be underestimated. Cyclohexane has the same 6 carbons, but its δD is 16.8. We can't see benzene in the table but it is 18.4. As we go to increasingly "strong" solvents, δD increases. These parameters are "cohesive energy density" so the more compact the molecule (benzene's density is 33% higher than hexane's) the higher δD. Another way to think about it is in terms of polarizability (that's where van der Waals forces, or, specifically, the London force components, come from) which is related to refractive index – $n_{benzene} = 1.5$, $n_{hexane} = 1.37$.

We can then look at δP. It seems reasonable that we go from chloroform's 3.1 through acetone's 10.4 through DMF's 13.7 up to acetonitrile's 18. You would not be able to guess the correct numbers, but the trend is reasonable, and strongly linked to dipole moment.

Similarly, δH is reasonable. Toluene's value of 2 is at first surprising, but aromatics are significant H-bond acceptors. Acetone's 7 is reasonable for the H-bonding C=O and we go up to ethanol's 19.4 which is less than methanol's 22.3 which is much less than water's 42.3.

Given that H-bond is a "polarity-related" effect it is sometimes asked if we could just use δD and a super-δP. But look at acetonitrile. It is highly polar but a weak H-bond acceptor, while ethanol is not especially polar but is a strong H-bonder. As solvents they behave very differently, and 50+ years of experience have shown that we really do need all three parameters. Why not four parameters, splitting H-bonding into Donor and Acceptor? Because no one has ever created a self-consistent and useful scheme using them. HSPiP includes the option to use a split δH but it is little used.

For those who worry that these numbers are just some arbitrary fitting values obtained by Hansen in the late 60's, for solvents the square root of the sum of the squares of the three values must equal the enthalpy of vaporization, the δD values must relate to refractive index and δP to dipole moment. HSP are extensions of Hildebrand theory which is related to Flory-Huggins and the χ parameter (Dissolution_Flory-Huggins) which in turn are related to lattice theory. Yes, HSP are not precise, yes, you have to use your formulation knowledge to get the best out of them (a feature, not a bug) but, no, they aren't arbitrary.

The Sphere

The "Score" column in the table contains some 0 ("bad" solvents) and 1 ("good" solvents) values. The scores were found by putting a small amount of the substance of interest (in this case PLA, poly lactic acid) into each of 27 tubes and adding the same volume of 27 different solvents to each tube, shaking and leaving for some hours. The scores in this case were found using the human eye. Other ways to distinguish between good/bad (including numeric measures such as measured solubility or relative sedimentation time in a centrifuge) achieve the same thing. Using a fitting function, the green sphere has been created with all the good solvents (in blue) inside and all the bad ones (in red) outside. The small green dot in the middle represents the HSP of the PLA.

What can we do with such a value? This is where the HSP Distance comes into play. Here's the distance formula once again:

$$D^2 = 4\left(\delta D_1 - \delta D_2\right)^2 + \left(\delta P_1 - \delta P_2\right)^2 + \left(\delta H_1 - \delta H_2\right)^2$$

If we ignore the factor of 4 (there for technical reasons), D is simply the Pythagorean distance in 3D space. It is rather obvious that the bad solvents are far away and that some good solvents are closer to the centre than others. And, yes, those closer to the centre are better solvents. Those at the edge really are borderline. An increase in MW or crystallinity might make the sphere smaller, so a previously good solvent becomes bad. This uncertainty at the edge is a feature, not a bug.

The RED column (Relative Energy Difference) shows D/Ra where Ra is the radius of the sphere. By definition, a RED > 1 means insolubility *for that specific substance when measured to a specific criterion for good solubility*. If your criterion for solubility is 0.01% of a 5000 MW polymer, the radius will be much larger than if your criterion is 10% solubility for the same polymer with a 500,000 MW. HSP are a pragmatic tool to be applied to your specific problem. The core HSP value does *not* depend on your definition. The RED is your way of deciding how close you need to be to that core value in order to obtain the solubility you desire.

Other options

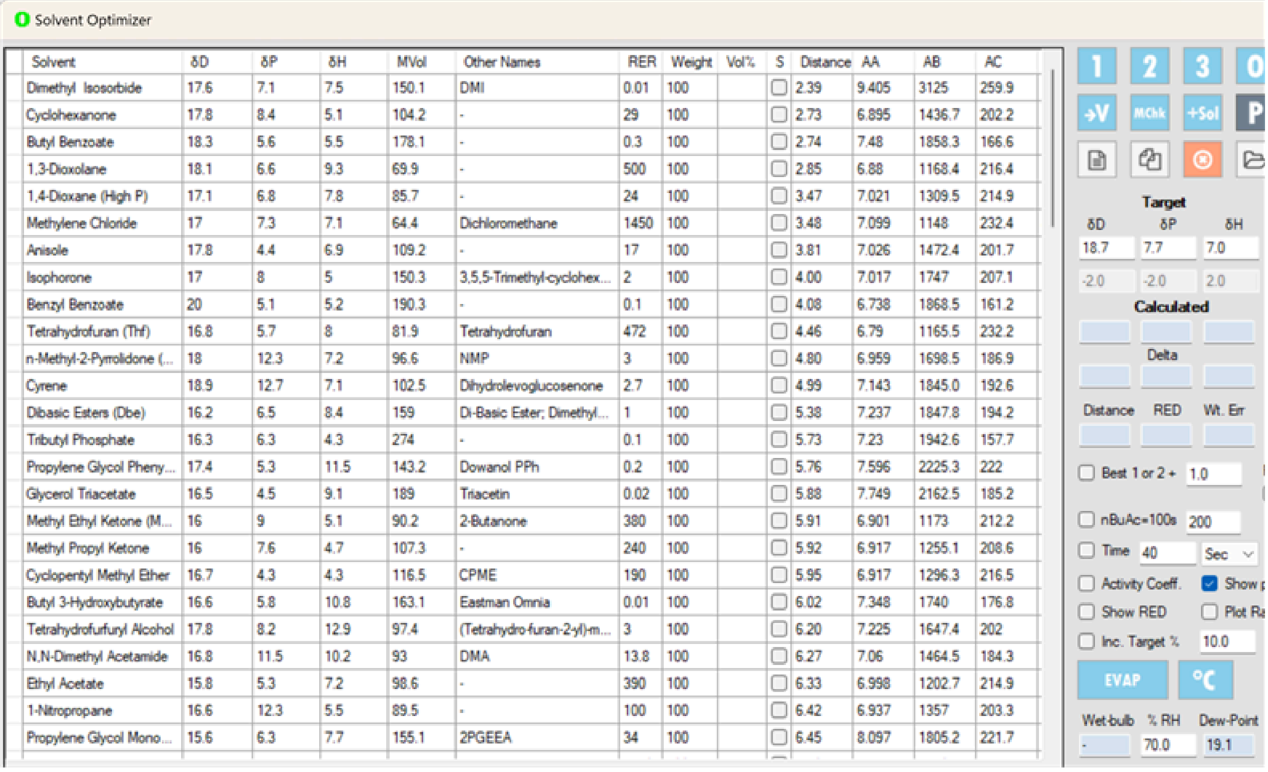
Because HSP can be used for soluble polymers, swellable (crosslinked) polymers, APIs, nanoparticles (see Dispersions_Solubility if you disagree) and more, it is unlikely that one fitting procedure (against a specific objective function) works well for everything, so there are choices of functions, each with their trade-offs. You can use scores of 1 (best) to 6 (worst) instead of 0/1, you can use numerical data rather than scores. And you can correct for temperature (HSP values decrease with increasing temperature).

The point of these options is to allow the pragmatic formulator to work in a way that makes sense to them.

Solvent blends

The solvents used in the measurement of HSP values are not chosen because they are nice or green. They are chosen because they span HSP space. There is no point in testing hexane, heptane and octane, or ethanol, propanol and butanol. You choose a representative solvent from each class then, in addition to typical ketones and esters use solvents such as NMP, chloroform, DMF, DMSO that are in unique parts of HSP space. You (or your robot) are using only a few ml of each solvent in capped tubes, so the safety risks are manageable. Thanks to these solvents you get a good HSP value. But the fact that pyridine happens to be a close HSP match to PLA is of no real-world use to you – you aren't going to use it for a formulation.

So you take your target HSP value (e.g. the one for PLA), assemble a convenient list of common solvents, then sort them by Distance to find, hopefully, one solvent that meets your requirements for cost, safety, volatility, odour … Here is such a list using HSPiP's Solvent Optimizer:



Solvent Optimizer

Solvent	δD	δP	δH	MVol	Other Names	RER	Weight	Vol%	S	Distance	AA	AB	AC
Dimethyl Isosorbide	17.6	7.1	7.5	150.1	DMI	0.01	100		☐	2.39	9.405	3125	259.9
Cyclohexanone	17.8	8.4	5.1	104.2	-	29	100		☐	2.73	6.895	1436.7	202.2
Butyl Benzoate	18.3	5.6	5.5	178.1	-	0.3	100		☐	2.74	7.48	1858.3	166.6
1,3-Dioxolane	18.1	6.6	9.3	69.9	-	500	100		☐	2.85	6.88	1168.4	216.4
1,4-Dioxane (High P)	17.1	6.8	7.8	85.7	-	24	100		☐	3.47	7.021	1309.5	214.9
Methylene Chloride	17	7.3	7.1	64.4	Dichloromethane	1450	100		☐	3.48	7.099	1148	232.4
Anisole	17.8	4.4	6.9	109.2	-	17	100		☐	3.81	7.026	1472.4	201.7
Isophorone	17	8	5	150.3	3,5,5-Trimethyl-cyclohex...	2	100		☐	4.00	7.017	1747	207.1
Benzyl Benzoate	20	5.1	5.2	190.3	-	0.1	100		☐	4.08	6.738	1868.5	161.2
Tetrahydrofuran (Thf)	16.8	5.7	8	81.9	Tetrahydrofuran	472	100		☐	4.46	6.79	1165.5	232.2
n-Methyl-2-Pyrrolidone (...	18	12.3	7.2	96.6	NMP	3	100		☐	4.80	6.959	1698.5	186.9
Cyrene	18.9	12.7	7.1	102.5	Dihydrolevoglucosenone	2.7	100		☐	4.99	7.143	1845.0	192.6
Dibasic Esters (Dbe)	16.2	6.5	8.4	159	Di-Basic Ester; Dimethyl...	1	100		☐	5.38	7.237	1847.8	194.2
Tributyl Phosphate	16.3	6.3	4.3	274	-	0.1	100		☐	5.73	7.23	1942.6	157.7
Propylene Glycol Pheny...	17.4	5.3	11.5	143.2	Dowanol PPh	0.2	100		☐	5.76	7.596	2225.3	222
Glycerol Triacetate	16.5	4.5	9.1	189	Triacetin	0.02	100		☐	5.88	7.749	2162.5	185.2
Methyl Ethyl Ketone (M...	16	9	5.1	90.2	2-Butanone	380	100		☐	5.91	6.901	1173	212.2
Methyl Propyl Ketone	16	7.6	4.7	107.3	-	240	100		☐	5.92	6.917	1255.1	208.6
Cyclopentyl Methyl Ether	16.7	4.3	4.3	116.5	CPME	190	100		☐	5.95	6.917	1296.3	216.5
Butyl 3-Hydroxybutyrate	16.6	5.8	10.8	163.1	Eastman Omnia	0.01	100		☐	6.02	7.348	1740	176.8
Tetrahydrofurfuryl Alcohol	17.8	8.2	12.9	97.4	(Tetrahydro-furan-2-yl)-m...	3	100		☐	6.20	7.225	1647.4	202
N,N-Dimethyl Acetamide	16.8	11.5	10.2	93	DMA	13.8	100		☐	6.27	7.06	1464.5	184.3
Ethyl Acetate	15.8	5.3	7.2	98.6	-	390	100		☐	6.33	6.998	1202.7	214.9
1-Nitropropane	16.6	12.3	5.5	89.5	-	100	100		☐	6.42	6.937	1357	203.3
Propylene Glycol Mono...	15.6	6.3	7.7	155.1	2PGEEA	34	100		☐	6.45	8.097	1805.2	221.7

Target

δD	δP	δH
18.7	7.7	7.0
-2.0	-2.0	2.0

Calculated

Delta

Distance RED Wt. Err

☐ Best 1 or 2 + 1.0

☐ nBuAc=100s 200

☐ Time 40 Sec

☐ Activity Coeff. ☑ Show

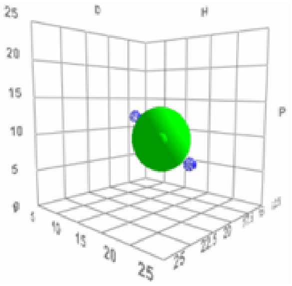
☐ Show RED ☐ Plot R

☐ Inc. Target % 10.0

Wet-bulb	% RH	Dew-Point
-	70.0	19.1

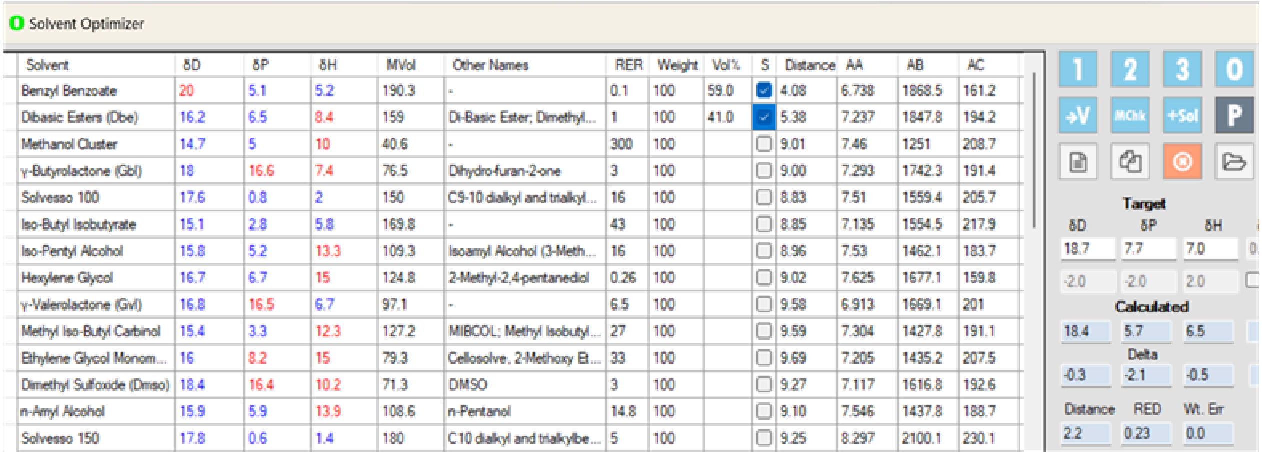
We find that DMI, cyclohexanone, butyl benzoate and 1,3-dioxolane are all reasonable matches to the "target" HSP of [18.7, 7.7, 7.0] which was calculated for PLA. The problem is that DMI is expensive and of low volatility (the RER column is Relative Evaporation Rate, with nBuAc = 100), some people don't like the odour of cyclohexanone, butyl benzoate is also involatile while dioxolane is too volatile, with a flash point (you need to scroll the columns to find it) of -6°C which makes it difficult to use safely.

If your solvent list was of a smaller set of green solvents, it is unlikely that you could find a single solvent that was a good-enough match.



That's where solvent blends become important. The HSP of an x:y blend of two solvents is the x:y ratio of their δD, δP and δH values. To make an extreme point, lets look at the situation where we have two bad solvents, outside, but on opposite sides of, the sphere. A 50:50 blend would give an HSP right in the centre, implying that two bad solvents could become a good solvent blend. Although the idea is anti-intuitive, it turns out to be correct. And if you can make a good solvent from two bad solvents, even more can you make a blend from two OK-but-not-good-enough solvents, providing their respective deficiencies (e.g. one's δH is too low, the other's δH is too high) are complementary, allowing an optimum intermediate blend.

When Charles Hansen developed HSP he realised that they predicted that two bad solvents could make a good one. This made no sense, so he did some experiments with bad solvent pairs to disprove his theory so he could then find a better one. To his surprise, the blends worked as the theory predicted. People remain surprised! In tutorial classes, people can find the HSP of nail polish. Butyl benzoate and benzyl alcohol are two of the test solvents and they are both found to be bad. From the Sphere fit they can then calculate that a 50:50 blend of the two is an OK solvent. They are always astonished when they test the blend and find that it works.



Solvent	δD	δP	δH	MVol	Other Names	RER	Weight	Vol%	S	Distance	AA	AB	AC
Benzyl Benzoate	20	5.1	5.2	190.3	-	0.1	100	59.0	☑	4.08	6.738	1868.5	161.2
Dibasic Esters (Dbe)	16.2	6.5	8.4	159	Di-Basic Ester; Dimethyl...	1	100	41.0	☑	5.38	7.237	1847.8	194.2
Methanol Cluster	14.7	5	10	40.6	-	300	100		☐	9.01	7.46	1251	208.7
γ-Butyrolactone (Gbl)	18	16.6	7.4	76.5	Dihydro-furan-2-one	3	100		☐	9.00	7.293	1742.3	191.4
Solvesso 100	17.6	0.8	2	150	C9-10 dialkyl and trialkyl...	16	100		☐	8.83	7.51	1559.4	205.7
Iso-Butyl Isobutyrate	15.1	2.8	5.8	169.8	-	43	100		☐	8.85	7.135	1554.5	217.9
Iso-Pentyl Alcohol	15.8	5.2	13.3	109.3	Isoamyl Alcohol (3-Meth...	16	100		☐	8.96	7.53	1462.1	183.7
Hexylene Glycol	16.7	6.7	15	124.8	2-Methyl-2,4-pentanediol	0.26	100		☐	9.02	7.625	1677.1	159.8
γ-Valerolactone (Gvl)	16.8	16.5	6.7	97.1	-	6.5	100		☐	9.58	6.913	1669.1	201
Methyl Iso-Butyl Carbinol	15.4	3.3	12.3	127.2	MIBCOL; Methyl Isobutyl...	27	100		☐	9.59	7.304	1427.8	191.1
Ethylene Glycol Monom...	16	8.2	15	79.3	Cellosolve, 2-Methoxy Et...	33	100		☐	9.69	7.205	1435.2	207.5
Dimethyl Sulfoxide (Dmso)	18.4	16.4	10.2	71.3	DMSO	3	100		☐	9.27	7.117	1616.8	192.6
n-Amyl Alcohol	15.9	5.9	13.9	108.6	n-Pentanol	14.8	100		☐	9.10	7.546	1437.8	188.7
Solvesso 150	17.8	0.6	1.4	180	C10 dialkyl and trialkylbe...	5	100		☐	9.25	8.297	2100.1	230.1

Target

δD	δP	δH
18.7	7.7	7.0
-2.0	-2.0	2.0

Calculated

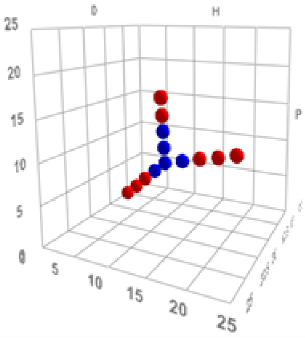
18.4	5.7	6.5

Delta

-0.3	-2.1	-0.5

Distance	RED	Wt. Err
2.2	0.23	0.0

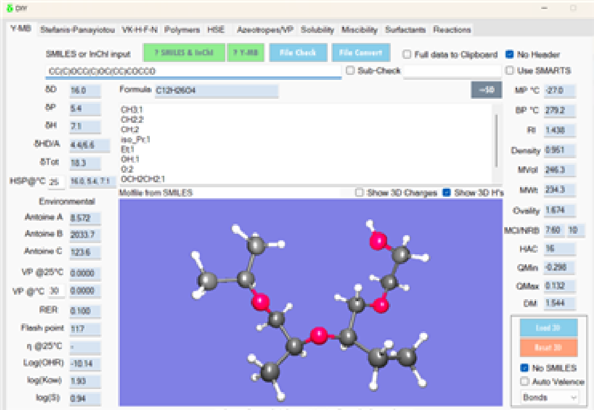
In the PLA example, if you wanted to use DBE as a relatively nice solvent, but which has a Distance of 5.4, you can find (manually or by pushing some buttons) that a 41:59 mix with benzyl benzoate (Distance of 4) gives a blend with a distance of 2.2. The colour coding shows that the δD values are high (red) and low (blue) compared to the target, the other way round for δH, with nothing useful in terms of δP.,



Solvent blends provide another capability. It can be very inconvenient to assemble, say, 15 different solvents. An alternative approach is to carefully choose 4 solvents and make systematic blends of them. If the first solvent is known to be good, and the other three are known to be bad, then you might find good (blue) and bad (red) solubilities if you try the 13 blends, easily created by some automated system. The resulting sphere won't be perfect (the blends cover only a limited range of HSP space) but (via a built-in tool, or using common sense) you can identify a few key test solvents in the missing parts of HSP space to refine the fit. This "grid" method is very popular and especially well suited for robotic systems.

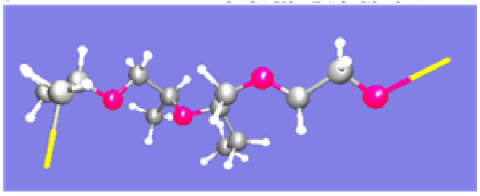
Getting HSP values

We have seen that HSP values for solids (polymers, crystalline solids, nanoparticles) can be measured using the sphere method. Low volatility semi-solids can be measured using Inverse Gas Chromatography (see https://www.stevenabbott.co.uk/practical-chromatography/hsp.php for an explanation, and read the free IGC eBook from the same site). But new solvents are hard to measure and if you need the HSP of lots of molecules, you are unlikely to carry out lots of sphere tests.



SMILES or InChI input CC(C)OCC(C)OC(CC)COCCO
☐ Full data to Clipboard
☑ No Header
☐ Sub-Check
☐ Use SMARTS
δD 16.0
δP 5.4
δH 7.1
δHD/A 4.4/6.6
δTot 18.3
HSP@°C 25 16.0 5.4 7.1
Formula C12H26O4
CH3:1
CH2:2
CH:2
iso_Pr:1
Et:1
OH:1
O:2
OCH2CH2:1
Molfile from SMILES
☐ Show 3D Charges
☑ Show 3D H's
Environmental
Antoine A 8.572
Antoine B 2033.7
Antoine C 123.6
VP @25°C 0.0000
VP @°C 30 0.0000
RER 0.100
Flash point 117
η @25°C -
Log(OHR) -10.14
log(Kow) 1.93
log(S) 0.94
MP °C -27.0
BP °C 279.2
RI 1.438
Density 0.951
MVol 246.3
MWt 234.3
Ovality 1.674
MCI/NRB 7.60 10
HAC 16
QMin -0.298
QMax 0.132
DM 1.544
☑ No SMILES
☐ Auto Valence

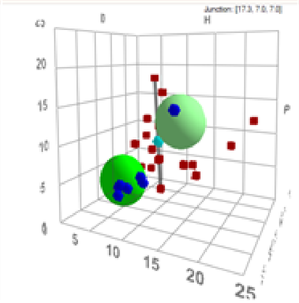
You can then use HSPiP to get a reasonable estimate simply from the SMILES structure. [If you are not familiar with SMILES, there is an excellent guide on Wikipedia.] Suppose you are thinking of some new green ether-alcohol, you can enter its SMILES (in this example it's CC(C)OCC(C)OC(CC)COCCO) and get both its 3D structure and an estimate of its HSP [16, 5.4, 7.1] plus handy estimates of many other properties. If the properties are not to your liking, you can try another variant.



If, instead of it being a small molecule, it was some new green polymer, you can enter the Polymer SMILES: XCC(C)OCC(C)OC(CC)COCCOX. The Xs are used to indicate the polymer chain links and you can see them in yellow in the image. In this case the calculated HSP are [17.4, 2.6, 2.2]. We have gained δD, as is typical with polymers, and lost both δP and δH because the terminal -OH group has become another ether link.

Complex formulations

We often find that we need solvents that must be good for some component of a formulation and bad for another. HSP make it easy to explore the trade-offs.



In this example we have two components that are well-separated in HSP space (for ease of illustration). Each has one or more good solvents (blue) from the list of solvents being used. Most of the chosen solvents are bad for both. The line and blue dot show how you can create solvent blends (there is a slider to allow this to happen) to find a blend that meets your specific requirements. This ability to work in 3D to explore multiple components interacting with multiple potential solvents is a powerful way to create advanced formulations.

Just remember that each copy of HSPiP comes with the advanced neural network of the person using it. The numbers provided by HSP are not the endpoint. They are the beginning of an interaction with your own chemical knowledge.

Not just HSPiP

It has been convenient to use HSPiP to illustrate the core science of HSP. You can find many useful apps and lots more explanations at both https://www.stevenabbott.co.uk/practical-solubility/hspintro.php and https://www.hansen-solubility.com/. On the Hansen-Solubility site you can download some useful Excel sheets that do many of the basic calculations for you and there is a selection of YouTube tutorials to deepen your understanding.

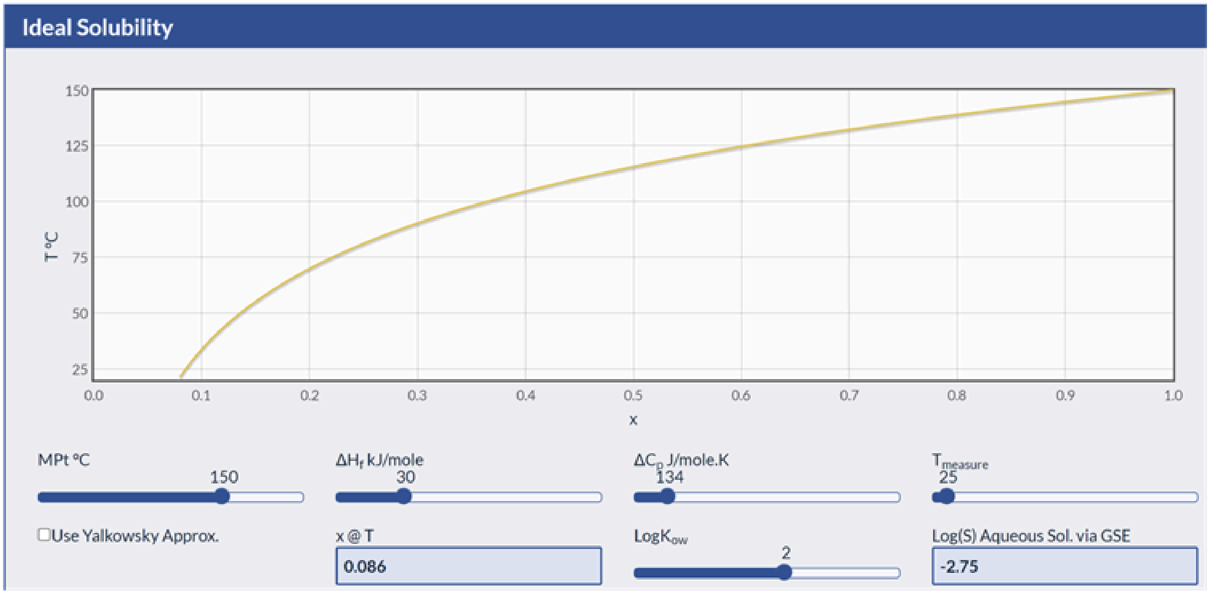
The 3D Optimization discussed in the previous section is available as an app: https://www.stevenabbott.co.uk/practical-solubility/HSP-3DO.php.

Dissolution Ideal solubility

Links

Sun Screens, SkinCare, Pharma Formulation, Microencapsulation

However good your solvent, for a crystalline solute you can't get better than the ideal solubility which is related to the energy required to break down the crystalline structure. The theory used here, although "standard", turns out to be tricky (with unknown parameters) and not wonderfully accurate. But it's better than having no theory at all.



MPt °C 150

ΔH_f kJ/mole 30

ΔC_p J/mole.K 134

$T_{measure}$ 25

☐ Use Yalkowsky Approx.

x @ T 0.086

$LogK_{ow}$ 2

Log(S) Aqueous Sol. via GSE -2.75

https://www.stevenabbott.co.uk/practical-solubility/ideal.php

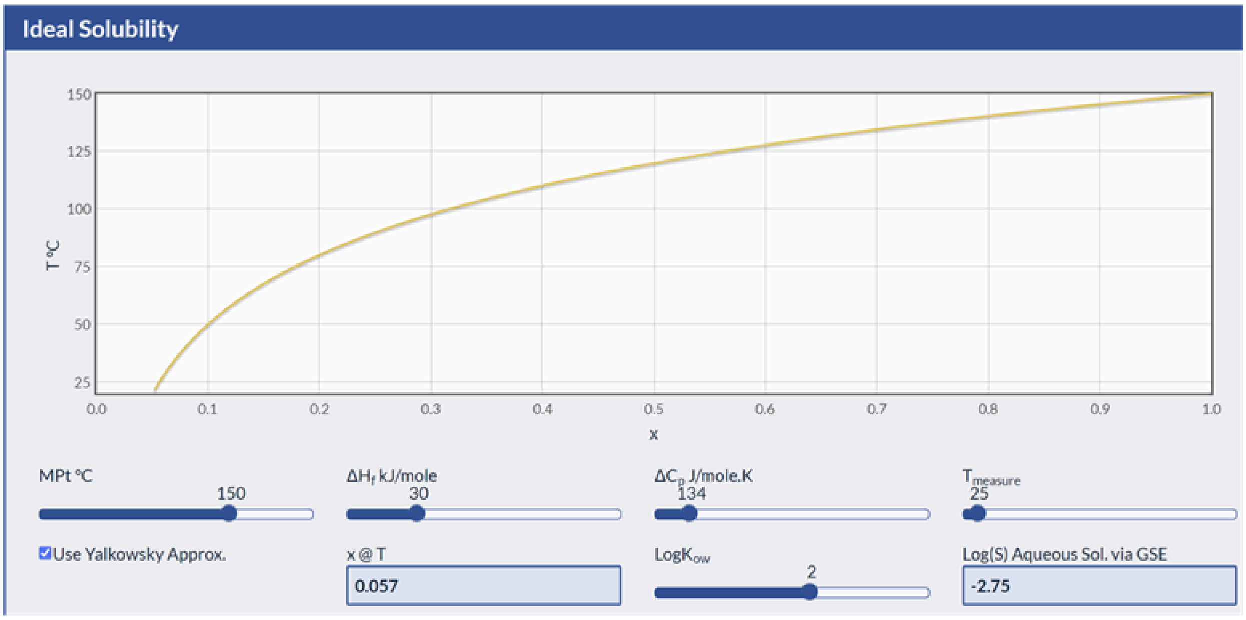
The next app will be more useful, but this version captures the essence of ideal solubility. For a solute with a MPt of 150 and a reasonable ΔH_f of 30 kJ/mole and the hard-to-measure ΔC_p of 134 J/mole/K the solubility at temperatures up to the MPt are shown. At room temperature the solubility is just below 0.1 mole fraction. If you need 0.2 mole fraction there is *nothing* you can do, because this is the solubility with an ideal solvent – one where the activity coefficient = 1. Real solvents won't be ideal, so solubility will be less than this.

Of course you could change the definition of "ideal" and if the solute is a base, add an acidic solvent. But those sorts of tricks are obvious. Let's stay with this definition of ideal.

Why is solubility limited? Because you first have to "virtually melt" the solute at your chosen temperature. The higher the MPt and the larger the ΔH_f, the harder it is to virtually melt the solute, so the lower the solubility. If you assume the validity of lattice theory (https://www.stevenabbott.co.uk/practical-solubility/lattice.php) then an ideal solvent is one where the solute-solute, solvent-solvent and solute-solvent interactions are identical. But first they both have to be liquids! That's why we have to virtually melt the solute.

A key reason ideal solubility theory isn't too helpful (in addition to some questionable assumptions) is that the correction for heat capacity is complicated and most of us don't know the relevant values.

Irritated by the fact that for 1000s of compounds with known solubility, the ideal solubility calculations were either too hard or too inaccurate, Yalkowsky introduced his famous approximation. Select the Yalkowsky option and you get this curve:



MPt °C 150

ΔH_f kJ/mole 30

ΔC_p J/mole.K 134

$T_{measure}$ 25

☑ Use Yalkowsky Approx.

x @ T: 0.057

LogK_{ow} 2

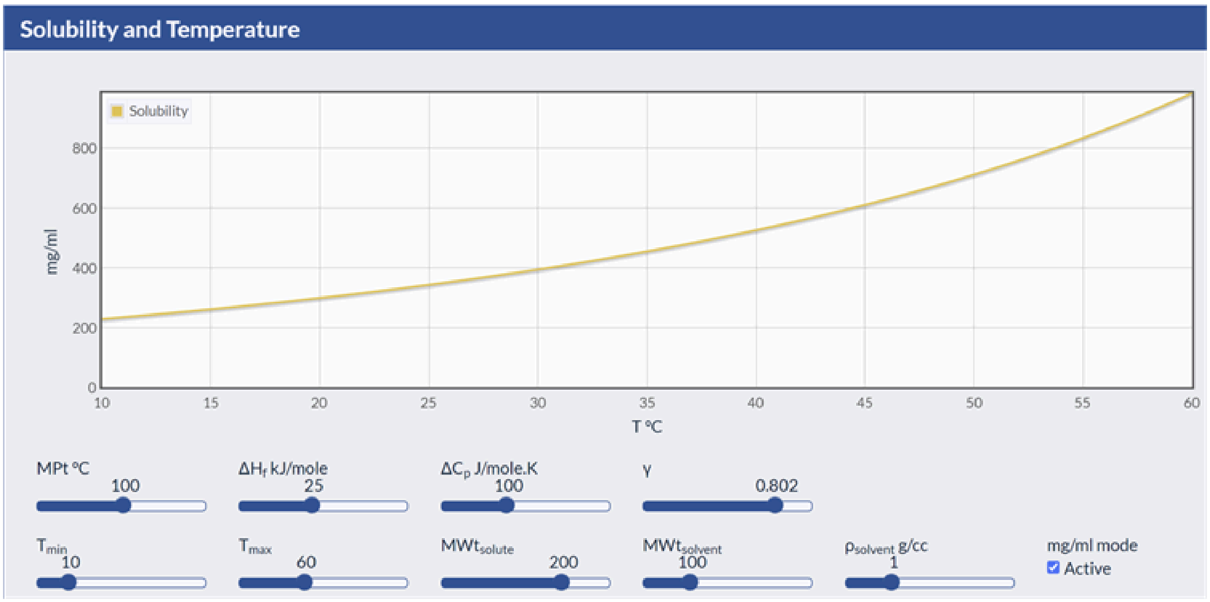
Log(S) Aqueous Sol. via GSE: -2.75

This isn't too far away from the standard ideal equation, but uses only the MPt, T_m to calculate the mole fraction x:

$$ln(x) = -0.023\left(T_m - T\right)$$

The full ideal solubility equation, for those who are interested, is found on the web page.

Non-ideal solubilty



MPt °C 100

ΔH_f kJ/mole 25

ΔC_p J/mole.K 100

γ 0.802

T_{min} 10

T_{max} 60

MWt_{solute} 200

$\text{MWt}_{solvent}$ 100

$\rho_{solvent}$ g/cc 1

mg/ml mode ☑ Active

https://www.stevenabbott.co.uk/practical-solubility/Solubility-T.php

This app plots that data in a more user-friendly format and adds the activity coefficient γ to take into account real-world solubility effects. It also allows you to plot in more user-friendly mg/ml units.

Where do you get γ values? Professional solubility software such as COSMO-RS readily gives values. Or you can make good-enough (relative) estimates via Dissolution_Hansen Solubility Parameters.

I'm being polite about Ideal Solubility. It really is a poor intellectual construct that has wasted too much effort around its crazy assumptions. There is a better way to think about it, but it will be some time before there's an app that can bring the superior approach to life. As soon as it's available, the alternative approach will be added to the FST.

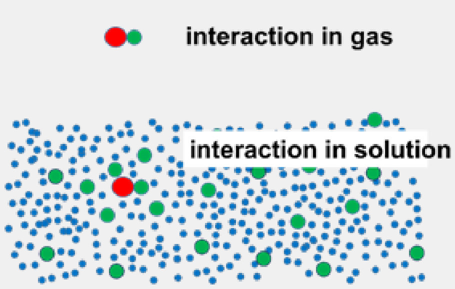
Dissolution Kirkwood-Buff

Links

Pharma Formulation

This chapter is courtesy of Dr Seishi Shimizu at U York, UK.

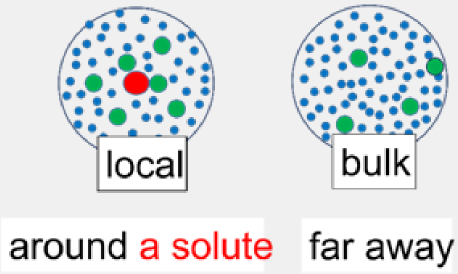
What does “interaction” mean in solution?



Thinking about a pair of molecules and its interaction (e.g., hydrogen bonding, charge interactions, and van der Waals forces) is enough in gas.

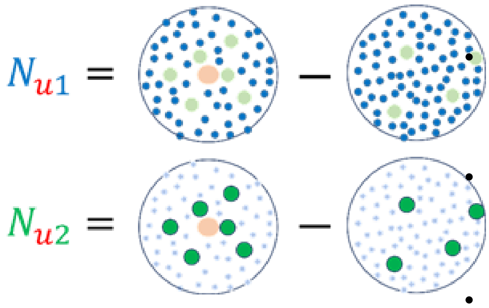
But around the pair in solution, there are many other molecules continuously moving about.

So, we are forced to think about the distribution of molecules to measure interactions.



The Kirkwood-Buff integral measures the interaction using the change of concentration from the bulk. Here you can see that on average there are more of the green molecules near the solute than are found on average.

Excess numbers and Kirkwood-Buff integrals



- If red-blue interaction is attractive, concentration of green around red is higher than in the bulk. This difference is the excess number of green molecules around a red molecule, N_{u1}.
- If red-green interaction is attractive, concentration of green around red is higher than in the bulk. This difference is the excess number of green molecules around a red molecule, N_{u2}.
- When the interaction is repulsive, N_{u1} or N_{u2} will become negative.

There is a limitation to this approach and how this could be overcome

- Blue molecules are more concentrated in bulk than green. This means N_{u1} can be larger than N_{u2} simply because there are more blue molecules than green. This poses difficulties when we need to compare red-blue and red-green interactions.
- To compare blue-red and green-red interactions, we need to account for the bulk concentrations, c_1 and c_2.

Therefore, the measures of interactions are given by the Kirkwood-Buff integrals as

$$G_{u1} = \frac{N_{u1}}{c_1}$$

$$G_{u2} = \frac{N_{u2}}{c_2}$$

Experimental determination of the Kirkwood-Buff integrals

- Kirkwood-Buff integrals can be determined from solubility measurements, when combined with density and activity data.
- Kirkwood-Buff integrals can be estimated from molecular dynamics simulations.
- Thus, Kirkwood-Buff integrals links bulk measurements to molecular distributions and experiments to simulation.

Specific examples are available at https://www.stevenabbott.co.uk/practical-solubility/kb.php. Many of the apps allow you to load your own “normal” datasets and analyse them via KB theory.

Dissolution Solubilizers and Hydrotropes

Links

Deodorant Sticks, Laundry Liquids, SkinCare, Pharma Formulation, Microencapsulation

Although we have reasonable theories for classical solubility and for oil-water microemulsions, there is a large area of formulation space where we rely on "solubilizers" to get solutes into water. Until recently, the explanations for these effects were incoherent. With the recent Kirkwood-Buff approach we can at least understand the effect of a solubilizer even if we cannot, yet, predict it.

Confusing terminology

Getting something into solution via non-classical solubility has been described by many names. Hydrotropes, co-surfactants, pre-ouzo effect, "surfactant-free emulsions", kosmotropes/chaotropes, entrainers (for scCO2, see below) and solubilizers all appear in the literature. For the rest of this chapter we shall just call them solubilizers as that's a handy generic name that describes the purpose, without the extra baggage associated with other names.

Bad explanations

There are two aspects to bad explanations:

1. They are factually wrong;
2. They don't help.

The world of solubilizers has suffered from both aspects. Lots of words have been expended on things like "water structure" as "explanations" for these effects, yet no one has been able to formulate better based on those wrong and unhelpful approaches.

Typical examples are:

- Water structure. The solubilizer changes the structure of the bulk water, allowing the solute to dissolve;
- Pseudo-micelles. The water makes the solubilizer form pseudo-micelles which, just like real micelles, then dissolve the solute;
- Complexes. The solute and solubilizer form specific complexes that, for some reason, aid solubility.

The water structure hypothesis is especially popular because (a) we know that water *is* structured and (b) we can give fancy names such as kosmotrope and chaotrope to molecules on the grounds that they make or break water structure. Sadly, after a century of the use of these terms, no coherent formulation understanding has emerged – for the good reason that water structure does not play a role in solubilizer effects.

A good explanation

We have 3 components: the water, 1, the solubilizer, 2 and the sol*u*te u. Just like lattice theory (see Dissolution_Flory-Huggins and https://www.stevenabbott.co.uk/practical-solubility/lattice.php) we can think of all possible interactions. In this case we have 1-1 (water structure), 1-2, 1-u, 2-2 (pseudo-micelles), 2-u and u-u.

Because we are trying to solubilize something with a low solubility, we know that 1-u interactions are unfavourable (if they were favourable, u would be soluble) and, at the low concentrations we can obtain, u-u interactions are irrelevant. So we can focus on the other interactions.

Using assumption-free statistical thermodynamics via Dissolution_Kirkwood-Buff theory we can nicely compute how these interactions play out in terms of solubility or, rather, rate of change of chemical potential of u, $\delta\mu_u$ with respect of the concentration of 2, δc_2. The G terms are Kirkwood-Buff Integrals, KBI, which describe how much two things like being next to each other *compared to their statistical average*. So a large positive G_{ij} implies that the molecules are positively attracted to each other:

$$\frac{\delta\mu_u}{\delta c_2} = -RT\frac{G_{u2} - G_{u1}}{1 + c_2\left(G_{22} - G_{21}\right)}$$

Let’s take the equation slowly. For good solubilization we want the rate of change of chemical potential with concentration to be as large and negative as possible. RT is what it is. Strangely, the concentration of the solubilizer is on the bottom, which means the higher it is the less negative. Clearly that c_2 term can’t drive the solubilization.

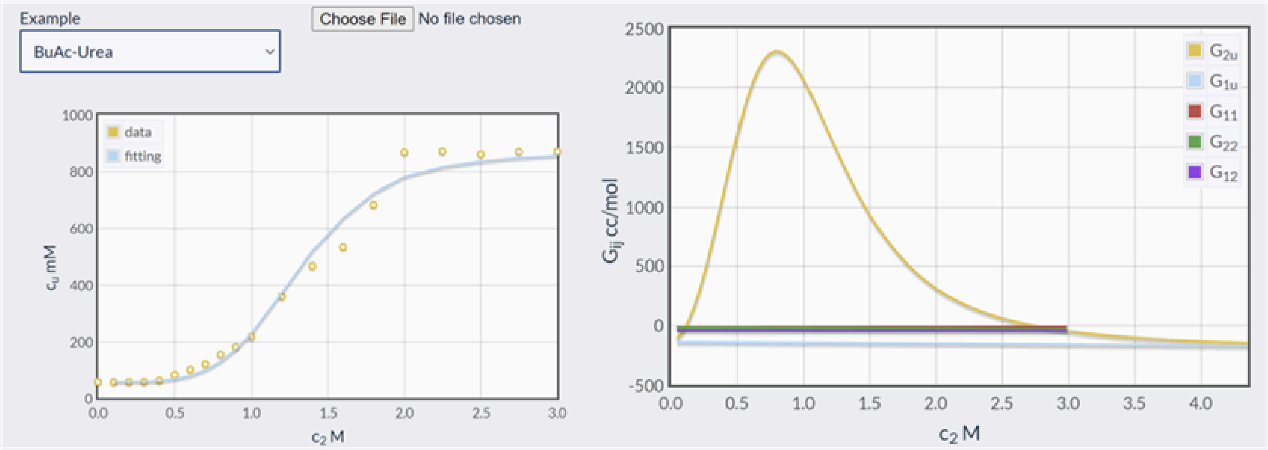
While we’re looking at the bottom, note that a large G_{22} decreases the effect. G_{22} is a measure of the amount of “pseudomicelles” in the system. Contrary to the popular explanation, pseudomicelles make things worse.

We already know that G_{u1} is small because the solute doesn’t much like the water, which is why we’re adding a solubilizer.

Before we reveal the true answer, look at the role of G_{11} the water structure term. It’s not in the formula! Water structure (kosmotrope or chaotrope) has no influence on solubilization, despite a century of opinions that it was *the* explanation.

Finally we see the solubilization is driven by G_{u2} the tendency towards mutual clustering of the solute and solubilizer. This is very different from the idea of a “complex” between them. When you plot G_{u2} against c_2 (as we shall do in a moment), you see G_{u2} rise and fall. We can understand the “rise” part – as you get more solubilizer there’s a bigger chance of mutual clustering. Why does it fall? Because G_{u2} is the *difference* of numbers in the vicinity from the statistical average, and as c_2 increases, the average increases so the difference decreases.

If we go to the app that describes typical solubilizer behaviour, we can get lots of information. But the dominance of the G_{u2} curve and the irrelevance or even counter-effect of G_{22} is also apparent.



Example: BuAc-Urea

Extracts from https://www.stevenabbott.co.uk/practical-solubility/kb-hydrotropes.php

The left-hand graph from the app shows the solubility of butyl acetate in water as urea (a well-known solubilizer) is added. The right-hand graph shows the G_{ij} values. The only large values are from G_{u2} (for KBI, $G_{u2} = G_{2u}$), with all others being irrelevant.

What the equation above cannot explain is the sigmoidal curve characteristic of these solubilizers. That's a bit more complicated – it's the solute-induced solubilizer clustering. G_{u2} becomes large because the solute encourages more solubilizers to gather around it.

The KB approach is assumption-free, not at all hard to implement and gives clarity of explanation in complicated systems. Yet it's still rarely used. Many are comfortable using hand-waving non-explanations and feel threatened by ideas that clearly disprove everything they've said for the past decades. But younger minds (and the brightest of the older generation) tend to adopt approaches that work, so the KB approach is spreading steadily.

scCO2

Looking at the solubilizer equation you can say "Yeh, just another equation". But it has a beauty to it in that it's attempting to balance all possibilities in a simple form. It would be nice if there was another example of such an equation from a different world.

As is well-known, supercritical CO2, scCO2, is a potentially interesting solvent because you can release your solute simply by letting the CO2 disappear when you release the pressure. The problem is that CO2, liquid or supercritical is a useless solvent – maybe similar to pentane. There's not much you can dissolve in pentane. So scCO2 extraction uses "entrainers", molecules at a low level (a few %) which dramatically increase the solubility. Water would be the most popular entrainer, and it indeed works for caffeine, if it were itself more soluble in scCO2. Because of its low solubility, people use molecules like ethanol or acetone instead.

The entrainer effect has been explained in, you guessed it, 3 potential ways:

- CO2 structure. The entrainer changes the structure of the scCO2, allowing the solute to dissolve;
- Pseudo-micelles. The scCO2 makes the entrainer form pseudo-micelles which, just like real

micelles, then dissolve the solute;
- Complexes. The solute and entrainer form specific complexes that, for some reason, aid solubility.

In reality, the equation driving entrainment is (see the app at https://www.stevenabbott.co.uk/practical-solubility/scco2.php):

$$\frac{\delta\mu_u}{\delta c_e} = -RT\frac{G_{ue} - G_{u1}}{1 + c_e\left(G_{ee} - G_{e1}\right)}$$

Yes, that's a copy/paste of the equation above, using e for the entrainer rather than 2 for the solubilizer. The entrainer effect is dominated by solute-entrainer interactions, entrainer-entrainer interactions reduce the effect and CO2 "structure" isn't even included in the equation.

In this case, (see the Simple scCO2 app, https://www.stevenabbott.co.uk/practical-solubility/simple-scco2.php) we can analyse the large literature on scCO2 solubilities to work out what constitutes a good entrainer for a given molecule, i.e. what would create a large G_{ue}? The answer is simple. If the solute has no potential for H-bonding, there's no significant entrainer effect from anything. If the solute is capable of H-bonding, then any entrainer with the right functionality to form H-bonds will help.

As with solubilizers, a simple equation brought the clarity necessary to get to the heart of the entrainment effect. The H-bonding idea had been proposed some decades earlier, but the scCO2 community expended its energy exploring the wrong ideas, wasting those decades.

Nearly every scCO2 paper starts with "scCO2 is a green solvent …", which is fine. But to be green you must not waste resources. As pointed out by Abbott and Shimizu in a paper and a book chapter, a few 10's of smart scCO2 entrainer experiment could have sorted out how and when scCO2 might be a useful solvent for a given process. Instead, we've had 100's of experiments using ad hoc approaches, assumptions, analyses and wrong explanations that have wastefully produced no usable common understanding. It was genuinely painful to go through all those papers to dig out the nuggets of understanding contained in the Simple scCO2 app.

The KB-χ

The fact that the same beautiful equation describes different processes suggested to Dr Seishi Shimizu a universality to the KB equation. If we cut out the bits which will seldom be important for these sorts of problems we end up with:

$$\frac{\delta\mu_u}{\delta c_2} = -RT\frac{G_{u2}}{1 + c_2G_{22}}$$

We can imagine the ratio of G_{u2} to $1 + c_2G_{22}$ as capturing the same essence as the Flory-Huggins χ value; the balance of competing attractions. Calling this the KB-χ provides a unifying name for the wide world of solubilizers.

Pre-ouzo, surfactant-free emulsions and more

Take a solution of anethole or thujone (ouzo and absinthe respectively) in 40% ethanol/water and add a lot of water. It becomes cloudy as the oily flavour molecules become insoluble. But if you add smaller amounts of water, the solution becomes "sparkly", there are nano-sized fluctuations near insolubility. If you do this for something like a fragrance then you have a "pre-ouzo" formulation. If you like to use a different terminology and, maybe, have less sparkle, then you can make a "surfactant-free emulsion".

In all these cases you have water, something reasonably water soluble such as an alcohol and your desired molecule such as a fragrance.

The sparkle show that the molecules are gathering into fluctuating clusters. Another word for Kirkwood-Buff theory is Fluctuation Theory, and the KB approach is ideal for describing what is going on. The relevant formulae are more complex because these wild fluctuations are near the borderline of phase separation (always a risk with these formulations) but once more they do a much better job than the non-alternatives.

Although these formulations have attractive names, there's no escaping the fact that you have to add lots of what amounts to a co-solvent, which diminishes their attractiveness as an alternative to ethanol (though this is an important consideration in some cultures) or to surfactants. For example, in a fragrance formulation, unless the added smart, green solubilizer is volatile, it's left sitting on the skin, which consumers may not appreciate.

A smart, efficient microemulsion created via Surfactancy_HLD also has a sparkle from nano-sized scattering.

Evaporation Basics

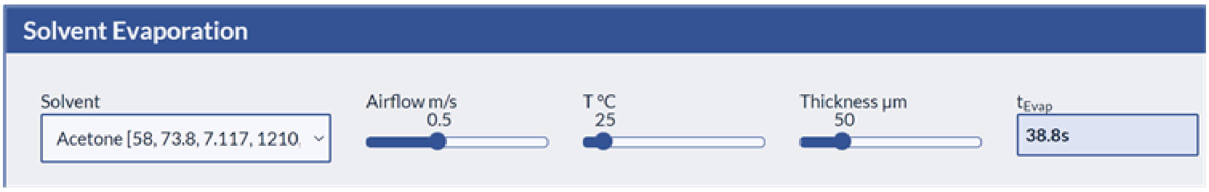
Links

Sun Screens, SkinCare, Fragrances, Surface Cleaning, Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint, Inkjet Inks

The evaporation rate of solvents depends on:

1. The temperature;
2. The vapour pressure of the solvent at that temperature;
3. The air velocity.

This app contains the Antoine Coefficients (see Evaporation_Temperature and Antoine Coefficients) of many common solvents – which let you calculate the vapour pressure at your chosen temperature. It also include the effect of air flow. With zero air flow there is almost zero evaporation.



Solvent Evaporation

Solvent: Acetone [58, 73.8, 7.117, 1210,

Airflow m/s: 0.5

T °C: 25

Thickness µm: 50

t_{Evap}: 38.8s

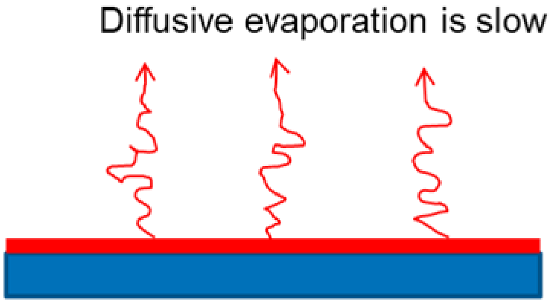
https://www.stevenabbott.co.uk/practical-coatings/evaporate.php

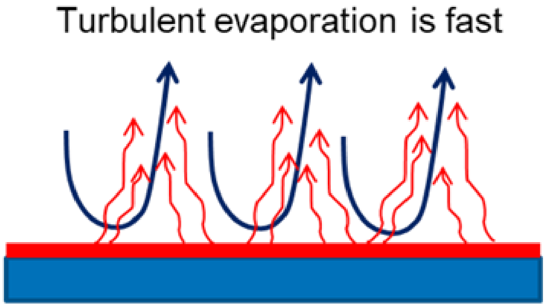
Solvent evaporation depends on the (temperature dependent) vapour pressure of the solvent (blend) and on the air speed across the surface of the solvent. The dependence of the vapour pressure P on temperature T is governed by the Antoine Equation and its three parameters, A, B and C:

$$\log_{10}P = A - \frac{B}{C+T}$$

The app has the Antoine Coefficients for the common solvents and the more powerful HSPiP package has many more.

The dependence on air flow is not obvious. The problem is that diffusion of the solvent vapour is very, very slow. For realistic evaporation, the vapour has to be swept away by turbulent air flow:





Although those who apply typical coatings don't realise it, a typical lab air flow over a typical surface might be 0.25-0.5 m/s, more than enough to give reasonable evaporation rates. Obviously drying with little air flow is both inefficient and dangerous in terms of solvent vapour build-up. In real dryers, air flows are much larger.

The equations used in the app have a complicated dependence on Reynolds Number, Schmidt Number and Sherwood Number from which the "Mass Transfer Coefficient" is calculated. You can read them in the app. If you have to do calculations without the app and have an evaporation rate at a known velocity, v, the standard approximation is that rate is proportional to $\sqrt{v}$. This tells us why we need so much airflow – to double the drying speed needs 4x the air velocity.

Those who speak of "IR driers" or even "Laser driers" love to talk about the efficiency with which their heat energy is absorbed. But heat can't dry. Drying is removal of molecules and only air flow can do that. If you want an "IR heater" then that's fine. If you want an "IR drier" make sure the device comes with plenty of airflow.

Diffusion limited

At some point, the rate at which solvent leaves the coating becomes limited by the rate at which the solvent can diffuse out of the coating. This is a complex issue involving not just standard Fickian Diffusion Diffusion_Basic Diffusion but also Diffusion_Concentration-Dependent Diffusion.

As discussed in the Evaporation_Diffusion limited chapter, if your system becomes diffusion limited, the velocity of air above the coating is irrelevant to drying speed. The only way to speed up diffusion-limited drying is to raise the temperature. For a formulator, a key consideration is to keep the coating "open" (e.g. a high-boiling solvent, a better HSP match) for longer, delaying the transition to diffusion limited.

Evaporation Diffusion limited

Links

Water-based Adhesive, Adhesive, Solvent-Based Paint

> Evaporation by sweeping away air molecules from the surface is the fastest way to remove solvent. But at some point, the rate at which solvent can diffuse to the surface is slower than the rate at which the solvent is swept away. Now we're in the realm of diffusion-limited evaporation.

A simple, unusable equation

There's a relatively simple equation for the rate of diffusion-limited evaporation, E, from the surface:

$$E = \frac{K_1}{h} e^{\frac{-K_2}{C}} e^{\frac{-K_3}{T}}$$

This tells us that the E depends on three terms:

1. A constant K_1, plus a 1/thickness (h) term as a reminder of the obvious fact that thinner coatings provide a reduced barrier to diffusion;
2. An exponential term, depending on K_2 which gets smaller (slowing evaporation) as the solvent concentration, C, gets smaller because this is a classic case of Diffusion_Concentration-Dependent Diffusion. "Smaller" means 2 or 3 orders of magnitude, which is why it is so hard to get rid of the last residues of solvent;
3. An exponential term, depending on K_3 which gets larger (increasing evaporation) as T increases, which tells us the obvious: to get rid of residual solvent you need to increase T as much as possible.

In an ideal world we would all know K_1, K_2, K_3 but in practice almost no one knows what they are because it is too much work to determine them. That's why there isn't an app.

Detecting diffusion limited evaporation

The effects of Evaporation_Enthalpy of Vaporization mean that in a hot oven with a good airflow, the temperature of the drying surface is low. For a water-based coating, the oven can be at 120°C and the surface will only be in the mid 40's C. In general, you want maximum airflow, so you put plenty of energy into your fan system.

As soon as the surface temperature starts to climb you know either that you are fully dry (unlikely, except for very thin coatings, hence the 1/h term in the equation) or that you are in diffusion-limited mode. Because of the K_3/T term, you now want the oven to be at the maximum reasonable temperature … but fan speed is now irrelevant, so you can save energy by turning the fans down to whatever minimum flow gives you the temperature you need.

This simple switch from high to low airflow should be obvious. But this chapter is written because it is all too common to see final zones of drying systems with maximum fan speed “to drive out the remaining solvent”. Nothing the fan can do can make the solvent diffuse faster through the coating.

It is also all too common to save on temperature measurement inside the oven, so users often have no idea when the system switches from evaporation-limited to diffusion-limited. This absence of knowledge makes it hard to optimise the system for speed and efficiency.

A good example is when drying a temperature-sensitive coating. If it must never go above 60°C it seems dangerous to set the oven temperature to 120°C. But if during that part of the process the coating is at 47°C (which is the case for a water-based coating thanks to the large enthalpy of vaporization of water) there is no reason not to get the most drying in the shortest time. What is important is knowing when diffusion limited drying kicks in, because then the temperature will rise dangerously high.

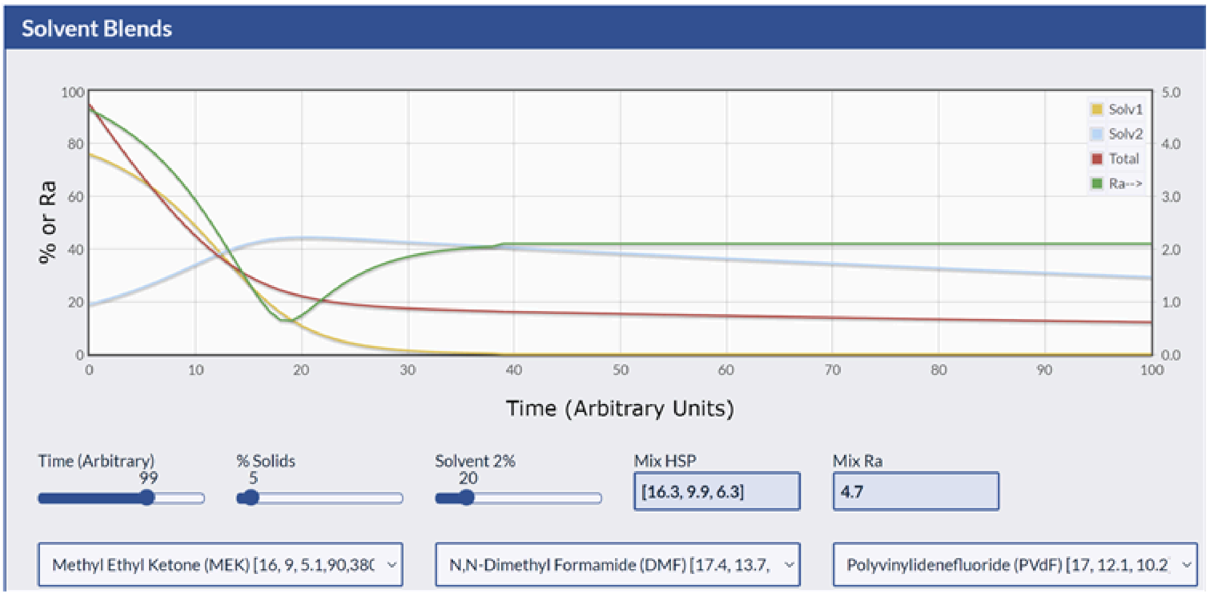
A counter example of drying too fast thanks to a poor choice of solvent is connected to the phenomenon of “skinning”.

Skinning and the myth of “drying from the reverse side”

Suppose you have a mixed solvent system where the combination is OK for the formulation, but where the better solvent evaporates faster. You might have chosen this fast-evaporating solvent to make drying faster, but the result is to make things worse. The idea of a blend is a good one, based on Dissolution_Hansen Solubility Parameters but you really need the worse solvent to evaporate faster. In the case of the better solvent evaporating faster, the surface layer might be rapidly depleted of good solvent, leading to the formulation crashing out and becoming diffusion limited. This is “skinning” and is a big barrier to rapid, high-quality drying. Usually skinning is identifiable too late – when the coating emerges from the oven, sometimes with blisters if the internal temperature of the coating exceeds the boiling point of the solvents. With an internal temperature monitor it is easy to see a sudden increase in coating temperature that *should* be a sign of drying being nearly complete but common sense will tell you that the onset is far too early.

There are two approaches to solving skinning – one wrong, one right:

1. The *wrong* way is to try to “dry from the reverse side”. This is a common idea which is literally meaningless. Drying can *only* take place from the surface and applying heat from the reverse side cannot alter this fact. To everyone’s surprise, the thermal conductivity of most reasonable wet coatings is so large that there is no significant temperature difference across the coating (it *might* be large enough to encourage Evaporation_Marangoni but is irrelevant to efficient drying).
2. The *right* way is to change the solvent blend so that the worst solvents evaporate faster, allowing the good solvents to keep the formulation “open” for as long as possible. This allows the whole system to dry faster (even if you end up using solvents with a higher boiling point!), and also leads to reduced stress in the coating and to higher gloss when you ensure that the very last bit of solvent to leave is the very best solvent for the formulation.



Time (Arbitrary) 99

% Solids 5

Solvent 2% 20

Mix HSP [16.3, 9.9, 6.3]

Mix Ra 4.7

Methyl Ethyl Ketone (MEK) [16, 9, 5.1,90,38(

N,N-Dimethyl Formamide (DMF) [17.4, 13.7,

Polyvinylidenefluoride (PVdF) [17, 12.1, 10.2

https://www.stevenabbott.co.uk/practical-solubility/solvent-blends.php

The example of MEK:DMF and PVdF shows how the "bad" MEK evaporates quickly leaving the "good" DMF behind, with the starting Ra (a measure of the HSP Distance between solvent blend and polymer) at a relatively high value because of the 80% MEK, going to a minimum at 26% MEK then reaching a "good" value of 2 once the MEK has all gone. This strange example exists because as described in Dissolution_Dissolution Kinetics if you try to dissolve PVdF powder in DMF you get a horrible gel, but if you add the DMF to a poor dispersion of the PVdF powder in MEK, it dissolves nicely.

Evaporation Enthalpy of Vaporization

Links

Fragrances

When we look at Evaporation_Basics, the main concern is the vapour pressure of the solvent and the airflow trying to sweep it away. It is obvious that higher boiling solvents will evaporate more slowly than lower boiling ones.

But when we are doing serious drying, using large airflows, we often find that high boiling solvents disappear surprisingly quickly … compared to water.

Evaporative cooling

The enthalpy (or latent heat) of vaporization or of evaporation (the terms are interchangeable) is in J/mole or, more practically, J/kg, telling us how much energy is required to change a liquid into a gas. We have a (literal) feel for the concept because even a warm breeze when we come out of the water after a swim can make us feel cold because our skin provides the energy needed for the water to evaporate and be swept away.

If you place an aqueous coating into a drier with a good airflow, even if the air is at 140°C, the coating itself (assuming it's not on a thick, solid surface which conducts heat quickly) might only be in the 40-50°C range. Why? Because if the temperature was higher the water would evaporate faster and soak up more energy, and if the temperature was lower, the slow evaporation would not require so much energy and the temperature would rise. Technically the coating is at the "wet bulb temperature".

A handy approximate formula gives us the answer for water with a BP, T_b = 373°K. It tells us that the temperature of the coating, T_c is given by:

$$T_c = 137\left(\frac{T_b}{373}\right)^{0.68}\log_{10}\left(T_{air}\right) - 45$$

For T_b=373 then we have 137 $\log_{10}$(413)-45 = 313°K = 40°C.

This has positive and negative implications:

- +ve If you have a sensitive aqueous coating that must not exceed, say, 60°C, you can use a hot oven to dry more quickly while remaining in a safe temperature range
- +ve High boiling solvents with a low enthalpy of vaporization are surprisingly easy to dry in an oven with a fast airflow. In a reasonable drier, a solvent like NMP with a BP of 204°C evaporates ~2x faster than water because its 550 J/g is so much smaller than water's 2400 J/g. The coating dries at ~95°C in the same oven where water is at ~40°C.
- -ve Although water is green, the fact that its enthalpy of vaporization per gram is 3x to 5x higher than many less green solvents significantly reduces its attractiveness in terms of PCF (Product Carbon Footprint) unless the energy needed to evaporate the solvent is itself entirely renewable.

If you need to model the drying process, the science is well-known (see Evaporation_Heat and Mass Transfer) but tricky to implement. The much-used drying model in the TopCoat package from RheoLogic is excellent though (a) you have to pay for it and (b) I wrote the model so, just as with the links to HSPiP in which I also have a financial interest, you need recognise that this advice is potentially biased.

Evaporation Heat and Mass Transfer

Links

When we provide flows of hot air to help evaporate a solvent we are mostly engaging in heat transfer, supplying (costly) energy, in order to achieve mass transfer, which is moving the solvent from the coating. Direct calculation of mass transfer is less intuitive than calculation of heat transfer, but fortunately a trick allows us to jump directly from heat to mass.

Energy in …

We are providing plenty of W of energy via a heated airflow which can be astonishingly ineffective at delivering the heat to the coating. This is because of the no-slip boundary condition which says that the velocity of a fluid at the surface (which is where we want a high velocity) is zero. We can improve the situation by providing turbulent flow, but this needs clever design of delivery of the hot air.

From elementary considerations of Evaporation_Enthalpy of Vaporization we know how much heat we need to put in to be able to evaporate the solvent.

Mass out …

It might seem simple to combine the W in and the enthalpy of vaporization to give us a mass of evaporating solvent per unit time.

But it's not. The problem is the uncertainty of the transfer of heat from the hot airstream to the surface of the coating (and then into the coating itself via thermal conduction). In some circumstances it is possible to calculate the HTCF, Heat Transfer Coefficient, but in most cases the system is too complex so, instead, we have to measure it.

This still leaves the problem of how the HTCF translates to the MTCF, Mass Transfer Coefficient.

The Gutoff method

The pragmatic approach is to use the method developed by Gutoff for his pioneering work on practical models for industrial dryers.

Given, at any instance, a temperature difference, ΔT, between the dryer temperature and the coating, given a HTCF, a constant K which we will discuss shortly, a partial pressure, P, of solvent at the coating temperature, a heat capacity C_P and weight W_T of the coating+substrate then the change of temperature with time $\frac{\delta T}{\delta t}$ is given by:

$$\frac{\delta T}{\delta t} = \frac{HTCF.\Delta T - \frac{HTCF}{K}P}{C_P W_T}$$

K contains the tricky conversion between our known HTCF and the required MTCF, the mass transfer coefficient and is given by:

$$K = \frac{Le.\pi.M_{Air}C_{PAir}}{\lambda M_{Solv}}$$

Here we have the heat capacity of air, C_{PAir}, atmospheric pressure, $\pi = 100$ kPa, the molecular weights of air and of the solvent, M_{Air} and M_{Solv}, the latent heat of vaporisation, λ and the Lewis constant, Le which is the ratio of thermal diffusivity and mass diffusivity. Calculating Le is a challenge but happily, to a good approximation, Le = 1 for typical dryers. Happily, too, the heat capacity of air is 1 J/g to the level of accuracy required. The λ and M_{Solv} values are calculated from the solvent data and the current % of each solvent.

So, at each position in the oven we know δT/δt. The next position is assigned a temperature of T+ Δt.δT/δt, where Δt = ΔL/V is the time taken for the web to move by the step length ΔL at web velocity V in m/s.

So far, all we've done is changed the temperature of the web. There is a second equation that tells us how much weight of solvent is lost in that same time step. It uses the same ΔT, the difference between oven and web temperatures.:

$$\frac{\delta W}{\delta t} = \frac{HTCF.\Delta T - C_P W_T \frac{\delta T}{\delta t}}{\lambda}$$

So we use the calculated δT/δt (modified by heat capacity and total weight, W_T of coating+substrate) not only to increase the web temperature in the first equation but also to calculate the solvent loss in this time step.

Clearly, implementing all that for a given coating system is not trivial and there isn't a convenient app for it. The important point is that drying involves a subtle balance of heat and mass transfer which in turn depend on solvent parameters plus the weight of coating and substrate.

In the end, for rapid drying of a given coating formulation, all you can do is optimize the HTCF, so let's see how to measure it.

The Gutoff approach is tricky to implement. The much-used drying model (grounded in the Gutoff approach) in the TopCoat package from RheoLogic is excellent though (a) you have to pay for it and (b) I wrote the model so, just as with the links to HSPiP in which I also have a financial interest, you need recognise that this advice is potentially biased.

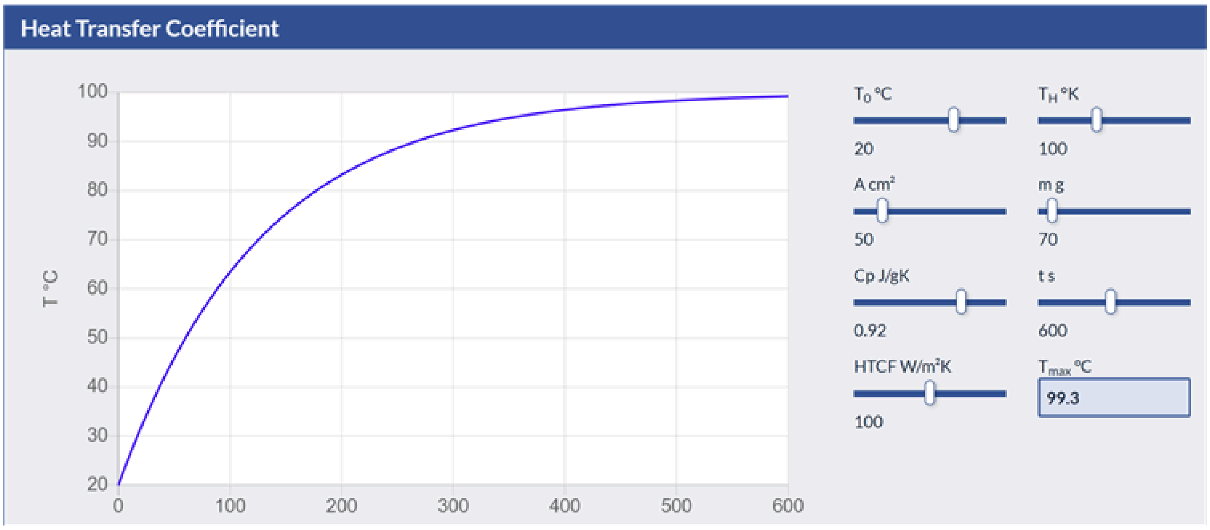
Measuring HTCF

The HTCF is in W/m²K. To measure it you take a modest cuboid of a material with known density and heat capacity (a lump of Al is perfect) and from its dimensions you know the area, A, exposed to the heat. You put a thermocouple inside it, ensure it's equilibrated at room temperature then place it into the dryer and plot the temperature with time.

From the simplified equation:

$$\frac{\delta T}{\delta t} = \frac{HTCF.\Delta T}{AC_P W_T}$$

it is easy to fit the curve with something like Excel. Or you can slide sliders in the app to match your experimental data:



Heat Transfer Coefficient

T_0 °C
20

T_H °K
100

A cm²
50

m g
70

Cp J/gK
0.92

t s
600

HTCF W/m²K
100

T_{max} °C
99.3

https://www.stevenabbott.co.uk/practical-mechanical/Heat-HTCF.php

Although the basic setup and data analysis are simple, measuring the HTCF for your specific equipment will pose some practical problems, so you need some ingenuity. If, for example, you want to measure HTCF in the various zones of a coating machine, although you might be able to fix your block of Al to the web, and it might be possible to get it through the machine, how do you measure the temperature? The 21st century has lots of ingenious, low-cost devices that you can solve problems like this. Be assured that having the ability to measure the HTCF of every zone, under a range of conditions makes the pain needed to find a suitable technique very much worthwhile.

Evaporation Humectants

Links

Hair Conditioners, Sun Screens, SkinCare

> It is surprisingly tricky to work out how much of each humectant is needed to keep a required amount of water within the end product.

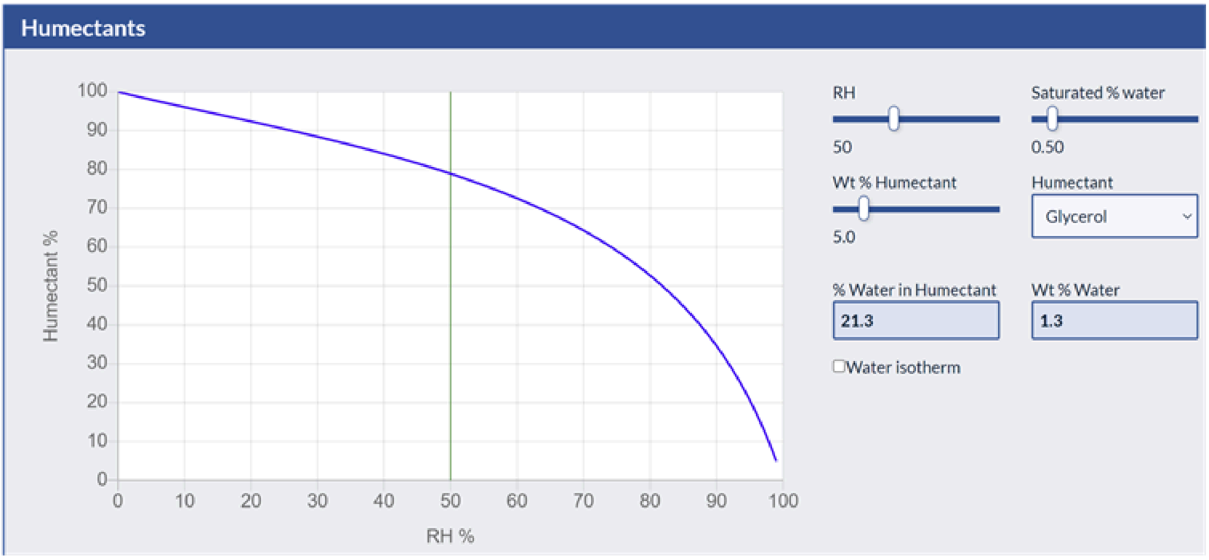
Norrish

The first thing we need to know is what % water remains in a humectant in contact with air at a given RH. The Norrish equation works in terms of mole fractions, x, with x_w being water and x_h being humectant with, conveniently, $x_h = 1 - x_w$. The equation involves a constant K, known, in principle, for each humectant:

$$x_w = \frac{RH}{e^{-Kx_h^2}}$$

Knowing the MWts of water and the humectant you can work out the wt % of water in the humectant at any given RH.

Now assume a small saturated % of water in the system (which might be 0) and assume (the errors are small) that his is linear with RH. Now provide the wt% of humectant in the system and you get the final water weight %.



RH: 50

Saturated % water: 0.50

Wt % Humectant: 5.0

Humectant: Glycerol

% Water in Humectant: 21.3

Wt % Water: 1.3

☐Water isotherm

https://www.stevenabbott.co.uk/practical-coatings/Humectants.php

In this example we have glycerol at 50% RH. It takes on 21% of water. We've added 5% of glycerol so that's ~1.05% water from the humectant. There's also 0.25% from the system itself, 50% of the saturated value. In total we have 1.3% water.

Assumptions and alternative views

Maybe the assumptions are too simplistic, or maybe we aren't adding humectants simply to hold equilibrium amounts of water. Here are some possibilities.

- Somehow the humectant must be nicely integrated into the formulation but behave according to the Norrish equation to estimate the % water. This sounds unlikely.
- A variation on that assumption is that the humectant doesn't react into the system. If it reacted then the humectant effect must disappear.
- That the Norrish equation is symmetrical for absorption and desorption. Typically, desorption isotherms show higher amounts of water than absorption (hysteresis). Maybe we add humectants to *delay* a move to equilibrium moisture content.
- The water sorption into the rest of the formulation is linear with RH up to your stated maximum. This is unlikely, but errors based on the assumption are small
- The extra water absorbed does good things to the formulation. This seems unlikely, unless we are adding a lot of humectant to highly water-sensitive formulation. Outside extreme cases, why are we adding these humectants?
- During drying of a formulation, keeping things "open" helps everything to relax into a high-quality coating. Maybe (as per the desorption hysteresis) that's all we're doing.

What is clear is that there is a great lack of clarity about why humectants are added to the formulation. The simple assumption that "it's to increase the water content" does not seem to reflect the multiple circumstances for which they are used, nor the relatively modest increases predicted from the simple formula.

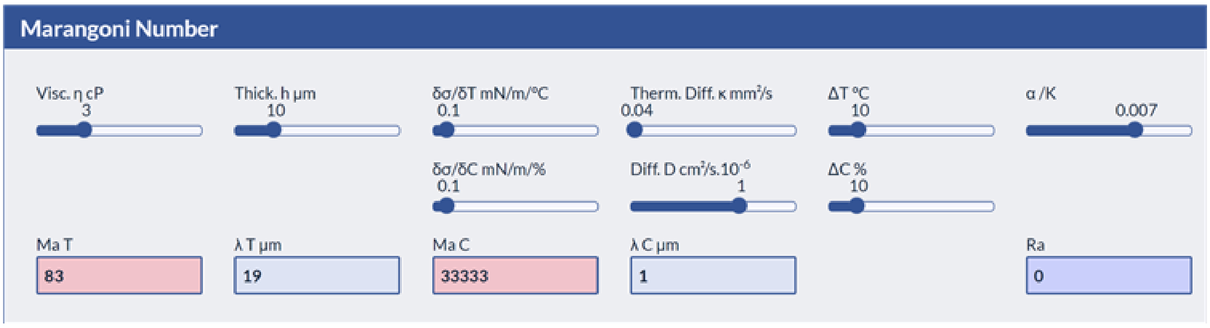
Users of the app have requested extra humectants to be added to the list. Given a reliable humectant Humidity_Water vapor isotherm it's not hard to fit it to the Norrish equation and add it. However, even for something as simple as glycerol there are varying isotherms. For example, a paper giving the isotherm for triethylene glycol gave their isotherm for glycerol and PEG400. Neither was compatible with the isotherms found in other sources. So TEG has not been added to the app.

Evaporation Marangoni

Links

Solvent-Based Paint

Many coating defects, from odd streaks, through orange peel to full-on hexagon patterns (yes, these really happen) are explained by the Marangoni effect. By understanding it, fixes can readily be found. There are active debates about whether this should be called Gibbs, Bénard, Rayleigh, in various combinations but it is common practice to focus on the general effect and call it Marangoni rather than worry about subtle distinctions.



Marangoni Number

Visc. η cP 3

Thick. h μm 10

δσ/δT mN/m/°C 0.1

Therm. Diff. κ mm²/s 0.04

ΔT °C 10

α /K 0.007

δσ/δC mN/m/% 0.1

Diff. D cm²/s.10⁻⁶ 1

ΔC % 10

Ma T 83

λ T μm 19

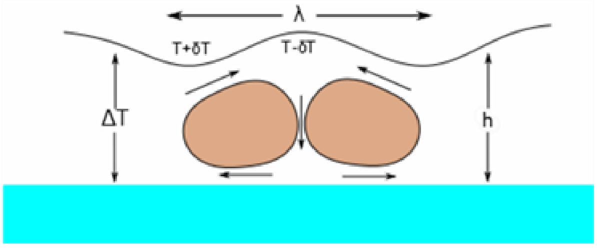
Ma C 33333

λ C μm 1

Ra 0

https://www.stevenabbott.co.uk/practical-coatings/marangoni.php

In the diagram we see that there is a temperature difference ΔT (maybe by evaporative cooling) between the base of the coating and the surface and that there is a small difference δT between one part of the surface and a neighbouring part, λ/2 away. That small δT creates a difference in surface tension, σ, (higher T means smaller surface tension) which depends on δσ/δT which is unique for each solvent. The surface ripples to accommodate the differences in surface tension.



Such small ripples might well go unnoticed, but they drive a flow from bottom to top. This too might remain unnoticed, but competing areas of rising and falling flow tend to organise themselves spontaneously into coherent structures, ultimately hexagonal cells, and the human eye can spot such defects very clearly. Most "orange peel" is a Marangoni effect – a proto-hexagonal effect can often be spotted in the semi-random defects.

I was a junior formulation engineer supervising the first coating trial of a fancy new machine. Looking at the coating end of the machine, everything was looking great … till those at the other end asked why the coating was covered with hexagons. We had no idea. Going to the wise head of research, when she heard the word "hexagon" she needed to hear no more. "It's Marangoni". And, after a brief explanation of the effect we could very quickly implement a root-cause cure.

An exactly analogous effect occurs when *concentration* differences at the surface create surface tension differences. This is especially common when using ill-matched solvent blends where the two solvents have very different surface tensions and very different evaporation rates allowing large concentration differences to build up spontaneously.

To avoid this you can:

- Use a solvent (blend) that shows a very low change of surface tension.
- Match the evaporation rates of the solvents to avoid concentration differences.
- Inhibit the circulation via a high viscosity,
- Use a small thickness which makes it harder for the circulation to build up.

All these factors are combined into a Marangoni number, Ma, and the idea of a Critical Marangoni Number, typically 80, above which a spontaneous pattern is likely to appear.

We have two effects – the temperature-driven and the concentration-driven, with the former much more discussed and the latter much more common. The equations for Ma are similar. For the thermal effect:

$$Ma = -\frac{\delta\sigma}{\delta T}\frac{h\Delta T}{\eta\kappa}$$

where κ is the thermal diffusivity, i.e. how readily the top surface temperature can equalise with the base.

For the concentration effect:

$$Ma = -\frac{\delta\sigma}{\delta C}\frac{h\Delta C}{\eta D}$$

where this time, D is the diffusion coefficient, where a larger value equalises concentrations faster.

Although you *can* look up all the values for your system and get a great calculation of Ma, it’s generally good enough to think what you can do to decrease the root causes. For example, those using xylene/butanol mixtures tend to use xylene/2-butanol rather than xylene/1-butanol because the evaporation rates of xylene and 2-butanol are better matched than with 1-butantol, so ΔC is likely to be smaller. Or you might choose to check the surface tension values of potential pairs of solvents and choose the pair with the smallest difference.

Look for the unexpected structure

If you find a coating defect, it can be a good idea to look closely under a microscope at the start or finish of the defect. We sometimes see streaks that might have multiple causes, but if you see something like a small hexagon at the start, that’s a good sign that Marangoni is the root cause, so the defects can be fixed by higher viscosity, thinner coating or a better solvent match.

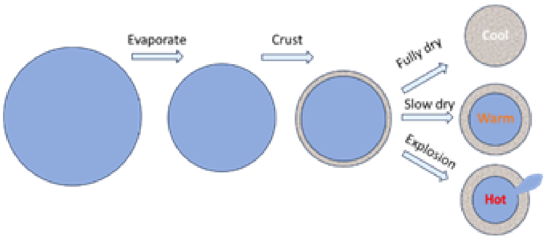
Attempts to coat a customer’s formulation kept failing because of “streak” defects. The customer was happy to blame our machine, but there was no obvious reason why the machine was causing the streaks. Down the microscope, a little hexagon at the start of each streak showed that it was a Marangoni problem. The customer had chosen a two-solvent blend using ethyl acetate as one of the components. The fact that it was more volatile than the other component hadn’t seemed a problem to them. Swapping to the less volatile butyl acetate immediately fixed the “streaks” problem.

Evaporation Spray Drying

Links

> Because spray drying needed to be described in the Microencapsulation chapter, this is a copy/paste of that text.

A formulator's view of spray drying



We have the same problem as with coatings. Spray drying, mostly from water, is mostly about getting a good atomizer and good airflows, neither of which is under the formulator's control. In this section we don't worry if we have a solution or an emulsion. The emulsion aspects are described in the next section.

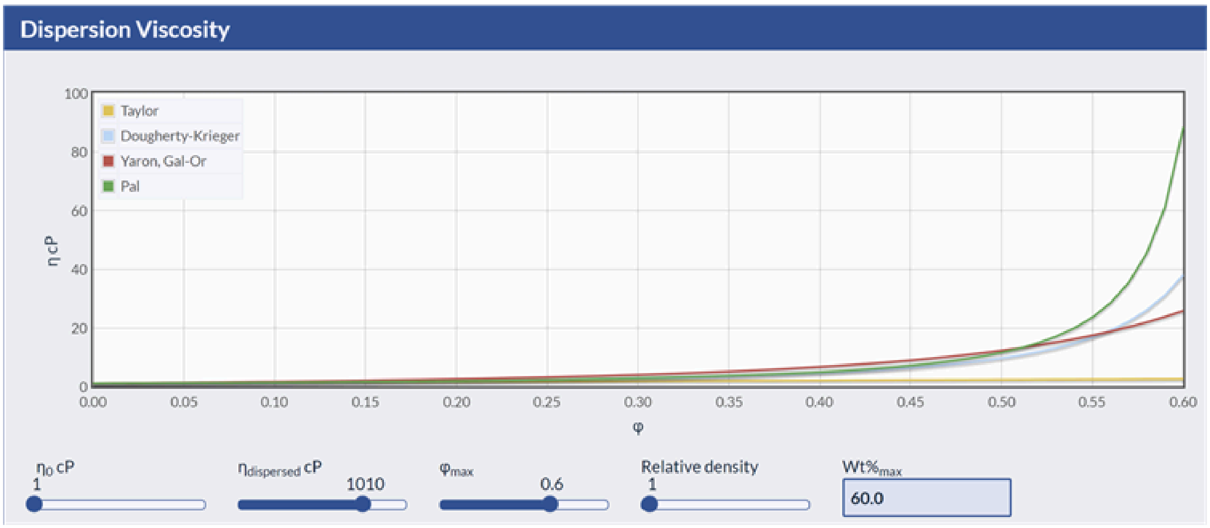
There are difficult tensions over which we *can* have some control.

- Production will always demand higher concentrations (less water to evaporate) and lower viscosity (easier atomization) which are usually contradictory.
- They also want a rapid transformation on the outside (crust formation) so that particles that touch each other or the wall don't stick, without that crust inhibiting rapid evaporation of the water. Such inhibition can slow down the process or, at worst, lead the water to explode as the internal temperature exceeds 100°C.
- We want maximum hot air temperature for rapid drying, but we don't want to destroy the contents of the particle, such as enzymes or friendly biome microbes.

Let's try to do something about those contradictions.

Maximum concentration, minimum viscosity

If our water contains particles or (it's the same thing) emulsion drops, then the same app shows the viscosity dependence on volume fraction, φ and on the viscosity of the dispersed particle (a few cP for a typical oil and for a solid particle 10,000 cP is good enough:


Dispersion Viscosity

η_0 cP 1

$\eta_{dispersed}$ cP 1010

φ_{max} 0.6

Relative density 1

Wt%$_{max}$ 60.0

https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

This tells us that without particle-particle interactions, just them getting in each other's way, the viscosity doesn't change much from water's 1cP till about 50% volume fraction when it reaches 10cP and is starting to grow rapidly.

But be careful. If your "water" starts at 10cP because of some soluble additive, then at 50% volume fraction the viscosity is now 100cP – the crowding effect is multiplicative, not additive. If you really need a few % more particles, do whatever it takes (and it's often relatively easy) to reduce your "water" viscosity by a few cP. This is an important and little-known trick.

If, however, your particles show significant interactions because, for example, you have a poor dispersant on solid particles, or because (and here is one warning about nanoclays for extra barrier properties) they have a high aspect ratio (length/thickness) then viscosities can shoot up.

We can explore these ideas in Flow_Particle Viscosity which includes this app:


High-Shear Particle Rheology

φ_c Percolation 0.27

φ_m Packing 0.63

N 1.6

Fractal Dim. d 2.8

Aspect Ratio A_f 5.5

$\varphi^* : \varphi_{c\ calc}$ 0.61 : 0.147

☑ Log(γ̇)
☐ Log Normal

● Viscosity
○ Stress

☐ High γ̇ Corr.
☐ Shear Thicken

λ Size ratio 1

k Fraction Small 0.35

φ_{max} Bimodal 0.63

E in kT 10

r radius nm 250

φ_{max} 0.5

φ values 0.28, 0.30,0.32,0.35,0.40,0.45

☐ Custom φ's

https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php

As you play with the sliders (the N and Fractal Dimension sliders describe particle association, there is also the Aspect Ratio slider) you can see two things:

- It is scarily easy to shoot up to super-high viscosities *at low shear*.
- But if your pumps are good enough to create high shear in the atomizer nozzle, then maybe you will be OK.

If you are going to rely on the very strong shear thinning of particle formulations, make sure you have tested many likely variations in production batch formulations. It is very easy to go from a 1000x low shear increase in viscosity to 10000x with a small change in dispersant or aspect ratio.

If you are dealing with polymer solutions then you need to be familiar with Flow_Polymer Viscosity as shown in the app:

R_g @0.5+40K nm 10
M_{mono} 100
Polymer MWt 28053
Polymer M_c 9759
Flory-Huggin χ 0.2
C_{max} g/dl 17.6
R_g @χ+MWt nm 11.4
Ne 32.5
C* g/dl 3.2
Ce g/dl 4.6
|η| dl/g 1.58
N.bonds 400
M_{bond} 100
Bond length nm 1.4
MWt 40000
R_g nm 11.4

https://www.stevenabbott.co.uk/practical-solubility/polymer-viscosity.php

The core idea is not so much that higher polymer concentrations lead to rapid increase in viscosity (we all know that), nor that high MW polymers are worse than low MW (again we know that). The key is MW compared to Mc, the "critical entanglement MW". You can have two similar polymers, with similar solubilities and similar MWs, but one of them will be much more viscous. This is because the number of tangles depends on MW/Mc. If the two polymers have Mc values differing by a factor of 2, then the one with the lower Mc is twice as tangled so will be more viscous.

In many applications (Adhesion_Entanglement) you *want* lots of entanglement, so you have a real fight when increasing concentration and MW. For spray drying you may not need so much entanglement so can choose either a lower MW or a polymer with a large Mc, so a high MW polymer (desirable for other reasons) becomes acceptable.

However, if you want a low diffusion coefficient for your encapsulated molecule in the polymer, all other things being equal, an entangled system leads to lower diffusion. Everything in formulation is a trade-off.

Another way to decrease viscosity of a polymer solution is to decrease its maximum solubility. Compared to a “happy” polymer with chains reaching out and tangling, the “unhappy” polymer is more coiled up on itself so isn’t so tangled. The obvious flaw in this approach is that you need a high solubility to get enough polymer into solution.

Fast drying, all the way to dry

Everyone who dries coatings wants maximum drying speed with minimum formation of a crust that inhibits further drying. See Evaporation_Basics and Evaporation_Diffusion limited for a more detailed discussion.

The key idea is that in “constant mode”, which is evaporation-limited drying, you are relying on a combination of heat and mass transfer, readily arranged in a modern drying system. As soon as you become “diffusion limited”, you can’t dry faster than the solvent gets through the crust, which becomes slower as the crust thickens, then there is nothing you can do other than increase temperature because diffusion goes faster at higher temperatures. High temperatures of a coating or a particle are generally not desirable.

So how do we keep our crust “open” enough to keep the drying evaporation limited? There are a few standard tricks:

- Add a “high boiler” (e.g. propylene glycol) in which your formulation remains at least partially soluble. This keeps the crust “open” for much longer.
- Use the formulation tricks for highest concentration with minimum viscosity to keep your formulation “happy” for as long as possible. You don’t care that the shell is now very high viscosity, you just want to make sure that there is still enough water for diffusion to be rapid and not rate limiting.
- Add some junk (e.g. rough particles) that stops the crust getting too compact. There’s an obvious downside if this makes the capsules too leaky, but with some thought you can get a reasonable compromise. If those particles are sticking out they will also reduce capsule-capsule and capsule-wall adhesion while the drying is incomplete.

This problem of the crust becoming too good a barrier is especially severe for those trying to use nanoclays as tortuosity barriers. There are no obviously good workarounds for this.

Keeping cool

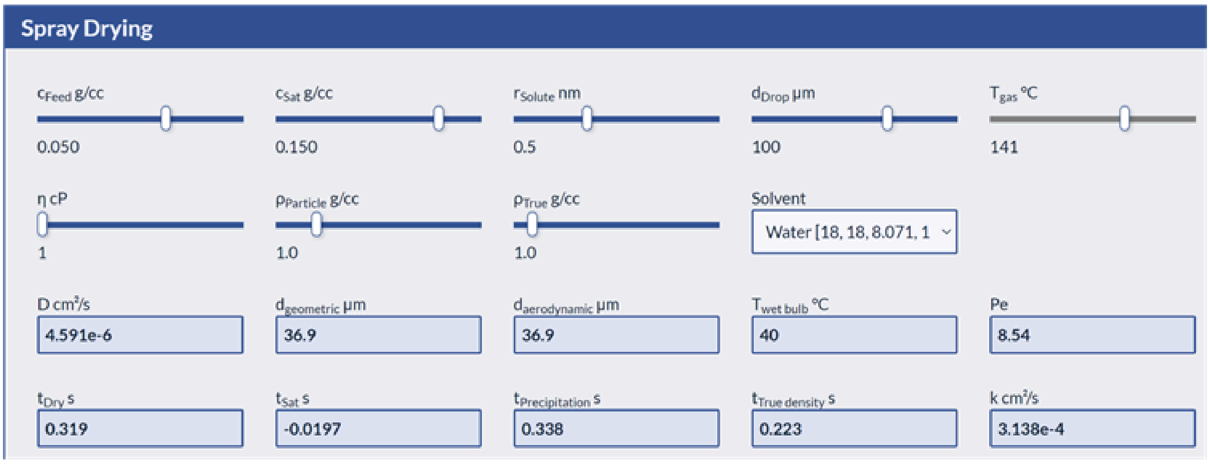
How do we spray dry heat-sensitive enzymes or microbes in 140°C (413°K) air, T_{gas}, without destroying them? A handy approximate formula (found in the app below) gives us the answer for water with a BP, T_b = 373°K. It tells us that the temperature of the microcapsule, T_{mc} is given by:

$$T_{mc} = 137\left(\frac{T_b}{373}\right)^{0.68}\log_{10}\left(T_{air}\right) - 45$$

For T_b=373 then we have 137 $\log_{10}$(413)-45 = 313°K = 40°C. Yes, our rapidly evaporating water is only at 40°C, so our enzyme or microbe inside the capsule is OK. If, however, we create a barrier crust then the capsule rapidly approaches 140°C and we have failed. Avoiding diffusion-limited drying really is important for spray drying.

A spray drying app

For those keen to explore some of the complexities of a pharma spray-drying system, this app gives more than enough to think about:



c_{Feed} g/cc 0.050
c_{Sat} g/cc 0.150
r_{Solute} nm 0.5
d_{Drop} μm 100
T_{gas} °C 141
η cP 1
$\rho_{Particle}$ g/cc 1.0
ρ_{True} g/cc 1.0
Solvent Water [18, 18, 8.071, 1
D cm^2/s 4.591e-6
$d_{geometric}$ μm 36.9
$d_{aerodynamic}$ μm 36.9
$T_{wet\ bulb}$ °C 40
Pe 8.54
t_{Dry} s 0.319
t_{Sat} s -0.0197
$t_{Precipitation}$ s 0.338
$t_{True\ density}$ s 0.223
k cm^2/s 3.138e-4

https://www.stevenabbott.co.uk/practical-coatings/Spray-Drying.php

The meaning of the multiple inputs and outputs are described on the app page.

Why only this one app? Because the spray drying community has not provided any other usable models.

Evaporation Temperature and Antoine Coefficients

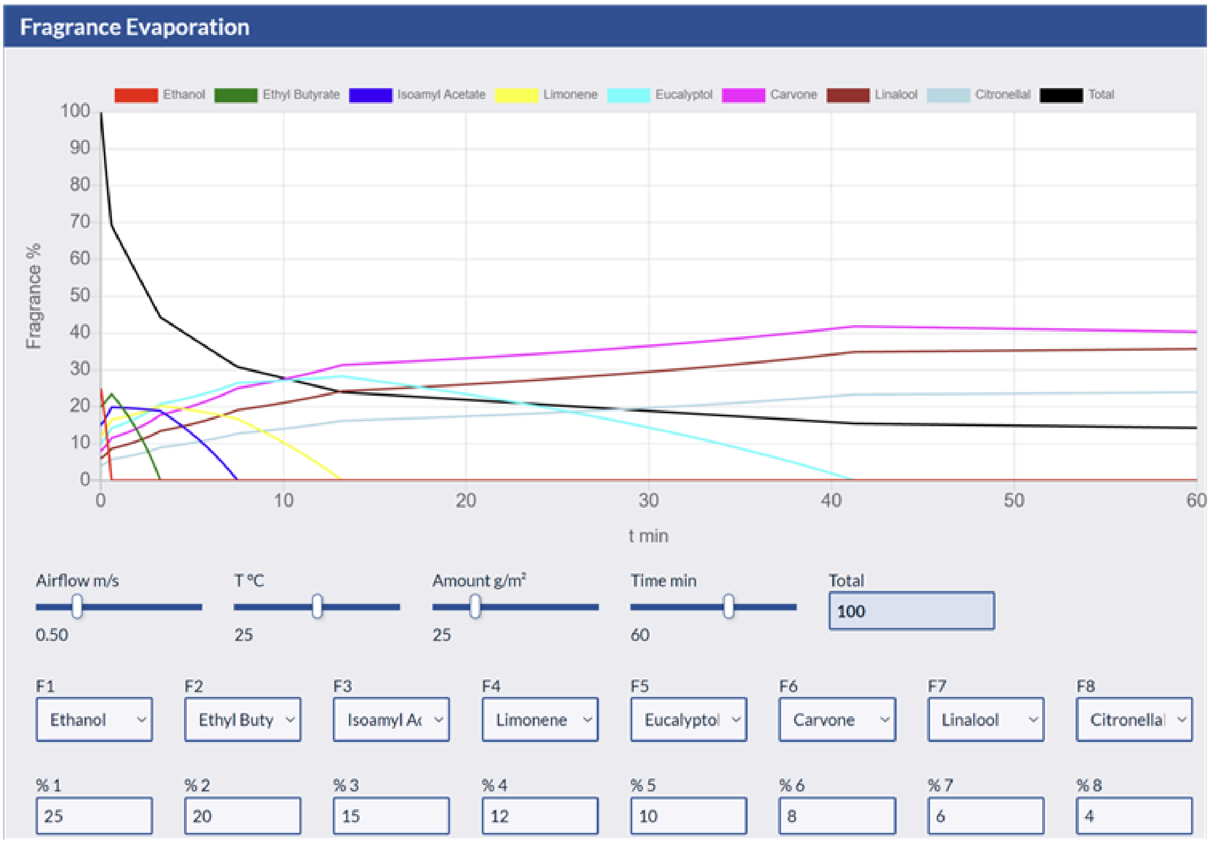
Links

[SkinCare](), [Fragrances](), [Solvent-Based Paint]()

> The easiest way to characterise the vapour pressures above a liquid at a given temperature is to use the Antoine Coefficients. These are merely fitting parameters to experimental curves and large datasets of values exist.

A fragrant example

Suppose you had a mix of 8 fragrance molecules and wanted to know how the high, medium and low notes (high, medium and low volatility components) changed over time. The app shows an idealised version:



Airflow m/s 0.50

T °C 25

Amount g/m² 25

Time min 60

Total 100

F1	F2	F3	F4	F5	F6	F7	F8
Ethanol	Ethyl Buty	Isoamyl Ac	Limonene	Eucalyptol	Carvone	Linalool	Citronellal

% 1	% 2	% 3	% 4	% 5	% 6	% 7	% 8
25	20	15	12	10	8	6	4

[https://www.stevenabbott.co.uk/practical-coatings/Fragrance-Evaporation.php]()

The ethanol, in red, has gone almost instantly. The limonene has all gone by ~14 min, the eucalyptol by 42min and the rest linger longer.

The app uses the Antoine Coefficients for each of the 8 molecules (chosen from a list of 40+) and calculates the individual vapour pressures, VP, as:

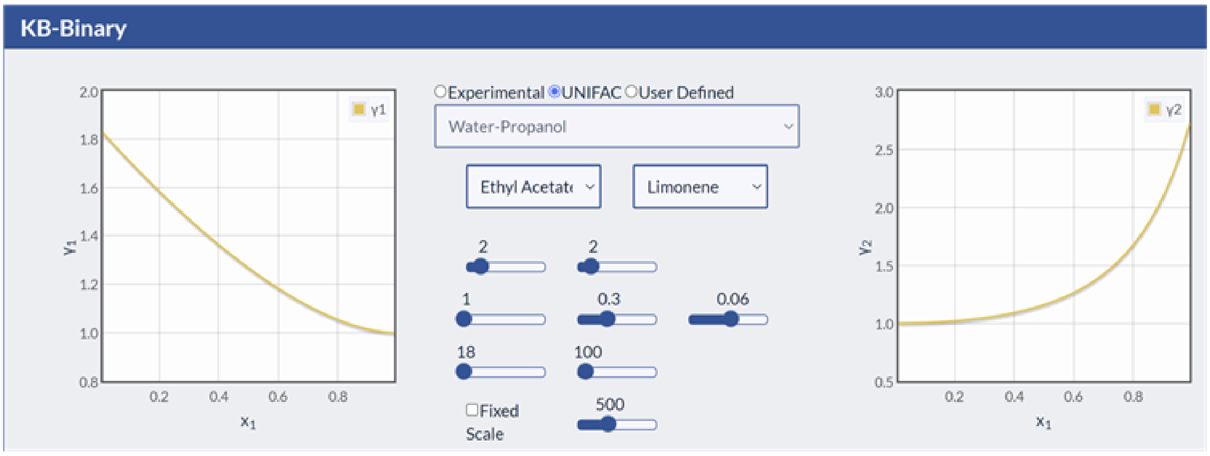
$$\log_{10}(VP) = AA - \frac{AB}{AC + T}$$

The 3 coefficients, AA, AB and AC are combined with the temperature T. The largest collections use values based on mm/Hg and °C temperatures. If you use values from different tables you have to be careful of the units.

You often find that different tables contain different constants for the same molecules. If you go to the trouble of plotting them, you find that the values are indistinguishable. As mentioned above, Antoine Coefficients are merely fitting constants and within inevitable experimental error, different authors decide on different constants to achieve the same end result.

It's not ideal

Real liquid mixtures are non-ideal so the vapour pressures will increase or decrease according to the activity coefficients at any given % mixture. The activity coefficients of binary mixes can be calculated with reasonable accuracy using UNIFAC:



https://www.stevenabbott.co.uk/practical-solubility/kb-binary.php

For example, at 50:50 mole fraction of ethyl acetate and limonene, the ethyl acetate has an activity coefficient, γ, of 1.27 and limonene 1.16, so ethyl acetate will be more present in the gas phase by a factor of 1.27/1.16 ~ 10%.

Beyond binary, you need more sophisticated software such as COSMO-RS. No attempt to introduce non-ideality was made in the previous app.

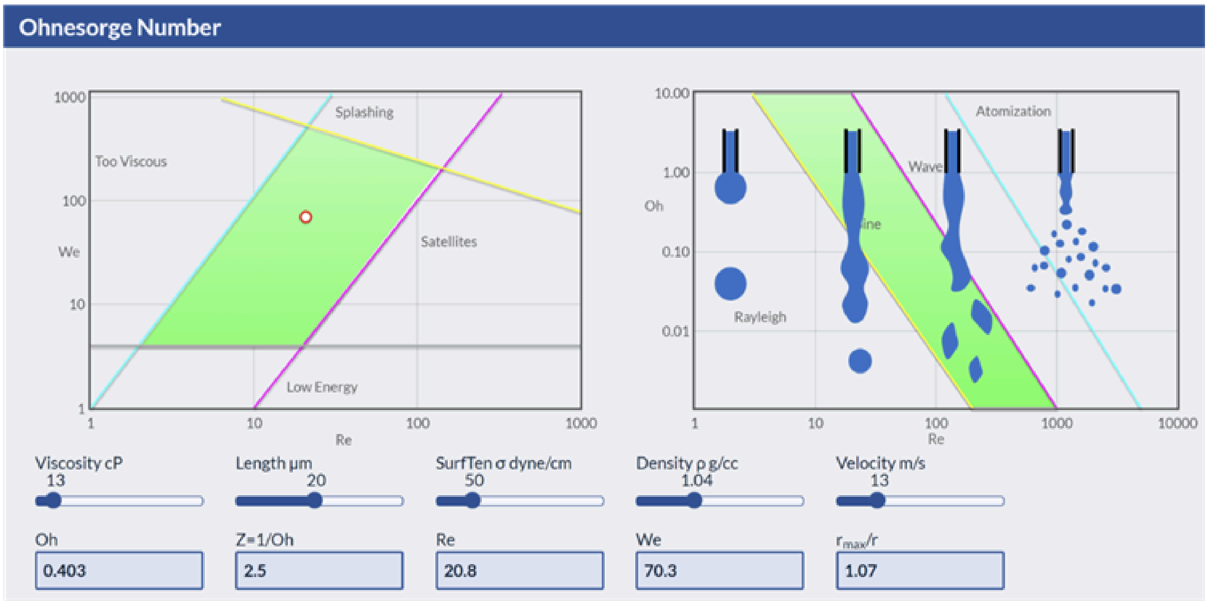
Flow Atomization

Links

This is a very short chapter

No useful theories

The first theory of multi-droplet formation from a jet of fluid, the Rayleigh instability is ~150 years old. Work from Ohnesorge and others in the 1930s sketched out a landscape of stability and instability in flows and we have two ways to look at dependencies on dimensionless numbers such as Reynolds and Weber:



Viscosity cP 13

Length µm 20

SurfTen σ dyne/cm 50

Density ρ g/cc 1.04

Velocity m/s 13

Oh 0.403

Z=1/Oh 2.5

Re 20.8

We 70.3

r_{max}/r 1.07

https://www.stevenabbott.co.uk/practical-coatings/ohnesorge.php

In the right hand graph of Ohnesorge versus Reynolds we see that as we go from low to high Re we go from dripping (Rayleigh) to orderly breakup (Sine wave) to disorderly breakup (Waves) and, finally, to atomization.

Because for flow velocity V, characteristic length (nozzle diameter) and viscosity η, $Re = \frac{V\rho l}{\eta}$, we reach the obvious conclusion that higher velocities and lower viscosities help atomization. Viscoelasticity risks producing strings rather than drops, but again that's rather obvious.

After that, we are without useful guides. Although formulae exist for some core drop sizes, they offer no insights into drop size distributions. Various fancy computational techniques exist for modelling atomization but these are of no practical help to the formulator.

Your choice of atomizer is constrained not by theory but by custom and practice and cost. For a simple hand spray pump a simple jet nozzle is fine and for spray drying a rotating co-air-flow nozzle is often required. In general we want to encourage the wave-to-atomization transition by imposing extra perturbations such as swirling motion or ultrasonics. But these extra effects are even less amenable to meaningful calculations so there is little the formulator can do to influence the outcomes other than reduce viscosity and viscoelasticity.

Even something as seemingly simple as a trigger gun spray for pharmaceutical or home use is complex physics and engineering (and intellectual property). When you start to squeeze, nothing should come out (otherwise it would be a dribble), then the full spray should emerge … till you stop pressing and the spray must cut off quickly rather than fail as a dribble.

For “normal” sprays, with “normal” fluids, there are a few rules of thumb. Let’s have a reference drop of size D_r in pure water with surface tension, σ, 72 mN/m, density ρ of 1 and a viscosity, η of 1 cP, sprayed with a pressure of P_r. So the new diameter D_n with surface tension σ_n, density ρ_n and viscosity η_n is given by:

$$D_n = D_r\left(\frac{P_n}{P_r}\right)^{-0.3}\left(\frac{\rho_n}{1}\right)^{-0.3}\left(\frac{\eta_n}{1}\right)^{0.2}\left(\frac{\sigma_n}{72}\right)^{0.5}$$

I apologize for this short chapter, but 150 years of research has not resulted in anything else that we can use.

> For many years I’ve known an expert on atomization. As an expert, he can draw on his experience and use his array of complex computational tools. Every few years I would ask if there are any useful new theories that could be used by formulators without his specialist knowledge. I asked again while writing this chapter and the answer is still “no”.

Trigger mechanisms more sophisticated build up pressure before opening then nice spray then quick shut-off

Flow Basic viscosity

Links

Dishwashing Liquids, Hair Conditioners, Hair Shampoos, Sun Screens, Soaps and Washing, Laundry Liquids, SkinCare, Surface Cleaning, Lipstick, Mascara, Pharma Formulation, Microencapsulation

> The basics of viscosity are easy to grasp. We just need to be sure we're clear about them before starting on more complex aspects.

Basics of viscosity

We have some fluid fixed at the lower surface and on top the fluid is moving with velocity V thanks to a stress τ, which is force/area, N/m² or Pa. The thickness of the fluid is Y. We can imagine the fluid as individual small layers (technically called streamlines). The stress, τ, needed to keep the fluid moving increases with V (obviously) and decreases with Y – because the bigger the gap between the stationary and fast fluid, the easier it is to move it. We can therefore say that:

$$\tau = \frac{\eta V}{Y}$$

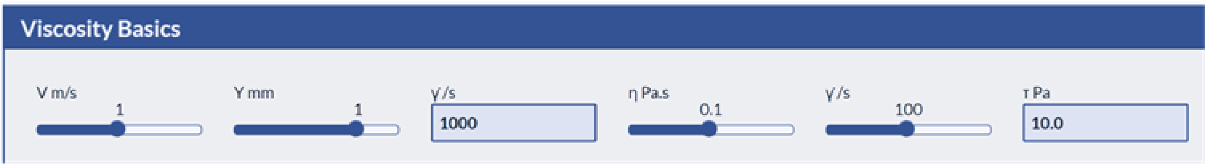
The constant of proportionality, η, is the viscosity. The V/Y term is the shear rate $\dot{\gamma}$ so we write that

$$\tau = \eta\dot{\gamma}$$

This equation was derived by Newton in the 17th century so this is "Newtonian viscosity" where η is constant with shear rate. For many systems, η is shear rate dependent, as discussed here Flow_Shear dependent.

The units of shear rate are $\frac{m/s}{m}$ which are simply 1/s. So viscosity is $\frac{Pa}{1/s} = Pa.s$. For a low viscosity such as water ~0.001 Pa.s we either say 1 mPa.s or 1 cP where cP are centipoise where 1 Poise is 1/100th of a Pa.s. It is *very* easy to get confused between Poise and Pa.s so try to avoid Poise. But cP are fine given that they are the same as mPa.s.

To get a feel for the different components you can use the app:



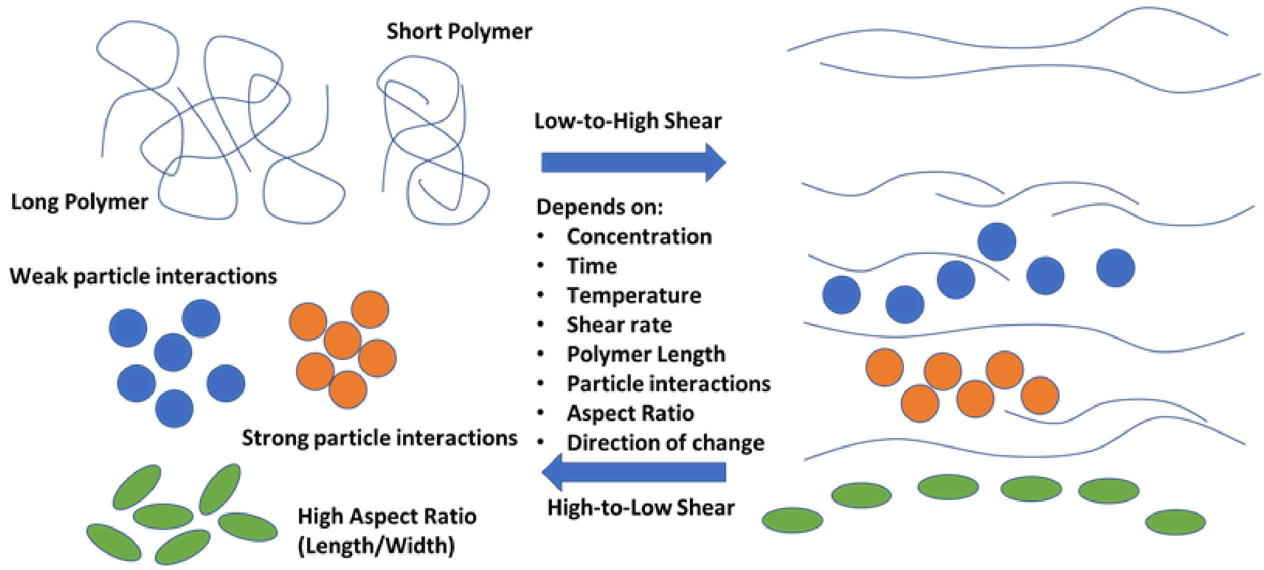
Viscosity Basics

V m/s: 1
Y mm: 1
$\dot{\gamma}$/s: 1000
η Pa.s: 0.1
$\dot{\gamma}$/s: 100
τ Pa: 10.0

https://www.stevenabbott.co.uk/practical-rheology/Viscosity-Basics.php

If viscosity is so easy …

… why is rheology so complicated? Because our systems are complicated.



We have short and long polymers, tangled and un-tangled, we have particles with weak and strong interactions with themselves and the polymers, we have different aspect ratios … and as we apply shear all sorts of different things can happen.

So our basic intuitions of viscosity take us only so far. After that we need Flow_Shear dependent and we need Flow_Oscillatory rheology.

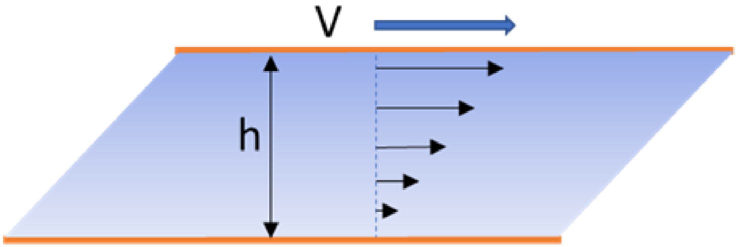
Flow Couette Flow

Links

Pharma Formulation

If you have a stationary fluid next to a moving wall then the fluid gets dragged along by Couette flow. The opposite effect, forcing fluid through a stationary gap, is described by Flow_Poiseuille Flow.

Couette Flow

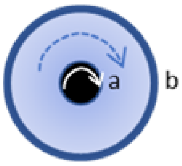


We have a stationary wall and, at a distance h, a wall moving with velocity V. It turns out that the velocity profile is linear, so at a distance x from the base, $v = \frac{Vx}{h}$

The shear stress, τ, is constant, simply $\frac{V}{h}$

Why do we need to know about Couette flow? Because there are times when such flows, for good or bad reasons, happen in your system. And because the Couette viscometer relies on the effect.

Couette Viscometer



A cylinder of radius *a* and length *l* spins at speed ω inside a cylinder of radius b containing a fluid of depth greater than *l*. The viscous drag requires a torque τ to overcome it. From this torque we can calculate the viscosity η via:

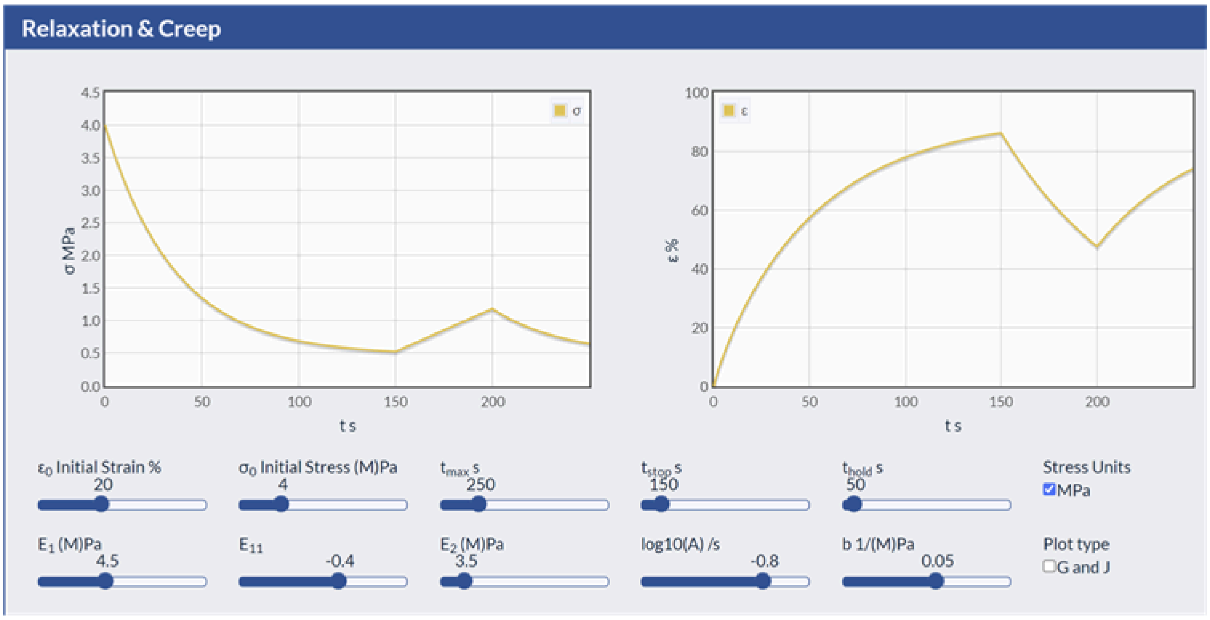
$$\tau = \frac{4\pi\eta l\omega a^2 b^2}{b^2 - a^2}$$

Flow Creep and Relaxation

Links

> Any material placed under load can show elastic deformation, which reverses when the load is removed. What we usually don't want is for the material to creep, to extend steadily under load so that it's permanently deformed. Here we see the theory of creep and its partner, relaxation.

Relaxation and creep



ε_0 Initial Strain % 20

σ_0 Initial Stress (M)Pa 4

t_{max} s 250

t_{stop} s 150

t_{hold} s 50

Stress Units ☑MPa

E_1 (M)Pa 4.5

E_{11} -0.4

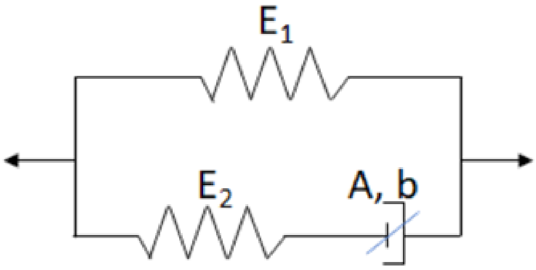
E_2 (M)Pa 3.5

log10(A) /s -0.8

b 1/(M)Pa 0.05

Plot type ☐G and J

https://www.stevenabbott.co.uk/practical-rheology/Relaxation-Creep.php



We have two springs, E_1 and E_2 which, under load, simply extend and would return to their original position if the load is removed. The "dashpot" next to E_2 is a device that slowly extends, irreversibly, under load. A dashpot is a purely viscous behaviour so doesn't class as creep.

The right-hand graph shows creep – the strain (% extension), ε, increases slowly over time under a constant stress σ_0. After 150s the stress is removed and the system partly returns, this is relaxation, then after a holding time of 50s, the stress is reapplied to give more creep.

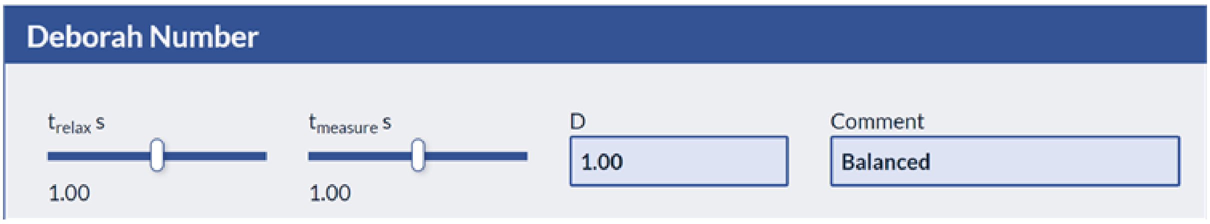
The left-hand graph shows relaxation. An initial strain, ε_0 is applied which creates an initial stress which, over time, relaxes. As with the creep, after 150s the strain is removed and the system starts to partially recover.

It's well-known that the creep can be described by a Kelvin-Voigt model of E_1 and the dashpot in parallel. While the relaxation can be described by a Maxwell model with E_2 and the dashpot in series. Unhelpfully, it is usual to describe either Kelvin-Voigt or Maxwell even though real materials show both types of behaviour. The app uses a combined model with 5 parameters. Although we have a simple E_2, instead of a simple E_1 we have a “strain hardening” spring, in which the modulus gets higher with extension, involving an extra parameter E_{11}. And instead of a simple dashpot with a single parameter, we have A which controls the frequency response (the typical dashpot parameter) plus a non-linearity term b.

The relevant equations are all shown in the app.

Complex behaviour

All materials creep and relax over some appropriate timescale. The Deborah Number D:



Deborah Number

t_{relax} s: 1.00

$t_{measure}$ s: 1.00

D: 1.00

Comment: Balanced

https://www.stevenabbott.co.uk/practical-adhesion/Deborah-Number.php

$$D = \frac{t_{relax}}{t_{measure}}$$

tells you whether your timescale is relevant. As the prophetess Deborah said: “The mountains flowed before the Lord”, meaning that over a sufficiently large timescale παντα ρες (panta res), everything flows.

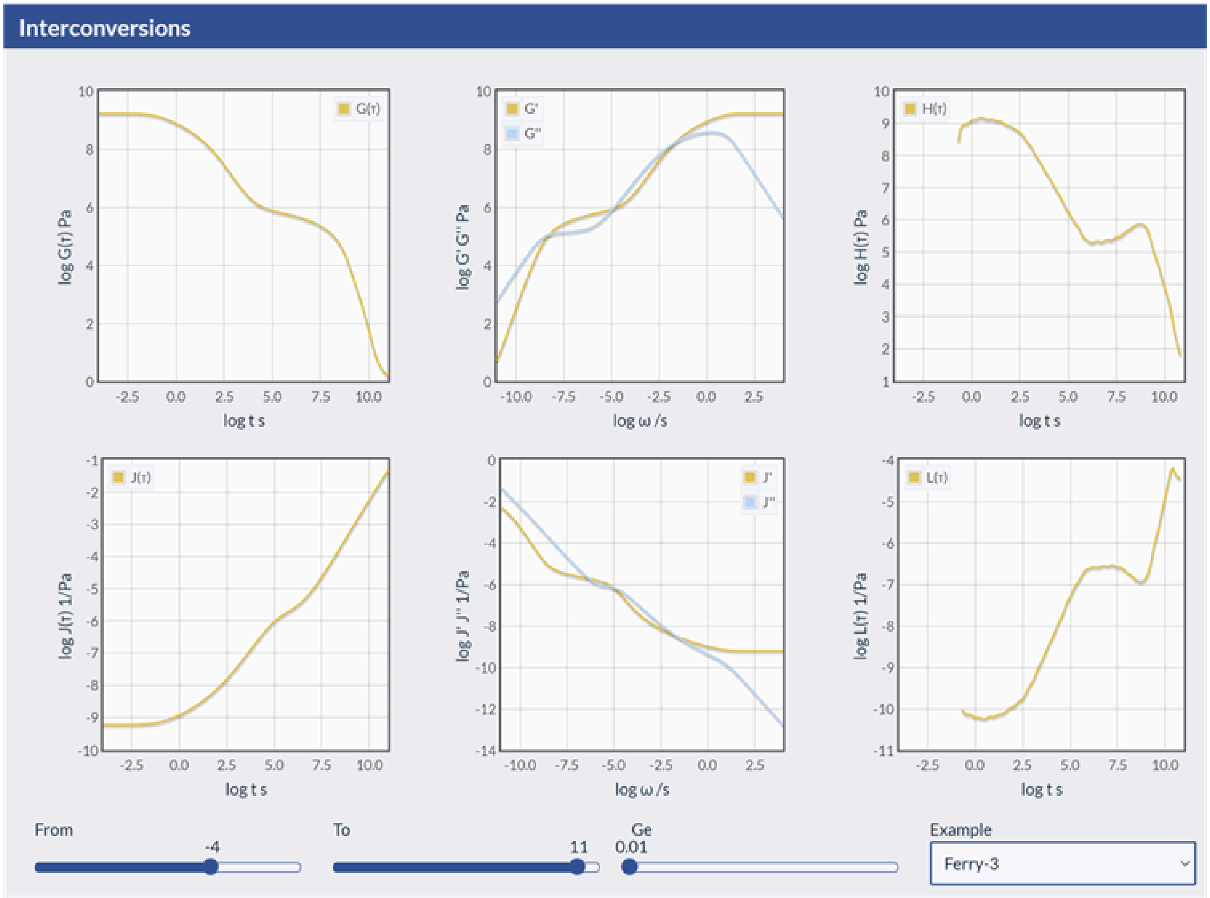
So if you measure your creep over 1s, 1min, 1hr, 1day, 1 month or 1 year you might see very different behaviour. As making measurements over 1 year is a bit tedious, you might want to do them at a higher temperature and use WLF Flow_TTS-WLF to recalculate at lower temperature, reflecting the fact that a mountain flows in human timescales if it gets hot enough to be a volcano.

Interconversions

But what happens if you can’t conveniently do a creep test on your rheometer or DMA? Maybe you can measure relaxation or G'/G". One of the wonders of rheology is that you can interconvert between measurements.

> I had no idea about this. I was watching one of the excellent TA Rheometer webinars when the speaker said “And of course you can use the software to interconvert between all the modes.” I had to stop the video and replay it to make sure I’d heard correctly. Science is wonderful.

If you have relaxation data you can convert it to creep etc. An app lets you see that in practice. The app lets you explore a wide variety of datasets. Your rheometer software should be able to do your conversions for you.



From -4

To 11

Ge 0.01

Example Ferry-3

https://www.stevenabbott.co.uk/practical-rheology/Interconversions.php

It’s common to assume that we should just read the latest reviews and papers to get the best science. But I simply could not understand how to do these interconversions. A polite request for help to one professor got the short reply: “My job is to create good maths equations, not to explain them to people like you”. Finally I went back to the 1980 3rd edition of the 1961 book by Ferry. There, the explanation was clear and simple.

Flow G' and G''

Links

SkinCare, Lipstick, Mascara, Water-based Adhesive, Adhesive, PSA, Solvent-Based Paint

For some reason, the most useful measures in the whole of rheology, G' and G'' create fear and confusion, especially when you add tanδ. In fact, they are easy, and you get a lot of information from measurements that your rheometer can routinely make. The app makes it simple to see what's going on.

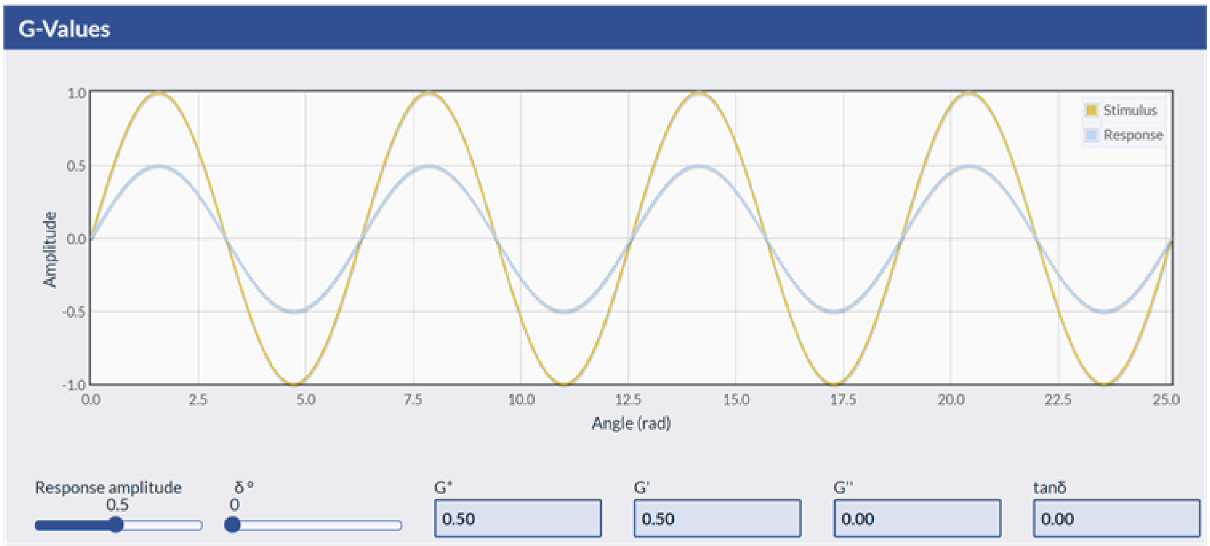
Elastic and Plastic

We know that under many circumstances (e.g. low temperatures and high speeds) materials can act elastically: as the stress (force per area) increases the strain (fractional extension) increases, and as the stress returns to zero, so does the strain.

Under other circumstances (e.g. high temperatures and slow speeds) materials can act as purely viscous materials – as stress increases the strain increases but as the stress returns to zero the strain remains the same. This is plastic deformation.

Equally obviously, materials under intermediate conditions will show some elastic and some plastic deformation.

G' is simply the measure of the purely elastic (shear) modulus and G'' the measure of the plastic/viscous modulus. And $tan\delta = \frac{G''}{G'}$ just tells us whether things are more (tanδ > 1) or less (tanδ < 1) plastic. Let's see that in an app:



Response amplitude 0.5

δ ° 0

G* 0.50

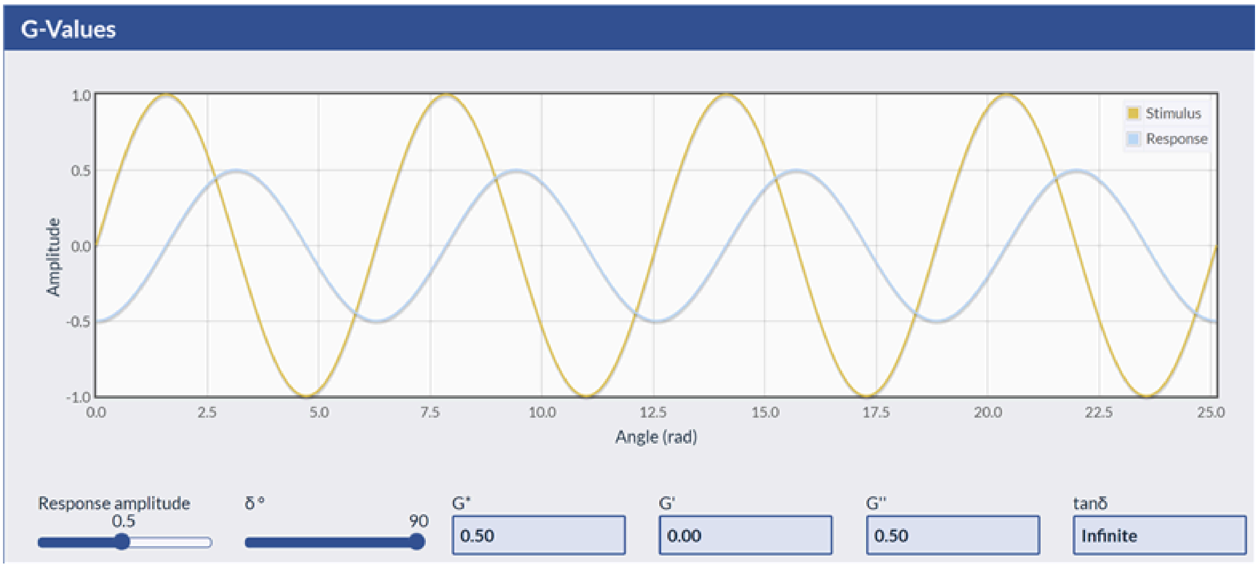
G' 0.50

G'' 0.00

tanδ 0.00

https://www.stevenabbott.co.uk/practical-rheology/G-Values.php

We have an oscillating stimulus, in gold, which produces a response, in blue. The magnitude of the response is not what concerns us, it's just some value of 0.5. What's important is that the response is exactly following the stimulus, the angular difference between them, δ = 0. This is a purely elastic behaviour so G' = 0.5, G" = 0 and tanδ = 0.



G-Values

Response amplitude 0.5

δ ° 90

G* 0.50

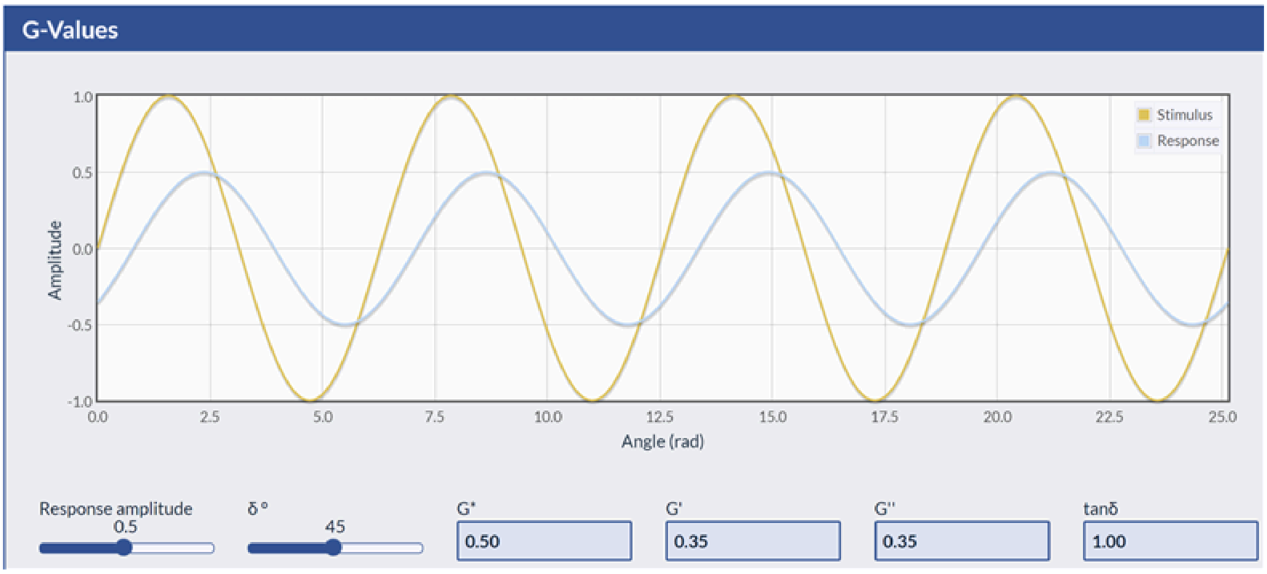
G' 0.00

G" 0.50

tanδ Infinite

Now we have a purely viscous sample where the response angle δ is 90° out of phase – meaning that the peak of the stimulus corresponds to a 0 value in the response, and a 0 of the stimulus corresponds to a maximum in the response. G' = 0, G" = 0.5 and tanδ is infinite.

Before continuing, why is a viscous response 90° out of phase? At the peak of the stimulus, the velocity of the oscillating motion is zero – the motion is changing direction. Viscous stress is viscosity times shear rate, which is zero. At the zero-crossing of the stimulus, the motion is fastest, so the viscous stress is largest.



G-Values

Response amplitude 0.5

δ ° 45

G* 0.50

G' 0.35

G" 0.35

tanδ 1.00

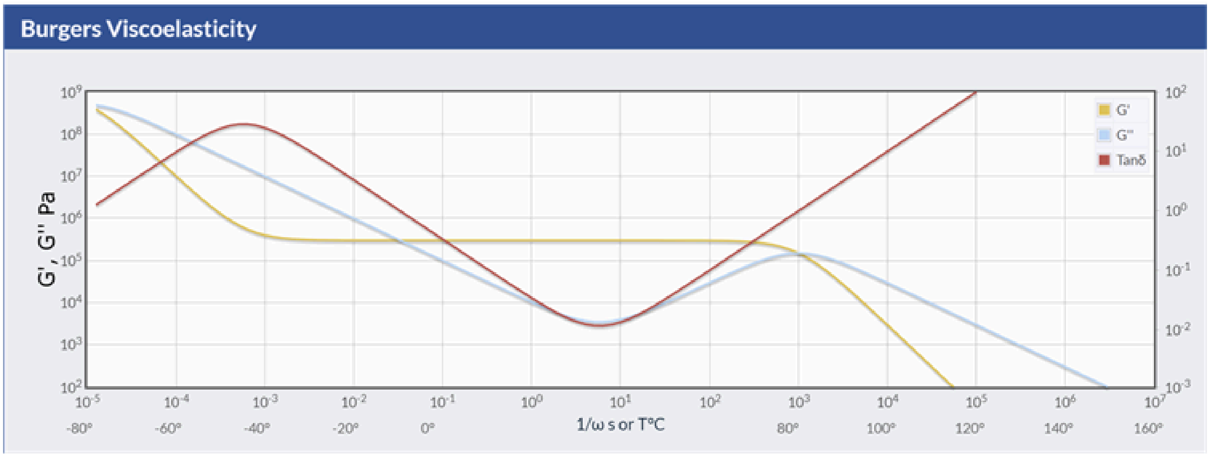
Now we have the intermediate state where there are equal amounts of G' and G". Their total, G* is still 0.5 but their individual values are 0.35. That's because $G^{*2}=G'^2+G''^2$. Now tanδ = 1 as there is an equal amount of elastic and plastic deformation.

That's it. You now know what G' and G" are and that tanδ isn't scary.

[Why is it called tanδ? The G^{*2} equation gives a clue. We're talking Pythagoras and right-angled triangles, with δ being the angle, G" the *opposite* side, G' the *adjacent* side and *tan=opposite/adjacent*.]

Although G' and G" are easy to understand, measuring them requires plenty of thought. The ideal setup is one where you have a method designed for your sorts of materials by your internal rheology expert. After training in use of that method (sample preparation, machine settings) you can then run your own samples.

Seeing what's going on



https://www.stevenabbott.co.uk/practical-rheology/Burgers.php

The specifics of this graph are not what's important. Instead think of this as the generalised behaviour of a polymer going from low to high T. At low T it's a very high modulus, both for G' and G". As it's warmed it reaches a rubber plateau where G' remains constant and G" steadily decreases, so tanδ also decreases. Their values cross at ~-10° in this plot and that is the conventional signal of Tg, the glass transition temperature. As we exceed Tg, the plastic modulus continues to decrease … until the polymer starts to approach its softening point. The plastic component starts to increase and crosses at the point when the elastic component finally fails as the polymer melts. In the end all that remains is the viscous component.

But you can look at the same graph in another way. You can keep T at 25°C and think of the measurements taking place at high frequencies or short timescales (on the left) and super-low frequencies, long timescales on the right. Because of Time-Temperature equivalence (see Flow_TTS-WLF), both views (or any combination) are correct. At a short timescale, a polymer cannot react to a stress, so it behaves as if it was a frozen lump of polymer measured at a normal timescale … and so forth.

This one graph teaches you a lot about what's going on inside the polymer, and behaviours such as brittleness or creep can be readily imagined. That's the power of this way of looking at materials.

It turns out that although plots of G' and G" are very common, there are two difficulties:

1. There are different ways of plotting them. I like low T to the left, high T to the right, so G' is highest on the left. Others prefer low frequency on the left, high on the right, so G' is highest on the right. Sometimes you have to find the legend to tell you which is G' and which is G", and sometimes tanδ is also plotted, usually with a Y2 axis on the right. You can therefore expect to be confused when someone shows you a rheology plot. It is entirely acceptable to say "Wait a

minute, I need to orient myself and work out what I'm looking at."

2. Although the basics aren't hard, working out what's happening, and why it's happening, isn't so easy. Again, it is acceptable to say something like "Wait a minute, are we saying that the plastic component is doing *this* at this temperature? Wouldn't that mean that Y is happening?" Your initial interpretation might be wrong, but you're starting a conversation that leads to deeper insights.

It's not just you who has these problems. We all do. So take your time, ask questions, and enjoy discovering new insights.

Sometimes you have two samples with very different behaviour and run the rheometer to see what G' and G" differences there are – and discover that the two samples are nearly identical. This sort of *negative* result is as useful as a positive one. If clear differences aren't visible in low-amplitude oscillations then you need to look at other parameters to explain the performance differences.

To get the data you use Flow_Oscillatory rheology.

E' and E"

So far, our sample has been sheared – maybe in an oscillatory rheometer or a DMA (Dynamic Mechanical Analyzer) in shear mode.

You can equally use a DMA in tensile mode, and determine the tensile elastic and plastic moduli, E' and E".

The values can be interconverted by taking into account the Poisson ratio, ν, which is by how much the material will shrink in one direction if stretched in the other. For a typical polymer ν = 0.3 and for a perfect rubber ν = 0.5:

$E = 2G(1 + \nu)$ and $G = \frac{E}{2(1 + \nu)}$

Who cares?

Many chapters of this book refer you to this chapter. G' and G" behaviour, the elastic and plastic behaviour, of our formulations are often key to unlocking puzzles of performance. The unnecessary fear of rheology and, perhaps, the historical difficulties of using older-style rheometers have caused formulators to avoid the topic. Although rheology can be complex, and although formulators don't have to be rheologists, understanding and using the insights of G' and G" is something formulators should embrace.

A project at a megacorp was going nowhere. Every time they fixed one problem, another one appeared. This was because there were trade-offs between G' & G" about which they had no idea because they didn't seek any rheological help. When, after much persuasion, they went to their rheologist, the results quickly showed why they'd kept failing and where the sweet spot was likely to be. After months of well-intention formulation tweaks without guidance from G' & G", within a week they had a working formulation.

Flow Ohnesorge

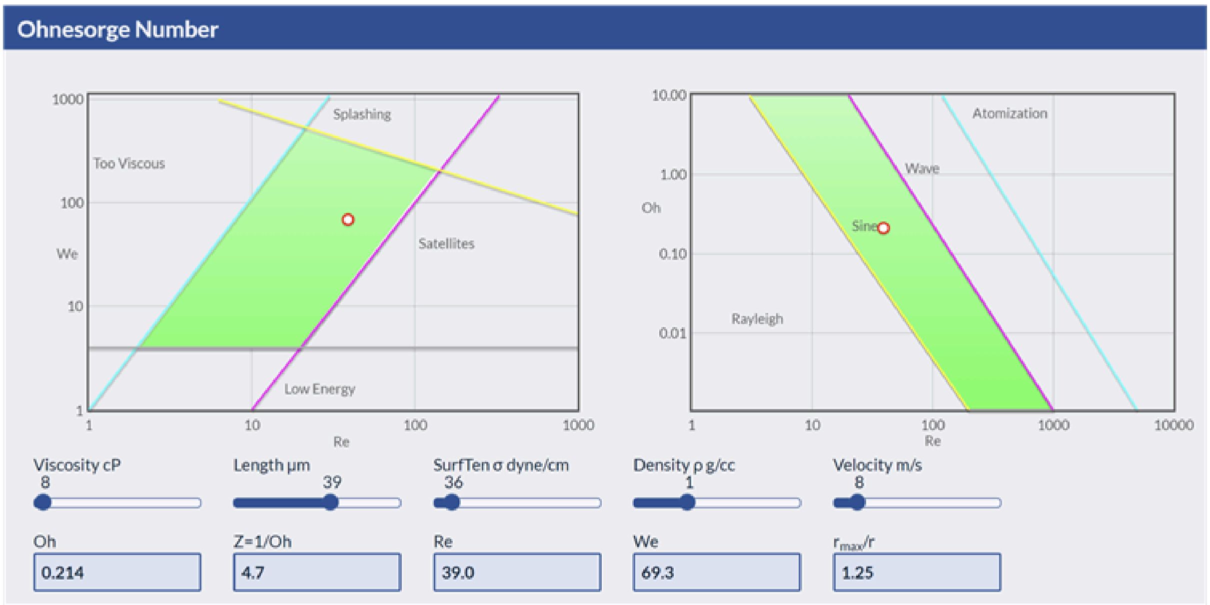
Links

[Inkjet Inks]

> Physics imposes severe restrictions on our ability to spray or inkjet. For example, there is only a narrow window for an inkjetable formulation, defined by a few dimensionless numbers, Reynolds, Weber and Ohnesorge. Inside, you get great drops, outside you don't. Or if, for example, we want atomisation then we need to be in a different part of dimensionless space.

The small inkjet window

Given a typical inkjet head, you can't print a good dot if the viscosity of the ink is greater than ~ 20cP or less than ~ 5 cP. Your surface tension can't be too high or too low. This knowledge is encapsulated via 3 dimensionless numbers, shown in the app:



Viscosity cP 8
Length µm 39
SurfTen σ dyne/cm 36
Density ρ g/cc 1
Velocity m/s 8
Oh 0.214
Z=1/Oh 4.7
Re 39.0
We 69.3
r_{max}/r 1.25

https://www.stevenabbott.co.uk/practical-coatings/ohnesorge.php

When the drop of viscosity η, surface tension σ, velocity V, density ρ and characteristic length l numbers are:

1. Reynolds, the ratio of inertial to viscous forces: $Re = \frac{V\rho l}{\eta}$
2. Weber, the ratio of inertial to surface tension forces: $We = \frac{V^2\rho l}{\sigma}$
3. Ohnesorge, the tendency for the drop to stay or fly apart: $Oh = \frac{\eta}{\sqrt{l\rho\sigma}}$

The left-hand graph shows the relevant regions in terms of We and Re. Typical trade-offs are when the viscosity is too high to make good drops or, in the other direction, satellite drops are created too easily. The right-hand graph shows them in terms of Oh and Re. The line between sine-wave break up (what we want for inkjet) and wave-like breakup is the one defined by Ohnesorge himself.

Atomization

The forces involved in inkjet are modest – as you would expect when you need to make millions of identical little drops. To get Flow_Atomization requires significantly larger drops and higher velocities, hence atomizers are much more about big pumps and high powers.

Sadly, if you go to the Atomization chapter you will discover that beyond this simple generalisation, there is little appable science to help us to spray better.

Flow Oscillatory rheology

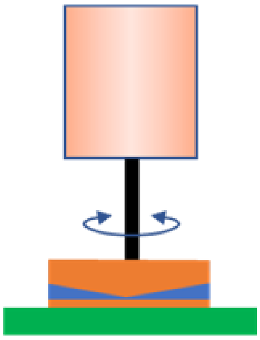
Links

Deodorant Sticks, SkinCare, Toothpastes, Lipstick, Mascara, Water-based Adhesive, Adhesive, PSA, Solvent-Based Paint, Pharma Formulation

Everyone is comfortable with classic, rotational, viscometry, Flow_Rotational rheology, including complex phenomena such as Flow_Shear dependent, yet oscillatory rheology, which is so powerful, is regarded as difficult and obscure – something for rheologists, not formulators.

Well, embrace oscillatory rheology. It's not difficult and it provides lots of formulation insights for relatively little work.

Oscillations

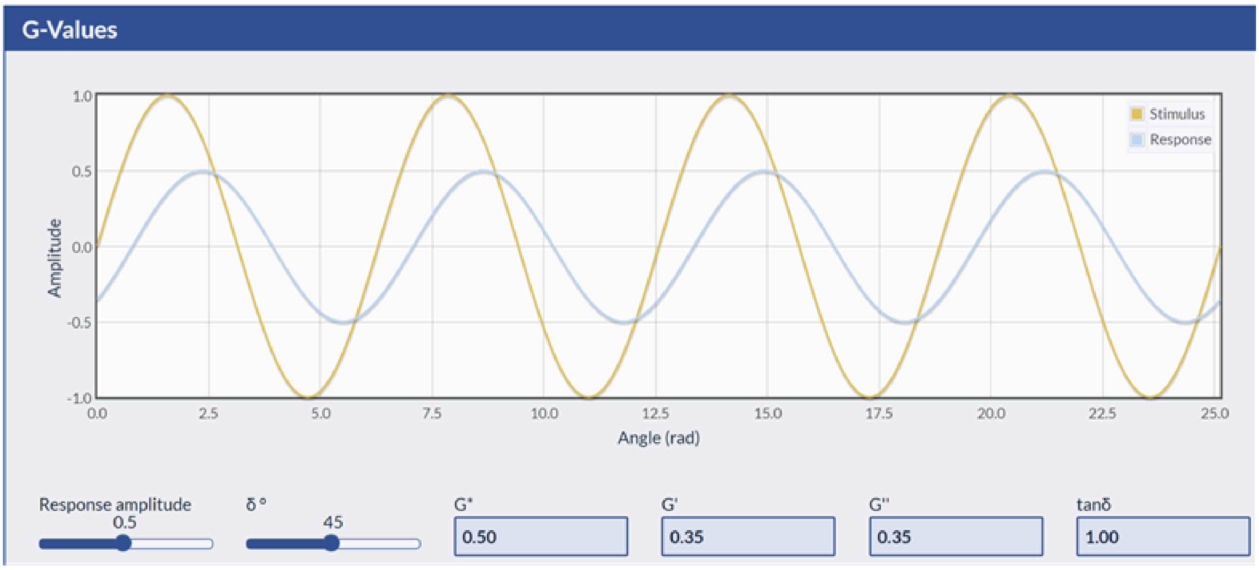


We have the liquid trapped between a lower plate and, in this example, a cone. An oscillatory stress is applied to the cone, and the strain, the amount of movement, is measured. Or an oscillatory strain is applied, and the stress needed to create that strain is measured. There are great debates about which type of machine is better, but the answer is that the best machine is the one you have, and with modern technology the differences between the two are not so important.

The oscillations give modest strains, maybe just 0.1%-0.5%. If you try for smaller strains, the signal becomes too weak, and larger strains can take you into the non-linear region where analysis gets difficult. Experts are happy working with LAOS (Large Amplitude Oscillatory Shear), but most of us can't cope with it.

What do we learn?

As described in Flow_G' and G" the stimulus (either stress or strain) provides a response (strain or stress) which we can interpret in terms of the elastic response, G' and the plastic (viscous) response G". In the image we have an equal fraction of each.



Response amplitude 0.5

δ ° 45

G* 0.50

G' 0.35

G" 0.35

tanδ 1.00

It's great that with just a few seconds of oscillations we can work out the complex behaviour of our formulation. The fraction of elastic and plastic response tells us a lot about, say, entanglements. But if we are going to the trouble of setting up these measurements, we should try to extract the maximum value.

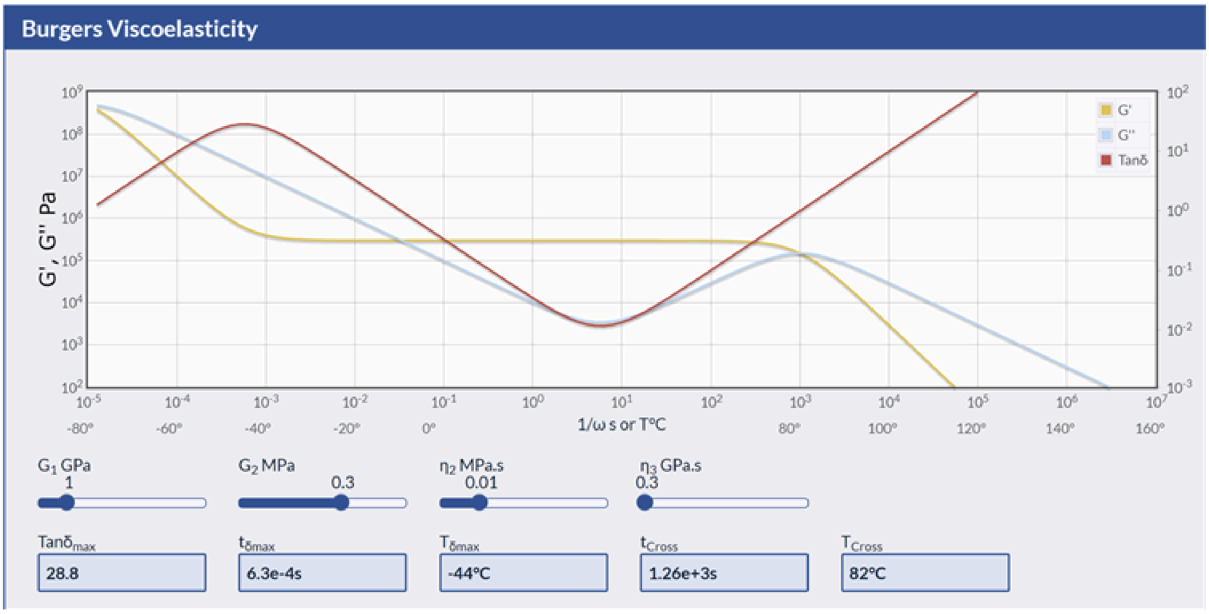
Temperature equals Time

A modern rheometer allows us to do these oscillations at a range of temperatures, thanks to a Peltier plate that might easily go from -20 to +100°C. As you change the temperature, the magnitudes and relative proportions of elastic and plastic flow will change. This gives you valuable insights into what your formulation is doing at a fundamental level. At low T the behaviour will be more elastic, as motion becomes "frozen" and at high T it will be more plastic. The crossing point where G' = G" is typical of the Tg of the formulation – the glass transition temperature above which large motions become possible.

The rheometer also lets you do these oscillations at a range of frequencies, typically from 0.01 up to 100 /s. Now you can see the *dynamic* behaviour of your formulation. Typically it will show plastic behaviour at low frequency, because the polymers and particles have time to move. At high frequency it will behave in a more elastic manner as the system doesn't have time to move internally.

The paragraphs on temperature and on timescale show similar effects. Indeed, TTS (Temperature Time Superposition) or TTE (Temperature Time Equivalence) (they mean the same thing) is a profound aspect of physics and can be captured across a wide range of materials and formulations via the WLF (Williams, Landell, Ferry) formula, Flow_TTS-WLF.

By using WLF to combine sweeps done (intelligently) at a range of temperatures and frequencies, you can work out the G' and G" behaviour over a wider range of temperatures or frequencies (it's your choice) than is possible to measure on any practical rheometer:



G_1 GPa 1

G_2 MPa 0.3

η_2 MPa.s 0.01

η_3 GPa.s 0.3

$Tan\delta_{max}$ 28.8

$t_{\delta max}$ 6.3e-4s

$T_{\delta max}$ -44°C

t_{Cross} 1.26e+3s

T_{Cross} 82°C

https://www.stevenabbott.co.uk/practical-rheology/Burgers.php

The app simulates the behaviour of a classic pressure sensitive adhesive over a range of 240°C and 10^{-5} to 10^7 Hz, even though the real measurements were made over a much smaller range.

Variations on a theme

Suppose you apply a paint to a surface with a brush. This is a high shear, high speed process for which oscillatory rheology is not much use.

But you now want to ensure that any lines in the coating will level out and that paint on a vertical surface won't sag, i.e. flow down as imperfect waves in the surface. You *can* use careful low shear rate rotational viscosity to pick up things like yield stress (Flow_Yield Stress) and thixotropy (Flow_Thixotropy), see https://www.stevenabbott.co.uk/practical-rheology/Levelling-&-Sag.php, but you can equally look at G' and G" as there is going to be a strong link between levelling/sag and these parameters. But paints take time to build up their structure after rapid processes, so a conventional G'/G" measurement might be looking at the wrong behaviour. Instead, you give a strong rotational stress to the paint in the rheometer then switch to oscillatory mode and watch how G' and G" change over time. A swift recovery to large G' values will indicate a low tendency to sag, but also a poor levelling behaviour. Slow recovery will encourage levelling but promote sag. You aim for a compromise.

You can model this recovery process as a way of looking at thixotropy:

G'_0 0.1

G'_∞ 1.2

Power law n 0.2

$Tan\delta_0$ 2

$Tan\delta_\infty$ 0.1

t_{max} s 300

☐Log(t)

☐Log(G or η)

https://www.stevenabbott.co.uk/practical-rheology/Thixotropic-Recovery.php

Rotational or Oscillatory?

Historically, formulators have used rotational viscometry. The machines are easier to use and ideas such as shear thinning and thixotropy can be explored without too much difficult. But this is the 21st century when equipment is more affordable and usable. So the clear answer to the question of Rotational or Oscillatory is “Both” or, maybe, “A rheometer because it can do both”. Each mode provide deep insights into your formulations. If you only have a viscometer, you are missing out on oscillatory insights. With a rheometer you can gather both types of data and use their relative strengths to get a full picture of what’s going on inside your formulations.

Oscillatory rheology is not too hard to understand, but there are many subtle details needed to get good data from a rheometer. The ideal setup is one where you have a method designed for your specific needs by your internal expert. They know how to answer questions like: Do you use plate-cone or plate-plate? Do you put sandpaper on a plate to stop sample slippage? What’s the best strain to choose? What’s the optimum sample thickness? How do I trim off stuff oozing out from the plates? After training in use of that method (sample preparation, machine settings) you can then run your own routine samples.

Flow Particle Viscosity

Links

Toothpastes, Microencapsulation

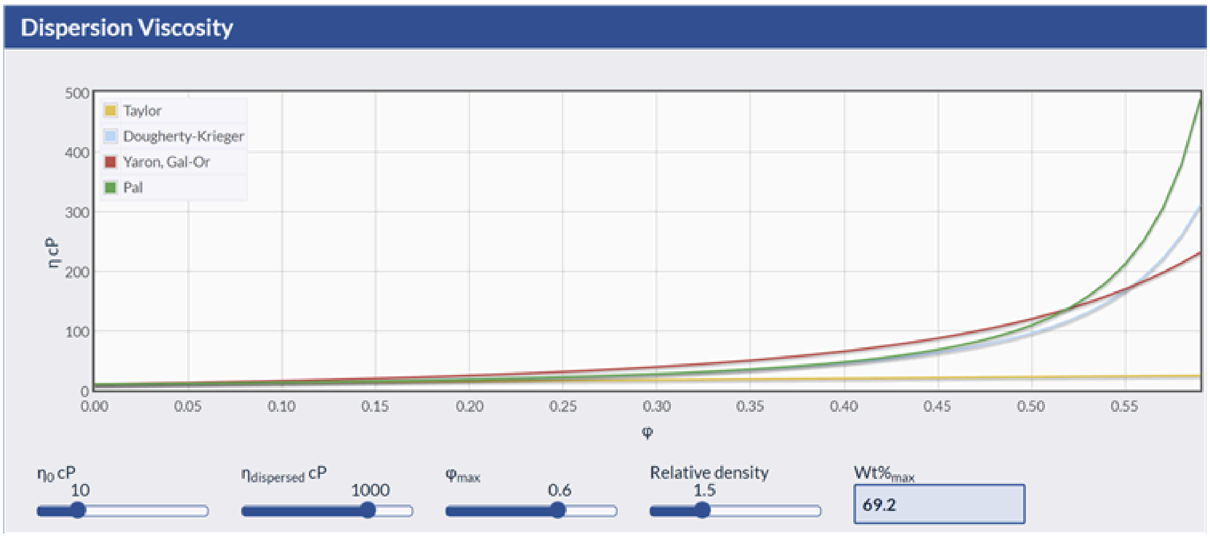
The particles/pigments in formulations have a big effect on viscosity. The *weight* fraction is not relevant to these effects, it is the *volume* fraction, φ.

When φ is low, say up to 40%, the effects on viscosity are not large – unless there are specific particle-particle or particle-polymer interactions. The effects are described with the Dispersion Viscosity app which applies to solid particles and to emulsion particles.

At higher φ values the viscosity is dependent both on shear rate and on subtle particle-particle interactions.

The topics are discussed in their respective Dispersion chapters, with links provided below.

Low-to-Medium φ



η_0 cP 10

$\eta_{dispersed}$ cP 1000

φ_{max} 0.6

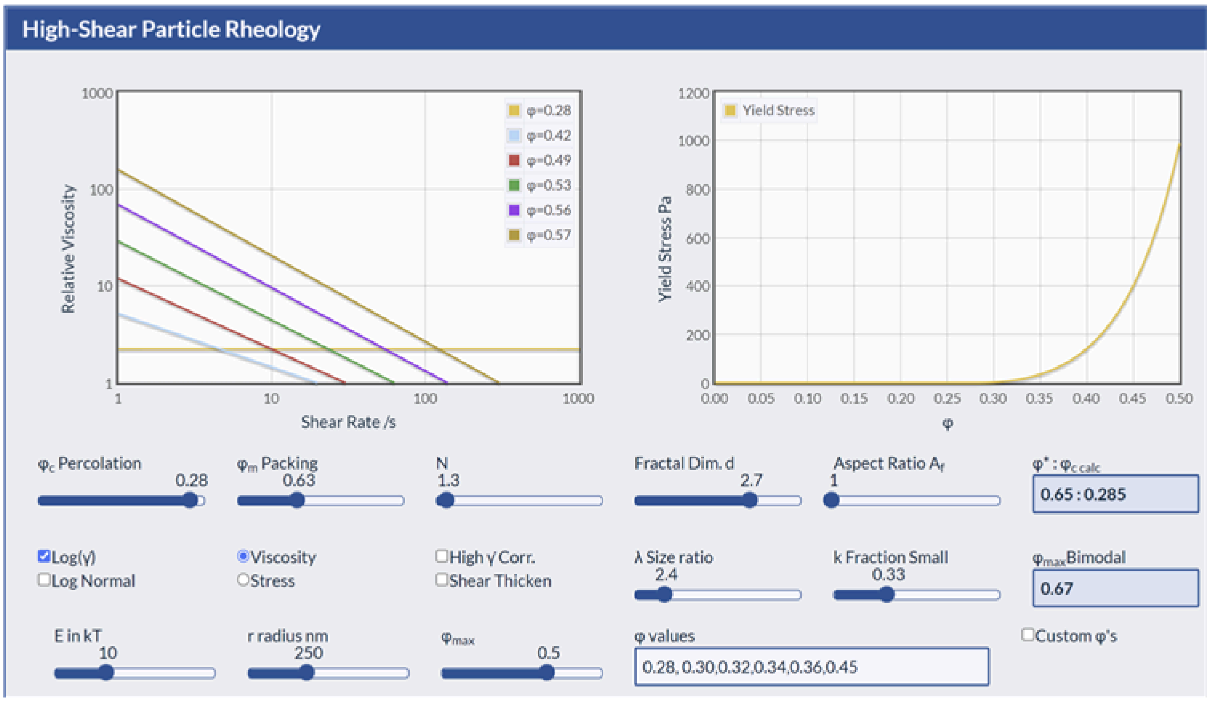
Relative density 1.5

Wt%$_{max}$ 69.2

https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php

You can read the full story at Dispersions_Rheology (Low shear)

High φ



φ_c Percolation 0.28

φ_m Packing 0.63

N 1.3

Fractal Dim. d 2.7

Aspect Ratio A_f 1

$\varphi^* : \varphi_{c\,calc}$ 0.65 : 0.285

☑ Log(γ)

☐ Log Normal

● Viscosity

○ Stress

☐ High γ Corr.

☐ Shear Thicken

λ Size ratio 2.4

k Fraction Small 0.33

φ_{max} Bimodal 0.67

E in kT 10

r radius nm 250

φ_{max} 0.5

φ values 0.28, 0.30,0.32,0.34,0.36,0.45

☐ Custom φ's

https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php

You can read the full story at Dispersions_Rheology (High shear).

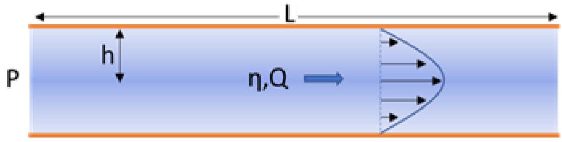
Flow Poiseuille Flow

Links

SkinCare, Pharma Formulation

When you push a fluid through a narrow gap of size h that you are fighting a $1/h^3$ (or $1/h^4$ if it's a pipe) dependency on the pressure required. Going from 100µm to 99µm makes little difference, going from 2 to 1µm increases the back pressure 8x or 16x. If one of the walls is moving then we have Flow_Couette Flow.

The Poiseuille Equations



The pressure, P, needed to create a flow Q of a fluid with viscosity η through a pipe of radius h and length L is:

$$P = \frac{8Q\eta L}{\pi h^4}$$

And for a coating flow between a gap of height h

$$P = \frac{12Q\eta L}{h^3}$$

In both cases the effects are, as intuition suggests, linear with the required flow, the length and the viscosity, so doubling any of those merely doubles the pressure required. But halving the radius or gap gives a 16x or 8x increase in required pressure.

The velocity gradient is parabolic; at a distance x from the wall, velocity is proportional to x^2, i.e. it goes from zero velocity at the wall (the *no slip boundary condition*) to a maximum at x = h.

The impacts of Poiseuille are significant.

- If you save a bit of money by buying a 12mm garden hose rather than a 18mm version, your flow is reduced 5x.
- If [you might like to open Excel and do this calculation yourself] you have a pump pulling a liquid of viscosity 100cP from a tank down a 3m long 12mm pipe to feed a coater delivering 50gsm across 1.5m at 20m/m then the pump has to pull ~0.9 bar – which isn't a problem. If you go up to 25m/m then the pump has to pull ~1.2bar … and your coating is full of air bubbles because the below-atmospheric pull has caused cavitation. This is a real example that happened to me. Fortunately, the fittings on the tank and pump allowed engineers to quickly swap over to an 18mm pipe (back pressure 0.2bar) and the production run proceeded without air bubbles.
- If you are applying a drop of adhesive it at first seems easy to squeeze it out into a wider blob to fill the gap, but soon you find that it's impossible to squeeze it thin enough, wide enough. This is Poiseuille at work but is generally called Stefan's Squeeze law Flow_Stefans Squeeze.
- Getting blood to flow down small capillaries is basically impossible because of Poiseuille. But our bodies manage it no problem. That's because Poiseuille assumes zero flow velocity at the walls

(see the No Slip Boundary Condition discussion in Cleaning_Boundary removal). A capillary has a special polymer liquid brush at the wall meaning that the blood proceeds as a plug flow – uniform velocity – with much lower back pressure needed.

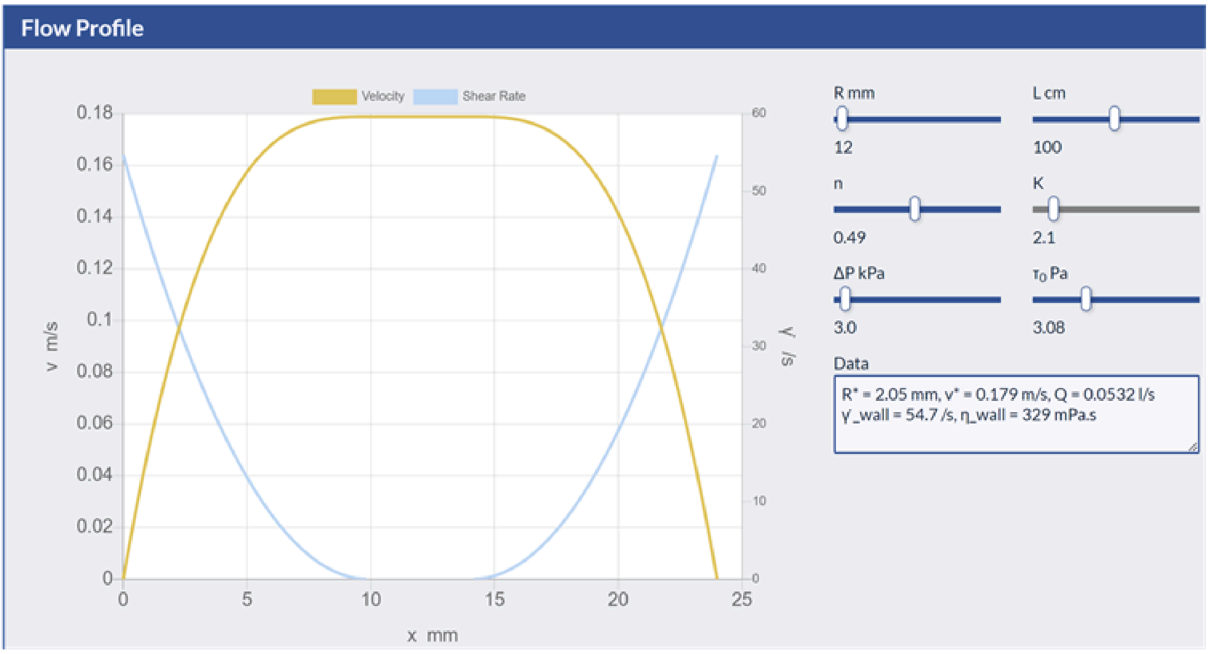
There is a wonderful paper called *From Red Cells to Snowboarding: A New Concept for a Train Track.* The train track is one packed with goose down. It is a rare example of academic playfulness combined with deep technical insight. Novel ways to overcome the no slip boundary condition are very much needed.

Non-Poiseuille flow – Shear-Thinning fluids

The parabolic shape of Poiseuille flow arises in Newtonian liquids. For shear-thinning fluids the shape is more complex. If we assume (see https://www.stevenabbott.co.uk/practical-rheology/Flow-Curves.php) a Herschel-Bulkley fluid with a yield stress τ_0 and a power law K, n dependence on shear rate so that stress τ is given by:

$$\tau = \tau_0 + K\dot{\gamma}^n$$

Then via a complex set of equations, described in the app, we can work out the velocity profile:



R mm 12

L cm 100

n 0.49

K 2.1

ΔP kPa 3.0

τ_0 Pa 3.08

Data

R* = 2.05 mm, v* = 0.179 m/s, Q = 0.0532 l/s
γ̇_wall = 54.7 /s, η_wall = 329 mPa.s

https://www.stevenabbott.co.uk/practical-rheology/Flow-Profile.php

Here we have a semi-plug flow. If you set the yield stress to 0 then you find the classic parabola.

The app mentions that the velocity profile can be measured via ultrasonics and that by reverse calculation you can measure the rheology of a fluid in real time inside a steel pipe in production. Such on-line rheology is very powerful because subtle changes in formulation which might not be spotted via standard QC can show up as a batch continues flowing through production.

Flow Polymer Viscosity

Links

Toothpastes, Water-based Adhesive, Adhesive, Solvent-Based Paint, Emulsion Paint, Emulsion Polymers, Microencapsulation

You have two problems with polymers:

- Getting enough solubility to have the concentration you require;
- Having a low-enough viscosity at the concentration you require.

The first can be solved via Dissolution_Hansen Solubility Parameters. The second requires understanding of the interaction between the MW of the polymer and Mc, the Critical Entanglement MW. The app is a quick guide to the complexities of the topic.

Polymer Viscosity

R_g @0.5+40K nm 10
M_{mono} 100
Polymer MWt 40000
Polymer M_c 10000
Flory-Huggin χ 0.25
C_{max} g/dl 15

R_g @χ+MWt nm 13.0
Ne 33.3
C* g/dl 3.0
Ce g/dl 4.1
|η| dl/g 1.67

N.bonds 400
M_{bond} 100
Bond length nm 1.4
MWt 40000
R_g nm 11.4

https://www.stevenabbott.co.uk/practical-solubility/polymer-viscosity.php

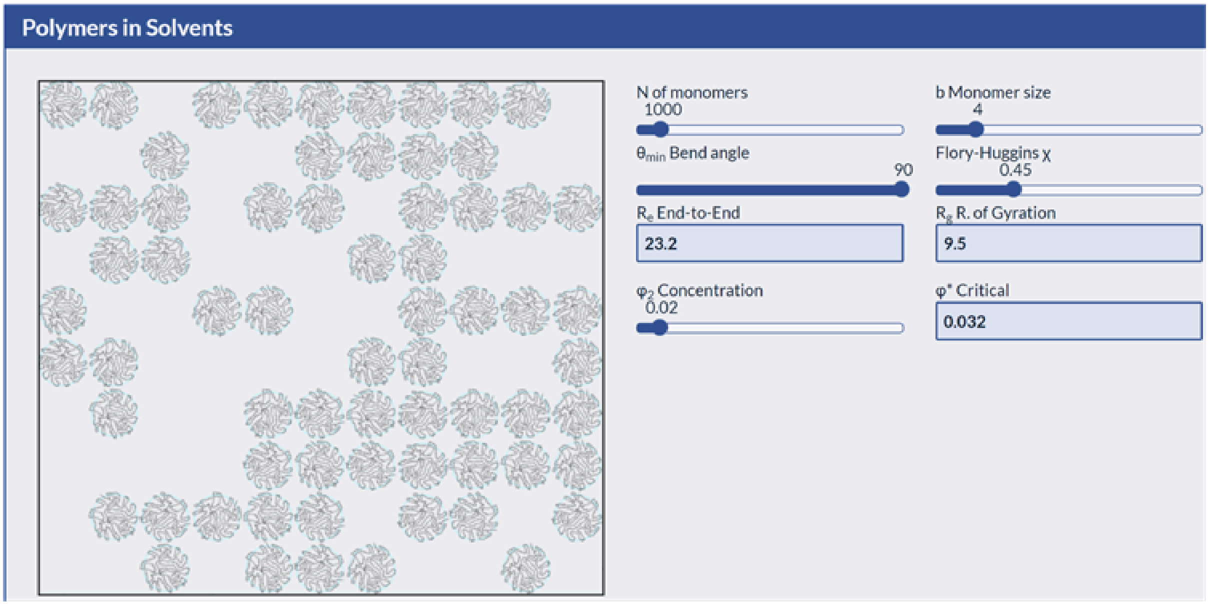
The viscosity of a polymer solution depends on its concentration and MW, of course. But the MW on its own is meaningless because what really matters is the ratio $\frac{MWt}{Mc}$ where Mc is the “critical entanglement MW”. Some polymers tangle more easily than others (i.e. at a lower value of Mc) so for a given MW the viscosity is higher because of the greater number of tangles.

The other factor is the “happiness” of the polymer in the solvent as judged by the Flory-Huggins χ parameter. When it is “unhappy” (χ > 0.5) it is curled up on itself and isn’t so viscous. In a good solvent (χ < 0.5) it expands so can tangle more, and the viscosity is higher.

The app lets you explore these basic ideas. Some of the inputs are a bit obscure and some of the calculated values are of nerd interest only, so you'll need to read the text of the app. There is a list of M_C values for a modest range of polymers, but its reliability and applicability to any specific version of a polymer of interest is unknown.

A complementary app lets us have a peek at what's going on inside the solution.

Polymer physics



N of monomers 1000

b Monomer size 4

θ_{min} Bend angle 90

Flory-Huggins χ 0.45

R_e End-to-End 23.2

R_g R. of Gyration 9.5

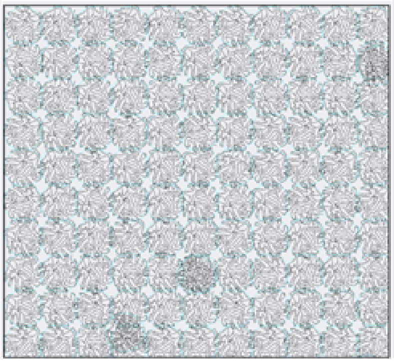
φ_2 Concentration 0.02

φ^* Critical 0.032

https://www.stevenabbott.co.uk/practical-solubility/polymers-in-solvents.php

Here we have an idealised set of polymer blobs. On average there is plenty of space between them – their volume fraction φ_2 (the 2 means the solute in the solvent, 1) is 0.02 which is below the critical fraction φ^* where they are guaranteed to be touching … and therefore having the first significant increase in viscosity. "Significant" in the eyes of polymer physicists because this is the transition into the "dilute" region which, for them, is considered to be dangerously concentrated and crowded. We can slide φ_2 to a larger value, or we can reduce φ^* by making each polymer blob larger. We can do that by:

- Increasing the MW or, in terms of the app, N the number of monomer units;
- Increasing the size of a monomer unit, i.e. using a different polymer;
- Decreasing the average bend angle, θ_{min}, between segments so that the polymer is more stretched out;
- Decreasing the Flory-Huggins χ to make the polymer happier in the solvent and, therefore, more stretched out.



In this screen shot we see overlap … those darker patches. That's because N was increased from 1000 to 2700 leading to a decrease of φ* to 0.019, leaving φ_2 at 0.02. The same result could have been achieved by tweaking any of the other 3 parameters, or some combination of them. The point is that in theory, and in real life, there are multiple ways that you can start to get your polymer crowded upon itself and, therefore, starting to build up viscosity.

Relating back to the earlier app, if, Polymer B has a monomer size b which happens to be longer, and a φ_{min} which happens to be smaller than the equivalent Polymer A, even though their Ns are the same, then B will be more easily tangled than A. This means that B's M_c is *smaller* than A's so for a given MW, B will be more tangled and viscous.

Because all this is standard polymer physics, and polymer physicists tend to give up when they are in these zones of "high" concentration (0.02 volume fraction!), there is no extra slider that allows you to see what happens at the sorts of real volume fractions (> 0.1) that are of interest to us.

What this means for your formulations

Our ability to formulate numerically with polymer solutions at the real-world concentrations that interest us is inhibited by:

- The failure of polymer physics to give us intellectual tools that work in our concentration range of interest.
- The failure of major polymer suppliers to give us M_C values for the polymers we wish to use. They give us M_N, they give us M_W but they don't give us M_C. This is a scandal.

What this means is that you have to adopt a do-it-yourself mentality.

Via Dissolution_Hansen Solubility Parameters get an idea of a good and not-so-good pair of solvents for your polymer.

Create a solution at a level that seems near the solubility limit of each solvent. It doesn't matter what that limit is, but you must know the φ values (you need to correct for densities) at the respective limits.

Because it should be a rule to never measure *a* viscosity, use your rotational rheometer to measure the flow curves (viscosity versus shear rate) across a comfortable range of shear rates, Flow_Shear dependent.

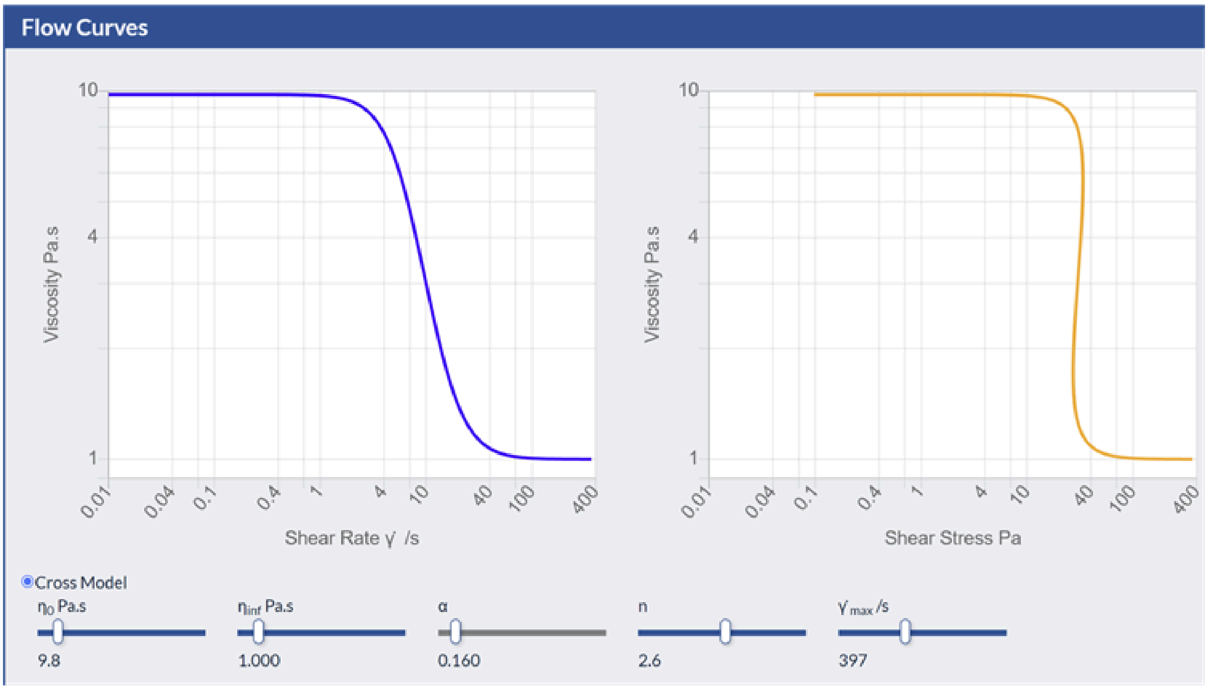
Now repeat these flow curves for stepwise dilutions of your respective solvents – that's why we started with concentrated solution.

If you have samples of two different MWs of your polymer, repeat this exercise with the 2nd MW.

You now have a rich dataset that will allow you to map out the key effects of concentration, solubility and MW. Doing any of these sets of experiments for the first time is tedious and time-consuming. Not doing such sets on a regular basis, once you've overcome the learning curve, is a sub-optimal approach to formulation.

One more thing

All this only tells you about low-shear viscosity. As we (usually) need the formulation to flow at some point then we need a good flow curve – plot of viscosity versus shear rate and, as it's a good idea to think in this way, how the viscosity depends on the shear stress:



● Cross Model

η_0 Pa.s 9.8

η_{inf} Pa.s 1.000

α 0.160

n 2.6

$\dot{\gamma}_{max}$ /s 397

https://www.stevenabbott.co.uk/practical-rheology/Flow-Curves.php

Here the viscosity plummets at 10/s which may or may not be what you desire. How do we predict this behaviour from the polymer-solvent properties? At the time of writing, I don't know. So you need good access to a good rotational rheometer to get these data.

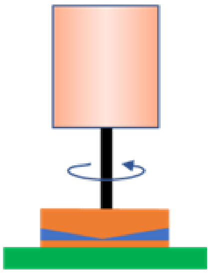
Flow Rotational rheology

Links

SkinCare, Toothpastes, Lipstick, Mascara, Water-based Adhesive, Adhesive, Solvent-Based Paint, Inkjet Inks, Pharma Formulation

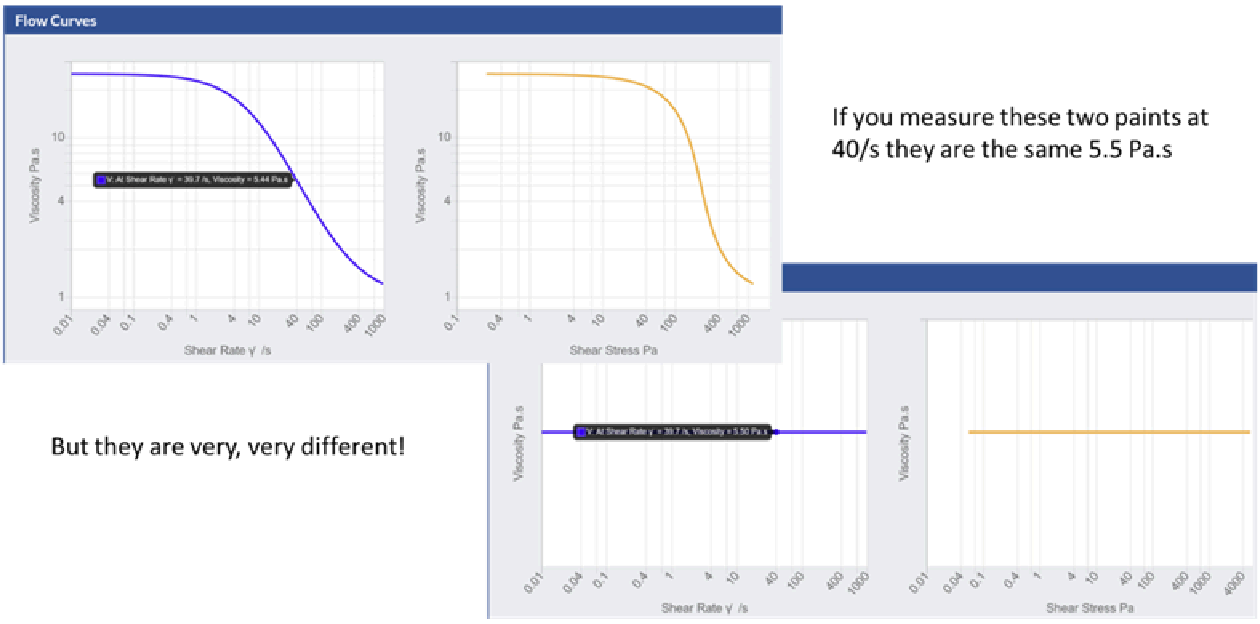
> Rotational *viscometry* can often be misleading while rotational *rheology* provides deep insights into your formulations. When you combine rotational rheology with the oscillatory version, Flow_Oscillatory rheology then you gain even deeper insights.

Never measure a viscosity…



There are lots of viscometers out there, such as cups (https://www.stevenabbott.co.uk/practical-rheology/Cup-Viscosity.php), and simple rotating discs ("Brookfield"), neither of which gives a genuine viscosity measurement, then there are scientifically meaningful Flow_Couette Flow and capillary viscometers. The rheometer shown in the diagram is working in rotational mode but can also work in oscillatory mode (Flow_Oscillatory rheology) so having a rheometer capable of doing both is highly recommended.

For formulators the advice is clear: "Never measure *a* viscosity", meaning that any single measurement can be misleading. You can have two formulations with the same, measured, viscosity under your specific technique, yet they can have totally different rheological properties and, therefore, performance as formulations. The diagram shows an extreme example. The first is a shear thinning liquid and the second is a Newtonian liquid. If you measure each of them at 40/s rate you would say that they are the same – which obviously they are not!



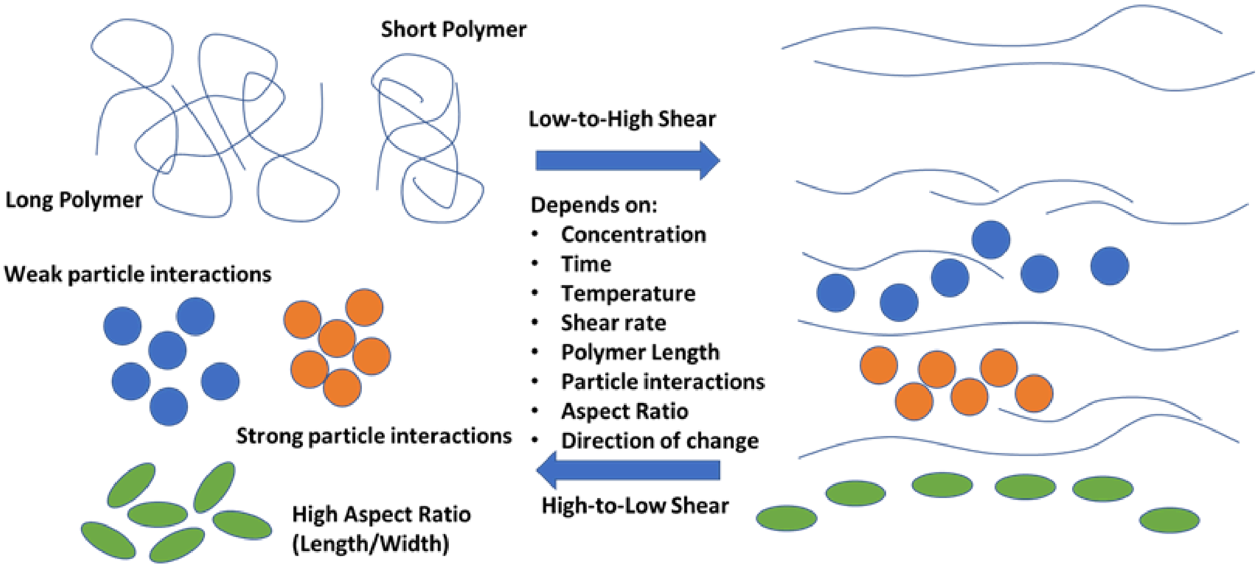
If you measure these two paints at 40/s they are the same 5.5 Pa.s

But they are very, very different!

For QC a well-thought-out single viscosity test is acceptable, but you have to be sure that the sorts of errors that appear in production can't produce single-viscosity values that fool QC.

… Because it’s complicated

Why are viscosity and rheology so complicated? Because our systems are complicated.



We have short and long polymers, tangled and un-tangled, we have particles with weak and strong interactions with themselves and the polymers, we have different aspect ratios … and as we apply shear all sorts of different things can happen.

The most common complication, as implied by the shear arrows in the diagram, is shear-dependent viscosity: Flow_Shear dependent. Then we have the complications of thixotropy, Flow_Thixotropy. We have to remember that although thixotropic systems are shear-dependent, there is a big difference between them. Classic shear-dependence depends only on the shear rate. Thixotropy depends on the time spent at a given shear rate.

As the diagram also implies, we don’t just have issues of polymers, we have particles as well, Flow_Particle Viscosity.

At low shear rates we also encounter the phenomenon of Flow_Yield Stress which can be important for cosmetic creams and foods like mayonnaise, where, in both cases, you don’t want the fluid to flow when sitting as a blob. In other cases, yield stress can be a nuisance as it stops fluids flowing when we want them to.

Flow Shear dependent

Links

Deodorant Sticks, SkinCare, Toothpastes, Water-based Adhesive, Adhesive, Solvent-Based Paint, Pharma Formulation

We know that formulations shear thin – viscosity gets lower with higher shear. A good viscometer or, better, a rheometer can measure viscosity at shear rates from low, say 0.1/s to high, 1000/s – called the Flow Curve. This range covers typical conditions of formulation manufacture and use. To compare formulations you need to describe the flow curve with a few parameters. The Cross model is convenient. It has two intuitive parameters:

1. η_0 the viscosity at very low shear;
2. η_{inf} the plateau at high shear .

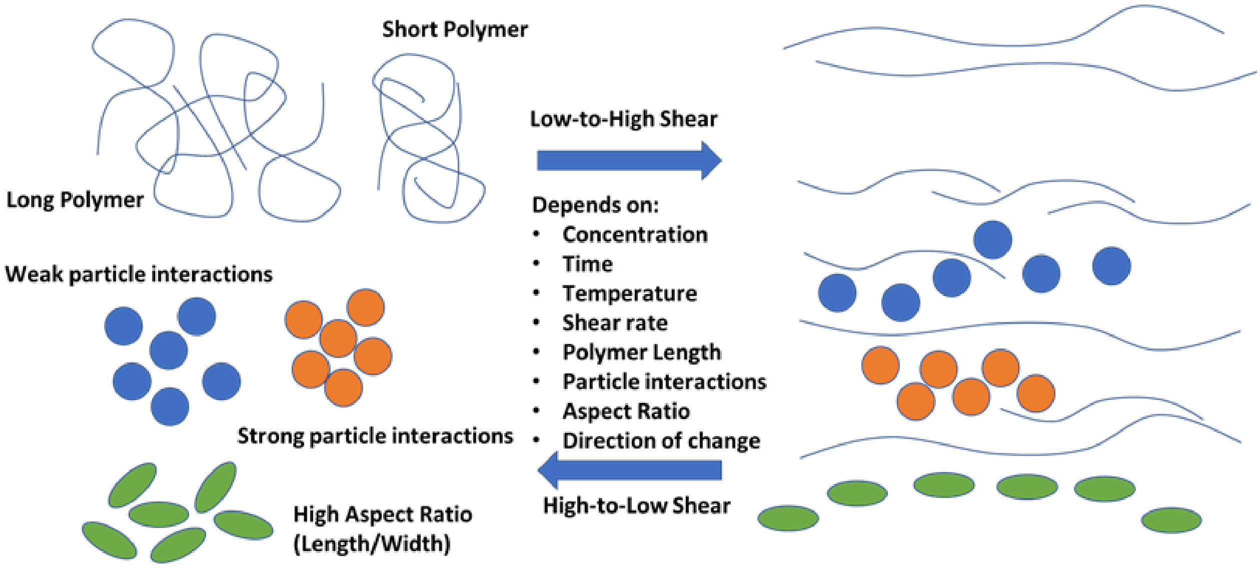
The curve from high to low is described by two more parameters:

1. n, which describes the "sharpness" of the transition;
2. α, which describes the position for the fall-off of viscosity.

With those 4 parameters for each formulation, you have a good basic description of key parameters. Other parameters are discussed below.

Why do we have shear-dependent behaviour?

Because our systems are complicated.



We have short and long polymers, tangled and un-tangled, we have particles with weak and strong interactions with themselves and the polymers, we have different aspect ratios … and as we apply shear all sorts of different things can happen. The app lets you see lots of the different behaviours.

Shear Dependence

η0 Pa.s 10

ηinf Pa.s 0.1

α 0.052

n 1.3

n (Power Law) 0.5

γ̇/s 100

η Pa.s 1.1

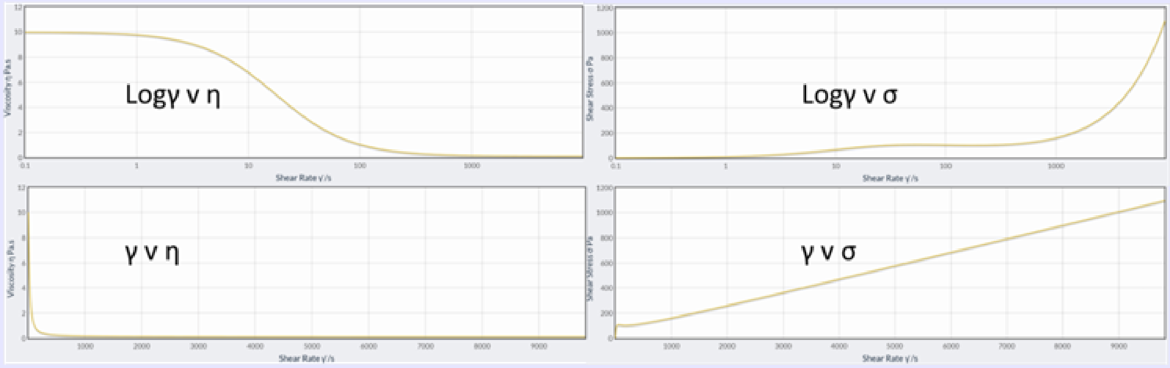
τ Pa 113.9

Power Law
☐Show Power Law

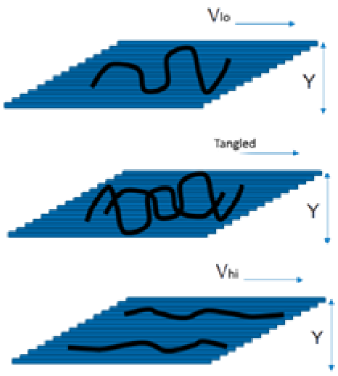
● Log(γ̇) ○ γ̇ ○ Log(σ) ● Viscosity η ○ Stress σ

https://www.stevenabbott.co.uk/practical-rheology/Shear-Viscosity.php

It is instructive to select the non-log option of shear rate to see the real shape of these curves. Log plots are often deceptive. You can also plot with log shear stress σ on the x-axis, or σ on the y-axis something that rheologists love to do. These alternative plots are common in the literature so it is good to get used to seeing them. Much of our fear of, and confusion with, rheology comes from seeing the same core data plotted in different ways. These are all the same data. Use the app to get used to seeing the same phenomena in different ways; each view has its merits.



We know that the simple viscosity relationship discussed in Viscosity Basics was derived by Newton and fluids which obey that law are Newtonian fluids. The cause of viscous drag is familiar to those who drive in busy traffic. If a driver in a slow lane moves into a faster lane that causes the cars behind to slow down to accommodate the new arrival.



For ordinary molecules in liquids their random movement between streamlines causes the same slowing down (viscous drag). When polymer molecules are concerned, if one part of the polymer is in one streamline and another part is in a different streamline then the polymer starts to be stretched, causing a bigger drag between streamlines. The effect is larger if concentrations and MWts are large enough to create tangles. However, with faster streamlines (with V going from lo ◊ hi) the polymer chains tend, on average, to become more and more aligned with the flow so they cause proportionally less drag between the streamlines. So higher shear leads to lower viscosity in many polymer systems. For such non-Newtonian systems the graph of viscosity versus shear rate, γ, can be conveniently described using the Cross formula:

$$\eta = \eta_{inf} + \frac{\eta_0 - \eta_{inf}}{1 + (\alpha\gamma)^n}$$

where η_{inf} is the viscosity at (essentially) infinite shear, η_0 is the viscosity at zero shear and α and n are fitting constants.

Systems where the viscosity decreases with shear are *pseudoplastic*. Shear thinning should not be confused with Flow_Thixotropy. Although most thixotropic systems are pseudoplastic, not all pseudoplastic systems are thixotropic!

The same formula handles dilatant systems where the viscosity gets higher with shear. The explanation for that is a log-jam effect where particles/polymers can't respond (reorientate) quickly enough when bumping into each other.

The shear thinning behaviour of systems with high particle loadings is described in Flow_Particle Viscosity.

For a more detailed look at these systems, the Flow Curves app https://www.stevenabbott.co.uk/practical-rheology/Flow-Curves.php provides more insight. In particular you can describe the low-shear regime via a Power Law model, used in the understanding of Coating_Levelling Theory.

Flow Stefans Squeeze

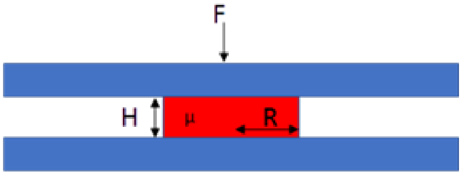
Links

Water-based Adhesive, Adhesive

How hard can it be to squeeze a drop of, say, adhesive so that it creates a thin, even layer? It turns out to be surprisingly hard. This means we have to find workarounds which have their own problems.

H^3 and R^4

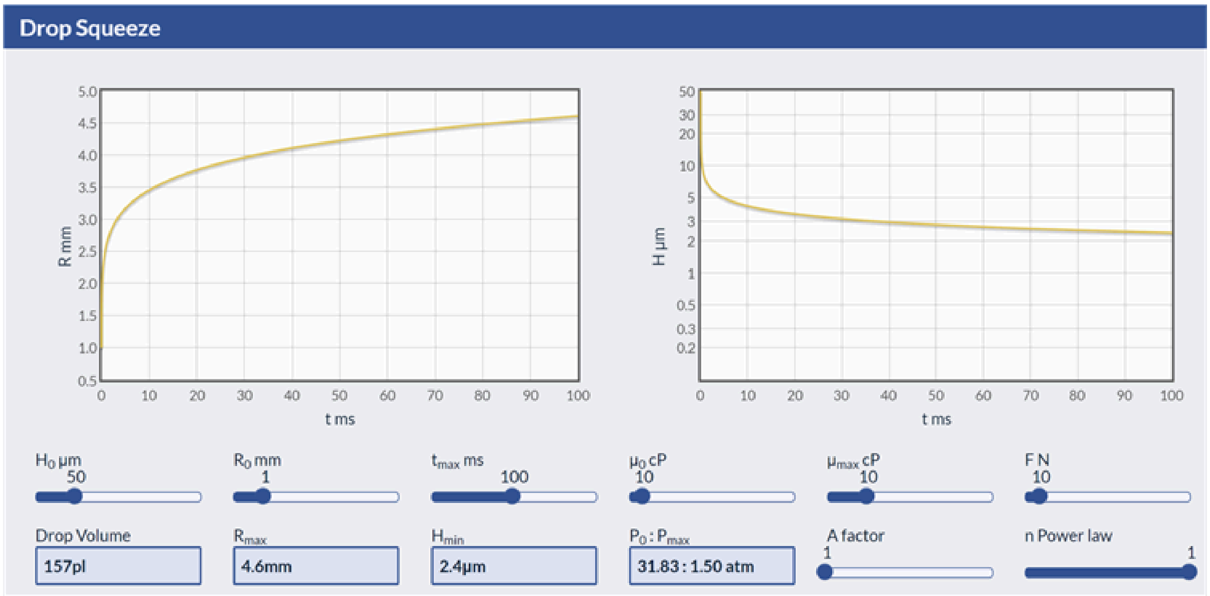
If you have a phenomenon with a linear or even squared dependence, you are probably OK. If you see a cubic or quartic dependence, you know you are in trouble.



Here we have a drop of radius R, thickness H and viscosity μ which we are squeezing with a force F. How quickly does the thickness reduce and the radius increase? The effects of F and μ are linear, so follow your intuitions. The problem is, as Stefan showed in 1874, that the squeeze slows down very rapidly because of the H^3 and R^4 terms.

$$\frac{\delta H}{\delta x} = \frac{3FH^3}{3\pi\mu R^4}$$

Because we are interested in both H and R, the app shows them both:


Drop Squeeze

H_0 μm 50

R_0 mm 1

t_{max} ms 100

μ_0 cP 10

μ_{max} cP 10

F N 10

Drop Volume 157pl

R_{max} 4.6mm

H_{min} 2.4μm

$P_0 : P_{max}$ 31.83 : 1.50 atm

A factor 1

n Power law 1

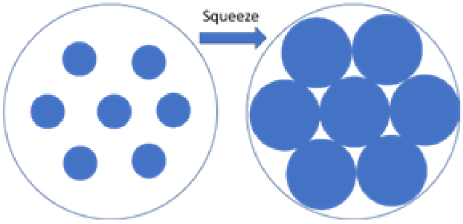
https://www.stevenabbott.co.uk/practical-adhesion/drop-squeeze.php

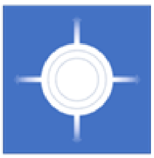
In both graphs we see that getting a thinner, wider drop is easy at first, but things rapidly slow down because of those cubic and quartic effects. When you need, e.g. for a Adhesion_Butt a thin layer of adhesive, you simply cannot do it with a single drop.

Breaking the problem into multiple drops gives a dramatic improvement. If the individual drops are 1/5 the radius of the single drop, then the problem is 1/625th of the original.

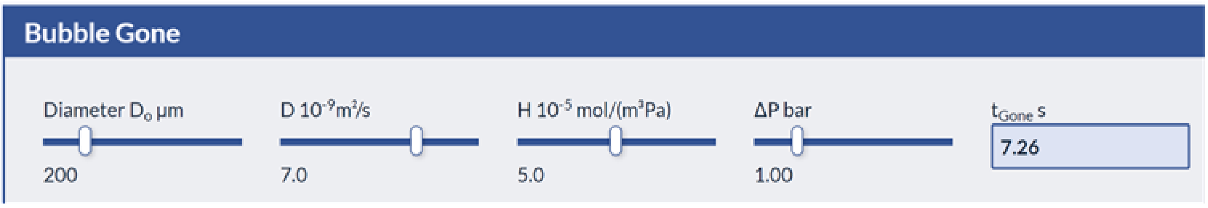
This, however, leads to a follow-up problem.

Joining the dots





We've made great progress in filling the space with the 7 dots instead of 1 larger dot which would already have failed. But what about those trapped bits of air between the dots? Although there's no app to deal with that geometry, you can get a good idea of how easy/hard it will be for the air in the round bubble to diffuse, under pressure, into its surrounds:



Bubble Gone

Diameter D_0 μm: 200

D 10^{-9}m²/s: 7.0

H 10^{-5} mol/(m³Pa): 5.0

ΔP bar: 1.00

t_{Gone} s: 7.26

https://www.stevenabbott.co.uk/practical-adhesion/Bubble-Gone.php

In practice, it turns out that these multi-drop blobs of adhesive form coherent films most of the time, though the industrial adhesives community spends lots of time debating whether drops, stripes, stars etc. are preferred for a specific geometry.

While writing my book *Sticking Together* I found that I needed to refer to Stefan at some point … then at another, then another… I'd not realised that it's such an important part of adhesion, and for other parts of formulation space. It subsequently turned out that the question of how many drops you use, in which geometry, is a huge question for the adhesives industry. There seems to be no appable algorithm about the big question which is whether a bubble will get trapped, Bubble-Gone was the closest relevant calculation I could find.

Flow Thixotropy

Links

SkinCare, Water-based Adhesive, Adhesive, Solvent-Based Paint

> Thixotropy must *not* be confused with shear thinning effects described in Flow_Shear dependent. Yes, thixotropes shear thin, but they take time to recover their original viscosity when the shear stops. A thixotropic paint can have an otherwise unusable high low-shear viscosity because when stirred, its low-shear viscosity is now low to make painting and flow-out of brush/roller marks easy – but can return to the high viscosity value to avoid sag as described in Coating_Levelling Theory.
>
> Thixotropy confuses everyone because it is hard to get a meaningful measurement.

η_{start} Pa.s 90
η_0 Pa.s 10
η_{inf} Pa.s 0.01
α 2
n 1
Max $\dot{\gamma}$ /s 200
Scan rate $\dot{\gamma}$/s 0.3
τ_{break} 2.8
τ_{build} 0.02
$\dot{\gamma}$ power 1.5
S Pa/s 13.5
S_{rel}% 15.3
☑ Log($\dot{\gamma}$)
● Viscosity η
○ Stress σ
☑ Log(η or σ)

https://www.stevenabbott.co.uk/practical-rheology/Thixotropy.php

Thixotropy is the time-dependent change in viscosity, different from the shear-rate dependent viscosity (see Flow_Shear dependent) that is normal for pseudoplastic systems. It is complex to study because each time you look at it, you are changing it! In addition to this rotational rheometer technique, thixotropy can also be studied using an oscillatory rheometer, described below.

Clearly thixotropy is important for paints and coatings. Too little and it's difficult to control the application. Too much and the formulation is hard to get to flow when you need it.

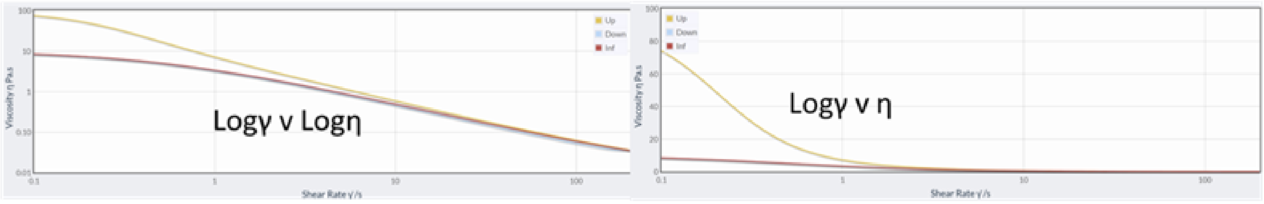
There are plenty of papers with theoretical treatments of thixotropy hysteresis loops. I've chosen for the app an approach that is mathematically tractable and theoretically clear. This starts with the Flow_Shear dependent app, using the Cross model to create the "equilibrium" line which is what you would measure if your thixotropic fluid had been infinitely well-sheared and had no time to recover. The plot goes up to your chosen maximum shear rate.

The Up line (the flow curve when you *increase shear rate*) is what you would expect from your thixotropic liquid which starts at near-zero shear (0.1/s) with a viscosity of η_{start}.The rate at which it approaches the equilibrium line depends on three factors.

1. How fast you scan your shear rate - the slower you scan, the longer it takes to get to your maximum shear rate and, therefore, the sooner (visually, in terms of shear rate) you reach the equilibrium line.
2. The timescale for breaking up your structure, τ_{break}. The longer this is, the more thixotropic your system will be.
3. This is combined with how much τ_{break} changes with shear rate, expressed as $\dot{\gamma}$ power. If this is 1 then the time decreases proportionally to shear rate

In principle, your "down" curve (now the flow curve with *decreasing shear rate*) will be very different from both your up curve (of course) and your equilibrium curve. This depends on τ_{build}, the timescale for re-building your original structure. This is assumed to be shear-rate independent. You only see significant effects (i.e. different from the equilibrium line) with fast scans and if τ_{build} is quite large.

You may be surprised that the curves don't look too much like the ones we always see in simple explanations. The reason is that there are lots of ways of looking at the data, and lots of "equilibrium" curves, so you have to first get your equilibrium curve right, which might need one sort of scale (e.g. log-linear viscosity), then view the thixotropy via a log-log plot with shear stress instead of viscosity. When testing the app against published papers I often thought the results were wrong, but then found I was using the wrong scale in the app. Here are the same data plotted with log shear rate but with log and linear viscosity. Neither is right or wrong, nor is it right or wrong to choose linear shear rates (not shown). The app lets you choose whichever view gives you the most insight … and ability to compare to the literature.



Thixotropic Area

The area S between the up curve and down curve (though here it's between up and equilibrium) is some sort of indication of the thixotropy of the system. However, the area has no absolute meaning because it depends on scan rate and maximum shear rate.The relative area S_{rel} which is S divided by the area under the up curve (S_{max}) is said to be rather more independent of the experimental variables. Both values are calculated for you. When you are happy with this sort of scan, it becomes routine to compare thixotropic behaviour, probably via S_{rel}.

Another way to look at thixotropy

G'$_0$ 0.1

G'$_\infty$ 1.2

Power law n 0.2

Tanδ$_0$ 2

Tanδ$_\infty$ 0.1

t$_{max}$ s 300

☐Log(t)

☐Log(G or η)

https://www.stevenabbott.co.uk/practical-rheology/Thixotropic-Recovery.php

If you use Flow_Oscillatory rheology you can analyse thixotropy by first breaking up any thixotropic structure by a rotational pre-shear, then use gentle oscillations (e.g. 1 Hz) at low strains (e.g. 1%) to see how long structure takes to recover.

As explained in the app, G' follows a recovery curve starting from the original G_0' via:

$$G' = G'_0 + At^n$$

You characterise your formulation via you G_0', the constant A and the power law n, though I have no experience with how to use those parameters. My uninformed view is that in the long run this would be a better approach to managing thixotropy, especially when issues such as levelling and sag are also measured via oscillatory techniques.

Flow TTS-WLF

Links

SkinCare, Lipstick, Mascara, Water-based Adhesive, Adhesive, PSA

One of the remarkable laws of physics is that in many situations time and temperature are equivalent, allowing TTS (Time Temperature Superposition) which can conveniently be carried out using the WLF (Williams, Landell and Ferry) equation.

WLF allows us to do many formulation tricks, so it's worth getting to know the simple formula and how to use it.

Polymer properties

Take a polymer and measure its elastic modulus, E' and its tendency to move under stress, E". See Flow_G' and G" for more information, noting that G values are shear moduli while E values are tensile ones.

How well do the two values you've measured, E' and E" characterise the polymer? The answer is "poorly". We know that below the polymer's glass transition temperature, Tg, the properties change, making it stronger but potentially more brittle; similarly as it reaches its melting point, the polymer gets weaker and under a tensile load it can flow gently, i.e. it shows creep, Flow_Creep and Relaxation..

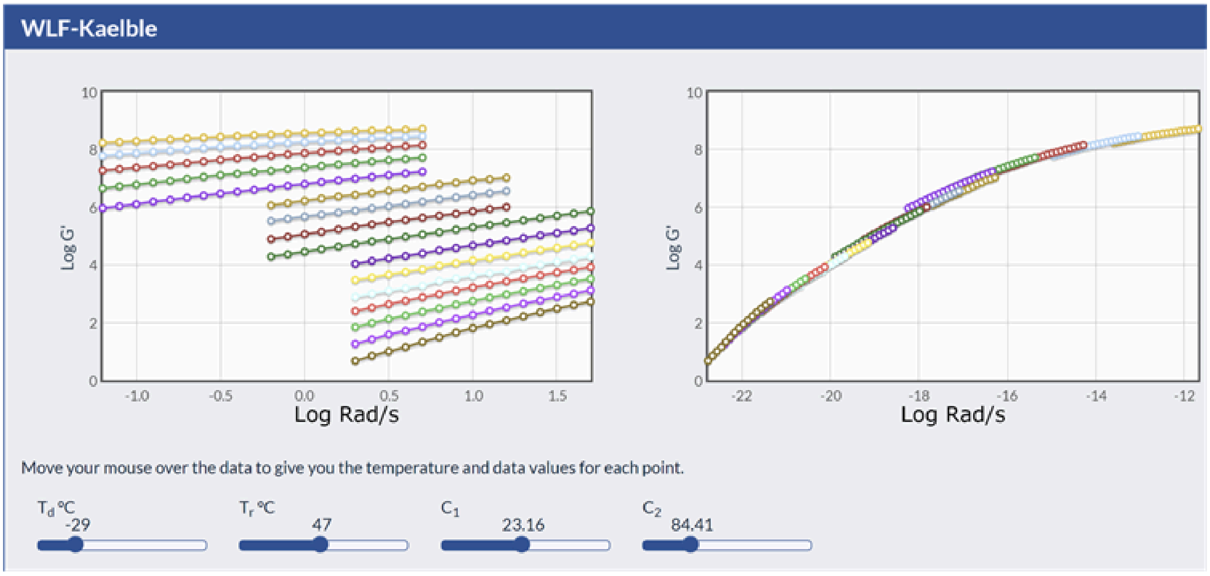
But even if we knew these properties at different temperatures, we'd still not characterise the polymer properly. We know that if we hit the polymer really quickly it can shatter – it's brittle at those speeds. And if we leave it under tensile load for a long time, it can creep.

The preceding paragraphs show not only that properties change over temperature and over time, but that the resulting phenomena are the same – there's no difference between, say, brittleness viewed by low temperature measurements at normal speeds or by high speed measurements at normal speed. That's at the heart of TTS – if you measure some properties at a given temperature and a given timescale, you can work out what the properties would be at a different timescale at that temperature or a different temperature at that timescale.

The equation for doing this transformation is WLF:

$$\log\left(a_t\right) = -\frac{C_1\left(T-T_r\right)}{C_2+T-T_r}$$

To understand what the mysterious a_t is we can first see the equation in action:



Move your mouse over the data to give you the temperature and data values for each point.

T_d °C -29

T_r °C 47

C_1 23.16

C_2 84.41

https://www.stevenabbott.co.uk/practical-rheology/WLF-K.php

On the left is a collection of sweeps (in this case of G') done at various frequencies from $10^{-1.25}$ to $10^{1.75}$ rad/s and (colour coded) temperatures. On the right are the same data linearised via WLF. You see that they now cover a *virtual* timescale range of 10^{-23} to $10^{-12.5}$ rad/s all at the reference temperature of 47°C.

The a_t value is the amount by which the timescale (or its inverse, frequency) has to be changed to translate a value at one temperature to produce the same value at the reference temperature.

One of the upsides of understanding WLF is that at a glance you can look at a rheology graph and know whether it has been TTS corrected. No one can measure G' at 10^{-23} rad/s, so the value *must* be a virtual value via WLF.

The constants C_1 and C_2 are just fitting constants, but they have some sort of meaning. C_1 is the range of frequencies spanned by the relevant data, in this case 23 orders of magnitude. C_2 is the temperature range over which the properties change by half C_1. In the app a more sophisticated variant of WLF is used so the meanings of C_1 and C_2 are not so clear.

Gathering and fitting WLF data

Because of the abstract nature of the WLF equation and the non-intuitive a_t, there is a universal fear of attempting to gather and fit WLF data. Because the app can load any reasonable WLF dataset (provided as a .csv file in the format specified on the app page), you might be tempted to simply use the app. And your rheometer will also have WLF so you don't even have to think about it – just press a few buttons.

But it is highly recommended to do at least one fit yourself in something like Excel. The simpler WLF app, https://www.stevenabbott.co.uk/practical-rheology/WLF.php, provides a link to an Excel dataset (it's real data, but without details of the sample being tested). You can go step-by-step through the columns to see how the calculations are made. Finding the fitting parameters involved some trial and error (go ahead, try your own fitting values), but again wasn't so hard. There is no way I know to fit the parameters using the Excel Solver.

Here's how that Excel dataset was created. During some work on an unfamiliar rheometer we couldn't find the WLF package in the rheometer's software. In my hotel room that evening I had to work out how to do the WLF transform in Excel. Although it's trivial, it seemed very hard. In the morning I presented the fit – and they also had an overnight email explaining where to find WLF in their software. Fortunately the results were the same.

That spreadsheet has been downloaded many times from my WLF page because everyone else finds their first WLF to be as hard as I did.

Flow Yield Stress

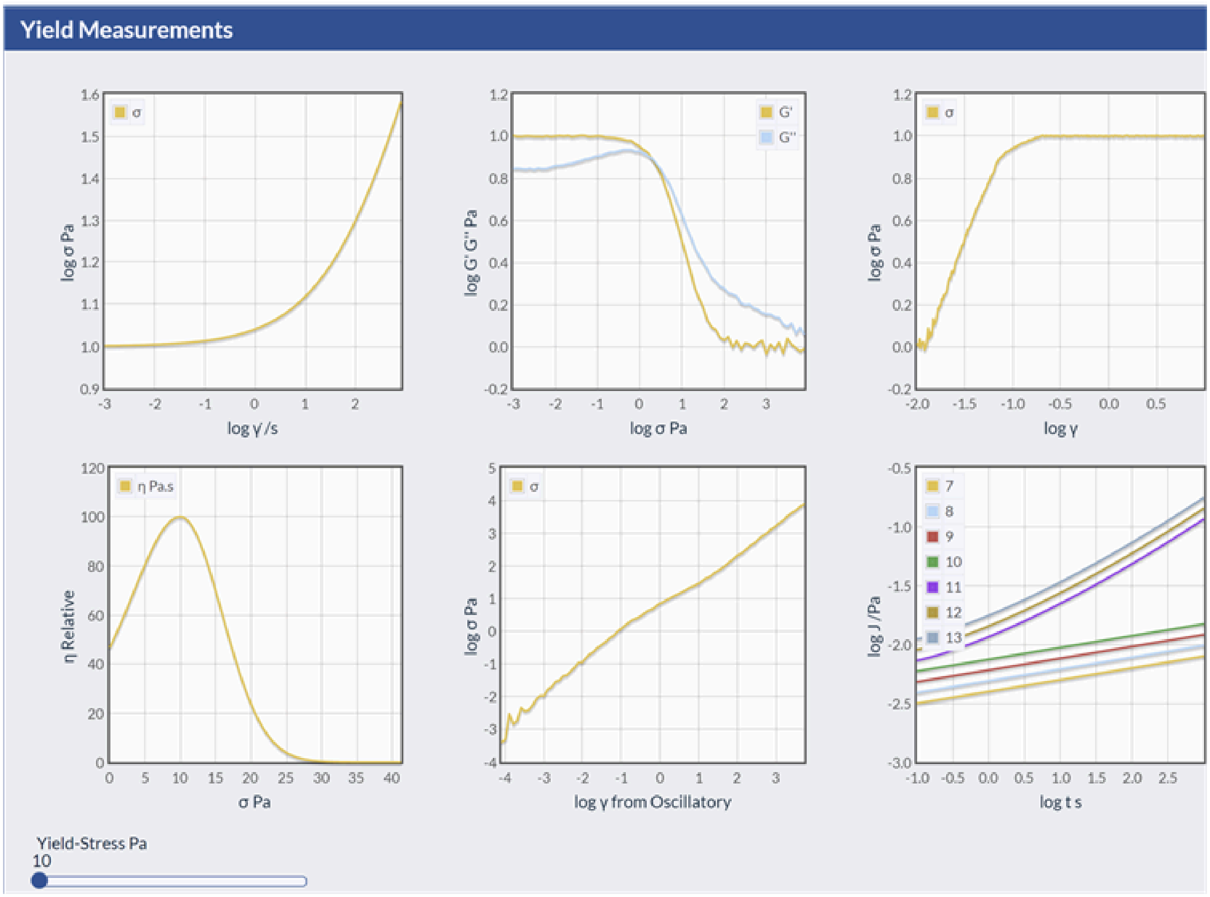
Links

Deodorant Sticks, Sun Screens, Soaps and Washing, SkinCare, Toothpastes, Lipstick, Mascara, Water-based Adhesive, Adhesive, Solvent-Based Paint, Pharma Formulation

> Although we expect liquids always to flow, even if slowly, some formulations have a “yield stress”. Small values of stress (e.g. from gravity or surface tension), below the yield stress, produce no flow. This is perfect for fighting sag as described in Coating_Levelling Theory and for producing a luxury feel for SkinCare products.
>
> It is difficult to get reliable yield stress values because measurements need to be made with small forces in delicate rheometers. What’s important is to understand the idea and find a measurement technique that with modest effort gives reasonable values for those products where yield behaviour might be important.

For some formulations you want them to flow under all circumstances. For others you want a blob of cream on your hand, or some paint on a surface to remain in place, resisting the force of gravity. This resistance is called yield stress. Yet with some modest extra stress (rubbing with a finger, moving with a paint brush) you want the formulation to flow easily. So the yield stress must be tunable to your specific needs. This means that you need to measure it.

Unfortunately, yield stress is hard to measure precisely and there are at least 6 ways to measure it. Why are there so many ways? Well, experts have their different opinions and some even argue that yield stress isn’t a real phenomenon so there’s nothing to be measured! The app text gives some explanations of the different techniques. And don’t worry, even if the purists are right and yield stress doesn’t exist, formulators over the decades have gained deep insights into their systems by assuming that it does exist and that they can measure it with their choice from the different techniques:



Yield-Stress Pa
10

https://www.stevenabbott.co.uk/practical-rheology/Yield-Measurement.php

There are two reasons for showing this app:

1. You will often find the same problem being addressed via yield stress measurements using different techniques, so it's good to recognise that these different graphs are all showing exactly the same yield stress value.
2. If you are having trouble getting good values from your current technique, you might be encouraged to swap to another one. They each have proven useful to some people in some regions of formulation space.

So don't worry too much about whether your technique is giving the definitive values (it probably isn't); find whichever technique is easiest and most reproducible on your equipment for your range of formulations, and find the yield stress values for a variety of products (your own and competitors'), then find a correlation with good overall performance versus yield stress. You can then tune future formulations to these optimal values.

These 6 ways were described in an excellent academic paper (cited in the app). I had a chance to meet the professor behind the work. I accidentally referred to him as a rheologist. "No, I'm not a rheologist; I'm someone who is passionate about using rheology to solve problems." I like this distinction. I hope that FST readers will become more passionate about using rheology to solve their own problems.

Fragrance Activity Coefficients

Links

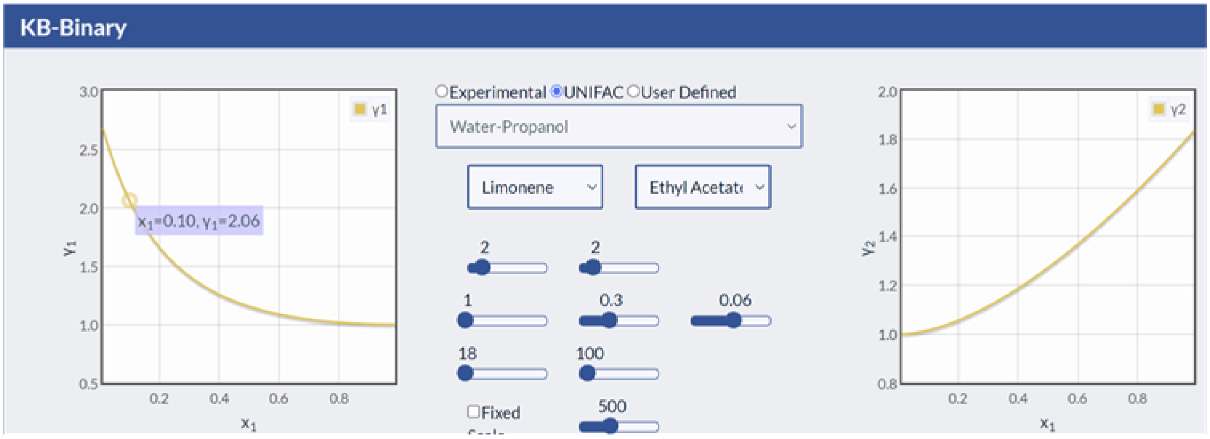
[Deodorant Sticks](), [Soaps and Washing](), [Laundry Liquids](), [Fragrances](), [Surface Cleaning](), [Microencapsulation]()

> The evaporation of the components of a complex fragrance under the assumption of ideal mixtures is described in [Fragrance_Vapour Pressure](). The activity coefficient effects (non-ideal mixtures) when “the same” fragrance is provided in different formats (shower gels, candles, hand creams …) are so significant that “the same” fragrance is often very different.

Activity coefficients

If you measure the vapour pressure of a volatile molecule above an ideal mixture, it will be exactly equal to the value calculated from its standard vapour pressure and its mole fraction in the mixture. If, however, the vapour pressure was, say, 1.5x larger than expected, you would say that the activity coefficient of this molecule at that mole fraction is 1.5.

You can build up an intuition of the sorts of activity coefficients you might expect from similar or dissimilar molecules via the Binary Mixtures app:



KB-Binary

○Experimental ●UNIFAC ○User Defined

Water-Propanol

Limonene

Ethyl Acetate

☐Fixed Scale

https://www.stevenabbott.co.uk/practical-solubility/kb-binary.php

Ignore the greyed-out box that says Water-Propanol. This is limonene (1) and ethyl acetate (2) calculated via the standard UNIFAC method.

The left-hand graph shows that small amounts of limonene (e.g. $x_1 = 0.1$) are unhappy in excess ethyl acetate, with the mouse readout showing that the activity coefficient, $\gamma_1 = 2.06$ – there would be 2x the amount of limonene in the vapour phase than expected from its mole fraction. As you increase the mole fraction of limonene it is happier, so at $x_1 = 0.5$ $\gamma_1 = 1.15$, only 15% more than ideal.

Activity coefficients are non-symmetrical. When ethyl acetate has $x_2 = 0.1$, $\gamma_2 = 1.71$, it’s less unhappy in limonene than limonene is in ethyl acetate, though at $x_2 = 0.5$, $\gamma_2 = 1.27$ so it’s *more* unhappy.

This simple example shows that the chances of calculating the activity coefficients in a complex fragrance are small. So why are we bothering with this chapter?

The same fragrance in different media

Suppose I came up with an exclusive “Essence of FST” fragrance and wished to provide this wonderful aroma experience across an FST product range of shower gels, candles and hand creams. Now suppose that the fragrance included limonene and ethyl acetate. In the candle, the activity coefficient of the alkane limonene will be ~ 1 while the ethyl acetate will be ~ 2. In the hand cream, which we can assume to contain plenty of polar molecules, the limonene might be, say, 1.5 while the ethyl acetate will be, say, 1.2.

If we put the same limonene:ethyl acetate ratio into each formulation, the perceived aroma would be ethyl acetate rich in the candle and limonene rich in the hand cream.

So now you understand the problem faced by major brands when they want to establish a common fragrance across a range of products. They need a different fragrance formulation for each product. Sometimes this can be via adjustments of ratios within a single formulation, but for something like a shower gel it might not be possible to accommodate the super-large activity coefficients for some components, so the fragrance has to be “re-imagined” – to give the same overall impression without using the same ingredients.

I once met a “nose”, one of only a few 100 people on the planet with exceptional abilities to analyse and create fragrances. When I asked her if I could smell some ambergris she dipped a little strip into her sample and I held it under my nose … and smelled nothing. “You have to wave it under your nose!” Sure enough I was able to experience the amazing complexity of this combination of ambroxide and ambrinol.

Head space analysis

Given that calculations of activity coefficients of a 20-component fragrance mixture are impossible and then calculations on the fragrance in a candle or a hand cream are even more impossible, what can you do about it?

Head space GC is the least bad option. Comparisons of the fragrance in its normal carrier (the so-called neutral molecules such as dipropylene glycol), in a candle and in a hand cream would give a good idea of how key fragrance components shift in relative amount in the vapour phase.

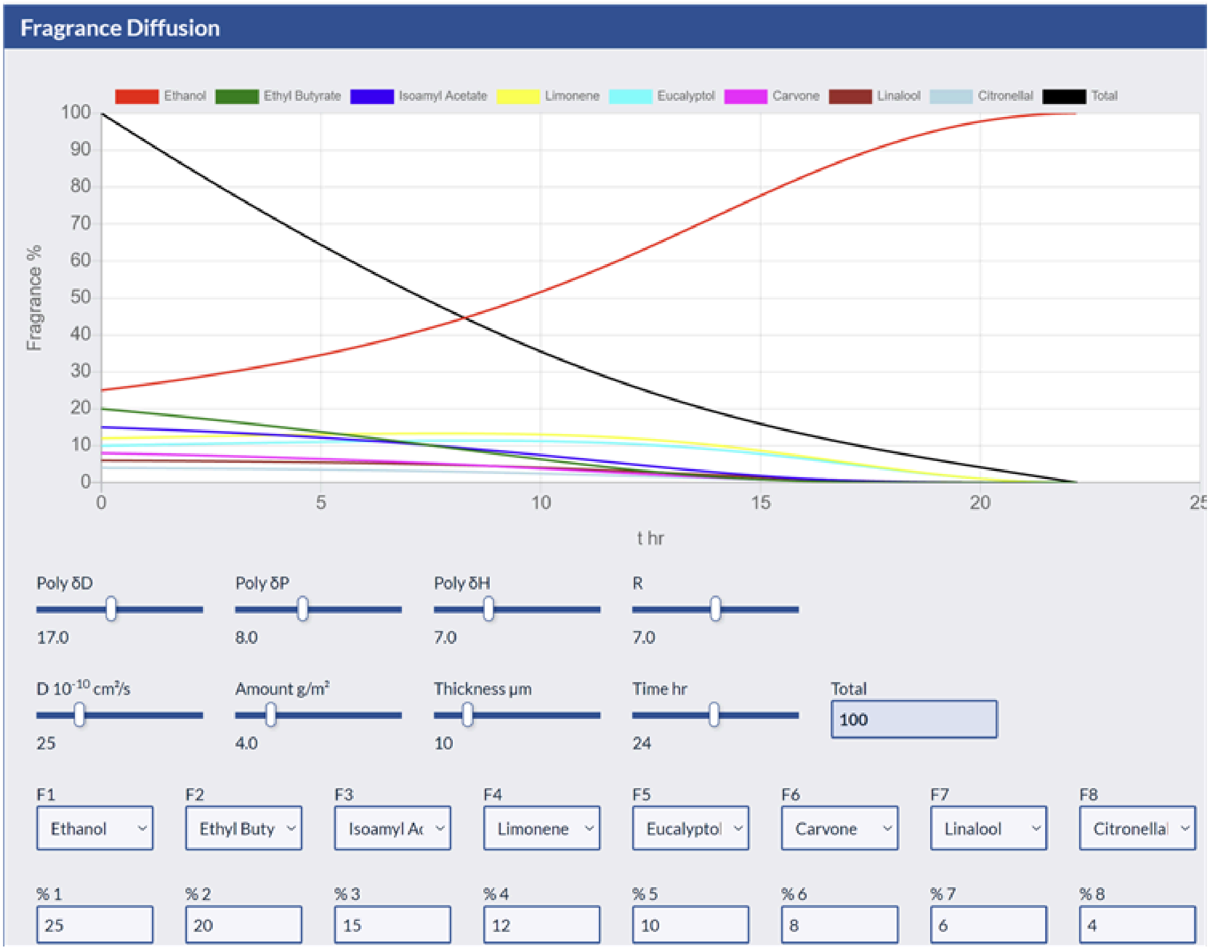
Fragrance Barrier Properties

Links

[Laundry Liquids](), [SkinCare](), [Fragrances](), [Microencapsulation]()

> If you have a fragrance or flavour formulation you might want it to be released only when and where you want, so you might want a barrier in place. Unfortunately, for many practical polymer barriers, different fragrance molecules will diffuse through the barrier at different rates – so although you might keep the majority of your fragrance, some components might disappear and, therefore, change the fragrance or, in the case of food & drink, give rise to “flavour scalping”.

Loss via partition



Poly δD 17.0

Poly δP 8.0

Poly δH 7.0

R 7.0

D 10^{-10} cm²/s 25

Amount g/m² 4.0

Thickness µm 10

Time hr 24

Total 100

F1	F2	F3	F4	F5	F6	F7	F8
Ethanol	Ethyl Buty	Isoamyl Ac	Limonene	Eucalyptol	Carvone	Linalool	Citronellal
%1	%2	%3	%4	%5	%6	%7	%8
25	20	15	12	10	8	6	4

https://www.stevenabbott.co.uk/practical-solubility/Fragrance-Diffusion.php

At time t = 0 we have a fragrance formulation (the same as used for [Fragrance_Vapour Pressure]()) trapped within a polymer barrier of known thickness and with known [Dissolution_Hansen Solubility Parameters](). The laws of [Diffusion_Basic Diffusion]() tell us that the rate at which any molecule diffuses through depends on the diffusion coefficient, D which, for simplicity is provided as a single value, D_{100} for a nominal MW

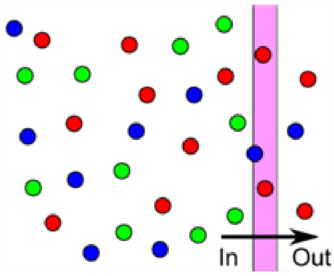
= 100 molecule and where individual values are $D = D_{100}\left(\frac{100}{MW}\right)^n$. In the app, for simplicity n = 1. For many polymers n = 2 and in some cases there's a strong exponential dependence. However, fragrance molecules cover a relatively narrow range of MW values so the value of n is not so important.

More important is that diffusion depends on the concentration gradient, so the higher the concentration (partition coefficient) of each molecule on the fragrance side, the greater the flux of that molecule.

We estimate the partition coefficient based on the HSP Distance – so the closer the HSP values (you can find them in the app's source code) to the selected polymer, the larger the partition coefficient and, for a similar-sized molecule, a faster loss of that molecule.

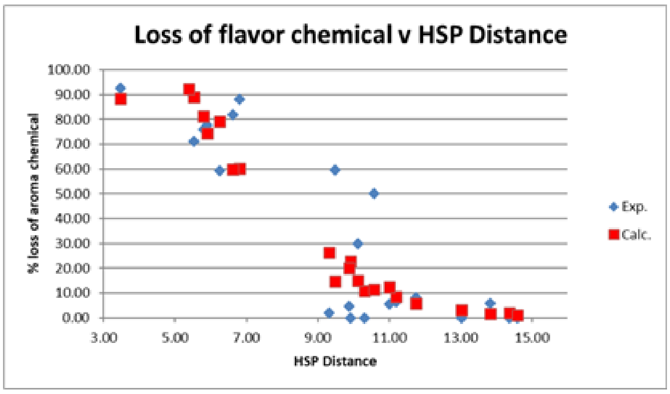
> At a time when PLA was a relatively new polymer I had to predict its fragrance barrier properties in a book chapter. It was no problem predicting (correctly) that it's a good barrier for hydrophobic limonene and hydrophilic acetic acid. I also predicted that PLA packaging would be no good for cinnamon-flavoured products as there's a close HSP match between the HSP of PLA and cinnamaldehyde.

In the example shown, the polymer is more similar to a typical fragrance molecule than it is to ethanol, so after ~20hrs what remains is mostly ethanol.



If we have 3 components where red has a small HSP Distance from the barrier polymer, blue has an intermediate distance and green is very distant, the red components partition strongly and are lost more quickly than the blue. The green which hardly partitions into the barrier is the slowest to be lost. The thickness of the barrier and even the MW effect of the different molecules are *much* less important than these partition effects, which can affect diffusion by factors of 100s or 1000s.

Real world flavour scalping



The image shows an Excel plot of ~20 components of an aqueous drink flavour stored in an alkane polymer. After a modest storage time (days, not the months required for product storage), a GC analysis showed that some components were 90% lost, others in the 70-80% range and some not at all. The effects of this flavour scalping were severe – the flavour was nothing like the original. Experts in food science were baffled.

But from the estimated HSP values of the specific molecules and the known HSP of the alkane polymer it was easy to calculate the expected loss depending on partition coefficient (from HSP Distance) and a correction for the MW. Although the match isn't perfect, it was sufficiently good to confirm the root cause of the problem.

Calculations showed that a swap to a different polymer would have significantly reduced the problem. A multi-layer package might also have worked. But at that stage in the project it was impossible to change the package … so the product had to be abandoned.

The take-home lesson is that a few simple HSP calculations and a basic knowledge of diffusion science can avoid some very expensive mistakes.

Fragrance Vapour Pressure

Links

Deodorant Sticks, Soaps and Washing, Laundry Liquids, SkinCare, Fragrances, Surface Cleaning, Microencapsulation

> Even a simple fragrance might contain 20 components, a high quality perfume might contain 100. However many components, there is still the issue that the fragrance must change over time as the High (or Top) notes (volatile) disappear, the Middle notes (intermediate) take over and finally the Low notes dominate. Using a simple 8-component mix and choosing from a variety of volatilities, you can see how things work out over time.

H, M, L

We seem to like our fragrances to hit us quickly with their High notes, then we like things to settle down to a more stable Medium sensation, while the Low tones linger on as a memory of the fragrance we enjoyed. One of the many complexities of the perfumer's art is to ensure that a fragrance that is changing all the time retains a psychological constancy. How that is done cannot be captured in an app.

Here we look at the key facts of the individual vapour pressures of the pure components, via their Evaporation_Temperature and Antoine Coefficients. Vapour pressure translates into molar concentrations, so we need the MW of each component to know the relative masses in the vapour phase. Then we need the rate at which the components will be removed by air flow. That allows us to work out the instantaneous removal of vapour components, which changes the amount in the fragrance, so that the whole system evolves over time.

Airflow m/s 0.50

T °C 25

Amount g/m² 25

Time min 60

Total 100

F1	F2	F3	F4	F5	F6	F7	F8
Ethanol	Ethyl Buty	Isoamyl Ac	Limonene	Eucalyptol	Carvone	Linalool	Citronellal

%1	%2	%3	%4	%5	%6	%7	%8
25	20	15	12	10	8	6	4

https://www.stevenabbott.co.uk/practical-coatings/Fragrance-Evaporation.php

For illustrative purposes, we have a mix with larger amounts of the volatile components (left) leading to smaller amounts of the least volatile. You see the volatile components disappear rapidly (causing the strange looking discontinuities in the various curves) and the low volatility components relative % increase to the ~40:35:25% of Carvone, Linalool and Citronellal after 60min.

The app contains the Antoine Coefficients, AA, AB, AC of 40+ components that you can choose. The individual vapour pressure is given by:

$$\log_{10} VP = AA - \frac{AB}{AC + T}$$

You can find the values in the source code of the app.

When I first had to write a fragrance VP/HSP/Diffusion program I allowed the users to include 20 fragrance molecules – this seemed far more than anyone might need. My client laughed at my ignorance. “20? 20? A simple fragrance has 40 and more complex ones can easily have 100”. I quickly changed the code to meet their needs.

Air flows and evaporation

The details of how air flow translates into removal of solvent are described in Evaporation_Basics. The key message is that the rate of removal is (approximately) proportional to Airflow^{½}, so a perfume on one’s wrist will evaporate slowly when sitting still (airflow ~ 0.05 m/s) and 4.5x faster when walking with the arm swinging at 1 m/s.

Real formulations

The app assumes that the Fragrance_Activity Coefficients are all equal to 1, i.e. they are an ideal mixture. For a mixture of fragrance molecules, this assumption is not too bad – they cover a modest range of chemical functionalities. But fragrances exist in carriers and are placed into things like soaps, candles, creams … Now the activity coefficients can be higher. Whether the *relative* activity coefficients are much different within a sensible formulation (e.g. one designed for candles will use less hydrophilic molecules than a fragrance designed for a shower gel) seems generally to be unknown.

Those with access to COSMO-RS can obtain relatively accurate activity coefficients at least for the simpler systems and it can be useful for simulating a candle or cream environment.

Gelling Networks and Percolation

Links

Deodorant Sticks

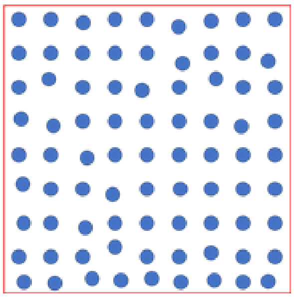
Many gels are formed from a low volume fraction, φ, of self-associating particles. While it's easy to see how a lump of particles might appear at the bottom of the tube Surfactancy_Emulsion Creaming and Flocculation, Separation_Settling and Centrifugation, it's not obvious how the whole tube can be gelled via a "percolated network".

"Particles"

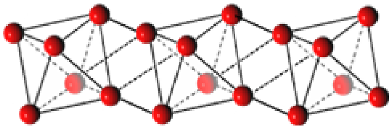
Gelation is a generic phenomenon. When it takes place within a polymeric system (e.g. Thickeners_Associative Thickeners) we don't have much difficulty visualising a tangled polymer network. Here we are interested in generic particles. They can be classic nanoparticles such as silica or nice polystyrene spheres much studied in colloid science, but they can also be semi-liquid such as blobs of emulsion polymers. For the purposes of this chapter, their nature is unimportant, because we want to understand the phenomenon in general. As we shall see, the existence of these gels depends on subtle particle-particle interactions described by virial coefficients so the precise nature of the particle itself (solid or semi-liquid) is less important than the particle-particle interactions.

From ergodic to icosahedral via isostatic taking in percolation

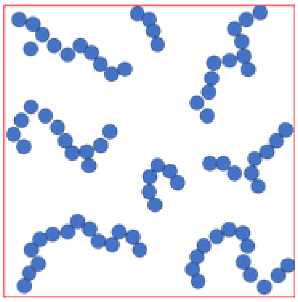
It's sometimes fun to get to use fancy words. But to understand gelling, these 4 words are rather useful, even though unfamiliar to most of us.



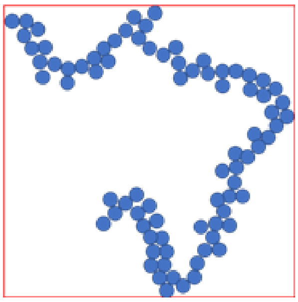
Let's start with "ergodic". This simply means that any particle within our relatively low volume fraction (it's 15% in this and subsequent diagrams) is able to reach any part of the container. Although any given particle might take a long time to explore the whole container, if the system is ergodic it *can* reach any part. When we first mix up our particles, we have created an ergodic fluid. Its viscosity will be a bit higher than the background fluid (see Flow_Particle Viscosity and https://www.stevenabbott.co.uk/practical-rheology/Low-Shear-Particles.php where you find that if φ = 15% the viscosity is doubled.



To have a gel, our particles must be "isostatic" a technical term from Maxwell saying that the net force on any particle is zero – so there is a 3D balance of forces between particles. The way to achieve this is via particles forming an icosahedral network. Given that we don't have enough particles to form such a network filling the whole system, it can only be formed as icosahedral chains as implied in the diagram. This alerts us to the obvious fact that we must have particle-particle interactions to allow such networks to form.



So our particles have to start forming these sub-chains. Here we have the same 15% of particles that have started to create icosahedral sub-chains. [Yes, with x-ray diffraction you can pick up the signal for these icosahedra]. The viscosity will be somewhat higher as these sub-chains might interfere with each other's motion, but it's certainly not a gel. Obviously this setup is not very isostatic because all those free ends aren't experiencing a balanced force. The shift from "viscous" to "gel" appears if we leave the system alone, allowing time for ends to meet, producing long chains.



Finally, the sub-chains have assembled into a *percolated* structure. This term is confusing. As explained in Dispersions_Rheology (High shear) and https://www.stevenabbott.co.uk/practical-coatings/percolation.php, when you go above φ = 28% for spheres (or less for ellipses), there is a (statistical) path of particle-to-particle contacts across the container and viscosity starts to increase. This definition of percolation is respectable and much used. For gelation we need a different definition, equally respectable and much used, which is that there is a path from any particle to any other particle. In both cases, the container is spanned and the viscosity is higher but simple particle systems only achieve the 2nd type of percolation at ~ 75%. The power (for good or bad) of percolated gels is that this 2nd type is achieved at much lower volume fractions – because of the particle-particle interactions and the icosahedral chains.

Why should you care?

It's nice to have fancy words, but do they help us either achieve gelation when we want it, or avoid it when we don't want it?

My personal answer to the question is Yes! I spent a miserable few weeks totally confused by gelation and percolation. Nothing made any sense. Everyone waved their hands but no one provided a formal view … till I found the right technical language and could construct my own diagrams to make sense of it all.

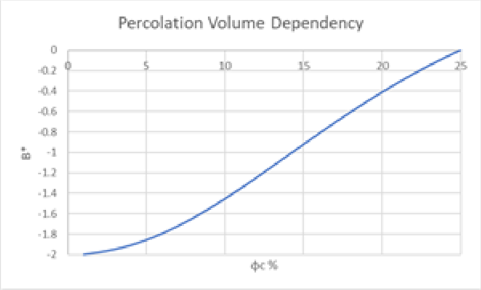
To avoid gelation, the trick, obviously, is to have minimal particle-particle interactions, using a good dispersant (Dispersions_ODC). And to get a good gel, the opposite trick is to avoid large particle-particle interactions that will lead to lumps of stuff flocculating out of the liquid.

Now imagine that you have a nice, un-gelled dispersion for some application. You check it after 1 hour and it's still nice. You come back the next day and it's a gel. It's easy to imagine that something new happened to cause the system to gel. Maybe the pH changed, maybe there's some subtle chemical reaction at the surface of the particle. More probably it's simply that the step-by-step build up of icosahedral chains takes time, and viscosity changes are relatively minor so you assume that nothing is happening. If you check the viscosity, the shear of the measurement might well break up the sub-gel chains so you will always confirm that everything is OK, even though it's not. We'll come back to the timescale of full gel formation once we've explored a little AHS theory.

What % particle do you need for gelation?

There is a respectable theory called Adhesive Hard Sphere (AHS) which assumes that spheres go from zero to strong interaction when they touch – a simplistic idea that works well-enough. The strength of the particle-particle interaction at contact is described by a "reduced virial coefficient", B*, which defines a Baxter temperature, $\tau = \frac{1}{4(1-B*)}$. Negative B* values imply particle-particle attraction. We can see how this attraction affects gelation because the critical percolation threshold, φ_c is related to τ via:

$$\tau = \frac{1-2\varphi_C+19\varphi_C^2}{12\left(1-\varphi_C\right)^2}$$



We can, connecting φ to B* via τ, see the % particle needed for strong percolation. Although most of us won't know our B* values (should you want to do so, you can measure them using light scattering dependency on φ), it's interesting to see that the percolation volume for gelation can readily reach just a few % with adequate particle-particle interactions – which is either good if you want gelation with low additions of particulates, or a problem if you had hoped for a relatively low viscosity particle dispersion below the 28% threshold for the weaker form of percolation.

Another way to think of the percolation threshold is via the mean cluster size, S. If we have a volume fraction, φ, less than φ_C, then $S = \frac{1}{\left(1-\frac{\varphi}{\varphi_C}\right)^2}$. In other words, when $\varphi = \varphi_c$, S is infinite, every particle is connected to every other particle.

Timescale

As particles assemble into larger clusters, their "relaxation times" (a proxy for their size) increase according to a $t^{0.57}$ dependence, a curious value because these are these are fractal clusters rather than pure spheres which would give a classical square root effect, $t^{0.5}$. So to double in size takes somewhat under 4x the time; growth is relatively slow. Once critical clusters have assembled then the network grows in a t^1 timescale, relatively quickly. That's why these gelations can be confusing. Nothing much happens for a long time, then obvious gelation seems to come out of nowhere.

It's even more confusing than that. There is another type of timescale. The beautiful academic work that resolved those timescales took hours of careful measurement in idealised conditions. We generally want our gels to form much faster. If the timescale for forming particle-particle bonds is short then a non-optimal gel will form quickly, without the chance to build up an isostatic network from repeated making/breaking during diffusion of particles and clumps of particles. We might be lucky and the non-optimal gel works fine during the product lifetime. Or we might be unlucky and, when subject to some shock, the gel might collapse.

Why stop now?

Conveniently, our particles assembled into a percolated gel network of those icosahedral strands. Although we've agreed that having particle interactions that are too strong can lead to a flocculated mess, there's no obvious reason why the particles should stop attracting each other. The annoying phenomenon of syneresis (Gelling_Syneresis) is at least partially due to the particles continuing their self-attraction. The non-optimal gels might be especially prone to sudden syneresis. Although syneresis is hard to understand and control, this network chapter provides a good basis for thinking through the issues.

The fact that these gels have no reason to stop is a feature, not a bug. In general we make gels because they are easy to break when we require – a skin cream easily spread or a soft spoonful of yoghurt. It is a feature with downsides, which is why we have the syneresis chapter, but it's still a feature.

Gel dynamics

These gels are fragile in both directions. They can carry on clumping, as in the previous paragraph, or be destroyed by shear. The classic tests are via Yield Stress measurements, Flow_Yield Stress which can be via Flow_Rotational rheology or via Flow_Oscillatory rheology. By going rotational you can also look for Flow_Thixotropy. By going oscillatory you can use Flow_G' and G" at different frequencies to compare the elastic component (G') and the plastic, flow component (G"). The advantage of the frequency sweeps is that you directly get the dynamics of the particle/gel interactions which will presumably be useful for you in use.

That last phrase is a reminder that we're here not to contemplate the theoretical wonders of gels, but to find ways to create them with the desired characteristics. Let's take two gels, say yours and a successful product from a competitor. If preliminary feedback from internal tests is that your gel doesn't, say, "feel" so good, what do you do to change it? Given the complexity of gels, there's no point blindly trying different formulations hoping that one will hit the competitor's sweet spot.

Instead we do the smart mapping, discussed in other parts of FST. Any such smart mapping will include G'/G" measurements at different frequencies. They are easy to do. If, at some reasonable frequency such a 1 Hz your G' is very different from the competitor's that will probably come at no surprise – the gels will behave obviously rather differently if you tilt them in a tube. You can probably imagine how to increase particle-particle interactions if you need to increase G' significantly or decrease them if your gel is too stiff. But you are a good formulator and probably had things about right – your gel was "the same" as the competitor's. That's when the frequency behaviour becomes important. The "feel" of a gel when rubbed with a finger is connected with timescale in the 10s or 100s of Hz. When you find that the competitor's G' is 1/10 or 10x that of yours at some higher frequency, that gives you a clue about what needs to be fixed.

This is where smart mapping starts to help, *assuming you have the habit of measuring these key parameters on all reasonable samples*. You might have tried some different formulation that showed the sort of high frequency G'/G" found in the competitor's product. That formulation will have failed for other reasons, but you now have a technical clue about how to get into that part of formulation space.

Smart digitalization

> Senior corporate people love "digitalization" till they survey the landscape of bad measurements of irrelevant data carried out on under-specified formulations for no reason other than some fad for measuring X for reasons now forgotten. They spend millions on robots that can create 1000s of worthless datapoints … and find that ML, AI etc. cannot create understanding from worthless data.

Smart digitialization requires the team to think of what physics is likely to correlate strongly to the product properties. If there is a direct link from physics to property, life is simple. In complex systems like gels, we know in advance that we don’t have the theories, the apps, the models to provide these direct insights. So we choose a minimax, the minimum number of types of measurements (e.g. frequency sweeps of G' & G") likely to give the maximum insights into how to steer a formulation towards success. Because we can’t know in advance what is “good” and what is “bad”, we need the measurements on a range of known good and bad products, internal and competitors’, to start building up a picture to feed to our AI.

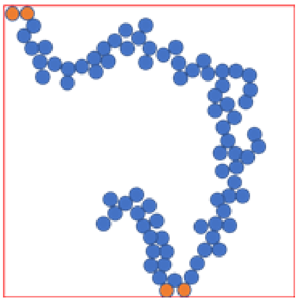
Gelling Syneresis

Links

Deodorant Sticks

You have a nice gel, then over time you see a layer of liquid forming on the surface, or, in a transparent vessel, you see the gel pulling away from the edges, giving a floating blob of denser gel. Sometimes everything seems fine till something (active or passive) causes the gel to detach from the walls. The effect is called syneresis (Greek for "coming together"). It is frustratingly hard to understand and control.

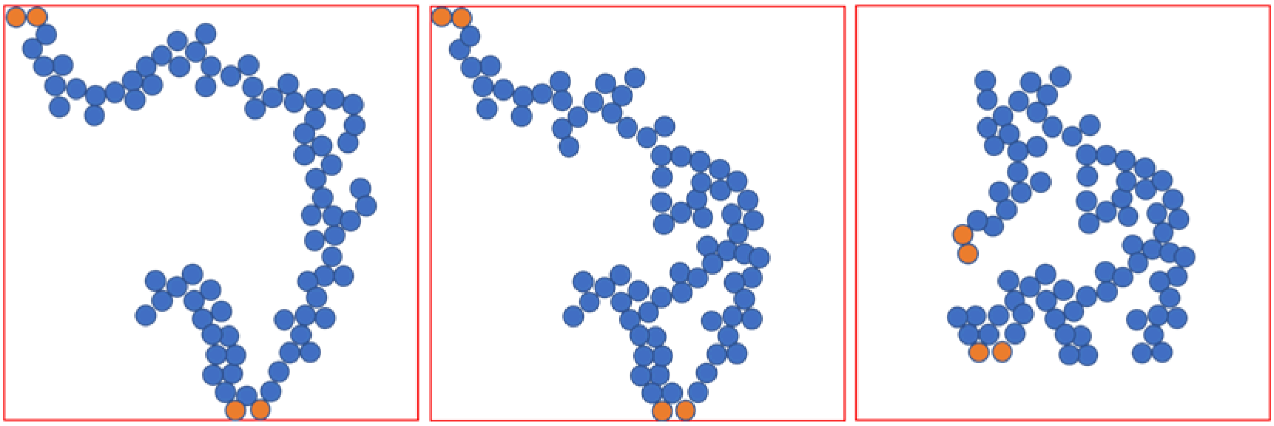
What's a gel?



The sorts of gels that are relevant to syneresis as described here involve remarkably few "particles" (for this chapter we're not too concerned about whether they are solids or semi-solids) that have assembled into a network. In Gelling_Networks and Percolation we saw how we could achieve a volume-spanning fully percolated (all particles have a route to all other particles) with even 1% volume fraction if the particles had sufficient mutual attraction, though the one in the diagram is relatively unsubtle as it has 15%. In this diagram, the particles are contacting the walls at the orange blobs. This may be simple touching or there may be (there commonly is) extra adhesion forces to the walls. The complication of such adhesion is a key aspect to the frustration of understanding and fixing syneresis issues.

That chapter had a warning that too little attraction (obviously) stops formation of a gel, while too much creates a flocculated mess at the bottom of the container. There was also the warning that the particles might want to continue to self-assemble … and the only way to do that is by removing particles from an area with average concentration. We can see this happening in two steps

Syneresis



We start with the original diagram on the left. In the middle we see that the particles have become more concentrated, leaving larger areas of free liquid that may or may not be visible. Whether or not there is adhesion at the wall, the compaction will lead to some elastic stresses along those parts of the chains that extend to the wall. Finally, on the right, the stresses pull in the rest of the chain and our gel is now much more locally concentrated, leaving lots of free liquid, obvious to the eye. This is syneresis. If there had been adhesion at the walls, it might take more build-up of self-association before overcoming the adhesion. For example, in a cylindrical container, suddenly spinning it imposes a shear force at the walls and this might be enough to overcome the adhesion and cause the syneresis to become visible.

For those studying the physics of syneresis, adhesion to the wall is an unwelcome complication. If the walls are treated to inhibit adhesion, then the syneresis can take place more smoothly. By tuning the fluid to be density matched with the particles, gravitational effects are removed. By using a temperature sensitive surface treatment (e.g. pNIPAM which changes its hydrophilic/hydrophobic tendencies sharply at ~32°C) the particle self-attraction can be tuned, allowing reversible syneresis.

Under these ideal circumstances you can spot that unlike the slow $t^{0.57}$ or t^1 behaviour in assembling the gel, syneresis is fast, t^2 behaviour called “ballistic motion” where the mutual self attractions cause an efficient directional motion rather than the more random assembly processes. In the non-ideal circumstances involving wall adhesion, syneresis is fast because of the built-up stresses across the chains. For those who knife cheese curd, this fast response is a feature, not a bug, as it allows rapid separation of the cheese solids.

Gravity

Take a tube with coloured, somewhat dense particles and leave it for a while. The particles settle – and no one is surprised, see Separation_Settling and Centrifugation . Now make the fluid viscous, or increase the fluid density via some additive. The settling is slower. Now start with particles that form a weak gel. It settles. Is that settling or is it syneresis? I would call it settling, but there are plenty of papers where this is called syneresis. Now watch it settle … then suddenly collapse – that probably is syneresis. The literature is confusing because different types of gels are being analysed from different viewpoints. One take on gravitational collapse is that the gel has its own internal strains which are resisted by the network … until gravity causes sufficient movement for local resistance to be overcome locally, triggering a collapse of the whole structure driven not by gravity but by the fact that the particles are in a lower-energy state when they are compacted.

Preventing syneresis

The literature on avoiding syneresis is maddeningly vague. Against vague hypotheses of what is causing syneresis there are vague assertions that the problem is fixed by “increasing osmotic pressure”, “strengthening the network”, “making thicker strands”, “changing the relaxation behaviour”. If the experiments aren’t controlling for adhesion to the walls then a “fix” might be nothing to do with the network and only an increase of wall adhesion. There’s nothing wrong with fixing a specific syneresis problem by increasing wall adhesion. The problem is fixing it via some presumed network effect, unaware that it’s a wall effect.

The term “osmotic pressure” is thrown around in syneresis discussions. In some it means the pressure from mutual attractions of the particles – a terminology that is technically true but rather unhelpful. In others where the system is a complex gel with small pores, osmotic pressure is a stabilising influence if the stuff in the pores (e.g. a polymer) is too large to get out through pore openings. Others use osmotic pressure as some form of magic that causes syneresis to be reduced or increased depending on the form of handwaving being used.

Things are worse with gravitational syneresis. Papers have little problem showing that “strengthening” the network fixes the problem. They might equally say that adding an external viscosity agent or a density modifier is a fix for syneresis. It’s a fix for the problem (good) but adds to the confusion about syneresis (bad). If your gravitational syneresis is not far removed from classic sedimentation that’s OK. If it appears as sudden (unwelcome) event, maybe that’s real syneresis.

How about applying rheology to understanding cause and effect. Again, the results are maddeningly vague. In addition to the confusions over adhesion and gravity, the large variety of rheological measurements makes it impossible to build a coherent picture. You can, for example, measure yield stress by at least 6 techniques (see Flow_Yield Stress), you can measure G' and G" (Flow_G' and G") at different oscillation rates and different maximum strains.

Those making gels out of particles large enough to be seen in a microscope can use the particles’ own motion (microrheology) to measure G'/G" as the gel matures. Not surprisingly, G' increases sharply relative to G" at the gel point. Good to know in terms of gel formation but not directly applicable to syneresis.

My reading of the results is that you can find some hand waving explanation to fit the results, providing no reliable methodology for fixing a similar syneresis problem in the future. A high G' (strong elastic modulus) might be good in some circumstances, yet be an indication of strong interparticle forces that might collapse the whole gel. Too much of anything is bad.

Yes, we should use rheology to probe our gels – it’s a powerful set of tools for answering questions. But we need some proper questions to answer.

Let’s try to ask some clearer questions. This means treating the walls of the container to avoid adhesion of the gel. If this gives instant syneresis then you already know that you either have to rely on wall adhesion in the product or use the fact that your gel is hopeless as a starting point for re-formulation.

As an example of a clear question, let’s use the φ_c formula from the network chapter. We can get a gel at a desirable low value for the critical volume fraction, φ_c, if there is strong particle-particle interactions. A φ_c of 1% is very attractive in terms of efficiency, but that leaves a large amount of empty volume for that small fraction of highly self-attracted particles to collapse into, with little redundant resistance to collapse from its super-thin network. Reducing the particle interaction strength (methods to do this in your own system should be obvious) would lead only to a non-gel, so you have to reduce the strength *and* increase the particle concentration to above the new higher φ_c.

If you *can* create such systems and you then use rather too much rheology to compare and contrast them, you might be able to extract a reliable rheological signal that relates to a reasonable hypothesis around φ_c.

An alternative hypothesis, from the gelling network chapter, is that the balance of timescales for particle motion and particle bond making/breaking is wrong. By definition, a well-balanced system will allow particles to move and to make and rearrange bonds in comparable timescales. If bonds are too weak and/or bondmaking too slow, no gel is formed. If the bonds are strong and formed quickly then the network will be mechanically sub-optimal (not at all isostatic) and will contain a lot of built-in stress, so will be susceptible to breakage under shock. This might happen with “raspberry” particles, those that are relatively rough with high inter-particle friction.

The opposite case is rapid bond creation with weak interactions. Now a gel can form nicely thanks to the rapid movement into an isostatic network. There is little remaining stress so the gel is quite impressive … till a shock is imparted. If the particle-wall interactions are strong then the resulting gel is sub-optimal because it is biased towards the walls. Some wall shear will create a disconnect and the resulting release of stress rapidly causes a “ballistic” collapse of the gel.

There’s one more thing. The discussion so far has assumed that interparticle attractions are a constant. Syneresis will also appear if some change in temperature, pH, chemistry or partitioning leads to a change in these attractions that allows the network to re-distribute itself to a lower-energy state.

Conclusion

If you are exasperated at this point, well, decades of syneresis research have not led to reliable toolkit ideas. The term is too vague, yet each person using it (or avoiding using it) has their own definition which might make their own analyses and recommendations misleading to others.

All you can do is avoid easy, but misleading stories and do what you can to pinpoint why your system is not resilient to the shocks and stresses that make syneresis appear. This means a few things:

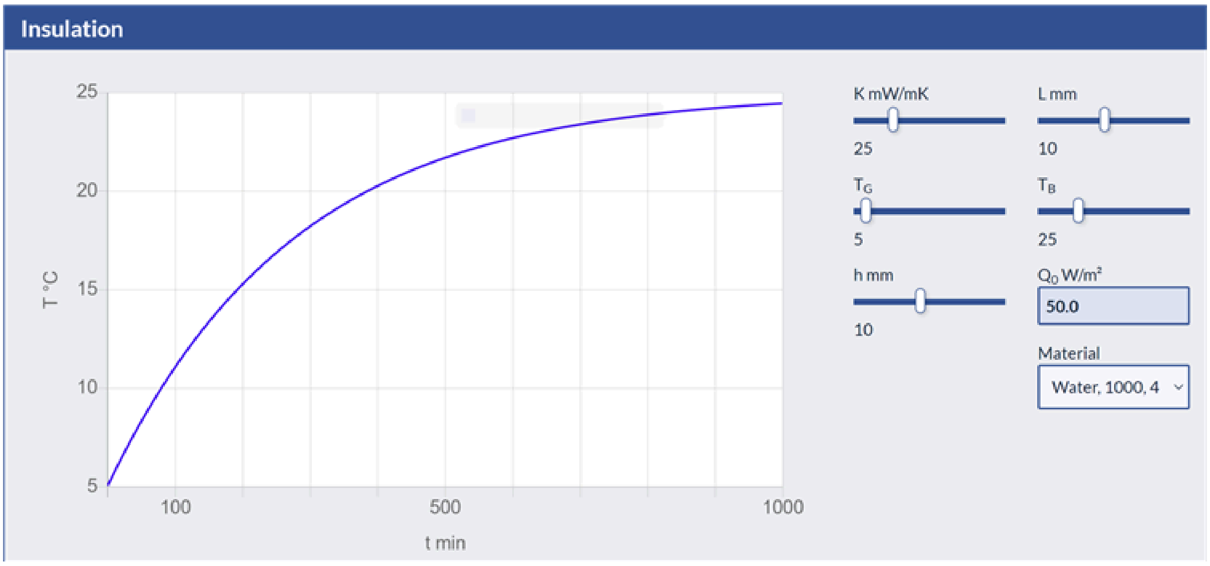
- Being clear about what particle-particle forces are holding things together.
- Estimating whether they are strong or weak.
- Checking on concentration-dependence to see if you are in an unlucky zone just below a zone with enough reliable network connections for a comfortable stable zone, or in a zone where there are too many interactions and the whole assembly collapses in on itself.
- Thinking of likely timescales for bonds making and breaking, on whether there is high “friction” stopping particles from moving or whether they are relatively fluid and can slide past each other.
- Using simple “osmotic” thinking for systems where there is an “inside” and “outside” in your gel and where a component, such as a polymer, can keep a high osmotic pressure inside so the liquid doesn’t want to escape to create syneresis.
- Checking that the syneresis isn’t happening due to some change in the balance of forces due to pH, chemistry etc.

Heat Conductive

Links

Because it is relatively easy to model conductive heat flow through a single layer we first look at an example of insulation. Then we see what happens over more general multiple layer systems.

A simple equation



K mW/mK 25

L mm 10

T_G 5

T_B 25

h mm 10

Q_0 W/m² 50.0

Material Water, 1000, 4

https://www.stevenabbott.co.uk/practical-mechanical/Insulation.php

As a practical example of heat flow and what it entails, we set up an insulation panel of thickness L and thermal conductivity K. We want to preserve some “good” temperature, T_G from some “bad” temperature, T_B, where good might be 20°C in the house and 0°C outside, or 5°C in a “good” fridge and 25°C in the “bad” room temperature, as in the screenshot. The temperature difference is ΔT

The heat, Q, in W/m², that can flow across the insulation is given by:

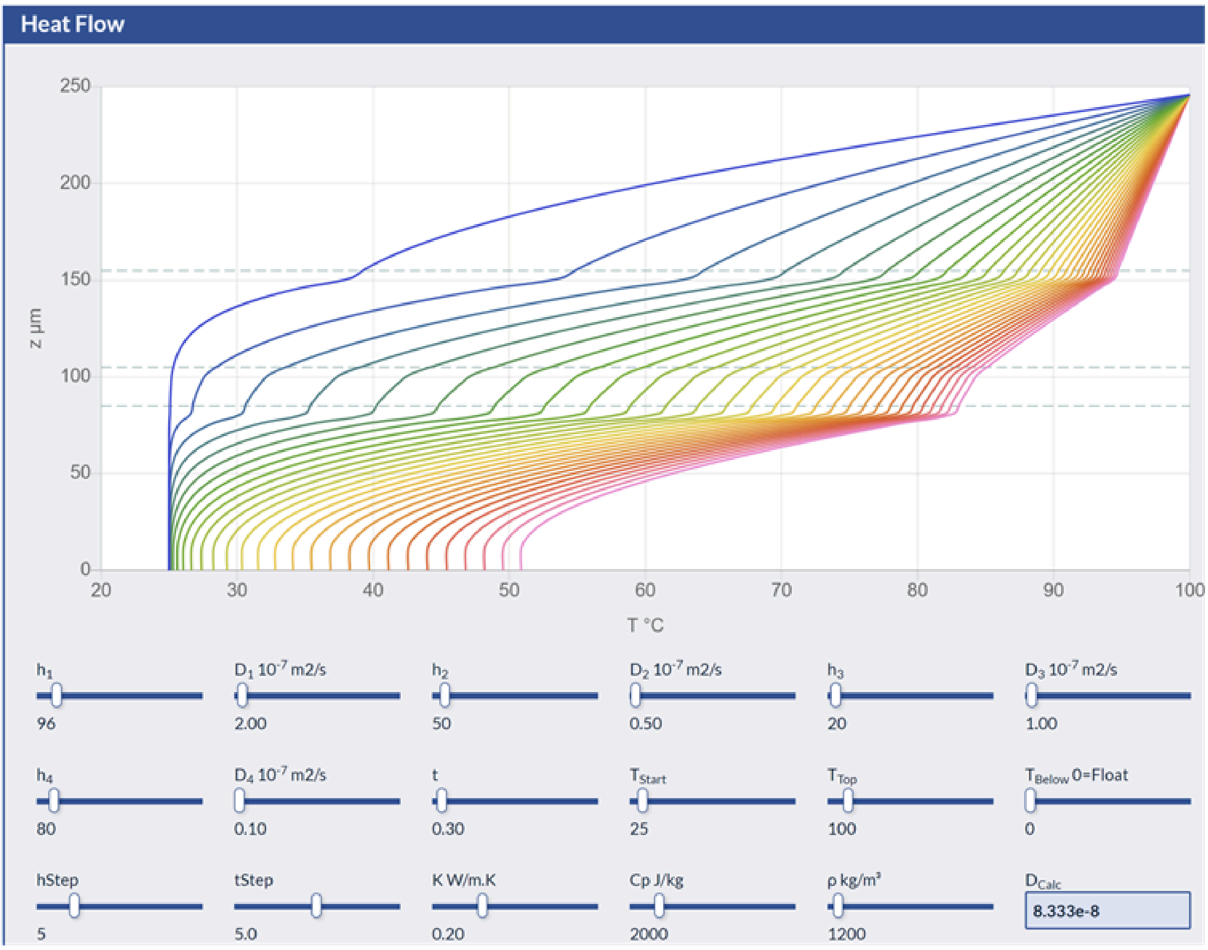
$$Q = \frac{K\Delta T}{L}$$

This heat flux is unhelpful without us knowing where the heat is going. Let’s assume we have a material (in the screenshot it’s water) of thickness h. From its density ρ and heat capacity Cp, we can work out the temperature increase per second, $\frac{\delta T}{\delta t}$, per W of heat.

$$\frac{\delta T}{\delta t} = \frac{Q}{h\rho Cp}$$

As the temperature rises, ΔT decreases, so Q decreases as does δT and the rate of heating is slower for a thicker material, with a larger density and heat capacity.

More complex conductive heat flow



h_1 96

D_1 10^{-7} m2/s 2.00

h_2 50

D_2 10^{-7} m2/s 0.50

h_3 20

D_3 10^{-7} m2/s 1.00

h_4 80

D_4 10^{-7} m2/s 0.10

t 0.30

T_{Start} 25

T_{Top} 100

T_{Below} 0=Float 0

hStep 5

tStep 5.0

K W/m.K 0.20

Cp J/kg 2000

ρ kg/m³ 1200

D_{Calc} 8.333e-8

https://www.stevenabbott.co.uk/practical-mechanical/Heat-Flow.php

As we saw before, the heat flow, Q, across a distance h depends on the temperature difference ΔT and on the thermal conductivity, K:

$$Q = \frac{K\Delta T}{h}$$

In this example we have multiple layers so we have a net Q, Q_{net}, given by the flows in and out of each layer:

$$Q_{net} = \frac{K_{in}\Delta T_{in}}{h_{in}} - \frac{K_{out}\Delta T_{out}}{h_{out}}$$

The rate of temperature change is given by:

$$\frac{\delta T}{\delta t} = \frac{K\Delta T}{h\rho Cp}$$

We can simplify the formula to:

$$\frac{\delta T}{\delta t} = \frac{D\Delta T}{h}$$

where the *thermal diffusivity* is $D = \frac{K}{\rho Cp}$. Although both K and D are commonly known, D saves us having to look up ρ and Cp for the material into which our heat is flowing. It also means that for this 4-layered app we need only 2 values to specify each layer rather than 4. There is a good range of D values in Wikipedia.

For a multilayer problem, we need to have an infinite source of heat in at the top, at a constant T_{Top} and we can choose T_{Below} to be some fixed temperature or (in the app set it to 0 to) make it float as if it was in contact with low conductivity air. Then given the starting temperature and thermal properties of each layer we just step through in time applying the formula for $\frac{\delta T}{\delta t}$.

The real problem is showing the results. The screenshot looks very confusing until you start using the mouse to discover that the graph is rainbow coloured in time – blue is short time, red is long time.

> Many years ago we wanted to have a high temperature next to a thermal head and a low temperature 50 µm away from it. Given that our heat pulses were only in the msec range, we assumed this was easy, but our experimental results showed high temperatures in the 50 µm zone. That's the first time I did a thermal conductivity calculation. The result was so obviously wrong that I spent hours trying to find the bug … till I realised that it was correct. Heat flow can be very anti-intuitive.

Thermal contact resistance

The app assumes that the top layer is heated by a source at constant temperature able to deliver whatever Q the calculation requires. In reality there is always a "thermal contact resistance" between, say, Cu heating plates and the polymer surface. Interestingly this can be calculated using the app in Mechanical_Friction, because friction depends on contact and a contact mechanics calculator is part of that chapter. The "contact" for friction is the same as "contact" for thermal contact resistance.

> I had never needed to bother about contact resistance … till I had to model heat sealing, https://www.stevenabbott.co.uk/practical-mechanical/Heat-Seal.php. I had high-quality data from the literature and my results showed temperature rises far higher than experimentally observed. That's when I learned the need to include thermal contact resistance in such calculations.

Heat Convective

Links

Convective heat transfer is harder to model than conductive and radiative, but the app gives some idea of what is going on for convection in air and in water.

A simple equation

Newton showed us a long time ago that convective cooling in a “still” environment (i.e. we don’t have forced fluid flow) is easy to understand. The rate of heat loss per unit area, Q, for temperatures T_H and T_C is $Q = h\left(T_H - T_c\right)$ or, more compactly, $Q = h\Delta T$. All we need is the heat transfer coefficient h. We know that it’s larger for water than for air. But getting numbers for this intuition is hard.

Rayleigh and Prandtl

Convection involves viscosity, η, heat capacity Cp, thermal conductivity, k, density ρ, thermal expansion coefficient β, thermal diffusivity α and gravity g. Convection also takes place over a length scale L which is, hopefully, obvious to the problem in hand. The parameters each have complicated dimensions, making general formulae tricky. It turns out that calculations are simpler if we use dimensionless numbers, and for convection we need the Rayleigh number, Ra given by:

$$Ra = \frac{\rho\beta\Delta TL^3g}{\eta\alpha}$$

We also need the rather simpler Prandtl number, Pr given by:

$$Pr = \frac{Cp\eta}{\kappa}$$

We then combine these for various convective cases. Assuming gentle convection with laminar flow we have:

Vertical convection:

$$h_V = \frac{k}{L}\left(0.68 + \frac{0.67Ra^{0.25}}{\left(1 + \left(\frac{0.492}{\mathrm{Pr}}\right)^{0.563}\right)^{0.444}}\right)$$

Horizontal convection, Hot Above:

$$h_{HA} = \frac{k}{L}0.54Ra^{0.25}$$

Horizontal convection, Hot Below:

$$h_{HB} = \frac{k}{L} 0.27 Ra^{0.25}$$

App calculation

Because it is far too tedious to find all the required parameters for a specific fluid, and because the calculations are in any case inexact, and L is not always well-defined, and because we should be using parameters defined at the *film temperature* which is some average of the hot surface and cold fluid, the app simply provides general guidance based on parameters defined at "normal" temperatures for either air or water.

Heat Convective

L mm: 50
ΔT°C: 10
Fluid: Water
Mode: Vertical
Ra: 5.29e+5
Pr: 6.97
h W/m²K: 206
Q W/m²: 2063

https://www.stevenabbott.co.uk/practical-mechanical/Heat-Convective.php

As you would expect, Q decreases from Vertical to Horizontal-Above to Horizontal-Below:

Heat Convective

L mm: 50
ΔT°C: 10
Fluid: Water
Mode: H-Above
Ra: 5.29e+5
Pr: 6.97
h W/m²K: 175
Q W/m²: 1748

Heat Convective

L mm: 50
ΔT°C: 10
Fluid: Water
Mode: H-Below
Ra: 5.29e+5
Pr: 6.97
h W/m²K: 87.4
Q W/m²: 874

Putting convective Q into context

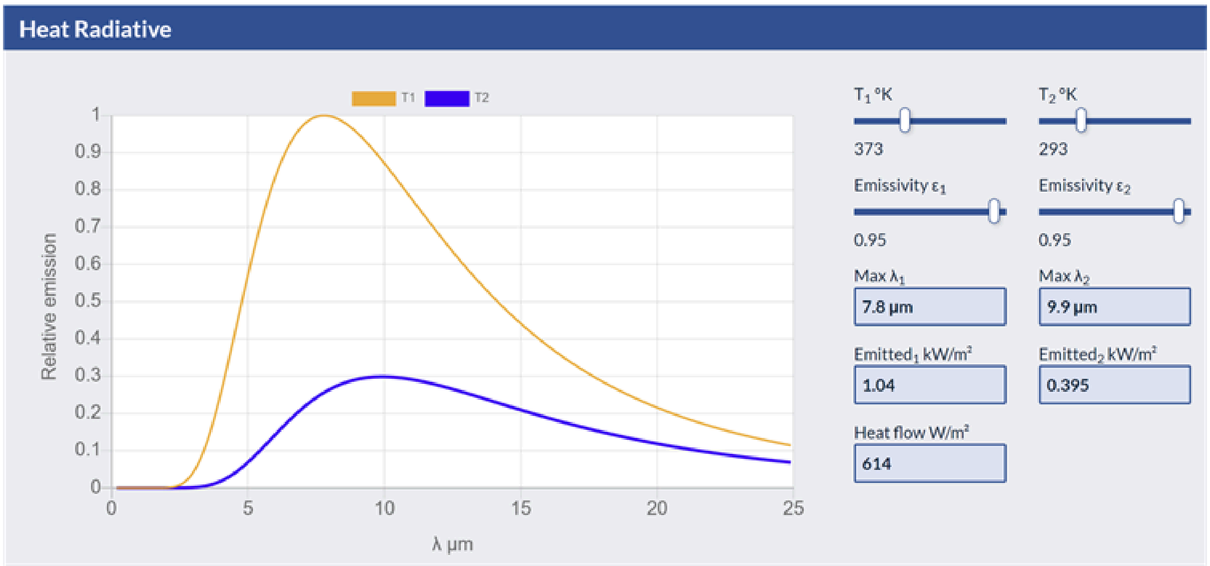
Compared to conductive heat flows, convective flows are small. A good example is in heat sealing, https://www.stevenabbott.co.uk/practical-mechanical/Heat-Seal.php. You can get to, say, 120°C in 0.5s in the clamped jaws, but, as you can find in the app, in 5s of convective cooling the temperature drops only to 90°C. During that long time the softened polymer in the seal is at risk of damage. So you have to provide rapid convective air flow (not covered in these apps) or provide conductive cooling via direct contact.

Heat Radiative

Links

Radiative heat transfer is easy to describe in equations, with the app making it easy to put the equations into practice.

Wavelength-dependent emission



T_1 °K 373

T_2 °K 293

Emissivity ε_1 0.95

Emissivity ε_2 0.95

Max λ_1 7.8 μm

Max λ_2 9.9 μm

Emitted$_1$ kW/m² 1.04

Emitted$_2$ kW/m² 0.395

Heat flow W/m² 614

https://www.stevenabbott.co.uk/practical-mechanical/Heat-Radiative.php

Any object gives off thermal energy over a spread of wavelength, λ. Planck's law describing these emissions is best described in terms of frequency $\nu = \frac{c}{\lambda}$, using the speed of light, c, to do the conversion. At a temperature T °K, invoking Planck's constant h and Boltzmann's constant k we have the intensity, I at frequency ν given by:

$$I_\nu = \varepsilon\left(\frac{2h\nu^3}{c^2}\right)\frac{1}{e^{\frac{h\nu}{kT}} - 1}$$

The emissivity, ε, varies from 1 for a pure radiative black body (which covers most "normal" surfaces) to something like 0.02 for a shiny metal.

We see how they differ for our two chosen temperatures, with the emission normalised to the highest of the two curves. The hotter one is in orange, the colder in blue.

Radiative heat flow

The total flux emitted by the each of the two surfaces is also calculated. They come from Stefan-Boltzmann and invoke Stefan's constant, σ = 5.67e^{-8}.

$$Q = \varepsilon\sigma T^4$$

Because we are often interested in heat flux (usually heat loss) between two (planar) surfaces we can calculate this, taking into account the separate emissivities as:

$$Q = \frac{\sigma\left(T_1^4 - T_2^4\right)}{1/\varepsilon_1 + 1/\varepsilon_2 - 1}$$

This calculation of total flux is done for you in the app. It’s complicated enough, but, as the next paragraph shows, it can get more complex than that.

If, one day, you happen to need to calculate the radiative heat flux between a hot drum and some coffee beans roasting in the drum it turns out to be a bit tricky. Q is W/m², but what is the relevant area when roasting beans? It seemed obvious that the large internal area of the drum would be the relevant value, but the calculated radiant heating was unrealistically high. It turns out that if you have a cylinder inside a cylinder or a sphere inside a sphere, the area used is the smaller one, and that you have to add an emissivity correction based on the relative areas. As the drum is a cylinder, it needed an equivalent cylinder diameter of the beans being roasted. Knowing the mass of the beans and a typical loose density when being roasted, the volume of the virtual cylinder could be calculated, giving results that made sense.

Should you find yourself in need of such a calculation, you can find the formula near the end of the Coffee Drum Roasting app: https://www.stevenabbott.co.uk/practical-mechanical/Coffee-Roasting.php.

I now know some world-class baristas who are passionate about finding the right science to help them brew ever-better coffee. Their initial contact was about espresso crema but since then we’ve worked on many interesting problems, including how to make a cold, non-dairy latte. The coffee roasting app evolved from what seemed to be a minor discussion about a specific phenomenon in roasting (“the flick of doom”) and is now a powerful tool challenging much of coffee roasting convention. It was they who demanded that I add the radiative heat calculations (a difficult challenge!), as the relative importance versus conduction and convection in roasting is a contested issue.

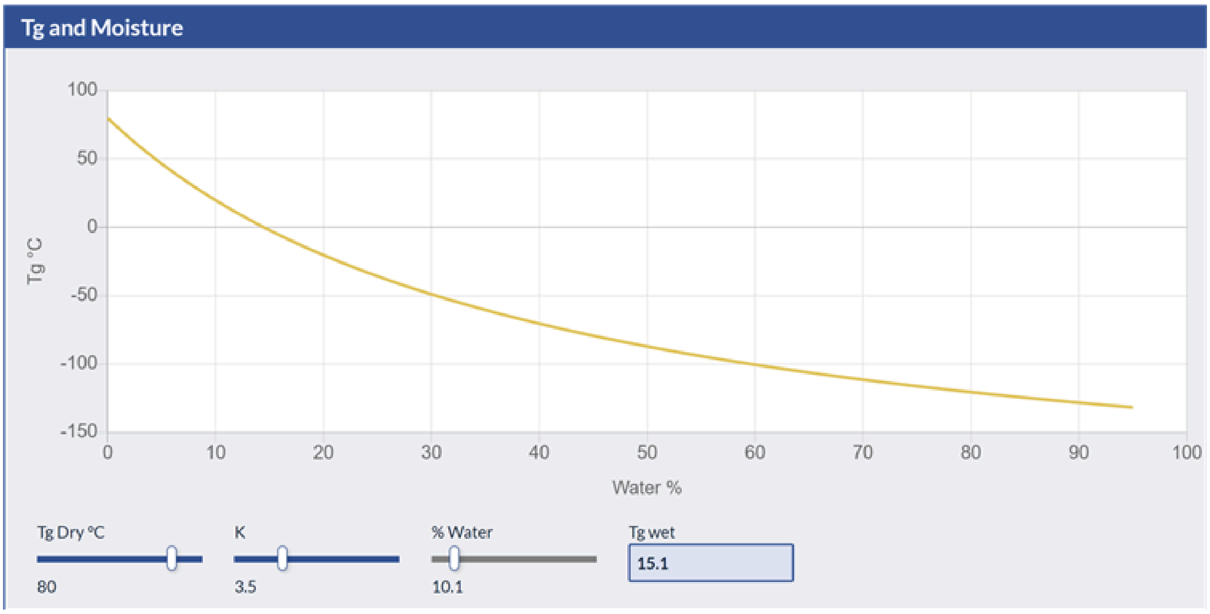
Humidity Water mechanical isotherm

Links

Water-based Adhesive, Emulsion Paint

We can plot how the moisture absorption of a powder depends on the RH (Relative Humidity, or a_W, water activity); that's a sorption isotherm. Less common, but of great importance, is the mechanical isotherm – how the mechanical properties depend on the moisture content.

T_g effects



https://www.stevenabbott.co.uk/practical-mechanical/Tg-W.php

The prime mechanical effect of moisture on the powder material itself is the effect on T_g, the glass transition temperature. The standard equation is a sort of Fox equation (https://www.stevenabbott.co.uk/practical-solubility/polymer-fox-equation.php) taking into account T_{g0}, the T_g at 0% water and T_{gw}, that of 100% water which is -135°C. Where *w* is the fraction of water then, via a constant K which typically has to be measured for each material (but is commonly available as tables of values for foodstuffs) then:

$$T_g = \frac{(1-w)T_g + wKT_{gw}}{(1-w) + wK}$$

What do you do with this knowledge?

Science is universal. I was involved in some intense debates about how best to model the powder clumping tendency of a household cleaning product. It's a complex problem. Eventually we found work (cited in the app) funded by Nestlé covering a similar issue in the processing of food powders such as granules of instant coffee. This approach broke the problem into three steps, each of which could be modelled in an app. Instead of complexity we now had clarity and comparative simplicity.

First you need to know *w* at a given relative humidity, and you can get that from the Isotherm app: Humidity_Water vapor isotherm . Then, by knowing T_g you can work out how likely it is that the powder particles will stick together, and you get that from Particles_Sintering.

This T_g approach to food processing and baking can also be found in the extended work of Louise Slade and Harry Levine.

Humidity Water vapor isotherm

Links

Water-based Adhesive, Emulsion Paint

When a dry sample is exposed to a low water RH (Relative Humidty) (or a_W, water activity), after some time it reaches an equilibrium of water content. Increase the RH and water content goes up. Plot a curve, and that's the adsorption isotherm. Now decrease the humidity in steps, that's the desorption isotherm. Why measure these isotherms? Because the properties of many formulations depend on them.

Fit for purpose

M 0.09

C 10.8

K 0.64

Measured RH 50

% Water 11.1

https://www.stevenabbott.co.uk/practical-mechanical/Sorption.php

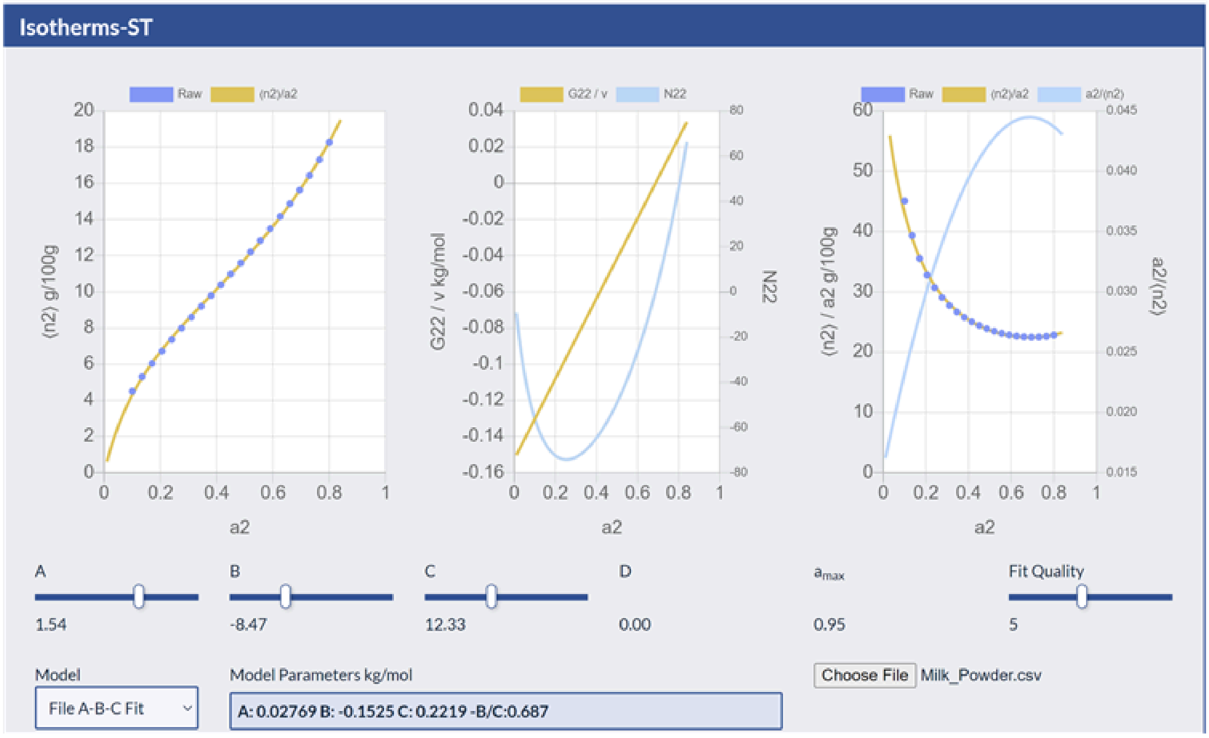
So we have a curve. The original datapoints aren't shown (we shall see them later); instead it has been fitted to some equation which we assume is a good fit. In this case, the equation is the famous GAB model (Guggenheim-Anderson-de Boer) using the three parameters M, C, K. It is supposed to be a refinement of the BET model (Brunauer-Emmett-Teller) but it turns out that each of them makes laughable assumptions about most relevant sorption isotherms. They assume planar surfaces that get covered with a monolayer of water than then subsequent layers of water build up onto that monolayer. The fact that people use GAB for analysing the sorption isotherm of, say, milk powder (that's the curve in the graph), makes you wonder where the monolayers might be.

The analysis via BET or GAB is assumed to give you a "surface area" defined by the monolayer coverage, with other parameters intended to inform you of other, vaguer notions. Again, by noting that GAB is used on milk powder, which definitely doesn't have a planar surface and certainly does some *ab*sorption as well as *ad*sorption, we have to ask why people are using an equation based on such poor assumptions..

What is not under debate is whether GAB gives a good fit to many isotherms. The answer is “Yes” for an interesting reason: GAB is right for the wrong reasons.

For no good reason, the world of Evaporation_Humectants relies on the Norrish equation for fitting the data. In fact, there are more than 80 published isotherm equations, many of which can give the same quality of fit to the data, even though they are supposed to be based on different mechanisms. It’s a mess. Fortunately, we have a way out of the mess.

Same data, correct fit



A 1.54

B -8.47

C 12.33

D 0.00

a_{max} 0.95

Fit Quality 5

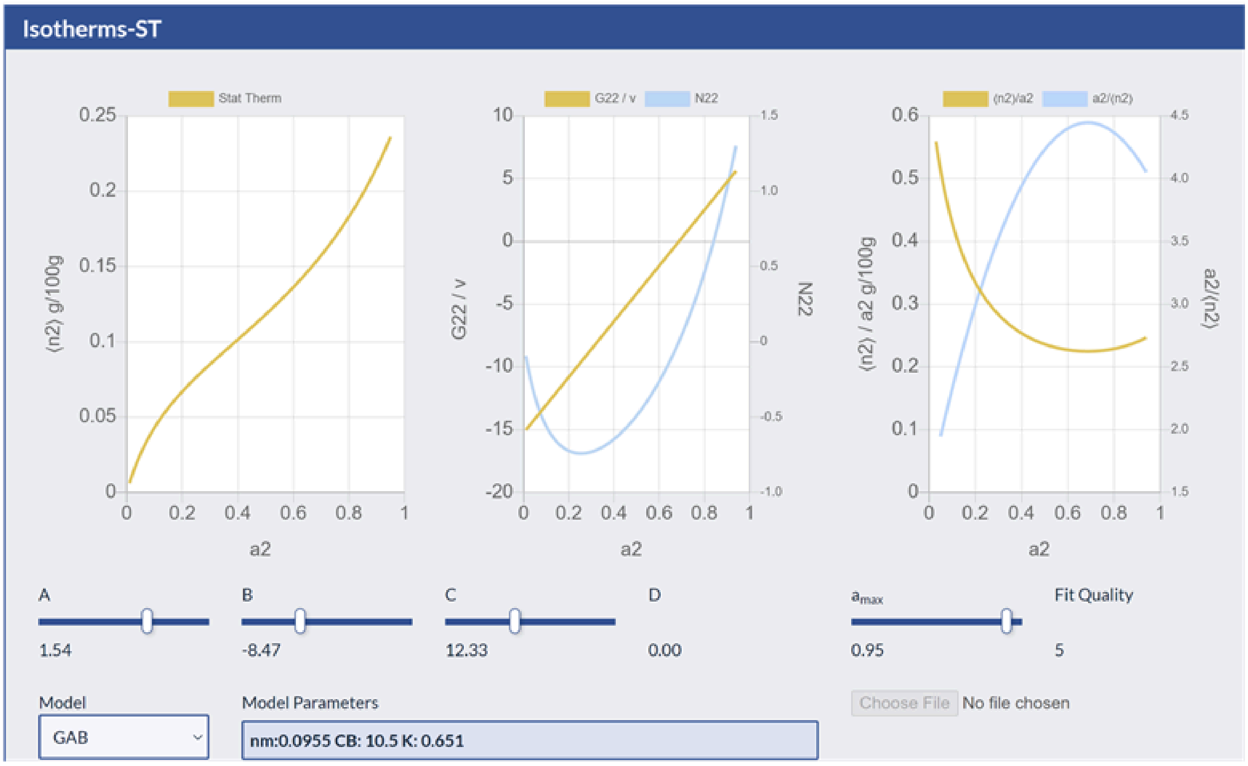
Model: File A-B-C Fit

Model Parameters kg/mol: A: 0.02769 B: -0.1525 C: 0.2219 -B/C:0.687

Milk_Powder.csv

https://www.stevenabbott.co.uk/practical-solubility/Isotherms-ST.php

There are the datapoints, fitted to the ABC model. We can fit the data to GAB and find the parameters used in the previous model, with M called nm, C being called CB and K the same:



A: 1.54

B: -8.47

C: 12.33

D: 0.00

a_{max}: 0.95

Fit Quality: 5

Model: GAB

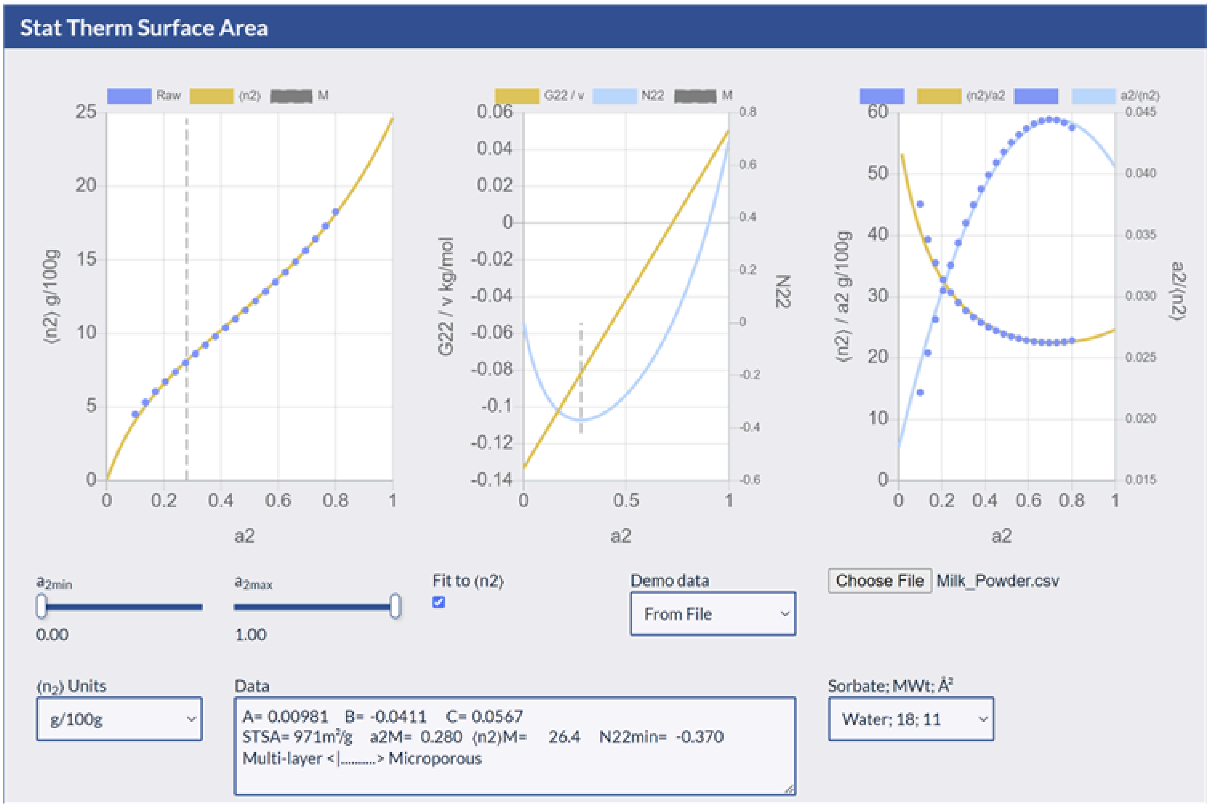
Model Parameters: nm:0.0955 CB: 10.5 K: 0.651

Why is ABC the correct fit? The clue is in the app name – Isotherms-ST. The ST stands for Statistical Thermodynamics which is an assumption-free way of analysing the data. The other two plots reveal key stat therm data *directly from the raw data* so they aren't "fitting", they just say what's going on. The ABC values are, indeed, fitted values but (a) they are provided for convenience and (b) the assumptions behind the fitting are minimal; no absurd ideas of monolayer filling of planar surfaces. ABC doesn't even differentiate between *ab*sorption and *ad*sorption.

The meaning of the ABC parameters is described in the app and in the papers quoted there. If all you want is a nicely-fitted curve for other reasons (which we will come to) then it doesn't matter if you use GAB or ABC, for the very good reason that GAB (and BET) has the same functional form as ABC. We can directly (algebraically, it's not "fitting") translate GAB (or BET) parameters into real ABC parameters. That's how the GAB curve was created in the app.

BET surface area

Those who use the isotherms to work out the surface area of their powder will automatically choose to use "BET surface area" because this is something that can equally be obtained via N2 at 77°K. However, nitrogen at 77°K is arguably somewhat different from water at 25°C so maybe we should enquire about the effective surface area available to water. Here it turns out that for many/most samples typically analysed via BET/GAB, the values are of dubious value. ABC can provide more reliable values, some of which are, under specific circumstances, identical to BET – again for the good reason that BET is a sub-set of ABC. The difference is that ABC provides a measure of how meaningful that surface area is – and it's often not very meaningful. Let's see it with the STSA app, Statistical Thermodynamic Surface Area:



a_{2min} 0.00

a_{2max} 1.00

Fit to (n2) ☑

Demo data: From File

Milk_Powder.csv

(n_2) Units: g/100g

Data

A= 0.00981 B= -0.0411 C= 0.0567
STSA= 971m²/g a2M= 0.280 (n2)M= 26.4 N22min= -0.370
Multi-layer <|.........> Microporous

Sorbate; MWt; Å²: Water; 18; 11

https://www.stevenabbott.co.uk/practical-solubility/STSA.php

We are told that the STSA is 971m²/g. We are also told that this is a bogus value. Directly beneath the STSA is an indicator of where the isotherm lies on the range from “Multilayer”, meaning that the water builds up on itself long before there is anything like monolayer coverage, over to “Microporous” where the water is clearly building up inside micropores, giving no meaning to “surface area”. In this case (as is very common with food/water isotherms) we are firmly in the multi-layer zone, probably because there’s plenty of *ab*sorption of water into the milk powder. So ABC tells us *not* to use the STSA as an indicator of any sort of true surface area.

Would a finer milk powder have given a larger STSA? Probably – there is at least *some* portion of real surface area in that value. But we can confidently say that pretending to know the surface areas via the isotherm is not a good idea.

The stat therm basis of the ABC isotherm also allows cooperative isotherms to be modelled correctly. So instead of 80+ isotherms attempting to capture IUPAC Types I to VI isotherms, there is one core, correct, theory. Will those working on isotherms jump over to the stat therm approach? They should, but they won’t. Scientific inertia is powerful. Now, if you are a keen BET, GAB, Norrish … user, ask yourself if you will change. The assumptions behind them really are nonsense. The stat therm apps are as user-friendly as possible – you can load your old data and see it afresh. So give it a go.

Desorption and hysteresis

It is often the case that the desorption isotherm deviates significantly from adsorption, typically with the sample holding on more strongly to the water. This “hysteresis” can be best analysed via statistical thermodynamics, but at the time of writing, this analysis is not published. As FST is a living book, this section will be updated.

Temperature effects

Again, stat therm allows us fresh insights into temperature effects in isotherms and, again, at the time of writing, this analysis is not published. Again, as FST is a living book, this section will be updated.

Using the isotherm for other purposes

Having a numerical value (via any fitting formula) for the % water absorbed at a given RH allows us to provide at least one answer to the question: "who cares what the isotherm is?"

If we have some milk powder, what we care about is whether the powder will start to self-adhere and become unusable both for the end user (no one likes opening or using a container full of caked milk powder) and for the producer. If a silo containing tons of milk powder is allowed to go to too high an RH, the powder self-sinters (Particles_Sintering) because it has become too weak to resist the pressure inside the silo. That's why we have a parallel chapter, Humidity_Water mechanical isotherm, linked to the GAB-based app that opens this chapter. Naturally, the mechanical isotherm app is itself linked to the sintering one.

Mechanical Friction

Links

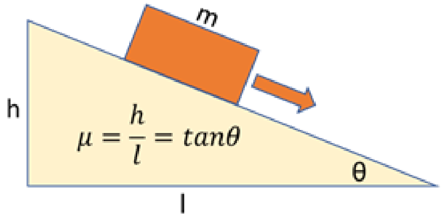
Friction seems so simple but frustratingly it can only be measured, not predicted from first principles. This is because the true contact area between two surfaces, a key aspect of the frictional force, is unknown.

The force, F needed to overcome the friction of an object mass m (times gravitational force, g) on a surface is given by:

$$F = \mu mg$$

Where μ is the friction coefficient. This can be the *static* coefficient, needed to get an object moving or the *kinetic* (or *dynamic*) coefficient to keep it moving, which is usually lower because less force is needed to overcome sliding friction. It seems non-intuitive that the contact area does not appear in the equation. The reason is that whatever causes the frictional force depends on the contact area (so doubling the area would double the frictional force) and on the mass per unit area, which halves as the contact area doubles. The effects cancel out – at least for "standard" friction. We see later how the assumptions can break down.

Although μ can be measured using some force sensor dragging a sample along a surface, this isn't so useful for measurements in real-world situations.



Instead, you have a relevant weight (its mass or contact area don't matter) sitting on your surface and tip the surface till the weight starts to slide. By measuring the height, h and length, l, you can calculate tanθ = h/l and this *is* the static friction coefficient!

$$\mu = \frac{h}{l} = tan\theta$$

It is interesting to take a block with length l, width w and height h, each very different from each other, and repeat the experiment with different sides in contact. The measured tanθ remains constant however you arrange the block.

This simple friction coefficient test is surprisingly little known. I've used it many times to solve problems where friction coefficients were required but everyone thought you needed a "real" machine, which they didn't have. The first time I used it was to measure the friction coefficient of cheese on different packaging films. Yes, it's an important value for those who are frustrated by the difficulty of sliding a block of cheese back into a packet.

The μ found at onset of motion is the *static* friction coefficient. If you slightly reduce h till the mass stops moving, you can find a slightly lower h (and therefore tanθ) and immediately get the smaller *kinetic* coefficient. You can stop and start as many times as you wish to check that your h values are reasonably constant, giving you reliable μ values.

Friction coefficients

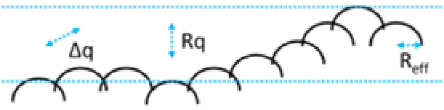
If friction was about numbers of atomic interactions across the interface, then a perfectly flat surface, such as mica, should have a friction coefficient = 1, which it does. If forces between surfaces can be spread/ dissipated over a larger area then the assumptions behind friction theory break down, so rubbers start off at μ = 1 and can be higher than that. If there is some difficulty in getting atoms to interact because they are dissimilar then friction coefficients should be lower; indeed, take metals A and B: $\mu_{A\text{-}A}$ and $\mu_{B\text{-}B}$ are usually higher than $\mu_{A\text{-}B}$. Because nothing much likes to interact with perfluoro surfaces, friction coefficients onto Teflon are often < 0.1.

If the surface molecules can fall apart under the frictional forces then μ will be lower. That's why waxes work well and, even better, why liquid layers act as lubricants – for as long as they stay trapped between the surfaces rather than being squeezed out.

That is a lot of hand waving. If you look at large tables of A-B friction coefficients, it is hard to come up with any more refined understanding, especially as quoted values can themselves be variable.

Contact mechanics

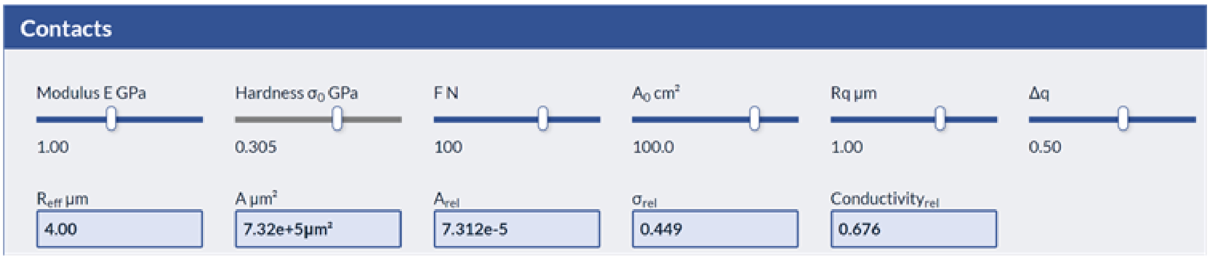
As discussed in Mechanical_Surface Roughness, you should never measure just the *amplitude* of roughness, you should equally measure its *wavelength.*



Strangely, any surface can be made equivalent to one where every element has an effective radius, R_{eff}, and the RMS wavelength can be measured as Δq and the RMS amplitude as Rq. There is a relationship between the three properties, which is useful if your measurement device gives you only two of them:

$$\Delta q = \sqrt{\frac{Rq}{Reff}}$$

From this we can calculate for a given "real" area, A_0, the actual contact area, A and from that the relative contact area A_{Rel}. It is quite shocking to see how little of the surface is really in contact:



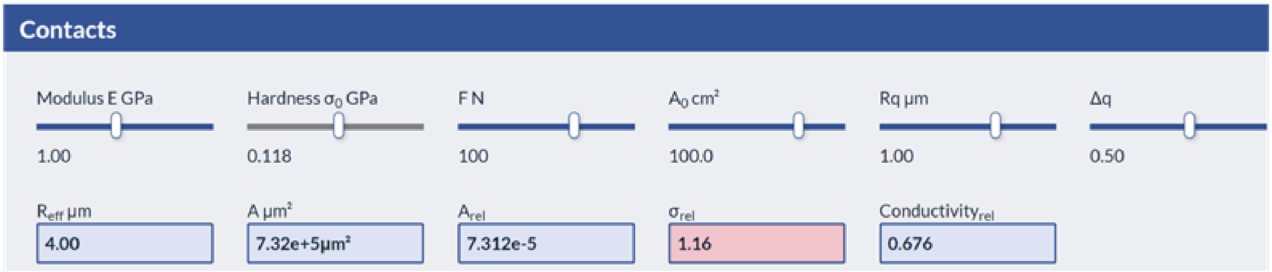
Modulus E GPa 1.00
Hardness σ_0 GPa 0.305
F N 100
A_0 cm² 100.0
Rq µm 1.00
Δq 0.50
R_{eff} µm 4.00
A µm² 7.32e+5µm²
A_{rel} 7.312e-5
σ_{rel} 0.449
$Conductivity_{rel}$ 0.676

https://www.stevenabbott.co.uk/practical-mechanical/Contacts.php

In this example, for a material with a modulus of 1 GPa, the actual contact area for a nominal 100 cm^2 area is $7.3e^{-3}$ cm^2, i.e. less than 0.01%.

This area is given by $A = \frac{2F}{E^{*}\Delta q}$ where E* is a modulus correct for the Poisson ratio, assumed to be 0.3.

The effective load rather than being $\sigma = \frac{F}{A_0}$ becomes $\sigma = \frac{F}{A}$ so can easily exceed the hardness of the material. [Hardness is not the same as modulus: see Mechanical_Hardness.] By reducing the hardness from 0.305 GPa in the example to 0.118, we find that the hardness is exceeded ($\sigma_{rel} = \frac{\sigma}{\sigma_0}$) and now we get plastic deformation, meaning that assumptions about friction coefficients are no longer valid.



Modulus E GPa 1.00
Hardness σ_0 GPa 0.118
F N 100
A_0 cm² 100.0
Rq µm 1.00
Δq 0.50
R_{eff} µm 4.00
A µm² 7.32e+5µm²
A_{rel} 7.312e-5
σ_{rel} 1.16
$\text{Conductivity}_{rel}$ 0.676

As mentioned in Heat_Conductive the same calculation allows you to estimate the electrical and/or thermal contact resistance, calculated as relative Conductivity in the app.

Kinetic friction, stick slip and squeaks

It always take some time for a system to change from static to kinetic friction and back again. A surface will have some natural waviness, and the device pulling the object isn't perfectly rigidly coupled. Take this all together and you get the famous stick-slip condition and squeaks from things like brakes. This is all calculable, though it needs some heroic degree of knowledge of your system, as the app explains.

μ_{Static} 0.30
$\mu_{Kinetic}$ 0.20
Load kg 10.0
t s 60
c kg/s² 1.0
η kg/s 1.00
τ s 1.00
D m 10.0
V_0 m/s 0.50
V_a 0.0
V_b 0.0
Resonance 0.05Hz, 0.32rad/s

https://www.stevenabbott.co.uk/practical-mechanical/Friction.php

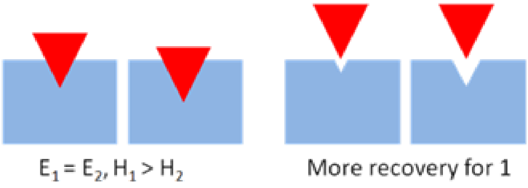
Mechanical Hardness

Links

PSA

Hardness is resistance to permanent marking so is an important parameter. We know that a coating has a Mechanical_Modulus defined as *elastic* resistance to small stresses. Hardness measures resistance to *plastic* stress, i.e. permanent flow under, say, a scratching load.

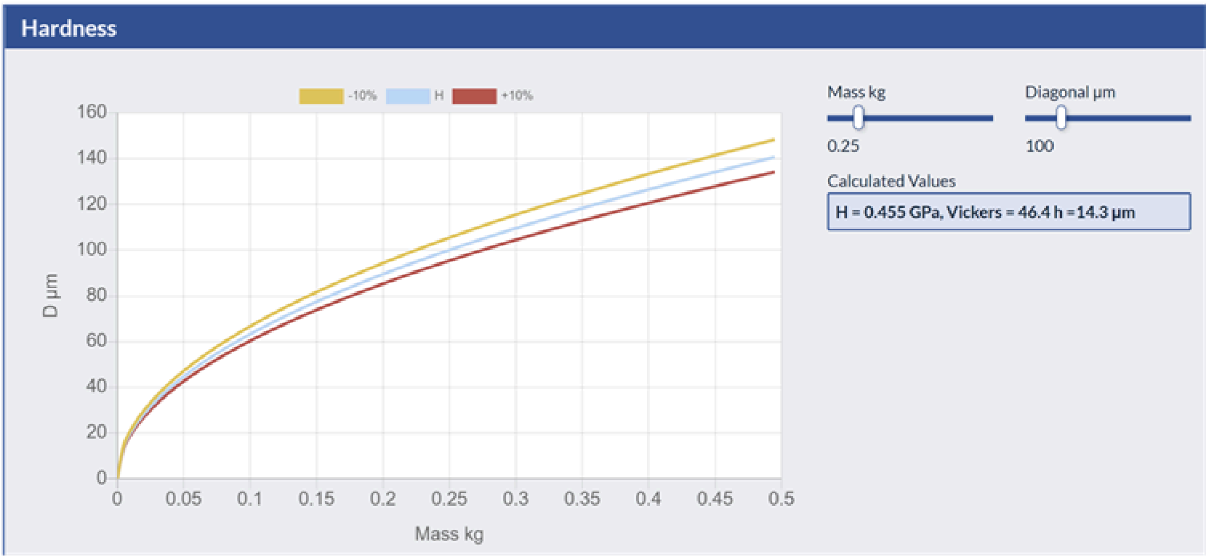
You can buy hardness measurement machines, but a simple device lets you measure hardness anywhere, e.g. inside an oven if you want to know how hardness changes with temperature.



Hardness is a very important parameter that many of us are unfamiliar with. This is unfortunate because it is of importance in many situations where we might be tempted to use the more familiar "modulus". The definition of hardness seems to make little sense:

$$Hardness = \frac{Force}{Contact\ Area}$$

However, a simple example helps make it clear. The same tip is pushed with the same force into materials with the same modulus E_1=E_2 but with different hardnesses $H_1 > H_2$. The tip enters further into 2 and on withdrawal it is clear that the tip has left a bigger permanent impression on the substrate - i.e. there has been more plastic flow and less elastic recovery. Finally, we can see that the contact area with the tip is higher for 2 than for 1, so Force/Contact_Area is less for 2 than 1, so the Hardness is lower for 2. It's as simple as that!



Mass kg 0.25

Diagonal µm 100

Calculated Values

H = 0.455 GPa, Vickers = 46.4 h =14.3 µm

https://www.stevenabbott.co.uk/practical-coatings/Hardness.php

One way to measure hardness is by pushing a Vickers square pyramidal diamond tip into the surface with a known load, F. You measure the diagonal width, D, of the mark you've created and calculate the hardness via:

$$H = 1.85\frac{F}{D^2}$$

Depending on whether F is in N and D is in m or F is in kg and D is in mm, you get H in units of GPa or Vickers units. The app provides both and also tells you the depth of the indentation.

The graph gives you a feel for how accurate your values might be given your chosen load and your likely errors in the measurement of D.

The problem with Vickers Hardness machines is that they are best suited for seriously hard materials such as metals. The formulation world tends to avoid measuring hardness because the official alternative to Vickers seems to be a nanoindenter, described shortly. And such machines can't conveniently be used in, say, an oven or at a customer's facility.

In my experience there are two "good enough" methods for measuring hardness of typical polymers and coatings, usable in a wide variety of formulation setups.

1. Attach 3 Vickers tips to a plate and, with a suitable added weight place the device on your surface, wait ~2 minutes then measure the 3 indents after ~1 minute. Those waits allow time for (a) plastic deformation and (b) elastic recovery. You might choose different times, but these are an acceptable starting point.
2. Instead of Vickers tips use spherical ball bearings. Measure the radius of the indent and instead of the Vickers equation use a Brinell sphere test.

The official "Brinell Hardness Number" BHN for a ball of diameter D giving a circular impression of diameter d, is given by the following, along with the real hardness, H:

$$BHN = \frac{2F}{\pi D\left(D - \sqrt{D^2 - d^2}\right)} \text{ or } H = \frac{4F}{\pi d^2}$$

The beauty of these methods is that with 3 tips or balls you have automatic balance in applying your load, you have 3 datapoints for statistics, and you can make your loads and devices to whatever is needed to give quick, reliable values. You don't even need to do accurate calculations – you can have a QC test saying "In spec is an impression greater than X and less than Y". The first time I used a Brinell tester I had never heard of Brinell. I just found a plate, 3 ball bearings and a mass that reliably gave us 3 nice indents for material we knew to be in-spec, and where some deliberately out-of-spec samples (too hard and too soft) showed up clear differences. When we developed a different product, we simply created a different plate with different ball bearings and mass, giving us a similarly quick and easy QC test.

Nanoindenters

A fancy version of a hardness tester is a nanoindenter. This gives lots more information, but is more expensive and delicate, though you can do other things with a nanoindenter than just measure hardness.

E GPa 2

H GPa 0.5

P_{max} mN 10

v 0.3

E_{calc} GPa 1.82

Elastic % 54.51

δ_{max} nm 1907

Plastic δ nm 1040

H_{calc} GPa 0.11

https://www.stevenabbott.co.uk/practical-coatings/nanoindentation.php

The nanoindenter gives a full test of elastic and plastic effects and their relevant parameters. A sharp (Berkovich) tip is pushed into the coating with an increasing force and the displacement is measured. The force is decreased and the displacement continues to be monitored.

The app is backwards - you input the modulus, E, and hardness, H, (along with a maximum load P_{max} and a Poisson ratio, ν, typically 0.3) and the nanoindenter trace is simulated, and the extra parameters such as elastic recovery are calculated. The idea is to give a feeling for what a real-world experiment might look like.

As before, the calculated hardness is based on the definition of:

$$Hardness = \frac{Force}{Contact\ Area}$$

Obviously the assumptions behind the creation of the nanoindenter curve are only approximations. The aim is to build up your intuitions of how things change, not to do hardness calculations, which your nanoindenter will happily perform for you, once you can get hold of one.

The calculation assumes a diamond indenter of modulus 1141GPa, ν=0.07. The contact area is calculated using standard Berkovich parameters.

It was obvious how to make a hard hardcoat – increase the crosslink density and add lots of nanoparticles. We could increase *modulus* (see the neat trick Mechanical_Modulus for measuring modulus of a thin coating) but we couldn't increase *hardness*, resistance to scratches. A chance came up to try out a (then new and exciting) nanoindenter. The expert running the machine could see that we chemists didn't understand the difference between modulus and hardness. "You know, don't you, that hardness is force divided by contact area?" We admitted that we didn't know this, and that the definition made no sense. It was a lightbulb moment. Most of our samples had a high elastic modulus, but the better ones showed less plastic deformation, so contact area was smaller and they were harder. Changing our focus to plastic flow led directly to new nanoparticle formulations where the dispersant reacted into the matrix, to reduce flow around the particles. For years we’d heard the word “hardness” and confused it with “modulus”. Cross-discipline interactions are vital for progress.

Mechanical Modulus

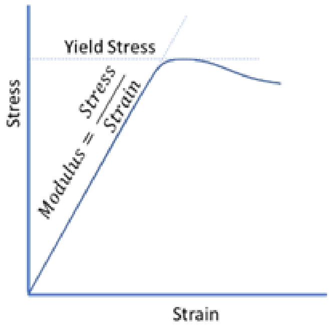
Links

PSA

Mechanical Hardness is resistance to permanent marking, via *plastic* deformation so is an important parameter. The resistance to *elastic* deformation is an essential material property called modulus, defined as Stress/Strain where Stress is the force you apply and Strain is the amount the sample stretches. It's good to know both properties as a measure of the "strength" of your materials and coatings.

Classic modulus measurements of bulk materials use large samples clamped into Instrons. For coatings, the modulus can be easily measured if you can make a thin sample on some sort of release coating and do a simple droop measurement!

Classic modulus

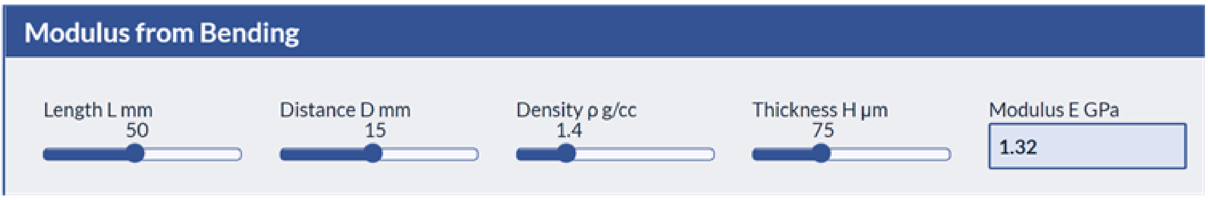


You clamp a sample into the jaws of an Instron-like machine and apply a *stress* (force per unit area) and measure the strain (% change in length). For a while the stress is proportional to strain and the gradient, Stress/Strain is the modulus. Stress is in Pa, strain is unitless, so modulus is in Pa, more usually MPa or GPa.

Eventually the material starts to deform, the graph becomes non-linear and at the maximum stress ("yield stress") the sample yields.

Modulus from Bending

To those of us who don't have a DMA, it is difficult to measure the modulus of a thin sample directly. This trick using the bending of a thin sample is super-useful.


Modulus from Bending

Length L mm 50

Distance D mm 15

Density ρ g/cc 1.4

Thickness H μm 75

Modulus E GPa 1.32

https://www.stevenabbott.co.uk/practical-coatings/modulus.php

Take a thin strip of your coating, thickness H. Put it onto the edge of a sample holder and extend it by length L. Under its own weight it droops by distance D. If you know its density ρ then you can calculate its modulus, E:

$$E = \frac{1.3L^4\rho g}{H^2D}$$

It’s as simple as that!

Getting a good sample of known H and finding a way to measure D accurately is tricky at first, but your organisation can quickly find ways to do it routinely.

I’ve used this trick many times on a large variety of samples, from the hardcoats (modulus ~ 2 GPa) mentioned in Mechanical_Hardness to inkjet films in the 10 MPa range. For obscure reasons, we needed the inkjet modulus at various temperatures so we just did the test in freezers, fridges and ovens. Not super-accurate, but the results were fit-for-purpose, which is good enough.

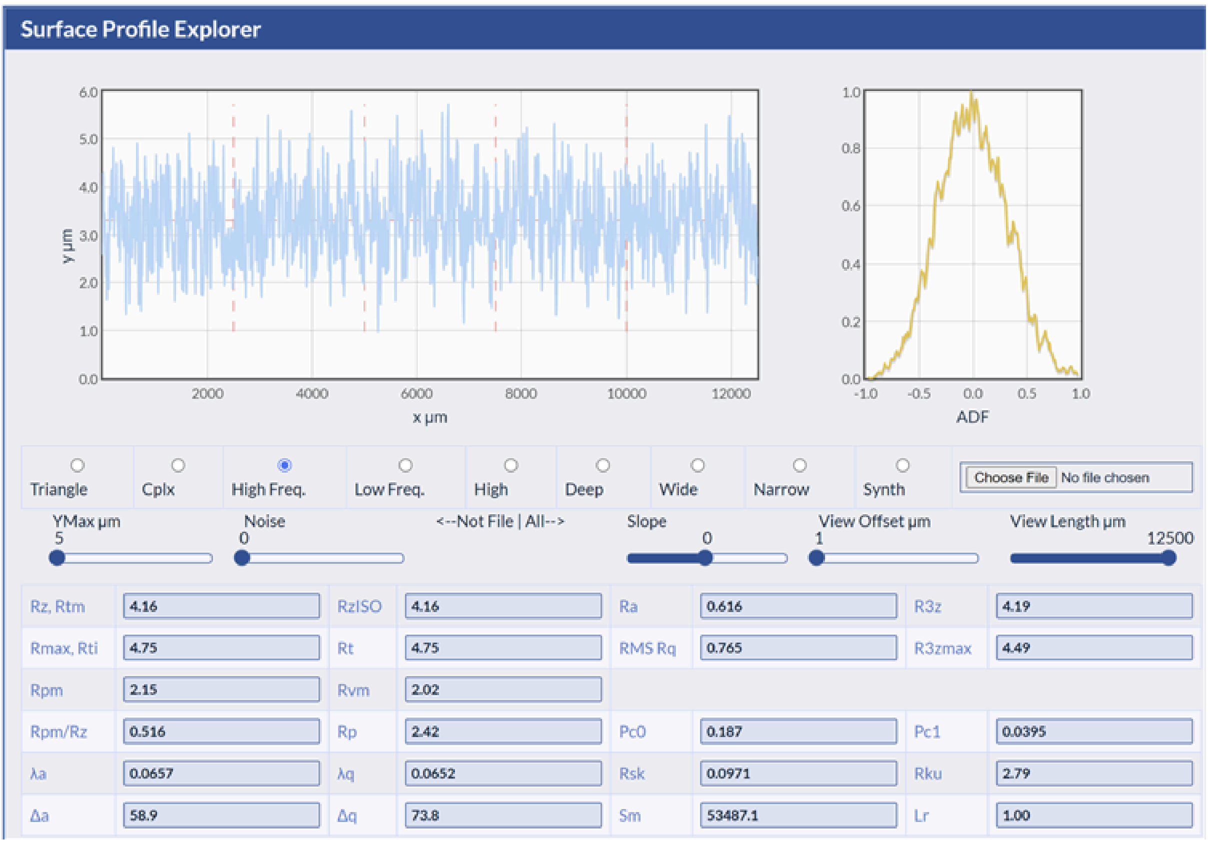
Mechanical Surface Roughness

Links

Adhesive

Surface roughness is important for the look, “feel” and behaviour of a surface. Modern measuring devices can give you many parameters, but for many surface effects only two, Rq and λq, the first an *amplitude* and the second a *wavelength* are really important.

With the aid of the app, which lets you explore different surfaces, you will quickly understand what the important roughness parameters mean.



○ Triangle ○ Cplx ● High Freq. ○ Low Freq. ○ High ○ Deep ○ Wide ○ Narrow ○ Synth

YMax µm 5

Noise 0

Slope 0

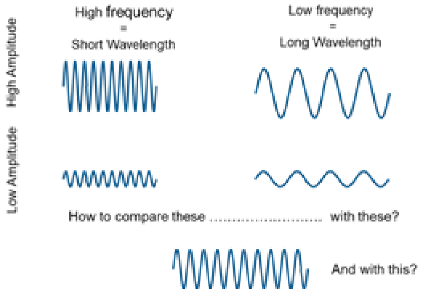
View Offset µm 1

View Length µm 12500

Rz, Rtm	4.16	RzISO	4.16	Ra	0.616	R3z	4.19
Rmax, Rti	4.75	Rt	4.75	RMS Rq	0.765	R3zmax	4.49
Rpm	2.15	Rvm	2.02				
Rpm/Rz	0.516	Rp	2.42	Pc0	0.187	Pc1	0.0395
λa	0.0657	λq	0.0652	Rsk	0.0971	Rku	2.79
Δa	58.9	Δq	73.8	Sm	53487.1	Lr	1.00

https://www.stevenabbott.co.uk/practical-coatings/surface-profile-explorer.php

Surface roughness is important for Optics_Gloss, Absorbency_Porosity and Darcy, Mechanical_Friction and the “feel” of a surface.



Most people are familiar with roughness *amplitude* as measured in "R" values like Rz, Ra or Rq. But this is only part of the story. Look at these surfaces. Two pairs have the same amplitude but are obviously very different. The difference is in the *frequency* or, more usually, *wavelength*. So whenever you measure a surface you *must* measure both amplitude ("R" values) and wavelength ("λ") values. Only then can you fully understand what is going on. It is extraordinary that most people extract only amplitude values from their surface profile devices … partly because many devices don't provide the wavelength values easily or at all. The strong advice is to never buy a surface profile machine that cannot provide wavelength data.

The origin of the surface roughness app was desperation. A supplier delivered a new batch of a product with an in-spec surface roughness but which obviously failed even a casual visual inspection. I wrote a surface analyzer program to process the raw digital data and added all the common parameters, including the λ values we'd never heard of. The R values showed no difference between products, the λ values were very different. That was a long time ago. The Surface Profile Explorer has been online for many years. But *still* people are unaware of the need for both R and λ values.

The app lets you compare/contrast many different types of surfaces and get values calculated for the standard R, λ and Δ measurements surface machines can provide. There is a description on the app page of what each parameter means. Modern methods tend to get areal roughnesses rather than linear values, but the learning from the linear app carries over to the areal values.

You should set a rule that when you measure a surface, *always* get a digital readout of the surface scan – don't just rely on the numbers produced by the machine. Sometimes those numbers are wrong because of a defect or problem that is obvious to the human eye, but not to the computer doing the calculations. Necessarily, this means in addition to not buying a machine without wavelength data, don't buy a machine that can't provide you with the raw data.

Again, this is a rule based on desperation. We were looking for differences between some other surfaces and the standard meters weren't giving useful information. We'd just acquired a digital version of the meter and could now look at the raw data. It was instantly obvious that some large glitches, unrelated to our issue, were distorting the measurements. Removing the glitches allowed us to get the data we needed to solve our problem. The human eye is often much smarter than any digital analysis.

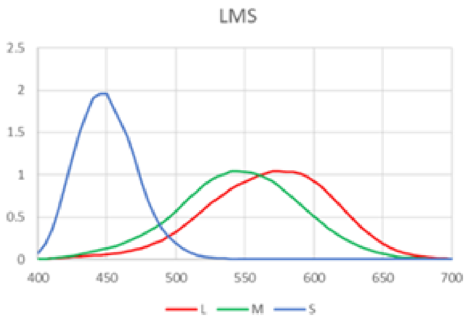
Optics Colour

Links

SkinCare, Lipstick,

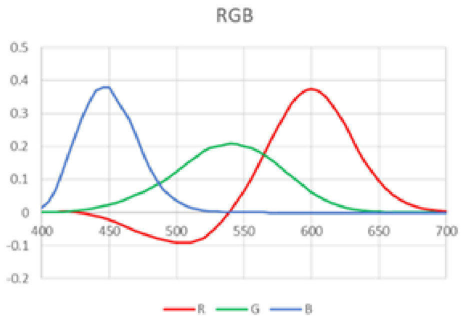
Colour is much more complicated than we might like. The objective measurement of colour is hard enough, interpreting it is also hard. Then the subjective element is also hard – a perceived colour depends on the colours around it. Here we provide some essentials.

What we see



The graph shows that our eyes have 3 colour cones that have response curves to light over 3 wavelength ranges, Long, Medium and Short. Why nature has selected such an odd, unequal set of ranges is unclear [there are conflicting theories, none of which I find convincing] – but it makes colour science more difficult. (*Note that imperfections in the plot and those that follow are from the standard datasets*):

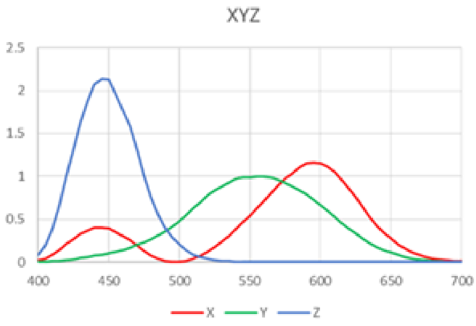
RGB



Humans prefer to think in terms of Red, Green and Blue, but this means that the RGB elements, include some weird negative values obtained by human colour-matching tests that involved back-projection of some red to give the negative values. These negative values are one source of colour science complexities.

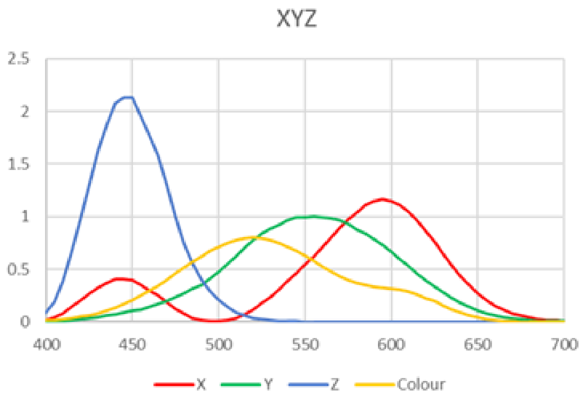
Because of these complexities, we are forced to use XYZ colours.

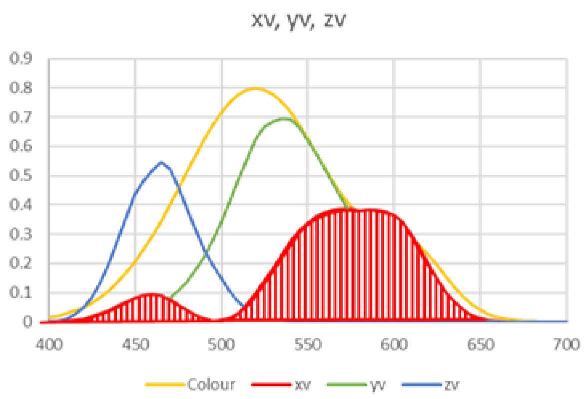
XYZ



The letters X, Y and Z are deliberately chosen to have no obvious meaning. They are just a mathematical re-normalization of RGB. We are even further from real colours, but XYZ coordinates allow us to specify any human-visible colour accurately, which is why they are used. The arithmetic behind them is not especially hard but giving the formulae here won’t help my purpose which is to take you through the chain of logic which takes us ever further from those photons hitting the receptors in our eyes. The next step looks at a real colour with absorption across the spectrum.

A real colour in XYZ





The real colour has some absorbance over the visible wavelength range. We can integrate the amount of light in the X, Y and Z ranges by multiplying the X, Y and Z curves with the colour curve. This gives us three values: xv, yv and zv. To quote the colour we add the three values to give Tot and provide the colour triplet [X=xv/Tot, Y=yv/Tot, Z=zv/Tot].

RGB and CMYK

Although we know that describing additive colours via Red, Green and Blue cannot capture the full colours of human vision and that the subtractive colours, Cyan, Magenta and Yellow suffer from the RGB problem plus the limitations of CMY pigments, we still have to work with them. [See the Kubelka-Munk app, https://www.stevenabbott.co.uk/practical-coatings/Kubelka-Munk.php for some hints of issues with imperfect pigments]. Adding in BlacK to reduce the amount of C, M & Y for grey levels adds a further complication.

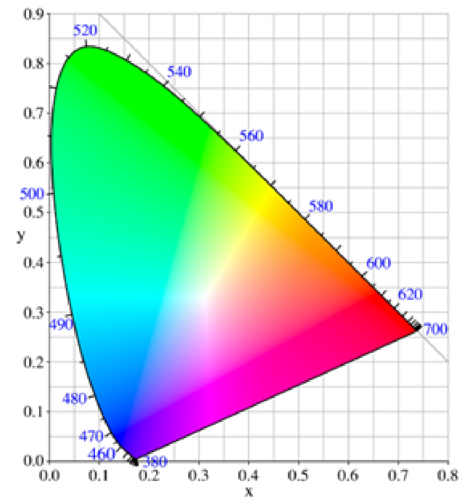
Many years ago, a colleague and I decided to spend a few hours learning colour theory. How hard could RGB and CMYK be? Decades later, and now, even with the resources of modern technology, I still find that I need a few days of full immersion in the science before I can handle a colour science issue. Everyone agrees that colour science is really, really difficult and definitely cannot be picked up in a couple of hours.

How white is white?

The amount of light detected by our colour sensor across the wavelength range depends on the colour of the white light used to illuminate it. D50 light is relatively yellow, with a colour temperature (“black body temperature”) of 5000°. D65 light is whiter, with a hotter colour temperature of 6500°.

From now on, we assume that your measurements are made with your choice of D50 or D65. It doesn’t matter which, as long as you specify it. It’s easier if you then stay in one mode or the other, but it’s possible to interconvert if, say, you’ve done everything in D50 and you need to communicate with someone who does everything in D65.

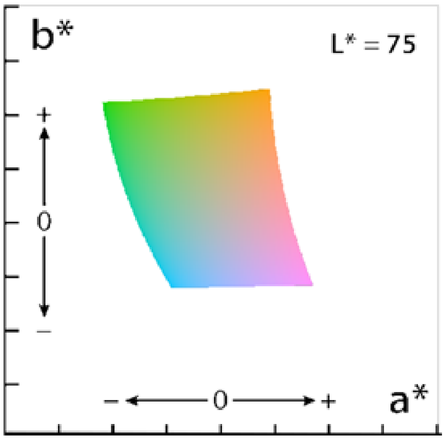
CIE



[X,Y,Z] triplets are perfect for conveying colour values, but useless for us understanding them. So there is yet another conversion formula to take XYZ into something we can work with. We create $x = \frac{X}{X+Y+Z}$ and $y = \frac{Y}{X+Y+Z}$ and by assuming that the "luminance" of the colour is simply Y we can create the famous 2D CIE colour space. Colour space images are courtesy of Wikipedia:

What a wonderful diagram, but also how useless. Look at how much of the colour space is taken up by subtle differences between greens, and how the blues are crammed into one corner. It is amazing that such a useless diagram is seen as so fundamental to colour science.

L*a*b*



So there is yet another conversion of CIE space into a more equal space using L (luminance) and, again, meaningless letters a and b. The * designation is an historical feature. This one is shown with a luminence of 75. You need a different chart for any other luminance. So although we all use Lab (or a specific version of Lab, there are many variants), and although it is useful for colour measurements, it's not at all useful for humans to be able to choose specific desired colours. Some say that a* is red-green and b* is blue-green, but this is unhelpful. Lab is just another attempt to do the impossible – compact the complexities of colour into a simple set of numbers. However, Lab is really useful for colour differences.

Colour differences

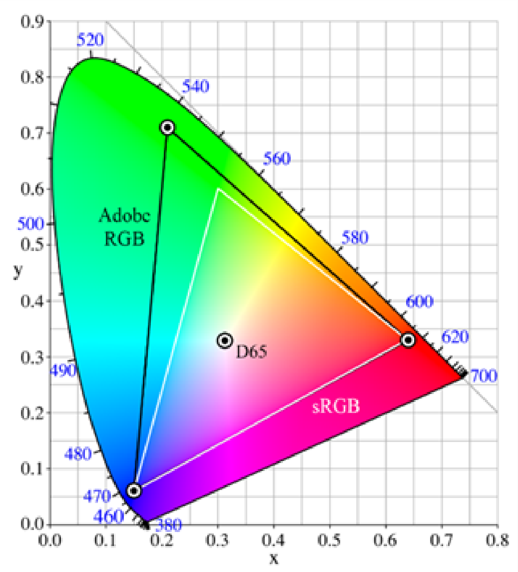
If you have two colours with values L_1, a_1, b_1 and L_2, a_2, b_2 then you can calculate the colour difference, ΔE via

$$\Delta E = \sqrt{\left((L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2\right)}$$

The convention is that colours differing by less than ~2.5 are not distinguishable by eye.

But, of course, there are slightly different definitions of ΔE depending on different Lab spaces. So you might choose to quote ΔE_{00} rather than ΔE_{ab}.

Gamuts

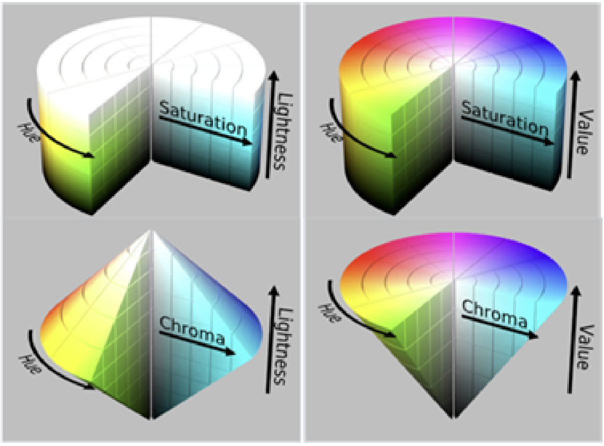


By definition, CIE and L*a*b* cover the range of human vision. But our computer monitors and printing inks are limited by the spectrum of the phosphors (emissive displays) and absorbers (subtractive colours) so we cannot reproduce them all. We therefore need to be aware of the colour gamut, the range that we *can* reproduce. There are a number of standard gamuts, including sRGB and Adobe RGB.

Note that this is defined with D65 white. A typical gamut for our laptops is sRGB. From the graph it looks very bad. Surely we need something like Adobe RGB! But because of the distortions of the CIE plot, sRGB is good enough for most of us. And because we are so used to sRGB, we don't notice that some shades of, say, green, can't be produced.

In any case, we are getting more and more detached from real colours. When you take a photo on your smartphone, the digital processing makes the sky bluer, the grass greener and the red apple redder – because we prefer reality to look that way.

HSV and HSL



If we have to *choose* a colour it is hard to use CIE or L*a*b*. So via yet another transformation we can create HSV (Hue, Saturation, Value) or HSL (Hue, Saturation, Lightness) spaces and select colours via some pseudo-3D chooser:

We can plot the values as cylinders: HLS (left) and HLV (right). Or as conical chroma plots.

If any of these plots was wonderful then we would all use it. But they each have their compromises, so we generally choose whichever is the default in our graphics software.

What is the real colour we see?

As is well known from many optical illusions, from the famous dress, and has been observed with joy or frustration by artists, illustrators and amateur home decorators, you can measure an objective colour with your equipment, confirm that it is *this* sort of blue but to most people it looks *that* sort of blue.

Similarly, if you are looking at, say, an apple as you walk from bright sunlight into a room with artificial lighting, the colour stays constant to your eyes, yet the XYZ coordinates will have changed significantly.

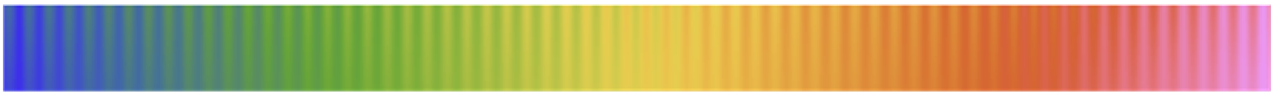
The human vision system is *not* a camera with an objective measure of the light falling onto pixels. It is a complex system designed to extract maximum relevant information for minimum expenditure of processing power. We know that the colour of the apple is constant, so of course the colour doesn't change as we move between different lighting conditions. Our heuristics for judging a colour surrounded by other colours are good for most of what we do, and if they get fooled by clever optical illusion setups, that's fine.

Rainbow colorization

It is common to colorize graphics (I do so in a number of my apps) and common to just use a standard rainbow like this, from https://agilescientific.com/blog/2017/12/14/no-more-rainbows:

As described on that page, this is a truly awful method – it has big perceptual blind spots, it over-emphasises some areas, it's bad for people with colour blindness, and it fails to reproduce in greyscale.

Having learned of the evils of rainbow colorization I went to https://colorcet.com/gallery.html and the super-useful datasets provided in many formats, and chose this "least worst rainbow":



If you use any of my color-coded apps, you'll hopefully find that rainbow sufficiently familiar, yet also providing a better range of information.

Optics Gloss

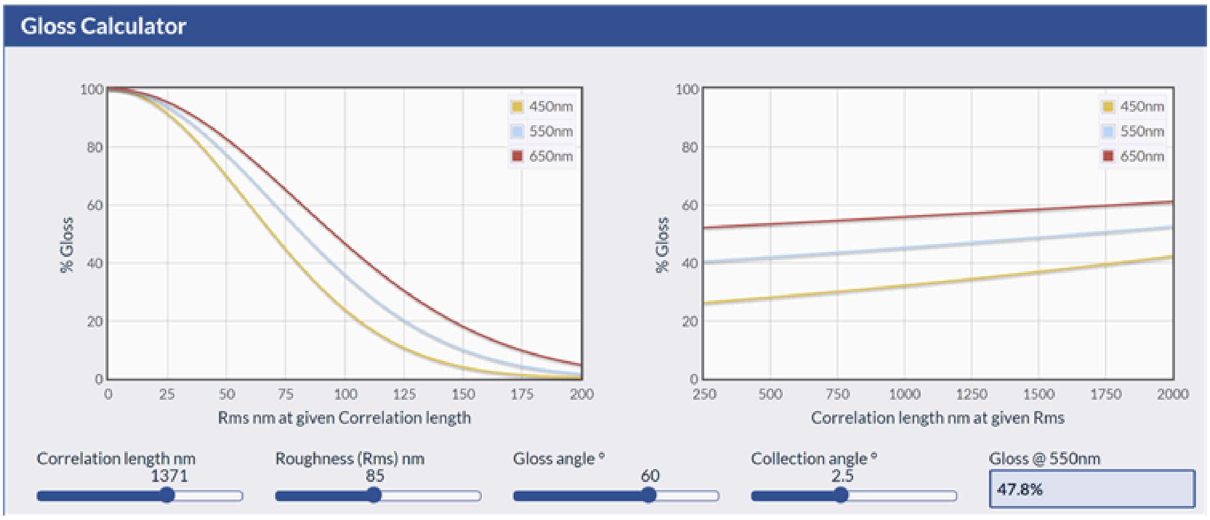
Links

Solvent-Based Paint, Emulsion PaintLipstick

Controlled gloss or matt finishes are super important for coatings. The values are measured with a gloss meter, at different angles. These values depend on two properties of the surface:

1. The roughness *amplitude;*
2. The roughness *wavelength.*

The important thing is to have a surface roughness measuring device that can give you *both* values. Only then can you begin to understand how to control gloss properly.



Correlation length nm 1371

Roughness (Rms) nm 85

Gloss angle ° 60

Collection angle ° 2.5

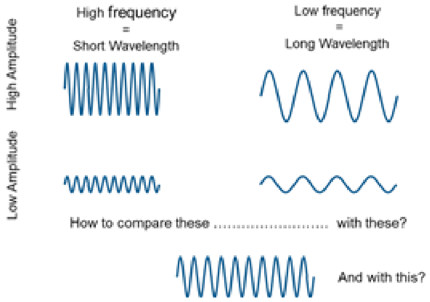
Gloss @ 550nm 47.8%

https://www.stevenabbott.co.uk/practical-coatings/gloss.php

We all know that the gloss of a coating or surface depends on the roughness of the surface (it also depends on the scattering from the bulk, see Optics_Scattering and Opacity.) Frustratingly, it seems hard to match roughness directly to gloss. This is for two reasons.

1. There are numerous “R” measures of "roughness" amplitude, see Mechanical_Surface Roughness, so it is not obvious which is most likely to correlate to gloss.
2. Another key factor which is often not measured turns out to be crucial: the frequency, or wavelength or (scientifically) the correlation length of the roughness.

The diagram shows how two surfaces with the same roughness *amplitude* will be different because they have different *wavelengths*:



The first issue can be addressed easily – out of all the different amplitude ("R") measures, gloss correlates strongly with Rms, the root-mean-square roughness, sometimes called Rq. Note that this is different from (and generally higher than) the common Ra value and usually much less than the equally common Rz. You *can* work with any convenient R value as they are often strongly correlated, but it is safer to use Rms as your working value.

The effect of the wavelength of the roughness is easily explained. Suppose (absurdly) that a given measured roughness undulates over 10s of µm. At any given point the deviation of the slope of the undulation from zero is small, so the light is reflected as if the undulation were not there. So a large wavelength or small frequency gives essentially no reduction in gloss. This is because the gloss meter measures light over a small but significant range of angles of reflection, typically 2.5°, the "collection angle". You can alter this in the app if you wish, but leave at 2.5° if you don't have any other data. If, conversely, the same roughness occurs over a nm range then the slopes are very steep so the angular deviation (scattering) is large.

The basic calculation takes a roughness and a correlation length and computes the gloss value if measured at 550nmm in the middle of "white light". The graphs show:

- How changing Rms at your given correlation length changes gloss at 3 different wavelengths;
- How changing correlation length at your given Rms changes gloss at those 3 wavelengths.

Hopefully, by understanding how the two parameters interact you will be able to find rational strategies for reaching and controlling your desired level of gloss or matt.

One question remains. How do you get a correlation length from your surface roughness scan? In the absence of a proper correlation function analysis, just chose the λa value , again see Mechanical_Surface Roughness or, even better, the λq value which is based on Rms.

It is sadly the case that many surface roughness devices just output a few "standard" numbers, often neglecting anything other than Ra and Rz and being generally unaware of frequency/wavelength issues. If you are seriously interested in surface optics, you cannot work without a proper digital analysis of the data in terms of Rms, something like λq and, even better, some sort of Fourier analysis from which you can derive the proper correlation length, or the exponential fit to correlation length which is the value used in this app.

Accurate calculations of gloss are highly complex. Fortunately there is a simplified equation:

For light of wavelength=λ, surface of Rms=σ, with a measurement collection angle=Δ and a gloss angle=θ the gloss (assuming an incoming refractive index of 1) is given by:

$$Gloss = exp\left[-16\left[\frac{\pi\sigma}{\lambda}\right]^2 cos^2\theta\left\{1-2\left\{\frac{G(a)}{\lambda}\right\}sin\Delta cos\theta\right\}\right]$$

G(a) is a function depending on the correlation length, a, given by:

$$G(a) = \left(\frac{2}{\Delta}\right)\mathrm{atan}(\Delta a)$$

Optics Scattering and Opacity

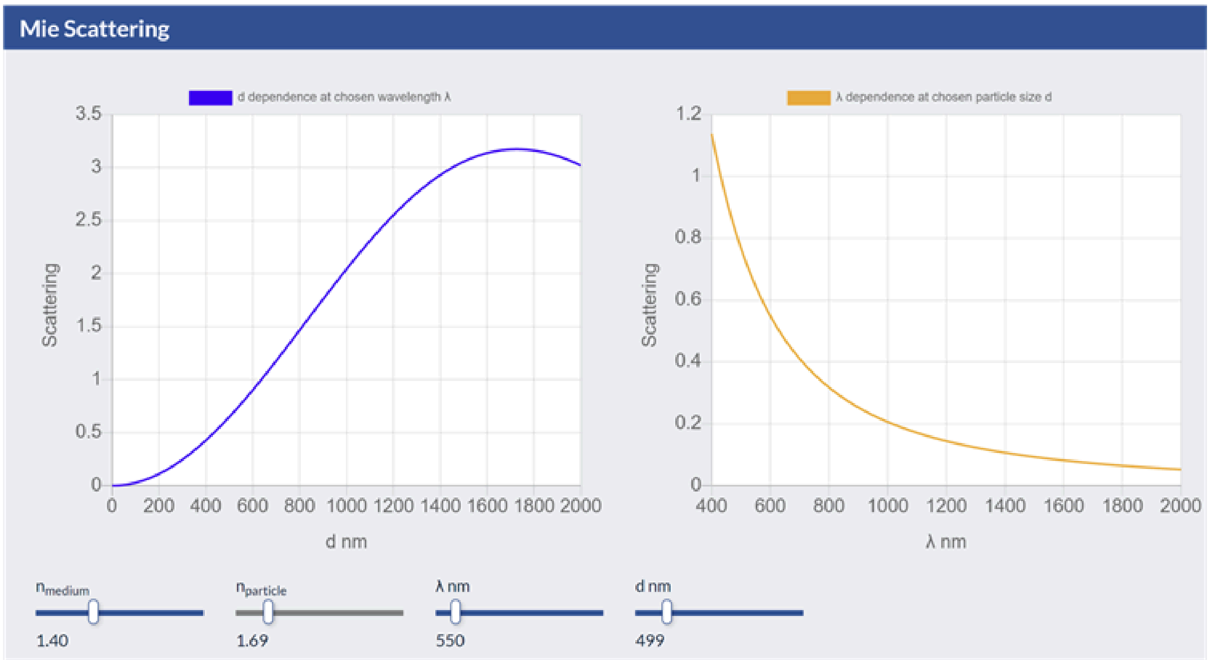
Links

Deodorant Sticks, Sun Screens, SkinCare, Solvent-Based Paint, Emulsion PaintLipstick Mascara

> For systems like inks, paints and sunscreens we often want light scattering to provide opacity and "hiding power", while for clear coatings we want to avoid scattering. The basic theory provided here is full of assumptions, but is the least bad option available.

Mie Scattering

If you have well-isolated particles in a uniform medium you can apply Mie scattering theory. This is too complex for an app, but there is an agreed "good enough" formula we can use. But any formulation with enough particles to be interesting will not conform to the assumptions behind Mie theory. Despite this, everyone uses Mie theory as a guide, so here it is:



n_{medium} 1.40

$n_{particle}$ 1.69

λ nm 550

d nm 499

https://www.stevenabbott.co.uk/practical-coatings/mie.php

We have some particles of diameter d, in this case 500nm. They have a refractive index, $n_{particle}$, of 1.69. They are in a medium with refractive index, n_{medium}, of 1.4. We can see the amount of scattering (the numbers have some meaning, but just think of them as relative values) in two ways:

1. If you are interested in scattering at a standard wavelength, λ, typically 550nm you see how the scattering depends on particle size.
2. If you are interested in a specific particle size, you can see how scattering depends on wavelength.

When you play with the sliders you find that for a given particle size or wavelength, scattering increases as the difference in refractive indices increases.

The simplified formula (van Hulst) that is conventionally used requires a parameter p:

$$p = 4\pi\left(\frac{n_{particle}}{n_{medium}} - 1\right)\frac{d}{2\lambda}$$

From this we can calculate the scattering Q:

$$Q = 2 - \frac{4}{p}\sin(p) + \frac{4}{p^2}\left(1 - \cos(p)\right)$$

If you want good opacity for white light then go for the highest refractive index particle you can find, typically TiO2, and use the app to find the particle diameter that gives the maximum. For TiO2 it is well known (and the app confirms it) that particles of 300nm are optimal. If we assume that ZnO's refractive index is ~2 then the optimal particle size is ~900nm.

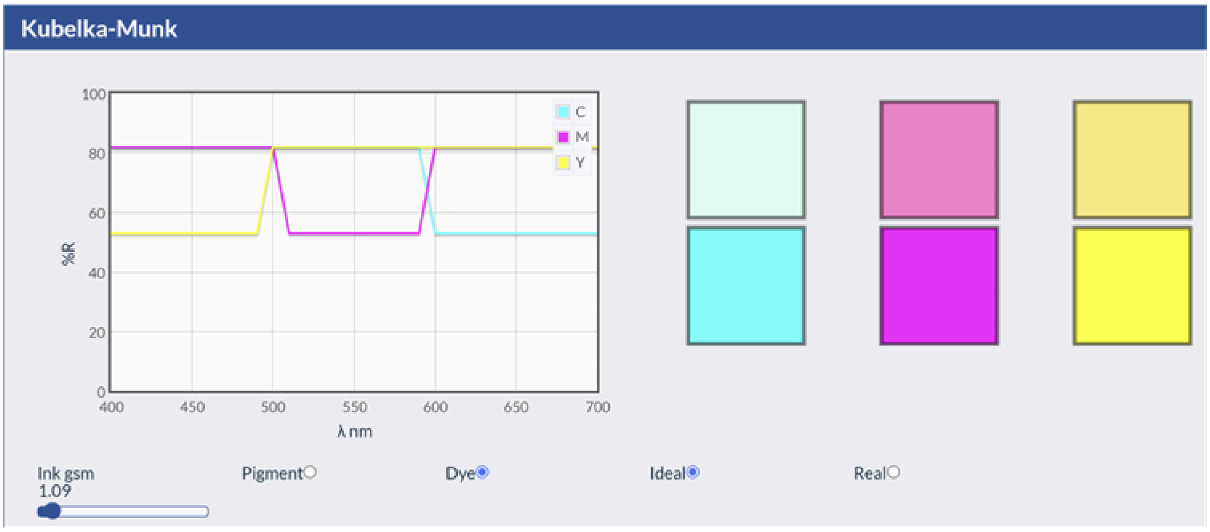
If you want transparent coatings then SiO2 is highly preferred as its refractive index is in the 1.4 range so is a good match for typical formulations that are also in the 1.4 range.

Scatter and colour

In paints it is common to use lots of TiO2 to provide the opacity and provide the colour via relatively small amounts of pigments. This works well because popular paint colours tend to be lighter shades, so the whiteness of the TiO2 is a benefit.

For printed colours such as inkjet, you need all the pigment you can get, so both colour and scattering are produced by the same pigment. There is a trade-off. Smaller pigment particles provide more colour per unit mass – the centre of a large pigment particle can absorb no extra light as it's already been absorbed by the outer few nm. But smaller particles scatter less. Getting the pigment to the correct trade-off size is part of the ink designer's craft.

When the inks are printed onto a scattering background, such as paper, you lose colour intensity because it gets drowned out by the scattered white light. You can calculate this effect via Kubelka-Munk theory:



Ink gsm 1.09

Pigment ○ Dye ● Ideal ● Real ○

https://www.stevenabbott.co.uk/practical-coatings/Kubelka-Munk.php

It's obvious here that the dye-based ink at 1.1 gsm is very much drowned out by the white background!

Particles Basic behaviour

Links

> Anyone who handles particles/powders struggles to control them. But after decades of particle research, we still have few reliable tools to understand their behaviour. The science here is at least some core basics for the formulation team to agree on.

Adhesion

There are a few core aspects to particle adhesion.

A sphere of radius R and a work of adhesion W (usually assumed to be the “surface energy”, typically 40 mN/m or 40 mJ/m^2, has an adhesive force F_{adh} given by:

$$F_{adh} = 1.5\pi RW$$

If the particle is stuck onto a vertical surface, will gravity pull it off? The gravitational force is:

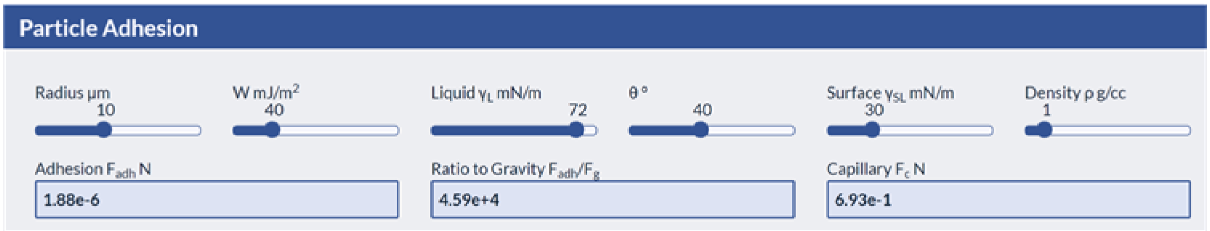
$$F_g = \frac{4}{3}\pi R^3 \rho g$$

and if you do the calculations for typical densities then for anything smaller than ~1mm, gravity is insignificant.

As is well known, if humidity is high (say above 50%) then you can start to get capillary condensation around the particle, giving an extra adhesion force. Where the surface tension of the liquid is γ_L, the contact angle is θ and the surface/liquid energy is γ_{SL} then:

$$F_c = 4\pi R\gamma_L \cos\theta + 4\pi R\gamma_{SL}$$

We can get the calculations from the app:



Particle Adhesion

Radius μm 10

W mJ/m^2 40

Liquid γ_L mN/m 72

θ ° 40

Surface γ_{SL} mN/m 30

Density ρ g/cc 1

Adhesion F_{adh} N: 1.88e-6

Ratio to Gravity F_{adh}/F_g: 4.59e+4

Capillary F_c N: 6.93e-1

https://www.stevenabbott.co.uk/practical-adhesion/particles.php

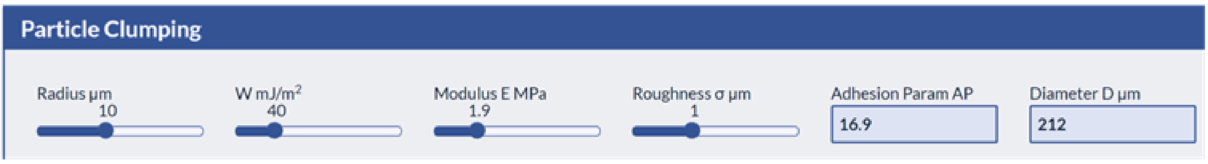
Those calculations are for smooth surfaces. Add a roughness σ for a particle with a (reduced) modulus E, then the adhesion is reduced by a factor of AP (Fuller-Tabor’s Adhesion Parameter), which, if larger than 10 essentially means no adhesion:

$$AP = \frac{E\sigma}{W}\sqrt{\frac{\sigma}{R}}$$

There’s also a tendency to clump into macro-particles of diameter D:

$$D = 0.064\left[\frac{2E[2R]^{\frac{5}{2}}}{W}\right]^{\frac{2}{3}}$$

Again, these are calculated on the second app on the same page:



Particle Clumping

Radius µm 10

W mJ/m² 40

Modulus E MPa 1.9

Roughness σ µm 1

Adhesion Param AP 16.9

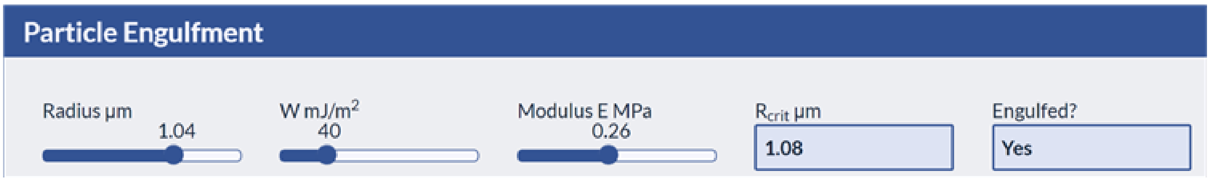
Diameter D µm 212

Engulfment

Sometimes particles can disappear spontaneously into a relatively soft surface. It seems an unfamiliar phenomenon, but in fact is well known to those who have tried to matt a surface and found that the particles they add tend not to give the desired rough surface.

If the radius of the particle is smaller than R_{crit} then the particle will be engulfed:

$$R_{crit} = \frac{7W}{E}$$



Particle Engulfment

Radius µm 1.04

W mJ/m² 40

Modulus E MPa 0.26

R_{crit} µm 1.08

Engulfed? Yes

Cleaning by blowing

Given that the pure surface adhesion force is so small, it should be easy to remove a particle by blowing over it. Given that the adhesion force is proportional to R, smaller particles should be really easy to remove! Yet we observe how hard it really is. This is because of the No Slip Boundary Condition, Cleaning_Boundary removal, which says that the air velocity at the surface is zero and rises relatively slowly above the surface. A small particle not only gets little air flow, but the removal force goes as R^2 so is less for the smaller particles. If you can get the air into a turbulent mode then the removal force increases.

Some convenient equations give us the drag for laminar flow in air of density ρ and viscosity η a distance x from where the air hits the surface:

$$F_{drag} = 2.6\rho R^2 v\sqrt{\frac{\eta}{x}}$$

For turbulent flow the equation is:

$$F_{drag} = \frac{19.8\rho(2R)^{1.44}v^{1.36}\eta^{0.54}}{x^{0.14}}$$

The app gives some idea of whether you are likely to remove a particle or not. Whether you are in laminar or turbulent mode depends if $Re = \frac{2Rv}{\eta} > 5$.

Particle Removal

Diameter μm 10

Velocity m/s 10

Re 5.5

W mJ/m^2 40

Distance x μm 10

F_{drag}lam. 1.50e-9

F_{drag}turb. 4.67e-7

F_{adh} N 9.42e-7

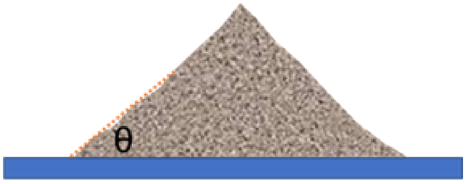
LamDrag/Adh 1.59e-3

TurbDrag/Adh 4.95e-1

https://www.stevenabbott.co.uk/practical-adhesion/partical-removal.php

Hoppers

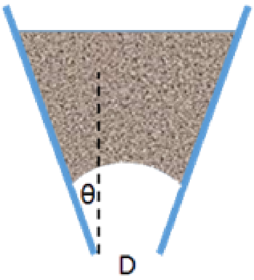
It is well known that near many hoppers across many industries their stands a big hammer. When the powder in the hopper gets jammed, the hammer is used to bash on the sides. The fact that this is still common in the 21st century is a sign that the science of powders in hoppers is not as useful as we should like. So when you read the following, expect to be disappointed. If there was some better, more useful science, it would be included here. Of course you can always use a big computer model with particle dynamic calculations, but if you have that sort of computational power and the data to feed into the model, you won't be reading this section.

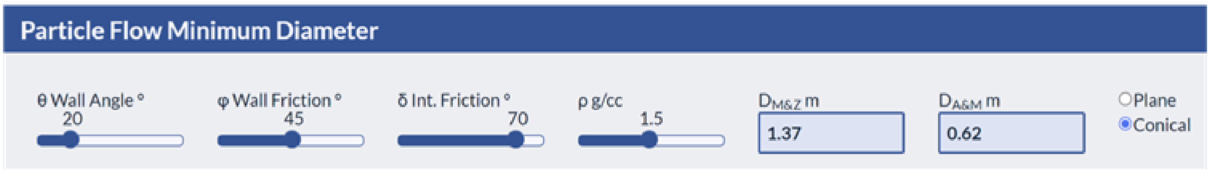


A simple metric for a powder is the angle of repose, θ, obtained when the powder is poured gently onto a surface.

We can consider it to be equivalent to a measure of μ, the coefficient of friction, (Mechanical_Friction) where $\mu = \tan(\theta)$

Obviously you want your hopper to be at an angle greater than θ. Confusingly, hopper angles are defined with respect to the vertical, so we want the hopper 90-θ to be greater than the angle of repose θ. In addition we have friction with the wall itself, so we have the "wall friction angle" φ as an extra value. However, the powder can stick to itself with another angle, δ, which is the "internal angle of friction". We can combine all those, with some horrendous algebra, to solve a core problem: what is the minimal opening, D, to a hopper that allows the powder to fall out rather than bridge?





Particle Flow Minimum Diameter

θ Wall Angle ° 20

φ Wall Friction ° 45

δ Int. Friction ° 70

ρ g/cc 1.5

$D_{M\&Z}$ m 1.37

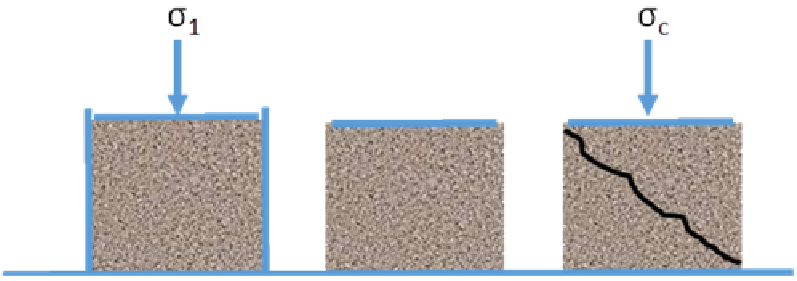
$D_{A\&M}$ m 0.62

○ Plane
● Conical

https://www.stevenabbott.co.uk/practical-adhesion/particle-flow.php

The final factor required for the calculation is the density ρ. Quite what this means is debatable, but the errors in the calculations are so large (as academic debates confirm) that worrying about ρ is not too important.

This is all OK if the powder is free flowing. But under the weight in the silo the powder might compact. If you know the powder's degree of compaction under pressure, you can, in principle, do something about it. But although the following thought experiment is simple, doing it isn't:



You trap some powder in a box and compress it with stress σ_1. You removed the sides of the box and press till you reach σ_c at which it cracks. You just need to do this for a series of σ_1 values, but because you need to pack the box each time people avoid this and use something like a Jenike box instead. The app page described the circuitous route needed to extract data.

Surface energy

The unhealthy obsession with surface energies being important for things like adhesion has led to a large expenditure of time using Washburn tubes (the choice of some) and Inverse Gas Chromatography (the choice of others). If you are tempted by Washburn tubes, read the caution in Absorbency, Porosity and Darcy. If you are interested in IGC, read my free eBook https://www.stevenabbott.co.uk/practical-chromatography/the-book.php.

That's it

Are you disappointed that we have nothing more to say? Well, it is really the case that decades of particle work has given us very little to put into your toolkit. Yes, you can read books with lots of general tips and tricks, but when it comes to reaching into a toolkit to solve a specific issue, there is very little available. But if you know something that I don't, let me know. I would be delighted to expand this chapter.

Particles Porosity

Links

The porosity of our particles and powders has a big effect on the overall properties. So how do we measure porosity?

2.5 Techniques

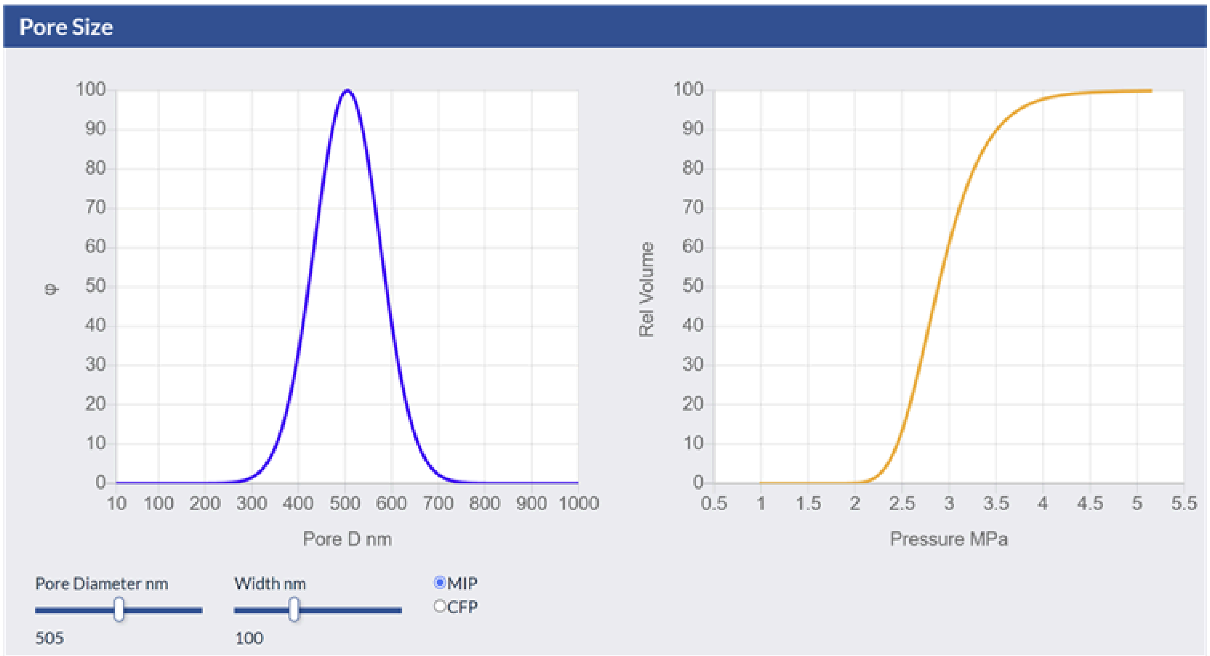
There are 2.5 standard methods

1. Mercury Intrusion Porosimetry (MIP) for micro- and meso-pores
2. Capillary Flow Porometry (CFP) also for micro- and meso-pores
3. Nitrogen porosimetry for nanopores

Why 2.5? Because #3, nitrogen porosimetry, is so full of assumptions and models that it's not really a valid technique – except if you compare values between samples that you have analysed via the same subjective technique. So here we discuss MIP and CFP.

The techniques work on the principle that the pressure P of a liquid inside a pore of diameter D depends on the surface tension of the liquid, σ, and the contact angle θ via:

$$P = \frac{4\sigma cos\theta}{D}$$



Pore Diameter nm 505

Width nm 100

● MIP
○ CFP

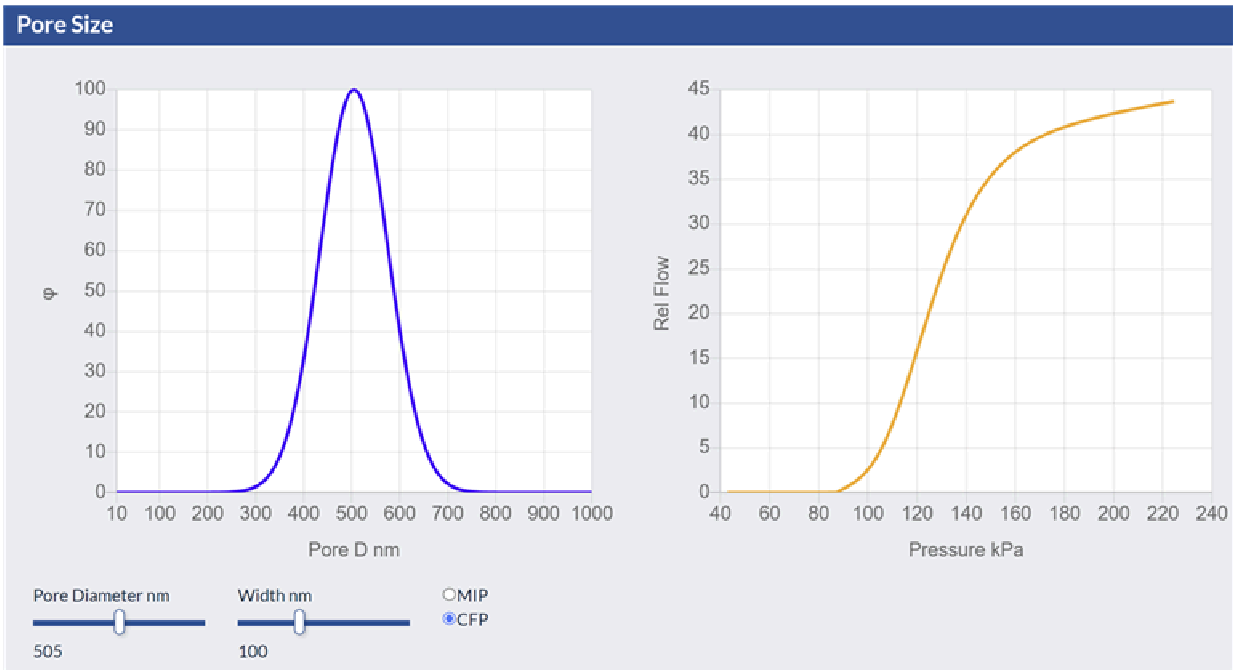
https://www.stevenabbott.co.uk/practical-coatings/Pore-Size.php

For MIP, where σ = 480 mN/m and θ = 140° the technique is simple in principle. Evacuate the sample then apply mercury under pressure, monitoring the volume uptake as the higher pressure increases the number of pores into which the mercury can flow. The app lets you create a simplified pore size distribution, the graph on the left, and see, on the right, what you would obtain in a real experiment.

The volume is only “relative” – in real life you would know the volume of particles and the volume of injected mercury so could get the total pore volume as well as the size distribution.

In the app the pore size is limited to ~ 10nm because below that the pressures become super-high, requiring more specialised equipment.

The same pore distribution measured with CFP is at first sight similar:



Pore Diameter nm 505

Width nm 100

○MIP

●CFP

But because the fluorocarbon used has σ = 16 mN/m and θ = 0° the pressure range is lower. And because the technique works by emptying the filled particles, at the highest pressures, the flow rate continues to rise as there is no longer any blocking fluid.

Filling the pores

The filling of a porous material is described via Darcy’s law, discussed here: Absorbency_Porosity and Darcy.

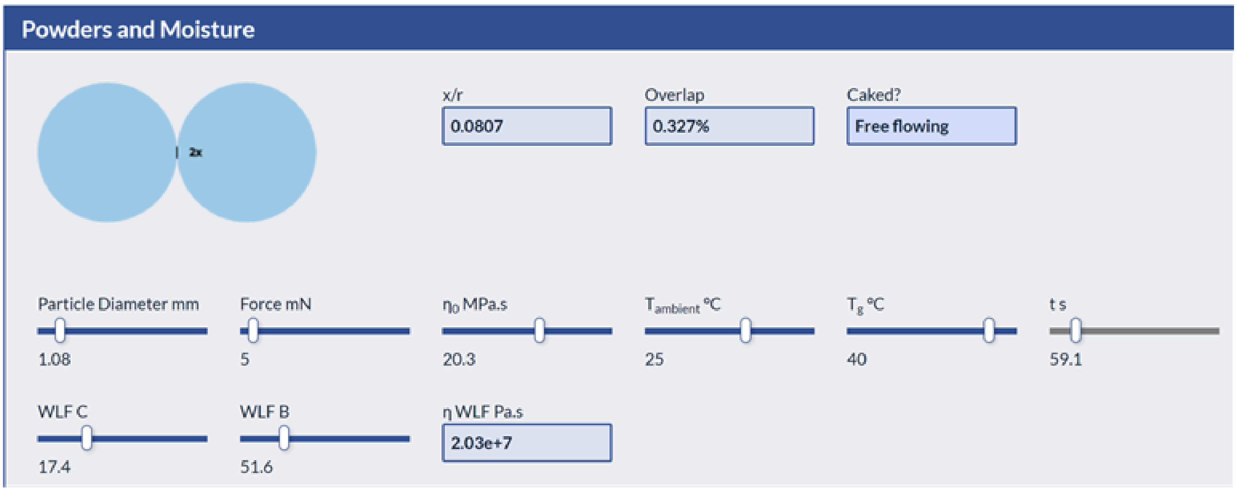
Particles Sintering

Links

Water-based Adhesive,Emulsion Paint

> We have chosen the relatively neutral word “sintering” to cover many aspects of particles coming together via extended contact and some applied pressure. This covers food powders, water-based paints and adhesives and “real” sintering of ceramics. Each application needs some flexibility in interpreting the specific formula used here; what’s important is to identify and control the key aspects influencing the effect for good or bad.

From free-flowing to caked



x/r: 0.0807
Overlap: 0.327%
Caked?: Free flowing
Particle Diameter mm: 1.08
Force mN: 5
η_0 MPa.s: 20.3
$T_{ambient}$ °C: 25
T_g °C: 40
t s: 59.1
WLF C: 17.4
WLF B: 51.6
η WLF Pa.s: 2.03e+7

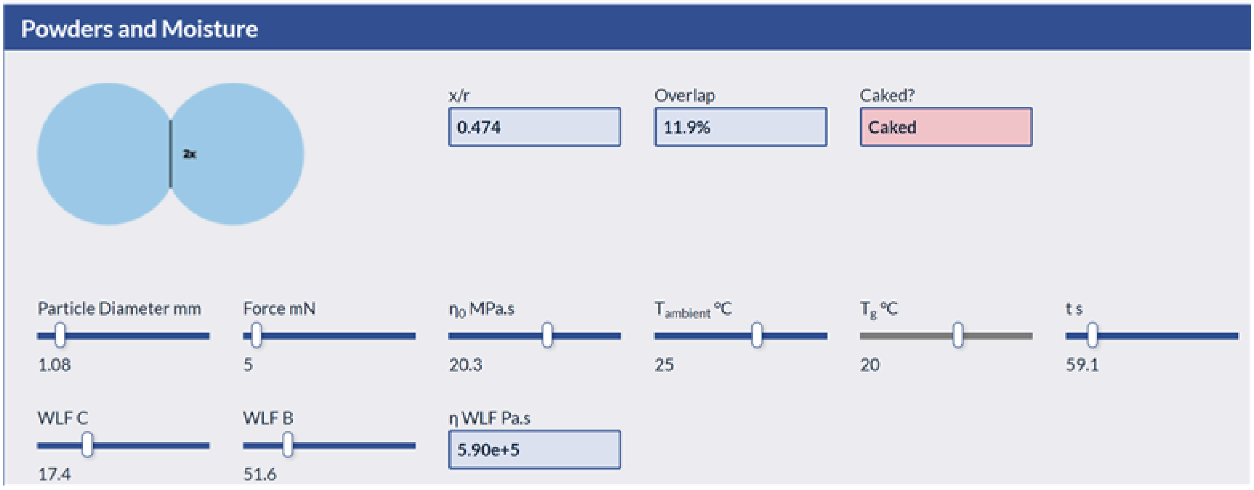
https://www.stevenabbott.co.uk/practical-mechanical/Powders-W.php

We have some 1mm powder particles and there’s a 5 mN force between them, for ~60s. The “viscosity” of the particles, η_0, is 20 MPa.s (that’s Mega, not milli). The glass transition temperature T_g is 40°C and because we are at 25°C we’re below T_g so our actual viscosity, η, is the same 20 MPa.s

From the Rumpf equation we can work out the % overlap (x/r) between particles after time t, where 2x is the contact width and r is the radius. It depends on the force F, the surface energy γ and our viscosity η:

$$\left(\frac{x}{r}\right)^2 = \left(0.8\frac{\gamma}{r} + 0.4\frac{F}{\pi r^2}\right)\frac{t}{\eta}$$

Now let’s expose the powder to humidity. It absorbs moisture according to its isotherm, Humidity_Water vapor isotherm. Because of the water, the T_g is reduced according to Humidity_Water mechanical isotherm, let’s say it’s fallen from 40 to 20°C. Now our powder is caked:



Powders and Moisture

x/r: 0.474

Overlap: 11.9%

Caked?: Caked

Particle Diameter mm: 1.08

Force mN: 5

η_0 MPa.s: 20.3

$T_{ambient}$ °C: 25

T_g °C: 20

t s: 59.1

WLF C: 17.4

WLF B: 51.6

η WLF Pa.s: 5.90e+5

That reduction in T_g reduced the viscosity from 20 MPa to 0.6 MPa and that was sufficient for sintering to give a 12% overlap.

That's the scary thing about sintering – relatively small changes can make a large difference. This works both ways. The assumption so far is that we want our powder to stay free-flowing. Sometimes you *want* particles to sinter sufficiently to give you a stable "green" shape which then can sinter properly by, say, heating in a furnace.

The sintering mindset

You may have a different word for describing particles getting stuck together. That's fine. The reason for this chapter being on "sintering" is that there's a good equation (Rumpf) that captures the core inputs and is valid across a wide range of particle issues. Using such a generic term allows you to focus on the generic science principles before getting sucked into the details of your own subset of issues.

For example, mud-cracking of particle-based formulations is unlikely to happen in formulations that sinter before the system becomes too dry. A couple of mud-cracking apps are discussed in Emulsion Paint.

> Science is universal. I was involved in some intense debates about how best to model the powder clumping tendency of a household cleaning product. It's a complex problem. Eventually we found work (cited in the app) funded by Nestlé covering a similar issue in the processing of food powders such as granules of instant coffee. This approach broke the problem into three steps, each of which could be modelled in an app. Instead of complexity we now had clarity and comparative simplicity.

Particles Size Distribution

Links

Toothpastes

Particle size distributions are difficult to measure reliably, difficult to interpret and the numbers quoted for a given particle might be inappropriate for the intended purpose. Via the app you can quickly understand why it is all so confusing … and how to get the numbers you need.

A brief comparison of measurement methods helps you find the least bad way for your particles.

Radius$_1$ nm 100

Height$_1$ 90

Width$_1$ 31

Radius$_2$ nm 378

Height$_2$ 10

Width$_2$ 290

☐ Log Normal

Max Plot Radius nm 914

D[1,0] : D[4,3]/D[1,0] 483nm : 1.85

D[3,2] : D[4,3]/D[3,2] 821nm : 1.09

D[4,3] 892nm

D50 898nm

Steepness 80

https://www.stevenabbott.co.uk/practical-rheology/distribution.php

We have a deliberately bi-modal distribution with 90% of the particles at 100nm and 10% at 378nm. These peaks are shown in red, the large one at the small radius and the small one at the large radius. The larger particles also have a wider distribution. These peaks represent the *number* of particles measured at each size.

Now shift to the blue peaks. These show the opposite trend – a small peak at low radius and large peak at large radius. These peaks show the *mass* (or volume) of particles at each size. Because mass goes as r^3, a relatively small number of particles represent a large mass.

We can see this in the cumulative curves. The orange number curve shows that most of the particles are below 150nm, while the cyan mass curve shows that most are above 350nm.

The green lines represent the *area* equivalent data, i.e. based on r^2.

To describe the distribution in a single number is unhelpful, but very common. Note that graphs of distribution are often in radii while quoted numbers are usually in diameters, so there's a factor of 2 from what you see in the plot. There are many possibilities. D[1,0] is the number average (mean), D[3,2] is the area average and D[4,3] the volume average. A common quoted value is D50 which is the value where *half* the volume-weighted particles are below that value.

The Steepness value is used especially in the world of mineral particles and is defined (as a percentage value) as the ratio of D30 & D70, representing the points where the cumulative volume value exceeds 30% and 70% respectively.

Measuring particle size distributions

A brave European project once asked several well-equipped labs to measure the particle size distributions of a number of powders in a “round robin” event. The results make depressing reading because there was a large variation not just between methods but between different groups using the same method. We’ll reveal the agreed least bad method at the end.

- Direct microscopy: TEM, SEM, Light microscopes
    - Advantages: “Real” measurements that include shape data.
    - Disadvantages: Relatively small sample size. Occlusion of particles. Bias towards a “nice” image
- Sieving
    - Advantages: A real measurement
    - Disadvantages: Suitable only for large sizes. Limited numbers of sieves, so crude data
- Light scattering:
    - Advantages: Works for small spherical particles in high dilution
    - Disadvantages: Doesn’t work at higher concentrations or with larger particles or with complex shapes and size distributions. Many assumptions between signal and results, so size distributions can often be unreliable.
- Analytical centrifugation (with real-time data)
    - Advantages: Works across a wide range of samples, sizes, distributions. Fewer assumptions in data interpretation
    - Disadvantages: Can be fooled by different particle densities and shapes

Although no technique is applicable to every type of particle, and each technique is preferred in different specific circumstances, the analytical centrifuge had the best combination of versatility and reliability of reported distributions. Although light scattering has had a reputation for being the default option, its many flaws for real-world particles suggest that the default option should shift to centrifugation.

Separation Filtration

Links

Although there is a well-known filtration equation, there are too many variables in it to make it useful, other than a guide to what you might do to keep flow rates acceptable.

The filtration equation

This is a variant of the Darcy equation, Absorbency_Porosity and Darcy. We have particles of density ρ, specific surface area S which for spheres of diameter d is given by $S = \frac{6}{d}$, at concentration C_S kg/m³, packing together with a porosity (void fraction) ε. Currently there is a volume V of filtrate already collected across a filter of area A. We have a constant k = 4.17 (from Carmen-Kozeny) and are applying a pressure P across the filter cake. The filter itself has a resistance R. We are interested in the increase of filtrate volume per unit time:

$$\frac{\delta V}{\delta t} = \frac{PA^2\varepsilon^3\rho}{\eta\left(kC_sV(1-\varepsilon)S^2 + RA\right)}$$

This is mostly intuitive. Higher pressure, larger filter area, higher porosity, lower viscosity, larger particles (smaller S) and a lower-resistance filter medium speed up filtration. The problem, of course, is that the rate of filtration decreases because of the 1/V dependency. If you increase the pressure to increase the flow, the chances are that the porosity will decrease, making matters worse.

And as is well-known, if your Particles_Size distribution includes lots of small particles, these can easily fill up pores between larger particles, decreasing ε. The fine particles can also start to clog up your filter medium, increasing R. This leads to a key message of this short chapter: if you need to filter, do what you can to reduce the fraction of fine particles before you start the filtration.

Spend time with the world of high-quality coffee and you find that fines are a big problem for optimal brewing which involves filtration through the espresso puck or through filter papers. It is often noted that coffee pods can give good, strong coffee using maybe only 60% of a typical espresso dose. The reason is that the roller mills used for the coffee pods can make fine coffee particles with low amounts of fines. Extraction is efficient without filters blocking. These roller mills aren't practical for daily use, and even high-quality burr grinders can't get the same balance of small particles with low levels of fines.

Filtering aerosols

It can seem strange to us that a, say, "1µm" filter can take out a high fraction of nanoparticles from the incoming air.

To understand this we first need the definition of "x µm" then we need to see how a fibre-based material captures particles.

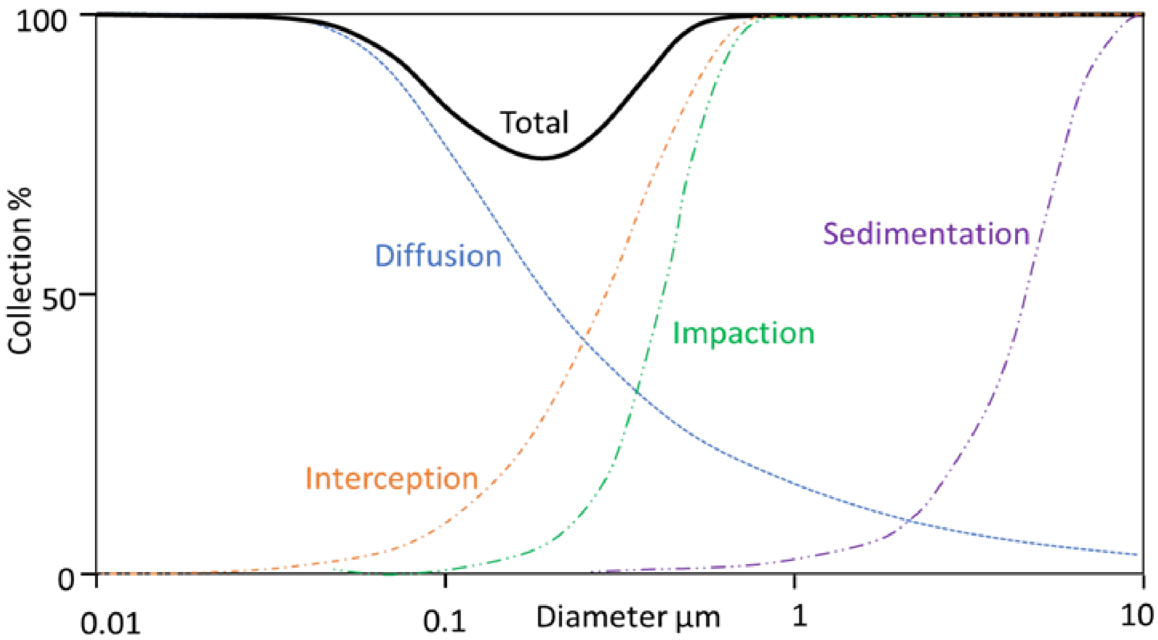
For a pure filter material with uniform x µm holes the filter size is x µm. But a typical fibre filter used in, say, aerosol masks doesn't have "a" hole diameter. Instead one side of the filter is placed into a liquid of surface tension γ with a contact angle (usually chosen to be near 0) of θ, and a pressure, P, is applied which is just large enough to create a bubble. From the simple capillary force equation the effective filter size, D, is:

$$D = \frac{4\gamma cos\theta}{P}$$

This value is purely a convenient fiction, useful for checking that masks are OK; that D is large enough to allow breathing, and small enough to provide a good filter.

A well-known graph shows how such a filter captures particles via:

- Sedimentation – the particles are so heavy they fall onto the filter;
- Impaction – the air channel is wide but changes direction so suddenly that the particle carries on and hits the filter;
- Interception – the particle cannot fit through the small gap in which it finds itself;
- Diffusion – the particle wanders randomly through the air and happens to hit the fibre material.



Obviously missing from this is a line for "electrostatics". That is because it is really tricky to design, model and rely on electrostatics which can so easily change with, say, relative humidity.

It's good to know that when we put on relatively simple masks to handle dusty particles, this combination of mechanisms does a good job of limiting risk, with each mode being especially good for particles that other modes might miss. It is the diffusion mechanism that captures most of the small particles from a mask porous enough to allow us to breath easily.

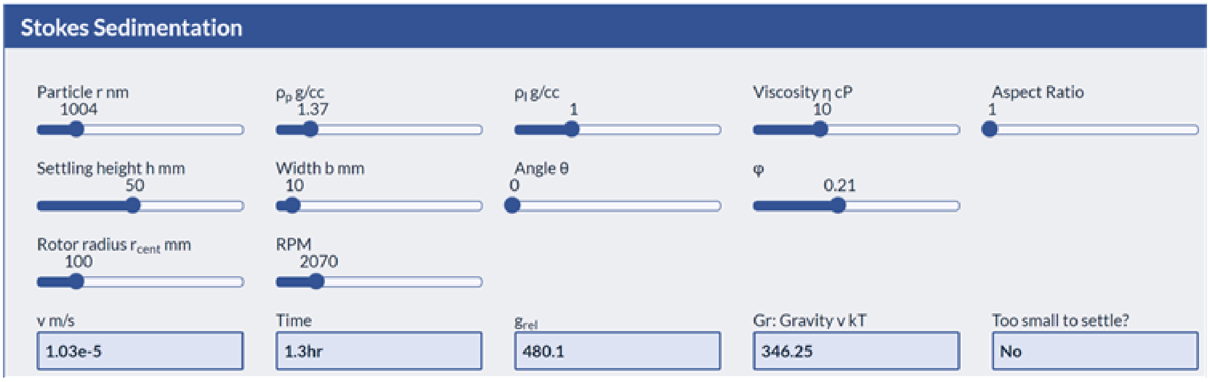
I tried to find the formulae for producing those graphs from reasonable inputs, but was defeated by their complexity.

Separation Settling and Centrifugation

Links

> Given some simple equations and a few basic parameters, knowing for how long you need to let your particles settle, or how fast to centrifuge your sample to achieve separation is straightforward.

The basic settling and centrifugation equations



Stokes Sedimentation

Particle r nm 1004

ρ_p g/cc 1.37

ρ_l g/cc 1

Viscosity η cP 10

Aspect Ratio 1

Settling height h mm 50

Width b mm 10

Angle θ 0

φ 0.21

Rotor radius r_{cent} mm 100

RPM 2070

v m/s 1.03e-5

Time 1.3hr

g_{rel} 480.1

Gr: Gravity v kT 346.25

Too small to settle? No

https://www.stevenabbott.co.uk/practical-solubility/stokes.php

We have spherical particles of radius r, density ρ_p in a liquid of density ρ_l and viscosity η. We want them to settle a distance of h. Without centrifugation the settling force is 1g. With centrifugation in tubes sitting at a distance r_{cent} from the centre of rotation of the machine spinning at our chosen RPM we have an extra g force, g_{rel} which is 1 with no rotation. The time, t, taken to settle depends on settling velocity v, so $t = \frac{h}{v}$. The velocity is given by:

$$v = 2.18(\rho_p - \rho_l)\frac{g_{rel}r^2}{\eta}$$

The relative force of gravity is given by:

$$g_{rel} = 1 + 1.118e^{-3}r_{cent}RPM^2$$

Some extra features

If the particles aren't spherical so their aspect ratio (length/width), A > 1, then they behave as if they were spherical particles with a larger radius, r_{eff} given by:

$$r_{eff} = r\sqrt{\frac{\operatorname{atan}\left(\sqrt{A^2 - 1}\right)}{\sqrt{A^2 - 1}}}$$

If the volume fraction of particles, φ, is large then the particles slow each other down. The standard Richardson & Zaki formula (in the Eulerian frame of reference) is a reduction in velocity of $(1 - \varphi)^{5.65}$.

If the tubes are held at an angle θ away from the gravitational pull then the Boycott effect takes place such that in a tube of width b, the effective velocity is increased by $\frac{h}{b} sin\theta$. The effect is surprisingly large and very real. Boycott first observed it in the 1920s for settling of red blood cells.

Like most people, I'd never heard of the Boycott effect. After a post about an improvement to a settling app, someone commented that it would be good to add the Boycott effect. It was great fun to learn of something completely new, to find the theory and to share it with many others for whom it was also new. It also shows the power of open science – there's always someone smarter than you out there from whom you can learn cool new stuff.

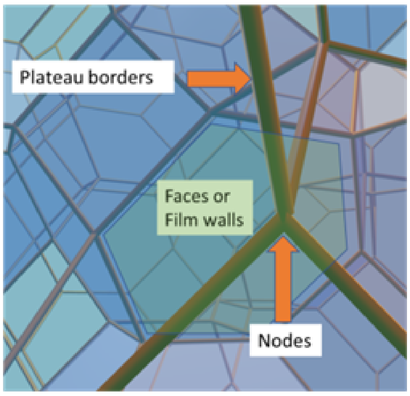
Surfactancy Anti-foaming

Links

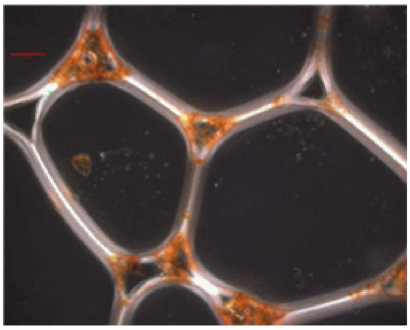
Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids

A well-known anti-foaming theory is of no real help, and a profound analytical technique seems too difficult to use in a formulation lab. So anti-foaming (even with silicones) is less about hard science and more about experience and, sadly, luck.

The core problems



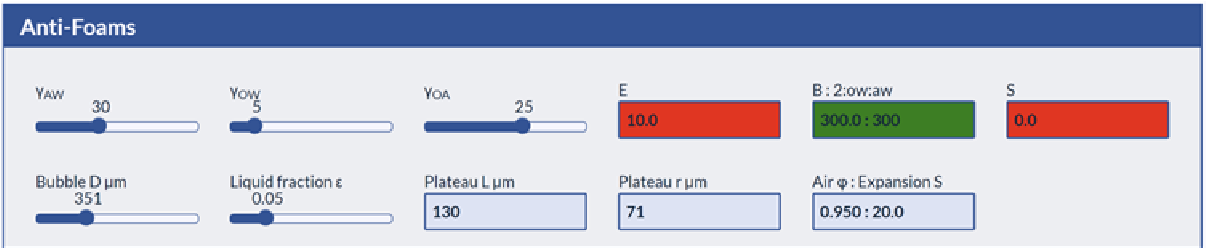
You want some blobs of oil to go to the foam wall and destroy it. If the blobs of oil are too large, there are too few of them to make a difference. This, therefore, excludes simple addition of hydrophobic oils such as mineral oil or silicones which quickly form such blobs. To get small oil drops you either have to have a not-very-hydrophobic oil, which might not destroy the foam wall, or you need to emulsify the oil, which means that it is protected by a surfactant layer so isn't destructive. Or you can break up drops via faster agitation, which is likely to induce more of the foam you are trying to destroy. One way round these contradictions is to engineer a complex system, which is how modern silicone/silica anti-foam, discussed below, can be effective, but still with plenty of trial and error.



There is a further problem. Within a foam, 90+% of the liquid is in the Plateau borders and the Nodes, so very little of the anti-foam has a chance to destroy the film walls. This beautiful image, reproduced with kind permission by Dr Illy, shows how oily particles in an espresso crema aren't acting as an anti-foam, they are just sitting in the borders and nodes.

Those drops of oil sitting in the nodes will do nothing until the liquid has drained so much that the drop is compressed against the wall and can then destroy the foam. This is the reason behind the phenomenon of anti-foams seeming to have no effect for a while, then suddenly the foam disappears. The calculation of the radius of these borders, in the second row of the Anti-foam app is discussed below.

Although there *is* a well-known theory around entry, bridging and spreading coefficients calculated in the top row of the app:


Anti-Foams

γAW 30

γOW 5

γOA 25

E 🟥 10.0

B : 2:ow:aw 🟩 300.0 : 300

S 🟥 0.0

Bubble D µm 351

Liquid fraction ε 0.05

Plateau L µm 130

Plateau r µm 71

Air φ : Expansion S 0.950 : 20.0

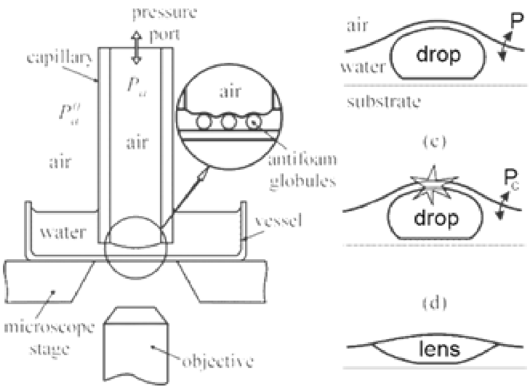
https://www.stevenabbott.co.uk/practical-surfactants/anti-foams.php

it has largely been discredited, because it focusses on oil drops touching the foam wall while the arguments above point out that this hardly happens. At best it tells you when failure is guaranteed by relative surface energies, but as failure is normal, even for those predicted to be good, that’s not much help. There is another reason why it doesn’t work.

The anti-foam app also takes a pessimistic view of anti-foaming. An earlier version had implied that anti-foam science was useless. I then had the chance for a delightful working lunch with a world expert on anti-foams. After considerable debate, I returned home and changed the app to say that the science was “mostly useless”. The expert, with proven success in industry, taught me a lot about specific issues. My pessimism is about the lack of usable science for us non-experts.

A subsequent discussion with an expert on modern silicone anti-foam systems allowed me to expand this chapter, but the core problem of lack of usable science remains.

The Entry Barrier



Using their Film Trapping Technique, FTT, the Denkov group in U Sofia showed that it can be remarkably difficult for an oil drop to push through an “entry barrier” at the foam surface:

If we could properly understand the entry barrier, measure it routinely, and test the barrier with different anti-foam oils we might be able to make some progress. Unfortunately, the FTT depends on the considerable skill and knowhow of the Sofia team and hasn’t yet become a routine test in other labs.

Silicone anti-foams

Throw in some silicone oil … and the chances are that it will not help de-foam. As discussed above, the oil will just gather as a blob, maybe breaking a few local foam walls but having little impact on the foam in general. What we need it lots of small droplets of silicone oil – and for that we need good surfactancy. If the foaming surfactant was a good emulsifier for the silicones, it might be too good – hiding the silicone from the foam. The short discussion in the Surfactancy_HLD chapter suggests that the only good surfactant for a silicone oil is a silicone surfactant. If the silicone surfactants (often complex silicone-PEO-PPO systems) are ignored by the rest of the system (as seems likely) then the silicone oil droplets can be nicely emulsified by the silicone surfactant, giving a semi-independent system able to deliver the silicone oil to the foam walls. There is good evidence that this semi-independence is real.

One way to control the silicone oil against breaking up too much into drops that are too small and too shielded to be effective against foam, is to make the “oil” a polymer, maybe a “rubber” from some silicone crosslinking. Those who use PDMS, which is high MW silicone, generally expect it to be rubbery via crosslinking. Obviously, highly-crosslinked PDMS could produce excellent, stable, (say) 5 µm silicone spheres, but these (especially when wrapped around by PEO/PPO chains) will be mostly harmless to the foam.

The size of the silicone drops/particles has to be optimized for each application, again for reasons that are unclear. The second row of the app is relevant to the problem of drops being trapped in the Plateau borders. If your bubbles are small and the foam is dry, the length and radius of the Plateau borders are both small, so small drops will be able to act. Large bubbles in a wet foam will need very large drops/particles to start to disrupt the borders.

What about the thickness of the foam wall? It is generally agreed that the capillary pressure sucks water out within time frames of seconds so they are all effectively sub 100 nm thick. There seems no obvious link to drop/particle size and the very thin walls that are formed so quickly.

Even with blends of oils, polymers, rubbers and fancy surfactants, the anti-foam can be useless. The final ingredient is some relatively sharp hydrophobic (silane-coated) silica particles that, using simplistic language, can puncture the foam wall, allowing the silicone oil to do its thing. Why not use these hydrophobic spikes without all the silicones? One common explanation is that they are difficult to disperse in a way that gets them to the foam walls. Another is that even if they pierce the walls, the surfactant can just flow around it. So the spikes need to bring the silicone along with them to finish off the job they started. Neither explanation seems satisfactory.

For a functioning anti-foam we need the oil, surfactant and silica all to be together, all being carried to every part of the foam, all able to destroy the foam walls and not get stuck together as hydrophobic islands nor get stuck in the Plateau borders and nodes, nor get stuck on the walls of our process equipment to be removed, with great difficulty, when the deposit gets too thick. How much the famous low surface tension of the silicone oil plays in this is unknown. We can’t do side-by-side comparisons with mineral oil equivalents as there are many other differences between successful formulations.

What might go wrong? Everything. The real surfactant might interfere with the silicone surfactant or the silica particles. Other parts of the formulation, such as, builders, enzymes or bleaches in a detergent, or fibres and coatings in paper pulp processing, might divert the silicone surfactants in some strange way. And, as discussed in Fragrances, one fragrance package might be fine while another wrecks the anti-foam performance by, presumably, partitioning unhelpfully at the PEO/PPO/Silicone boundary and wrecking the curvature needed to emulsify the silicone oil. Or maybe the fragrance changes the chameleon-like nature of (flexible) silicones that can happily present a relatively hydrophilic side if the methyl groups choose to fold inwards (perhaps around the fragrance molecules). A “hydrophilic” blob of silicone oil will no longer be a threat to the foam wall.

As mentioned above, those skilled in the art, with plenty of experience, can have a feel for how to navigate around this complex space, and how to find plausible hypotheses for failure in a specific environment, with plausible fixes for the problem. Yet, as these experts freely admit, their attempts can still fail. It’s even tougher for the rest of us who don’t have this specialised knowledge.

How might we change this? In the spirit of “smart mapping”, used elsewhere in this book, you can imagine systematic experiments that could explore simpler regions of anti-foam space which, when combined against results from real-world experiments, might lead to a more fundamental approach to problem solving.

But there’s one more problem.

Testing an anti-foam

The chapter on Surfactancy_Foaming shows that there aren’t many good rules for deciding which surfactants are low foam – it depends strongly on the context of how you make the foam. A surfactant might be useless in one context and create a splendid foam in another. This makes testing for anti-foam properties especially hard. The only test that really matters is in-use. Putting on lots of washing machines and testing with different anti-foam formulations is a necessary qualification step, but hopeless for development. So “all” you need is a relevant foam test. But because the general approach to foaming is a confused mix of tests that are often poor or irrelevant, there is a lack of rational debate as to which tests map best onto which real-world anti-foam challenge. Do you use a Ross-Miles, a Foam Analyzer, a kitchen blender or kitchen mixer, a shake test, some timed sequence in a SITA tester, and so on? It was amusing to see a video saying that “the” test for foam was ASTM D1173. Yes, this is the test for those interested in foams analysed according to ASTM D1173 (I could have chosen other standards) but might be, and probably is, totally irrelevant for your specific foaming issue.

The trick, therefore, is to find a quick and easy test that maps well onto your specific challenge. You can focus your formulation efforts to get the most information for the least effort. The chances of success in the field are greater than if you had spent time working on, say, ASTM D1173 if it happens to have near-zero correlation with the foaming mechanism in your system.

You already have a selection of your own and competitor formulations with known field results with specific surfactant systems (e.g. the 2016 and 2018 versions of well-known detergents). If your test does a decent job at distinguishing good from bad then you not only have a fast-track method for future development, you have a reliable method for testing hypotheses: “If I replace X with Y in the successful 2016 de-foamer formulation, will it make things better or worse, if so, why?”

The fixation with ASTM standards gets in the way of optimising other complex formulations such as Adhesive. It’s not that ASTM standards are bad (they’re not), but that they are often irrelevant to your specific problem and tie up too much precious formulation resource. Once you have a quick test that lets you formulate rapidly against criteria relevant to your problem, the chances are high that the product will then pass the appropriate ASTM test or, even better, the end use test which is what really matters.

The best anti-foam …

… is to not create bubbles in your process.

Analyse why there is air, how you can avoid or reduce it, and, if you have to have air, how you can avoid individual bubbles (usually harmless) from reaching the surface, or joining up to form a foam, which includes trying to avoid larger bubbles breaking up into smaller ones. A single bubble of 500 µm is much less harmful than the same volume of air in 125 100 µm bubbles, not just because the isolated bubble is less protected against damage, but because foams from smaller bubbles (see other foam chapters) are significantly more stable than larger ones.

By shifting the discussion from trying to find the perfect anti-foam (which doesn’t exist) to rational ways to mitigate the production of foam the chances of a long-term solution to your problems increase significantly.

Happily, others have reached the same conclusion. and many have, indeed, found ways to solve their foaming problems without all the problems they used to have with antifoams.

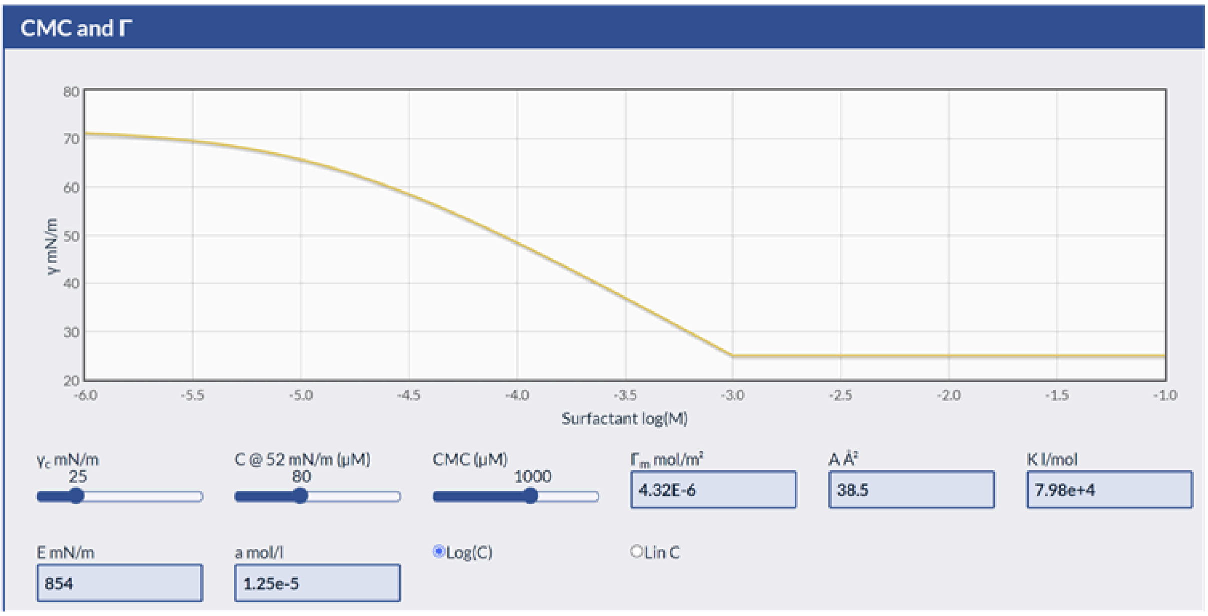
Surfactancy CMC and Langmuir

Links

[Dishwashing Liquids](), [Hair Shampoos](), [Soaps and Washing](), [Laundry Liquids](), [Surface Cleaning]()

> Critical Micelle Concentration, CMC, is featured heavily in surfactant science and yet is of little interest to the formulator. It is best measured with carefully purified surfactants in pure water, and yet our formulations are with commercial-grade surfactants and we have lots of other ingredients.
>
> If CMC was important, we'd all use long-chain ethoxylates which have low CMC values. In reality we use lots of C12 sulfates with high CMC values.
>
> If you measure CMC using a plot of surface tension versus concentration, the other information from the experiment is more interesting than the CMC.

The Langmuir isotherm



γ_c mN/m 25

C @ 52 mN/m (μM) 80

CMC (μM) 1000

Γ_m mol/m² 4.32E-6

A Å² 38.5

K l/mol 7.98e+4

E mN/m 854

a mol/l 1.25e-5

● Log(C)

○ Lin C

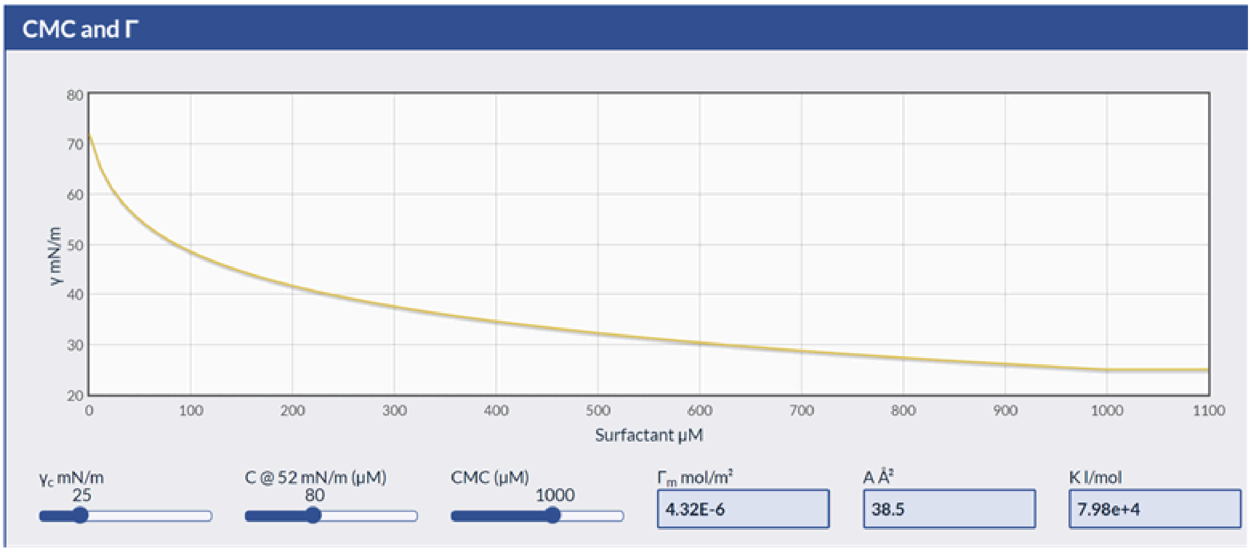
https://www.stevenabbott.co.uk/practical-surfactants/cmc.php

When you plot surface tension, γ, versus surfactant concentration, c, the resulting plot can be analysed via the Langmuir-Szyszkowski plot. It depends on the starting value, γ_0 which is 72 mN/m for pure water, on a surface excess concentration Γ_m and an equilibrium constant K:

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

The app takes your experimental γ_c, the CMC and the concentration at which the surface tension has decreased by 20 mN/m to construct the whole curve and extract the important Γ_m and K values.

What is significant about 52 mN/m? This is when you approach full surface coverage, an idea that shocks those who think that CMC is the point of full surface coverage. But CMC knows *nothing* about surfaces. You can measure CMC via bulk properties such as conductivity or scattering. The relative *un*importance of CMC is shown when you plot the same data on a linear concentration scale:


CMC and Γ

γ_c mN/m 25

C @ 52 mN/m (μM) 80

CMC (μM) 1000

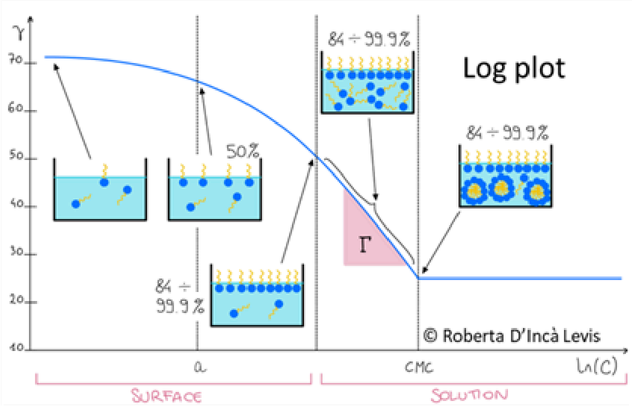
Γ_m mol/m² 4.32E-6

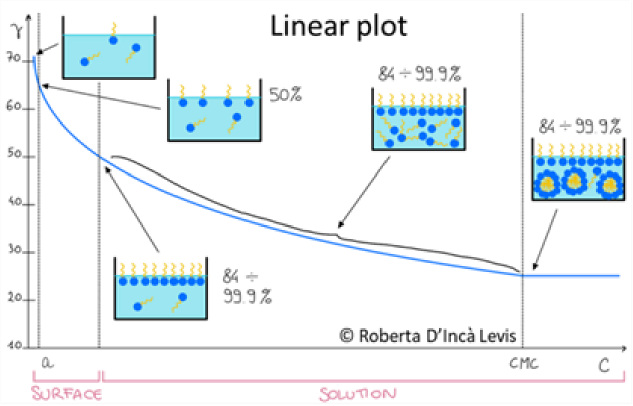
A Å² 38.5

K l/mol 7.98e+4

Now you see that the interesting stuff takes place early on and it is actually hard to spot where the CMC is, because nothing much changes at the point.

This can all be explained by a pair of beautiful diagrams kindly supplied by **Roberta D'Incà Levis**.





We see the steady increase of concentration in the bulk, leading eventually to micelle formation, but, especially clear in the Linear plot, we see that the surface is near saturation a long time before micelle formation. As a bonus, the Log plot shows where Γ is determined, over the relatively narrow range where the surface goes from rather crowded to very crowded.

Γ_m, K and A

Given that the CMC is uninteresting in most respects, if you want to do these titrations with your (pure) surfactant (curves from commercial grade surfactants offer few deep insights), then learn to interpret the surface excess concentration Γm, the partition coefficient, K, and the surfactant head area A.

It turns out that values for surface excess don't differ strongly over most common surfactants. The partition coefficients follow intuitions – more hydrophobic surfactants have a higher coefficient. If you want a lot of surface stability from a little surfactant then, arguably, a large head area might help. But it's usually more complicated than that. A is also used for calculating the Critical Packing Parameter, Surfactancy_CPP and phases .

If the world of surfactant science had de-emphasised CMC and had focussed, for example, on the Cc values for Surfactancy_HLD theory, we would all be formulating much more effectively.

γ_c

For most surfactant uses, γ_c is of surprisingly little importance. It is measured, carefully, over timescales of minutes. Most surfactant uses require the surfactant to act in milliseconds or seconds. So for formulators, Surfactancy_Dynamic Surface Tension is far more important. In a nice irony, one way to ensure a fast reduction in surface tension is, as that chapter shows, to choose a surfactant with a relatively high CMC.

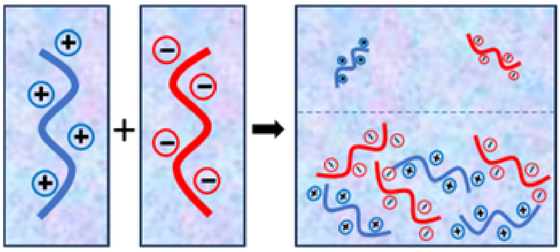
Surfactancy Coacervation

Links

Hair Shampoos, Soaps and Washing, Surface Cleaning, Pharma Formulation

Coacervation is one of those formulation techniques we take for granted in our respective areas of expertise, such as Microencapsulation or creating all-in-one Hair Shampoos. It is unique because we want it for controlled "failure" to keep things in solution. Yet we have no usable theoretical tools and little practical guidance about how to design our own formulations. So this was a very short chapter. Thanks to a brilliant lecture by Dr Tobias Halthur of CR Competence AB, the chapter has been extended to include a summary of all charged/uncharged surfactant/polymer interactions showing that the coacervate effect is a generalisation of a few key ideas … and that if we try to get the best of both worlds, we generally make things worse.

Plus and minus

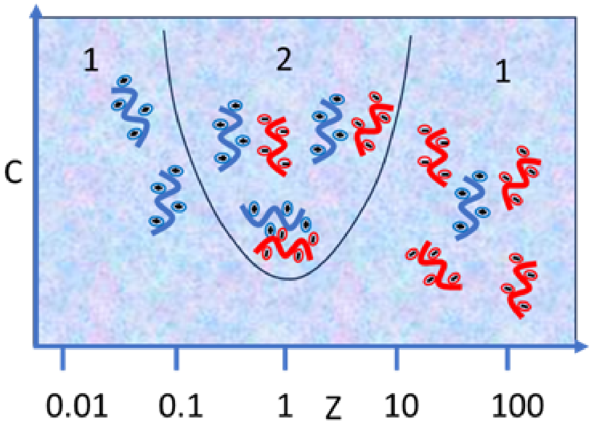


We have at least one charged polymer (e.g. a polyacrylic, a polysulfonate, a polyquat) and an oppositely charged molecule that is either a polymer or something a bit complex like a surfactant. And we may have some added salt.

Under the right conditions, the "complex coacervate" appears most commonly as some sort of scattering or turbidity, maybe along with some visible phase separation, depending on the size and density of the drops. The scattering droplets aren't so much a neutral pairing crashing out of solution, but pairings that contain a lot less water, but still plenty of it, in water that contains a lot less of the starting pair. The droplet detection limit depends on the wavelength of interrogating radiation – x-rays see the onset of coacervation before optical turbidity meters or the human eye. If (e.g. for surface cleaners, Surface Cleaning) you want the coacervate at the surface, clever FTIR experiments (Analytical Techniques) can extract insightful signals.

It is strange to say "right conditions" because this means "failure" to remain as an homogenous solution. In the classic case of 2-in-1 shampoos we want to convert a homogenous solution containing lots of surfactant, some charged polymer and some oil into a phase-separated state which deposits charged polymer and oil (tangled in the polymer) onto the surface. The transformation must happen when the whole system is diluted. In the ideal case, the deposit is at least kinetically stable but the small surfactant molecules wash out, leaving the charged polymer and oil to do what the formulator requires.

This phase separation is best shown via ternary phase diagrams, as discussed in the general overview below. Before then we can look at a binary phase diagram often shown in the context of coacervates:



The classic coacervate plot shows the ratio of charged species, Z, against concentration, C. When this is 1 you have the two phases. When it is small or large you have a single phase, in between, the higher the concentration of each (at the same ratio) the more likely you are to have the coacervate. You can get similar plots with change of the salt concentration (or, more likely, ionic strength). With any of these changes you get more or less turbidity or a different FTIR signal. For example, a gross imbalance of the charges (wrong ratio) means that there's little reason for a new species to be formed. And at high salt concentrations, charge effects are so shielded that little of interest can happen. If you change not the polymers but the % charged species in the polymer, then effects are unpredictable. And if you've had great success with polymer A and opposite charged B and want to change A to C or B to D, you are probably starting from ground zero.

In 1957 Voorn/Overbeek theory was developed to describe the phenomena. You can readily generate V/O parameters to fit your data (I could produce a V/O app!), but that's no help because just about every assumption behind V/O theory is wrong. Clever work by smart theoreticians has produced (according to reviews) at least 3 powerful theories that could/should explain coacervates but, as the experts acknowledge they are each wrong/limited in at least one major way and none of them is usable in any real-world context.

What can you do about it?

A few chapters reach the same sad conclusion that there are no useful toolkit tools. Sometimes this is because researchers have been vainly using the wrong approach and things can be fixed with a switch to a more productive one. For coacervates a lot of very smart people have worked hard to come up with better predictive tools but it seems a genuinely tough problem that awaits some genius breakthrough to give us something for our toolkit. Sure, we understand the *general* phenomena, but we have no workable, appable tools to make development less painful.

Although what follows still doesn't help directly with coacervates, the logical approach that I learned from a lecture by Dr Tobias Halthur provides some structured thinking to the relevant issues. It also shows that coacervates are one example of a range of inter-related phenomena. Generally we want the best of both worlds as these systems are often intended to give us controlled viscosity in our formulation. If A gives viscosity, but not enough, and the same with B, maybe A+B will give us what we need. Sadly, the opposite is often true. As Dr Halthur acknowledges, a lot of the ideas below are explored in the book *Surface Chemistry of Surfactants and Polymers* by Kronberg, Holmberg and Lindman. Of course, all blame for oversimplifications of this complex topic goes to me.

The right viscosity behaviour

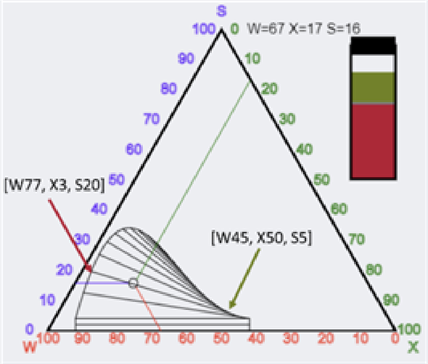
It is a common formulation requirement to obtain a relatively high low-shear viscosity along with a low high-shear viscosity. The chapters on Thickeners_Associative Thickeners, Thickeners_Polymeric Thickeners and Thickeners_Wormlike Micelles cover some key themes.

The main theme is that entanglement provides viscosity. With conventional high MW polymers, entanglement is easy to obtain, but although shear thinning occurs (Flow_Shear dependent) it is not the rapid decline that is often required in paints, and personal care products. HEURs and Wormlikes are entangled, but now we get the desired properties because the hydrophobic forces that create the tangles are easily broken by shear … then recover rapidly.

If you can get the desired properties in your specific formulation with one of those three approaches, then life is easy. When they each fail, for different reasons, it's tempting to think that a combination of two of them might solve the problem. The point of this addition to the Coacervates story is that many such attempts are guaranteed to fail, for reasons that are not at all obvious.

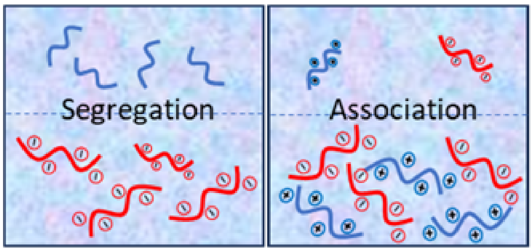
Ternary phase diagrams

Most of us are uncomfortable with ternary diagrams. The Surfactancy_Phase Diagrams chapter attempts to reduce our discomfort and the screen shot from https://www.stevenabbott.co.uk/practical-surfactants/pde.php shows how to read the values at the point shown by the circle. The trick is to read from a line that is parallel to the face of the triangle *opposite* the component of interest and from the scale that reaches 100 at the component of interest. The line relevant for Water is horizontal, so it intersects the blue and the green scales. But the blue scale reaches 100 at W so you read that it's 46%. The Surfactant line is parallel with the right edge of the triangle and because it is the red scale that reaches 100 at S, you read off 27%.



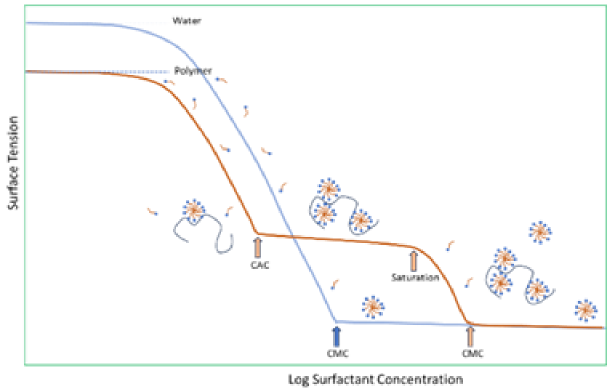
The next thing we need to be comfortable with is drawings of "tie lines". Before we do, note that in this phase diagram (from https://www.stevenabbott.co.uk/practical-surfactants/pde2.php), the order of W, S, X has been shifted from the previous plot and the helpful grid showing the axes has been removed. It is frustrating that this view is different from the previous one, but you have to accept that everyone plots their phase diagrams the way they happen to like them. You have to get used to reading whatever is put in front of you.

This time, the circle is at [W67, X17, S16] and is in the middle of a group of lines within a boundary. These are tie lines. The entire region inside the boundary is unstable – the system spontaneously breaks into 2 phases. We often say that "the surfactant (or polymer) crashes out of solution" which implies a region with pure solution and a region of pure surfactant/polymer. That is almost never the case. Instead you have a region with a high water content and a low level of surfactant/polymer, then a region of high surfactant/polymer and lower water content. The tie lines tell you the contents of the two phases. In this case at one end you have 77% water and 3% X (maybe our polymer) and at the other end it is 45% water, 50%X and 5% S.



One final detail about tie lines. In the example above they are more-or-less horizontal. These are "segregation" tie lines. If they are more-or-less vertical then they are "association" tie lines. We see the difference in this diagram. The segregation type simply gives region of higher and lower concentrations of the respective ingredients. The association type has regions of higher and lower concentrations of *both* ingredients. If you want a single phase, you may not be concerned whether phase separation is via segregation or association. But if you find two phases, although it's obvious from the diagrams what is going on, to optimize a real-world formulation, understanding which type of failure you are looking at, gives you some root cause and cure ideas.

Clues from the CAC



We are all familiar with the critical micelle concentration, CMC (Surfactancy_CMC and Langmuir) which we can obtain by plotting just about any parameter such as surface tension or conductivity versus surfactant concentration. Here we plot surface tension in blue. At the start, the value is that of the water. When the plot reaches a plateau, extra surfactant is going into micelles rather than reducing surface tension further.

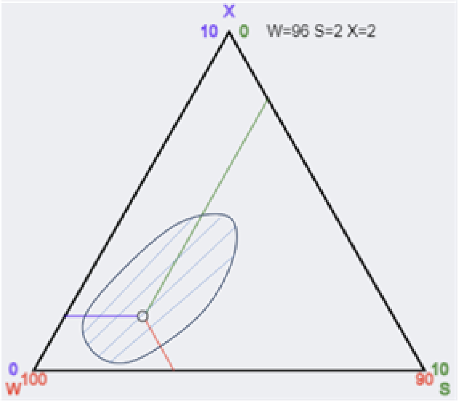
In the orange curve we start with a solution of our desired polymer, and the polymer itself has reduced the starting surface tension. As you add surfactant, at the Critical Association Concentration, CAC, extra surfactant goes into associating with the polymer, so the measured parameter remains constant. The association is in the form of local micelles formed next to the polymer chain. The CAC is *not* dependent on the polymer concentration. As you continue to increase the surfactant concentration, at the saturation point (which *does* depend on polymer concentration) the extra surfactant is free, so you carry on down the curve to the CMC.

Why do you care about this? If you attempt this and find an unchanged CMC, you know there is no significant surfactant/polymer interaction. If you find a reasonable CAC you get a good idea about surfactant polymer interaction and can use data such as the saturation concentration to work out surfactant/polymer interactions. If CAC << CMC then you probably have an undesirable phase separation and it may not be worth exploring the system much further.

How do different types of surfactant interact with different types of polymers? We can answer this in a simple table:

Surfactant	Polymer	CAC?	Comment
Ionic	Opposite charge	CAC << CMC	Coacervate
Ionic	Polar nonionic	CAC < CMC	Usable
Nonionic	Polar or charged	No effect	No effect
Zwitterionic	Polar or charged	No effect	No effect
All	Hydrophobic	CAC < CMC	Usable

Phase diagrams



The sorts of phase diagrams of interest to us are really small corners of the full triangle. Here we go from 100% to 90% water and 0% to 10% of surfactant S and polymer X. The (associative) tie lines are in the range of 98% to 94% water and just a few % each of S and X. So when in the following we talk about single phases we mean “single phase in the region of interest”. Outside this region, where we might have 20% surfactant, we are likely to have a single phase because the effects of interest here are swamped, so there is no need to explore the larger diagram.

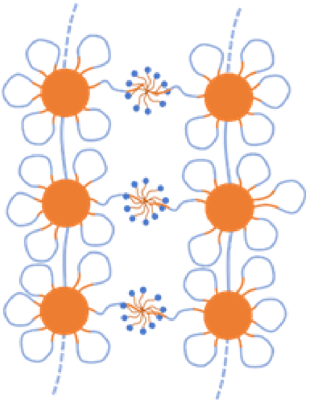
It is a general rule that polymers of different types tend to phase separate. There is usually some enthalpic penalty (“unlike isn’t compatible with unlike”) and entropic gains from mixing are too small to compensate. For the purposes of this discussion, we can call a micelle a polymer, in that it requires some extra entropic boost to promote miscibility with a real polymer.

As an example, an ionic surfactant’s counterions are relatively constrained when ordinary micelles are formed, and are more free when the surfactant associates with a polar nonionic polymer. This extra freedom allows a single phase to be formed. If you add salt which reduces the constraints of the counterions around the micelle, then the entropic gain is smaller and you get segregative phase separation. If, for whatever reasons, you *have* to have a surfactant and a polar nonionic polymer, then avoid extra salt and, if possible, choose a surfactant with a relatively small tendency to produce large micelles.

You *can* get some entropic gains from oppositely charged surfactant and polymer when the charge densities aren’t too high, so CAC < CMC. But stronger interactions give associative phase separation. This takes us back to coacervates, and we’ll discuss some more once we’ve finished this overview.

When surfactant and polymer are the same charge, there’s no chance for entropic gain, so there’s phase separation.

Similarly, nonionic surfactants and polar polymers have no reason for entropic gain, so there’s phase separation.



This leaves us with one useful category: surfactants of any type with hydrophobic, or hydrophobically modified, polymers. Because the system is driven by hydrophobic interactions, there is plenty of room for compatibility. Indeed, if you take a typical HEUR-like associative thickener (see Thickeners_Associative Thickeners) and start adding surfactant, the viscosity can increase. The micelles can start linking up the HEUR in new ways. But too much of a good thing is always bad. Although viscosity increases for a while, the micelles start interfering with the HEUR mechanism, so the viscosity starts to decrease.

The way out of this problem is to have micelles that themselves can tangle. A combination of wormlike micelles and HEUR-like systems can give higher viscosities than either system on its own.

Back to coacervates

Here we show three possible coacervate zones. Assuming that we want a phase separation (for example to deposit a hair conditioner) how do we organise the system to deposit in one of our chosen zones?

The fancy answer is that we do systematic phase diagram explorations checking for turbidity over a range of different WXS ratios, shown with the blue dots. We do these while varying ratios of key parameters such as:

- Surfactant tail length
- Charge concentration on the polymer
- Salt concentrations

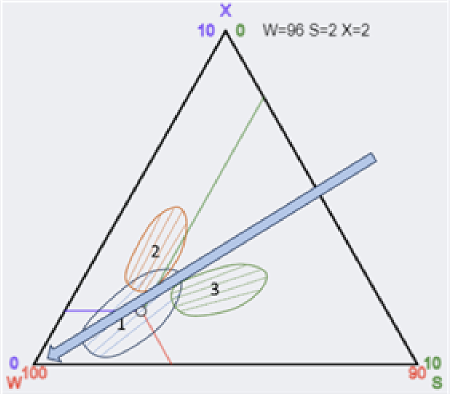
Even with that sparse blue dot matrix, that's 45 experiments per variant, so with "hi" and "lo" values for each variant we're up to 270 experiments. In practice, most of us aren't going to do this.

The preceding sections provide some heuristics for some minimum experiments to reach a satisfactory outcome.

Increasing the surfactant tail length makes both the CMC and the CAC lower. Because this is a coacervate CAC, maybe 2 orders of magnitude smaller, that means a shift in the direction of 2.

A higher charge concentration leads to stronger attraction of the surfactant. This will shift one edge in the direction of 2 but will also extend the recovery in the direction of 3. There will also be an extension of the phase separated zone towards the X direction.

Although salts can encourage phase separation for ionics + neutrals by reducing entropic gains, they also reduce the attractions of the oppositely charged surfactant and polymer for coacervates. This will lead to a smaller 2-phase zone, somewhere between 1, 2 and 3.



These intuitions will not be precise, but they are still helpful. We know that we will be diluting our formulation from somewhere outside the small subsection of the ternary diagram and that it will reach the 100% water corner. If we happened to have formulated with a coacervate in regions 2 or 3, then our formulation will fail to deliver whatever package we intended. Instead of a the careful analysis of the phase diagram via 270 experiments, the idea would be to identify potential dilution curves and test for the appearance of coacervates. This has the advantage that the experiments themselves are much simpler – start at some interesting concentrations of salt, potential surfactants, potential polymers and measure turbidity during titration with water. With one or two formulations identified as interesting, add a few well-chosen blue dot WXS experiments to be confident that there is sufficient margin of error in the dilution curve. Region 1 in the diagram is robust against modest starting-point errors. If you had a success from just touching Region3, then the extra experiments would confirm that this would not be a reliable formulation strategy.

It's not easy

The first draft of this chapter emphasised that there are too few usable tools to help us formulate these complex systems. The extra insights in the extended version still don’t provide the sorts of tools that could make your life much easier. Instead, the hope is that with some extra logical background, it will be somewhat easier to think through why things go wrong or go right and what adjustments might be needed to accommodate changes required by other parts of the formulation.

If any reader has some better ideas they would like to share with the community, I’d be delighted to update the chapter.

Surfactancy CPP and phases

Links

Dishwashing Liquids, Hair Shampoos

Critical Packing Parameter, CPP, is really only relevant for concentrated surfactants in water and salt solutions. In principle it can help explain the various liquid crystal phases we observe in such solutions, though careful academic work shows that the predictive power is not very impressive.

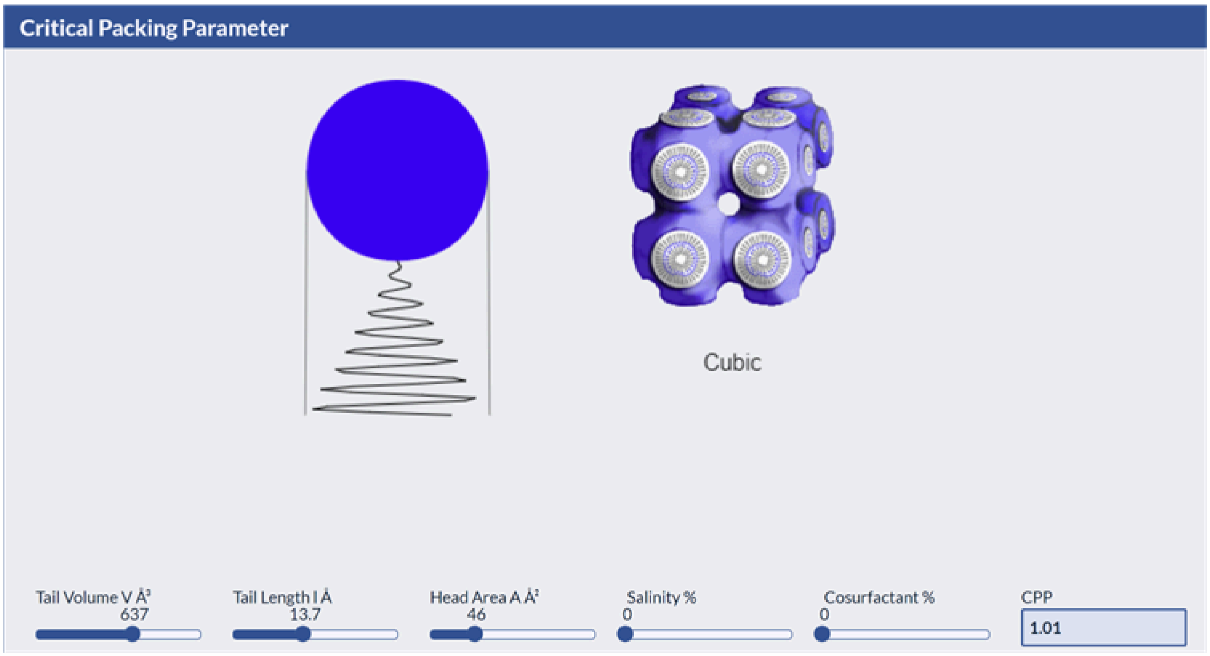
It can be argued that it is useful for understanding salt and cosurfactant effects that can change a surfactant solution from low to (super) high viscosity.

A, V and l

A surfactant molecule has a head area of A, a tail volume of V and a tail length of l. We can then calculate the CPP via:

$$CPP = \frac{V}{Al}$$

If we draw this as a cartoon and happen to have V ~ Al, so CPP~1 we find:



Critical Packing Parameter

Tail Volume V Å³ 637

Tail Length l Å 13.7

Head Area A Å² 46

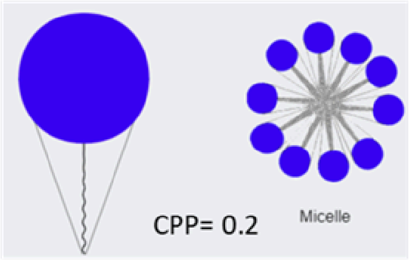
Salinity % 0

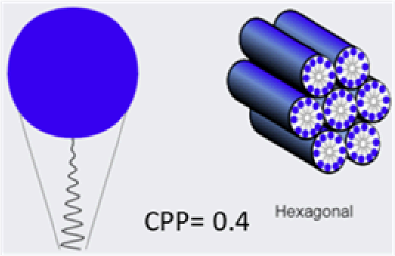
Cosurfactant % 0

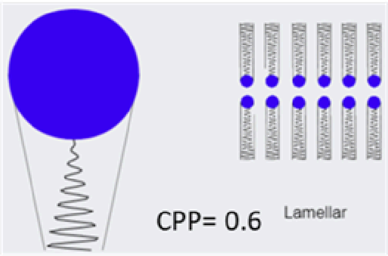
CPP 1.01

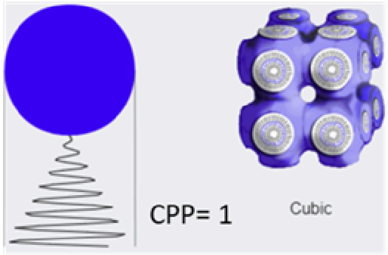
https://www.stevenabbott.co.uk/practical-surfactants/cpp.php

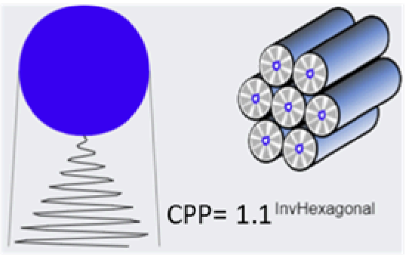
When the shape is “balanced” then the molecules tend to pack in a cubic lattice. A range of examples that you can find in the app (in this case by changing only V) are:

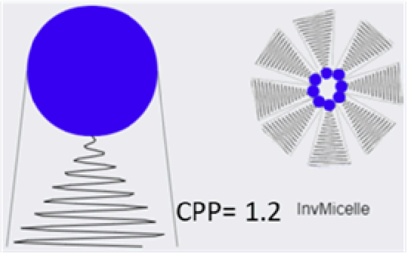












If you keep the same V and l but change A then you can go from spherical micelles to an hexagonal arrangements (highly viscous) then onto lamellae, cubic (again viscous) the inverse phases. This is often done with simple surfactants like SLS or SLES which have relatively large effective head areas because of charge-charge repulsion. As you add NaCl, the charge-charge repulsion decreases, the head area decreases, so CPP increases.

The graphics showing the structures were kindly generated by Seth Lindberg at P&G.

My assertion that CPP is a worthless idea has been regularly challenged. Eventually I had to admit that it had *some* use in thinking through wormlike micelles. More recently I had to double the number of applicable domains from 1 to 2. That's a small number of plausible uses for an idea that features so heavily in surfactant science courses.

The hexagonal structures can elongate into Thickeners_Wormlike MicellesThe highly viscous wormlike phases give the impression of luxury formulations, appreciated by those making Soaps and Washing and Hair Shampoos.

Packing parameters

CPP is all about concentrated surfactants in water. The idea is essentially worthless for understanding emulsion behaviour because we have oil which interacts in unknown ways with the tail regions of different surfactants and, as we know from Surfactancy_HLD, systems have complex behaviours and any surfactant can, in the right circumstances, be o/w or w/o. An alternative approach to HLD uses the idea of Packing Parameters which reflect the current geometry of the system. There is unfortunate confusion between the valid PP approach and the invalid CPP approach to emulsions.

If you want to describe HLD effects by saying that the oil, salt and temperature change the PP, that's fine, but because there is no viable way to predict PP values in real formulations the idea is noted simply for completeness.

Surfactancy Critical Capillary Number

Links

There are at least two cases where a Critical Capillary Number, the ratio of intertial versus surface tension forces, has a big impact on surfactant-related formulations.

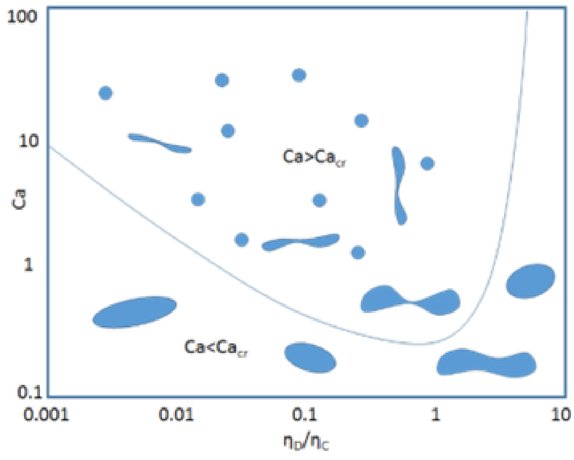
Capillary numbers

Across the whole of fluid flow phenomena the same “dimensionless numbers” keep cropping up. One of these is the Reynolds number, familiar in terms of the transition from laminar to turbulent flow. Another is the capillary number, Ca, which is the ratio of inertial to surface tension forces. For a fluid with velocity U, viscosity η and surface tension γ:

$$Ca = \frac{U\eta}{\gamma}$$

The dimensions of the units (m/s, Pa.s, N/m) cancel out so Ca is dimensionless.

Ca and Emulsification



Many of us have tried, and failed, to make an o/w emulsion when the oil happens to be viscous. [The same logic applies to w/o emulsions but we will focus on o/w]. It’s not that it’s difficult, it’s impossible (at least for normal dispersion methods). The problem is shown via a graph:

The Y-axis is the capillary number required to create an emulsion. The definition of Ca is based on the interfacial tension (IFT), γ, the radius of the desired drops, R, the viscosity of the continuous phase, η_c and a velocity gradient G which depends on your disperser speed.

$$Ca = \frac{\eta_c G R}{\gamma}$$

It is no surprise that you need a certain value of G, and therefore of Ca to be able to break up the oil drops. If the viscosity of the oil, η_d for dispersed phase, is approximately the same as η_c for the water continuous phase then the emulsification is efficient. If your aqueous phase contains some additives that raise its viscosity and the oil stays the same viscosity, you have to work harder to get an emulsion. At a ratio of 0.1 you need relatively almost 10x the capillary number.

The surprise is when you try to disperse a viscous oil. Once it reaches 4x the viscosity of the aqueous phase, the required Ca is infinite – you can’t disperse it.

For those who need to create an emulsion of a viscous oil in water a standard method is to use Surfactancy_Emulsion Inversion.

As we shall see in the next section, an alternative to increasing G by increasing the energy of the disperser, is to decrease the IFT.

Removing oil

Although oil removal is a general problem, a specific example is removing remnants of oil from an oil well, a process called EOR, Extended Oil Recovery. A similar process is removing chemicals from contaminated soil.

If you pump a surfactant solution through the narrow pores which are coated with oil or contaminant, it follows the general rule that low viscosity solutions simply punch a hole through higher viscosity liquids and the results of your efforts are some slightly oily surfactant solution. But if the capillary number is greater than a critical value, the removal becomes efficient.

☐ Horizontal

U mm/min 1

r µm 10

Δρ g/cc 0.1

η cP 10

φ_0 0.9

φ_{hi} 0.2

Log(Crit) -4

n 4

IFT mN/m 4

Ca 4.17e-5

B 2.45e-5

CB 6.62e-5

https://www.stevenabbott.co.uk/practical-surfactants/removal.php

Here, φ is the amount of oil remaining and CB is a combination of capillary and Bond numbers which, for a horizontal process, is simply the capillary numbers – the Bond number captures extra buoyancy forces for vertical extraction.

Because increasing Ca via velocity or viscosity requires greatly increased pumping energy, it is more effective to increase Ca by reducing the IFT. The best way to reduce the Surfactancy_Interfacial Tension and Rigidity is via smart use of Surfactancy_HLD. It is no coincidence that much of the scientific work on HLD came from the oil industry.

Surfactancy Dynamic Surface Tension

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, SkinCare, Surface Cleaning, Water-based Adhesive, Emulsion Polymers

> For most formulation purposes, it is not surface tension but *dynamic* surface tension (DST) which matters. Classic surface tension is measured over minutes, formulations require responses in milliseconds or seconds. Being able to measure and understand DST is, therefore, rather important.

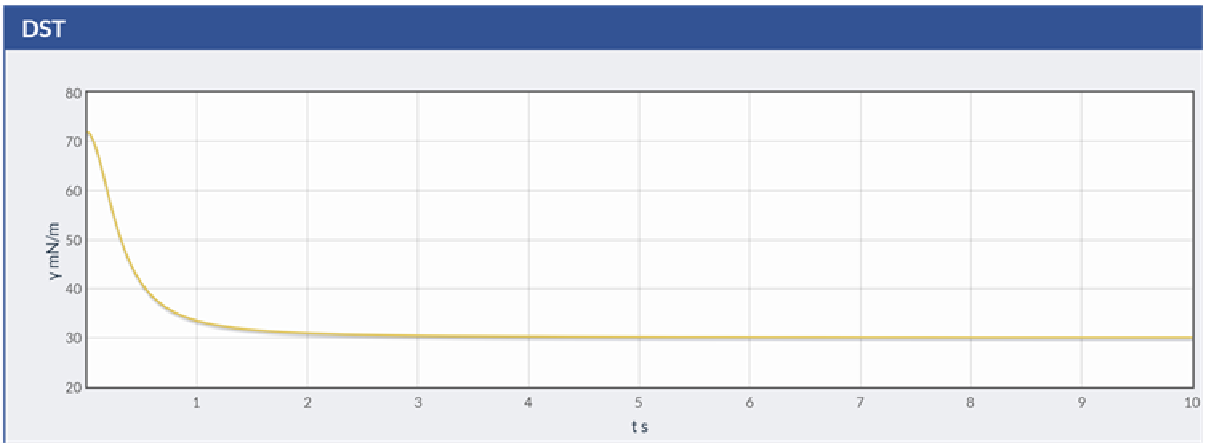
Classic DST

Countless academic papers on DST consider sigmoidal plots like these:

https://www.stevenabbott.co.uk/practical-surfactants/dst.php

Such plots can be analysed via complex theories, yet the results of the analyses seem not to offer many usable insights. This frustration was sensed by Rosen who (as per the app) provided a set of parameters that could at least identify power laws (how sharply the surface tension decreased) and timescales (central point of the rapid decline).

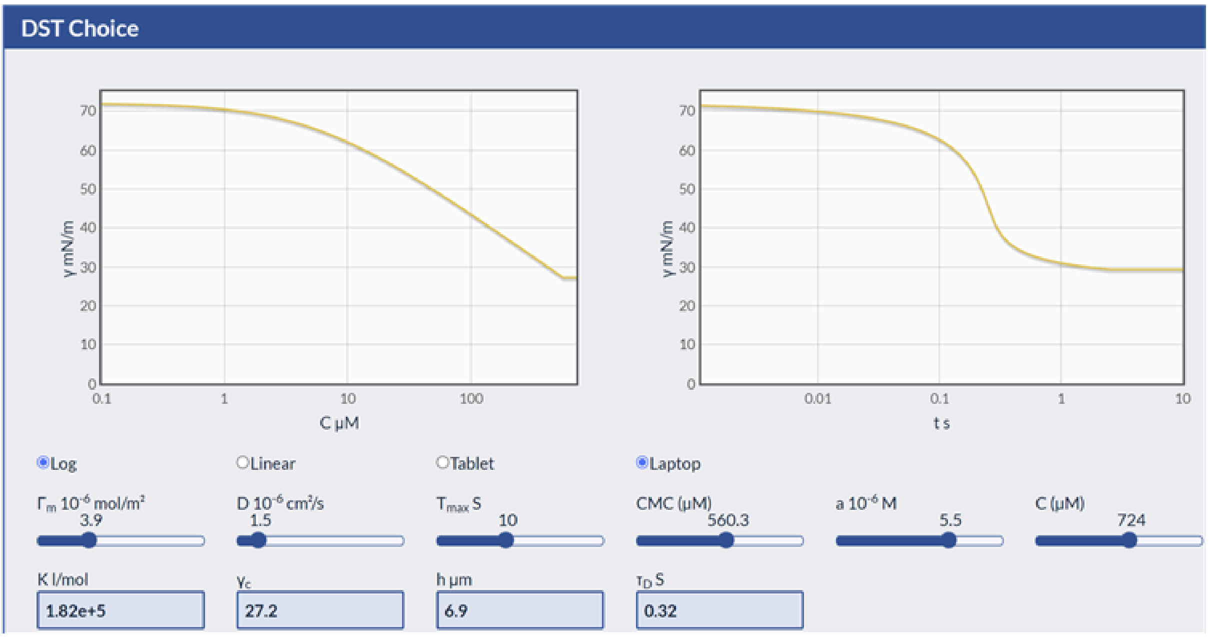
For the formulators this isn't much help. If you use a linear, instead of a log plot with the same data, it doesn't look too interesting:



Changing the X-axis scale to 1 second (not shown) we can see that the surface tension halves in about 0.3s.

If you continue to read the Rosen analysis included with the app, there is a long chain of reasoning which is hard to master. It would be nice to have an alternative view.

DST Choice



● Log ○ Linear ○ Tablet ● Laptop

Γ_m 10^{-6} mol/m² 3.9

D 10^{-6} cm²/s 1.5

T_{max} S 10

CMC (μM) 560.3

a 10^{-6} M 5.5

C (μM) 724

K l/mol 1.82e+5

γ_c 27.2

h μm 6.9

τ_D S 0.32

https://www.stevenabbott.co.uk/practical-surfactants/dst-choice.php

Instead of Rosen's long chain of logic based on plausible parameters, we have a direct calculation based on parameters described in the Surfactancy_CMC and Langmuir chapter. We have Γ_m, and a (which is 1/K from the CMC app) and the CMC. That is the left-hand graph. An additional factor is the diffusion coefficient D which varies surprisingly little over a broad range of surfactants.

From those inputs we can calculate the DST behaviour. The value of τ_D is the characteristic diffusion time which defines when the surface tension starts to fall sharply. This depends on D (of course) and also on h, which is the depth of liquid needed to provide enough surfactant molecules to reach the surface during that time. The larger h, the larger (worse) the τ_D.

$$\tau_D = \frac{h^2}{D}$$

For good DST behaviour we need a small h which is given by the core surfactant parameters and C which is the concentration of free surfactant:

$$h = \frac{\Gamma_m}{a\left(1 + \frac{C}{a}\right)}$$

For a fast DST you need a large C. But if you have a "good" surfactant with a small CMC, C hits a maximum at that point – after that, any extra surfactant goes into creating micelles. So for good DST performance you need a "bad" surfactant with a high CMC.

This explains why coating formulators regularly add "acetylenic" surfactants which, at concentrations in the range of 0.1% (large by normal surfactant standards) produce excellent DST behaviour. And, yes, these have high CMC values.

The calculation of DST with time is done by a "standard" equation from Ward and Tordai. None of the standard references tell you how to implement this complicated differential equation. Fortunately, Prof Paul Stevenson needed to use this equation and his bright graduate student, Xueliang Li, found a relatively easy way to implement it. They then kindly published their method which is, indeed, easy to implement.

It is frustrating that so much potentially usable science is unused because smart academics don't take the trouble to tell the rest of us how to implement their smart theories. And wonderful when you find those who do.

Measuring DST

The standard modern technique is a bubble pressure tensiometer. A bubble of known radius R is created at the end of a capillary. There is a maximum pressure of $P = \frac{2\gamma}{R}$. The bubble can be created over timescales from 10ms upwards, with the maximum pressure, and therefore γ, decreasing as the speed of creation decreases. It is, as always, somewhat more complicated than that, but the principle is simple and direct.

Surfactancy Emulsification

Links

Sun Screens, SkinCare, Water-based Adhesive, Emulsion Polymers, Pharma Formulation

> The language of "macroemulsions", "nanoemulsions" and "microemulsions" is confusing. The first are in the µm size range, the second in the nm range and both are only kinetically stable. The third are also in the nm range but are thermodynamically stable, you can create them by simple mixing.

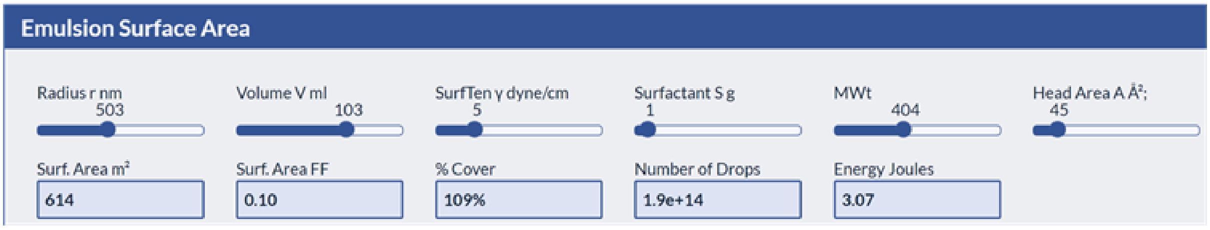
Macroemulsions

Put a disperser into a continuous phase (say water) containing a dispersed phase (say oil) and the oil drops are broken into smaller ones (for an exception when the oil is viscous see Surfactancy_Critical Capillary Number). These drops readily coalesce Surfactancy_Emulsion Coalescence.

Now put a suitable surface active agent into the mix so the interfacial tension (IFT) is reduced and multiple things can happen:

- The driving force for droplets to coalesce (their surface tension) is reduced.
- There can be a charge or steric barrier around the oil drop, reducing the tendency to coalesce.
- There are Marangoni effects that move surfactant molecules to fill any random holes in surfactant coverage, helping to protect the drop.

It remains unclear which of these is the most important. What *is* clear is that the thermodynamic energy required to create the drops is very small, yet the kinetic energy used to create them is very large – the process is massively inefficient. A simple app captures much of this:



Radius r nm 503

Volume V ml 103

SurfTen γ dyne/cm 5

Surfactant S g 1

MWt 404

Head Area A $Å^2$; 45

Surf. Area m^2 614

Surf. Area FF 0.10

% Cover 109%

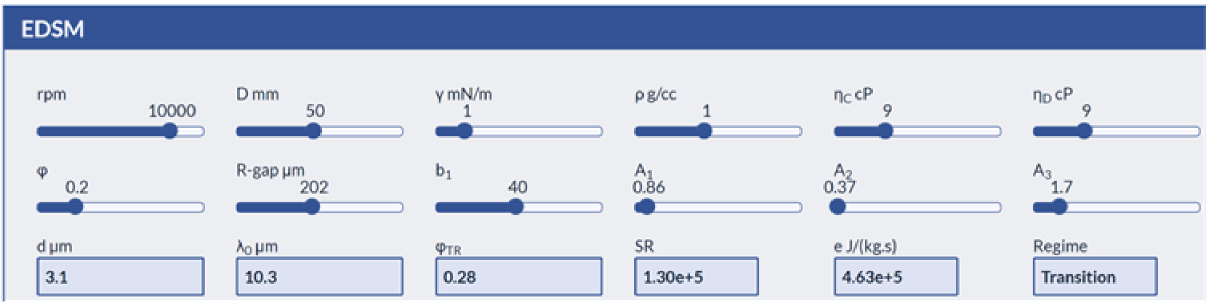
Number of Drops 1.9e+14

Energy Joules 3.07

https://www.stevenabbott.co.uk/practical-surfactants/emusa.php

If the emulsion is 100 ml of 1µm drops with 1g of surfactant of MW 400 and a head area of 45 $Å^2$, then you've created 2.10^{14} drops, with a total surface area of $600m^2$, equivalent (this is a popular unit for emulsions) to 1/10 of a football field, and have covered the drops to 109%. Perfect. The energy needed to create that surface area is 3 J, but you would have used kJ of disperser energy.

To give a feeling of what it takes to create small emulsion drops, the Emulsion Drop Size Model brings together some complex physics to provide some estimates:



rpm 10000
D mm 50
γ mN/m 1
ρ g/cc 1
η_C cP 9
η_D cP 9
φ 0.2
R-gap µm 202
b_1 40
A_1 0.86
A_2 0.37
A_3 1.7
d µm 3.1
λ_0 µm 10.3
φ_{TR} 0.28
SR 1.30e+5
e J/(kg.s) 4.63e+5
Regime Transition

https://www.stevenabbott.co.uk/practical-surfactants/edsm.php

You must read the text of the app to know what all the inputs and outputs mean, but the take-home message is that "it's complicated". So it's not surprising that macro-emulsification remains more of an art than a science. With a key exception.

For those who have an efficient dispersing system and can spend the necessary time to reach equilibrium, the previous model relating surfactant to surface area does a reasonable job of telling you what you can hope for. If, in the example, the surfactant had been present as 0.5 g, giving 50% theoretical coverage, then, by simple arithmetic, the final drop size will be ~ 2µm instead of 1µm.

Choice of surfactant

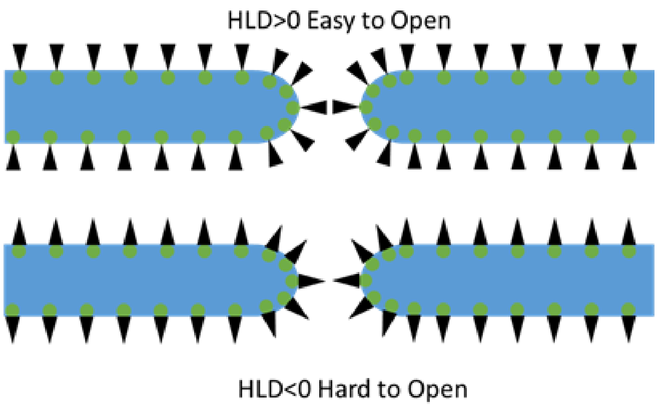
There are a few unhelpful historical "rules" to help choose the surfactant. Bancroft's rule says: "The phase in which an emulsifier is more soluble constitutes the continuous phase." Those who love HLB say that you get o/w with high HLB and w/o with low HLB. As is well known, both rules have so many exceptions as to be worthless.

Using Surfactancy_HLD it is simple. Knowing the temperature, salinity, oiliness of the oil (EACN), you need a surfactant with a characteristic value, Cc that gives an HLD in the range -0.5 to -1 for an o/w or 0.5 to 1 for a w/o. It's as simple as that.

Because we know that HLD values closer to 0 give low IFT, and because generation of drops is *much* easier at low IFT, an efficient way to create an o/w emulsion is to have HLD ~ -0.1. Unfortunately, the low IFT also means that there is strong Surfactancy_Emulsion Coalescence. Hence the trick common to ethoxylates to raise the temperature to a point where HLD ~ -0.1, obtain a fine emulsion relatively easily, then rapidly cool the emulsion to HLD ~ -1, where the emulsion is stable.

The reason that HLD lets us formulate emulsion rationally is that it is a *curvature*-based theory. There is complete agreement that -ve curvature gives o/w and +ve curvature gives w/o; after all the Winsor R value is all about this. What is *not* clear is *why* the curvature has this effect.

The great Langmuir thought he'd spotted the obvious reason – the surfactants at the interface curve in the same way as the emulsion drop. This is obvious but wrong. The curvature is over the nm scale while a 1µm drop is flat when viewed at that scale. An ingenious way to relate nm-scale curvature to µm-scale drops is to focus on the moments when, during their creation, drops have a chance to break up o/w or w/o. That breakup will take place starting with a nm-scale hole. If that curvature is the right way round (leading, say, to an o/w net result) then the hole will open and extend. If it's the wrong way round, it will not open:



In the top pair, the curvature naturally favours the opening from (for simplicity of drawing) a large tail and small head, giving a w/o emulsion (HLD >0). With an HLD < 0, it would be hard to open the water drop, so oil drop openings will be favoured so you get o/w.

This "Kabalnov wedge" theory is controversial, and others might reach similar conclusions via other routes. What's not in doubt is that curvature rules emulsification and the HLD approach to curvature is the most practical tool available to use.

Nanoemulsions

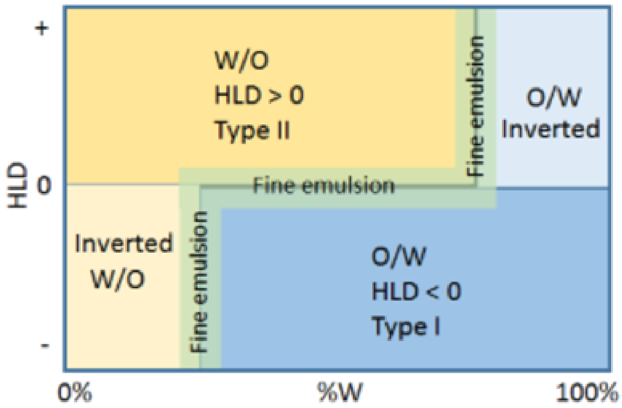
Nanoemulsions have a key advantage that their rate of Surfactancy_Emulsion Creaming and Flocculation is very low, so if they are also well stabilised against Surfactancy_Emulsion Coalescence they can be great emulsions. The three classic problems for nanoemulsions are that:

- They take a lot of energy in the disperser system to create the small drops and, as the Emulsion Surface Area app emphasises;
- They require more surfactant to cover the extra surface area;
- They are prone to Ostwald ripening.

The first reason is why it is popular to make nanoemulsions via clever Surfactancy_Emulsion Inversion schemes where "inversion" can be used loosely.

A classic example is the ethoxylate trick of getting to HLD = 0 via temperature (so there is no necessity to invert by going to HLD > 0) so that very small droplets at super low IFT are easy to make. With a rapid cooling quench to restore HLD to ~ -1, the IFT increases so the barrier to coalescence is raised.

There are similar tricks that can be played with intelligent use of the HLD equation and different EACN values for the oil (start with a low EACN to create HLD ~ 0 then "quench" with a high EACN oil to reach HLD ~ -1), or do a similar trick with a surfactant, starting with a higher Cc then quenching with a low Cc molecule.



The alternative route is to use catastrophic phase inversion where you start with a small amount of water so the system is forced to be w/o then add enough water to create the inversion to the correct o/w state. The reason this works is that *any* flipping of curvature, via HLD or catastrophic phase inversion, goes via a point of low IFT. The smart formulator works out how to do the flipping through knowledge of the Salager diagram that shows both inversion points with the fine emulsions that they can create. This diagram from the https://www.stevenabbott.co.uk/practical-surfactants/inversion.php app makes this point.

There’s nothing you can do about the need for more surfactant – the surface area has to be covered.

The issues around Ostwald ripening are covered in that chapter: Surfactancy_Emulsion Ostwald.

Microemulsions

Creating a microemulsion is easy. Set up the ingredients so that when combined HLD = 0, have enough surfactant to allow all the dispersed phase to be solubilized, and carry out a normal low-energy mixing to allow the microemulsion to form.

A frequent criticism of this strategy is that it requires “too much surfactant”. Yes, a naïve formulation with no attention to the NAC part of HLD-NAC might require too much surfactant. By attending to those aspects of head area and tail length that relate to efficiency and by either adding “linkers” to extend the head and tail lengths or using “extended surfactants” that have a neutral (propylene oxide) middle section, efficient systems can be created. You should also acknowledge that it is easier to create efficient microemulsions with smaller rather than larger EACN values.

A problem in practice is that a new batch of the “same” surfactant or oil might have a different Cc or EACN so your carefully-tuned formulation might not work. That is why a culture of “scans” to measure Cc and EACN values is important for success – you can readily adjust your surfactant or oil blend to accommodate batch-to-batch variations.

What spoils many such schemes is that surfactants can form strange phases with other surfactants. This means that an otherwise perfect blend based on Cc values might fail because the surfactants tie each other up in a nasty phase. It has also been found that although microemulsions are thermodynamically stable, the order of addition can create kinetically stable alternative states that block the formation of the microemulsion.

So this chapter comes with a negative guarantee. I cannot guarantee that you can get a perfect emulsion by following a few basic rules – our state of knowledge of surfactant interactions is too limited. But I can guarantee that if you don’t follow the rules, your chances of success are a lot smaller.

Surfactancy Emulsion Coalescence

Links

Water-based Adhesive, Pharma Formulation

Coalescence, the joining of smaller emulsion drops to make bigger ones has two impacts. The first is obvious, when we see a fine emulsion fall apart into bigger and bigger drops that cream easily. The second is hidden, when the two drops we've just created with our disperser coalesce quickly, requiring another pass of the disperser to get them to re-form.

We can't model the processes exactly, but we can learn a lot from the theory and the app.

Diffusion, Collisions, Coalescence

To coalesce, the drops must collide and to collide they must diffuse through the continuous phase. So we start with the diffusion coefficient, D, for a drop of radius r in a medium of viscosity η and where kT is Boltzmann constant times temperature:

$$D = \frac{kT}{6\pi\eta r}$$

The coalescence rate, CR, when you have *n* particles per unit volume is given by a two-part equation. The first is the collision rate and the second represents the chances of a collision resulting in coalescence, involving the activation energy, E, the barrier over which the drops have to go in order to merge:

$$CR = 8\pi Drn^2 . e^{-\frac{E}{kT}}$$

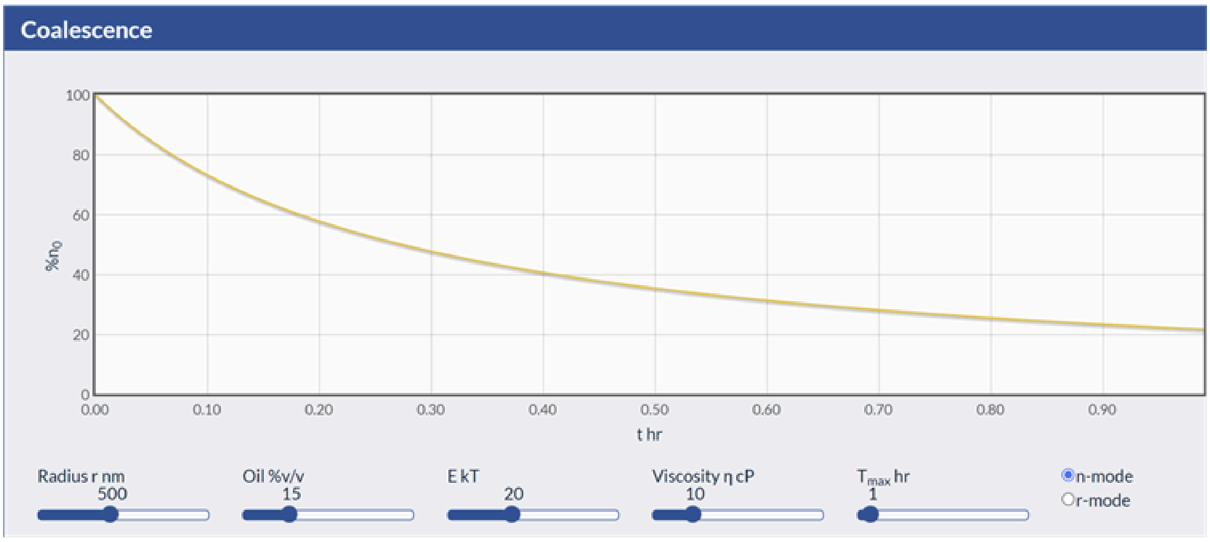
Substituting for D we end up with:

$$CR = \frac{4kT}{3\eta} n^2 e^{-\frac{E}{kT}}$$

Because the radius cancelled out, it looks as though it is not important, but because $n = \frac{V}{V_{drop}}$ there is a $\frac{1}{r^6}$ dependence, one of the many reasons it is harder to make small emulsion drops. From the coalescence rate we can calculate the number of particles at time t, n_t, starting from an original number $n_0$via:

$$\frac{1}{n_t} = \frac{1}{n_0} + t\left(\frac{4kT}{3\eta}\right) exp\left(-\frac{E}{kT}\right)$$

We can then plot $\frac{n_t}{n_0}$ over time:



Radius r nm 500

Oil %v/v 15

E kT 20

Viscosity η cP 10

T_{max} hr 1

● n-mode

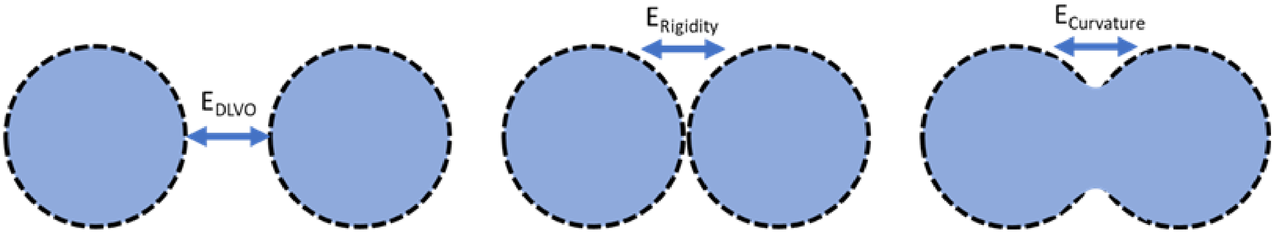
○ r-mode

https://www.stevenabbott.co.uk/practical-surfactants/coalescence.php

In this screen shot, the number of drops decreases to 20% of the original over 1 hour. If you change the radius, the % oil or the viscosity, timescales change somewhat. But change E from 20 to 18 kT and coalescence is almost complete in 20 min.

That is the strength and weakness of the model. It is good that it makes us focus on the size of the activation barrier, but it is hard to know how to engineer the system to ensure a large E.

We can imagine some terms that make up E. To help us we need to look at two drops joining:



There are 3 sorts of barriers:

1. Charge or steric repulsion we can lump under Dispersions_DLVO, which depends on the net charge or the steric repulsion;
2. Interfacial rigidity which depends on the surfactant packing in the interface;
3. An effect around the need to create a fresh curved interface which depends on the IFT – with a low IFT leading to low stability.

Effects 2 and 3 are discussed in Surfactancy_Interfacial Tension and Rigidity

Although we can say that $E = E_{DLVO} + E_{Rigidity} + E_{Curvature}$ and although we can appify many aspects of the above (https://www.stevenabbott.co.uk/practical-surfactants/estability.php), it is more important to understand the principles.

For DLVO stabilisation, a large charge on the surface and a low salt concentration helps for o/w but fails for w/o because (as discussed in the app) charge effects don't work in environments with low dielectric constants. For steric barriers, a large ethoxylate chain can provide the stability for o/w and a large hydrophobic tail provides the stability for w/o.

Interfacial rigidity is relatively little studied. One common observation is that “bad” surfactants, such as polymerics, which have problems of slow kinetics when forming emulsions are “good” once at the interface as they are hard to remove.

Although we tend to think of coalescence as applying only to emulsion drops once formed, very rapid coalescence during emulsification means that you can’t make the emulsion. So it is worth stating here that the IFT effect is also relevant during Surfactancy_Emulsification because, as the “Kabalnov Wedge” shows, the curvature at the narrow bridge when drops are prospectively coalescing might be the one way round to encourage it or the other way to discourage it. If you are trying to create an emulsion and, at some point, had 50:50 o/w and w/o drops, if the curvature encouraged w/o drops to coalesce then the resulting emulsion would be o/w.

Surfactancy Emulsion Creaming and Flocculation

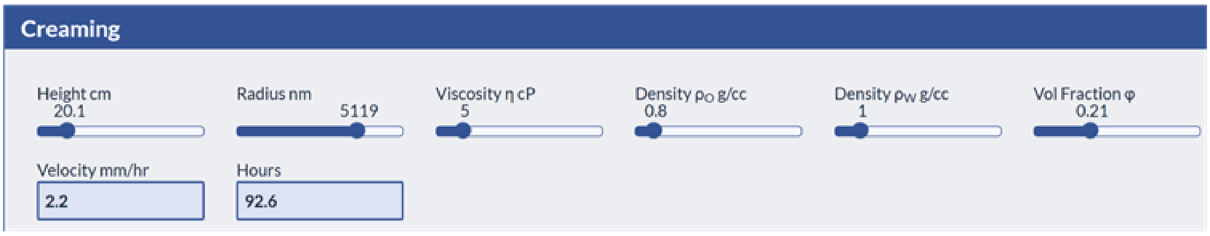
Links

Water-based Adhesive, Pharma Formulation

The rise of oil emulsion drops to the surface of an o/w emulsion is “creaming”. If you are interested in the settling of water drops in a w/o emulsion, just invert the direction.

Although we start with simple creaming, we swiftly move on to flocculation.

Stokes Equation


Creaming

Height cm 20.1

Radius nm 5119

Viscosity η cP 5

Density ρ_O g/cc 0.8

Density ρ_W g/cc 1

Vol Fraction φ 0.21

Velocity mm/hr 2.2

Hours 92.6

https://www.stevenabbott.co.uk/practical-surfactants/creaming.php

Creaming is governed by Stokes equation. The velocity v depends on the radius r, the viscosity η and the density difference between oil and water, Δρ. In addition, the Richardson & Zaki dependence on volume fraction φ is included:

$$v = \frac{2gr^2\Delta\rho}{9\eta}(1-\varphi)^{5.65}$$

To avoid creaming, increase the viscosity, increase the volume fraction of oil and decrease the drop radius. Because the drop radius is likely to increase because of coalescence, this simple equation is only of use to get an idea of what to expect at the start of creaming. If you can reduce v at the start, the side effects of coalescence are also reduced.

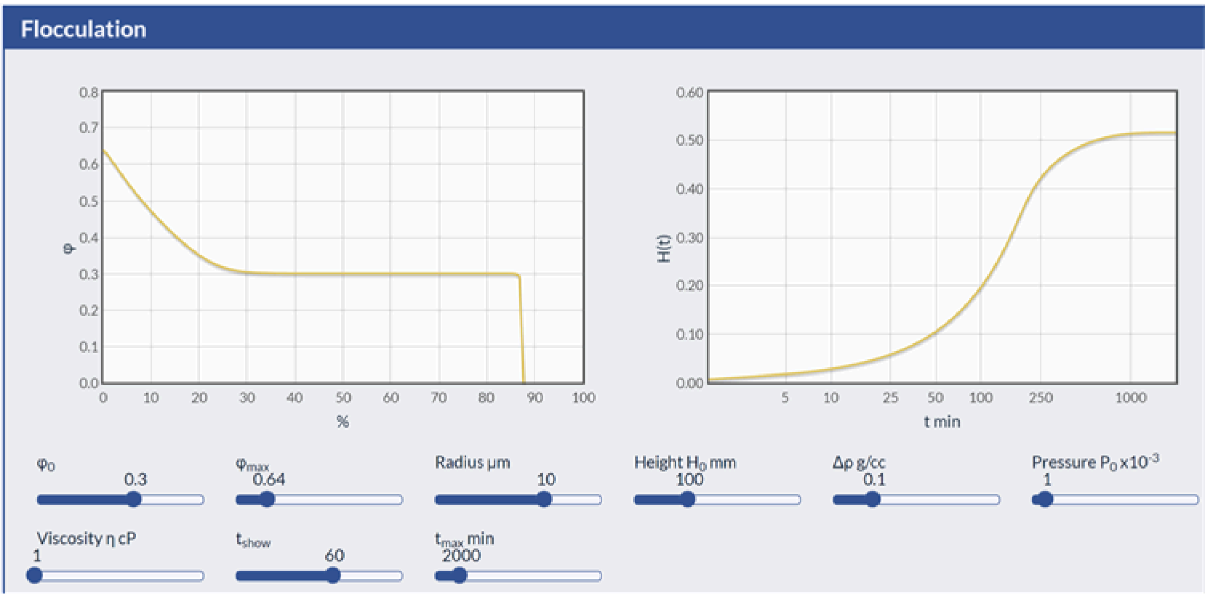
Flocculation

As the drops rise, they become more crowded and can behave either like a “floc”, i.e. a strong assembly of individual drops, or they can coalesce into larger drops. The practical difference in behaviour isn’t significant at this level of simulation – in each case you get drops that want to rise faster, blocked by the steadily-increasing concentration of similar flocs/drops.

As explained in the app text, the complex process can be modelled via a sophisticated technique from U Sofia. It requires two views:

1. The volume fraction of oil from the top (0%) to the bottom (100%) at a chosen time

2. How the height of the clear zone increases over time



φ_0 0.3

φ_{max} 0.64

Radius μm 10

Height H_0 mm 100

Δρ g/cc 0.1

Pressure P_0 x10^{-3} 1

Viscosity η cP 1

t_{show} 60

t_{max} min 2000

https://www.stevenabbott.co.uk/practical-surfactants/flocculation.php

On the left we see that after 60min an originally 30% tube has cleared below 87% of the tube, is unchanged up to 30% and rises to the close-packed maximum of 67% at the surface.

On the right, the clear zone has risen from 0% up from the bottom to 50%, where it is plateauing.

It takes a while to get used to interpreting the plots, but after a while it makes sense. Flocculation *is* complex, so it is better to carry out lots of virtual experiments before trying to interpret real ones.

The algorithm in the paper by Prof Gurkov's team at U Sofia was hard to implement. Prof Gurkov, when asked for help, found the text of their code written in old-fashioned, simple, C for a computer that no longer existed. Fortunately, old-fashioned C is easily translated into modern JavaScript and the app was quickly written.

Readers will have noticed a number of references to the surfactant science group at Sofia. They have contributed massively to our understanding of surfactant formulation and have always been generous with their help when I wanted to appify their work.

Surfactancy Emulsion Inversion

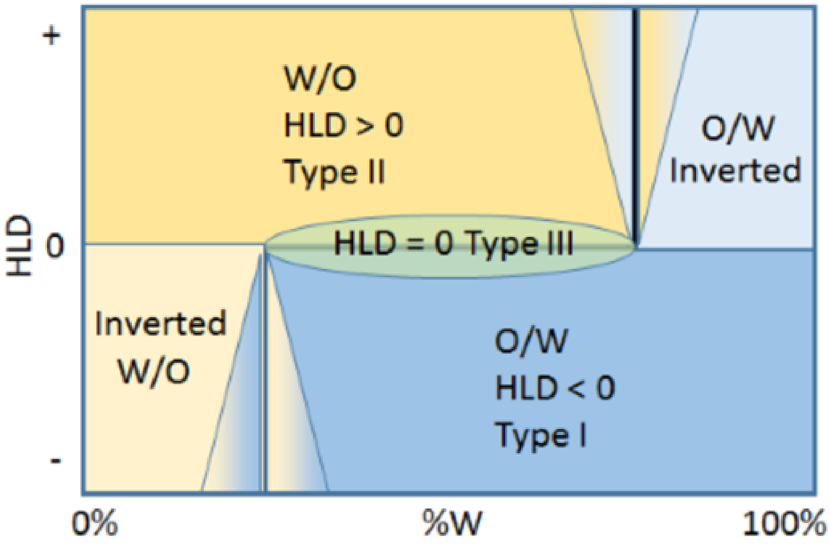
Links

Emulsion Paint, Emulsion Polymers

Emulsion inversion is a powerful technique for producing products that might otherwise be impossible to create. The downside of this power is that it can wreck the formulations of those who don't understand and embrace its power.

From one core diagram it is possible to develop a deep understanding of a phenomenon that causes so much confusion.

Two Types of Inversion



The Y-axis of this diagram from Salager shows that you can invert an emulsion by changing the Surfactancy_HLD around 0. If you start with 50:50 oil to water with a surfactant system with HLD < 0 you will have an o/w emulsion. If you increase the HLD (e.g. for an ethoxylate system, heating it, for an ionic system by adding salt, then you go via a Type III microemulsion into a w/o emulsion when HLD > 0.

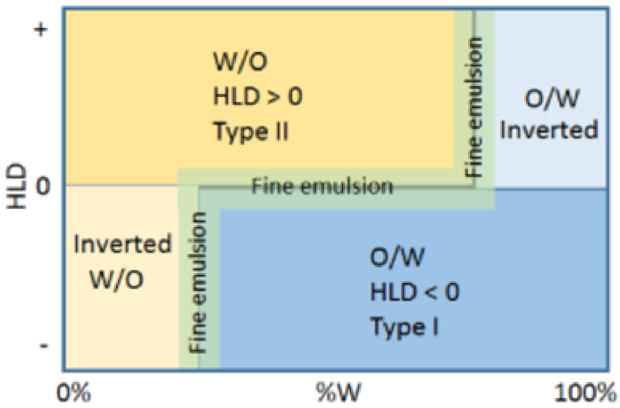
But even if HLD < 0, if you have only a small fraction of water, there isn't enough continuous phase of water to make o/w, so you switch to w/o. Theoretically this should happen around 70% oil (and the inverse, making o/w when HLD > 0, at 70% water) because you are beyond the close-pack limit of oil (or water) droplets.

This effect when there is insufficient continuous phase can be called "catastrophic inversion".

In real life, the point of catastrophic inversion is uncertain as it depends on how far you are from HLD = 0, on the direction of the transition and on how good or bad your disperser is.

Because those who, historically, specialised in catastrophic inversion seemed generally unaware of HLD theory, the literature is very confusing. They would change some parameter and see an unexpected inversion – unexpected to those who were unaware that that parameter (a different oil, change of salinity, change of temperature, change of surfactant blend) might have changed the HLD.

Low IFT



It is well-known that at HLD = 0 it is very easy to create a fine emulsion, because the interfacial tension, IFT, is very low. Such an emulsion is also unstable, because of the low IFT, so you start your emulsification at this HLD = 0 point then do something (e.g. cool for o/w emulsions via ethoxylates) to take HLD away from 0, allowing a kinetically stable emulsion.

Because IFT is controlled by curvature, it is also low at the catastrophic inversion point, where the curvature suddenly flips. So emulsification is very fast, and then requires you to keep going beyond the inversion point to enter a stable domain with a higher IFT.

Accidental inversion in production

In the lab we can add all our ingredients in whichever order we prefer and start our disperser, so we tend to get the results we expect. But in production the stirrer might be in the bottom 1/3 of the vessel and the order of addition of ingredients may be fixed by different constraints. So it is possible to start up the stirrer with the wrong oil-to-water ratio. If that happens to be in the right zone to create an emulsion that is the inverse of what is required, and if the stirrer starts up rapidly, you might create an inverted emulsion. It can be very hard to re-invert an emulsion, especially if the mixing/stirring is not very efficient.

This gives a rule for the starting of mixing in production. Know the oil-to-water ratio, know the HLD, know which way round is right and which is wrong. If at startup you are already the right way round, you can start your stirrer at full speed while adding the other ingredients. If you are the wrong way round, then start the stirrer very slowly and keep it slow till you are safely in the correct zone.

After a long day of teaching surfactant science theory there was an “open questions” session. A question appeared about rules of starting production with fast or slow mixing each giving different problems for different products. Total panic – how to answer such an obscure question. A wise co-trainer, seeing my panic, whispered “The inversion Z-diagram”. That’s why this “accidental inversion” section appears in the FST. You can’t know everything, you have to be “out there” learning from the real world and, if you are lucky and have a wise colleague, you, and others, learn new things.

Emulsions of viscous oils

In the Surfactancy_Critical Capillary Number chapter you find that it is impossible (at least for conventional dispersers) to create a good emulsion if the viscosity of the dispersed phase is more than 4x that of the continuous phase.

If you need an o/w emulsion with a viscous oil, start with the easy problem of creating a w/o emulsion using your desired emulsifier which is good for o/w, but with a small fraction of your total water. You will be able to make a w/o inverted emulsion. As you now increase the amount of water you reach the catastrophic inversion point where IFT = 0, and you immediately get a fine o/w emulsion. Flood that with more water to reach your desired ratio with oil and the emulsion is stable.

The moment of inversion is obvious. Because the oil is highly viscous, the w/o emulsion is also viscous. But at the inversion point, water becomes the continuous phase so viscosity reduces drastically.

In some industries this way of making emulsions is routine. To those who have never heard of it, it is an awesome way to make a seemingly impossible product.

Theory is fine, but it’s scary if you’ve never needed to use it. A company asked me to help them achieve an “impossible” emulsion via inversion, in a system I knew nothing about. They also knew that I’d never worked on a real inversion project. The Salager theory is so beautiful that I was able to make some predictions that seemed reasonable.

Failure of predictions in complex systems is normal – real life doesn’t always follow simple models. You learn from the failure. But when, as happened in this case, the predictions work out OK, it’s most satisfying.

Surfactancy Emulsion Ostwald

Links

Water-based Adhesive, Pharma Formulation

> If you make a fine emulsion you often find that the few larger drops (see Size Distribution below) get larger while the small get smaller. This effect, which regularly spoils otherwise great emulsions is called Ostwald Ripening. A few tricks can help reduce the damage that Ostwald can cause.

The big get bigger

https://www.stevenabbott.co.uk/practical-surfactants/ostwald.php

The app shows that if the “average” size (see Size Distribution below) starts at a radius, r_0, of 250nm then after 24 hours this particular emulsion will have droplets over 800nm.

The “pressure” inside a drop is $\frac{2\gamma}{r}$, where γ is the interfacial tension, IFT. If some of the oil inside the drop happens to migrate to a drop with a larger r, the overall energy of the system decreases. The original drop now has an even smaller r, to the relative driving force increases. It’s a one-way journey, the big get bigger and the small get smaller … till they disappear under the super-high pressure once they get into the low nm range.

The rate at which this can happen depends on how easy it is for a molecule to leave one drop and migrate to the bigger one. If the solubility of the oil, c, is high, then this process is easy. So one rule for avoiding Ostwald ripening is to use insoluble oils. Hexadecane emulsions (we will return soon to hexadecane) do not change size over time. Obviously the time taken also depends on the diffusion coefficient, D – the higher it is, the faster the ripening.

The equation captures these factors:

$$r_t^3 - r_0^3 = \frac{8\gamma DcVt}{9RT}$$

V is the molar volume of the oil and we have RT, gas constant time temperature on the bottom. In the app, T is assumed to be 298°K.

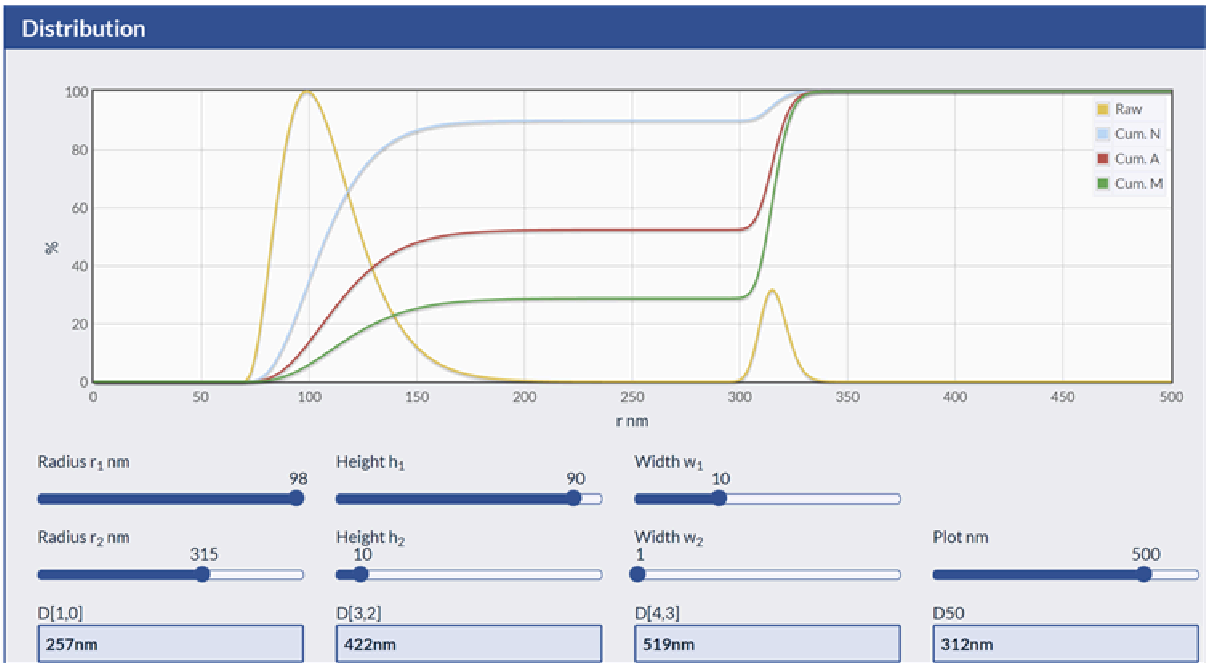
As expected from the driving force, decreasing γ, having a better surfactant, decreases the speed of ripening. The other reason for having a good surfactant, which spends its time efficiently at the interface, is that there is little risk of surfactant micelles carrying the oil between droplets.

The Hexadecane trick

As mentioned above, hexadecane has a such a low solubility in water that there is no drop-to-drop transmission, so there is no ripening. The trick is that if you add a small amount of hexadecane to your real oil, which might have a higher solubility and unacceptable ripening speeds, the ripening starts … then stops. What happens is that the hexadecane concentration builds up inside the smaller drops as the oil leaves, and this concentration difference between smaller and larger drops gives a chemical potential difference (some call it an osmotic pressure difference) that balances the surface energy difference, at which point the ripening process stops.

If you have water-in-oil emulsions then a little bit of salt in the water has the same effect in stopping Ostwald.

Size distribution



Radius r_1 nm 98

Height h_1 90

Width w_1 10

Radius r_2 nm 315

Height h_2 10

Width w_2 1

Plot nm 500

D[1,0] 257nm

D[3,2] 422nm

D[4,3] 519nm

D50 312nm

https://www.stevenabbott.co.uk/practical-surfactants/distribution.php

Discussion about Ostwald requires analysis of the emulsion drop size. The “average” size, mentioned above, is ill-defined. Looking at the app, the size distribution has at least 4 numbers, each of which is valid and, in the right circumstances, meaningful. D[1,0] is the “number average” diameter, based on counting the total number of drops. D[3,2] is the “area average” and D[4,3] is the “volume average”. Although there are a small number of larger drops, because volume is $\frac{4}{3}\pi r^3$, each drop contains much more oil than the smaller drops, so the average in terms of volume or weight is, in this case, 2x larger than the number average. D50 is the size where half the drops have a mass less than it and half more.

This digression into particle size is important for discussions of Ostwald. You can't understand the phenomenon unless you:

- Have a good measure of particle size (most likely an analytical centrifuge);
- Understand how to interpret the particle size data.

The ripening equation describes the "average" radius, but which one. The r^3 term in the equation tells us that it must be the number average radius because the effect depends on its cube.

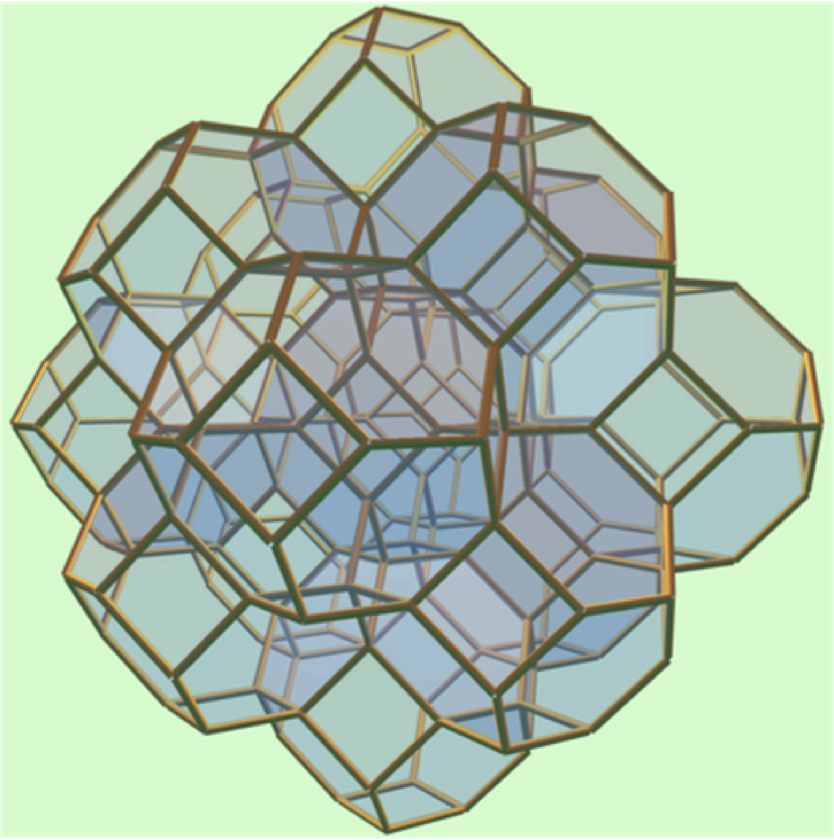
Surfactancy Foam Drainage

Links

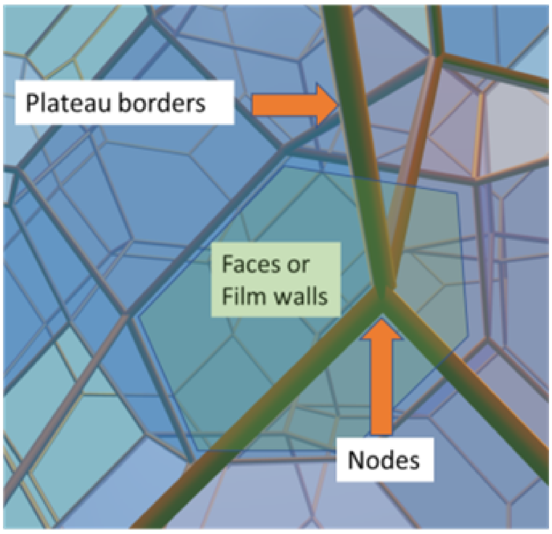
Dishwashing Liquids, Hair Shampoos, Soaps and Washing

Foams can easily fail by drainage – all the water runs out of the foam which then collapses. How this happens, and how to avoid it, is non-intuitive but makes sense once explained.

Where is the water that drains?



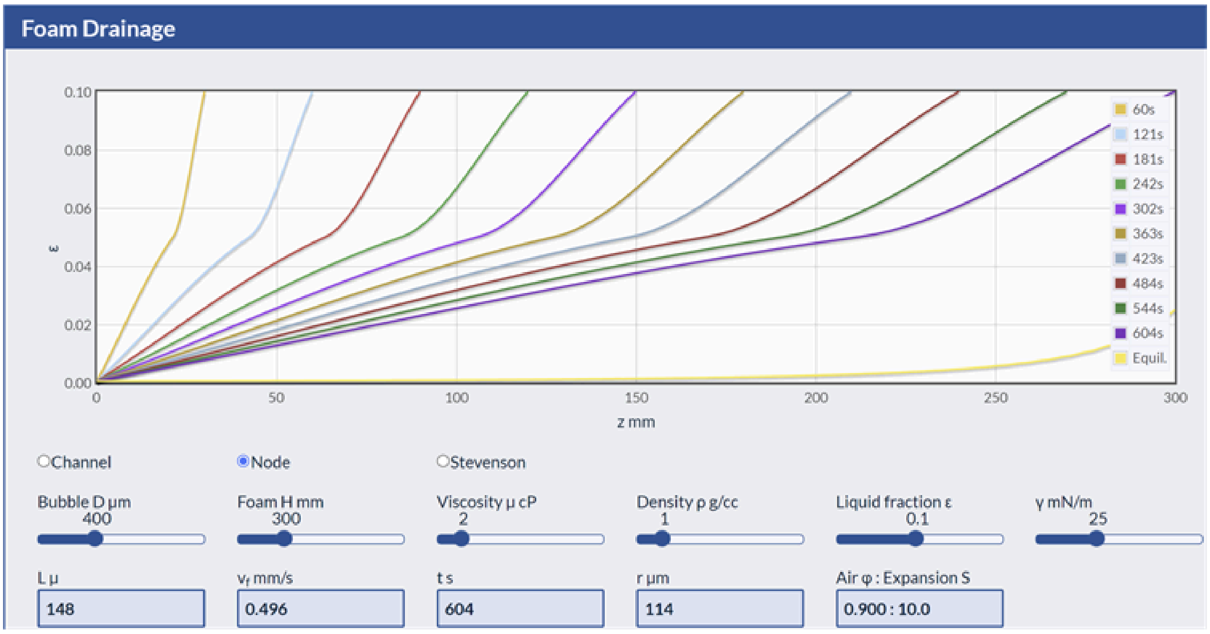
This image from https://www.stevenabbott.co.uk/practical-surfactants/Foam-3D.php gives an exaggerated view of the Plateau borders at the edge of each foam wall and of the nodes where the borders meet.



We can see this in more detail from this close-up. The key fact is that although the *volume* of the foam is made up by the faces/walls, the amount of water in those walls is <1% of the total. So when we worry about drainage, we need to worry about water running down drainpipes (the borders) that intersect at the nodes.

Yes, once a lot of water has drained, then the walls start to drain and collapse, but our focus should be on the main drainage. Although the precise choice of surfactant might affect the "disjoining pressure" (see https://www.stevenabbott.co.uk/practical-surfactants/foam-dlvo.php), that keeps the walls from collapsing, because the drainpipe drainage really controls the foam stability, it is a surprise to find that the surfactant has relatively little to do with stability.

What controls drainage?



○Channel ●Node ○Stevenson

Bubble D µm 400

Foam H mm 300

Viscosity µ cP 2

Density ρ g/cc 1

Liquid fraction ε 0.1

γ mN/m 25

L µ 148

v_f mm/s 0.496

t s 604

r µm 114

Air φ : Expansion S 0.900 : 10.0

https://www.stevenabbott.co.uk/practical-surfactants/foam-drainage.php

Unlike most apps, the lines in this graph never change (except when you switch modes, explained below). When you change the key parameters, the change is in the times recorded. Also, because the theory behind the calculations is yricky, we delay discussing it until we find out what is and is not important for good drainage behaviour.

First, we need to know what the graph means. Let's look at the 60s line. It's saying that although the foam started with 0.1 fraction of liquid at the top of the tube, z=0, that wet front has sunk down to ~25mm from the top, so that the foam from 25mm upwards has steadily decreasing amounts of water. After 2 min, the wet front has fallen to 60mm from the top, and so forth.

How can we increase the time for the front to move downwards, i.e. to reduce drainage? As hinted above, we don't control the time via the surfactant. Instead we control it via the liquid fraction, ε, the viscosity, µ, and the bubble diameter, D:

- Decreasing ε by a factor of 2 increases the time by √2 (for channel mode, below, this is 2x)
- Increasing the viscosity μ by a factor of 2, increases the time 2x
- Decreasing the bubble diameter, D, by a factor of 2, increases the time 4x

Generally you make a foam with the wetness (ε) required for your purpose and your viscosity is similarly dictated by other factors in your formulation. The D^2 dependence tells us that for a long-lasting foam, you should focus on creating small bubbles. See Surfactancy_Foaming where you might discover that we are often bad at focussing on this key parameter for foams.

Before discussing how drainage happens, we need to check one more factor.

Surfactant effects

The nature of the surfactant doesn't feature in the above list. Certainly the surface tension of the foam makes no difference except to the uninteresting final curve which is the equilibrium one.

But there is one option in the app which *is* dependent on the surfactant. The plot shown is in "Node" mode where the walls of the channels are so mobile that they are essentially liquid, so water descends in "plug" flow and overall flow gets limited by the flow capacity of the nodes. If you switch to "Channel" mode, the walls behave like classic pipes with viscous drag. This means that the drainage is generally slower.

A rigid surfactant or a floppy surfactant with a classic rigidifying additive such as myristic acid is therefore preferred if you want to reduce drainage and have a tougher foam wall that will resist the effects of drainage.

In the app, the differences can be hard to spot because they have different dependencies on ε. This means that at large ε values, channel flow can be a bit faster than node mode.

The theory

The theory itself turns out to be simple in concept, but had to be developed over many years by different groups to model the whole process as everything depends on everything as the drainage continues.

The drainage literature is confusing, partly because the problem is difficult, partly because different academic groups have their favourite ways of addressing the issues, using different notations and theoretical approaches. But sometimes you get lucky. An email to Prof Saint-Jalmes, who had a theory somewhat different from other approaches, resulted in a generous explanation of how the theories were complementary and how his approach was especially convenient for creating an app that related nicely to his team's extensive experimental data.

Here we just look at the core theory. The diameter of the pipes is directly related to the size of the foam bubbles. If you halve the bubble size, you halve the diameter of the pipes. Although flow down a pipe goes as $\frac{1}{r^3}$, halving the bubble size also decreases the pipe's length by a factor of 2, so the net effect is a $\frac{1}{r^2}$ effect. The common theories concentrate on the pipe length L which conveniently is given by L = D/2.7. We want a small L because over a given height of foam, the smooth flow through the pipe is interrupted more often.

So our core equations for the velocity of the drainage front, v_f, are, depending on density ρ and gravity g:

$$Channel:\ v_f = \frac{0.007\rho g L^2 \varepsilon}{\mu}$$

And

$$Node:\ v_f = \frac{0.002\rho g L^2 \sqrt{\varepsilon}}{\mu}$$

Again, the details don't much matter and there's not much you can do with μ nor with ε (which changes as the foam drains, making the calculations more complex), so for stable foams you probably want a stiffer surfactant system with channel-dominated flow, but mostly you should focus on getting smaller bubbles which give the smaller length L, with more interruptions to smooth flow. It's as simple as that.

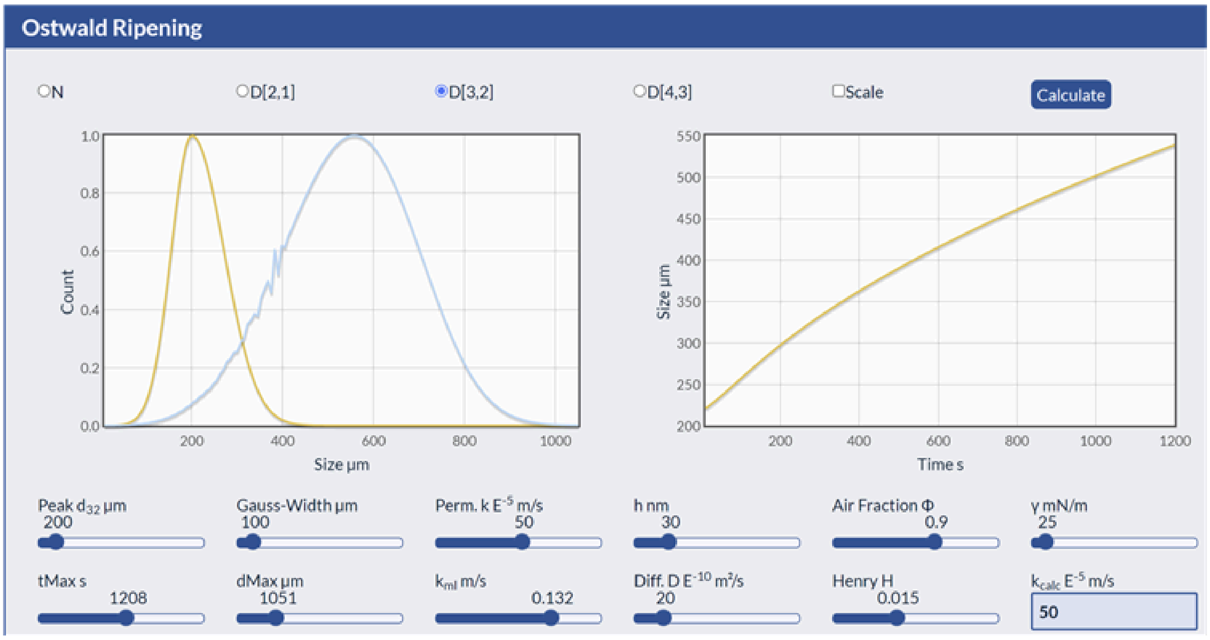
Surfactancy Foam Ostwald

Links

[Dishwashing Liquids], [Hair Shampoos], [Soaps and Washing]

> We often see foams decaying via a decrease in the number of small bubbles as the large bubbles get larger. This is Ostwald ripening, and the ways to tackle it are limited but clear.

Changing bubble size distribution



Peak d_{32} µm 200
Gauss-Width µm 100
Perm. k E^{-5} m/s 50
h nm 30
Air Fraction Φ 0.9
γ mN/m 25
tMax s 1208
dMax µm 1051
k_{ml} m/s 0.132
Diff. D E^{-10} m²/s 20
Henry H 0.015
k_{calc} E^{-5} m/s 50

https://www.stevenabbott.co.uk/practical-surfactants/foam-ostwald.php

The chapter on Surfactancy_Emulsion Ostwald uses a simple Ostwald formulation. Here we can take advantage of the sophisticated work of the team at U Sofia who provided a complex algorithm and much help to look at a more realistic case where there is a bubble size distribution.

The basics of Ostwald are simple. The pressure inside a bubble is given by $\frac{2\gamma}{r}$, so smaller bubbles have a larger pressure. If some of the gas in the small bubble moves to a larger bubble, the larger bubble gets larger and its pressure decreases. The next effect is that gas continues to move from smaller to larger bubbles.

In a foam the process is complicated by the fact that as the bubbles change size, they pack differently so there are kinetic barriers to the process happening.

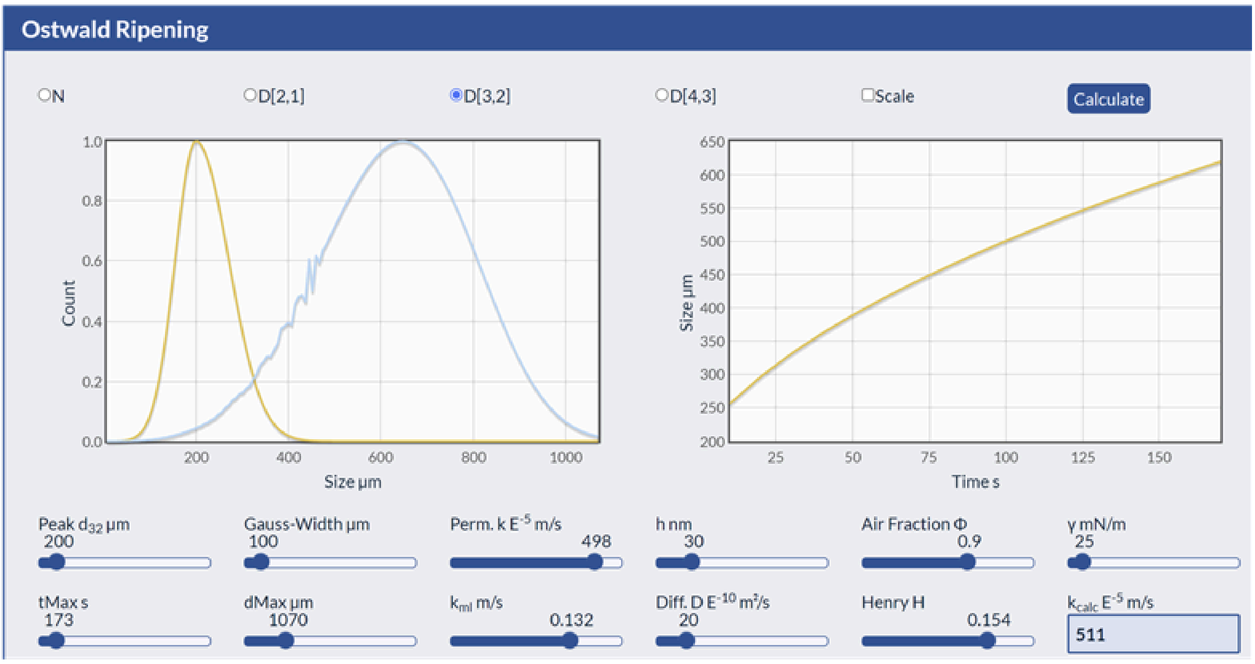
The app gives you, in the left-hand graph how the bubble size distribution changes over your specified time period while the one on the right shows the rise of the “average” bubble size. The word average is in quotes because there are various definitions of average depending on number, area or volume averages, as described below.

You define the starting distribution with a peak value and a starting width which produces a skewed Gaussian.

Slowing Ostwald

The driving force is internal pressure. So a better surfactant, with a lower surface tension, will give a slower growth. As you can't change these values by more than a factor of 2, the effect on Ostwald is relatively small.

The ability for a gas molecule to move from a smaller to a larger bubble depends on the solubility of the gas in the water. CO2 has a much higher solubility than air, so CO2 foams (beer, espresso crema) ripen very quickly. We can simulate this via the Permeability, k which is 50 in the screen shot given that we are using air with a Henry constant ~ 0.015. If we add CO2 so that Henry becomes 10x larger, then k becomes 500 then we find that we get the same degree or ripening in 170s rather than 1200s.



○N ○D[2,1] ●D[3,2] ○D[4,3] ☐Scale

Peak d_{32} µm 200

Gauss-Width µm 100

Perm. k E^{-5} m/s 498

h nm 30

Air Fraction Φ 0.9

γ mN/m 25

tMax s 173

dMax µm 1070

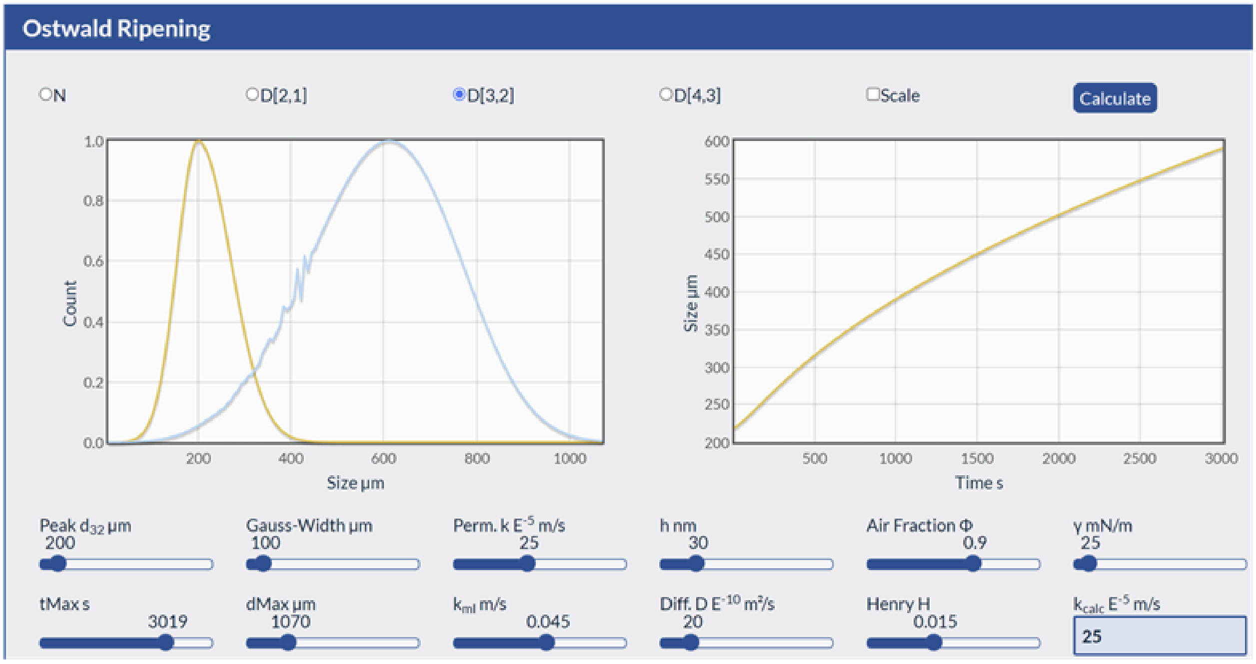
k_{ml} m/s 0.132

Diff. D E^{-10} m²/s 20

Henry H 0.154

k_{calc} E^{-5} m/s 511

But if you can't change your gas from, say, air, how can you slow things down? The key is to reduce the permeability of the monolayer itself, k_{ml}. A factor of 2 decreases k by a factor of 2 and it now takes 3000 s to reach the same increase in size:



Peak d_{32} µm 200

Gauss-Width µm 100

Perm. k E^{-5} m/s 25

h nm 30

Air Fraction Φ 0.9

γ mN/m 25

tMax s 3019

dMax µm 1070

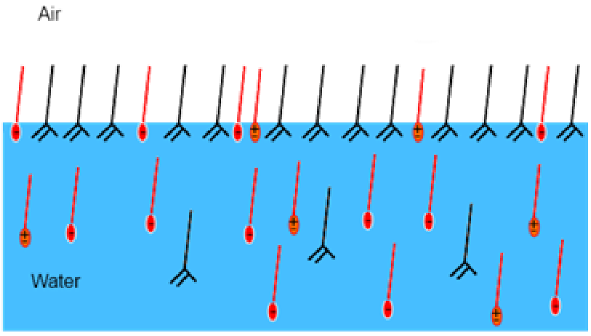
k_{ml} m/s 0.045

Diff. D E^{-10} m²/s 20

Henry H 0.015

k_{calc} E^{-5} m/s 25

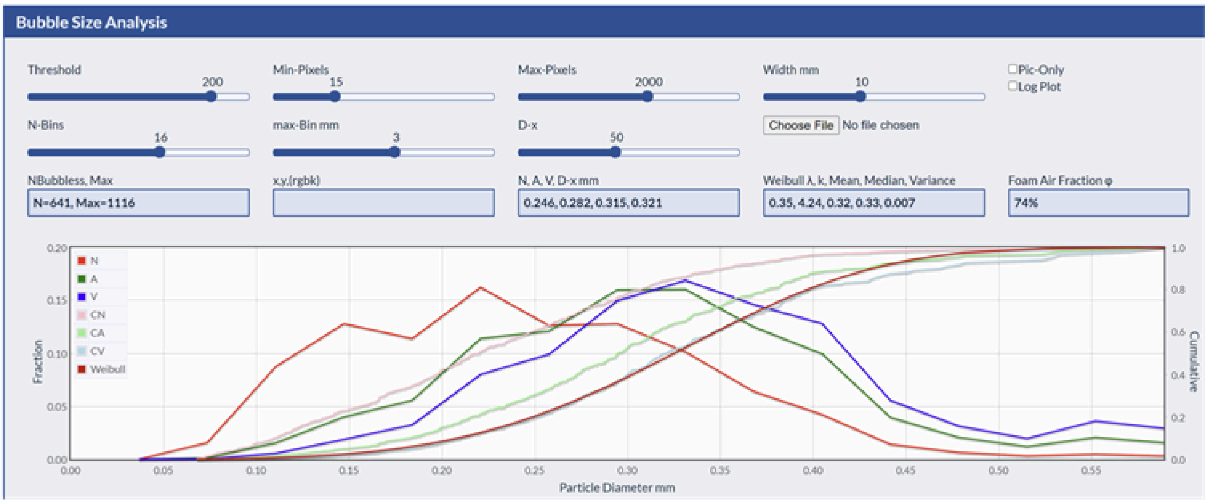
How do we reduce k_{ml}? This depends on the stiffness of the surfactant layer. The standard trick used across the personal care industry is to add some myristic acid (or similar) to the formulation.



The black myristic acid is useless at forming the foam, that's the job of the red surfactant which might be SLS. As soon as the foam is formed, the highly insoluble myristic acid prefers the interface and, because it's neutral, packs more tightly than the anionic surfactant. The foam elasticity (https://www.stevenabbott.co.uk/practical-surfactants/foam-elasticity.php) increases and the ability of gas molecules to permeate through the foam wall decreases.

Bubble size

Using a method described in the text of the app, (it needs a 90° prism, an LED light and your smartphone camera) it's not so hard to get a good picture of your bubbles. Load your image into the app and you get the bubble size distribution.



Threshold 200

Min-Pixels 15

Max-Pixels 2000

Width mm 10

N-Bins 16

max-Bin mm 3

D-x 50

NBubbless, Max: N=641, Max=1116

x,y,(rgbk):

N, A, V, D-x mm: 0.246, 0.282, 0.315, 0.321

Weibull λ, k, Mean, Median, Variance: 0.35, 4.24, 0.32, 0.33, 0.007

Foam Air Fraction φ: 74%

https://www.stevenabbott.co.uk/practical-surfactants/Bubble-Size.php

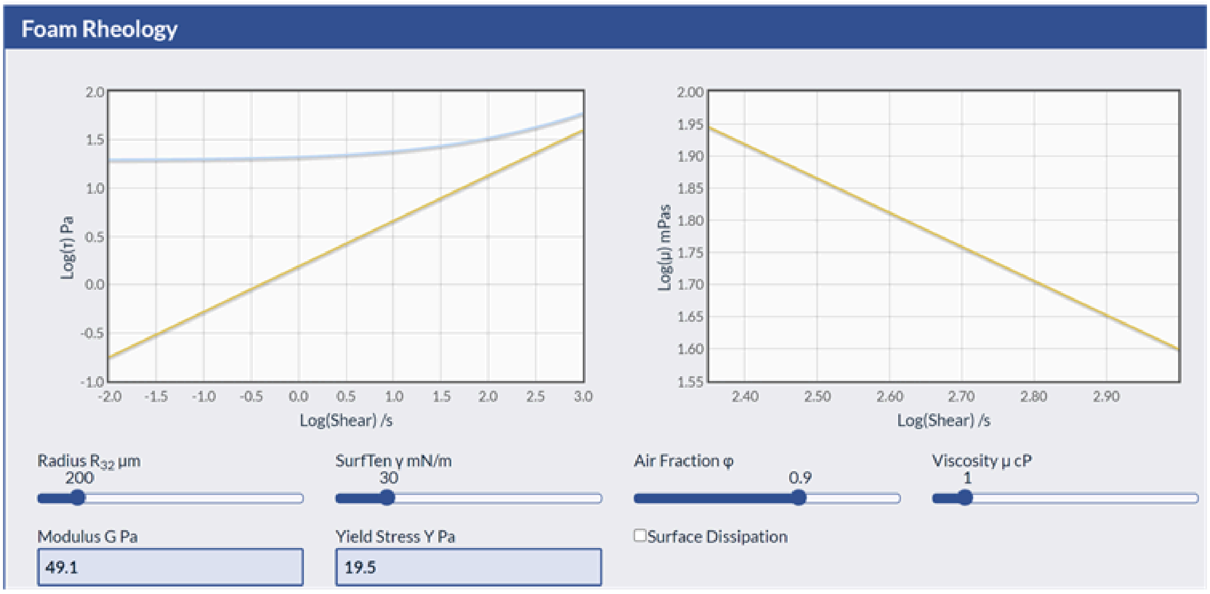
It is worth noting that when I needed to know how to measure the bubble size, it was, once again, the U Sofia group who told me how to do it. They'd not invented the technique, but they'd worked out how to do it effectively. Via the app, the technique has spread widely and is much appreciated. Occasionally someone tries a 60° prism, but that fails. And although I say "use a smartphone camera" it's not so easy to get a good, even, in-focus image. Because the quality of image analysis is 90% the quality of the original image, invest some time and experimentation into getting a high-quality image; it's worth it.

Surfactancy Foam Rheology

Links

[Dishwashing Liquids](), [Hair Shampoos](), [Soaps and Washing]()

Foam rheology is about Modulus, Yield Stress and shear-dependent Viscosity. Four simple inputs give us the behaviour, which in turn tells us that for a strong foam, small bubbles is the smart way to go.



Foam Rheology

Radius R_{32} μm 200

SurfTen γ mN/m 30

Air Fraction φ 0.9

Viscosity μ cP 1

Modulus G Pa 49.1

Yield Stress Y Pa 19.5

☐Surface Dissipation

https://www.stevenabbott.co.uk/practical-surfactants/foam-rheology.php

Controlling foam behaviour

The basic "feel" of a foam is captured by the 3 components of its rheology:

1. Modulus – how springy it is if we try to compress it;
2. Yield Stress – its resistance to flowing;
3. Viscosity – how viscous it is when you get it flowing.

When the air fraction is φ, the surface tension is γ and the bubble radius is R then the modulus, G, is given by:

$$G = \frac{1.4\varphi(\varphi - 0.64)\gamma}{R}$$

The yield stress, Y, is given by:

$$Y = \frac{0.5(\varphi - 0.64)\gamma}{R}$$

The equations for two types of shear-rate-dependent viscosity are provided in the app. The two types are

1. Sliding friction – bubbles just moving past each other smoothly;
2. Surface dissipation – bubbles expand & contract during motion and therefore dissipate energy.

You can click between them to see their differences. It is easy to take a sliding foam from simple surfactants and change it to a surface dissipation foam by making the foam walls more rigid via a small % addition of myristic acid or similar poor surfactant-like molecules that prefer to pack at the interface once the basic surfactant has created the foam.

Yet again, these core equations come from the team at U Sofia who have provided so much great, usable foam science.

Make small bubbles

Although surface tension makes a difference, the practical range of control is limited. Viscosity is usually close to water unless you add glycerol, so again there's not much room for change. Going to small bubbles is, therefore, the only way you can readily produce a "stronger" foam. The trick for making small bubbles is to focus less on the surfactant and more on your foaming mechanism. As shown in Surfactancy_Foaming, a good foaming technique can boost the foam more easily than playing with the formulation.

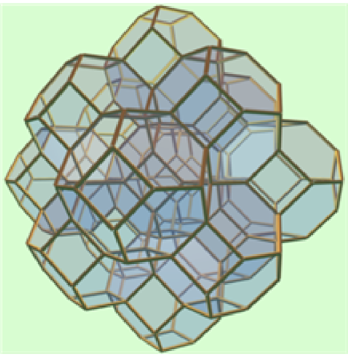
Surfactancy Foaming

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, Toothpastes

There is no mystery to foaming – any bubble in any liquid can create a "foam". Once there is sufficient density of bubbles (a foam fraction above ~74%) a real foam is in existence. Now the question is how stable the foam is. Most "reasonable" surfactants and plenty of nano-scale particles or lumps of protein can provide the stability. So now we're on to real foam volume and quality. This is less dependent on the surfactants than is imagined.

Create small bubbles, quickly



This image from https://www.stevenabbott.co.uk/practical-surfactants/Foam-3D.php shows what we are trying to achieve – lots of bubbles, maybe with volume fraction over 90% packed into a stable structure. As the description above sets out, just about anything can be made to foam and with the right amount of air being chopped down to small-enough foam bubbles, a credible foam can be created via any reasonable surface active agent that can create a reduction in air-water interfacial tension or provide a semi-solid array ("Pickering foam") of small particles.

This description is very different from the standard descriptions of lists of surfactants that are said to be strong or weak foamers, of special synergies of co-foamers and the general confusion about foam quantity and stability.

It is also different from analyses of the stability of isolated bubbles where analyses of disjoining pressure dependence on DLVO (see https://www.stevenabbott.co.uk/practical-surfactants/foam-dlvo.php), elasticity (see https://www.stevenabbott.co.uk/practical-surfactants/foam-elasticity.php) are of interest. These effects are real, but seem to play a minor role in real foam stability.

We can state this because a ground-breaking paper from the Denkov group in U Sofia, referenced in the foam-making app https://www.stevenabbott.co.uk/practical-surfactants/foam-making.php, used high-powered theoretical tools to analyse extensive foam creation data from a wide range of "normal" surfactants and concluded that there was no real difference between the quantity and quality of foam produced, so the theoretical tools provided no significant insights.

Instead, they pointed out, the key to making a good foam is to ensure a rapid reduction in bubble size through proper application of shear forces. As soon as you have small bubbles then Surfactancy_Foam Drainage decreases (proportional to radius squared) and the Surfactancy_Foam Rheology shows higher modulus and yield stress (proportional to 1/radius).

For those who want plenty of stable foam, concentrating on the foaming mechanics is more important than worrying about surfactant formulations. For example, a weak and useless fire-fighting foam created via inefficient aspirator nozzles, can be replaced with stunning stable foams using the CAF (Compressed Air

Foam) technique. A bubbly hand soap wash can be turned into a rich shaving foam using a Japanese foaming net (rapid rubbing with a fine mesh that efficiently breaks up bubbles) or a simple latte whisk.

Bad foaming needs better surfactants

If you create a foam by forcing bubbles through a sintered disk, you are automatically guaranteeing a poor foam because there is insufficient shear to make the relatively large bubbles smaller. So carrying out detailed analyses in this "foam analyzer" is an analysis of an unnecessarily failed foam. Just because a piece of equipment is flashy and complicated doesn't mean that it produces data relevant to your application.

If you use a kitchen blender, or repeatedly invert cylinders of foaming solution in other "standard tests" you again are working with a foam that is less good than it should be.

So in these cases, the surfactant system matters. Getting good Surfactancy_Dynamic Surface Tension effects, and using ill-defined foaming synergies become necessary and tedious tasks. The problem is ill-defined and despite decades of work we seem to be no nearer to finding helpful scientific rules.

If your organisation is demanding that you make another me-too foaming product, but this time with a more luxurious foam, try to innovate by *not* doing what everyone else is doing. Use your science and ingenuity to create a novel system that makes a great foam from a me-too formulation. Astonish them with the quality of a foam from a Japanese shaving net, or bring in a cheap latte whisk. Once colleagues see that it's the foamer, not the foam, they might be persuaded to embark on a fresh approach.

Take inspiration from the amateur YouTubers (some of whom I've enjoyed helping) who are keen to make great foams for cement blocks, children's parties, protecting crops from frost and other amazing applications. Invariably they evolve methods that naturally create very small bubbles, or break up bubbles efficiently from large to small.

Surfactancy HLD

Links

Dishwashing Liquids, Hair Shampoos, Soaps and Washing, Laundry Liquids, SkinCare, Water-based Adhesive, Emulsion Polymers, Pharma Formulation

> If we are formulating with surfactants and oils, using different salt concentrations and temperatures, it is a good idea to use a formula that captures all 4 factors, allowing us to move around surfactant space as we adjust our formulations for different end uses, for changes of raw materials and for going greener. Unlike all other systems, HLD lets you do exactly that.

Hydrophilic Lipophilic Difference

Let us have a temperature T, a salinity S in equivalents of g/100cc of NaCl, and oil with an oiliness EACN and a surfactant with a characteristic value, Cc. We can calculate the HLD via the simple formula:

$$HLD = F(S) - k.EACN - \alpha(T - 25) + Cc$$

While k ~ 0.16 for most systems, the other parameters are specific to surfactant types:

- For ethoxylates F(S) = 0.13S, $\alpha = -0.06$
- For ionics F(S) = ln(S), $\alpha = 0.01$
- For APGs, Spans … F(S) = 0, $\alpha = 0$ and for polyglycerols $\alpha \sim -0.01$

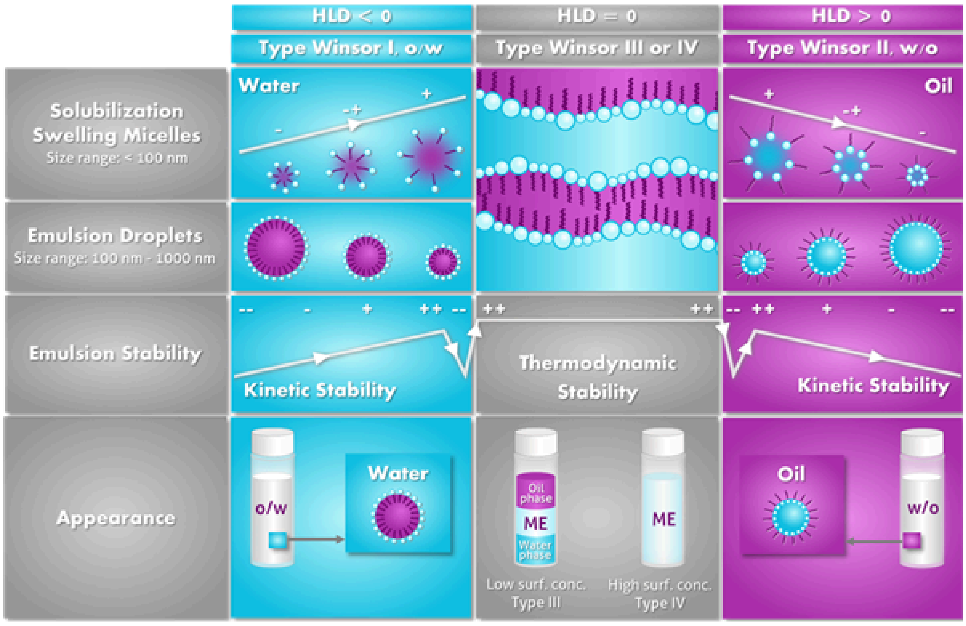
Even if we don’t know what HLD means, these bullet points immediately alert us that formulating with ethoxylates, ionics and sugar-based requires different tricks for each class, especially the sugar-based which have no dependence on salinity or temperature.

> **Battling HLB**
>
> When HLB was first invented, it did a good job of describing some attributes of the then new ethoxylate surfactants. But it didn’t take into account their large T-dependence and didn’t have a coherent way to deal with oils. Somehow it got extended to all surfactants, despite the fact that there was no objective way to measure their values. In the 1970’s the first steps towards HLD were taken, but the HLB orthodoxy blocked adoption of the obviously better idea. If HLB had died back then we would have had 30+ years of data gathering on HLD and we would all be in a better place. Killing off HLB has been extraordinarily difficult, but its demise is accelerating, and it will, hopefully, be soon forgotten.

The meaning of HLD

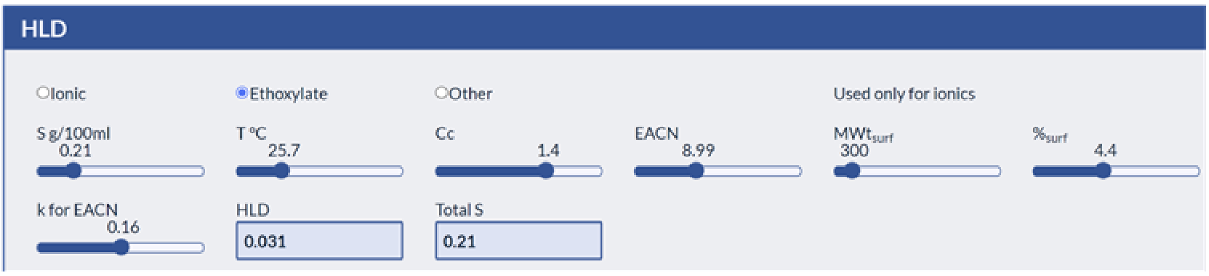
When HLD = 0 we have a system balanced between oil and water. The Surfactancy_Interfacial Tension and Rigidity is super low making it easy to create microemulsions and to break up oil drops for emulsification and for cleaning oils from surface.

- When HLD < 0 we have oil-in-water curvature of the interface. If you want to make o/w emulsions, it’s a good idea to know that your HLD < 0.
- When HLD > 0 we have water-in-oil curvature of the interface. If you want to make w/o emulsions, it’s a good idea to know that your HLD > 0.
- Because for ethoxylates we can increase HLD with T we can take an HLD < 0 o/w system at room temperature to the HLD = 0 inversion point at a higher temperature, and, if we choose, into HLD > 0 to make a w/o system at even higher temperature. This simple fact means that to call any surfactant (especially an ethoxylate) ready for o/w or w/o is misleading.



Using a graphic kindly provided by VLCI in Amsterdam, who are experts in formulating with HLD, we can see the whole of HLD space in one go, including the types of emulsions (called Winsor Types I, II, III or IV), their visual appearance and solubilization capabilities. It is impressive both that a single theory encapsulates so much and that VLCI’s graphic is able to convey that information so elegantly.

Although the calculation of HLD is trivial, it is easier to get an app to do it, taking into account the different coefficients for the different types of surfactant:


HLD

○Ionic ●Ethoxylate ○Other Used only for ionics

S g/100ml 0.21
T °C 25.7
Cc 1.4
EACN 8.99
MWt_{surf} 300
$\%_{surf}$ 4.4
k for EACN 0.16
HLD 0.031
Total S 0.21

https://www.stevenabbott.co.uk/practical-surfactants/hld.php

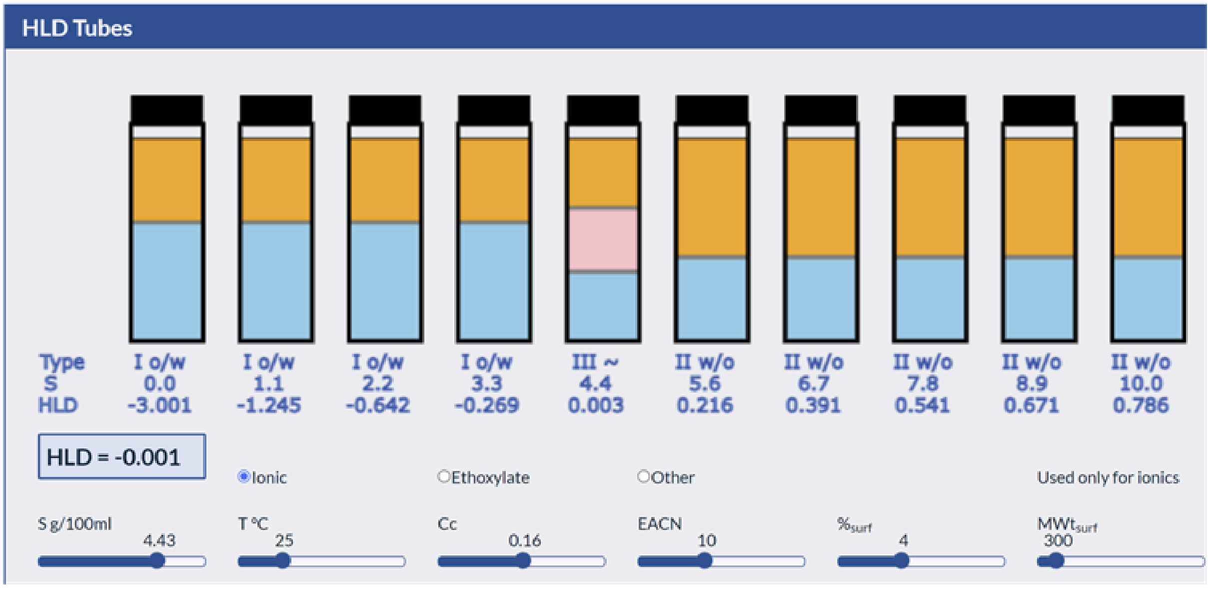
Sometimes you find a theory out of desperation. I needed a microemulsion for a cosmetics project and from the literature I could create one using decane as the oil. It had the right properties in tests but I needed to change to IPM, isopropyl myristate. All it needed was a change in the surfactant package as IPM is a different oil from decane. A search of the literature showed that HLB, CPP, Winsor R etc. were useless. Oh, there’s another theory called HLD. A couple of weeks exploring

HLD in the literature and in the lab resulted in a perfect IPM microemulsion. Obviously everyone would want to use such a simple and usable theory, so I put the first spreadsheets, programs and apps into the public domain. That was in 2010 and people (though far fewer) are *still* saying that they prefer HLB or that HLD is “too hard”.

What is the Cc of my surfactant?

Sometime in the future, all surfactant suppliers will provide the Cc values not just of each surfactant but of each *batch* of surfactant. The last point is important. Commercial surfactants are mixtures and even small changes of the chemicals in the mix can change the Cc by a significant amount. This explains the often-found failure of a new batch of the “same” surfactant, or how the “same” surfactant from a different supplier can behave differently *even though standard analyses say they are the same*. One of the reasons for the existence of HLD is that we can rationally formulate both large changes (as in the diagram) and small tweaks for different batches of a surfactant.

There is a list of currently known Cc values on https://www.stevenabbott.co.uk/practical-surfactants/cc.php but for *your* surfactant you either need a good value from the supplier or measure it yourself. Here’s how to do that:



Type	I o/w	I o/w	I o/w	I o/w	III ~	II w/o	II w/o	II w/o	II w/o	II w/o
S	0.0	1.1	2.2	3.3	4.4	5.6	6.7	7.8	8.9	10.0
HLD	-3.001	-1.245	-0.642	-0.269	0.003	0.216	0.391	0.541	0.671	0.786

HLD = -0.001

● Ionic ○ Ethoxylate ○ Other

Used only for ionics

S g/100ml 4.43

T °C 25

Cc 0.16

EACN 10

$\%_{surf}$ 4

MWt_{surf} 300

https://www.stevenabbott.co.uk/practical-surfactants/HLD-Tubes.php

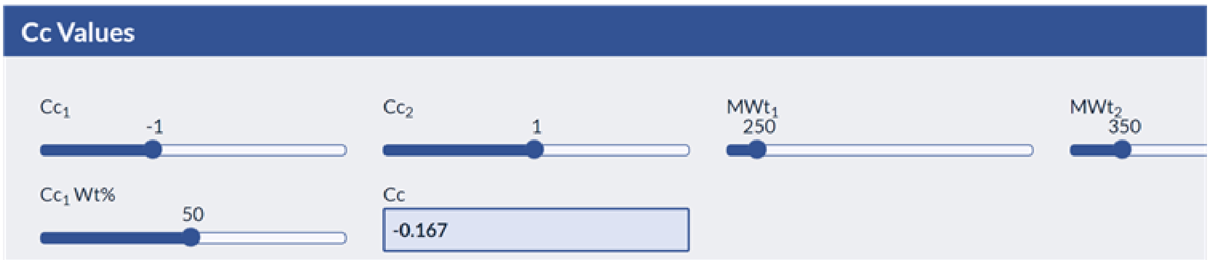
You have a set of tubes containing 50:50 oil and water. You systematically vary one of the variables (salinity, EACN, Cc) across the tubes and you look for the 3-phase tube where HLD = 0, the swap between o/w and w/o. Because you know all the other variables, you know the Cc of your surfactant.

In the early days of trying to popularise HLD, a small amount of funding gave the chance for myself and VLCI to attempt to measure the Cc values of ~20 surfactants, kindly provided by some companies who were willing to explore this relatively unfamiliar approach to surfactant science.

This was a mad thing to try to do. We had no apps for guidance, we'd only done a few such scans before, and surfactants are often uncooperative We made lots of errors out of ignorance (we now know never to mix SLS and AOT – they hate each other) and we had a few failures. But over a few weeks we got reasonable values (or upper/lower limits) on ~ 15 surfactants.

The secret of success was hard work, self-criticism, iterative learning from failure … and having the wise advice of Prof Edgar Acosta available by email and (back then) Skype.

Although scans are rarely as perfect as the screen shot, they are done routinely. If, for some reason, your preferred scan gives strange surfactant phases then you have to try an alternative. One variant is to make a scan with a known good surfactant that gives clean scan tubes, then to add a small amount of your new surfactant. After shaking and equilibration, the 3-phase tube might have shifted by 1. This gives you a new Cc for your mix and from a simple calculation based on mole fraction of the two surfactants, you get the unknown Cc value.



Cc Values

Cc_1 -1

Cc_2 1

MWt_1 250

MWt_2 350

Cc_1 Wt% 50

Cc -0.167

https://www.stevenabbott.co.uk/practical-surfactants/cc.php

In this simple example, the Cc of a 50:50 wt% mix of surfactants with values of -1 and 1 does not give Cc = 0 because #1 is a lower MW so there is a larger molar fraction of #1, resulting in the negative Cc.

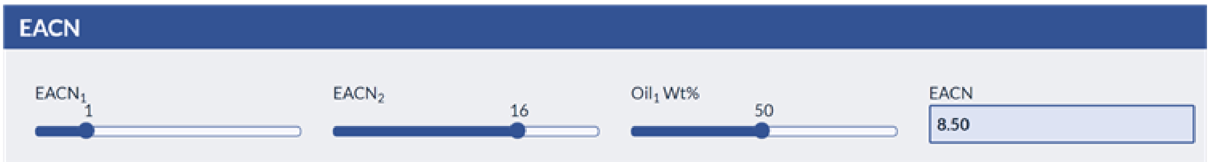
This mixture rule is very powerful. If you need a precise Cc for a great formulation but don't have (or can't afford, or don't like the green credentials of) a single surfactant at that value, you create a mix from 2 available, affordable and greener surfactants.

Other scans are becoming increasingly popular. The Acosta Fast method (see a description at https://www.stevenabbott.co.uk/practical-surfactants/fast-method.php) is one, the PIT-slope method from U Lille is another as is their clever salinity scan that uses smart mixing to increase salinity while keeping other ratios constant. Some find that conductivity measurements are faster for identifying phase changes. Choose whichever technique works for your specific circumstances. There's a big difference between a scan for a new, unknown surfactant, and a scan to detect subtle batch-to-batch variations of a known one, ionics are different from non-ionics, ethoxylates are different from everything else. There's no single right answer, but the worst answer is to keep going with a slow and painful technique that discourages you from getting the information you need.

What is the EACN of my oil?

The "oiliness" of octane is 8 and of dodecane is 12. Suppose we have an oil which is not a pure alkane but, in a scan, behaves like decane? We give it an Equivalent Alkane Carbon Number of 10. If you do plenty of scans with plenty of oils you can find values from 20+ down to -5. What is an alkane with -5 carbon atoms? Of course it's meaningless in terms of alkanes, but this linear scale lets us formulate a wide range of oils.

How do we know the EACN? Again, suppliers should tell you their values. Or you could look them up on the list on the site below, or you can measure them by scans, or you can create them from a rational blend (this time in terms of vol%) of two known oils:



$EACN_1$ 1
$EACN_2$ 16
Oil_1 Wt% 50
EACN 8.50

https://www.stevenabbott.co.uk/practical-surfactants/eacn.php

As with Cc values, EACN values of the "same" oil can vary – for example typical plant-based oils. So doing a scan to measure each incoming batch of oil tells you whether you need to tweak your formulation for the new batch, e.g. making a slight adjustment to the surfactant ratio.

Polar oils

Although HLD can handle salts, temperature, surfactants and oils, it has a problem with "polar oils". These are molecules like octanol that can be considered an oil or a surfactant but don't behave as either. A small amount (e.g. a fragrance molecule in a formulation) can make a large change to HLD. There happens to be a good theory for tackling polar oils issues. It is Tchakalova's CIT (Constant Interface Thickness) approach, captured in https://www.stevenabbott.co.uk/practical-surfactants/curvature.php. In practice, the approach is far too hard for standard formulators, so we have to check their effects on our individual formulations. Fortunately, the culture of scans that we use to measure EACNs, Ccs and batch variations can also be used for polar oils. One example of doing this is the "Pit Shift Method" which is the 4th app inside https://www.stevenabbott.co.uk/practical-surfactants/measure-eacn.php.

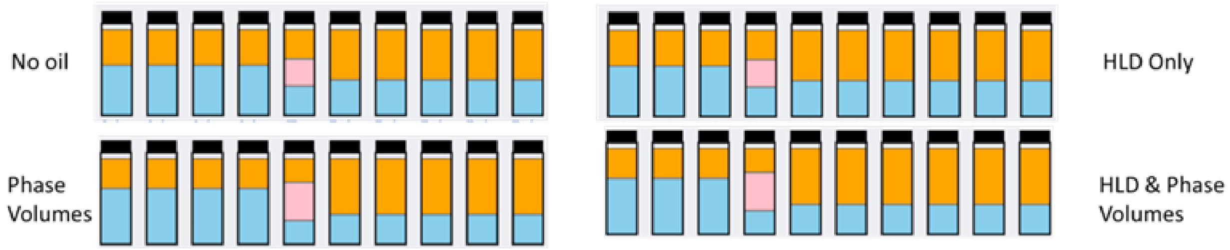
Take a scan and find that the transition tube is, say, #4. Now add a small amount (e.g. 0.5%) of the polar oil to each tube. After re-equilibration we might find a shift towards #5. Repeat with another 0.5% and now we are clearly in #5. With luck (and it's usually the case) you will find a linear relationship. This means that you can rationally modify your HLD to take into account the amount of polar oil required in your formulation. If you need to increase the amount for some other reason, you can rationally tweak, say, the Cc blend to accommodate it. A shift to a different fragrance mix? Re-do the 0.5% scans and get a new correlation.

Yes, it would be nicer if we could predict polar oil effects. But until we can, some quick scans can get us quickly to the right formulation.

Cosurfactants, linkers, extended surfactants

HLD takes you to the right part of formulation space, but how much surfactant should you add to get the required effect? HLD cannot tell you, you need the more powerful HLD-NAC theory, https://www.stevenabbott.co.uk/practical-surfactants/nacmore.php. A key factor in HLD-NAC is the ξ value which tells you the efficiency of that surfactant with that oil (sadly there is no general-purpose ξ theory so far). Here we don't worry about ξ but take on board the fact that for a given oil at the same HLD different surfactant blends at the correct Cc can have smaller or larger ξ values and, therefore, different efficiencies.

If the efficiency isn't as high as you would like then you can try adding some cosurfactants or linkers (the terminology if vague). Many formulators get into a mess trying this because any extra molecule can change both the HLD and the efficiency. Let’s show that with a diagram:



The additive might change the HLD only. If you weren’t aware of this change, you might think the additive makes things worse. It might change the HLD *and* the phase volumes. Again, if you weren’t aware of the HLD change you would think that the additive made things worse when in fact it is just what you required – but at a shifted HLD (which you can shift back via the Cc blend). Yes, if you are lucky the phase volumes get larger so your experiment would be a success. But formulators can’t rely on luck.

So each time you try an additive to improve the system, test the additive on some scan tubes and if you have an improvement in efficiency (good) with a shift in HLD (bad), you just shift the HLD back to where you wanted it.

How do cosurfactants/linkers work? By effectively extending the tail region and/or the head region to improve the ability of the surfactant to control larger phase volumes. Why not just use a surfactant with a longer tail? If you can, that’s the best option. But often a C14 surfactant will have the right solubility and phase behaviour while a C16 version is insoluble or makes undesirable phases with the other surfactant in the blend.

“Extended surfactants” are a different story. They have a similar head and tail region but add a “neutral” polypropylene oxide middle region that doesn’t mess up the surfactant behaviour while increasing the influence over the oil and water.

The disadvantage of the higher efficiencies of extended surfactants is that they are often slower to act.

Silicone surfactants and oils

The Cc and EACN numbers are linked via the traditional carbon chemistry of alkyl chains, aromatics and modest functionalities such as esters. Silicones inhabit a different world. Although they are “hydrophobic” that doesn’t mean that they are “oleophilic”; indeed, they are often insoluble in common oils.

This means that the simplest advice is “the best surfactant for a silicone is a silicone surfactant”. Or, in other words, “Don’t try to use HLD for silicones”.

This general separation of standard and silicone surfactants provides insights into why, as discussed in Surfactancy_Anti-foaming, silicone anti-foamers are generally more successful or, rather, less unsuccessful.

Surfactancy Interfacial Tension and Rigidity

Links

Laundry Liquids, SkinCare

Interfacial Tension, IFT, is the “surface tension” between oil and water. It dominates how the oil and water interact across the interface. When you are at an HLD = 0 balance (Surfactancy_HLD) the IFT isn’t just low but super-low, enabling many desirable actions in our surfactant systems.

Interfacial rigidity

The aligned surfactants at an interface will add some rigidity to the interface. This rigidity, E_r, has an influence on what happens when emulsion drops collide and along with the Marangoni stabilization effect (where surfactant molecules rush in to fix any temporary reduction in surfactant concentration at the interface) exerts control over emulsions and emulsification. It is obviously linked to curvature values, as it is harder to bend a more rigid interface. Yet it is little talked about and data on values is sparse. Instinct suggests that those surfactants with a high rigidity will give emulsions that are more stable against Surfactancy_Emulsion Coalescence, yet the relevant theory there uses an activation barrier, of unspecified origin, rather than interfacial rigidity.

The only rule of thumb is that for ionics it’s around 1kT and for ethoxylates around 4kT. We’ll use those values in a moment.

Interfacial tension

The energy to open up $1m^2$ of fresh interface (J/m^2) or the force exerted by the interface along 1m of, say, a Wilhelmy plate, (N/m) is the interfacial tension. For a stable emulsion we want this value to be high. To create fresh interface (e.g. cleaning oil from a surface, Surfactancy_Roll-up and Eötvös Number, or sweeping oil from rocks Surfactancy_Critical Capillary Number) we want a low value.

It is directly related to the curvature of the interface though we might prefer to use radius, R, as being 1/curvature.

$$IFT = \frac{E_r}{4\pi R^2}$$

When we invert an emulsion, going from o/w to w/o, we invert the curvature so it becomes 0. This means that the radius passes through infinity at that point and the IFT is 0. This point is defined via Surfactancy_HLD theory. If we apply the more sophisticated HLD-NAC theory we find that R doesn’t go fully to infinity, it is merely a large value. We see all this in the app:

○T ○EACN ○S ●Cc

●Ionic ○Ethoxylate ○Other ☐Non-log

S g/100ml 4.25

T °C 25.7

MWt 305

Cc 0

A Å² 46

L Å 13

EACN 8.8

N kT 1

HLD -0.030

ξ Å 102

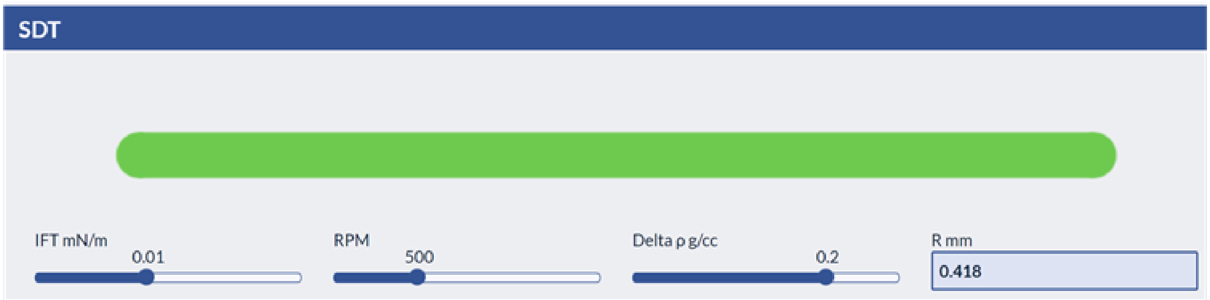
https://www.stevenabbott.co.uk/practical-surfactants/ift.php

In the screenshot, the system reaches HLD = 0 with a surfactant of Cc = 0. At that point the IFT is 0.006 mN/m. A typical o/w emulsion created at HLD < -3 will have an IFT > 1 so there is a reduction of a factor of 150 by tuning the formulation to HLD = 0.

Much less effort is required to create a fine emulsion when the IFT is super low, a fact used (but often ill-understood) in PIT (Phase Inversion Temperature) emulsification, in catastrophic phase inversion (Surfactancy_Emulsion Inversion a sudden flip in curvature unrelated to HLD, but still a high radius interface), and by those who use HLD to create emulsions via other rational variations of S, T, EACN and Cc.

Measuring IFT

If you can measure IFT with anything other than a spinning drop tensiometer (SDT), the values you are measuring are not too interesting. Because the real interest is with super-low IFTs you need a device that measures these low values.



IFT mN/m 0.01

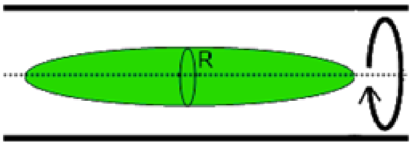
RPM 500

Delta ρ g/cc 0.2

R mm 0.418

https://www.stevenabbott.co.uk/practical-surfactants/sdt.php

We have a narrow tube containing our aqueous phase and we inject a small sample of oil, which will form a spherical drop. You spin the tube and the centrifugal forces force the drop into a cylindrical shape:



Via a microscope you measure the radius R and from it you calculate the IFT.

IFT γ is calculated from R, the density difference $\Delta\rho$ and the angular velocity ω via:

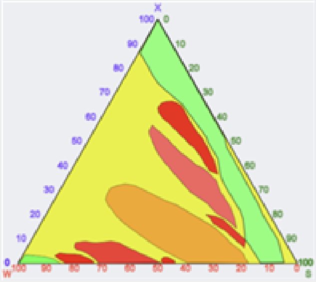
$$\gamma = \frac{\Delta\rho\omega^2 R^3}{4}$$

It sounds easy, but those who have worked with classical SDTs find them to be very hard to use. Fortunately, the newer generation of SDT devices are easier so such measurements can become more routine.

Surfactancy Phase Diagrams

Links

Hair Conditioners, Laundry Liquids, Pharma Formulation



Most of us avoid surfactant phase diagrams.

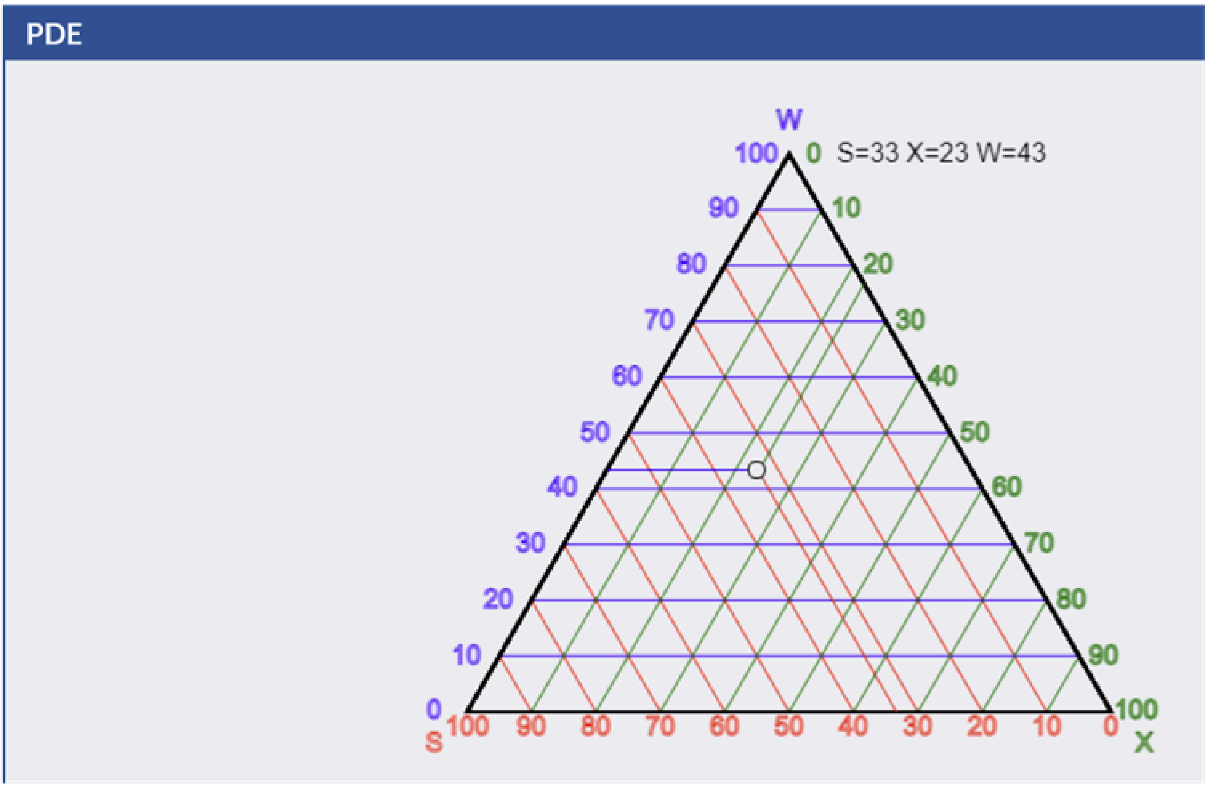
We find their ternary shapes confusing and examples like this look like too much hard work (which is probably true).

But we can do a *lot* with relatively simple diagrams and with the help of a few apps they turn out to be easy to create and easy to analyse.

What are we trying to achieve?

We seldom have just a single surfactant and water – generally we have at least 3 components, such as surfactant, water and oil. To navigate around that space a ternary diagram is compact, informative and, to most of us, confusing.

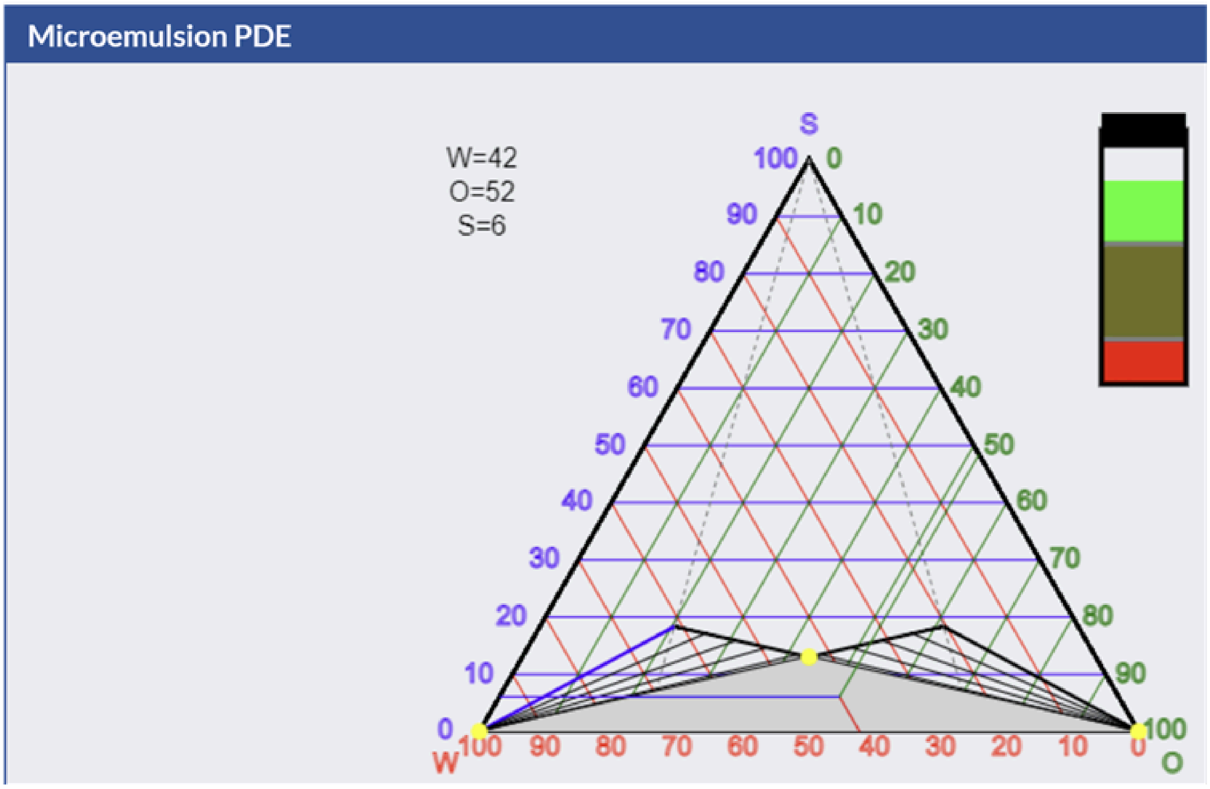
So first of all, get used to reading a diagram. The Phase Diagram Explorer lets you view a variety of different formats and lets you set your own level of difficulty in exploring it. Here we have chosen to have the most helpful output, telling us where our mouse pointer is … and why those are the values.



https://www.stevenabbott.co.uk/practical-surfactants/pde.php

I was having lunch at a P&G facility. We were discussing phase diagrams and I commented that I found them so difficult that I was writing an app so I could understand them myself. The stranger next to me was delighted. “I teach phase diagrams here, and everyone is confused by them. Can I help you to write the app”. The stranger was Seth Lindberg who is expert in both in phase diagrams and graphics. We agreed on a joint project. The phase diagram apps are correct because of Seth’s critique of iterative versions and are informative because of Seth’s graphics skills.

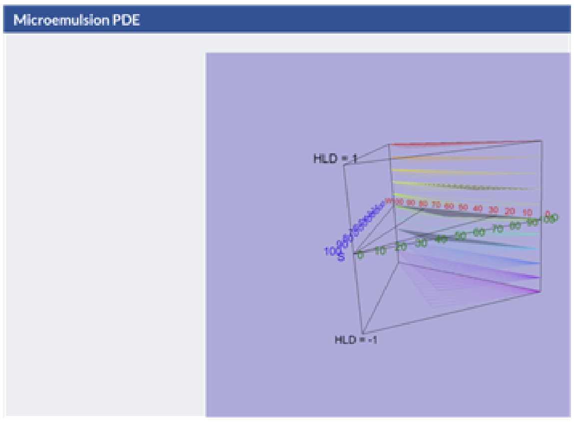
You can quickly graduate from this training app to one that shows the tie lines and shaded zones from microemulsion systems:


Microemulsion PDE

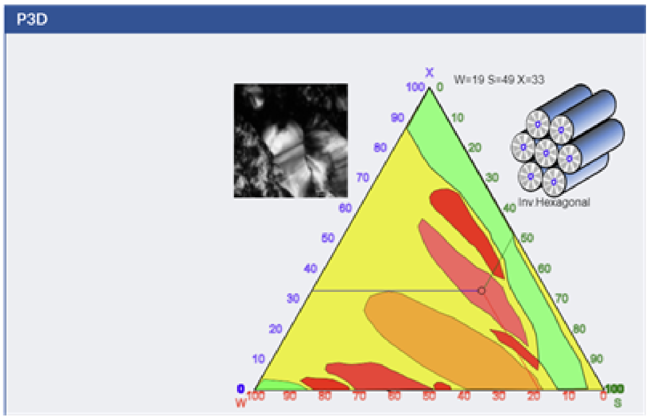
W=42
O=52
S=6

https://www.stevenabbott.co.uk/practical-surfactants/mpde.php

You not only get the Water, Oil, Surfactant readout (42:52:6) but you also see the 3 phases that appear in this region. Things that are mysterious to most of us, such as “tie lines” and the triangular areas without lines, become straightforward once you explore them with your mouse.

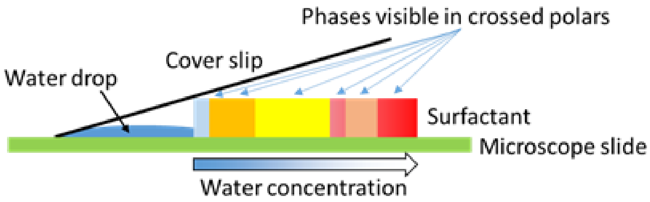

Microemulsion PDE

If 3 dimensions aren’t enough for you, you can even view data in 4D. This is a fancy “prism plot” and is about the limit of what can be achieved on a 2D screen: https://www.stevenabbott.co.uk/practical-surfactants/mpde-prism.php . It is a powerful example of how theory and visualization can come together. The theoretical microemulsion ternary diagrams can be seen over a range of HLD values so you can grasp what happens if, for example, the temperature changes for an ethoxylate system.



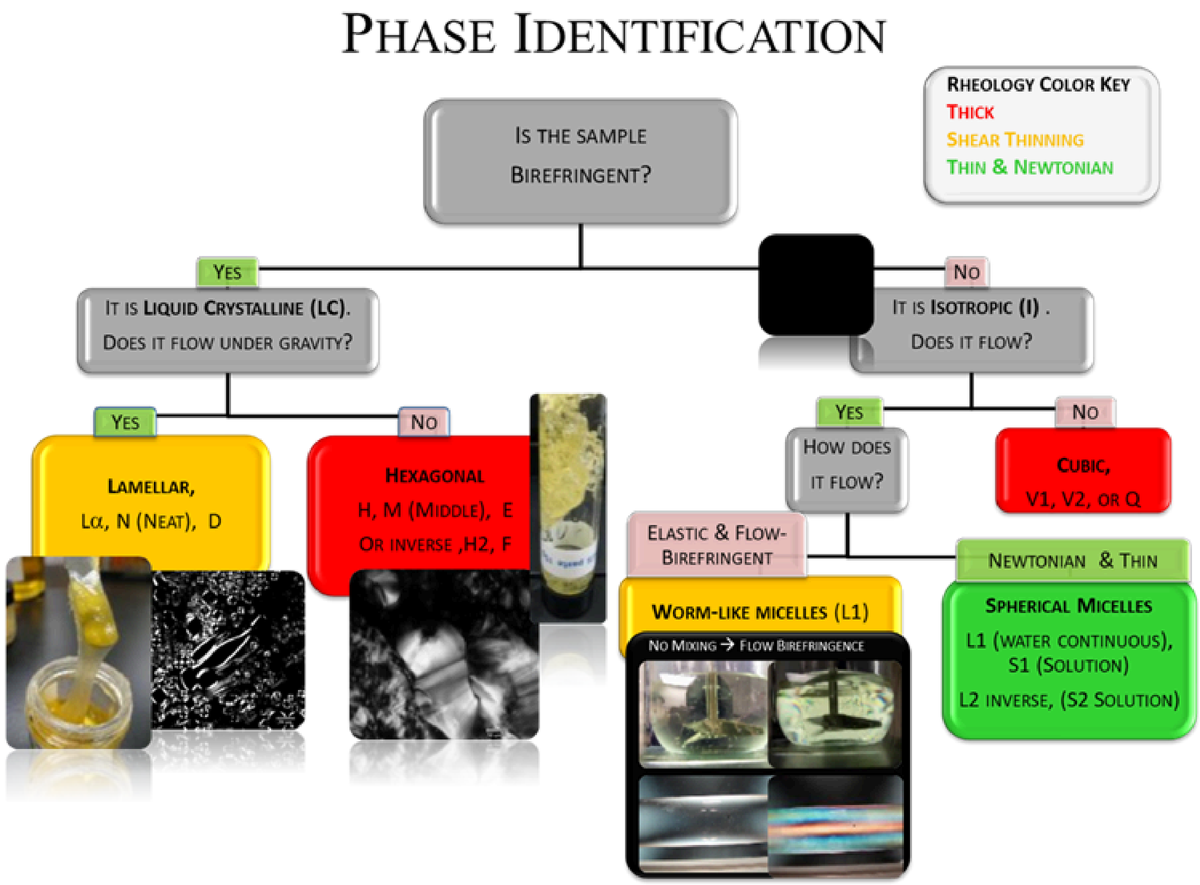
Additionally, what about the scary diagram in the introduction to this chapter which was extracted from https://www.stevenabbott.co.uk/practical-surfactants/p3d.php? Just move your mouse over it and you can see what the phase is, what it looks like in a polarizing microscope and how the surfactant molecules are packed. Although it's great if you have a polarizing microscope to explore complex phase diagrams, they aren't necessary for most of us who only need to explore a limited part of surfactant space.

Your own phase diagram



So how do you create your own phase diagram? If you happen to have a polarizing microscope, one neat trick is to place a sample of your surfactant onto a microscope slide, place a drop of water on one part and place a cover slip as shown. Looking down themicroscope you can quickly see if the surfactant likes to get into interesting phases as you go from low to high concentrations.

For more complex setups, you need to look at a few well-chosen tubes and work out what phases they contain. For this you *don't* need a polarizing microscope. The techniques described in the QPD (Quick Phase Diagrams) app, developed by Seth Lindberg at P&G have proved very popular because they are both simple and powerful. Here, using Seth's graphic, is how to know what phase is in your tube:



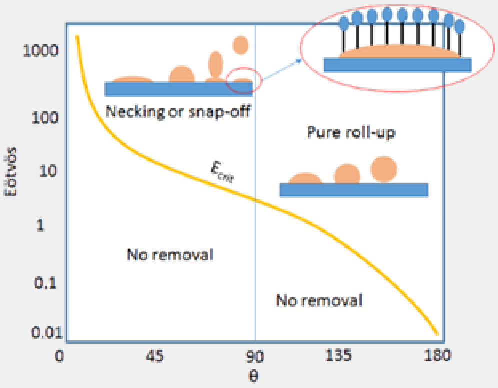
https://www.stevenabbott.co.uk/practical-surfactants/QPD.php

A few other tips and tricks (e.g. a light box with some sheets of polarizer) can be found on the Quick Phase Diagram web page.

Surfactancy Roll-up and Eötvös Number

Links

Dishwashing Liquids, Hair Shampoos, Laundry Liquids



Oil is most efficiently removed by surfactant from a surface via “roll up”, where the blob of oil just detaches or, snap-off, where a big chunk comes off. This can only happen when the Eötvös number, a ratio of gravitational and interfacial forces is above a critical number (the yellow line) which, in turn, depends on the contact angle of the oil with the surface.

The best way to get a high Eötvös number? Have a low interfacial tension. The best way to get a low interfacial tension? Formulate to HLD = 0.

Why do we want roll-up?

We know that surfactants can wrap around small blobs of oil, creating an emulsion. So *if* you can get the oil from your clothes or dishes into the water, the surfactant can wrap around it and keep it from the surface. But it takes a lot to overcome the problem of Cleaning_Boundary removal. If you can get a large blob of oil to simply rise into the water, that would be ideal.

The force lifting the oil depends on its radius, R, gravity g and the density difference Δρ between the oil and water. The retaining force is the interfacial tension, γ. The balance between them is the Eötvös number (named after a famous Hungarian scientist) E given by:

$$E = \frac{\Delta\rho g R^2}{2\gamma}$$

When $E > E_{crit}$, the yellow line in the diagram, you get roll-up or, at least, necking & snap-off which are better than nothing.

What is E_{crit}?

It is rather easy to clean oil from cotton. Cotton is so hydrophilic that the oil is hardly attracted, so the contact angle is high, and a small E will remove it. The same oil on a polyester fabric is much harder to remove – the fabric and oil are more compatible, the contact angle is low and a high E is required. For each contact angle there is a critical value of E, E_{crit}, above which the oil will be removed. The calculation of E_{crit} is described in the app.

Drop Volume μL 5

Density Diff. Δρ g/cc 0.1

Contact Angle θ° 67

IFT mN/m 0.27

Status: Unstable

E_{calc}/E_{crit}: 5.58/5.06

θ_{calc}°: -2

r/R mm: 1.61/1.75

https://www.stevenabbott.co.uk/practical-surfactants/eotvos.php

In this example, the contact angle is 67°, E_{crit} is 5 and the calculated E value from the IFT of 0.27 mN/m is 5.6, so the oil is "unstable", i.e. will be removed.

Getting a low γ

The only way to get a high E is from a low γ. We know from Surfactancy_HLD theory that when salinity, temperature, surfactant and oil are matched the Hydrophilic Lipophilic Difference is 0. And because we can go directly from HLD to Surfactancy_Interfacial Tension and Rigidity we can find that γ becomes super low (orders of magnitude) at this point.

So for good cleaning it's easy. Find a surfactant which, at the required washing temperature, and for the intended oil (typically a triglyceride fat with an EACN ~ 12) gives HLD ~ 0.

It's never so simple

Of course this relatively simple theory cannot capture the full complexity of cleaning oil from surfaces. At the very least, the mechanical motion with a dish washing brush or inside a washing machine will give a boost similar to a larger E_{crit}. But what is not disputed is that tuning your surfactant formulation to be closer to the ideal of HLD ~ 0 will greatly aid cleaning.

As discussed in Dishwashing Liquids, the formulations have been optimised for foaming rather than cleaning, so they are surprisingly ineffective. Formulations with HLD closer to zero *do* exist so you have a choice – nice foam or efficient cleaning.

Thickeners Associative Thickeners

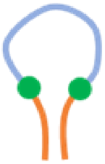
Links

Sun Screens, Emulsion Paint

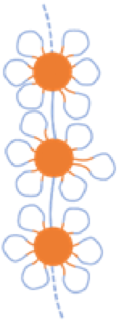
If we need real, permanent thickening of a system, then choose Thickeners_Polymeric Thickeners. When you want something to be able to go quickly from thick to thin and the reverse, then you need associative thickeners or, in special cases, Thickeners_Wormlike Micelles. Although ASE thickener aren't associative, their HASE variants are, so they are included here because their behaviour is so similar.

A fragile network

We start with the classic HEURs, Hydrophobically Modified Ethoxylated Urethanes, because they are relatively simple molecules with a rich, complex behaviour for good (the desired thickening behaviour) and bad (side-effects discussed later).



The orange parts are hydrophobic alkyl chains, say C16. They are attached by urethane links (green dots) to the blue ethoxylate chain. This hydrophobic-hydrophilic combination can form micelles, orange blobs in the next image, but with the ethoxylate loops sticking out. These molecules in the 10-35K MW range are called "telechelic" with "chelic" meaning "claw".

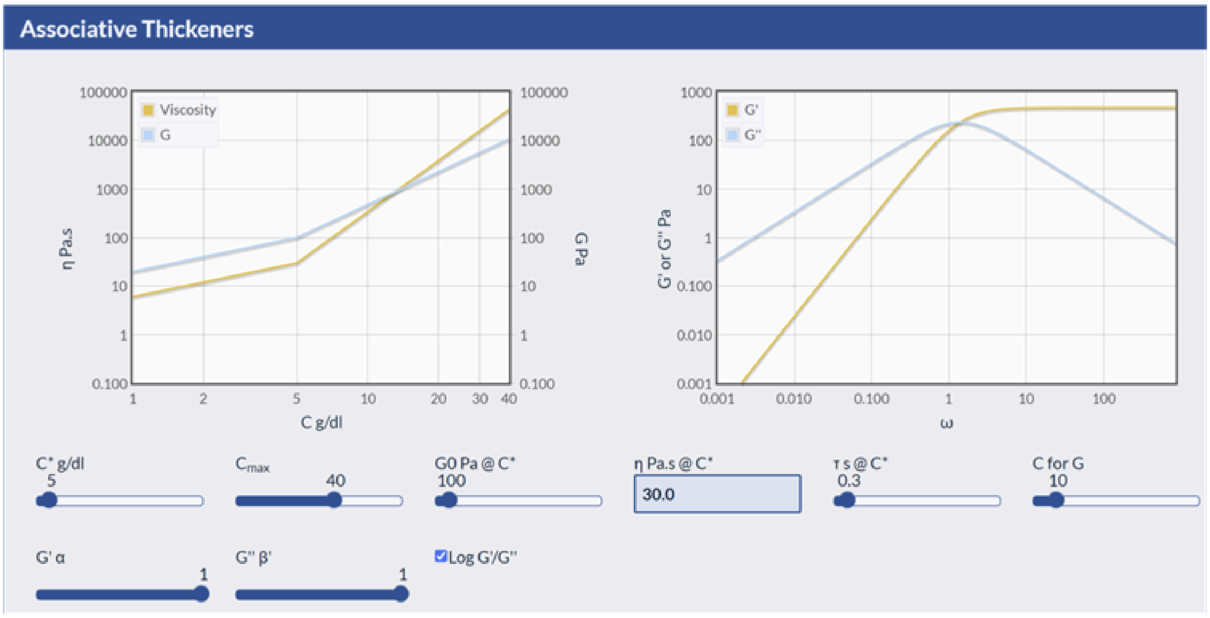


Some of the hydrophobic ends can go from micelle to micelle, allowing a network to build up. Here it's linear but it could equally be branched. At low shears, and short timescales these long pseudo-polymers can tangle and create high viscosities and a desirable yield stress Flow_Yield Stress so that a paint, coating or cosmetics is suitably "thick" for application, yet easy to apply under shear from an applicator or finger.

If it's a paint or cosmetic emulsion, why not simply thicken by increasing the volume fraction of emulsion particles? Because as shown in Flow_Particle Viscosity, viscosity rises alarmingly around the φ > 0.55 region so small changes can have a catastrophic effect on the performance of the formulation.

Thickener rheology

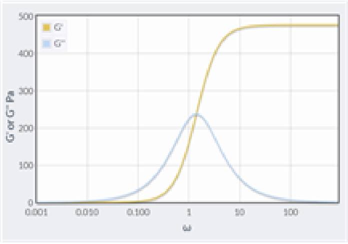
Most of us don't know what rheological behaviour we would like, but we would like to be able to have tests that can reveal internal details of good, OK and bad formulations. Standard low strain rheology does an especially good job for associative thickeners because we know in advance what an ideal system looks like – as shown in the app:



C* g/dl 5
C_{max} 40
G0 Pa @ C* 100
η Pa.s @ C* 30.0
τs @ C* 0.3
C for G 10
G' α 1
G" β' 1
☑ Log G'/G"

https://www.stevenabbott.co.uk/practical-surfactants/Associative-Thickeners.php

The left-hand graph shows what we might find if we measured low-shear viscosity at different concentrations of the thickener. There's a linear, uninteresting, increase up to some specific point, then the viscosity takes off. If you do such tests for a range of good/OK/bad thickeners in your specific system, you get a good idea of how they compare in this practical aspect.



The right-hand graph shows a typical "frequency sweep" where G' & G" (Flow_G' and G"). The G' is the elastic portion that gives you the gel-like behaviour you need. The higher G', the "stronger" the system. The log plot in the screenshot is deceptive (but is what is commonly shown in the literature) – this little image is the same data in linear format, showing more clearly the massive increase in viscosity around the 1/s timescale.

The timescale, ω, at which the response changes from fluid to gel-like is controlled by a relaxation time τ_R. You will want to control τ_R as much as you want to control the maximum G'

The reason we do these plots is because τ_R and G' are super important for formulation properties. For example, in paints you are interested in the performance on the brush or roller and in the subsequent levelling and sag: Coating_Levelling Theory. For cosmetics you will have other criteria such as the more subjective "feel". Although you might have a sense that one formulation is better than another, it's better to have numbers that can tell you whether the better formulation has a different τ_R, a different G' max or some sweet-spot combination.

Influencing the thickening

You can change 3 things in your choices of HEURs

1. Hydrophilic chain length – biggest impact on G' and τ_R. Going from C14 to C16 then C16 to C18 can give successive increases of 10x in each property.
2. Ethoxylate chain length – longer chains give somewhat greater thickening, but because the increase is apparent at the same *molar* concentration (same number of molecules), the effect is smaller when we measure it in terms of wt%.

3. Nature of the urethane bridge - the bridges affect hydrophobicity and rigidity so there are no obvious rules of how to choose

You can also choose, say, HASEs, Hydrophobic Alkali-Swellable Emulsion Polymers. These tend to be less well-defined than the HEURs so their rheological response is not so “sharp”. In the app, reduce the G' α and G" β values from their HEUR value of 1 to see how the rheological response broadens.

The other influence, for good or bad, on the thickening is surfactants. In a clean lab formulation you can add increasing amounts of your favourite surfactants to see the maximum G' or low-shear viscosity increase then, usually, decrease. For typical surfactants the effects are modest.

At first the tails of the surfactants displace one end of a HEUR, freeing it to link to another micelle. So the viscosity goes up. Eventually this becomes self-defeating – so many HEUR molecules are displaced that there’s no chance of a good network.

If you add “pre-wormlike” surfactants which, thanks to a Critical Packing Parameter ~ 1 (Surfactancy_CPP and phases) are keen to form strong micellar phases, then you can get larger increases. Indeed, if you add real wormlike micelles (Thickeners_Wormlike Micelles, with an app remarkably similar to the one in this chapter) then you are potentially doubling up on the thickening effect – with the warning that these complex systems can be unpredictable.

Real-world emulsion thickening

It’s a necessary first step to get used to HEURs, HASEs etc. in simple, clean formulations. Their behaviour is complex enough, and you have so many parameters to choose from.

As soon as they are applied to, say, real-world emulsion polymer formulations, complications appear.

The first is one you’ve prepared for. Those formulations may have their own surfactants which might be free enough to start interacting. Depending on their levels and their interactions with the HEUR (and you may have no idea of either) they might give you a thickening boost or (hopefully unlikely) will take you over the boost limit.

If you make your own emulsion polymers you can handle this rationally. If you buy them in, and the supplier is reluctant to reveal any useful information, you have to rely on intelligent guesswork.

Despite the uncertainties, at least this effect is only changing the basic rheology and with your background knowledge of good/OK/bad you can tweak your HEUR type or quantity to hit your sweet spot.

The other effect is very different – and is another good reason for using rheology. As described in Dispersions_Rheology (High shear) (which covers high volume fractions), the viscosity of an emulsion can increase dramatically if some sort of fractal particle association takes place. Those familiar with the terms “depletion flocculation” or “bridging flocculation” can imagine how the pseudo-polymeric associations of thickeners can act to attract the particles causing at best an increase in viscosity and at worst a complete collapse of the formulation.

The theories behind these effects are poorly described by any tool other than Dispersions_Scheutjens-Fleer theory. Read that chapter if you are keen to understand the effects and have some clue as to how to fix the problems.

Even if you don’t want to know about the theory, why should you worry about an increase in viscosity – after all we’re using thickeners, so some extra thickness sounds like a free lunch.

Put two formulations into a rotational viscometer and do a sweep of shear rate. The first formulation has no such particle/thickener interactions and you see the viscosity reduce sharply at low shear rates. The second formulation will start at a somewhat higher viscosity (not a big problem) but the fall-off is very slow. You now have a very different type of thickener. If you wanted this type you could probably have achieved it by increasing the volume fraction of emulsion.

Smart mapping

There is no realistic chance of you solving your thickener optimization challenges using pure theory – you have too many real-world uncertainties.

The alternative, trial & error plus “experience” has not been proven to be a great alternative.

What the theory tells us is that regularly carrying out the same experiments on a range of good/OK/bad formulations builds up a picture of:

- Viscosity versus concentration
- G' & G" over a range of frequencies
- Shear rate behaviour

Because each is related to a different aspect of formulation space you will be able to map how formulation ingredients interact across this complex space. With some sensible data analytics, you will be able to extract predictive tools that work in your formulation space.

It’s a minimax – minimum work for maximum payback.

Thickeners Polymeric Thickeners

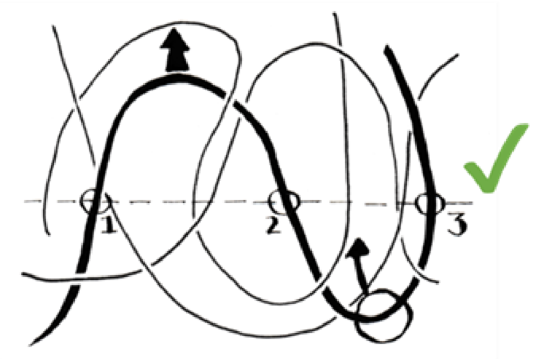
Links

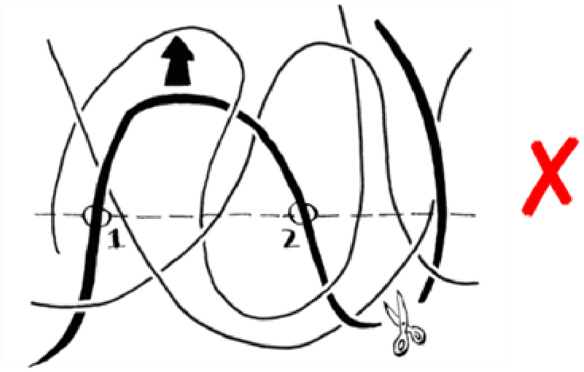
Sun Screens, Emulsion Paint

Thickening with a polymer is either a nuisance (when you want a high concentration of polymer in a formulation, but it makes it too viscous to handle) or a necessity (when you need a controlled, higher viscosity for some purpose). The theory is the same and covered in Flow_Polymer Viscosity. In this short chapter we focus on the general need to achieve a desired viscosity with the least added polymer and the fewest downsides.

Lots of tangles

Viscosity comes from entanglement of polymer chains. You may or may not be surprised that entanglement that causes viscosity is the same as causes adhesion, Adhesion_Entanglement, but that means we can borrow a diagram that shows that entanglement is 3 crossings:

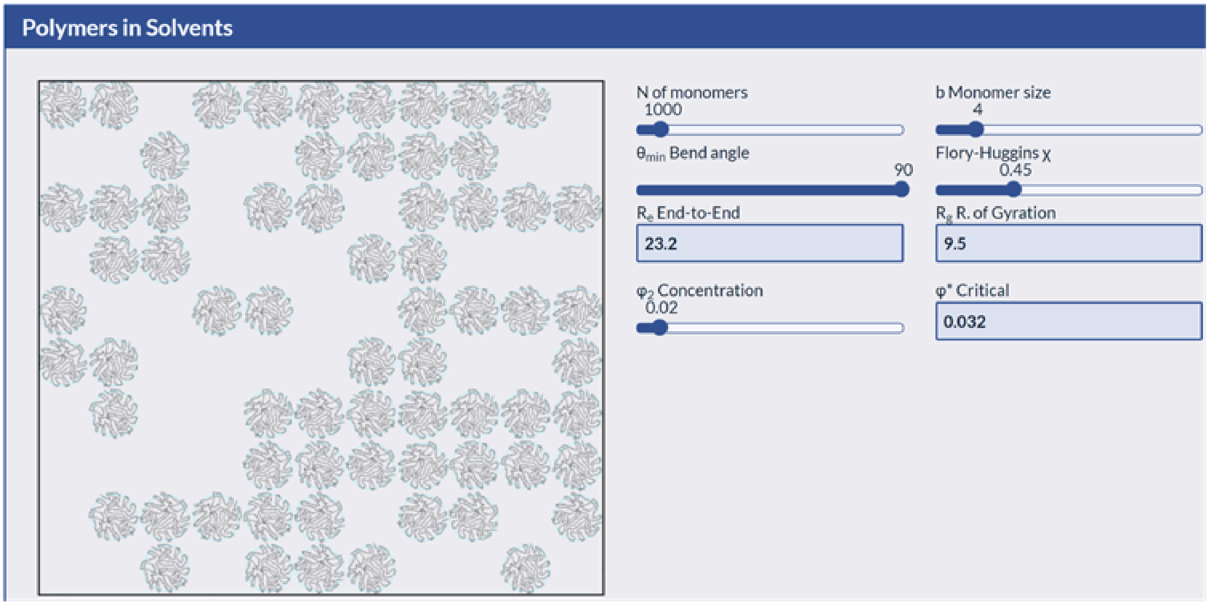




The dotted line is a real or virtual interface. In the left-hand image, the polymer chain of interest crosses it 3 times. This means that it is entangled. If you pull up on the big arrow at the top left, the lower loop gets tangled with the other loops.

In the right-hand image, the chain has been cut so it now only crosses 2 times. When you pull up, the loop is merely intermingled and can be pulled out. In terms of viscosity, this is a much smaller effect, closer to the φ* (volume fraction) and C* (concentration) situations discussed shortly.

Obviously if you have a larger MW you will get more tangles. But why do some polymers appear regularly as thickeners while others don't? Because good thickeners have a low M_c, Critical Entanglement MW. If you play with:


Polymers in Solvents

N of monomers 1000

b Monomer size 4

θ_{min} Bend angle 90

Flory-Huggins χ 0.45

R_e End-to-End 23.2

R_g R. of Gyration 9.5

φ_2 Concentration 0.02

φ* Critical 0.032

https://www.stevenabbott.co.uk/practical-solubility/polymers-in-solvents.php

you will find that in this very simplified view of blobs of polymers, the θ_{min} "bend angle" has a big effect on how much the polymer sticks out into the solvent, as does the monomer size b. Some polymers stick out a lot for their MW and so have a bigger chance to encounter other chains and, therefore, entangle. Although φ* is merely overlapping, rather than entangling, there's a good chance that a small φ* will lead to a small M_C

If you go to

Polymer Viscosity

R_g @0.5+40K nm 10

M_{mono} 100

Polymer MWt 40000

Polymer M_c 10000

Flory-Huggin χ 0.25

C_{max} g/dl 15

R_g @χ+MWt nm 13.0

Ne 33.3

C* g/dl 3.0

Ce g/dl 4.1

|η| dl/g 1.67

N.bonds 400

M_{bond} 100

Bond length nm 1.4

MWt 40000

R_g nm 11.4

https://www.stevenabbott.co.uk/practical-solubility/polymer-viscosity.php

you see that the viscosity starts to take off somewhat above a C* and massively at C_e which is when entanglement kicks in.

In that app you need to enter M_c so how do you know what it is? Well, your suppliers should give you these values routinely, but they don't. If they did you would know that a 100,000 MW polymer with M_c = 10,000 will have 10 tangles while one with M_c = 25,000 will only have 4. So the same MW of the first gives much higher viscosities, more tangles, than the other.

On that page you can find a list of all the M_c values I happen to know. Here it is in plain text:

Polymer M_c values

Polyethylene: 3300; Poly(vinyl fluoride): 3300; Poly(ethylene terephthalate): 3400; Poly(ether ether ketone): 3400; Poly(ether ketone): 3500; Nylon 6,12: 4000; Polyoxyethylene: 4200; Poly(1,4-butylene succinate): 4400; Poly(epsilon-caprolactam): 4800; Poly(epsilon-caprolactone): 4800; Poly(vinylidene fluoride): 5200; Polycarbonate (Bisphenol-A): 5200; Poly(propylene oxide): 5600; Polyoxymethylene: 5800; Poly(3-hydroxybutyrate): 5800; Polypropylene: 6400; Poly(glycolic acid): 6800; Poly(1,4-butadiene): 6800; Polylactic acid: 8600; Poly(vinyl alcohol): 10500; Polytetrafluoroethylene: 10800; Poly(vinyl chloride): 11500; Poly(acrylic acid): 13800; Poly(methacrylic acid): 15100; Polyisobutylene: 15200; Poly(vinyl acetate): 15400; Polyacrylonitrile: 15800; PDMS: 16600; Poly(methyl methacrylate): 20400; Poly(ethyl methacrylate): 27000; Poly(vinyl butyrate): 27600; Polyacrylamide: 28000; Poly(vinylidene chloride): 29200; Polystyrene: 31000; Poly(ethyl 2-cyanoacrylate): 59000;

It is, incidentally, no coincidence that all the great engineering polymers, PE, PET, PEK, Nylon, PC have relatively low M_c values – they are strongly internally tangled to be tough. PMMA and PS, both known to be brittle, have a high M_c. Superglue, Poly(ethyl 2-cyanoacrylate), is an especially weak polymer with a massive M_c!

Solubility

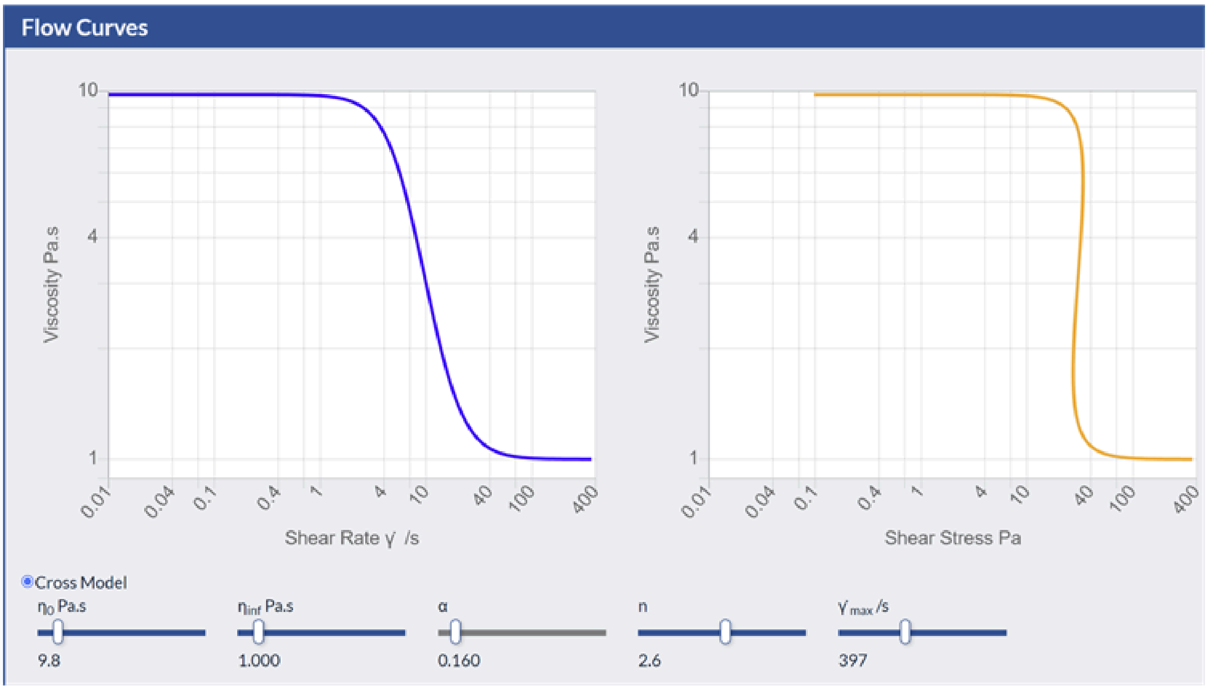
The other aspect of the previous app is the Flory-Huggins χ parameter, Dissolution_Flory-Huggins. When it's much less than 0.5 you can readily dissolve your polymer as long as it is largely amorphous – crystalline polymers, like PE, are insoluble because the solvent can't disrupt the crystalline structure.

What is critical is that at low χ values, as the first app shows, the polymer opens up (at $\chi = 0.5$ it's a tight, Gaussian ball), giving you greater practical entanglement and viscosity. The entanglement from a low M_c is only a potential – the polymer has to be opened up before it can tangle.

How do you match a polymer to a solvent to get a low χ value? Via Dissolution_Hansen Solubility Parameters, where you match the HSP values of your solvent (blend) and polymer. If it's an aqueous system, you simply choose your favourite high MW, low M_c, water-soluble polymer.

Flow behaviour

So far we've been concerned with low shear viscosity. As we (usually) need the formulation to flow at some point then we need a good flow curve – plot of viscosity versus shear rate and, as it's a good idea to think in this way, how the viscosity depends on the shear stress:



● Cross Model

η_0 Pa.s 9.8

η_{inf} Pa.s 1.000

α 0.160

n 2.6

$\dot{\gamma}_{max}$ /s 397

https://www.stevenabbott.co.uk/practical-rheology/Flow-Curves.php

Here the viscosity plummets at 10/s which may or may not be what you desire. How do we predict this behaviour from the polymer-solvent properties? At the time of writing, I don't know. So you need good access to a good rotational rheometer to get these data.

In practice

This means what you already know: there aren't too many useful thickeners which have the dual ability of providing large viscosity at a small concentration. Either you can't get a high enough MW or the M_C is too large, or the polymer you like, even if sufficiently soluble, isn't happy enough in the solvent to expand and tangle.

If your problem is the polymer/solvent compatibility, then via HSP you might be able to tweak your solvent blend to give you the lower χ value for greater entanglement.

And even if you get the right low-shear viscosity, there's no guarantee that you'll get the lower viscosity you require at a higher shear rate. There seems to be no substitute for a good rheometer and, perhaps, some building up of flow curves for different polymer/solvent/χ-value/M_C combinations – another potential example of "smart mapping" where you gather the same meaningful data over a range of relevant samples to help provide some useful rules in an area where theory, as in this case, lets you down in terms of pure prediction.

Thickeners Wormlike Micelles

Links

Hair Conditioners, Hair Shampoos, Soaps and Washing

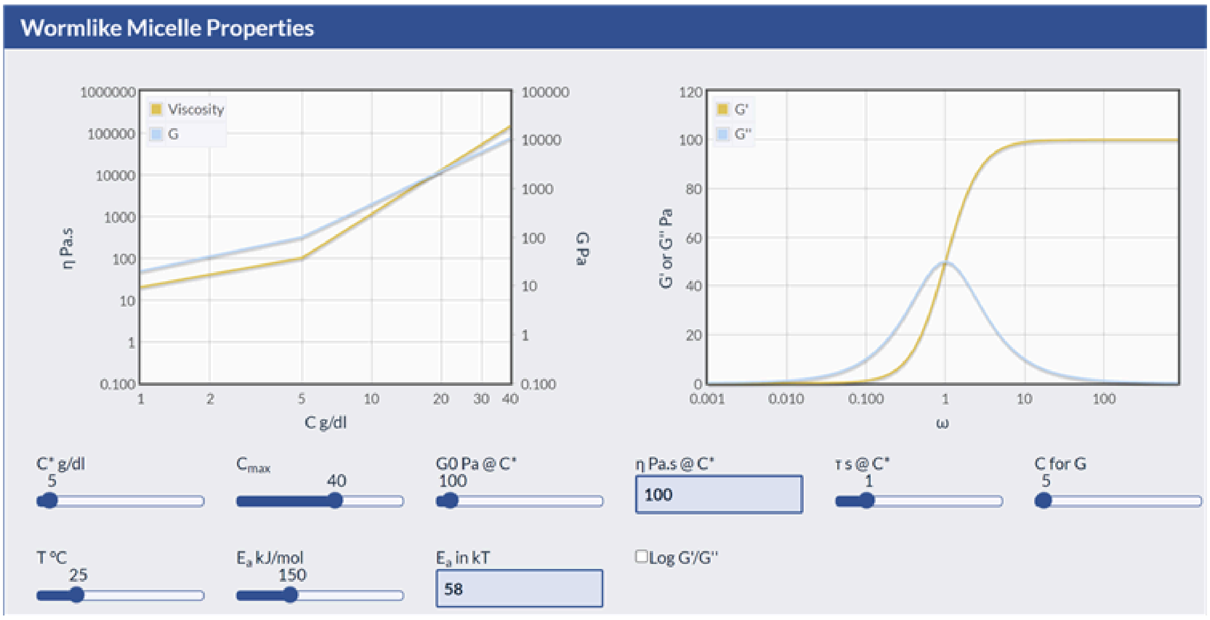
When you need a convenient thickener for a personal care product, adding some NaCl to a simple surfactant mix can do the job via the creation of wormlike micelles.

At the right combination of surfactant, concentration and salinity, some surfactants that were in typical spherical micellar solutions of low viscosity can transition to long "wormlike" micelles, shown in this image kindly provided by Seth Lindberg.

Although they are relatively delicate, if one micelle crosses another there is an entanglement force, so the system can be highly viscous. In soaps and shampoos this allows the formulator to provide a "luxury" viscous feel using only the surfactant and some sodium chloride.

Unlike Thickeners_Polymeric Thickeners, which can be too sticky, the relative delicacy of the wormlike micelles means that they can break up easily and, of course, viscosity decreases rapidly as water is added, reducing both the surfactant and salt concentrations. The reduction in salinity and surfactant concentrations changes the form of the micelles (usually back to spherical micelles) so the entanglement viscosity is lost.

Like other forms of entanglement (see Adhesion_Entanglement) the effect on properties such as viscosity depends on concentration to a high power (~3.5) above a critical entanglement concentration C*. We can capture such behaviour in an app:



C* g/dl 5

C_{max} 40

G0 Pa @ C* 100

η Pa.s @ C* 100

τ s @ C* 1

C for G 5

T °C 25

E_a kJ/mol 150

E_a in kT 58

☐Log G'/G''

https://www.stevenabbott.co.uk/practical-surfactants/Wormlike.php

C* is defined as a user input and the dramatic change in slope of viscosity is clear from this log plot.

Of equal interest is the rheology of these systems. Because the micelles are only loose clusters, if you probe them over long timescales, they are easily broken. Using the standard rheological measures G' (elastic) and G" (viscous) (Flow_G' and G"), we see from the right-hand graph that at low speeds (ω is the frequency of oscillation), viscosity is low, building to a high value at a specific speed which corresponds to the relaxation rate of the system. At higher speeds, the system becomes effectively an elastic solid, with a high G' and low G".

Flows like honey, wobbles like jelly

As explained by Neil Cunningham of the Centre for Industrial Rheology, the implications of the right-hand graph are that at low shear rates, gently tipping a bottle, the wormlike formulation should flow like honey (or, as in his excellent TA Webinar, corn syrup), while at high speeds it is elastic and wobbles like jelly.

The rheology curves are classic Maxwell equations. Given a reference G value, G_0, and a relaxation timescale τ, the values are given by:

$$G' = G_0 \frac{\omega^2 \tau^2}{1 + \omega^2 \tau^2}$$

$$G'' = G_0 \frac{\omega \tau}{1 + \omega^2 \tau^2}$$

The low shear viscosity, η, shown in the left-hand graph comes directly from the Maxwell equation:

$$\eta = G_0 \tau$$

The viscous and rheological effects are very temperature dependent via τ's dependence on an activation E_a in an Arrhenius equation:

$$\tau = Ae^{-\frac{E_a}{RT}}$$

Although the equations are useful, where do you get the values? The answer is that you use simple viscosity tests to get C* and standard rheometer measures to get τ. Then you have to reverse fit to match the graphs in the apps. Although this is unsatisfactory, it is still a relatively fast way to get core data on your system and how it changes with surfactants and salinity.

Upsetting the worms

So far we've thought of nice wormlike micelles in pure water with added salt. Our real formulation challenges involve plenty of other ingredients intended for other purposes. As has been noted by many, a standard Hair Shampoos can be nicely thickened, yet a simple marketing request to change the fragrance totally destroys the effect. One or more of the fragrance molecules gets into the hydrophobic micellar core and flips the subtle balance that had previously given the desired viscosity. Whether you can recover the viscosity by changing the salt concentration is unknowable in advance.

As there is no usable theory to predict the effects of additives to a wormlike formulation we resort to the standard approach in the FST: have a good rheometer with a well-defined method that lets you quickly scan formulations with systematic variations of salinity and ingredients. As you accumulate enough data, who knows, your AI might be able to provide you with the desired predictive power. This can't happen if you gather data randomly, with different methodologies, on unreliable rheometers. Not even an AI can disentangle such data.

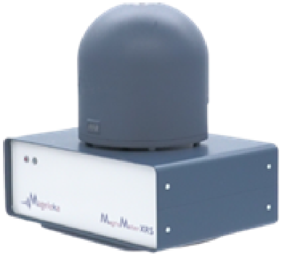
- Extras -

As the book progressed I needed to add one “extra” then a second. If readers find the need for more extras, I’ll be happy to add them.

Analytical Techniques

Here we look at some ideas that are slightly away from the standard uses of standard analytical techniques, providing some deep formulation insights. The inspiration for this chapter came from a lecture by Dr Tobias Halthur of CR Competence AB. He kindly provided some key images, below, so you get his expert view of the topics.

NMR Relaxation for particle/solvent interactions



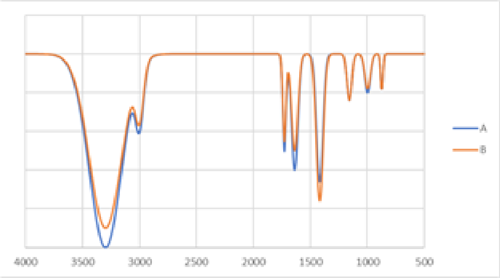
We add our particles to a solvent in whatever manner we choose, and put the results into a normal NMR tube. The tube is then inserted into a small, relatively simple benchtop NMR machine (this image is from Mageleka) that does *not* measure an NMR spectrum. Instead, it issues a chain of pulses and follows the relaxation of the pulses. A run with the solvent on its own establishes the standard, relatively slow relaxation time. If the solvent has little interaction with the particles, the measured relaxation time will be little changed from the reference. But strong solvent-particle interactions provide a large change in relaxation time.

Via a systematic comparison of the relaxation behaviour of the particles with different solvents you get a good idea of how the particle interacts with its surrounds.

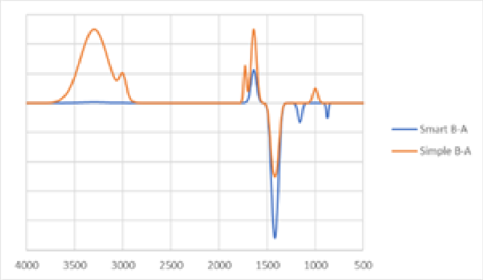
If the particle is covered by some sort of dispersant, the relaxation effects are more subtle. If the dispersant is nicely extended into the solvent, then the solvent will see a relatively fluid boundary and the relaxation time will change much less. Like all techniques, you have to have good hypotheses and use good sense to interpret the results.

You don't *have* to use the benchtop machine. A classic NMR machine can also work in relaxation mode. It's your choice of tying up your fancy NMR machine or having a benchtop machine for regular particle-solvent analyses.

FTIR with smart subtraction



A typical FTIR spectrum is relatively complex and if you need to do a before/after comparison of some process or reaction, or if you need to compare your formulation with, say, a similar formulation from a competitor, it can be difficult to extract meaning from a side-by-side comparison. Even in this simplified spectrum it's hard to know what is going on between sample A and B. Real, more complex spectra are even harder to compare.



The obvious alternative is to subtract one spectrum from another and look for differences. However, small differences in concentration or measurement setup can lead to too many differences, positive and negative.

Smart subtraction requires you to be confident that at least one peak in the spectrum should be an unchanging reference between the samples. So you adjust the subtraction so that this peak (or maybe 2 or 3 peaks) are effectively removed. In this example, simple B-A subtraction gives differences everywhere. If we know that the -OH peaks around 3400 cm^{-1} should be unchanged, we tune the subtraction so this area is flat. Now the resulting negative and positive peaks should represent real differences between the samples. With knowledge of what the peaks showing differences might be, it's possible to formulate a hypothesis of what's going on.

FTIR of aqueous complexes

Systems such as Surfactancy_Coacervation, Thickeners_Associative_Thickeners and Thickeners_Wormlike Micelles involve complex interactions between water, surfactants, polymers, salts and additives. For some users, the bulk properties are important, for others (such as the complexes left on the surface of Surface Cleaning) it is the surface effects that are important.

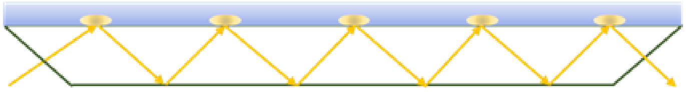
In both cases it might seem difficult to get meaningful FTIR spectra, but with the right setup, you can get a lot of useful information from these complex systems. The review by David Scheuing, *Fourier transform infrared spectroscopy in surfactant science: A personal view*, J Surfact Deterg. 2022;1–25, is intended for the practical formulator in this complex space. Although the review is for "surfactant science" the principles apply to similarly complex systems.

Information from bulk solutions

For bulk solutions the trick is to have a precision spaced (~25µm), temperature-controlled aqueous cell, with windows of water-insoluble CaF_2 or BaF_2. Temperature control is vital because the IR absorbance of water changes significantly even over 1°C. You take a reference spectrum of whatever "water" is relevant (typically you will have specified salts and pH), then use that, via smart subtraction, to remove as much as possible of the water spectrum in order to amplify the relatively weak and complex signals from your sample. The main peak ~3400 can be relatively easily nulled out. The peak that might cause more trouble with functionalities in your system is the one at 1640 which is itself influenced by interactions with solutes. Switching to D_2O may be necessary if you really need data in this region.

Although you can imagine that, say, carboxylate absorption for a soap should shift at the CMC (which it does), it is not obvious why a spectrum might provide any insights into the surfactant tail region; it is just some boring CH2 groups with a terminal CH3. However, as the review points out, the various wagging, bending, stretching modes change in relative intensity depending on the degree of ordering in the tail. The ordering of tails in a spherical micelle is different from tails in wormlike micelles, and there are diffierences between tails of systems with different curvatures at oil/water interfaces. The review has many specific examples of spectra from real-world systems showing the sorts of effects of interest: wormlike micelles, phase changes, ethoxylate behaviour at the phase inversion temperature, microemulsions, mixed surfactant systems.

Interfacial FTIR



Those familiar with ATR (Attenuated Total Reflectance) know that a crystal can be placed next to a surface and a spectrum easily gathered in the region where the light partially penetrates across the interface. But the spectrum is the average over the 1-2µm penetration depth, so is useless for picking up nm-scale effects such as coacervate deposition. However, using a multi-reflection crystal, you can get 10-25 reflections so the effective pathlength is 10-25x longer. This means that a signal from a nm-scale layer can be observed after subtraction of the bulk signal, as in the previous section.

This is fine if you want to see interactions of surfactants and polymers with the surface of your, say, Ge crystal. But generally we need a specific surface. There's no problem in principle of adding a layer thinner than the penetration depth. The limitation is your confidence in applying a layer to your expensive crystal and being able to remove it later. As the review notes, it's not so hard to add alumina (sputtering) or titania (from solution) to the surface.

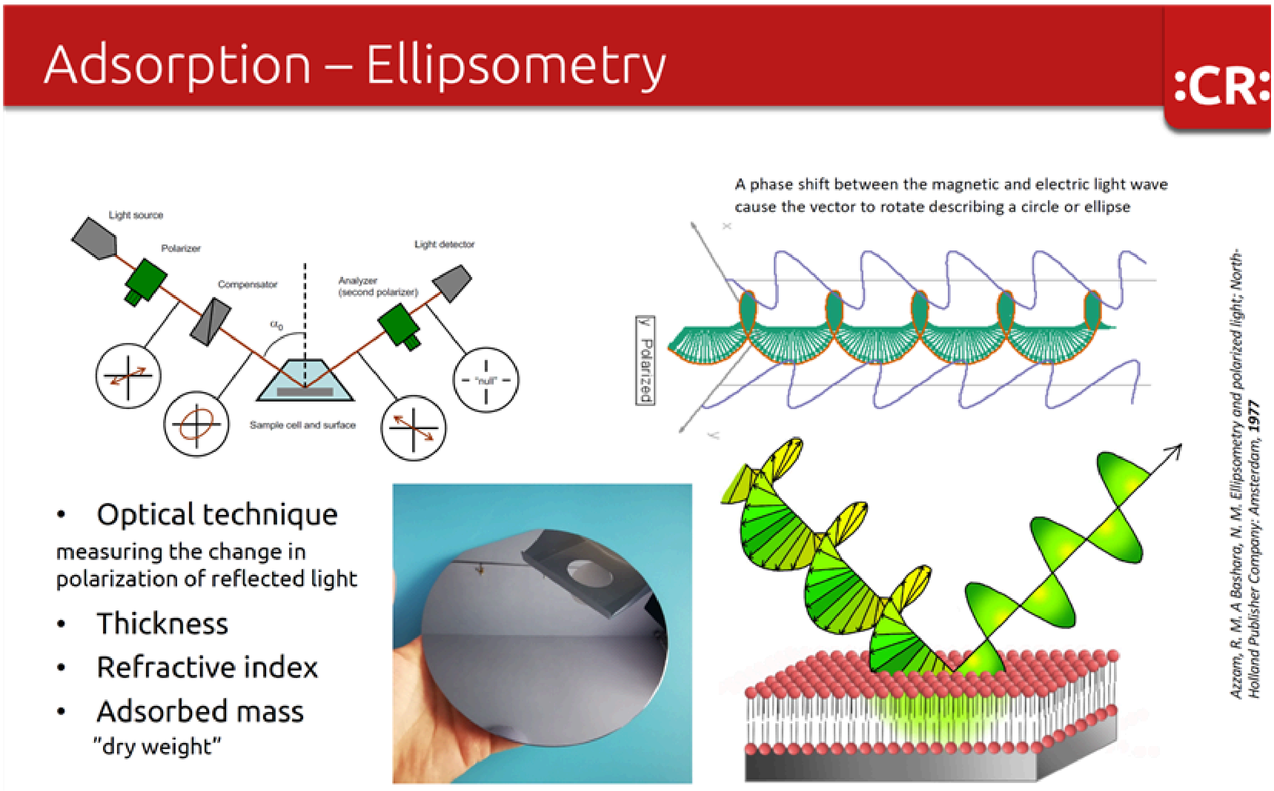
If you take the trouble to work out how to do this, you can be rewarded with insights that tend to over-turn the simplistic diagrams we like to make of polymers and surfactants at surfaces. For example, a monolayer of SDS resolutely refuses to provide the ordering we assume is the norm – the various CH2 modes don't change into those of packed layers. And when you add polymers and electrolytes, it is clear that what's obtained on the surface is a loose fluffy agglomerate, with the final form depending not just on the stoichiometry, but also on the order of addition. These structures have a large element of kinetic control, despite what the thermodynamics would suggest.

The ellipsometry and QCM experiments shown below very much confirm these sorts of conclusions.

A fascinating set of studies included in the review starts with a "thick" (~1 µm) layer of an oil and follows its removal via detergents in real time. Using oils of different melting points and ethoxylates with different HLD behaviour it was possible to gain deep insights into the processes involved.

Ellipsometry

This image, kindly provided by Dr Halthur sums up ellipsometry:


Adsorption – Ellipsometry

:CR:

A phase shift between the magnetic and electric light wave cause the vector to rotate describing a circle or ellipse

- Optical technique

measuring the change in polarization of reflected light

- Thickness
- Refractive index
- Adsorbed mass

"dry weight"

Azzam, R. M. A Bashara, N. M. Ellipsometry and polarized light; North-Holland Publisher Company: Amsterdam, 1977

If a beam of circularly polarized light is reflected from a smooth surface, the amplitude and the phase of the reflected polarized depend on the complex refractive index of whatever is on the surface, and on its thickness. “Complex refractive index” means, in normal language, its refractive index and its absorption coefficient for light at the given wavelength.

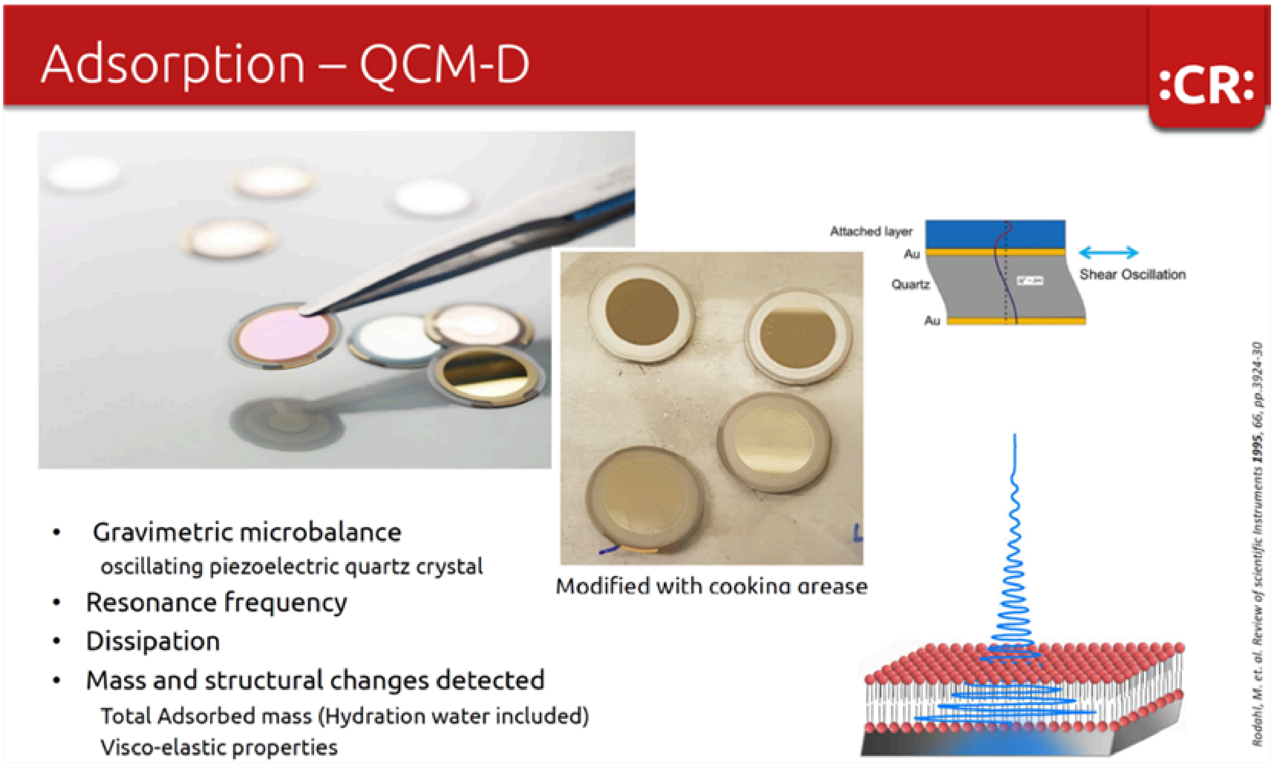
The classic technique is done with dry coatings in air and is sensitive down to low nm of thickness. With care and experience, the experiments can be done under liquid, which makes it more interesting to formulators. What you get from the measurements are:

- Thickness
- Refractive index
- Adsorbed mass – dry weight (even though that mass might be swollen)

A specific example is shown below, comparing/contrasting with the QCM-D.

QCM-D (Quartz Crystal Microbalance with Dispersion)

Another image from Dr Halthur helps us to grasp QCM-D:



Adsorption – QCM-D

- Gravimetric microbalance
  oscillating piezoelectric quartz crystal
- Resonance frequency
- Dissipation
- Mass and structural changes detected
  Total Adsorbed mass (Hydration water included)
  Visco-elastic properties

Modified with cooking grease

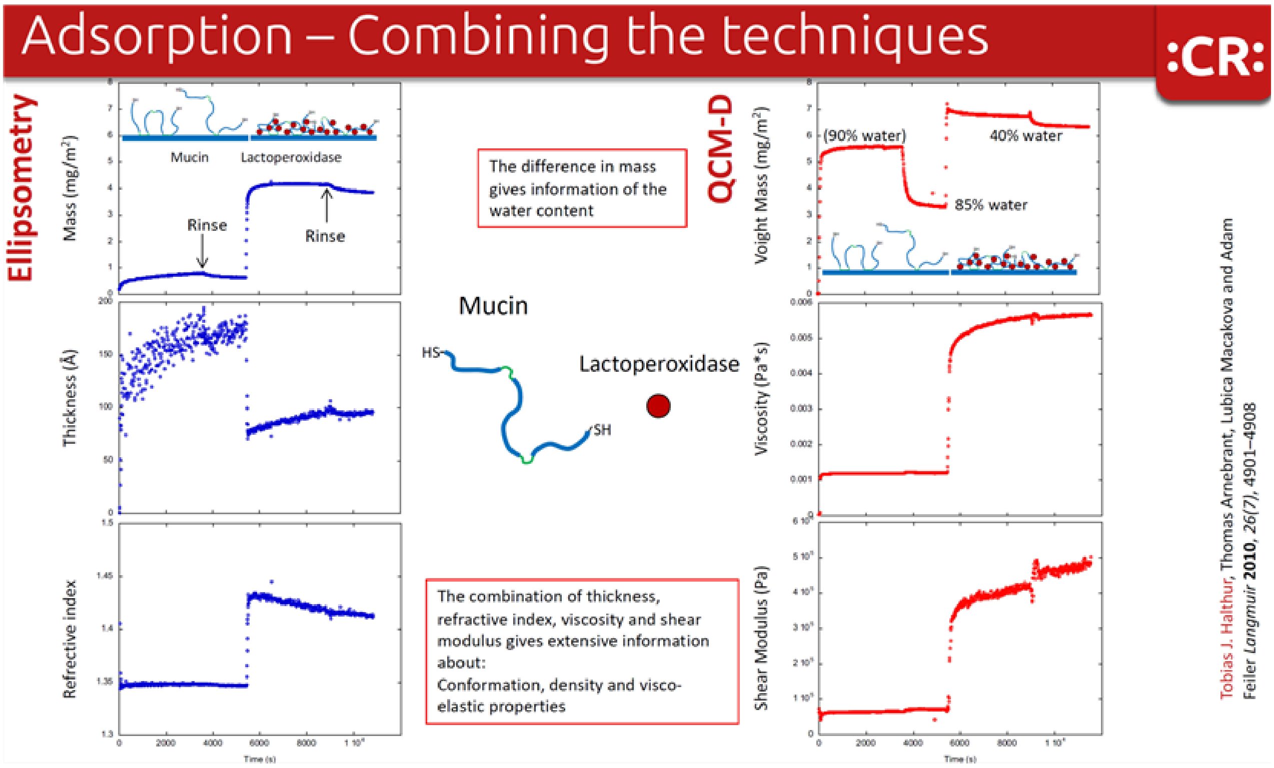
Rodahl, M. et. al. Review of scientific Instruments 1995, 66, pp.3924-30

Put a quartz crystal into your solution and get it to oscillate at its resonant frequency. Because of the viscosity of the liquid, this frequency will be different from that in air. If you have a QCM-D machine (D for Dispersion) you can see how the amplitude of the oscillations decay when the stimulating signal is switched off.

Now gently flow a solution of interest over the crystal. If something in the solution wants to accumulate on the surface, the mass of the crystal *in*creases, so the resonant frequency *de*creases. The effect is exquisitely sensitive so you can pick up nm deposits. . What you get from the measurements are:

- Total adsorbed mass (including the liquid)
- Visco-elastic properties

A further image from Dr Halthur lets us see a specific example (the reference is included in the image) of how the two techniques reveal complementary information:



Adsorption – Combining the techniques

The difference in mass gives information of the water content

The combination of thickness, refractive index, viscosity and shear modulus gives extensive information about:
Conformation, density and visco-elastic properties

Tobias J. Halthur, Thomas Arnebrant, Lubica Macakova and Adam Feiler *Langmuir* **2010**, *26(7)*, 4901–4908

Some mucin, a polymer with thiol groups at each end is deposited from solution and rinsed. It may or may not be attached by one of the thiol groups and we can imagine many ways that it sits on the surface – flat and compact or sticking up as a loose bundle. The ellipsometry measures its dry mass, which is relatively low – no surprise because it's a small deposit. The thickness is reasonably large, and the refractive index is very close to that of water. The QCM-D seems to tell a different story. The mass increase is large. But then we recall that the device records the total mass, including water absorbed along with the mucin. The low refractive index seen in the ellipsometry now makes sense because it is seeing a highly swollen mucin layer.

When lactoperoxidse is added, the mass increases, substantially, according to the ellipsometry, but the QCM-D shows only a small mass increase. The ellipsometric thickness shows a sharp decrease and the refractive index increases – all indicating that the complex of the two polymers is highly compact. Finally the viscosity and shear modulus from the QCM-D both show a significant increase, which is consistent with the data from the ellipsometer indicating a more substantial complex layer.

With beautiful results and excellent diagrams, it all seems rather straightforward. Like all sophisticated techniques, things aren't so hard if you have high-quality equipment and years of training. The point is not for you to rush out and become instant experts but to know that these capabilities exist and, as with all modern technology they tend to get more automated and more powerful so, maybe, more routine for when you need to better understand what's happening at an interface.

Beware Superduperometers

The best analytical technique is the one that gives you key information quickly. You ask the right question and the equipment gives you helpful answers.

However, it is frequently assumed that the best technique is some superduperometer, a machine so new, powerful, complex, expensive that it *must* be good. If you put your sample into the machine it will generate lots of amazing numbers you couldn't have obtained any other way. But if you are not asking it the right question and it's not giving you helpful answers, it's a waste of time.

Yes, there are times when a critical issue can only be resolved using a fancy machine. Go ahead, use it. The techniques highlighted in this chapter are, in the hands of those I know who use them, great examples of how to learn stuff you otherwise could not know. But academia and expensive corporate central labs are full of fancy machines that provide very little problem-solving insight for a lot of money.

And this is the 21st century. It is amazing what you can do with some ingenuity, an Arduino, some sensors, a 3D printer and a good hypothesis. Add some LED lighting, decent optics and a laptop or phone to record some images or video and you can get a feel for what is going on at some key interface. I measured the coefficient of friction of various cheeses on various packaging films on my kitchen table. An order to Amazon and a wait of a few days and I could measure some important thermal conductivities that a megacorp admitted would have taken months to organise. When a really nice pump system, required for a key experiment, got damaged in transit, we got the data we needed using a syringe and a steady hand. You can have controlled RH chambers with a fish tank, some water and some well-chosen crystals. Someone I know replaced a complicated robotic paint-brush cleaning system by using disposable 3D printed brushes and saved $1000's by using a standard oven rather than a fancy automated oven … by using the robot to press the On button on the cheap oven.

Two workhorse rheometers will probably be better for you than one fiddly, expensive one locked in an expert's lab. A routine FTIR measurement when I was visiting a megacorp was a nightmare because instead of a couple of lab machines they had a super-fancy one, run by an expert in a "Do not enter" lab which you had to book days in advance.

Another problem with superduperometers is that you often have to submit your sample and wait till the data comes back. Despite your careful explanation, what they measured turns out to be what they *thought* you'd asked, not what you *really* asked. A good rule is "Never get anything measured on a fancy machine without you being present". What is obvious to you isn't obvious to the operators, so they will miss pointing their ion beam gun at the one thing that could solve your problem and, instead, gather data to prove something you already know. This is a slow and expensive way to learn nothing.

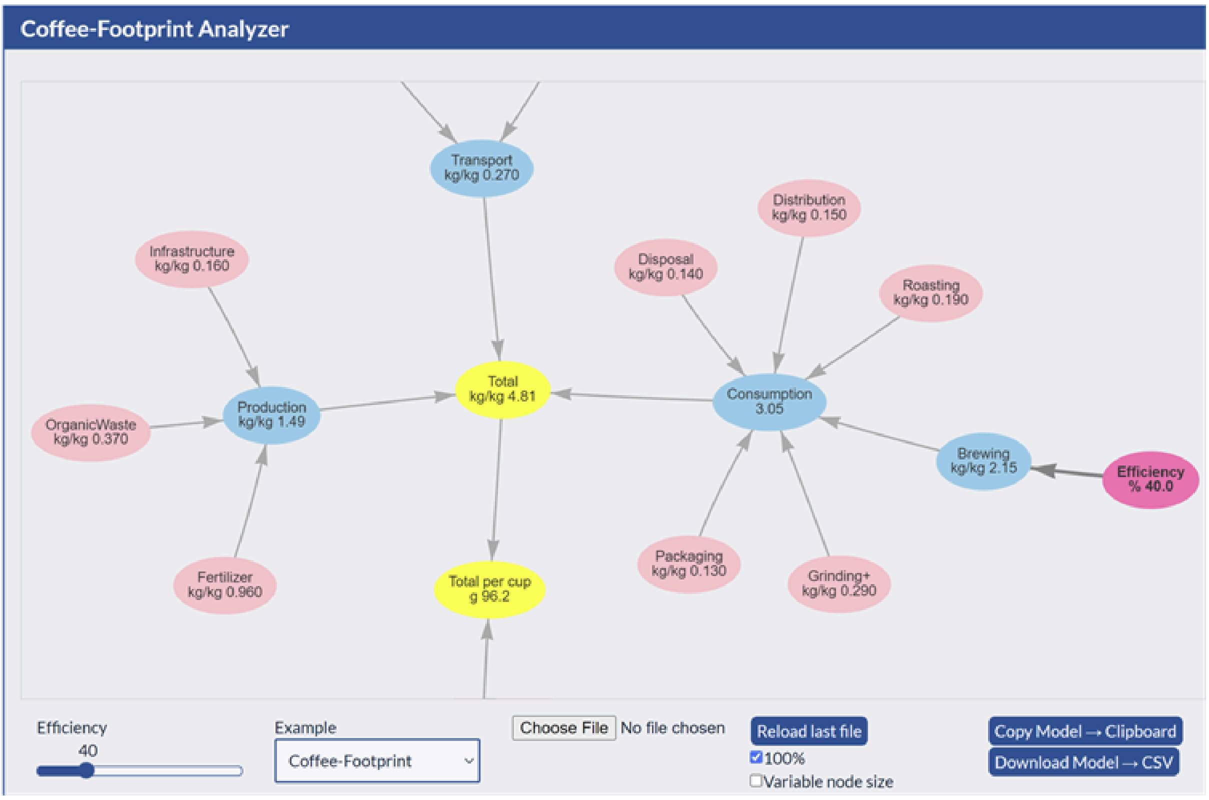
Find the simplest, quickest way to answer a clear question to help solve a real problem. Sit at the set-up and start asking the questions … and discover that there's a glitch in the idea. Modify, try again, find another glitch … and quickly you get the answers you need. Lean, agile, fail fast aren't the answer to everything, but for formulations they are much more likely to be helpful than any superduperometer.

Open Analyzer

Many readers will be formulating for greener products. It turns out that good intentions are not the same as good outcomes, so it’s important to check sooner, rather than later, that your precious resources aren’t going to be spent on something that is not only not better for the planet but actually worse than the current product or process.

A full LCA (Life Cycle Analysis) is the gold standard for this. But they are cumbersome and bureaucratic. Sometimes a less formal “what if” tool can give you the answers you need. Open Analyzer is a handy tool for early exploration of whichever part of the system is important to you.

Interactive What Ifs



Efficiency 40

Example Coffee-Footprint

☑100%

☐Variable node size

https://www.stevenabbott.co.uk/Open-Analyzer/Analyzer.php

Suppose you wanted to know how to reduce the CO2 footprint of a cup of coffee. You might want to reduce cost of transportation, roasting, packaging, fertilizers … things that are someone else’s responsibility. But if you assemble all the relevant parameters and play with the effects of, say, halving them (by clicking on any of the pink ovals) you find that most make little difference … compared to the efficiency with which you heat the water to make your cup of coffee.

I could tell you as a fact, or I could provide a full LCA, but having a model that you can interact with and try out “what ifs” is an efficient way to understand what’s going on.

But suppose you found this model was lacking some key parameters. No problem, you copy or download the model, alter it in Excel or Google Docs (the data are just simple .csv) then load your new version. Because it’s just an app, you can have two tabs open with a different model in each to be able to compare and contrast.

If you look at other examples you will find attempts to replace some conventionally produced chemical with a bio-based one, and learn that after maybe 10 years of hard work you might save 1 kT CO2/year, the equivalent of 1 second of current human emissions. Is it a good idea to put that much effort into such a small saving? Would it have been better to put your precious resources into something that could save 1 MT?

Another example is the famous “plant bottle” for fizzy drinks, using bio-based ethylene glycol. The Open Analyzer analysis shows that it is worse for the planet than the conventional one, a truth backed up by a full LCA. Happily, the originators of the plant bottle decided that using recycled PET was a much better strategy, which it is.

How to use Open Analyzer

The inputs

You need an Excel .csv file containing the Nodes. If click Copy Model on any of the datasets that interest you then paste into Excel, or if you click the Download Model, you will see that the format is relatively simple, though getting everything right is, I admit, tricky. Each node is identified by a unique id, then has a set of other properties. One key problem is that in the world of Javascript, everything is Case Sensitive. So when you specify an id you can't specify an Id or an ID. And if you have a node called MyNode you cannot have formulae based on myNode or Mynode (unless you confused yourself by creating different nodes with those id values!). If you have an optional file with the same name but with a .html extension then this assumed to be the HTML you wish to include on the page to help explain what is going on. If you write this in an ordinary text editor at the least provide an HTML headline using a few HTML tags

This is your title

and start a paragraph with the

tag.

Loading a Network

From the combobox you just select a model and the job is done. To load your own then (for obscure Javascript security reasons) you need to load the Nodes file first then, optionally, the .HTML file.

Sensitivity Analysis

If you select one of the inputs and provide a value, you automatically get a sensitivity analysis! Each time you change something, the values of every node are automatically put on the Clipboard using the minimum and maximum of that slider along with your current value. Just paste into Excel and you can analyse it any way you want. Why not add a Sensitivity Analysis function? Because it complicates the interface and doing it automatically gives you the maximum benefit for the minimum work

Your own input files

Here are what each of the inputs mean for Nodes:

- id: The name of the node.
    - If this id = xxx or XXX then everything in that line will be ignored. This lets you add full-line comments to your .csv file
    - If this id = hierarchy then the label field should be UD, DU, LR, RL to describe up/down, down/up, left/right or right/left hierarchical layout. For this you need an extra column called "level" where you give numbers starting at 1 for the different hierarchical levels. The quality of the view is not guaranteed as hierarchies are complex to get right. Clone the HomeCompost-H example to check how it works
- label: What appears in the box
- units: Units (optional) that precede the value on the lower line of the label
- title: Tooltip pop-up text. The column and individual tooltips are optional
- color: Choose between pink, blue, yellow, salmon, green, red, cyan, magenta and white. There are complicated reasons why you can't just enter your own colours. My convention is pink for inputs, blue for simple intermediates, salmon for aggregated intermediates and yellow for stand-out conclusions. You can choose your own scheme, but this one seems to work well over a range of models
- val: The value of a known input. This must always be the actual number. So if something is 1 million, it must be 1000000. If this node produces calculated values, put a ?
- conv: To display the number in the box, multiply by this value. So to see 1000000 as 1.00, set conv to 0.000001 or 1e-6. Use m in the label to indicate to the user that the value is in millions
- from:, to:, step: Most inputs make sense only over a range which you specify in converted units as from and to, with step showing the precision on the slider. So if you want to go from 100 thousand to 10 million in steps of 100 thousand and your units are in millions you would have from:0.1, to:10, step:0.1
- equn: This is the trickiest, but not so hard. Suppose you have two inputs called Factor1 and Factor1 and you wanted the node to be their sum. You just put {Factor1}+{Factor2}, i.e. you type the sum, but put the names in curly brackets so the software can work out what is what. You can use any Javascript code you like1. With even the smallest amount of effort (and attention to brackets) you can easily tell the meaning of 3*(Math.sqrt({Factor1})+Math.pow(4.2+{Factor2},2))
- comment: This comment is for your own use as a comment about that row. But if a comment is present it appears as a new line within the Tooltip, under the "title" text

Errors in your network

It is entirely normal (i.e. it has happened to me many times) to have some errors in your .csv file. For example if you have a node called MyNode you might include {Mynode} in a formula, or you might have typed (MyNode} with one curved and one curly bracket. When you load a model, the code attempts to alert you with messages that should make it easy to spot and fix the problem. You may get more messages than you might like (an error may have knock-on effects), but it is better to have too many than too few. When you correct the .csv, reload it and keep going till you get none of those sorts of errors. Of course, you will probably have other sorts of errors in your logic so you will need a few further cycles of refinement.

Capturing a different version of a model

You may have a model with one set of inputs then decide you want to keep a version with inputs you have just adjusted on screen. Click the Model → Clipboard button and the model is ready to paste into Excel and save as a new .csv. Or click Download Model → CSV if you prefer to have the file delivered directly to your download directory.

A database of input values?

Those in the green biotech world who want to use Open Analyzer to model their processes would love to have a set of standard values for many of the common processes. I could easily create a Standard Values database if I had the data. But I'm not a biotech expert. If anyone wishes to help create a database of values I will be happy to do the coding and, of course, their contribution would be fully acknowledged.