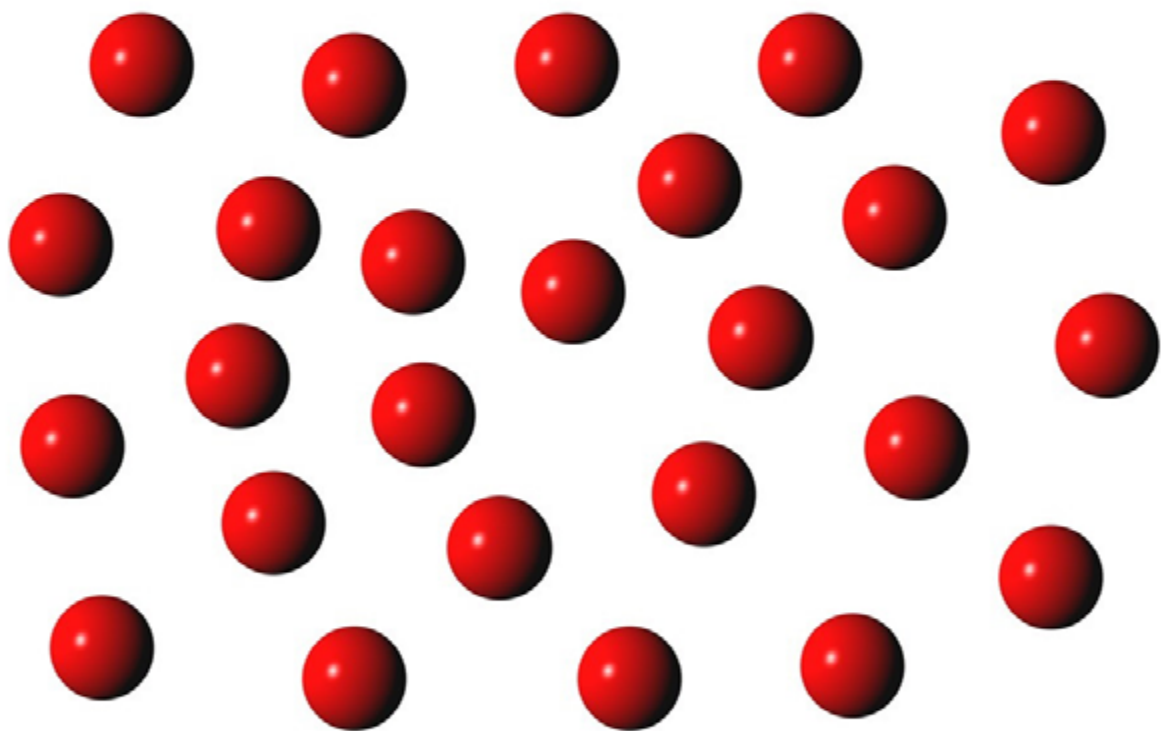


Particle Formulation Science

Principles & Practice



Prof Steven Abbott

Particle Formulation Science: Principles and Practice

Prof Steven Abbott

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

steven@stevenabbott.co.uk

www.stevenabbott.co.uk

Version history:

First words written 05 October 2018

Version 1.0.0 19 July 2020

Version 1.0.1 8 May 2022

Copyright © 2020-2022 Steven Abbott



This book is distributed under the Creative Commons BY-ND, Attribution and No-Derivatives license.

Contents

Preface	6
Abbreviations and Definitions	9
1 A few basics	10
1.1 ϕ	10
1.2 Size Distribution	11
1.3 Aspect Ratio	12
1.4 Percolation threshold	12
1.5 Dispersions, Colloids?	13
1.6 Tortuosity	14
1.7 Interparticle Distance	15
1.8 Measuring particle size	15
1.9 Lattice theory and the Flory-Huggins χ parameter	17
2 Controlling Stability	19
2.1 Fighting sedimentation	19
2.1.1 Centrifuges	20
2.2 Self-association	21
2.3 Fighting self-association	21
2.3.1 Lyophilic and lyophobic	22
2.4 Core DLVO	23
2.4.1 DLVO	23
2.5 Zeta potential ζ	28
2.5.1 Extended DLVO	30
2.5.2 DLVO as a religion	31
2.5.3 The secondary minimum	32
2.5.4 Summarizing DLVO	33
2.6 Solvation and solubilization	34
2.6.1 Looking for differences	35
2.6.2 Fundamental thermodynamics	38
2.6.3 Using solubility	40
2.7 Polymers on particles	40
2.7.1 Locked and loaded	41
2.7.2 Bridged	42
2.7.3 Simple depletion	42
2.7.4 Vincent depletion	43
2.7.5 Beyond bridging and depletion	44
2.7.6 Dissolving silicon in DMSO	44
3 Solubility of particles	46
3.1 HSP Basics	46
3.2 The Distance D	48
3.3 The HSP values	48

3.4	Measuring HSP	50
3.5	The radius of a polymer HSP	52
3.6	The Flory-Huggins χ parameter	53
3.7	Particle HSP values and radii	54
3.8	An objective measure	56
3.8.1	The transmittogram technique	57
3.9	Happiness from NMR	59
3.10	NMR and solubility theory	60
3.11	FTIR	62
3.12	So what?	62
3.12.1	HSP and Hamaker A_{12}	65
3.13	Going to the next level	65
3.14	Particle solubility? It's your choice	65
4	The practical power of SF/SCF theory	67
4.1	SF/SCF Theory	67
4.2	Where is the polymer?	70
4.3	Di-blocks, Combs, Brushes and A+B	72
4.4	Adsorption isotherms	74
4.5	Interparticle Interactions	75
4.6	Tails repel	78
4.7	Thermodynamics and kinetics	79
4.8	Using HSP for SF theory	80
4.9	Imagine a world where ...	82
4.10	SF of aqueous systems	83
4.11	Appendix: An SF master class	83
5	Rheology of particle systems	86
5.1	Low Shear and Yield Stress	87
5.2	Yield stress	91
5.3	High shear rheology	92
5.4	Self-associating particles	94
5.5	Thixotropy	95
5.6	That's not a long chapter!	97
6	Understanding the surface	98
6.1	Fancy techniques	98
6.2	Surface energy	98
6.2.1	Washburn tubes	99
6.2.2	Simple IGC	100
6.3	Surface areas	102
6.3.1	Via particle size	102
6.3.2	Via "direct" area measurements	102
6.3.3	Use with caution	103
6.4	Subtle surface properties	103

6.4.1	The 5-fold way	103
6.4.2	Surface chemistry via IGC?	105
6.5	Surfaces via HSP	105
6.6	Surfaces in real formulations	106
6.6.1	The reality of surfaces	106
7	Formulating in water	108
7.1	DLVO and zeta	108
7.2	Balancing a formulation	110
7.3	Neutral molecules matter too	112
7.4	Another very short chapter	112
8	Simulating Particle Solubility Effects	113
8.1	Particles in a box	114
8.2	A few integrations	114
8.3	A sliding scale	115
8.4	The potential of mean force	115
8.5	How stable is stable?	117
8.5.1	25 kT	118
8.6	Gels	119
8.6.1	From ergodic to icosahedral via isostatic	120
8.6.2	A different percolation	122
8.6.3	How they did it	124
8.7	So what?	125
9	Capillary Suspensions	127
9.1	Just a drop of oil (or water)	127
9.2	The basic rules	128
9.3	It's not so easy	131
9.4	Tips and Tricks	132
9.4.1	Understanding the secondary fluid	132
9.4.2	Understanding the particles	133
9.4.3	Creating the suspension	134
9.4.4	What can go wrong?	134
9.4.5	How do you know what you've got?	135
9.5	Making porous solids	135
9.6	Percolated polyhedra	136
9.7	Capillary foams	138
9.8	Summary	140

Preface

I once found myself saying to a Dutch¹ colloid scientist that after many decades of colloid science, we had very few usable tools that someone with a particle formulation problem could turn to. Given that there are many Departments of Colloid Science in universities, many books on Colloid Science and a few Nobel Prizes on the topic, it seems obvious that my remark must be wrong.

But the Dutch scientist didn't disagree too much, and the more I looked into it, learning lots of science along the way, the more I found it to be true. If I look at the literature, most of it is unusable. At one time I thought that maybe I was unlucky in the topics that interested me, or was being over-hopeful about what academic applied research should be achieving. With recent studies across wide areas of science showing that 70% plus of science is somewhere between wrong and useless², it is now clear that my own judgement was not far wrong. If I look at the formulation community, it is very rare to meet formulators who rely on strong science-based principles. In general they prefer to go with experience and intuition. I am now part of a European team promoting "Science-Based Formulation". This should be entirely unnecessary as all formulation should be science-based, yet experience, and the response we have received, shows that the approach is very much needed.

Fortunately, a wise and experienced senior research manager provided me with a key piece of encouragement. He had consistently found that in his labs in a large chemical company, those with a colloid science background were better formulators than those with more traditional chemistry or chemical-engineering backgrounds. He, too, found that they had rather few science-based principles; but their training had given them a general feel for the sorts of things that go on in a complex particle/colloid formulation and this made them better able to arrive at workable solutions.

This observation has freed me up to write in a rather different style. In my other books I felt I had to fight against the bad guys who'd ruined the topic for everyone: surface energy in Adhesion, HLB in Surfactants, "water structure" in Solubility, surface energy (again!) in Printing and, oh, yes, surface energy in IGC. In this book I will take the core science and find ways in which it can be made usable in everyday formulation challenges. I will still get excited about the use of the word "dispersion" which turns out to be ill-defined and will show that allowing the word "solution" to be used opens up a wide vista.

1 There is a strong Dutch tradition of colloid science so it's quite risky to make such a statement.

2 I have a peer-reviewed publication saying this: Steven Abbott, Solubility, similarity, and compatibility: A general-purpose theory for the formulator, *Current Opinion in Colloid & Interface Science*, 48, 65-76, 2020

At one stage the book just wasn't going where I wanted, but a series of those chance encounters that are so important to science led me first to some key NMR science of which I had been unaware and this in turn led me to Prof Terence Cosgrove at U Bristol who, in addition to helping me in NMR introduced me to Prof Frans Leermakers at Wageningen U. This then opened up the whole topic of Scheutjens-Fleer theory which turns out to be the key to addressing many of those things that have puzzled formulators. Thanks to Prof Leermakers' "SFBox" we can now directly explore steric stabilization, di-blocks, combs, brushes, depletion and bridging flocculation. Instead of hand-waving (which is about all we could do without this theory) we can calculate the effects directly. The access to this excellent and well-validated theory allowed me to return to the book with renewed energy.

"Particles" can mean just about anything. In this book I am restricting the meaning to small particles and those that have some intrinsic strength, so it's the harder parts of soft matter up to solid particles. This means that I am excluding emulsions, even though much of the physics applies equally well. "Formulation" can also mean just about anything. Here my concerns are with the more chemical/industrial applications rather than directly with food formulations, even though foods are clearly both chemical and industrial.

In all cases I am assuming a formulation with plenty of liquid, at the very least enough to more-or-less fully wet the particles. This means that I am not addressing the fascinating problems of handling dry powders, where the science involves a very different set of physics principles.

I am going to assume that readers have arrived at the book because they are already faced with formulation challenges. If you need a whole textbook on standard colloid science, this isn't the book for you. My working assumption is that *despite* some knowledge of aspects of colloid science, you don't have a good set of tools to solve your very real formulation issues.

As with my other books, where possible, key ideas and formulae are linked to on-line apps so you can explore things live. Just click on the link and you can immediately start exploring. The apps are standard HTML5/Javascript/CSS3 so they run on phones, tablets and laptops, are safe on corporate networks (the standards prohibit unauthorised access to your device), and are free and free of ads. Most of them were written in a different context, for example coatings or solubility so they are not specific to this particles book. However, with modern web design I've been able to bring them together in <https://www.stevenabbott.co.uk/practical-particles/> so you can access them as a coherent group if you wish.

The "where possible" caveat in the previous paragraph means that the most advanced implementation of SF theory at the heart of this book, along with advanced handling of HSP theory, are best done within the commercial HSPiP

(Hansen Solubility Parameters in Practice) software of which I am a co-author. I generally get annoyed with authors who use a book to promote their own software but in the context of a free eBook and so many free apps, and with the excuse that advanced SF theory really is complicated to handle, I hope the reader will forgive this lapse.

I would like to thank those who have helped me along the path of this book. Their insights, advice, data, disagreements and generosity are very much appreciated. In alphabetical order:

- Shalmali Bapat for her Transmittogram idea
- Prof Terence Cosgrove for NMR and SF guidance
- Dr Ralph Duempelmann for key insights on IGC
- Dr David Fairhurst & Dr Shin'Ichi Takeda for insights on NMR methodologies
- Dr Charles Hansen & Dr Hiroshi Yamamoto for all the HSP particle knowledge discussed over our years working together
- Prof Frans Leermakers for all the SF insights and access to SFBox
- Prof Dietmar Lerche & Prof Doris Segets for guidance on the LUMiSizer methodology and wise discussions on many other issues
- Dr Seishi Shimizu for key insights on nanoparticle thermodynamics and solubility science
- Titus Sobisch for his spirited defence of particles being "dispersions"
- Sander van Loon & Beverley Fricker of VLCI for help with particle-related HSPs

Responsibility for errors in describing their work is entirely mine.

Steven Abbott
Ipswich, 2020

Update May 2022

I had been totally unaware of Capillary Suspensions till April 2022. Chapter 9 is my introduction to these interesting fluids and an apology for not knowing about them when I wrote the rest of the book.

Abbreviations and Definitions

AFM	Atomic Force Microscope
BET	Brunauer–Emmett–Teller surface area measurement
CNT	Carbon Nanotubes
CTAB	The cationic surfactant Cetyl trimethylammonium bromide,
DLVO	Derjaguin & Landau, Verwey & Overbeek theory
DMSO & NMP	Dimethyl sulfoxide, N-methyl pyrrolidone; solvents
Flory-Huggins	The standard theory of polymer/solvent interactions via χ
FTIR	Fourier Transfer Infrared Spectroscopy
Hofmeister	Systematic effects from changing from Li to Na to K etc.
HSP	Hansen Solubility Parameters
HSPiP	Hansen Solubility Parameters in Practice (Software package)
IGC	Inverse Gas Chromatography
kT	The Boltzmann constant times (absolute) temperature
Lattice Theory	The idea that components in a solution sit on lattice sites
NMR	Nuclear Magnetic Resonance
PMF	Potential of Mean Force
RST	Relative Settling Time of particles in a centrifuge
SCF Theory	Self-Consistent Field theory, also known as SF theory
SEM	Scanning Electron Microscopy
SF Theory	Scheutjens-Fleer theory also known as SCF theory
SLS	The anionic surfactant sodium lauryl sulfonate
TEM	Transmission Electron Microscopy
TOF-SIMS	Time of Flight Surface Ionization Mass Spectrometry
XPS	X-Ray Photo-electron Spectroscopy
χ (chi) Param.	A term expressing mutual (0) or borderline (0.5) compatibility
ζ (zeta) potential	The effective charge at the surface of a particle

1 A few basics

As a reader of this book, you presumably want to get some particles, often nanoparticles, into some formulation to achieve some goals. Sometimes the particles themselves are the star: pigments in paints and ink, fancy nano-electronic particles. Sometimes they are a supporting role such as apparently humble carbon black adding necessary conductivity to fancy battery formulations. Sometimes they are just background stuff providing extra functionality such as ceric oxide nanoparticles in UV hard coatings or the right nanoclay in the right polymer to give added stiffness or tensile strength. Sometimes the aim is to 100% destroy everything in the formulation except the particles - that's what ceramics demand, but on the way they need the ceramic paste to be easy to work with.

Whatever your need, it is likely that you will be fighting to get enough particle into your formulation with the right solvents, dispersants, stabilizers, while keeping the formulation usable when you mix, pump, coat, print, paint, stir it or store it.

The reason for the fight is that if you can get 5% particles in, everyone will ask you to get 10% in. If you heroically manage 10% they will ask for 15%. You will always be at a limit. This book is my attempt to provide you with the relatively few key ideas that will help you to push the limit - and to understand the tradeoffs that are inevitable. This book isn't about nice idealised "formulations" with one particle and one solvent. It's about the messy world of compromises where the perfect solvent for the particles might be awful for the key polymeric co-component, or vice versa, or even where the perfect scientific solvent is unusable because of flammability or toxicity concerns.

Although the reader might want me to say "therefore you need this specific additive or that specific solvent", that's not how it works. My concern is with the relatively few principles that you can then put into intelligent practice.

Many readers will be able to skip this short chapter. I just want to make sure that all readers are familiar with a few basic concepts that permeate the book.

1.1 ϕ

The common language for talking about particles is in terms of their *volume* fraction, indicated as ϕ (phi). Obviously multiply by 100 to get volume percent. There is a good reason for the use of volume fraction. The *mass* fraction is easy to specify in formulations but can be wildly misleading. I once made a formulation by adding 2g of particles to 1g of liquid. It felt very strange to do so, but with a density of 9 g/cc this had a ϕ of less than 0.2

1.2 Size Distribution

Size is important for many reasons, so we had better be clear what we mean by size and why it is important. Our first problem is that many formulae focus on radii, yet particle size distributions tend to be specified in terms of diameters. I cannot choose one or the other, so the reader has to be alert; after all, a factor of 2 can make a big difference to an outcome.

The bigger problem is that we can easily imagine we are using the same definition of size yet be completely wrong. It is possible to say "This particle distribution is 95% below 300nm" and, for the same particles, "95% of the particles are greater than 1.5 μm ". Let's see why, using our first app. Just click the link and feel free to slide sliders and see what happens:



App 1-1 <https://www.stevenabbott.co.uk/practical-rheology/distribution.php>

We have a rather odd distribution with lots of particles with 100nm radius and rather few with a 925nm radius. If we count the number of particles we find (green line) that ~95% have a radius less than 150nm, so a diameter less than 300nm. If, however, we find what volume or mass of particles we have, then (orange line) ~95% of the particles are the relatively small number of large particles with a diameter of 1850nm.

Although we are unlikely to get such an extreme distribution, the point is still valid that we have to be very careful what we talk about. The number distribution, N, tells us very little about where most of our particles are, because volume and mass go as r^3 which means that the mass distribution, M, is dominated by the larger particles. If we want to add dispersants to our particles then we need to cover their surface area, so the intermediate A value is of interest.

The app provides a set of values that might be reported (as diameters) from a particle size analysis. The D[x,y] values capture respectively the number-

average, area-average and volume-average sizes. D50 is the size where there is a 50:50 split between smaller and larger particles, judged by mass.

As with so much of scientific graphing, you have to be very careful about scales. If you select the Log Normal option then you see the smaller particles massively emphasised, even though they are an almost irrelevant contributor to the particle mass. However, if you want to escape the restriction on nanoparticle regulations, this "irrelevant" portion can be highly relevant because by definition (or at least some definitions) nanoparticles contain 50% or more particles by number distribution.

1.3 Aspect Ratio

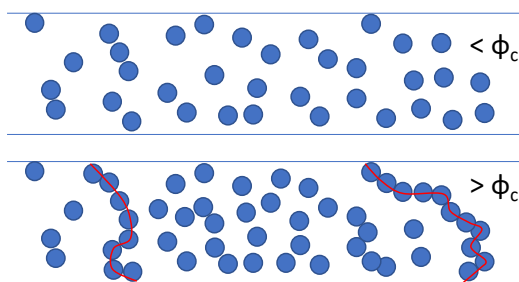
A lot of good things and bad things happen when particles move from spherical to some elongated shape with a high aspect ratio - longest dimension divided by shortest dimension.

Most of the time we have either more-or-less spherical particles or well-known extremes such as clays, carbon nanotubes, silver nanowires or cellulose nanofibres. The effects of an aspect ratio of 2-3 are usually modest and probably show up most in the yield stress effects discussed in the rheology chapter.

The problem with high aspect ratios can be summed up in a single word: percolation.

1.4 Percolation threshold

The key idea is of a percolation threshold ϕ_c which depends on the shape (aspect ratio) of the particle.



To explain percolation and the percolation threshold, imagine that you have some soil or rocks with spherical holes in it, with water beneath and dry ground above. If you have just a few holes, there is no guaranteed path taking the water from hole to hole through to the surface. As you increase the number of holes,

the water still cannot percolate through. At ϕ_c there are sufficient holes for a path for the water to be statistically likely to percolate through; this is the percolation threshold. Now make those holes into spheres. Below ϕ_c there is no continuous particle-to-particle network across the sample, so motion is relatively free. Above ϕ_c there is a network of touching particles so movement of the particles is greatly restricted. For spherical particles this is at $\phi=0.28$. As particles acquire a greater aspect ratio it becomes easier to find a percolation route. One of my apps shows this: <https://www.stevenabbott.co.uk/practical-coatings/percolation.php>. When the aspect ratio is ~ 100 , you need only 0.7% to get percolation.

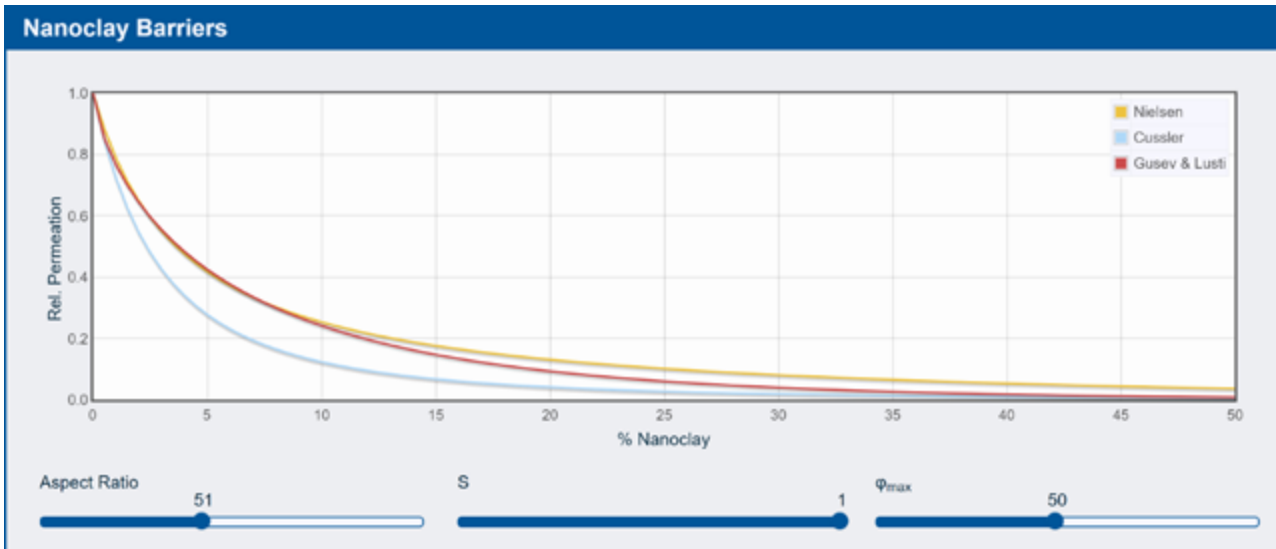
What are the bad aspects of a high aspect ratio and low ϕ_c ? Because the particles are touching across the whole formulation, a stress applied to one portion of a formulation is transmitted through the formulation, which is another way of saying that the viscosity will be high and/or there will be a high yield stress. Which is another way of saying that a dispersion of just a few % of carbon nanotubes or cellulose nanofibres will be difficult to work with. There has been a boom in methods for extracting cellulose nanofibres. The problem is that it is often difficult to use them in any real-world application because any concentration greater than a few % is impossible to handle. Drying off 97% water is not a great way to make large-scale products.

What are the good aspects? For those who wish to make a conducting material, a guaranteed conductive path for the electrons to travel only happens above the percolation threshold. An approximately spherical particle such as carbon black needs 28% to make the material conductive, and considerably more to provide good conduction. If you use the currently fashionable silver nanowires, just a few % is all you need.

1.5 Dispersions, Colloids?

What's a dispersion, what's a colloid? I've found the word "colloid" and the phrase "colloid science" so unhelpful that I seldom choose to use them, though there are times where they seem the right choice. As we shall see, the word "dispersion" comes with implicit assumption (which I will refute) that the word "solubility" cannot be used. Because the word "dispersant" is handy for describing molecules used to stop particles from crashing out, I'm happy to use "dispersion" for particles in solvents, as long as it's understood that solubility theories very much apply to them.

1.6 Tortuosity



App 1-2 <https://www.stevenabbott.co.uk/practical-coatings/nanoclays.php>

We have some nanoclay well-mixed in to some polymer coating and want to know how good a barrier it will be. Given that nothing can get through the clay particles themselves, the barrier properties depend on how difficult it is for things to wind their way past the particles. The path will be tortuous, so we talk of the tortuosity of the coating, the distance travelled around the barrier, L_B divided by the thickness of the coating, L .

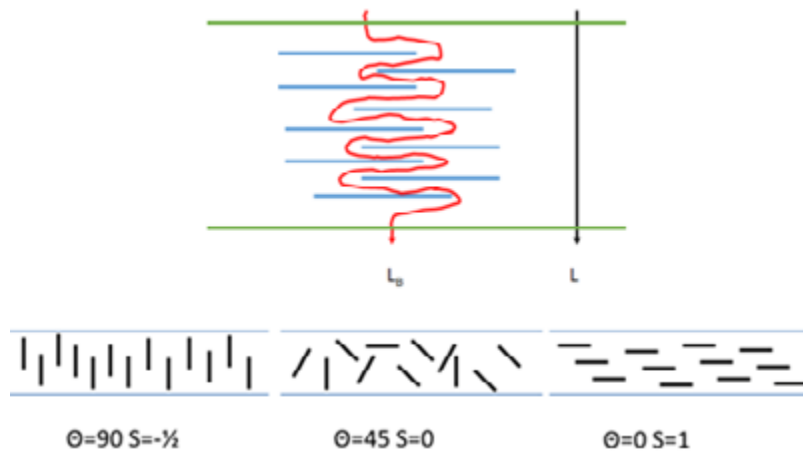


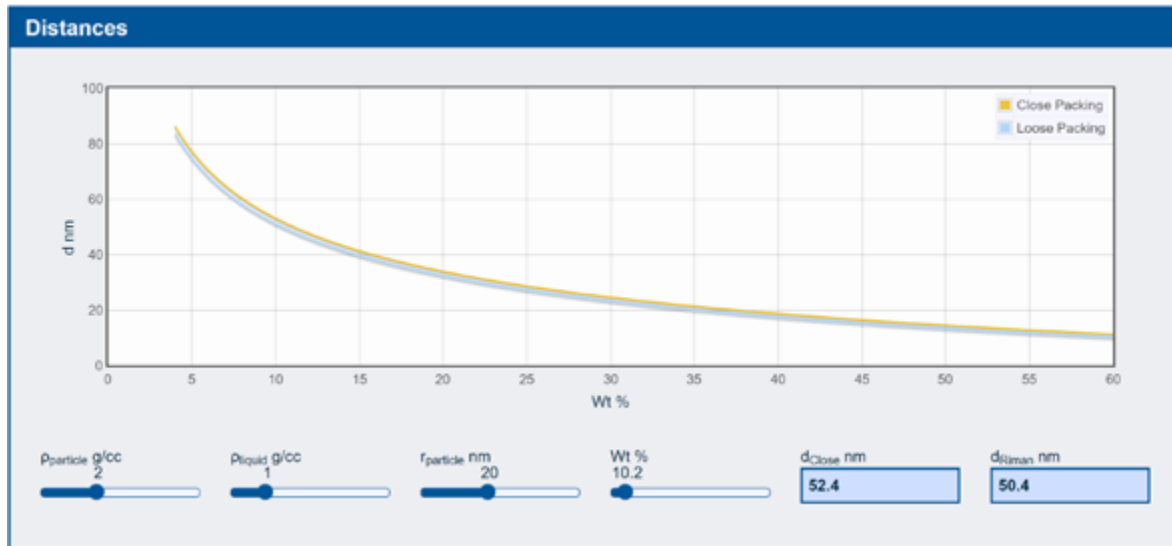
Figure 1-1 The tortuosity is L_B/L and the alignment goes from $S=-\frac{1}{2}$ to $S=1$

The barrier properties depend on three things, which you can explore in the app:

1. The aspect ratio of the clay - the bigger it is the further things have to travel to get round them.
2. The alignment of the particles - if they are all aligned vertically there is no effective barrier, if they are all horizontal the barrier is maximal. This is captured in the S value, which, for no obvious reason, goes from $-\frac{1}{2}$ to 1.

3. The volume fraction of the particles - this is obvious, though getting a beautifully well-dispersed high concentration of clay particles is a challenge. The relationship between volume fraction and relative barrier properties is disputed, so three different calculations have been provided. In any case, you have to be alert to the fact that tortuous barriers appear perfect until the permeant has traversed the path, at which point they are merely good.

1.7 Interparticle Distance



App 1-3 <https://www.stevenabbott.co.uk/practical-coatings/distances.php>

We have some nanoparticles of a given radius at a given Wt% and, from the densities, a known Vol%. It's often interesting to know what the interparticle distance (edge to edge, add $2r$ for centre-to-centre) is, to get some idea of how (un)likely it is that particles will encounter each other. The formula, for those who are interested, is given on the app page and the user can choose which packing approximation to use.

1.8 Measuring particle size

Many decades ago I had access to the amazing new world of laser-based particle size analysis. Whether it's laser diffraction or dynamic light scattering, the idea is that you put your sample into the laser and from readouts of the light intensity (at various times or angles) the particle size distribution can be calculated. In those days, lasers (and computers for the calculations) were not for the faint hearted and this equipment had to be run by experts, unlike the devices casually sitting on lab benches today doing the same job, but rather better.

I quickly became disillusioned with the technique. The manufacturers presented a glossy image of how wonderful and objective it was, whereas in reality the results were only as good as the assumptions you put into the calculations. In my case the results were more-or-less worthless because my particles were

not the idealised monodisperse spheres for which the technique is really rather good.

Jump forward many decades and I visited the booth of the same (though now vastly bigger) manufacturer at a big particle trade show. The supreme confidence in their amazing technology was apparent, yet the basic flaws remain. For most real-world particles the results are only as good as the assumptions you provide, i.e. the results are unlikely to be reliable. For example, the calculations have to assume that particles are spherical and are not interacting significantly with each other. The former assumption is often invalid (though rational corrections can be made if we know the real shape) and the act of diluting the particles sufficiently to remove particle-particle interactions can itself create artefacts.

An alternative methodology is to put the sample into a centrifuge (or ultracentrifuge if they are really small) and via some suitable real-time or stop/start measurement of light or x-rays along the tube, the rates at which the particles move down the tube can be measured and from these the size distribution can be calculated.

Like the laser technique there are plenty of assumptions behind the technique, with, again, the assumption of spherical particles and the assumption that the particles are not self-associating. Concentrations can be higher if the particles are not self-associating because the theory can take into account the general interference of particles with each other as they move down the tube. A methodology able to cope with the complexities of real formulations is discussed in the solubility chapter.

For those who have an SEM handy, the particles are simply placed in the microscope, an image taken and one's image analysis software measures the real shapes and sizes of all the particles. We've all seen such images in academic papers and conference presentations. What we tend not to see are the images where there are too few or too many particles, or where particles are overlapping making it impossible for the image analysis to do a rational calculation - how can it tell whether it's two smaller particles overlapping or one bigger one? If you have an even more expensive TEM machine then it becomes possible to get more 3D information to distinguish between single and overlapping particles. Neither technique is easy and both of them are too slow for routine lab work. Because you are trying to measure the size of billions of particles by taking images of 100s of them, the chances of statistical error and/or sampling bias are large.

Fortunately, you don't have to take my word for it that such techniques have multiple problems. EU regulations are very strict about nanoparticles so it is rather important that we have objective measures of how much (or little)

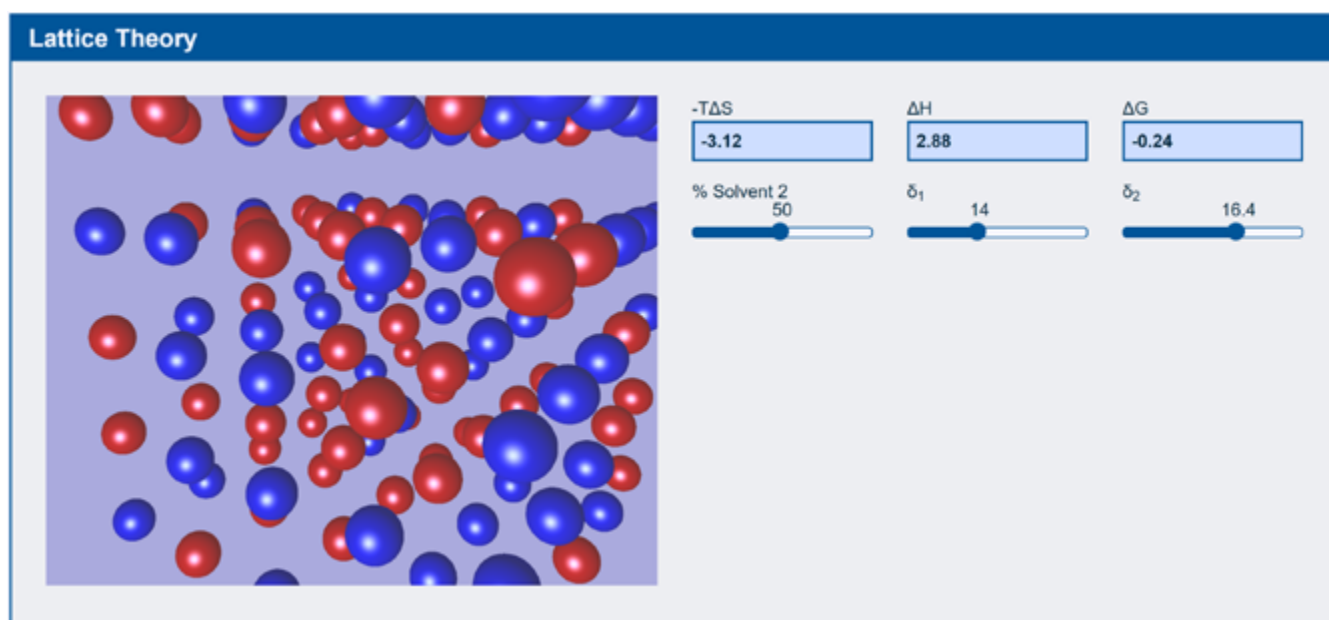
nanoparticle is present in a product. An EU project³ was, therefore, set up to compare and contrast measurement techniques, including via surface area measurements.

The team used highly-motivated, trained groups analyzing carefully-chosen samples that were typical of real-world samples without being especially complex. The outcome was disappointing. For well-behaved, simple, spherical particles most of the techniques could do a decent job - though many struggled with the nano end. For the more usual particles it was hard to be enthusiastic about any of the techniques.

It's unfortunate that particle size measurement of real-world samples is difficult. The important thing is to not spend too much time refining measures that have intrinsic flaws. Use the least inconvenient method that gives you the most practical insight for the least effort. If your whole formulation depends on some subtle particle size issues then it's likely that you have problems much bigger than the act of measuring the sizes. Far better to adjust the formulation so that minor particle size changes don't cause major problems.

1.9 Lattice theory and the Flory-Huggins χ parameter

As described in my Solubility Science book (it's free so you can download it and read for yourself), basic solubility theory places, say, solvent molecules on a regular grid, a lattice, and can calculate the entropy and enthalpy of the system.



App 1-4 <https://www.stevenabbott.co.uk/practical-solubility/lattice.php>

In the screen shot from the app we see a 50:50 mixture of solvent molecules. From a simplistic solubility parameter approximation we find that the entropy

³ https://publications.jrc.ec.europa.eu/repository/bitstream/JRC92531/jrc_eurocolour%20report_final.pdf

is favourable but that the enthalpy is not - and their balance, ΔG is only slightly negative - so the mixture is nearly unstable.

What is remarkable about lattice theory is that you can place many things on the lattice, such as particle surfaces and polymer chains, yet despite the fact that the idealisation is unrealistic, the calculations work out to be valid.

The mismatch between, say, polymer segments and solvent molecules can be captured (according to the theory developed by Flory and Huggins) in a single parameter, χ . When this is 0, the polymer and solvent molecules are thermodynamically indistinguishable. When it is 0.5 they are balanced on the edge of stability. Go above 0.5 and you are in unstable territory.

It's more complex than that, but for our purposes we just need to know that:

- Lattice theory is respectable
- We can calculate a surprising number of interesting effects if we use it
- Low values of χ (they can even be negative if there are donor/acceptor interactions) are good for solubility/compatibility, values above 0.5 are bad.

We will be using χ parameters in classic DLVO, in Hansen Solubility Parameters and the Scheutjens-Fleer (SF), Self-Consistent Field (SCF) theory.

So that's it for the basics.

2 Controlling Stability

Most of the time we want the particles in our formulation to remain uniformly distributed in space and time. So we don't want them to sink to the bottom or clump together. We start with the easier issue.

2.1 Fighting sedimentation

In general, large particles don't clump much but love to sink. Small particles don't sink much but love to clump together (and then they might sink). Let us look at the simpler problem first, that of sinking. A much more sophisticated version of this app taking into account Brownian motion and therefore more suited to nanoparticles is discussed later. It's worth noting here that the common distinction that particles are different from solutes "because particles sink" is fundamentally wrong: at the nanoscale everything sinks!

The Stokes app lets us explore the issues:



App 2-1 <https://www.stevenabbott.co.uk/practical-solubility/stokes.php>

We have a particle of radius $\sim 1\mu\text{m}$, with a density ρ_p of 2 g/cc in water with density ρ_l of 1 g/cc. The viscosity, η is 5 cP, which tells us that we have something else in the formulation to raise the viscosity, such as a bit of polymer. We have a volume fraction, ϕ , of 5% particles. The options connected to settling in a centrifuge are set to zero. We will return to them later.

A note on units

I tend to use "natural" units, those that formulators are most comfortable with. For Stokes settling it would seem odd to specify the density of water as 1000 kg/m^3 or for its viscosity to be in Pa.s. One of the tricks of writing apps is to take the inputs and immediately convert them to correct SI units. If you look at my Javascript (feel free!) you will see those conversions clearly noted, mostly for my own benefit as unit conversion errors are so common.

The app tells us that if we had a small tube with 25mm of sample, it would settle in $\sim 20\text{hrs}$. A moment's thought tells us that this is a bogus number, an issue we can address once we've sorted out the basic issues.

Most readers will know (of) the Stokes settling formula:

$$\text{Equ. 2-1} \quad v = \frac{2(\rho_p - \rho_l)gr^2(1-\varphi)^{4.65}}{9\eta}$$

This gives us the velocity, v , at which the particles fall. If we want to slow down settling and cannot change the densities of the particle or liquid then we either have to increase the viscosity or reduce the radius. Because of the r^2 dependency, relatively small changes are amplified, for better or (usually) for worse. If two particles come together, their combined drag increases by ~20% and their mass is doubled, so the velocity is ~1.6 greater. Because smaller particles tend to clump faster (for multiple reasons), trying to solve settling by going smaller creates its own dangers. That is one of the core issues that we have to fight.

One extra factor can help, though as with most things to do with particles, it contains a danger. If we increased the volume fraction, φ , the particles themselves interfere with each other and reduce the velocity. The $(1-\varphi)^{4.65}$ term that has been added to the standard Stokes equation is the Richardson & Zaki correction. The reason the 20hrs is a bogus number in the app is that as the particles settle, their effective φ increases so the rate of settling decreases. The reason the φ effect is a danger is that, as makes intuitive sense, the higher the concentration, the higher the rate at which any clumping of particles will take place.

The app includes one more calculation. When particles are small enough, Brownian motion (via the thermal energy kT , where k is the Boltzmann constant) is enough to overcome the gravitational sedimentation. The rule used in the app is that settling is stopped if $r^4\Delta\rho.g < kT$. If you slide the radius slider in the example above then it suggests that particles $<700\text{nm}$ will not settle - if you can stop them from self-associating. When we come to the powerful gravitational settling app we will see that this whole "it won't settle because of Brownian motion" argument is bogus. I was a bit shocked when a distinguished professor got very agitated after I casually mentioned it, as I had been entirely unaware that it is bogus. The calculation remains in the app because in real-world tubes that also experience vibration, convection currents and other disturbances, many formulations of smaller particles are stable enough for long enough.

2.1.1 Centrifuges

The Stokes formula includes g , the effective gravity. Normally this is just the gravitational constant, 9.81 m/s^2 . In a centrifuge it can be a million times larger. We will not, in general, put our formulations into centrifuges unless we want the particles to separate out. A specific use for a centrifuge is described later. The interesting aspect of the centrifugal effect is discussed in the solubility section below.

2.2 Self-association

A lot of what we do as formulators involved fighting the tendency for particles to self-associate. Terms commonly used for when self-association happens include:

- agglomeration
- coagulation
- flocculation
- coalescence (often restricted to "fluid" particles that become a single larger particle)

Despite my best efforts, I can find no systematic definitions of these terms and they seem to be used indiscriminately throughout the literature. I will tend to use flocculation simply because we also have the common phenomena (described later) that everyone calls bridging flocculation and depletion flocculation, rather than, say, bridging agglomeration or depletion coagulation.

2.3 Fighting self-association

All particles have a tendency to self-associate via the very short range van der Waals attractive forces. There are five ways to fight that attraction in the hope to keep the particles separated. The first two are a standard part of DLVO theory and the third is now tacked on as if Derjaguin, Landau, Verwey & Overbeek had thought of it:

1. Carefully match the refractive index of the particle and the solvent - the net attraction is then 0
2. Build a self-repelling charged shell around the particle
3. Attach a polymer shell to the particle that provides a steric barrier to attraction
4. Solubilize the particle
5. Add enough energy via some sort of mixer or disperser to rip the particles apart.

The first method is sometimes mentioned and sometimes found in academic papers where they can use highly-tuned solvent blends so they can focus on other aspects of the particles. The next two methods are the ones discussed at great length in all colloid books and are conveniently packed up as part of DLVO theory. The fourth method is usually regarded as being heretical and is never discussed, yet is actually very powerful. The fifth is routine and will be discussed in a different chapter.

Rather than waste a lot of your time with pages of solemn theory about DLVO (you can find more than enough in any colloid book) we will focus on the meaning of the key formulae, touching lightly on many of the parameters, then use the DLVO app to teach us the (very) few lessons that matter.

DLVO was amazing when it emerged, independently, from Derjaguin and Landau in Russia, and from Verwey and Overbeek in the Netherlands. It is generally presented as being awesomely complex and profound, with formulae that most of us cannot routinely implement. The aim here is to make it a rather obvious bit of background that any formulator would use as a sanity check before doing serious formulation work. Despite its strengths, I will be critical of DLVO theory because it turns out that much of it is limited or wrong.

As I am allowed to copy from my own books, what follows is more or less the same as in my Solubility Science book. However, I will first insert a diversion about a once-popular distinction.

2.3.1 Lyophilic and lyophobic

I regret to say that I have always been a bit hard on D&L and V&O, basically accusing them of over-selling their theory. Like most people I had never read their original texts, so I was unaware that V&O (I don't think I'll ever read D&L) were smart critics of their own ideas, very much aware of the limitations that others then glossed over. But the biggest clue to the limitation comes from the title of their 1948 book: *Theory of Stability of Lyophobic Colloids*.

At the time, there was a well-known distinction between lyophilic, literally solvent-loving and lyophobic, solvent-hating colloids. The lyophilic colloids were things like gelatin and were seen as being swollen polymers with their own features of solubility and stability, of no interest to V&O in terms of the problem that was really bothering them. The lyophobic colloids were those lumps of solid like gold or silica with no possibility of the solvent going into them and causing effects such as swelling. Explaining those was, at the time, a real problem.

If you start with a clear distinction and keep going with that distinction, life is easy. Our problem in the 21st century is that the distinction between lyophilic and lyophobic makes no sense. Many "classic DLVO" systems are particles with charged polymers around them. So they are sort of lyophobic cores with lyophilic externals. Which theory should be applied to them? Later I will discuss a paper where they dissolved (there's no better word for it) 5nm silicon (yes, silicon, not silica) in DMSO and NMP solvents.

Equally important, within half a page of describing the distinctive salt effect that can cause lyophobic colloids to flocculate, V&O point out that, actually, some lyophilic colloids can show similar effects (probably for similar reasons), so they were well aware that even then their distinction was not a sharp one.

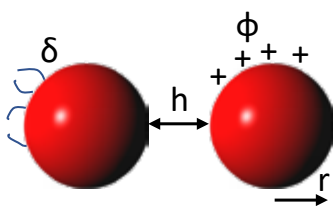
At the time, just being able to explain the pure, solid charged particles was an astonishing achievement, so worrying about blurred distinctions was not the focus of D&L and V&O. So when I am critical about DLVO theory, I am not

attacking D&L and V&O. It's those who packaged up DLVO and, in the words of Barry Ninham, discussed later, made a religion out of it.

2.4 Core DLVO

The approach adopted here is to present the formulae for each of the three effects, then to identify the key inputs required for a calculation, then to provide users with the app that handles the calculations. This is pragmatic DLVO theory - the maximum benefit for the minimum work. Those who want a deeper understanding of any of the components can readily find explanations on the Internet.

2.4.1 DLVO



The key parameters in each of the three equations are the radius of the particle, r , and the distance between particles, h . The diagram also shows δ , the length of the chains sticking out from a particle and ϕ , the charge on the particle.

The first term in DLVO is the van der Waals term though I have called it V_H because the strengths of the van der Waals attractions between two particles are calculated using Hamaker's method and the Hamaker constant A_{12} . The calculations are all done in units of kT , Boltzmann constant times (absolute) T :

Equ. 2-2

$$V_H = -\frac{A_{12}r}{12hkT}$$

Larger particles have a stronger self-attraction and as h becomes very small, V_H becomes enormous - i.e. if the particles ever get close-enough to touch, nothing can separate them. This reiterates the point mentioned in the HSP chapter that van der Waals forces seem very dull, yet the fact that everything is attracted to everything else via van der Waals means that they are the single most important force in solubility/stability. What are the values of the Hamaker constants for a given system? They can be measured with great difficulty but usually it is not worth the bother; a value of $1E^{-20}J$ (or 10 zJ for zeptoJoules) is good enough for most of us. A table of Hamaker values is given in the app, though they should be treated with great scepticism.

To reduce self-association, all you have to do is reduce A_{12} , so how is that done? To find A_{12} you need a Lifshitz theory calculation that can sum integrals of dielectric constants over an infinite frequency range. Fortunately, that's not our question, which is how we can minimize A_{12} and the answer is straightforward, match the (complex) refractive indices of the particle and the solvent. In the literature you can find, for example, careful experiments of polystyrene particles

suspended in decalin/tetralin solvent blends that exactly match the refractive index of the polymer and therefore allow experiments where the particles have no self-attraction. For the more complex formulations of concern to readers, the most straightforward approach seems to be to match the Hansen Solubility Parameters as discussed later.

The second term is the charge-dependent term that I have called V_D because it is based on the Debye-Hückel calculation of charges within ionic solutions. The formula is complex and contains the confusingly-designated k^{-1} parameter which is not to be confused with the k in kT . It also includes Avogadro's number N_A :

Equ. 2-3

$$V_D = -\frac{2\pi e_0 \epsilon r \phi^2 \ln(1 + e^{\frac{h}{k^{-1}}})}{kT}$$

Equ. 2-4

$$k^{-1} = \sqrt{\frac{e_0 \epsilon kT}{2N_A e^2 I}}$$

Fortunately the app deals with all this complexity. Just for reference, here is what the parameters mean. The k^{-1} is the Debye screening length (how quickly the effect of the charge disappears over distance) and depends on e_0 which is the permittivity of free space, e is the charge on an electron, ϵ is the dielectric constant of the medium and I is the ionic strength which (as per the app) depends on the concentration of salts in the solution and their charges Z_1 and Z_2 which would each be 1 and 1 for NaCl and 2 and 1 for MgCl₂ and so on.

The dependence on h is relatively gentle and long-range so if there is a significant charge ϕ , a low ionic strength I and a high dielectric constant ϵ there is a strong repulsion keeping things apart. So large salt concentrations are bad, any solvent other than water ($\epsilon=80$) is bad and, of course, you need a lot of charge on the particle. What, then, is ϕ ? Embarrassingly, the answer is that no one really knows, but if we use the measured zeta potential, ζ , that seems to be good enough. When we explore ζ you will see why there is so much vagueness.

The simplest form of charge stabilization is when the particle itself is charged. For neutral particles it is common to add a molecule where one neutral portion sticks nicely to the neutral surface and the other charged portion sticks out into the solution. The classic examples are surfactants with hydrocarbon tails in the 10-18 range and, typically, sulfonate salt head groups. As with so much of DLVO the problem isn't that these don't work - they work splendidly for academic studies of DLVO. The problem, as we shall see, is that most of us are not interested in pure charge-stabilized particles - we rather like to add other functional ingredients such as polymers, and these can cause chaos to the simple story.

Finally we have the steric term, V_s which was not part of original DLVO. Here we use the value of δ , the length of the molecule (we'll call it a polymer for simplicity though it could be a long surfactant chain) sticking out from the particle plus Γ which is the coverage of the particle by the steric barrier material, and a density ρ and $MVol$ that are of no great significance.

A more refined theory is available

In the chapter on SF theory we will find a more satisfactory way of dealing with what causes steric stabilization and how it can be made more robust [add lots of tails!] and what can mess it up [have lots of loops]. Our concerns aren't over a single δ , and the χ parameter effects are more interesting than those discussed here, but it's still good to be familiar with the simple approach.

The fact that the χ parameter (as discussed in the HSP chapter) appears is evidence that "solubility" and "dispersions" are, indeed, related:

Equ. 2-5

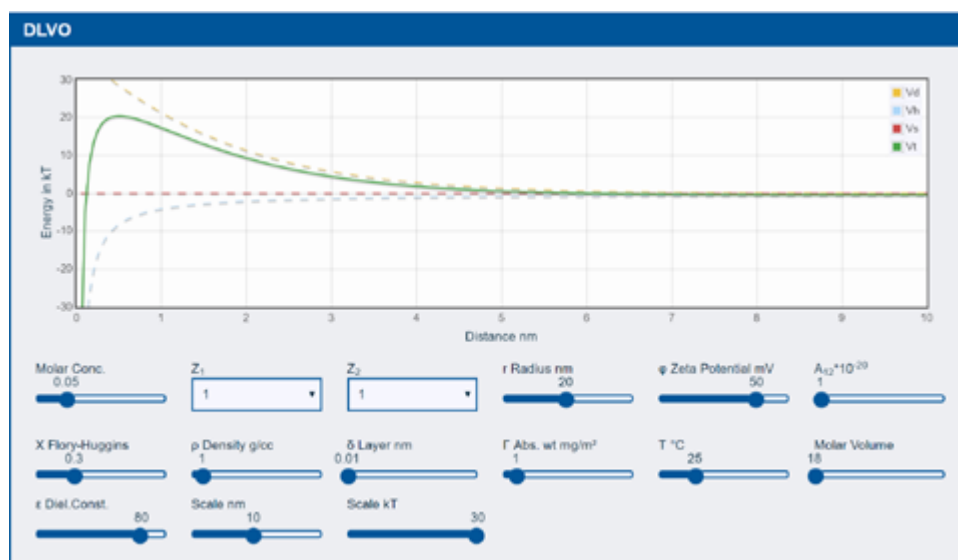
$$V_s = \frac{30N_A\pi r\Gamma^2}{\rho^2 MVol} (0.5 - \chi) \left(1 - \frac{h}{2\delta}\right)^2$$

The equation is only valid when $h < 2\delta$ for the obvious reason that there is no steric force until the steric chains are touching. V_s suddenly becomes very large and positive (repulsive) which is why steric stabilisation is so popular. However, there is a catch - that term with χ inside. As we will discuss in the HSP chapter, χ captures how (dis)similar solvent and polymer are. If they are identical (in terms of interactions with themselves and each other) then $\chi=0$ so the $(0.5-\chi)$ term has a maximum value of 0.5. As solvent and polymer become more dissimilar, χ increases. If you try doing this in the app, not a lot happens because the V_s term is so large that decreasing it to, say, 0.25 when $\chi=0.25$ is not significant. Indeed, a formulator can unknowingly end up with a solvent that takes χ close to 0.5 and still see no problem. The problem arises when the solvent changes to slightly above 0.5. Now the $(0.5-\chi)$ term is negative and V_s becomes negative (attractive), causing the whole dispersion to coagulate. Many of us have suffered from this large effect from a small changes without knowing why. Hopefully, knowing the root cause will help formulators to be in a safe, low- χ zone.

The effect of χ is put to good use in many elegant studies that strive to find out what's going on when particulate systems gel, something discussed in detail in the final chapter. Briefly, you coat, say, silica particles with stearyl alcohol chains, giving the required steric repulsion in an OK solvent. It happens that decalin is borderline OK. At, say, 40°C these particles are stable indefinitely, but when the temperature falls below 34°C, the theta temperature where $\chi=0.5$, then gelation can take place - very slowly at 33, faster at 25, faster still at 20 and so forth - the temperature effect becomes rather dramatic, as we find in the last chapter. Heat

the gel back over 34°C and the gel disappears - a reminder that not everything about particles is a catastrophic choice between stable and crashed out.

Now we can bring all the terms together, giving a total potential, $V_T = V_H + V_D + V_S$. The aim is to keep V_T as positive (repulsive) as possible for as short a distance as possible so that particles that happen to be approaching each other will bounce off rather than stick together. We can now discuss why V_T is expressed in terms of kT . The average particle will be moving with that Boltzmann energy of kT with fewer (exponentially) at higher kT values. The rule of thumb is for the barrier to be $20kT$ which means a probability of e^{-20} of overcoming the barrier, i.e. 1 in a trillion. If the balance of forces gives a minimum of a few kT at some significant distance, that indicates that the particles might pseudo-clump, which may be worrying but which can usually be overcome by re-mixing/shaking.



App 2-2 <https://www.stevenabbott.co.uk/practical-solubility/dlvo.php>

The first example shows a typical charge stabilised system with a zeta potential of 50mV in water and a suitably low molar concentration of ions. The particles reach a barrier of 20kT at around 0.5nm, perhaps dangerously close if there are imperfections somewhere in the system.



Figure 2-1 Pure steric stabilisation for the same system

The second example shows an uncharged particle relying on a 1.5nm protective shell, leading to a stable distance of 3nm.

The key advice for handling the app is to build up confidence by leaving all parameters alone except for r , ϕ and δ . Playing with these gives a good feel for the key effects, the shape of the V_T curve and the relative chances of clumping (barrier $< 20kT$), pseudo-clumping (a minimum of a few kT) or general stability at a safe distance (high ϕ) or at the 2δ distance where strong steric stabilisation kicks in. Those who rely on charges should then reduce the dielectric constant from water's value of 80 to that of common alcohols around 30 to see that water is probably the only environment where charge-based stability is viable.

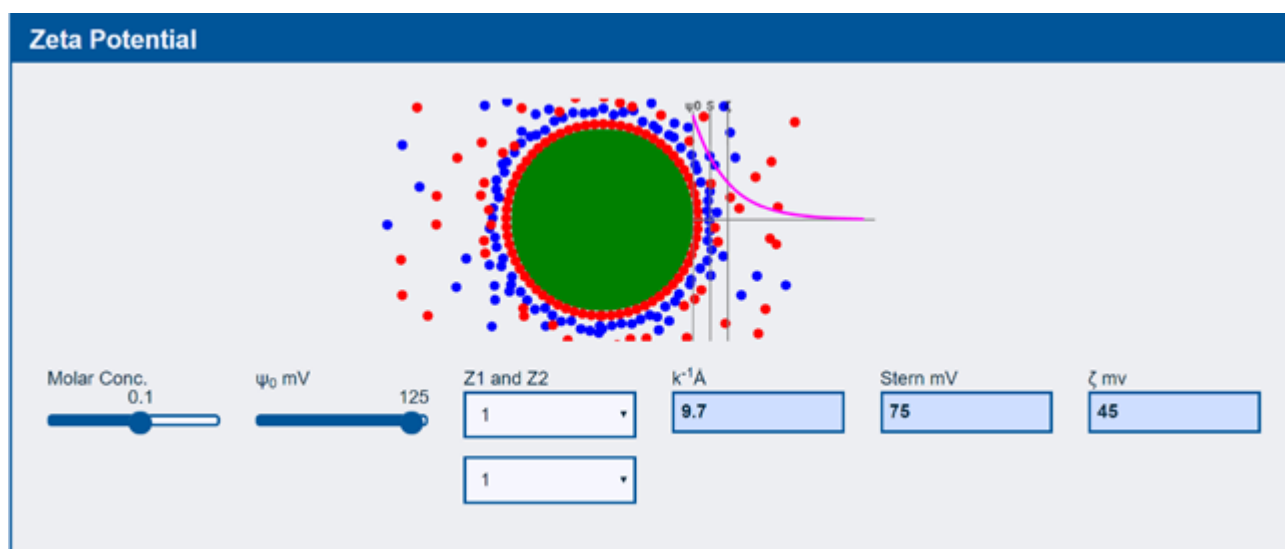
For those choosing to rely on steric stabilisation, the question is why it works at all. Imagine a sphere of radius r which self-associated via van der Waals attraction. Now put a really good polymer layer around it so that its radius is $r + \delta$. Why doesn't this clump because of van der Waals attractions between the polymers? That is a very good question and it turns out that people are still debating the issue. My personal favourite is that it is an "excluded volume" effect in terms of Kirkwood-Buff theory described in Practical Solubility. It turns out that the other common explanations based on "entropic" effects and, even, "osmotic pressure" effects are basically the same thing, seen from different viewpoints. The classic osmotic pressure explanation for steric stabilisation states that the local polymer concentration is larger if two particles get close and solvent needs to flood in to equalise concentrations. Unfortunately, none of these equivalent theories explain why the χ parameter is critical. The simplest explanation I can find is that the excluded volume/osmotic effects are fine if the polymer-solvent interactions aren't too different from polymer-polymer interactions, but as soon as the polymer is no longer happy in the solvent (which happens by definition at the θ point when $\chi = 0.5$) then we are back to the $r + \delta$ ball which just sticks together via van der Waals.

In any case, the SF theory discussed later gets around these contortions and is able to deal with the complexities of the mutual interactions of solvents, particles and the multiple polymer components (e.g. A-B di-blocks) used in practice. If you are a Nobel Prize winner like de Gennes you can just see via symmetry arguments that tails repel and coils of polymer attract.

Staying, for the moment, in the simpler DLVO world, we do not have to care too much about the detailed explanations as they are concerned with idealised systems and we have messy real-world systems. We *do* need to care that the 2δ protection distance and the sudden clumping when $\chi > 0.5$ are well-validated effects in model systems and are good-enough principles for formulation work.

2.5 Zeta potential ζ

The charge term φ in DLVO is rather complicated. So the convention is to use the measured ζ potential. A typical lab particle-sizer which provides r is also able to apply an electrical field across the measurement cell and from the way that the particles move in that field, ζ can be calculated. Given that the measurement is simple and pragmatic, it is good enough for most of us. Unfortunately, although one might think that "the charge on a particle" is straightforward and self-evident, experience shows that the value of the ζ potential even of a relatively simple particle such as silica can vary wildly not only with obvious things like pH but with subtle things like low levels of ionic impurities. It also depends on the speed of the technique by which it is measured - so there is a kinetic as well as thermodynamic components to it. It is, therefore, important to see why "the charge on a particle" is not a simple concept. The app shows this nicely.



App 2-3 <https://www.stevenabbott.co.uk/practical-solubility/zeta.php>

The input value ψ_0 represents the "real" charge on the particle, though this is a meaningless concept. This real charge is the ring of red dots around the particle and it attracts, naturally enough, a ring of negative charges in blue. And

we can immediately see why "the charge" is such a tricky concept. If that ring of negative charges were, in general, perfect then the charge of all particles would be exactly zero. That is clearly not what happens, but undeniably the measured voltage will be significantly less than the "real" charge. And, of course, that imperfect ring of negative charges will attract an even more imperfect ring of positive charges with an even more uncertain impact on the final charge. The usual hand-waving way out of this is to say that the first pair of charges forms the Stern Layer and that at a somewhat larger radius you have the swarm of charges that, on average, move along with the particle, as opposed to those ions that are sitting around in the bulk solution that do not move if the particle moves. The point where the moving cloud ends and the bulk solution begins is where the ζ potential is "measured".

Now you can see how relatively small effects can have amplified effects via their impact on these vague clouds of particles. Some large multi-valent ions that happen to be near the particle, or some contaminants at the true surface of the particle can each be imagined to have complex effects on the ionic cloud and, therefore, on the measured potential.

As mentioned earlier, a particle stabilized with a surfactant head sticking out might behave wonderfully for an academic paper, yet pose many problems when other ingredients start to interact with the system.

This rather jaundiced way of looking at ζ potential is deliberate because it leads to the key rule about ζ potentials: "Never measure a *single* ζ potential, measure the potential in a variety of conditions relevant to your formulation". In the past this would have been unwelcome news because measurement was so hard. With modern equipment there is no excuse for not following this rule. Suppose you deliberately change the pH across the range of any likely formulations. This might show the well-known effect where a large negative potential shifts through zero (the iso-potential point) to a large positive potential as the pH is lowered.

More interestingly, systematic measurements from varying the (low) concentration of an additive might show that the additive which was there for purposes not connected to the particle dispersion, might produce an unexpected and unwelcome large change of ζ value. It is precisely these unexpected consequences of well-intended additions of other components that the ζ rule is intended to reveal. If, under all reasonable formulation variations, your ζ is rock solid and if you have the right dielectric constant and a low level of salt additives, then you are likely to have few surprises in terms of particle stability. If you start to see significant variations then you can probably come up with a reasonable hypothesis of why additive X (e.g. via a minor ionic impurity) is affecting particle Y, and do something rational about it.

An especially common "minor ionic impurity" can be a surfactant. An anionic particle formulation showing robust DLVO stability can be instantly converted

to a flocculated mess by the accidental addition of a small quantity of a cationic surfactant. The first time this happened to me, I didn't even know that my additive included a low level of surfactant, so I had no reason to suspect that (as it turned out) it contained a cationic version.

2.5.1 Extended DLVO

The reader might wonder if the deficiencies of DLVO can be fixed with, say, Extended DLVO. With over 3000 hits in Google Scholar for "extended DLVO" it might seem strange that "it" has not yet been discussed. But there isn't "an" extended DLVO. There are lots of papers where the deficiencies of DLVO are fixed via whichever theoretical infrastructure appeals to the authors. So really they are "ad hoc extended DLVO" theories. My rule of thumb is that if there are lots of alternative theories then none of them is any good. If any one of them was really good, the others would rapidly disappear (and I might well have applied it).

Any time I look at extended DLVO papers I am struck by the fact that to a theory where there is already much that is uncertain, the problem is "solved" by adding a couple of complicated extra expressions and some values derived that can help fit the specific experimental data in that paper. What never seems to happen is a development of a coherent set of parameters that can be used widely, even by the academic community.

Good science builds on itself so that later users have a richer set of data on which to build, and for an even later generation to use.

As I've pointed out elsewhere, Hansen Solubility Parameters, a key part of this book, were not an especially novel idea. Others were working on similar schemes at the time. What made HSP so successful is that Hansen did the hard slog of producing a canonical set of values that allowed himself and others to build upon. I've read papers from around that time of schemes that seemed smarter. But they could never scale to a coherent set of usable parameters so never made it into general use. Hansen's original ~100 values quickly became a few 100 and reached a stable ~1000. From these it has been possible to build functional group schemes which, with various degrees of success, can predict other values. At the same time, the original 100 solvents made it possible to *measure* HSP values (discussed in the next chapter) for polymers and particles, so the technique could spread into many different research areas.

Extended DLVO doesn't exist as a coherent whole because either it is intrinsically incoherent or there was no one with the qualities or resources needed to create the required coherence.

2.5.2 DLVO as a religion

The great Australian scientist Prof Barry Ninham was a personal friend and colleague of both Derjaguin and Overbeek so knew DLVO theory intimately. Yet this hasn't stopped him being scornful of DLVO's dominance of the colloid field⁴. For the time, the theory wasn't so bad. But it makes the crucial assumption that the vdW term and the charge term can be logically separated which, according to Ninham is not true ("...the additivity of electrostatic and dispersion forces assumed in DLVO theory is inadmissible"). As another quote shows, even D and O didn't believe their own assumptions⁵:

"The theory has severe limitations, acknowledged by both Deryaguin and Overbeek: because a liquid between interacting bodies is assumed to have bulk properties up to a molecular distance from an idealised surface. Further, apart from the contamination issue, most surfaces are not molecularly smooth or chemically homogeneous and in addition solvent molecules may interact directly with the surface, for example via hydrophilic or hydrophobic effects in the case of water. Specific ion effects are also ignored. Lastly, the theory has further, amplified problems at high electrolyte concentrations, where only very short range surface forces are expected and these other factors can dominate."

Also, as shown by Ninham's one-time-colleague Israelachvilli (and many others subsequently), the mean field assumption behind DLVO breaks down when particles are separated by a few molecules. The steric portion was never part of the original DLVO and, as we shall see, has many problems of its own, all of which can be addressed much better via Scheutjens-Fleer theory.

The "bulk properties" and "mean field" assumptions sound academic niceties. We shall soon see, with some nanoparticles that by all definitions are soluble, that this isn't just a nicety.

Finally, Ninham is a passionate believer that many studies done in ordinary water are misleading because dissolved air in the form of nanobubbles dominates many of the effects put down to DLVO. The reader will have to forgive me for not being able to provide a coherent set of formulation rules when nanobubbles are involved - I simply have not been able to find enough good literature.

4 Barry W. Ninham, B. V. Derjaguin and J. Theo. G. Overbeek. *Their Times, and Ours*, Substantia 3(2): 65-72. This is an Open Access article and a fascinating read!

5 Barry W. Ninham, Richard M. Pashley, Pierandrea Lo Nostro, *Surface forces: Changing concepts and complexity with dissolved gas, bubbles, salt and heat*, Current Opinion in Colloid & Interface Science, 27, 25-32, 2017

Nanobubbles

For many years it was contentious to claim that nanobubbles, say 200nm diameter, 50nm high, sat on hydrophobic surfaces in water. Basic theory says that bubble pressure is proportional to $1/r$ so such bubbles simply could not exist - the internal pressure would be implausibly high. However, it turns out¹ that subtle issues of pinned contact angles and over-saturated gas concentrations (e.g. taking cold water that is saturated with air and bringing it to room temperature) allow these bubbles to be stable. Such bubbles form easily on hydrophobic particles and, therefore, add a totally different set of attraction/repulsion effects unknown to DLVO.

I first came across nanobubbles in a discussion over the fact that some well-known fragrance molecules that are insoluble in water become adequately soluble if all gas is first removed, in other words, the fragrance molecules separate if nanobubbles are present. This immediately suggests that we could have "chemical free" fragrances. However, it is very costly to remove all the air from water, and difficult to create a perfume spray that doesn't introduce air.

¹ Detlef Lohse and Xuehua Zhang, Surface nanobubbles and nanodroplets, Reviews of Modern Physics, 87, 981-1035, 2015

In Ninham's view, a perfectly good theory got hijacked by the academic community into some sort of religion that has blinded them to the harsh reality that it's not actually a very good theory, and doesn't help much with real-world formulations. An especially good quote, about the purported link between AFM measurements and DLVO captures Ninham's style:

"Indeed if anyone claims agreement with DLVO theory, his measurements are wrong. The foundations of the theory, are deeply flawed even of continuum solvent theory. They include pH, pKas, interfacial tensions, activities, interparticle interactions, zeta potentials, etc."

2.5.3 The secondary minimum

Maybe I'm being too hard on DLVO. There's one principle that it captures very nicely, is vindicated by careful academic experiments, and can sometimes help us sort out issues in real formulations. This is the principle of the secondary minimum, easily captured in the DLVO app, more easily with charge stabilization than with steric, but the principle is the same.

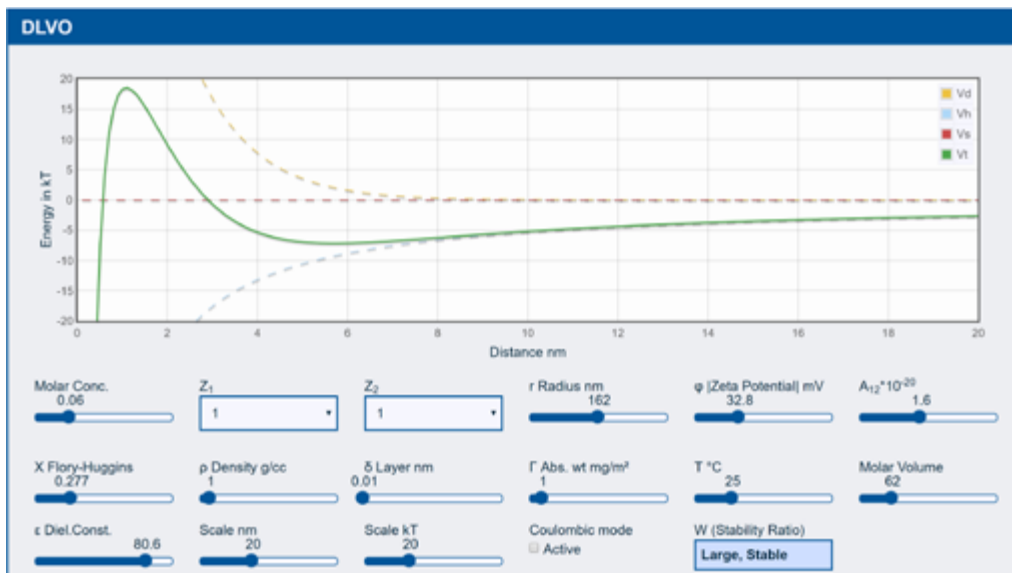


Figure 2-2 The secondary minimum gives gentle flocculation

The van der Waals interaction (the bottom dotted line) is pulling the particles together and by 5 nm they are attracted by ~7 kT. Then the charge stabilization kicks in, stopping the particles from fully sticking together - at least for a while. If thermal energy kicks the particles over the barrier then they are trapped in the deep primary minimum. Instead they are trapped in the shallow secondary minimum. This sort of flocculation is relatively easy to undo with a gentle shake, though the fact that particles are on average held closer together increases the probability of them going over the barrier. As we shall see in the rheology chapter, these minima create high yield stresses and produce strong shear-thinning behaviour.

We can estimate the stability via the Stability Ratio, W . In this example it is large, giving a stable formulation. If you reduce the zeta potential or increase the ionic concentration, W starts to decrease, threatening instability. What is W ? It is the ratio of the clumping speed (which depends on particle size and vdW attraction) if there is no barrier to the speed with a barrier. Typically if W is > 5000 then the system is "stable" under "reasonable" time-scales - whatever that means.

There are plenty of confirmations of the value of the insights from W when studying pure particles under ideal circumstances. As ever, it's harder to apply the idea in complex formulations.

2.5.4 Summarizing DLVO

The lessons are simple and obvious:

1. In aqueous systems, charge stabilization is fine as long as you avoid adding salts and, especially, multi-valent salts.

2. Beware of zeta potentials because rather small formulation changes can suddenly reduce them via some sort of poisoning or neutralisation or interaction with other charged or uncharged polymers.
3. Don't try charge stabilization in anything other than near-aqueous systems
4. For steric stabilization, have good coverage of a thick layer polymer that is locked to your particle as a "brush" - an idea we will explore in detail in the SF chapter. But too much of a good thing is a bad thing, as we will find when we come to highly-filled systems.
5. Beware of bad solvents and solvent blends that can suddenly tip your steric stabilization into full-on self-association or, alternatively, pay attention to the various χ values when you use SF theory..

In terms of the 4th point, I am struck by the number of formulators who assume that they have to throw in some dispersant rather than engineer their particles to have a chemically-attached brush, at some modest level that's both easy to achieve and very effective. I'll return to this point in Locked and Loaded.

The last point about solvents brings us nicely to the fourth method of keeping particles from self-associating.

2.6 Solvation and solubilization

There is general agreement that even a large protein or a large piece of DNA is "soluble". There is equal agreement that particles are, at best, "dispersed" because they will eventually settle out under gravity. So what is the difference between a dispersion and a solution? One of the great books in the colloid field is Mewis and Wagner's *Colloid Suspension Rheology* and they provide their view on the topic:

"Colloids in a suspending medium are a type of mixture. Note that we use the term dispersions, as the term suspension often refers to mixtures where the dispersed phase particles are greater than colloidal in size. Solutions, at the other extreme, refer to molecular mixtures – although polymer and protein solutions are often treated, in many respects, effectively as colloidal dispersions. The reader will encounter terms such as colloidal dispersion, suspension, and solution used interchangeably in the literature."

So a colloid is a type of mixture, and a solution is a molecular mixture, and soluble polymers are effectively dispersions. I have looked hard, but have failed to find any clear definition of the difference between solution and dispersion. We shall meet another definition later, which is provably wrong.

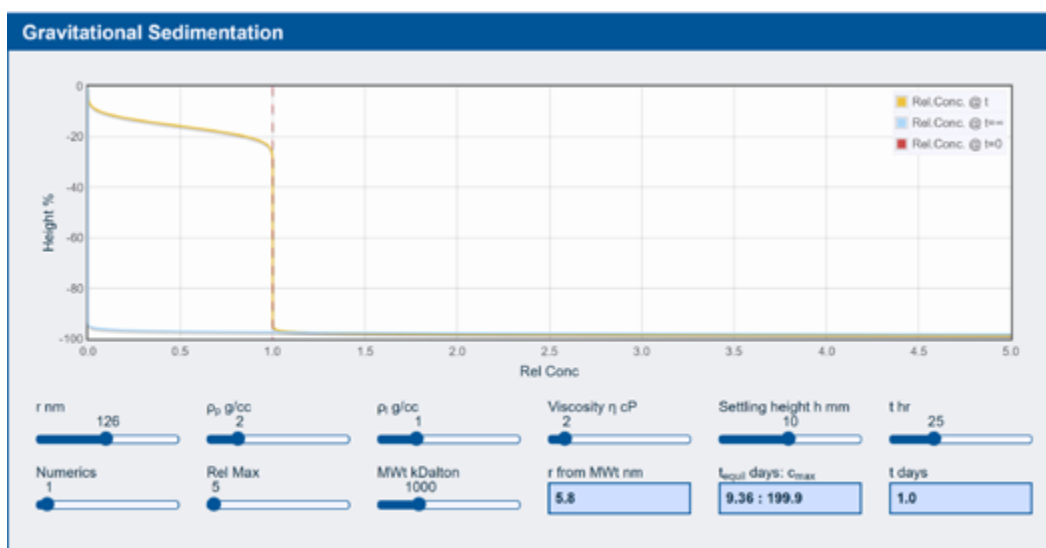
My explanation for this is simple: there is no difference, nor can there be a difference. What follows from this is that we can call upon the resources of solubility theories to help us formulate with particles. This is such an important

conclusion, and the reader will be so reluctant to accept it, that I must work hard to be convincing.

2.6.1 Looking for differences

The most obvious difference is that particles fall out of dispersions and soluble molecules stay dissolved forever. This distinction fails very easily. As pointed out above, and as we shall see next, smaller particles will stay suspended for a long time - as long as they are sufficiently stabilized against agglomeration. In the other direction, all it takes to make a protein or DNA molecule fall to the bottom of a tube is to increase gravity in a centrifuge. It seems unreasonable to define something as "soluble" merely by reference to the gravitational field of a planet.

In any case, settling under gravity is not a proof of anything because *everything* settles under gravity, even molecules. Nothing can resist gravity because it is universal. As we know, everything in a liquid environment also experiences a universal phenomenon, Brownian motion from molecular collisions. This gives rise to the erroneous idea that the only thing that stops smaller molecules or particles from settling is Brownian motion. When I mentioned this "fact" to a German professor I got a very severe lecture about how wrong I was. The fact is that settling via gravity is *thermodynamics*. What Brownian motion does is affect the *kinetics*. A small nanoparticle or polymer of the same size and density are equally resistant to settling because Brownian motion makes the process very slow. The "soluble" polymer behaves exactly the same as the "dispersed" nanoparticle. There's even a formula for it, developed in the 1920s by Mason & Weaver. An app tells us exactly what to expect from any particle or polymer - with exactly no difference between them.



App 2-4 <https://www.stevenabbott.co.uk/practical-solubility/Gravitational-Sedimentation.php>

What we see is the relative concentration, after 1 day, of some 250nm diameter silica particles (density ~ 2 g/cc) in 10mm of a solvent with the density of water and a viscosity of 2 cP. For most of the tube, the concentration is 1, i.e. the

original, which follows the vertical dotted line. In the top 20% of the tube, the concentration falls off rapidly, and the bottom few % of the tube has a high concentration. If you slid the t slider to ~9 days then the silica would have settled out to the high concentration near the bottom just visible as the blue line.

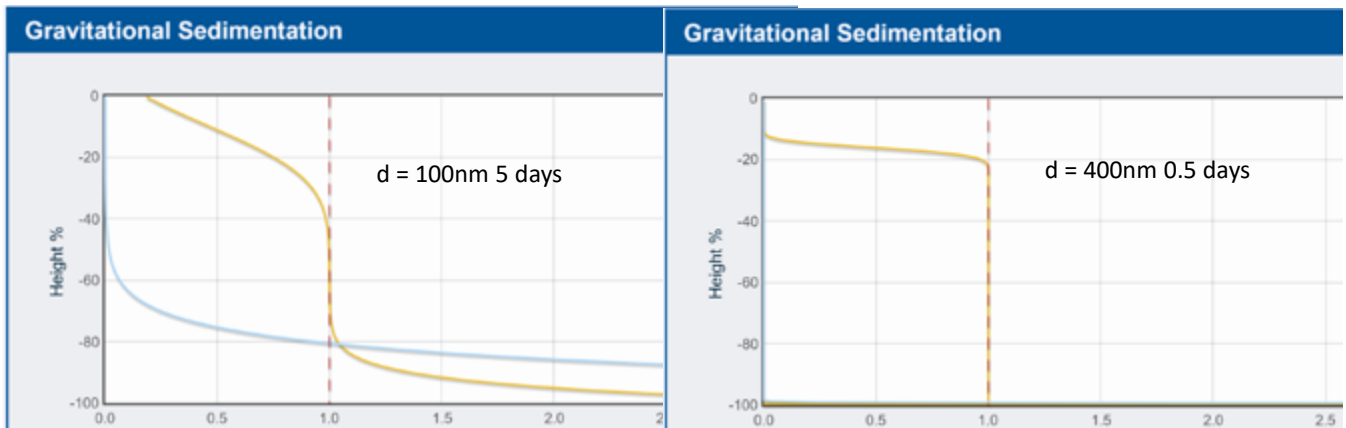


Figure 2-3 The same silica showing different distributions with different particle sizes and times

The app lets us see the effect of size and time on the distribution. Smaller particles take much longer to settle (of course) with a much more gradual concentration profile (not so obvious). The point of all this is that your decision on whether the particles settle or not depends on the time-scale of your observations, the care with which you avoid any convection currents, and the precision of your observations.

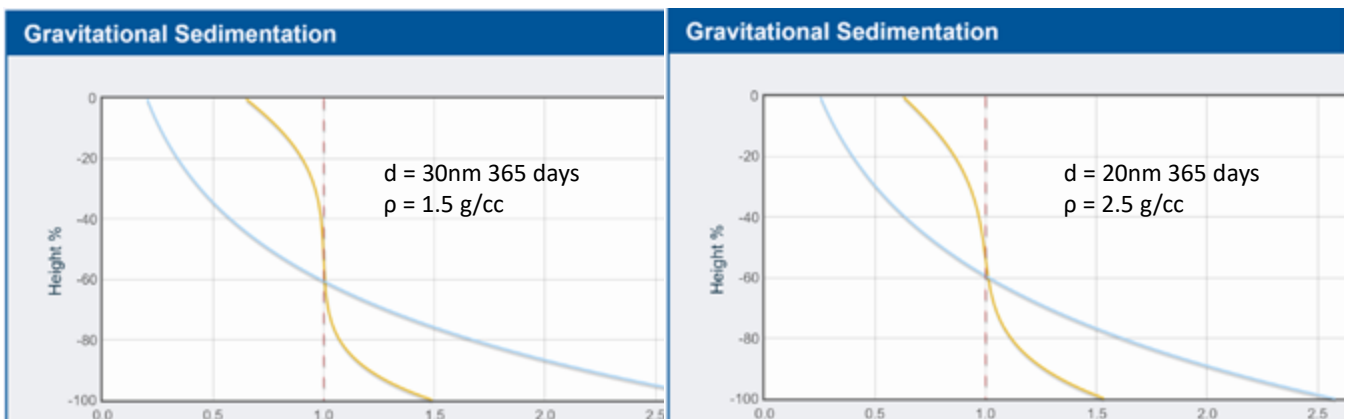


Figure 2-4 We get the same curves after 1yr of settling with a 12000 kD polymer, $\rho = 1.5$, $d = 30\text{nm}$ as with a 20nm particle with $\rho = 2.5$

If we are prepared to wait 1 year for our observations, we will find that a large polymer gives essentially the same curve as a somewhat smaller nanoparticle with a larger density.

Remembering that there is a distinction between thermodynamics (the blue equilibrium curve) and kinetics (the results after a long wait) it's amazing to

learn that with some clever experiments⁶ where the tubes are created from an osmotic membrane, the equilibrium values for those experiments (polymers or nanoparticles, it makes no difference) can be reached in a few days. If this "osmosedimentation" trick was easier to do, we would be much more at ease with the fact that gravity makes *everything* settle, not just particles.

We have a tendency to think of particles as objects that will clump together unless we actively protect them with charges or steric stabilizers. This is true. So what do biologists do when handling proteins? They take great care not to add too much or too little of the wrong salt, or they might avoid adding sugar to the solution because proteins can very easily be persuaded to clump together and fall out of "solution".

We can try a thought experiment. We take a polymer chain that is lightly crosslinked within itself, though not to other polymer chains. We have it in a good solvent where it swells happily and is clearly nicely soluble. We increase the degree of crosslinking. It gets harder to keep it in solution till a highly crosslinked polymer is impossible to get into solution, and the best we can hope for is some swelling by the solvent. Across that range of behaviour we can use well-known (Flory-Huggins and Flory-Reyner) solubility theories and never have to invoke dispersion science. Now let us go the other way. Start with a solid silica particle and gradually increase its porosity until it is a flimsy assembly of silica chains. Do the experiment in, say, hexane and at no stage would "solubility" seem the right word to use. Do the experiment in water and it would be hard to defend the idea that all those SiOH groups in the porous silica are not making the particle soluble.

Now try some serious thermodynamics. One feature of some solutions is that they have a (confusingly named) lower critical temperature (LCT), above which the solute phase separates. Solubility theorists love analysing such cases and can precisely explain the observed spinodal system. The LCT is a precise, reproducible value. After all, this is solubility thermodynamics. Now take some silica dispersions, heat them to a specific temperature and ... you get a phase separation. It is precisely equivalent to an LCT.

Our training from DLVO teaches us that we must erect barriers to stop the particles from coming together. Putting charges on the outside of particles really does stop them coming together - as does adding charges to a protein. Just as changing the pH of a dispersion can take it through an isoelectric point where the zeta potential is close to zero and the particles agglomerate, so a protein can be made to fall out of solution when it has a net zero charge. The steric barrier for a particle is regarded as some sort of magical force field, until the solvent is tuned to χ just larger than 0.5 when the particles agglomerate. Take

⁶ Alfredo T. N. Pires, Suzana P. Nunes, and Fernando Galembeck, *Osmosedimentation: Approach to Sedimentation Equilibrium under Gravity*, Journal of Colloid and Interface Science, 98, 1984, 489-493

a standard polymer through its theta solvent point (when $\chi=0.5$) and it also falls out of solution (undergoes a spinodal decomposition).

Our intuitions tell us that particles are different because they are large, while molecules are small. Yet some DNA molecules are large enough to see with a light microscope (so comparable to a typical particle). Yes, solubility theory shows that, other things being equal, larger molecules are less soluble than smaller ones for simple entropic reasons - they cause a lot of reduction in freedom for the solvent molecules without gaining much freedom for themselves. All this means is that if we have a larger molecule we have to try harder via enthalpy to keep the system stable.

This entropy/enthalpy argument gets to the core of why particles seem intuitively so different from molecules. A rigid particle the same size as a large molecule is going to be less soluble because reductions in entropy are larger. Because biological molecules usually need to be highly structured, they too lose out on entropy, so they have to compensate via enthalpy - with plenty of charged groups to make them soluble in water.

We can reverse the Mewes and Wagner view. They say that soluble polymers and proteins can be regarded as colloids. It works the other way - particles can be (and in my view *should* be) regarded as soluble.

2.6.2 Fundamental thermodynamics

Here's another definition of a dispersion from IUPAC, the International Union of Pure and Applied Chemistry, an authoritative body that someone like me should not argue with. A dispersion is a:

“[m]aterial comprising more than one phase where at least one of the phases consists of finely divided phase domains, often in the colloidal size range, dispersed throughout a continuous phase”.

By this definition, a dispersion is at least two phases - the particles and the bulk phase in which the particles sit. If we could prove that nanoparticle dispersions are a single phase, then suddenly we have no way to distinguish between a single phase solution and a single phase "dispersion".

Happily, such a proof has been provided by Shimizu and Matubayasi⁷ and because it is based on deep thermodynamics there's no argument against it. Ultimately it calls upon the Gibbs Phase Rule, one of those facts of the universe against which it is impossible to argue. What the paper does is take routine observations from the nanoparticle world, such as gravitational settling, osmotic

⁷ Seishi Shimizu and Nobuyuki Matubayasi, *Thermodynamic stability condition can judge whether a nanoparticle dispersion can be considered a solution in a single phase*, J. Colloid Interface Sci., 575, 2020, 472-479

pressure measurements or light scattering and show that the methods used day-to-day to interpret the measured data only make sense if everything is a single phase. And if it's a single phase, then it's no longer a "dispersion" if you believe that this implies that the particles are in a second phase.

They go on to show how the standard tools of deep statistical thermodynamics that are routinely used to analyse "solutions" work equally well for nanoparticle systems. By the end of the paper, there's no place to hide for those who want to deny that solubility theory applies to nanoparticles.

Particle in a box

One of the intuitions behind the idea that nanoparticles are a separate phase comes from the obvious fact of the huge size mismatch between particle and solvent. For thermodynamics the size is of no importance in a large-enough system, which means essentially all the systems of interest to readers of this book.

One of the subtleties of the Shimizu & Matubayasi paper is that it can give a definition of when a particulate system transforms from being a concentrated solution into a crashed-out mess. The definition depends on the size of the fluctuations away from the average of particle concentrations. When these are larger than the average interparticle distance then the system has moved from one to two phases.

An interesting corollary of this is that if your nanoparticle system is confined to a small space - say inside some nanofluidic device, then it is very much a 2-phase system and classic solubility science ceases to be directly applicable. Interestingly the same logic applies to large biological molecules in cells - to remain "soluble" requires more than just water (of which there is rather little in a crowded cell).

The authors are well aware that nanoparticle systems can crash out, just as polymers can crash out of a solvent, or a two-solvent system changes from miscible to immiscible when, say, the temperature is changed. Yes, nanoparticles phase separate when things go wrong, and, yes, things go wrong more easily for nanoparticles for well-understood reasons. But a nanoparticle phase separation is no different in statistical thermodynamic terms (i.e. assumption free, deep thermodynamics) from a polymeric or solvent phase separation.

Why do nanoparticles crash out more easily? Because they are large, and the laws of statistical thermodynamics show that, other things being equal, larger molecules/polymers/particles are less soluble because of their excluded volume, i.e. the obvious fact that where they are there's no room for other molecules.

If, for example, you try to mix hexane with a polar solvent, you might well find that they are miscible at all proportions. If you try decane you might find a small region where they are immiscible. If you try hexadecane, they might be mostly immiscible. The molecular interactions are the same - boring alkanes interacting with polar molecules. It's just that the hexadecane is larger, and its bigger excluded volume means that the thermodynamics of mixing get off to a worse start than with hexane.

Similarly, it might be no problem to dissolve a 10,000 MWt polymer in a given solvent, but impossible with 100,000 MWt, which phase separates.

Before Shimizu and Matubayasi it was possible to claim that nanoparticles were a separate phase and that they were a dispersion, as defined by IUPAC. That excuse has gone. Nanoparticles are governed as much by solution thermodynamics as small molecules, solvent blends, hydrotropes, polymers, proteins and DNA. It is no coincidence that it is also Shimizu and Matubasi who have united so much of the solubility (and phase separation) theory across those fields. Statistical thermodynamics of solutions is just that - an analysis of solutions, and it works just as well for nanoparticles.

2.6.3 Using solubility

The historical insistence that particles are "dispersed" rather than "soluble" has been deeply unfortunate. It has cut off the formulator from a broad range of tools that are taken for granted in the solubility world.

At the same time, as discussed in my Solubility Science book, the solubility world has done itself no favours by clinging to outmoded ideas such as water structure or by trying to explain everything via bulk entropy/enthalpy arguments rather than via molecular-level tools.

The good news is that the combination of proven solubility tools plus the tools from SF theory, plus recently revived molecular-level tools (based on Kirkwood-Buff theory⁸) allows us to formulate complex particle systems more effectively.

In other words, the reason for adopting solubility-based ideas is not so much because they are right (which they are) but because they work!

2.7 Polymers on particles

In classic colloids texts, polymers are an endlessly fascinating source of interesting effects, some of which stabilize particles and many of which destabilize. In the formulation world, polymers are things we have to add to get the other functionalities required by the overall product. With such different

⁸ Sadly, for this first edition such tools are not ready for use - watch this space!

mindsets, it is not obvious how a formulator can take the hard-won lessons from the pure science and apply them in the lab.

My view is that many of the problems disappear if the solubility approach is followed. Because solubility is the next chapter, in this section we will try to look at the polymer problem with a more conventional approach - to highlight the problems from the conventional approach which we can then largely solve by different means.

2.7.1 Locked and loaded

There is one simple principle we can all rely on. If you can cover your particles early on in the process with a polymer stabilizer and if one end of the stabilizer is strongly locked to the surface (these are "brush" systems) then your formulation is going to be much simpler to control and optimize.

We need to load the surface sufficiently so there are no unprotected spots where self-association can take place. We need to lock the polymer onto the particle so that we never have bits of the polymer floating off and cause chaos.

If the polymer is nicely soluble in the relevant solvent (i.e. χ parameter less than 0.5) the particles will be nicely stable, providing that the locking mechanism is not affected by the solvent. Chemical attachment is relatively permanent. Some sort of brush-comb where the comb loves the particle and the brush loves the solvent also works.

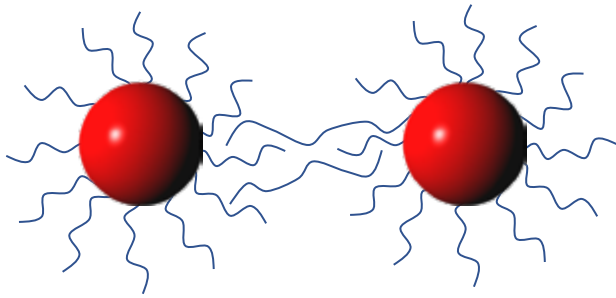
It is all very easy, and in terms of basic particle stability it makes large chunks of colloid science unnecessary as this simple principle of load and lock works reliably. Why is it not the default? My view is that it is a chicken and egg problem. Too few formulators demand chemically attachable brushes (or really effective brush-combs), so the suppliers don't supply them. Too few suppliers supply them or actively promote them (because customers aren't asking for them) so the formulators don't bother to ask for them.

There's also the well-known problem that it is more-or-less impossible for formulators to contact the sorts of scientists within the large organisation who now supply most of these products. They are shielded by the standard sales teams who lack the knowledge to have an informed conversation about the right chemistry for attachment and the right (Hansen) solubility properties to work well with the rest of the formulation.

We will explore the trade-offs necessitated by real formulations in the Solubility chapter.

2.7.2 Bridged

If you are foolish enough to just throw in some particle-loving polymer, it is likely that the result will be a frustratingly non-reproducible mess.

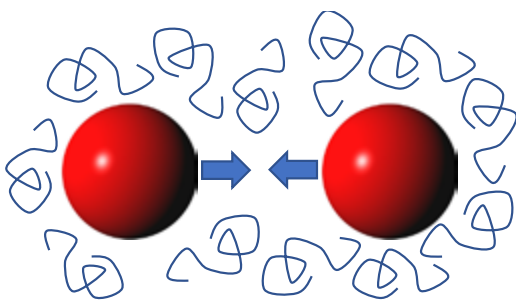


The reason for the mess is bridging flocculation, where, as conventionally described, one end of a polymer chain attaches to one particle and the other end to another particle. There really is no excuse for finding yourself in this mess because the cause and effect are so

clear. As it happens, this diagram is completely wrong, but we need to get to the SF chapter to discover why.

The reason that the mess is non-reproducible is that if you try hard you can get the polymer to stabilise the particles; with enough polymer surrounding each particle, there is no longer the possibility of bridging. The problem is that even if you start with enough polymer for full coverage, by varying the speed and energy of mixing then you can end up with just about every possible scenario, because kinetics and thermodynamics interfere with each other. If you really want strong flocculation, take 100% covered particles and mix them with 100% uncovered particles (yes, someone has done the experiments). Even though 50:50 coverage would be enough to keep them separated, the shock of the encounter is enough to crash them out of solution.

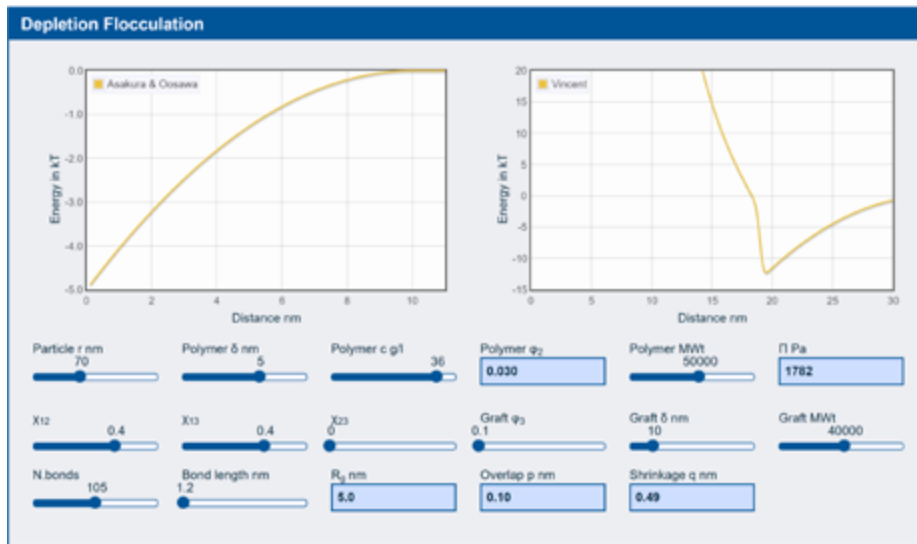
2.7.3 Simple depletion



Frustratingly, addition of a non-interacting polymer (one with less than $0.5kT$ interaction energy) which might be there for other purposes, can also cause the particles to crash out. The problem is that because the polymer chain don't interact with the particles, if the particles are close enough for the

polymer to become uncomfortable, the polymer simply moves away from the gap. This means that there is now a concentration gradient of polymer, which means an imbalance in chemical potential which has to be filled somehow, which is done by the particles coming together. The classical shorthand for this is to say that the depletion in concentration creates an osmotic pressure which drives the process, which is therefore called depletion flocculation.

For those who would like an app to get an idea of what is going on, the theory of Asakura and Oosawa is the one most widely known. You are welcome to try it out, plus the more powerful Vincent theory included on the same app page:



App 2-5 <https://www.stevenabbott.co.uk/practical-solubility/depletion.php>

It allows you to explore the effect of polymer radius, δ , (likely to be the radius of gyration R_g if you happen to know it) concentration and MWT plus the radius of the particle. You will see the calculated osmotic pressure and get an idea of the attractive energy, in terms of kT , between the particles as they get closer. The app page gives the relevant formulae.

Because the driving force is simple exclusion, you can get the same depletion flocculation effect by adding small particles or surfactant micelles.

The one positive thing about depletion flocculation is that it is a rather gentle force and flocs are readily re-dispersed.

2.7.4 Vincent depletion

Many descriptions of depletion flocculation stop at that point - which is unhelpful because the questions are only just starting. The first three questions lead to the one question we all want to know the answer to:

1. what about particles with steric shells around them?
2. what about the degree of coverage of particles?
3. what are the solvent effects with respect to the depleting polymer?
4. what is the critical concentration required to produce flocculation?

The wonderfully named "pragmatic theory" from Prof Brian Vincent was at one time proposed to help us along the way. As he points out, most of the systems *studied* had been hard spheres, though most systems *used* were steric stabilized, so it was time for a theory that was more relevant to the real world. It assumes the standard depletion energetics of the original hard sphere approach and modifies it with a mixture of the energies of the shells overlapping (repulsive) and the polymer overlapping with the shells. The key idea is that the

overlap energies can be worked out via the equivalents of Flory-Huggins theory which means that the various χ parameters are relevant. This, in turn, (though Vincent didn't do this) means that we can use our standard solubility theory (HSP) to better understand what is going on.

Although I managed to implement as much as I could of Vincent's theory, and readers are welcome to explore the app, the reality is that it never proved very popular. More importantly for us, SF theory came along making it possible to calculate the whole range of particle/polymer effects, not just the depletion subset. All this will make sense when we come to the SF chapter.

2.7.5 Beyond bridging and depletion

If you add enough extra polymer all sorts of good and bad things can happen. The polymers might interact so much with themselves and the polymers on the particle that everything is compatibilized. Or the polymers might self-organise and start excluding the particles which themselves start to self-organise. Or you can get gas-liquid phases. Such is the strange (to me) terminology used in theories such as those of Lekkerkerker which can create complex phase diagrams with relatively simple approximations about polymer, particle, solvent mixes. For a while, I got quite excited about these phase diagrams, even though they were hard for me to comprehend. But eventually I realised that if you have relatively large particles (the bulk of the interest in this book) and relatively small (i.e. normal) polymers then the phase diagrams say that with too much polymer or too much particle you get phase separation, which is what we already knew.

For those of us who need to formulate in the real world, then the SF theory we shall come to in a later chapter provides the most practical way to think through all these competing complexities.

2.7.6 Dissolving silicon in DMSO

If it looks like a duck, walks like a duck and quacks like a duck then it's probably a duck. So if you take some 10 nm silicon⁹ (yes, this is silicon, not silica) nanoparticles in a tube, add some DMSO, shake and get a bright red solution that is stable for 3+ years, but where you can precipitate it by adding an excess of a non-solvent like water, then there's no other word to describe what's going on other than saying that it's a solution of silicon.

As it happens, pure silicon nanoparticles aren't soluble in DMSO - you first have to treat them with diborane to make an outer molecular layer that happens to interact nicely with solvents such as DMSO or NMP.

⁹ Lance M. Wheeler, Nicolaas J. Kramer, and Uwe R. Kortshagen, *Thermodynamic Driving Force in the Spontaneous Formation of Inorganic Nanoparticle Solutions*, Nano Lett. 2018, 18, 1888–1895

The authors of the paper are adamant that this is a solution. They specifically take DLVO to task because the solvent in DLVO is just "a continuous dielectric medium" not something where real molecules interact (in ways that are commonplace in other solubility situations) with the surface. They show using conventional FTIR techniques that the carbonyl stretch of NMP in contact with the surface shifts by $\sim 40 \text{ cm}^{-1}$ as would be expected from a donor/acceptor interaction with the boron on the surface.

Although I was delighted when I found this excellent paper, I was also rather saddened by it. Why, in 2018, does anyone have to publish a paper saying that nanoparticles are soluble when there has never been a good reason *not* to use solubility language? It's partly habit, it's partly dogma, it's partly the unjustified pre-eminence of DLVO. And it's partly language: if you call everything with a particle a "dispersion", even if no one can define what "dispersion" means nor how to distinguish it from "solution", everyone just accepts that you can only use "dispersion science" and not "solubility science". Well, that's just not acceptable anymore.

3 Solubility of particles

At one time, trying to appease the particle community who were so strongly against "solubility" I started to adopt the word "similarity". As a short-term tactic this was helpful but I now regret it because scientifically I am convinced that we can adopt the full range of solubility science to help us formulate with particles. If you prefer to hear "similarity" each time I write "solubility", be my guest. Whatever the word, the point is that it works well in practice. As we shall see in the SCF chapter, the most powerful general theory usable by our community is based on the same core solubility ideas. In this current chapter we will emphasise a pragmatic solubility-based approach that serves our community very well. The SCF chapter then shows how a more profound methodology allows us a deeper understanding of the phenomena that interest us. SCF requires χ parameters and this chapter gives us a pragmatic way to understand and estimate χ parameters.

In the Solubility Science book I've written a whole chapter on the many different solubility theories that are out there. I can save you the bother of reading it because the conclusion is stark. For the sorts of messy real-world problems that interest us, there are only two solubility theories that are relevant, however splendid some of the alternatives might be in their pure solubility domains.

Of the two relevant theories, the powerful Kirkwood-Buff approach is still not ready for general-purpose use¹⁰. So here we use the rough and ready, and astonishingly successful Hansen Solubility Parameter (HSP) approach.

3.1 HSP Basics

If, at first, we stay safely with normal solvents and polymers we can quickly grasp the essence of HSP. Indeed, their development by Charles Hansen in 1967 was based on solvents and polymers. It was only later that Hansen tried out his own ideas on systems such as the pigments in paints and found that they worked just as well.

What we want is a numerical way to describe the chemical essentials of any solvent or polymer. Clearly this cannot be done with a single number - though this hasn't stopped countless bad publications using a single number, the Hildebrand Solubility Parameter, δ . However, we can all agree that every molecule must have a number describing its general van der Waals (vdW) interactions because all molecules, however dull, are attracted via vdW. These interactions depend on the electrons sloshing around a molecule or, in better language, the molecule's polarizability. Dull molecules have little polarizability (manifest as a low refractive index, RI), less dull molecules have more and when you start having aromaticity, halogens and sulfurs, the refractive index,

¹⁰ As mentioned in a previous footnote, I had hoped to have a whole chapter on Kirkwood-Buff for nanoparticles but there is not yet sufficient data and theory combined to make it worthwhile to write one.

the polarizability and the tendency to stick together (giving, for example, a high boiling point and enthalpy of vaporization) all increase. We can call this aspect of a molecule the Dispersion parameter and label it δD .

Next we need to capture the polarity of a molecule, something that is strongly related to its dipole moment. We call this the polar parameter, δP . Although there have been attempts to stop there, they make no sense because intuition and good science tells us that although a molecule like ethanol is polar, it also has the ability to interact via hydrogen bonds, while acetonitrile is highly polar yet only weakly hydrogen bonding. There is no way to capture "polar" and "hydrogen bonding" in a single value, so we add a third value, the hydrogen bonding parameter, δH .

And that's it. Three numbers to describe any solvent or polymer. For solubility theory we need to add (because of the entropic effect) the size of the molecule, better captured as its molar volume, $MVol$, than its molecular weight MWt . But I will not explore the impact of $MVol$ very much.

An obvious objection to a 3-parameter system is that hydrogen bonding should itself be described via two numbers - Donor and Acceptor. While everyone (including Hansen) agrees that this is a great idea in theory, in practice no coherent 4-parameter scheme has ever proved widely usable. A simple thought experiment explains why.

We often find ourselves making blends of solvents, because no single solvent can give us the properties we desire. Suppose one solvent is a pure H-bond donor and the other is a pure H-bond acceptor. If we mix them, what happens to their H-bonding abilities with respect to the solute? Do they associate so strongly with each other that the solute is ignored? If not, what fraction of their donor/acceptor gets used on the solute? Although there are complicated approaches to answering these questions, they have never been convincingly applied by formulators in the real world. So a 3-parameter system is much more usable.

We can even more quickly refute the need for a 4-parameter system. If it was necessary, the 3-parameter HSP system would have been a failure from the start. But HSP are more used today than they have ever been, so the 3-parameter approach must be good enough for many purposes. Of course there will be special cases (the Coleman-Painter polymer miscibility approach, discussed in the Solubility Science book, is a key example) where 4 parameters are required. And we need to be alert to donor/acceptor effects in any specific system. But in general, our 3 parameters work well thanks to the key fact that via them we can calculate the HSP Distance.

3.2 The Distance D

Here's the trick. We have two molecules (e.g. a solvent, 1 and a polymer 2) each with their 3 HSP $[\delta D_1, \delta P_1, \delta H_1]$ ¹¹, and $[\delta D_2, \delta P_2, \delta H_2]$. We calculate the Distance, D as a slightly modified (δD is privileged with a factor of 4) sum of squares of the individual differences:

Equ. 3-1
$$D^2 = 4(\delta D_1 - \delta D_2)^2 + (\delta P_1 - \delta P_2)^2 + (\delta H_1 - \delta H_2)^2$$

It is as simple as that. If D is small, because all three of the pairs of parameters are similar, then the two molecules are going to interact with each other as if they were interacting with themselves. In other words, a small D means that there is no enthalpic penalty in mixing. If there is a large difference between one of the pairs, or if they are all relatively different, then the two molecules don't like associating with each other as much as they like themselves, so there is an enthalpic penalty in mixing.

Before seeing how we determine the HSP values for a polymer and then for a particle, we first need to get a feel for the HSP values themselves.

3.3 The HSP values

A key advantage of HSP is that their values make intuitive sense. After a while, an HSP user can make a reasonable guess as to the values of a relatively straightforward molecule and not be too far wrong. We will look at the three parameters one value at a time.

- **δD** represents the dispersion or van der Waals component of a molecule and for simple molecules it represents the whole of HSP. We know that even a boring molecule such as hexane has a significant enthalpy of vapourisation and we know that δD must (via the square root of that enthalpy) be a modest, finite value. In fact for hexane that value is $\sim 15 \text{MPa}^{1/2}$. For solvents that have very low enthalpies of vapourisation such as silicones or fluorocarbons, δD can drop to 11 or 12. Cyclohexane is more compact (higher density) and more self-associating so its δD is nearly 17. Aromatics happen to have some H-bonding component but they are mostly δD and we know that they self-associate (π - π^*) and are harder to evaporate so it is no surprise that benzene's δD is 18. Once we start adding chlorine or sulfur atoms to an aromatic we can easily be up in the low 20s. And that's about it for δD , other than to point out that it has a strong correlation with refractive index which is, at a deeper level, a correlation with the polarizability of a molecule. Those molecules with more electrons able to move freely at the surface have a higher polarizability, a higher refractive index, stronger van der Waals interactions and a higher δD . δD is often dismissed as boring, yet it is often

¹¹ All HSP values are shown in this format, so [17, 8, 12] would mean $\delta D=17, \delta P=8, \delta H=12$

the dominant HSP component because *all* molecules have van der Waals attractions while polar and H-bond components, when present, can often be relatively small.

- δP is strongly related to dipole moment, so we know that hexane will have a value of 0 and that something with a large dipole moment such as acetonitrile will have a large δP . It turns out to be 18 and it then is no surprise that DMSO is ~ 16 , acetone is 10, THF is 6 and so forth. You might not have got those exact values, but the general trends make sense and there are no big surprises. A symmetrical molecule such as 1,4-dioxane has no dipole moment. In this case the correlation breaks down because dioxane is not a "non-polar" molecule.
- δH is, once again, zero for hexane and for methanol it is 22. This allows us to scale ethanol and propanol as 19 and 17 because their -OH is "diluted" with CH_2 groups. DMSO is a good H-bond acceptor with a value of 10 and acetone has a value of 7. When Hansen did his original experiments he found that he had to give a small value of δH to aromatics like benzene and toluene. This "mistake" was the cause of some criticism but Hansen stuck with experimental data and kept those values. It was, therefore, with some satisfaction when, 40 years later, the small but significant H-bonding abilities of aromatics were formally recognised by IUPAC in their definition of the H-bond. The values for simple aromatics are ~ 2 .

Thus all three sets of values conform to a chemist's intuitions. There is a great benefit to this. There are many systems for estimating HSP from molecular structures and many of them do a reasonable job in their comfort zone. But try a molecule with a somewhat different structure and the system can provide ludicrous values. The same thing can happen with measurements. You follow a measurement recipe, click a button to get an answer and you find some numbers that don't make intuitive sense based on your knowledge of the system. This allows you to go back over the estimation or measurement process to see if you've done something wrong, or, for measurements, it warns you that there is something odd about this specific sample. In one case Charles Hansen and I were able to spot that a sample of quantum dots was impure, simply because the measured value made no sense in the context of values of similar samples. The team who had made the quantum dots had to go to considerable trouble to find that, indeed, that batch was impure.

3.4 Measuring HSP

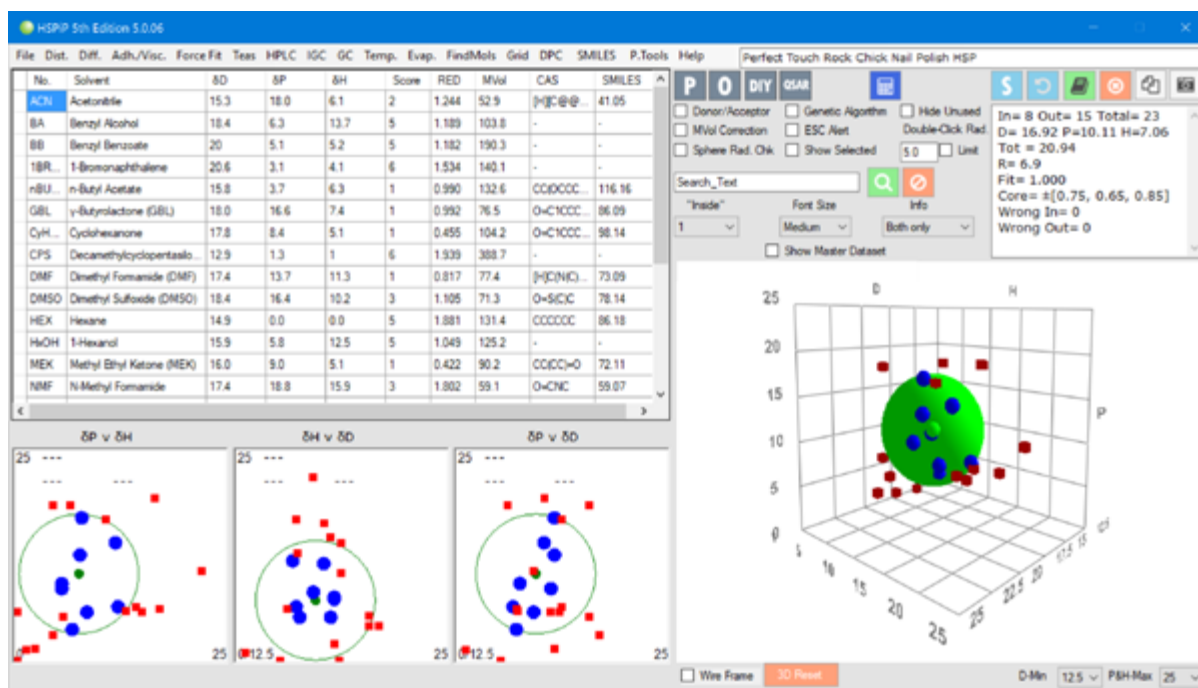


Figure 3-1 Measuring the HSP of ... nail polish

The measurement technique is quite general. Rather than demonstrate it with a pure polymer such as poly(lactic acid) here we have the measurement of the HSP of some nail polish (courtesy of VLCl in Amsterdam who have used this as a very successful training exercise). We have taken 23 solvents and into a tube of each of them we insert a glass slide coated with a dried sample of the nail polish. In very good solvents, within a few minutes it is clear that the nail polish is being attacked. In very bad solvents, even after hours there is no discernible change. Others are in between.

Using our judgement we score all the good solvents (i.e. the ones that mostly remove the polish) as a 1 and all the others as a zero.

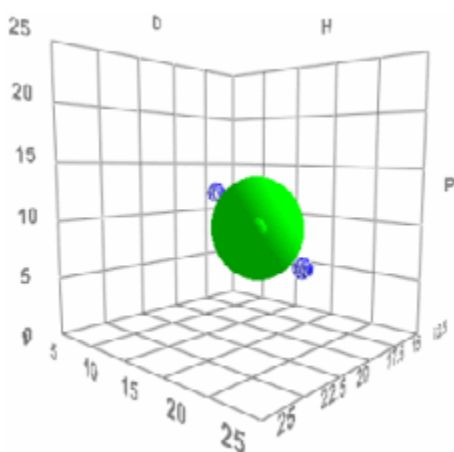
We then put the results into software (here it's HSPiP) or an Excel sheet¹² and using a fitting technique find the best sphere that (ideally) includes all the good solvents and excludes all the bad ones. For this nail polish the sphere has a centre at [17, 10, 7] which represents a practical average of the polymer (nitrocellulose) and the pigment which presumably has a dispersant which makes it compatible with nitrocellulose. We will explore the meaning of the radius of the sphere in a moment.

The measurement technique is perfectly general. As long as you have a set of solvents that are distinguished (roughly 50:50) between those in which the

12 Manuel Díaz de los Ríos, Eduardo Hernández Ramos, *Determination of the Hansen solubility parameters and the Hansen sphere radius with the aid of the solver add-in of Microsoft Excel*, SN Applied Sciences (2020) 2:676. The spreadsheet, developed from one that I originally created, is at <https://www.hansen-solubility.com/HSP-science/basics.php>

material is "happy" and those where it's "unhappy" then you can do this fitting process and obtain the HSP. The unscientific use of "happy" is deliberate. For linear polymers we might look for "soluble". For crosslinked polymers we might look for "swellable". For protective gloves we might look for "penetrates quickly". For particles we might look for a stable dispersion (I'm using the generic term even though I believe that solubility is at the heart of it) using the eye, or as described shortly, a centrifuge system. For nail polish we can observe that good solvents remove the polymer and the particles, bad solvents touch neither and intermediate solvents are struggling with the particles.

To show the power of the numbers we've obtained we can ask ourselves the question: "What other solvents could be used as nail polish removers?" The answer is simple. Assemble a list of solvents that you might class as green or safe or whatever other criteria appeal to you and calculate their distances from [17, 10, 7]. Those solvents with the smallest distance are going to be good solvents, provided they aren't large molecules because, as stated above, a high MVol is thermodynamically (and kinetically) linked to a lower solubility¹³.



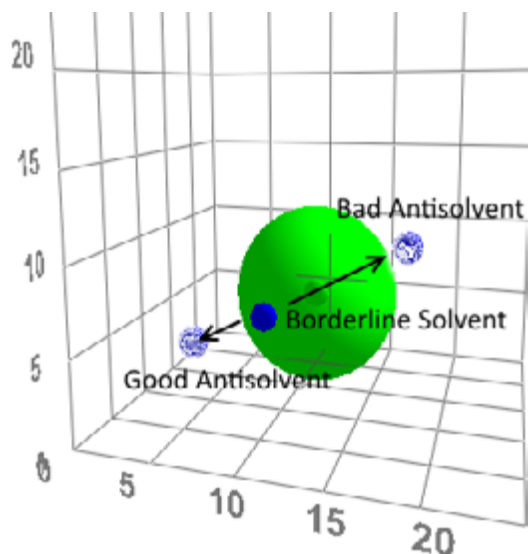
Suppose, however, that two solvents we like for other reasons are bad solvents. Our measurement allows us some formulation magic. The HSP of a blend of solvents is the volume averaged mean of their values. As the diagram shows, if the two bad solvents happen to be on the opposite side of the sphere and, necessarily, outside it, then a 50:50 blend would be a solvent at the centre of the sphere. Two bad solvents can create a good one.

This principle has been demonstrated countless times. In the case of the nail polish tests, neither benzyl benzoate nor benzyl alcohol is scored as a good solvent, but a 50:50 blend, as predicted from the 50:50 average of their HSP, is quite good. No solubility theory other than COSMO-RS (discussed in the Solubility Science book) makes it so easy to generate good solvents from bad ones, or turn not-very-good solvents into excellent ones.

This trick allows great formulation freedom. If your single best solvent turns out to be too expensive, (in)volatile, odorous or unsafe you can replace it with a pair of solvents tuned to the right mix of cost, volatility, odour and safety, along with the necessary solubility. You can be even smarter. In a fairly dilute solution of some formulation, necessary, for example, to provide a low-enough viscosity, you can have a lot of relatively poor solvent that is both low cost and relatively

¹³ It is very easy to find excellent green solvents to replace acetone or ethyl acetate for nail polish. But the green solvents are slow to evaporate and consumers prefer convenience to green-ness.

volatile. As the poor solvent evaporates, the HSP shifts towards the better one, keeping the formulation nicely in solution until the last moments.



We can also understand a different sort of puzzle. It is common to add anti-solvents to a formulation to get something to precipitate out for further processing. You can take two equally bad solvents yet find that you need to add 10% of one of them and 90% of the other to achieve the desired precipitation. The explanation is visualised in the diagram. The "good" solvent has been chosen (wisely) to be borderline. As you add the good anti-solvent, the HSP rapidly shifts to a value outside the sphere and precipitation is rapid. The bad anti-solvent

at first makes the solvent blend better, so you need to add a lot of it before the blend comes outside the sphere.

3.5 The radius of a polymer HSP

What is the significance of the radius? Or, to put it another way, what is the Distance that distinguishes between a good and bad solvent?

The answer at first sounds odd: the radius is whatever you decide is appropriate for your system. Suppose, for example, that you are looking for solvents that keep some larger particulates suspended and stable for a week. You test 20 solvents and only 3 of them do the job. Suppose, at the other extreme, that you just want the particles to last 30min in reasonable suspension. Now those same 20 solvents might give 10 that do a good job. The sphere for the first set will be much smaller than for the second. If you then search for alternative solvents, the Distance that a solvent must be to be classed as "good" will be smaller for the first case. This is a feature, not a bug, in the HSP system. By being pragmatic about the radius we gain the ability to formulate effectively for whatever purpose is of interest to us. If you need stability for a week your choice of solvents is much smaller than if you need stability for 30min.

As it happens, a radius of 7 to 8 is the thermodynamic value found from careful analysis of the stability of polymer solutions. The precise value depends on the MWt of the polymer and the MVols of the solvents. The theory is Flory-Huggins theory, mentioned briefly in the first chapter, and can be explored in my app: <https://www.stevenabbott.co.uk/practical-solubility/polymer-solubility.php>. But almost no one measures polymers this carefully. I once worked to get the radius of poly(lactic acid) and it was, indeed, 8 as expected from basic Flory-Huggins theory described in the next section. I used a standard PLA from a normal supplier. But this value turned out to be useless for the specific problem because

the polymer being used in that formulation was different. It had a high MWt and was highly crystalline. Its effective radius, judged by its solubility after a short exposure to solvent, was 4. Using the value of 4 we could create a great solvent blend for the formulation. Had we used the "scientific" value the appropriate blend (because it had to be near the radius) would have been useless.

Because we will make extensive use of χ for SF theory, let's look at it more closely.

3.6 The Flory-Huggins χ parameter

A polymer in a really good solvent is relatively well-extended. The polymer chain gently curves in a more-or-less random walk and you get a nice ball of polymer-in-solvent. It has an average radius, R_g which is the radius of gyration. R_g depends on the MWt of the polymer, doubling in size for a quadrupling of the MWt. Now decrease the quality of the solvent, so the polymer is less happy. It will curl more upon itself, so R_g will decrease. At a critical quality of solvent, called the θ (theta) solvent, we are about to tip into a regime where the polymer is clearly very unhappy in the solvent and solubility decreases catastrophically. The small amount of polymer that is dissolved is very much curled up on itself.

The relative happiness of the polymer in the solvent is captured by the Flory-Huggins χ (chi) parameter. When this is 0, the polymer is very happy, when it is 0.5 the polymer is at that delicate θ point and higher than 0.5 the polymer starts to crash out.

Because χ is such a fundamental value, there has been a whole academic industry of measuring χ values. As it turns out, many of these values are wrong, as beautifully described in a paper whose title doesn't pull any punches¹⁴: "Beware of the Flory parameter...". Fortunately, HSP provides us with a pragmatic way to work with the χ parameter. As pointed out many years ago by Charles Hansen, a typical HSP radius of 8 for a polymer, defining the limit of solubility, equates to a χ value of 0.5. Via some simple arithmetic we find that:

Equ. 3-2

$$\chi = \frac{MVol.D^2}{4RT}$$

So if we know the distance, D , between a polymer and a solvent of a known molar volume, $MVol$, we have a good idea of χ .

As we shall see in the SF chapter, if we know D for solvent and polymer and D for solvent and particle, we have a good idea of an important type of χ value called χ_s which describes the relative affinities of the polymer and the solvent for

¹⁴ G. Miquelard-Garnier and S. Roland, Beware of the Flory parameter to characterize polymer-polymer interactions: A critical reexamination of the experimental literature, Eur. Polym. J., 2016, 84, 111–124.

the particle. This enables us to do some really nice calculations of how well, or badly, a polymer might act as a dispersant.

But first we need to explore more about HSP and particles.

3.7 Particle HSP values and radii

Many scientists might have chosen simply to decide that particles cannot be soluble and that it would therefore be a nonsense to apply HSP to particles. But in the 1970s Charles Hansen had some formulation challenges while working for the PPG paint company and decided to see if HSP worked as well for his paint pigments as they worked for his polymer formulations. If this pragmatic test had failed, you would not be reading this book. But it worked fine, so for many in the paints industry, measuring the HSP of particles, and their radii, became commonplace.

For some reason, the idea did not spread across much of the particle world. Another Hansen experiment many years later measured the HSP of the C60 Buckyballs which, as everyone agrees, are "soluble" so the values were not controversial. Then along came carbon nanotubes (CNT). If you assumed that they, too, were soluble you could measure their HSP and, no surprise, they were not too far from C60. Since that time it has become more and more routine to measure the HSP of particles. While the radius of the HSP of a polymer can be mapped on to Flory-Huggins theory and the χ parameter, there is, currently, no theory of the solubility limits of particles. So we just have accept the measured radius values and work with them.

Let's look at CNT. In his PhD thesis at U Namur, Simon Detriche judged the stability of CNT in various solvents using the human eye and the blackness of the suspension as a criterion. Bad solvents were easy to see (I saw the samples myself) because the CNT were sitting in an unhappy clump at the bottom of the tube. Really good tubes were scored 1 and not-so-bad tubes scored 2.

If we allow just the 1's to be good then we get a rather small sphere, which is what we'd expect from hard-to-disperse CNT. If we are a bit more relaxed, because we want a larger range of solvents and the amount that is dispersed in the 2's isn't so bad, then we can refit the sphere including the 2's. The values aren't so different.

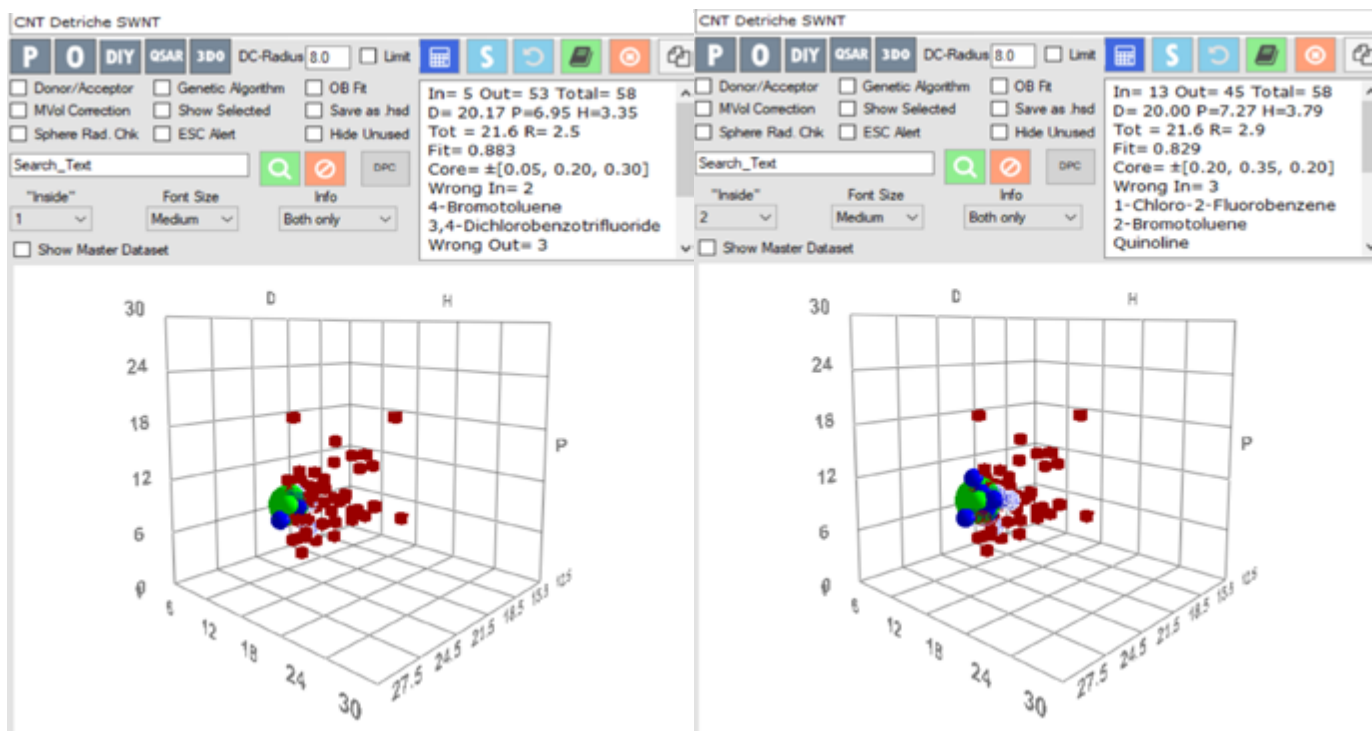


Figure 3-2 The radius of CNT depends on your definition of a good solvent.

Of course the radius is rather larger, it's gone from 2.5 to 2.9, but that's exactly what we expect. Although it seems that "the radius is what you want it to be" is sloppy science, it turns out to be a key part of HSP's ability to allow us to formulate particles rationally.

If you need to work with particles, it's not enough to know their HSP - you have to have some idea, via the radius, how much room for manoeuvre you have as you juggle other priorities with your solvents and solvent blends. If you *must* have a high concentration of CNT then a solvent blend with a distance of 3 will not be helpful, so a different blend will be needed. If you just want some general compatibility then 3 is probably OK, though be alert to formulation creep that takes you to 3.1 or 3.2...

It's your choice as a scientist: work only with pure theories that you can't apply or with pragmatic theories that work rather well. I know which I choose. However, there comes a very wise warning from a senior solubility researcher in a large chemical company. His day job (rightly) required him to work with the high powered COSMO-RS software for precise solubility calculations on well-defined systems, something for which HSP is not good. To my surprise, he was a great supporter of HSP being used in day-to-day lab formulations. The warning came like this: "My problem is not to get the people in the lab to use HSP. My problem is to stop them from using it when it is clearly outside its capabilities." He had to run an education campaign to make sure that users added an extra element to HSP which is a scientist's sceptical mind. I entirely agree with him.

3.8 An objective measure

The subjectivity of measuring the HSP of particles can often be an advantage. It can equally be a problem should you want, for example, to use the measurement as a quality control technique for something like carbon black to be used in batteries.

Why would anyone want to know the HSP of carbon black? It's a messy material that can be produced in many different ways. Assuming the particle sizes of different materials are the same, how do you choose between them? One way is to choose between a hydrophilic and hydrophobic version. But these words are far too vague for the sort of precision formulation required to make gigawatts of batteries. Hydrophobic can mean "insoluble in water, but soluble in everything else from hexane to methanol", or it might mean "only soluble in hydrocarbons". In other words, the term is near meaningless. Now suppose we have two hydrophobic samples, with HSP values of [17.2, 4, 2] and [17.2, 8.5, 11]. The one with the relatively low δP and δH is going to formulate very differently with respect to solvents and polymers than the one with the higher pair of values.

Now suppose your standard carbon black is known to be [17.2, 8.5, 11] but a new batch is measured to be [17.2, 6.3, 8.9]. Are you going to accept it, or should it be rejected? By knowing the HSP of other critical components in the formulation and, hopefully, by having explored HSP space early on in the development process, you and the supplier can reach a rational decision.

The decision to accept or reject a batch requires a reliable, objective measurement in the first place. An interesting methodology for making measurements more objective¹⁵ came from Prof Dietmar Lerche from LUM and Prof Doris Segets now at U Essen/Duisberg.

Their starting point is a centrifuge that allows you to follow the sedimentation of particles in real time. The system, and ways to visualize the sedimentation, are discussed below. Using this technique it is possible to derive an objective point at which you say that the particles have "settled" under the higher gravity within the centrifuge. Because the centrifuge measures a number of samples simultaneously it's possible to get Settling Times (ST) for particle in, say, 16 solvents very quickly. Because solvents have different viscosities and densities, the ST have to be corrected to produce a Relative Settling Time (RST).

We now have some objective data that allow us to distinguish between bad solvents with short RSTs and good solvents with long RSTs. But to create an objective sphere with a meaningful radius, it's important to create an objective view of which RST is a cut-off value between good and bad. The technique

¹⁵ Sebastian Süß, Titus Sobisch, Wolfgang Peukert, Dietmar Lerche, Doris Segets *Determination of Hansen parameters for particles: A standardized routine based on analytical centrifugation*, *Advanced Powder Technology* 29 (2018) 1550–1561

starts by sorting the solvents from best (largest RST) to worst (lowest RST). It then makes the guess that only the two longest RSTs are "good" and the others bad. It calculates an HSP value based on that assumption. Then it assumes that the first 3 are good ... and so on till the final guess is with only the two worst being bad.

The jump in HSP values from the 2-solvent value to the 3-solvent value is usually large, then from 3 to 4 is smaller until at some number the jump is rather small, indicating that the calculated value is in a stable regime. The jump then increases. The HSP is then taken to be from the number of good solvents that gives the smallest jump - though some common sense has to be used if there are artefacts in the fit. See for yourself:

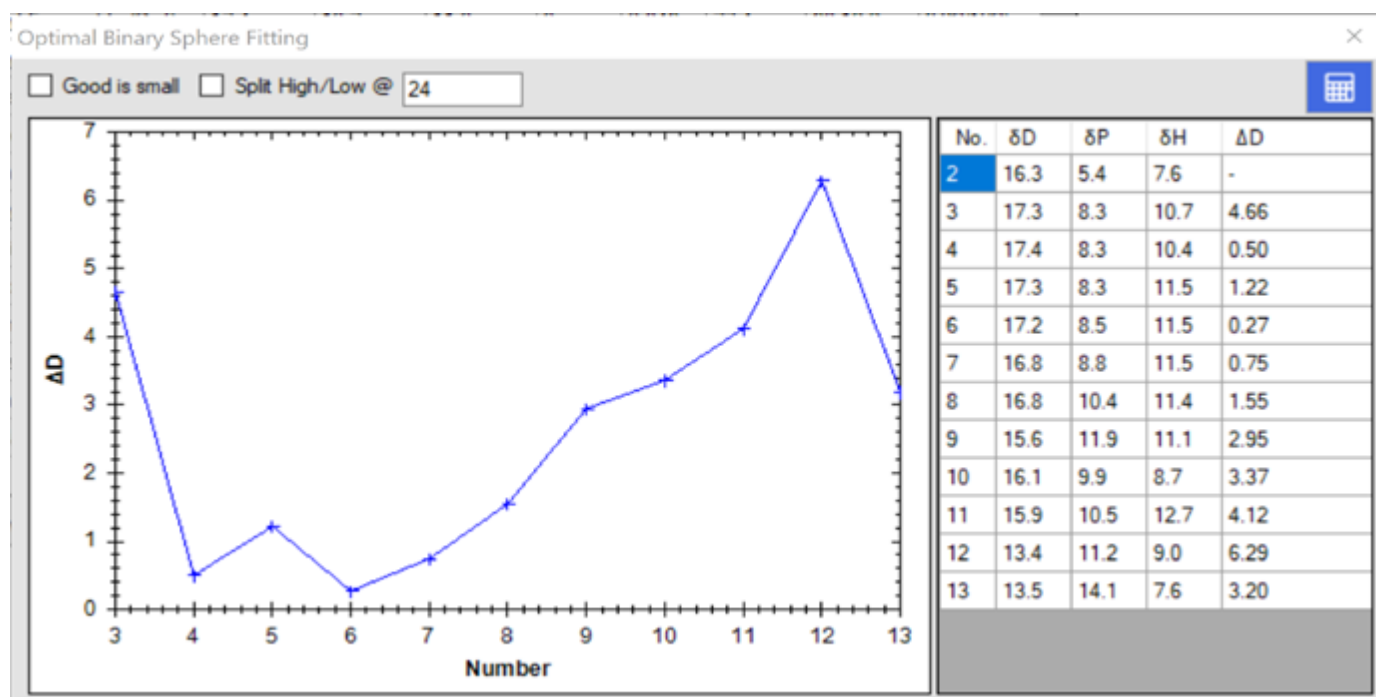


Figure 3-3 The optimal binary fitting of carbon black RST data. The value obtained by choosing the 6 best solvents as "good" seems sensible.

Although this technique was created for RST data from a LUMiSizer, it is perfectly general for any technique that creates numbers rather than just good/bad judgements. It can, therefore, be used on the NMR technique described after a short diversion into transmittograms.

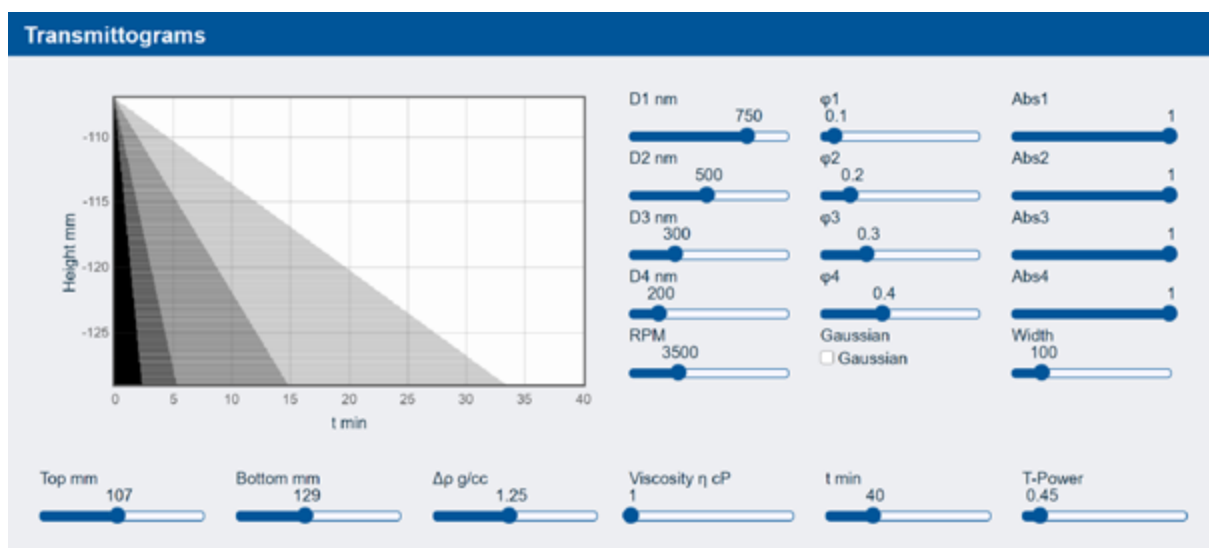
3.8.1 The transmittogram technique

Because tools like the LUMiSizer are becoming more common and make it easier to explore solubility effects on sedimentation, I have to admit to finding it difficult to understand the raw data from these devices, especially when the particles and their behaviour are more complex. This section describes a technique that makes it easier to visualise the data and, in my view, makes the technique more accessible to non-expert users.

Take a tube containing some particles and using some suitable technique, measure the light (or X-ray) absorbency along the length of the tube. Assuming sufficient particles and a suitable wavelength, define that level of absorbence as being "black". Now spin the tube in the centrifuge while (as in the LUMiSizer) following the absorbency down the length of the tube. At some suitable rotational speed the particles will move to the bottom of the tube at some suitable rate. Given a monodisperse particle in ideal circumstances, it will be relatively simple to track the transition zone from black to clear and from the speed of the zone's movement down the tube, the particle size (and the RST) can be calculated.

But what happens when you have multiple particle sizes, or a particle size distribution? There is no longer a single transition from absorbent to transparent so it is rather tricky to see what is happening from the dynamic data.

Fortunately, Shalmali Bapat, working in the lab of Prof Doris Segets at U. Duisburg-Essen came up with an ingenious way of visualising the optical density data via a Transmittogram¹⁶.



App 3-1 <https://www.stevenabbott.co.uk/practical-solubility/Transmittograms.php>

The app allows you to simulate what would be observed in a typical experiment. The example shown has a mix of 4 particles of diameters 750, 500, 300 and 200nm and with volume ratios of 0.1, 0.2, 0.3 and 0.4 respectively. After 35 min with this specific setup of length, density difference and rotation speed, the optical densities down the tube are shown at different times. To orient yourself, start at 25min. Most of the tube is transparent, but the bottom portion is somewhat absorbent from the smallest particles, all larger particles having reached the bottom earlier. Now look at the data after 10min. The top is clear, the middle has absorbence from the smallest particles and the lower portion has absorbence from both the 200 and 300nm particles. At 3min there is a small

¹⁶ Shalmali Bapat and Doris Segets, *Sedimentation Dynamics of Colloidal Formulations through Direct Visualization: Implications for Fuel Cell Catalyst Inks*, <https://pubs.acs.org/doi/pdf/10.1021/acsnm.0c01467>

clear zone then absorbance from 200, 200+300 and 200+300+500nm particles as you go from top to bottom.

The tube is only fully black in portions of the tube up to 2min when all particles are contributing to absorbance. As noted in the original paper, distinguishing the different absorbances can be tricky if a purely linear visual scale is used. In the app, the T-Power slider allows you to change the power-law dependency of the transmission visualisation. A value of 1 is linear, the value in the screen shot of 0.45 opens up the darker regions, a value larger than 1 makes it easier to distinguish detail in the lighter regions.

The Gaussian option creates a distribution centred at D1 with a chosen Width, with 10 different particle sizes. You very much need to slide the T-Power slider to see details in the different size ranges.

Now back to measuring HSP happiness.

3.9 Happiness from NMR

The HSP technique has nothing to say about why a particle is or is not happy with a solvent. Of course for systems which have a strong steric stabilization shell around them, the DLVO theory and the Flory-Huggins χ value makes sense, as does SF theory. But experience shows that the HSP values for most particles with stabilizer shells around them seem to be a mix of "particle" and "shell". And for particles without shells (such as the carbon blacks), the measured value must be a complex mix of interactions on the chaotic surface of an ill-defined material.

In the face of this general ignorance, Dr Shin'Ichi Takeda made the hypothesis that the NMR relaxation rate of a solvent might provide an objective measure for the overall interactions of the solvent with whatever mix of particle and shell happened to be on the surface. Strong interactions (from whatever cause) near the surface would lead to large changes in relaxation rates.

The idea builds on an important tradition of using NMR solvent relaxation times to study how dispersants interact with particles. We shall see shortly that the two approaches, when combined, bring immense power to the formulator, though this power also provides a caution about the naive use of the technique.

First we need to think through NMR relaxation times. If you are familiar with NMR spectroscopy then you know that in addition to the standard techniques to get discrete spectra, you can set various radio frequency pulse trains and measure how the signals from the excited hydrogen atoms "relax" over time. Those relaxation times can, for example, distinguish between rigid and flexible parts of molecules. Modern NMR machines, with superconducting magnets, can fill a room and are capable of measuring just about anything imaginable.

But forget all that. Just think of a little benchtop box with a modest magnet into which you insert a simple NMR tube and out of which you get not a spectrum but a curve showing how the general NMR signal from the solvent (the particles contribute insignificantly to the signal) decays over timescales of seconds. For the discussions that follow, we don't need to worry about whether we use spin-lattice T_1 or spin-spin T_2 times, though in practice the default is the faster T_2 but with a respectable, say $5T_1$, interval between pulses.

For those unfamiliar with NMR, it's a way of exciting hydrogen atoms in molecules with radio-frequency pulses and measuring the signal emitted by those atoms. If the atoms are equivalent to a bell, the pulses are there to strike the bell and the machine measures how the sound of the bell fades over time. The time to fade is called a relaxation time. A bad bell has a fast relaxation time, the sound fades quickly. A good bell carries on ringing for a long time, so has a slow relaxation time.

The key fact is that solvent molecules that are free to move, relax the signal slowly, those near a particle are more constrained (the particle, within the NMR measurement timescale, is considered an immovable solid) and relax faster.

Confusingly, what is used in the analyses is the relaxation *rate*. Solvent near the particle, and which interacts strongly with the particle, has a shorter relaxation *time* but a higher relaxation *rate*. We judge the happiness of the particle in the solvent from the relative relaxation rate which is the ratio of rate with and without particles. A large relative rate means strong solvent/particle interaction.

The technique then consists of knowing the relaxation rate of, say, 15 solvents, then re-measuring the rates with an added few % of particle. Good solvents will have a large relative rate (say 8x) and poor ones will have a small relative rate. How much is "a few %"? As the effect is a purely surface-related one, you might need just 1% for a 50nm nanoparticle and 10% for a 500nm particle. At the start you probably have a good idea of at least one good solvent, so you would test with 1, 2, 5% to see how little is needed to get a large-enough relative rate for the data to be statistically meaningful.

What happens with a really bad solvent? It visually clumps so badly you don't have to bother to do the measurement - the relative rate will be "low".

3.10 NMR and solubility theory

For some decades, the combined colloid partnership at the universities of Bristol in the UK and Wageningen in the Netherlands has used NMR relaxation for some classic analyses of particle-polymer interactions. This has been combined with the sophisticated Scheutjens-Fleer model of polymer chains on particle surfaces to allow us to get a good feel for what happens to dispersant polymers on surfaces. This will be discussed in the next chapter.

A typical experiment starts with the significant (say 8x) relative relaxation rate of the given solvent with the raw particle. Then polymer is added bit by bit and the relative rate decreases to 1, i.e. there is no measurable difference. This gives us the (well, it turns out to be "an") absorption isotherm - a graph of how much polymer needs to be added to cover any given fraction of the surface of the particle, up to "full" coverage - with the scare quotes there for a good reason.

For any particle-dispersant pair we should routinely know these absorption isotherms, but most of the time we don't. We *should* know them because we can distinguish dispersants that lock on quickly and effectively at low concentrations (desirable) from those that need a high concentration to fully cover the particle (undesirable). In general we *don't* know these isotherms because they are hard to obtain by normal means. For example, you can stir up some particle with solvent and a known concentration of dispersant, then centrifuge the particles and measure the dispersant concentration remaining in the supernatant. The problem is that it is hard to measure the concentration. For example, most dispersants don't have strong UV signals, and most solvents (other than water) have some UV absorption, so it isn't easy to do measurements via UV.

At first sight, therefore, it seems that the NMR method is the answer for relatively easy measurement of absorption isotherms. To a certain extent it is. But it's not the full answer. Remember that NMR relaxation is fast near an (effectively) stationary surface. Solvent near that part of the polymer chain that is stuck strongly to the particle surface is still near a stationary surface, though probably somewhat different, so the relaxation rate will change somewhat. Any of the polymer that is in loops or free-hanging tails is, as far as NMR relaxation concerned, fully mobile, so the technique cannot see it at all.

This means that the absorption isotherm is measuring only the initial coverage of the particle, that is relatively sharp and at low concentrations. The more flexible parts, that build up more slowly, are invisible to the technique. The classic centrifuge-based isotherm shows a slower build-up and extends to higher concentrations before the plateau of full coverage.

Arguably, the NMR-measured isotherm is the one of most interest to us because the subsequent slow build-up of polymer is of less significance in terms of desirable outcomes such as particle stability.

But the difference alerts us to a limitation of the HSP measurement technique. We were agnostic about whether a particle was happy in a solvent because of solvent-particle or solvent-dispersant interactions, we just wanted it to be happy. The problem is that if we take a particle covered with a good dispersant and measure the change in relaxation rate, we will see zero change and score the solvent as "bad". The sedimentation technique is *not* fooled, the dispersed particles will sediment more slowly.

Much though we would like a simple, "does it all" technique, the NMR method is clearly flawed when the particles have a good coverage of a flexible dispersant. Which raises the question of how much of the dispersant polymer is flexible for a given coverage. This is where things get really interesting and require the chapter that shortly follows.

3.11 FTIR

Another way to look at particle-solvent interactions is via FTIR. As was found with the silicon/boron nanoparticles mentioned previously, with NMP a strong interaction (in this case a donor carbonyl with an acceptor boron) leads to a shift in one or more of the IR bands. We can readily imagine that this could be a general-purpose way to distinguish good (strong IR band shift) from bad (small shift) solvents. But not all solvents have IR bands with clear strong signals that are easy to follow and there is no obvious way to say whether a 20 cm^{-1} shift in one band is more or less significant than a 40 cm^{-1} shift in another.

As far as I can tell, these limitations on FTIR mean that it is not a widely adopted technique for determining particle-solvent interactions, but maybe that will change with time.

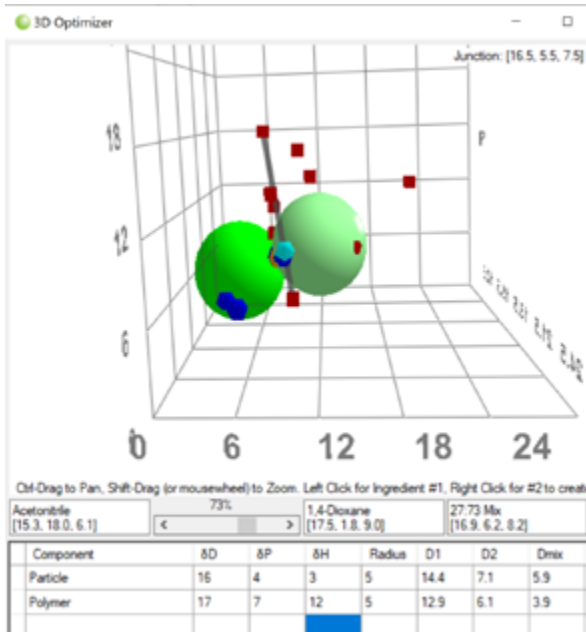
3.12 So what?

Via whatever means that works for you, we now have the HSP of a particle. What can we do with it?

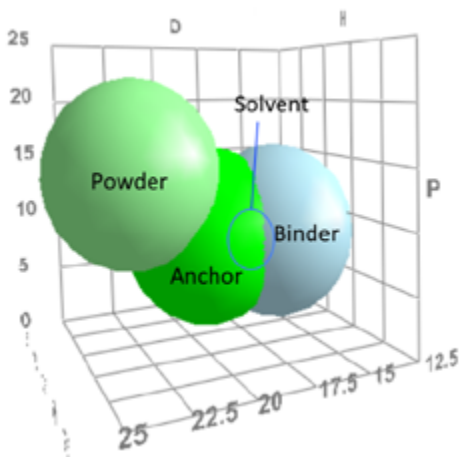
We already know that we can find an optimum solvent blend, perhaps using one or two solvents that individually are not so compatible with the particle. But why do we need such compatibility? For two reasons. First, because the rheology of a particle in a good solvent is likely to allow more formulation options, for example, a higher volume fraction of particles. The reasons for this are discussed in detail in the Rheology chapter. Second, because it is likely to be much easier to obtain a good, stable dispersion of particles within a polymer matrix if the HSP Distance is small. Given that many polymer dispersions are created via a solvent blend, the rheology and polymer arguments come together to make it easier to formulate satisfactorily.

An example is one I know first-hand from the world of ceramics formulation. Specific parts of the ceramic industry tend to end up in one formulation domain with a solvent and polymer combination that allow a high packing fraction of the ceramic particles along with the desired minimum of polymer that keeps the green ceramic together till it is fired. The problem is that any change that is made to that system in order to improve it, for example to swap to a greener solvent or to a different polymer with a different set of desirable properties is likely to produce an intractable mess. The formulation is within a local minimum in solvent/particle/polymer space and any random walk away from that minimum is likely to end up with a worse performance.

If, however, we know the HSP of the ceramic and of the old and new polymers we can check if the ceramic and the new polymer are compatible. If they aren't then it isn't worth trying to formulate in that space. If they are, then an optimum solvent would be one that is inside the spheres of both components or, at the very least, at a point that is close to being equidistant from the centres of the spheres. If a "scientific" solvent is in that spot (i.e. one that can be used for testing, even though it would not be practical in production), a quick test can confirm that this is the right part of solvent space. Then a rational blend of more suitable solvents can be devised to meet the solubility criteria as well as others such as cost, odour, safety.



It is difficult to visualize all the components in the Hansen 3D space. A tool like the 3D Optimizer in HSPiP makes it easier. Here we see the spheres of the particle and the polymer, we know that the Junction (i.e. the equidistant point) is at [16.5, 5.5, 7.5] and although close examination shows two solvents that might be OK (they are in blue), if you happened to want a blend of acetonitrile and dioxane, then 73% dioxane seems to provide the optimum possible from those two. Other, more friendly, solvents could be added to the available list to allow a better pair to be chosen.



A real-world application of this technique comes from Toyota¹⁷ who needed (for reasons unknown to me) a tantalum carbide ceramic coating. They had a complex set of solvent, polymer, particle interactions and via HSP they could clearly see how to navigate towards a smart optimum. Trying too hard in one way, dragged the stabilizing anchor off the particle. Trying too hard in another way meant that the binder was insufficiently soluble, but too hard in another and the particle was not sufficiently compatible. They even had to

manage a solvent blend to meet specific requirements for coating and drying the formulation. The visible differences in the final coated and sintered ceramic between good, medium and poor choices in HSP space were obvious.

Other examples for the need to match the HSP of particles and polymers come from the need to add particles to polymers for, say, increased hardness

17 Daisuke Nakamura, Keisuke Shigetoh, Akitoshi Suzumura, *Tantalum carbide coating via wet powder process: From slurry design to practical process tests*, Journal of the European Ceramic Society, 37, 2017, 1175–1185

or toughness. At one time it was found (as it happens, by Toyota) that adding some nanoclays to nylon polymers resulted in astounding improvements in performance. A bandwagon was created and countless PhDs threw nanoclays and other nanoparticles into polymers expecting amazing improvements in properties. The results were overwhelmingly disappointing. The reason was simple - they had given no thought to HSP matching. A specific example comes from a book chapter I wrote on PLA, poly(lactic acid). PLA is insoluble in water and is, therefore, hydrophobic. The popular nanoclays at the time were also hydrophobic (they contained lots of long alkyl chains). When the clays were mixed (with difficulty) into PLA, the results were invariably poor. The reason is that although PLA isn't soluble in water, merely labelling it as "hydrophobic" doesn't mean that it is compatible with, essentially, hexadecane (in which it's equally insoluble). Had they used the HSP values of PLA [18.7, 7.7, 7.0] they would have realised that to be compatible, the nanoclay required a surface with plenty of δD , δP and δH . They could have worked this out with 5 minutes' thought but instead they spent a 3-year PhD on a hopeless endeavour.

Some years later I heard an excellent talk on adding nanoclays to PET polymer. They had worked out which of the commercially available nanoclays would be the most HSP-compatible with PET and simple lab experiments confirmed the desired performance improvements. Unfortunately, that specific clay reacted with the PET during commercial processing so could not be used. It was obvious how to get a clay that would be the best of both worlds, but unfortunately it was not commercially available so the project ended. Yes, the project was a failure, but it was a fast failure, with eyes wide open. What is sad to see are the many slow failures where, from simple HSP, there is no hope of success but by not knowing this, teams carried on working to no avail.

My own experience with nanoparticles added to UV cure coatings to make them super-hard was humbling. We had worked out that good HSP compatibility was a requirement and we made excellent formulations using well-chosen dispersants, yet with disappointing hardness. We (along with many others, this was some time ago) had not realised that adding some functionality to the tail of the dispersant was necessary to get the real benefits of adding the nanoparticles. Getting the right dispersant that reacted at one end with the nanoparticle, had the right HSP to be compatible with the formulation, and with an acrylate functionality on the end of the tail was a challenge. We were lucky at the time to have good access to the technical experts inside one of the big dispersant suppliers so this tricky task turned out to be do-able. As discussed later, most of us now have little chance for such discussions because the top internal scientists are shielded from the outside world.

The message is that with the right HSP knowledge it is relatively easy to formulate in the right sort of space. You might still fail, because formulation life is hard, but failing intelligently against a reasonable hypothesis is superior to failing with no knowledge of why.

3.12.1 HSP and Hamaker A_{12}

If you insist on using DLVO then you need to know the Hamaker constant A_{12} . My general advice is to assume that it's $1E^{-20}J$. But what we really want to do is reduce it to near zero. If we assume that the particle is non-polar, for example one of the polystyrene particles much used in academic research, then as briefly mentioned, Lifshitz theory tells us that matching the refractive indices of the particle and solvent gives a zero value. Next we can note that δD is strongly correlated to refractive index. So a matched δD of particle and solvent should give minimal attraction. The standard δD for polystyrene is 18.5, and for cis-decalin it's 17.8 and for tetrahydronaphthalene it's 19.6. By simple arithmetic, a 60:40 blend is a good match, and the sometimes vague academic literature seems to indicate a 75:25 blend as being optimal, so we're not too far away with a simple calculation.

But what about more realistic particles with plenty of δP and δH ? My view, so far neither validated nor invalidated in the literature, is that a good overall HSP match is equivalent to a small Hamaker constant. If anyone can provide support for, or evidence against, this idea I will update this section accordingly.

3.13 Going to the next level

As we have seen, DLVO isn't all that useful to us in practice. HSP, although pragmatically very useful and a justification for using the word "solubility", clearly don't provide a solid basis for exploring the complex behaviours found in real-world formulations with, for example, complex comb dispersants.

When I started this book, I found that gap in understanding to be embarrassing, and I had planned simply to say "Don't blame me - the situation is what it is".

Happily, as so often happens during my adventures in writing books, while I was trying hard to understand something new to me (the NMR work) I came across references to SF/SCF theory of which I had never heard. When I tried to understand it I realised (a) that it was wonderful and (b) that I would never be smart enough to understand it properly and (c) without hands-on apps to explore the ideas, most of us would never be able to grasp its power.

Happily, again, Prof Terence Cosgrove, from whom I'd learned so much about NMR and the basics of SF/SCF, put me in touch with Prof Frans Leermakers and the result is the next chapter. I have thanked them in private - here I thank them in public.

3.14 Particle solubility? It's your choice

The habit of thought that particles are "dispersed" and therefore need "dispersion" science is deeply engrained. Saying that a particle is "soluble" is something some scientists will refuse to do. That's their (your) choice.

With "dispersion" thinking you have DLVO that, even on a good day, is acknowledged to have many flaws. Add to that surface energy, discussed in the Surface chapter, and that's the limit of the formulation tools available other than some modest choices around depletion flocculation. Of course, in water with charged particles and charged stabilizing polymers, things work well in DLVO terms. As long as you don't add too much salt or counter-ions that wreck the charged surface, or add polymers that themselves interact with the ions, life is fine. But if aqueous systems are your main concern, this book is not of so much relevance (though I do have a short chapter on the topic).

If, on the other hand, you choose to accept that solubility ideas work, then a rich domain of formulation opportunities opens up. For example, you can:

- Objectively measure the differences between the "same" materials such as carbon blacks, silicas, or titanium dioxides. Not just "hydrophilic/hydrophobic" but the full spectrum of δD , δP and δH values.
- Rationally work out which nanoparticle formulations will be more or less compatible with polymers for optimising polymer properties.
- Tune dispersants, solvents and solvent blends for optimal performance during processing and for optimal compatibility with the final matrix (e.g. polymers) once processed.
- Efficiently use (if available) the extra functionality on the end of the tail to react into a reactive solvent (UV curing if the tail functionality is an acrylate) or into a proto-polymer (epoxy or urethane if the tail functionality is an amine or alcohol), making the particles part of the overall network for maximum performance such as hardness.
- Understand the trade-offs between solvency for ingredients in your formulation and solvency for particle stability.
- Understand what happens to a particle formulation as the different components of a solvent blend evaporate at different speeds.

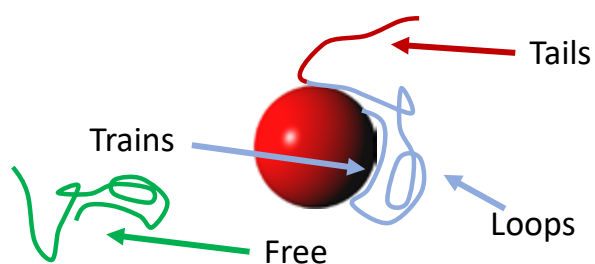
Given the profound incoherence of any attempt to define "dispersion", given the limitations of dispersion theory, it seems to me that the rational way forward is via solubility - especially when SF/SCF theory, linked to HSP offers so much.

4 The practical power of SF/SCF theory

4.1 SF/SCF Theory

The lattice theory of polymers in solvents is rather easy to understand¹⁸. To see what a polymer chain is doing, you just look at whether a lattice site is filled by a solvent or a monomer link. By counting up the relative numbers of such sites across the lattice, you get an idea of whether the polymer is happy or not in the solvent and by how much or little it extends into the solvent. The distribution depends on the relative monomer-monomer, monomer-solvent and solvent-solvent attractions, expressed via the χ parameter.

Scheutjens-Fleer¹⁹ (SF) theory or, to use their name for it, Self-Consistent Field (SCF) theory does exactly the same thing for polymers (dispersants) on particles, but this time with the specific constraints that the polymer and solvent compete for sites on the particle surface and that the surface is a solid wall for the polymer. To do the calculations, we need two χ parameters, the first is the standard one between polymer and solvent and the second, called χ_s , captures the relative attractions of the polymer and solvent to the particle surface.



From the calculations come the distribution of polymer chains which are sitting as "trains"²⁰ flat on the surface, "loops" of polymer out in the solvent but tied to the surface via trains, and "tails", the free ends of the polymer, plus polymer free in solution.

The theory is general and can handle, for example, block copolymers (where one block prefers to be in contact with the surface), comb polymers (a linear chain that likes the particle and short comb chains that like the solvent and create a brush effect) or polymers attached via chemical or strong physical bonds to the surface, real brush systems. We can also have two separate polymers, one which likes the surface and one which prefers the solvent.

Although it is, unfortunately, full of complex formulae that are far beyond my own grasp, the core idea behind the calculations is still simple: it's a Self-Consistent Field theory. So some people call it SF theory after the original authors, others

¹⁸ The app in the first chapter can be used as a reminder.

¹⁹ For English speakers, the closest approximation I can find for pronouncing their Dutch names are Schir-tyuns (or Schur-tyens or, to emphasise the hard Dutch Sch, Skirt-yens) and Fleer rhyming with beer. The "r" in Schir/Schur shouldn't really be there, but to my ears it's clearly audible in the sound file kindly provided by a Dutch linguist.

²⁰ Apparently Scheutjens liked the idea that polymer chains going along the surface were like trains running along a track. Like many people, I find the name unhelpful and easy to confuse with "tails", but that's the standard nomenclature so we learn to live with it.

call it SCF theory to reflect its generality and the work of many colleagues of Scheutjens and Fleer who helped develop its full potential.

SCF says that you start with some reasonable guess as to where the different parts of the polymer, solvent and surface are within a lattice and calculate the energy²¹ of the entire system. The energy depend on the sum of all the interaction potentials between neighbours. You then make a small alteration to the system, for example, moving one part of a polymer chain to the next lattice position, perhaps swapping it for a solvent molecule. If the energy gets higher, that's the wrong direction, if lower then the change is in the right sort of direction. You carry on making such changes so there is a self-consistency about all the interactions. In SCF language, the aim is to optimize the volume fractions of solid surface, trains, loops, tails and solvent to the correct potential energies, and to optimize the potential energies to the correct volume fractions.

The simple phrase "moving one part of a polymer chain" is not so simple in practice. Polymer chains have their own rules, well-known to those who have been introduced to polymer physics. Locally the segments tend to want to go in straight lines, but because the segments end up taking random walks (maybe constrained by bond angles and local rigidity) chains tend to loop back on themselves. The better the solvent, the happier the chains are to be more extended, the worse the solvent the more likely they are to coiled up. The effects are a mixture of enthalpy and entropy. Once one part of a polymer chain is on the surface of a particle the tendency of the random walk to loop back onto the surface or stick out into the solvent again depends on subtle balances of enthalpy and entropy. Happily for us, all these effects work themselves out naturally via the SCF calculations.

In the past, only SCF experts had the ability to access the power of the methodology. The core team behind SCF gradually built up many numerical tricks to increase the power and reliability of the technique, creating a complex set of codes that were understood by very few and used by those with access to typical academic computational infrastructures. It was simply impossible for the rest of us to be able to explore these ideas for ourselves. Although some of the simpler calculations can be put into apps, the code is far too complex for that to become possible for the range of calculations that interest us.

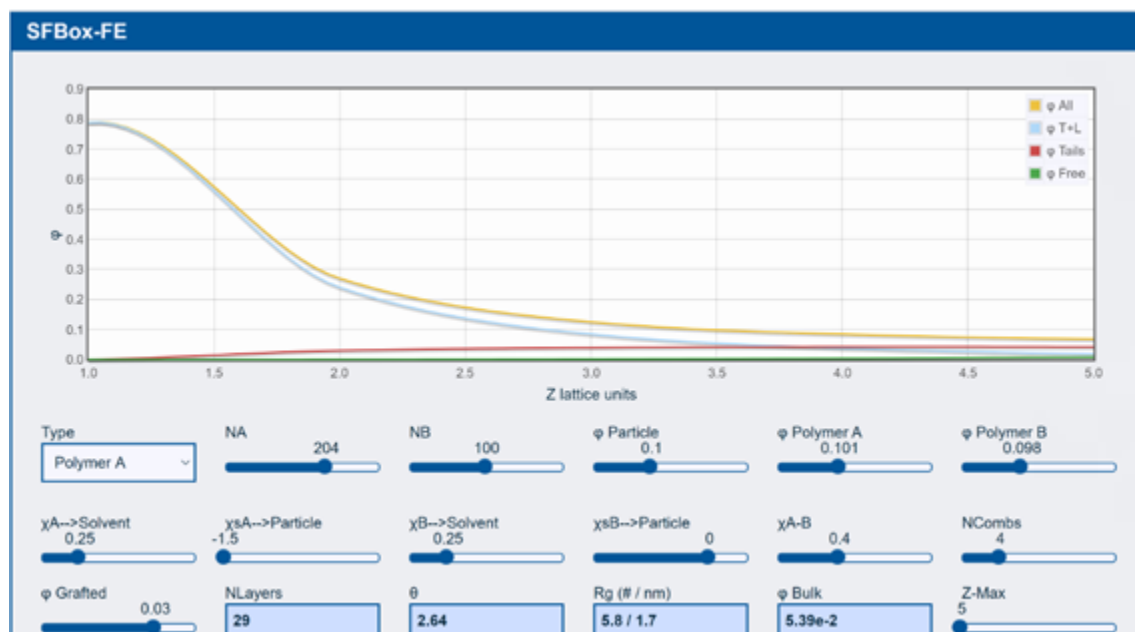
Fortunately, Prof Frans Leermakers from Wageningen U had created a "black box" version of the code so that academics could access it a little more easily. The code was called SFBox and there was a front end that made it possible to feed inputs and extract outputs in a more friendly manner. However, this was still usable only by those with a lot of expertise. With the considerable help from Prof Leermakers, I wrote two versions of SFBox Front End. The first, full-power

²¹ There are various energies such as grand potential or Gibbs (or Helmholtz) free energy used in these calculations. For our purposes "energy" is good enough.

version is included as part of the HSPiP software package where it fits in with the ability to measure the HSP of polymers and particles and allows calculation of the relevant χ parameters. Even this does not tap the full power of SFBox (which is general for many other types of problems), but it allows formulators to explore most of the key aspects of polymers and particles. Those who want to use SFBox itself can ignore the front end and provide their own inputs and interpret their own outputs. The SFBox manual is included in the downloaded package.

Because an average laptop today has more power than a supercomputer from the days of SF, many of the calculations can be done "live" as you slide a slider or change an option. You automatically get a graph of all the relevant calculated values. If things start to get too slow, you can switch to the mode where you need to click the Calculate button after setting your input parameters.

The second implementation is an app version which, of necessity lacks the full power but shows the basics of particle/polymer interactions. It works even on a phone because the SFBox itself is sitting on a server and the app is merely a front end.



App 4-1 <https://www.stevenabbott.co.uk/practical-solubility/SF-Theory.php>

For consistency, and because I can access all the bells and whistles, the screen shots that follow are from the HSPiP version, but you will find most of the key points when you run the app - until we come to particle-particle behaviour which is beyond the scope of the app.

Using this powerful tool we can bypass DLVO, and bypass the hand waving that inevitably surrounds issues such as steric stabilization and its opposites, bridging and depletion flocculation. For most of what follows we will make use

of χ and χ_s as if we knew what they are. Towards the end we will discuss how to get good-enough estimates via HSP.

Briefly, χ goes from 0 (good solvent) to 0.5 (borderline) to >0.5 (poor solvent). Confusingly, Silberberg who came up with χ_s decided that when the polymer was compatible with the particle and the solvent was incompatible, it would have a value of, say, -1.5 and when polymer and solvent were equally happy on the particle surface the value is 0. If the particle prefers the solvent to the polymer then, presumably, $\chi_s > 0$ but that doesn't seem to be a common situation.



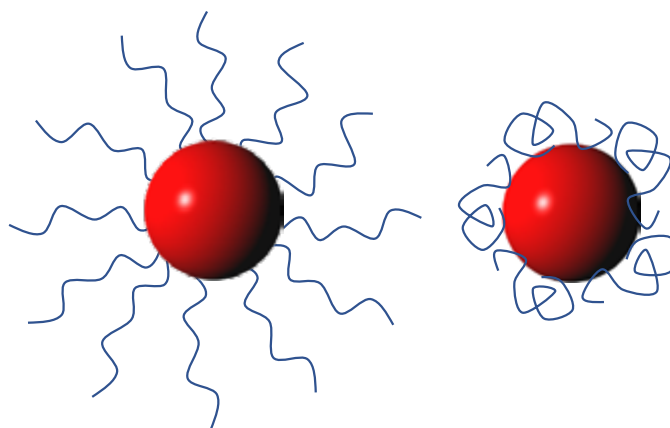
There's one important thing to note. For simplicity and convenience, all calculations are done as if the particle is planar, not spherical. This seems shocking till you look at scale at a typical 40nm diameter nanoparticle and a 0.4nm diameter solvent. The errors from assuming that the particle is planar are rather small.

We will take things one step at a time, using the SF calculations to guide us. What follows is about the science, not about running the program. The program itself has a Help button. If you click that (or you read the Help file that's included with the package) then you get all the necessary details about the many available options.

4.2 Where is the polymer?

Every time I've drawn polymer chains sitting on surfaces I've shown them as wavy lines sticking out into the solvent, as in the left-hand image.

This picture is entirely wrong! Under most circumstances, the polymer is sitting along the particle surface as trains and coiled up into loops, as in the right-hand image. The tails sticking out are a minority. We can



easily see this for ourselves with SF theory. We start with a polymer chain that is 200 lattice units long. This is our first problem - what does that mean? Thinking in terms of MWt, if we take a typical monomer unit to be somewhere between polyethylene (28) and PMMA (100), so calling it 50, the 200 unit polymer is 10,000 MWt - not large but not tiny. Alternatively, just multiply the number by 0.3 to get nm - so this is a chain 60nm long, which sounds quite impressive if you think in terms of the left-hand (wrong) diagram. The radius of gyration, R_g , of the coiled polymer would be something like $\sqrt{(200/6)} \sim 6$ lattice units or 2nm. This R_g calculation is an alert that this polymer is likely to be sticking out by ~ 6 lattice units (as in the right-hand diagram), not the 200 we might naively expect from our default picture of polymer chains on surfaces. Let's see what SFBox says:

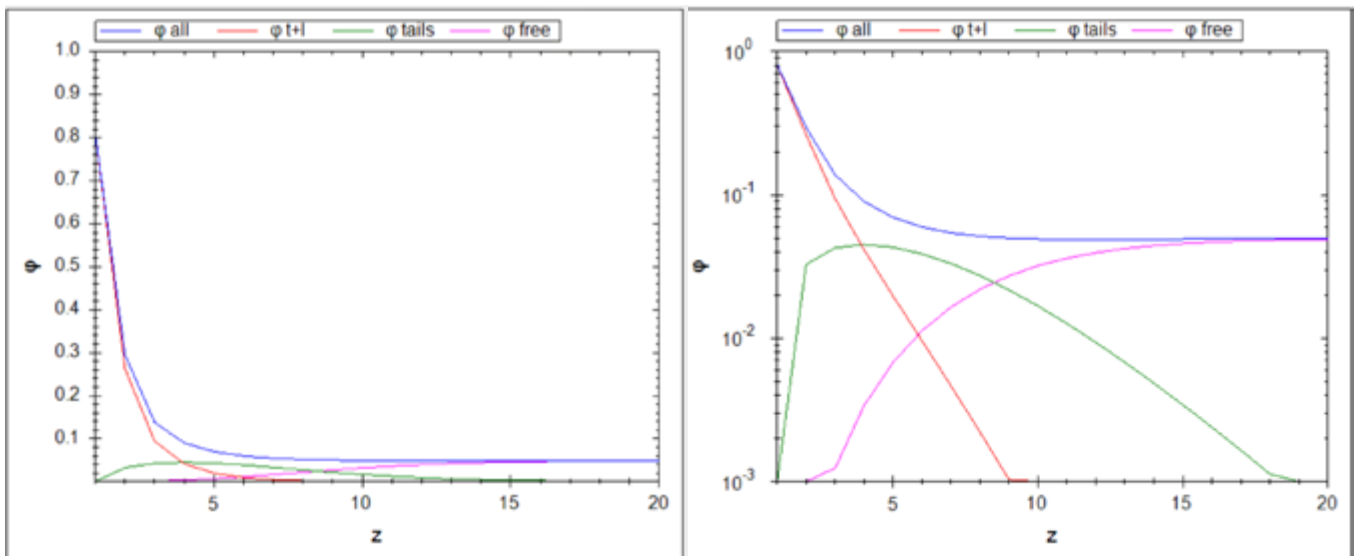


Figure 4-1 Example: Poly200. Linear plot on left, Log plot (same data) on the right.

If you are using the app, you can work out most settings, if you are using HSPiP then, as with all the examples, you can load them as .sfb files from the software so I don't have to tell you all the details. Here we happen to have 10% particle and 10% polymer (happy in the solvent with $\chi=0.25$) and the net result is a background polymer concentration of 5% (which you can read from the graph but is also provided as an output on the main form) which is the long blue+magenta line to the right of the plot. Blue is Total, Magenta is Free so obviously Blue=Magenta far enough away from the particle. The red line is the combination of trains+loops (it's unfortunate that tails also start with "t", so "t+l" is ambiguous), i.e. the bulk of the attached polymer. You can see that this reaches near-zero around the R_g value of 6 (which is also calculated for you on the main form). Well, that's what you read from the left-hand plot which is at linear scale. I strongly recommend that your default view is linear. When you look at the log plot to the right, everything looks much more exciting. Although you can just about see the small amount of (green) tails in the linear plot, they look much more interesting in the log plot. But here's the thing. For academics, these log plots are truly interesting because you get to see lots of subtle details that are great for academic papers. But for formulators, if you need a log plot to see it, it probably doesn't matter. In the plot shown, the plot is limited to a volume fraction of 0.001, i.e. 0.1%. In reality, anything below 1% isn't going to be of much interest, which is why the linear plot is more relevant.

As a simple exercise, double N , the number of segments, from 200 to 400. Generally, doubling a polymer's MWt would be expected to be fairly dramatic. A core lesson from SF theory is that the effects are modest. It's because of the square root in the calculation of R_g . To get a doubling of the amounts sticking out, you need to quadruple the MWt. The modest result is shown in the next figure.

If you are going to use a single polymer to stabilize your particles (and we'll soon see that this is a *bad* choice), and if you can't get so much benefit from changing the Mwt, what can you do? Simple! Just increase the χ parameter to, say, 0.6:

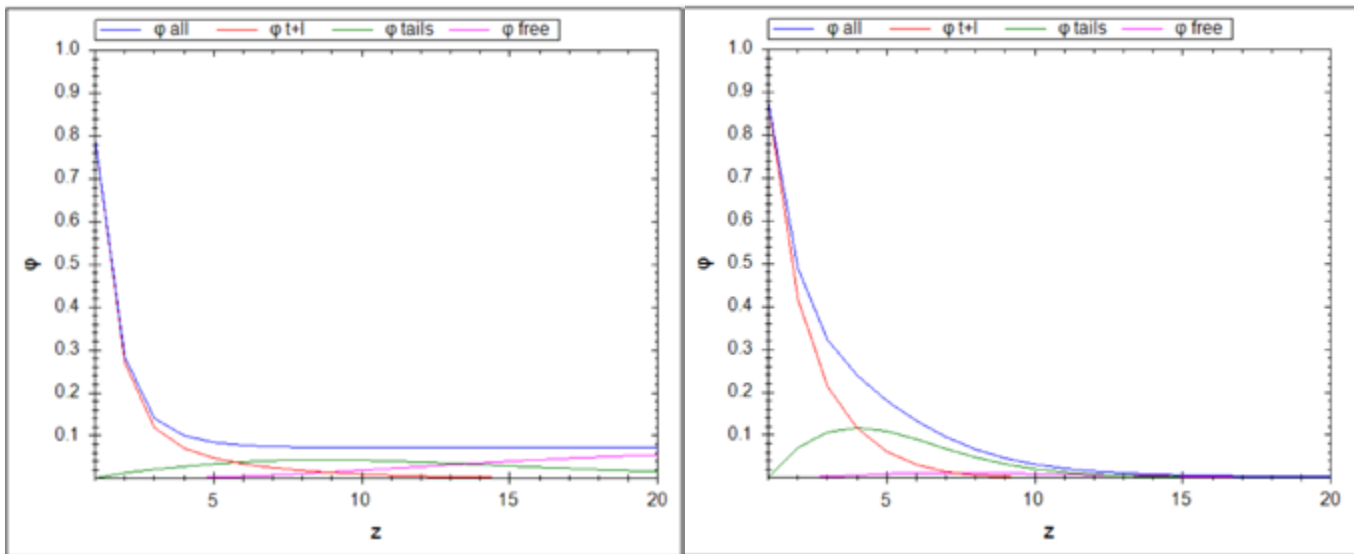


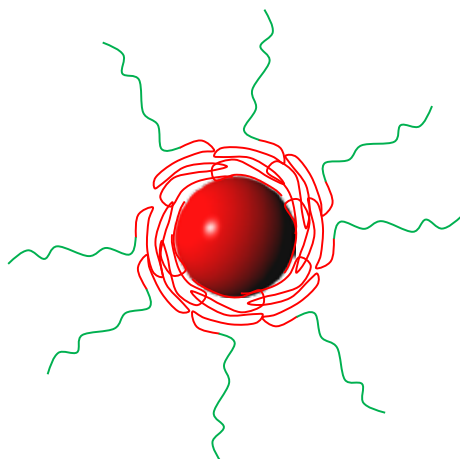
Figure 4-2 Example: Poly200. Left, 4x increase in segments makes little difference. Right, increasing χ makes a big difference - but isn't viable.

The effect is dramatic - with such a large χ the polymer doesn't want to be in the solvent so clumps onto the particle. The amount of free polymer has clearly slumped.

But there are two problems with this approach. First, you can't dissolve 10% of your polymer in such a poor solvents. Second, even if you could, the polymer, even from simple DLVO, will not provide any stabilization. We can check this later with SF.

With a single polymer you are caught between two problems. If the solvent is too good, then the polymer doesn't want to stay on the particle, if it's too poor, you can't get the polymer into solution and it doesn't provide protection.

4.3 Di-blocks, Combs, Brushes and A+B



If you load the DiBlock200 example (in the app have a 100A and $\chi = 0.5$, $\chi_s = -1.5$ and 100B with $\chi = 0.25$, $\chi_s = 0$) you will see what a difference a di-block makes. The red line is the A block which has a relatively high χ parameter with a large negative χ_s parameter meaning that it likes to lock onto the particles. The B block has the same 0.25 χ parameter as the previous single polymer, but a 0 χ_s parameter, meaning no special liking for the particle itself. The red line shows how the 100-unit A block is

behaving like the single polymer in the previous example. The green line shows how the 100-unit B block is sticking out into the solvent. The total number of segments, 200, is the same as for the single polymer, but the effect is dramatically different. If you can tune the solubility properties of particle, block A and block B with this happy balance, you have a good dispersant.

Using exactly the same A-B properties, we can create a comb polymer, typical of modern hyperdispersants. If you load Comb300 you will find that we take the same 100 segment A chain and add 4 B comb side-chains, each 50 segments long. The figure allows us to compare the results:

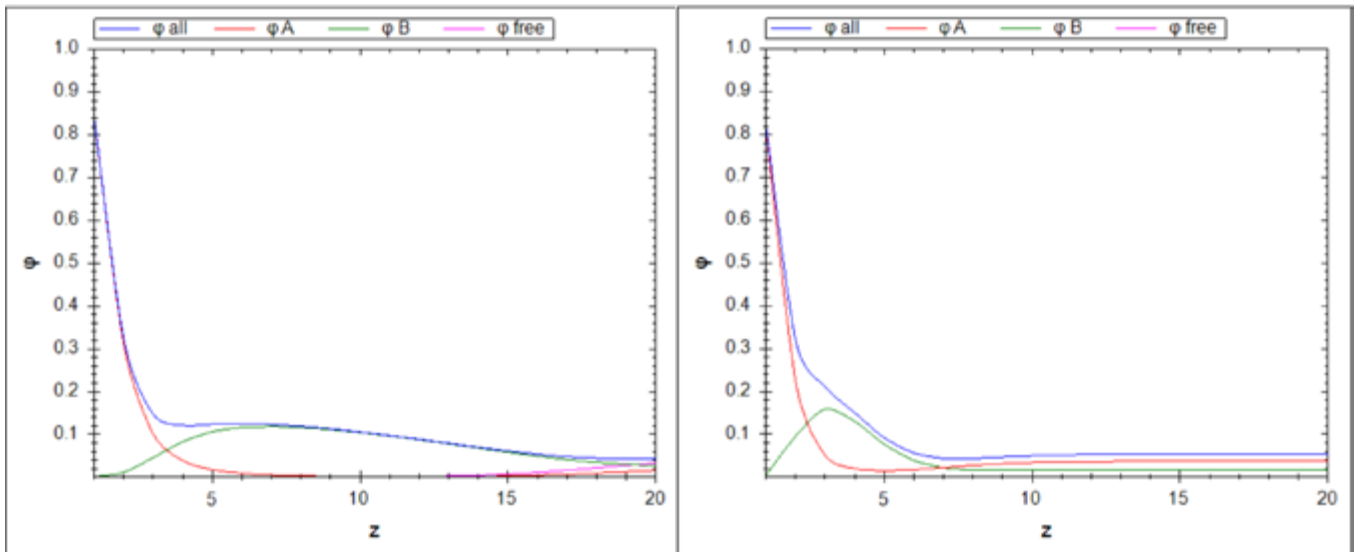
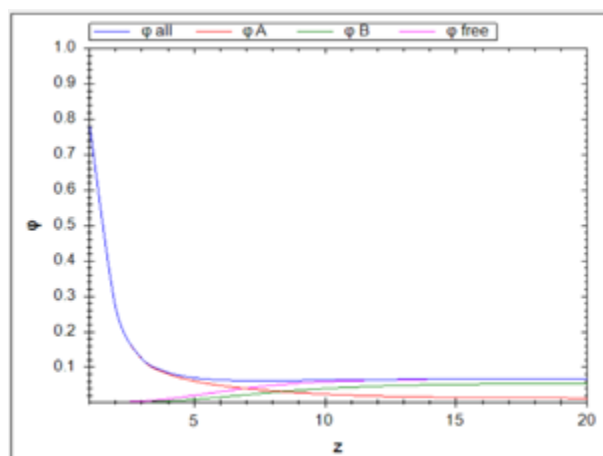


Figure 4-3 Example: DiBlock200 has the B block extending far into the solvent. Comb300, with 50% more B segments does not extend so far, but has other advantages.

In terms of extending into the solvent, the di-block has a clear advantage. We will shortly see why the comb polymers prove to be so popular.

Even better than a comb, if you can create it, is a brush where one end of the polymer is chemically attached to the particle. As long as the χ parameter with respect to the solvent is less than, say, 0.4, a brush is really going to refuse to self-associate, this is because a brush is "all tails", and as we shall see, tails repel.

Finally in this section, is the rather odd combination of the A and B polymers, but as independent molecules. In terms of classic particle/polymer science this makes no sense. In terms of practical formulation it is quite commonplace - we add different polymers for independent reasons,



often with catastrophic consequences, not evident here. When you load the 2Poly200, you find that we have A and B each of 100 segments, and each at 5% concentration to give the same overall polymer concentration as the original Poly200. The results are fairly obvious, A is behaving fairly normally at the particle surface (though, perhaps surprisingly, B stretches A out somewhat) then B hardly features near the surface and makes up most of the background concentration.

Before we get on to the real implications of all this, we need a quick exploration of adsorption isotherms.

4.4 Adsorption isotherms

I thought it would be very cool to be able to calculate the adsorption isotherms of the different polymers, i.e. how quickly you cover the surface of the particle as you gradually increase the concentration. However, as is clear from the academic literature, these isotherms are very boring because for any polymer with a reasonable χ_s the answer is that you get saturation cover within volume fractions below 10^{-5} , i.e. far below anything that would interest a formulator.

To explore this, open Poly200. Now slide the ϕ Polymer slider to lower concentrations. The slider is log-based so you can go to very low concentrations. At first the amount of polymer sitting in the solvent, shown as ϕ Bulk is a sensible value because you've exceeded the amount of absorbing space on the polymer. Then suddenly, ϕ Bulk drops to 10^{-6} or lower, showing that the particle is absorbing all the polymer you have provided.

This behaviour is entirely expected by the theoreticians who then worry about whether the curve saturates at 10^{-6} or whether there is some slight sub-saturation up to 10^{-5} . For the rest of us we know that if we get the polymer/particle basics right, we don't have to worry about whether the particles will saturate - they will.

Where we *should* worry, and where the academics are less interested²², is when the compromises necessary in our formulations lead to a combination of χ and χ_s values where we don't approach saturation even at relatively high concentrations. My personal journey towards SF theory started with a real-world problem where gradually it became clear that the root cause was a dispersant that was in this borderline zone. Had I had the SFBox-FE we could have worked out the problem in 10 minutes. Instead it took most of a day before arriving at the correct conclusion.

²² In a typical publication they rightly focus on standard cases such as $\chi=0$ and $\chi=0.5$. It's not typically their job to worry about the messy world of formulation.

There is a separate worry for which we really *do* have to thank the academics. For the moment we are focussing on the *thermodynamics of monodisperse polymers*. The problems of the *kinetics of polydisperse polymers* are a topic for later.

4.5 Interparticle Interactions

The way authors write the sections of a book might make it seem as though we sat down one day with a beautiful logical plan which unfolded chapter by chapter, section by section. The real process is much more confused. When I started this book I'd never heard of SF theory. When I started to work with it, I had to treat it as a form of magic which required inputs I could hardly decipher and gave results of which I could make little sense. At one point Prof Leermakers had to give me a stern lecture on some elementary aspect of SF theory and when he realised, with a sigh, that I was too stupid to understand these basics he had to write a lengthy document (which I append to this chapter) so I could make the SFBox-FE program work in the most usable mode ("canonical") rather than the mode most convenient for me to understand and program ("grand canonical").

I had one big dream that seemed to me to be impossible. Once my grasp of the basics was acceptable I plucked up courage to state my impossible dream. "Oh, that's easy, you just do a canonical calculation, then use the ϕ Bulk from that to do a series of grand canonical calculations with stepwise diminishing NLayers - I'll send you the script."

The next day, armed with the extra 4 lines of script, I was able to realise my dream and could calculate particle-particle interactions.

This is a very big deal. We know that DLVO can calculate steric stabilization if we happen to know (which we rarely do) how thick the stabilization layer is. We also know that things can easily go wrong if we start adding more polymer, for various reasons. At lower concentrations of polymer we can get bridging flocculation and/or depletion flocculation and at higher concentrations we might (or might not) get stabilization. As we saw in the early chapters, the theories available to us were very crude. For decades we have had Asakura and Oozawa's ideas for depletion flocculation, but these applied only to naked particles and polymers with zero interaction with the particles. An attempt by Vincent (the "pragmatic method") to handle particles with polymer brushes and polymers that interacted more or less well with those fuzzy particles was of modest success but was (as far as I can tell) never used by formulators. So for decades we have been formulating with crude ideas that most of us cannot implement.

With the Interparticle button on SFBox-FE we can now intelligently explore the whole gamut of polymers, di-blocks, combs and, even, 2 separate polymers.

With the χ parameters as our guide to the various interactions we can get a good feel for what is interacting with what, with what effect. More specifically, we can do the calculations in normal mode to see which bits of polymers are doing what around the particles, then click the Interparticle button to see how much or little protection (in terms of energy barriers) those polymers are providing. There is no Live mode with Interparticle as the calculations typically take a few seconds to run - not surprising given the complexity of the computations.

Let's do this with our simple Poly200 setup. On the left we see the (by now) familiar distribution of trains, loops, tails and bulk polymer, with $\chi=0.25$. On the right we see the effect, in terms of free energy, ΔG^{23} , as the particles approach each other. The barrier is very large, not surprising for this large excess of polymer which is strongly attracted to the particles. Included in the image are the results for when χ gets close to 0.5. As is found in practice (and seen in the standard DLVO app where things flip dramatically when χ exceeds 0.5), the barrier becomes much smaller at, say 0.49, then disappears for relatively small changes in solvent quality, taking χ over (in this case) 0.52.

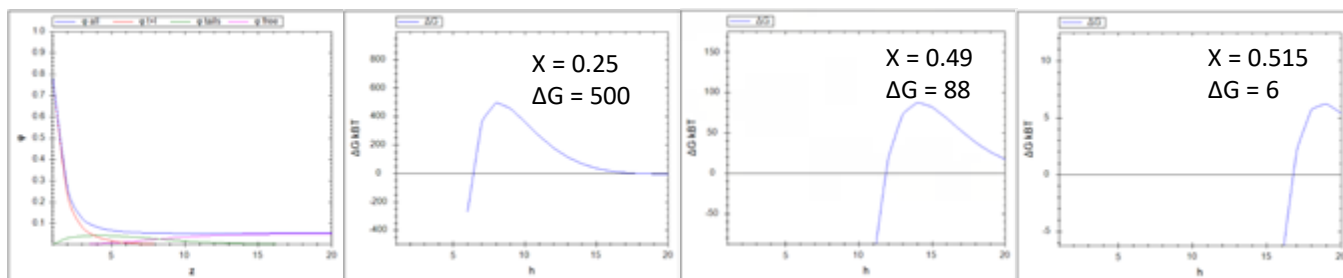


Figure 4-4 Example: Poly200 as normal, then Interparticle with increasing χ and decreasing barrier. With $\chi > 0.52$ there is no barrier.

The HSP view

In HSP terms, this means that as you gradually increase the % of "bad" solvent in a blend, the particles seem to remain OK but then suddenly crash out with a small extra % of that solvent. This is especially damaging for coatings from mixed solvents if the good solvent evaporates faster - the system suddenly transitions from happy to unhappy and the consequences are seldom good.

Why do we use di-blocks, combs and brushes for dispersing particles? The first reason, as discussed previously, is that we can tune A to be compatible with the particle and B to be compatible with the solvent, each doing what it does best. The second reason is that the barriers they produce are much larger. Naively taking the DiBlock200 and Comb300 files and clicking the Interparticle button creates barriers in the (unrealistic) 1000s of kT.

²³ It is debatable whether this should be, say, ΔA for Helmholtz free energy. But as the precise thermodynamics are not our concern, we'll stick with the familiar ΔG , in units of kT, as we have used for DLVO.

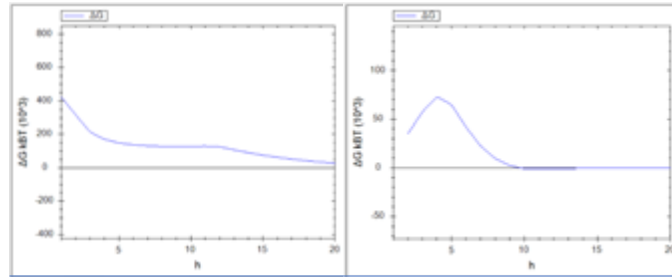


Figure 4-5 Example: DiBlock200 and Comb300 in Interparticle mode - producing enormous barriers.

Where things get really interesting, and relevant to many of our failures, is when we have two polymers - one attracted to the particle and one mostly in the solvent. This is where we can get stabilization (if we're lucky), depletion flocculation (if we're unlucky) or bridging flocculation if we're really unlucky.

What has historically been bothersome about such systems is that they are tricky to control in practice, yet we have few intellectual tools to think through what we are doing right or wrong - other than the broad generalizations about depletion and bridging, and rather uninspiring theory such as Asakura & Oozawa.

With SF theory we can start to think through things and gain deeper insights into which properties help or hinder stabilization. We are in for some surprises.

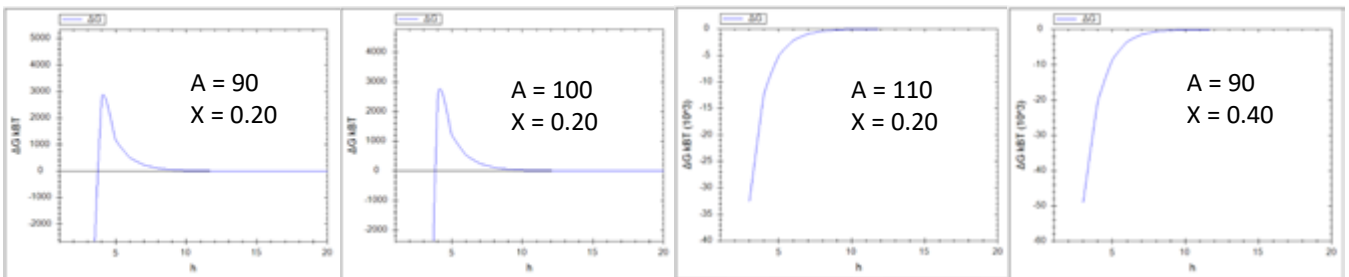


Figure 4-6 Example: 2Poly190. Short A chains are stable - but increasing them makes the system unstable, as does changing the χ parameter by not very much.

Before you play with the example, be aware that the numerics are very jumpy - you may have to fiddle with settings to get reliable results. That is an indication, as the results bear out, that these systems are themselves very jumpy - it doesn't take much to kick them from stability into instability.

The first surprise is that increasing the chain length of the polymer A on the particle makes things worse. Our instincts are that longer chains must be better for stability - there's more stuff sticking out to protect the particle. Our instincts are wrong in a profound way, discussed in the next section.

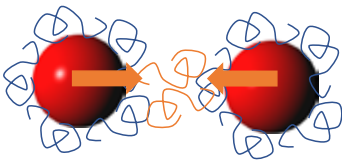
The second surprise is that modest changes to the χ parameter, which does nothing dramatic when you look at the polymer distribution, can upset the stability.

The third surprise, not shown here, is that the stability of the $\chi=0.4$ system increases (well, it becomes borderline) if you either decrease or increase ϕ_B .

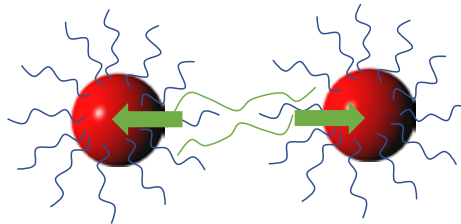
The fourth surprise is that decreasing the solvent compatibility of B (increasing χ) makes the system more stable, the reverse being true if χ is decreased.

No wonder that when we have two polymers in a particle system, we get surprises when we change some apparently independent parameter such as the χ of one of the components.

4.6 Tails repel



In the early days of SCF theory, Scheutjens proudly showed the results of his calculations to Nobel Prize winner Pierre de Gennes - who merely glanced at the results and pronounced that they were wrong.



When Scheutjens investigated his code, he found that he had made a programming error, with a wrong sign at one point. When he fixed the error, the results met de Gennes' approval.

The reason de Gennes merely had to glance at the results to know that they were wrong is that these systems obey symmetry rules²⁴. If you take one of the straightforward cases that academics like to investigate, the direction of the results - stability or instability - can be identified by symmetry. And the rule is:

Symmetrical systems attract, asymmetric systems repel.

If you want a stable particle system, make sure it's got plenty of asymmetry!

Translating this into polymer language:

Loops attract, tails repel.

When you push a bunch of loops together, the loops just see loops, so they see symmetry.

When you push tails together you start with no overlap, then more overlap then more overlap - this is not a symmetric system.

²⁴ If you trace just about anything in this field it tends to pass via de Gennes to another Nobel Prize winner Landau - in this case to Ginzburg-Landau theory. Although they could both indulge in super-sophisticated maths, they were generally happiest if their ideas came down to simple principles. Many academics hide their ignorance in complexification; these geniuses showed their wisdom in simplicity.

Asymmetric systems want to break their asymmetry, by splitting apart.

I have to admit to finding this all rather hard to understand. Even reading reviews by the experts doesn't leave me completely clear about the times that tails can be attractive and (when really pushed together) loops repel. This is even more difficult for me because the arguments seem to rest more on entropy than on enthalpy, and I always find entropic arguments hard to grasp.

There is one example where the entropic argument seems to be simple enough for me to understand and, maybe, captures the symmetry argument too. Imagine a loop attached to one particle. It comes close to the loop of a separate particle. Before that meeting it had a constraint - the solvent beyond the loop. Now that it meets, the loop can equally be on one particle as on the other, so it has a gain in theoretical freedom so a gain in entropy.

This is especially important for bridging flocculation. In the naive diagram that most of us draw, the particles have tails sticking out and the tails of the bridging polymer are nicely bridged into those tails. This is not how bridging flocculation works - it's loops interacting with loops.

We can equally imagine a tail that has plenty of freedom to move around, large entropy, not liking to be constrained by a surface. So "tails don't like surfaces" is a phrase I encounter in the expert literature.

If you have enough polymer getting in the way then whether it is a bridge or tail doesn't matter. Polymer physicists casually (and confusingly) refer to there being a "polymer melt" providing the repulsion.

But those aren't the sorts of conditions that are of great interest to formulators. Typically we are trying to use the least amount of dispersant and polymer to achieve our goals because generally those components are a nuisance in the final product. It is exactly in these low-polymer situations that the subtle balances of symmetry, entropy and enthalpy come into play.

Returning to situations with modest amounts of polymer, we can see why shorter polymers might (as in the example above) enhance stability - there is a larger fraction of stabilizing tails. But we can't push things too far - shorter polymers stick less well to the particles so we might not have enough polymer sticking out for the tails to repel.

4.7 Thermodynamics and kinetics

So far, we've had the luxury of monodisperse ideal polymers where we can tune the key properties in any way we wish. In real life we have messy polydisperse materials so things aren't straightforward.

If we give ourselves the luxury of infinitely long experiments then even our messy polydispersity simplifies itself with a straightforward rule:

Larger polymers stick better and will displace the smaller ones.

The converse is true for when we do things as quickly as possible:

Smaller polymers get to the particles faster.

That might not be so bad, except for the third rule:

Whichever polymers get to the surface first tend to stay there.

This set of conflicting rules makes life complex. If, for example, you want super-fast stabilization of particles as fresh surfaces emerge from milling, then small dispersants are ideal as they will quickly attach themselves. The problem is that they are not the stabilizers you need for long term formulations, so somehow you have to swap them over for the larger or more complex stabilizers. Given enough time this is possible, though the displaced polymers are now an unnecessary complication.

My view of problems like these is that the first step towards finding a solution is to be as clear and honest as possible about the problem. Solutions then tend to find themselves. For example, with clarity about optimal χ and χ_s values plus some way to measure them, using SF theory you can find the smallest possible polymer which, at the smallest possible dose, will provide the largest possible stabilization. If, as a bonus, the molecule is available in a narrow range of MWts, then there are fewer trade-offs of thermodynamics and kinetics.

One of the luxuries of modern polymer synthesis is that techniques exist to give exquisite control of di-block and comb polymers. This means that those claims from the manufacturers about their new technology might not just be marketing spin. If you can have an intelligent conversation with them about the relevant χ parameters then you can find out whether you get both aspects of a great dispersant - the right molecular structure with the right near-monodisperse low MWt.

Curiously, the hard part is not so much the science, but the ability to have the intelligent conversation. As the only language I know in which the molecular needs can be discussed meaningfully is that of HSP, it requires both you and the supplier to speak that language. 5 years ago the chances of that happening were very close to zero. Now the benefits of speaking the HSP language are becoming more widely known to users, so sooner or later those suppliers who cannot speak that language will lose out.

4.8 Using HSP for SF theory

There is a beauty in the way theoreticians can distil complex problems down to elegant principles. SF theory has been systematically validated via experimental

data on well-defined systems. This means that although polymers are not strings of units that sit on lattices and that particles are not a uniform lattice surface and χ values are an oversimplification of how solvents, polymers and surfaces interact, the basic ideas translate well between theory and reality.

For formulators to use SF theory, all it takes, therefore, is a good estimate of the number of lattice units to use for A and B polymers (and that's not so hard), along with plausible values for χ and χ_s .

For the sort of ideal polymers used in SF theory, we know that the limit of their compatibility with the solvent is when $\chi=0.5$, and for HSP we know that this is when the Distance, D, is ~ 8 . Via some simple arithmetic we can work out that if we know D and the molar volume (MVol) of the solvent and the universal gas constant R and our temperature T in Kelvin that, as discussed before:

Equ. 4-1

$$\chi = \frac{MVol.D^2}{4RT}$$

For most of the common polymers used as A and B parts of stabilizers we have good estimates of their HSP values. For any solvent or solvent blend we also know the HSP, so from the simple formula for D we can readily estimate χ .

What about χ_s ? We can use some simple logic.

If $\chi_{\text{Particle-Polymer}}$ is small and $\chi_{\text{Particle-Solvent}}$ is large then we know the polymer is strongly preferred by the particle so $\chi_s \sim -1.5$ in the strange terminology established by Silberberg. If both of these is small then $\chi_s \sim 0$. If the particle strongly prefers the solvent then don't even bother with the calculations, you are not going to have a polymer-stabilized particle system.

Now we have set the extreme cases, what about real life? Here we hit a problem. Some careful measurements of carbon nanotubes (CNT) showed that the effective radius, R_0 that defines the limit of solvent compatibility is small, let's say 3. So if a polymer has $D=2$ and a solvent has $D=4$ it doesn't feel right to apply the above formula which is based on the limit of $D=8$. Even if the solvent and the polymer have identical D values, there is still the MVol term in the χ calculation. For the same D, a small solvent molecule is going to be "happier" than a large monomer unit of the polymer, and vice versa.

So I propose a pragmatic formula, which can be revised once the particle community has evidence to support a better one. Where R_s is the particle's effective radius of compatible solvents, $MVol_s$ & D_s are, respectively, the MVol and Distance of the solvent to the particle, it's easy to calculate χ_{Sol} :

Equ. 4-2

$$\chi_{Sol} = \frac{2MVol_s D_s^2}{R_s RT}$$

To calculate the χ_{Pol} between polymer and particle we need the $MVol_p$ of the monomer unit of the polymer, the Distance between polymer and particle, D_p and then an effective Radius R_p . There is no obviously good way of knowing this, so we use whichever is the smaller of the standard R_0 for the polymer in solvents and R_s , $R_p = \text{Min}(R_0, R_s)$:

Equ. 4-3

$$\chi_{Pol} = \frac{2MVol_p D_p^2}{R_p RT}$$

From these two χ values we then compute χ_s via:

Equ. 4-4

$$\chi_s = 3(\chi_{Pol} - \chi_{Sol})$$

The factor of 3 in the χ_s formula changes a $\chi=0$ for the polymer and a $\chi=0.5$ for the solvent into a value of -1.5. We need to be sensible and limit the χ values to a maximum of 0.5, along with using our common sense. We are not pretending that these are perfect formulae. We are using them as guidance with respect to a theory which itself contains approximations. My experience with using HSP calculations in general is that common sense and experience must always be part of applying the approach, so this isn't something new. Equally it's not saying "anything goes" - the HSP values we are using capture key elements of our formulations so these χ values are significant if not (yet) precise.

4.9 Imagine a world where ...

Imagine a world where you, the formulator, could speak to the suppliers of di-blocks, combs and functionalised brushes, and where you could share the language of HSP for the particles, solvents and polymers in your formulation and where they are as at ease with SF theory as you hopefully will be. This would be a very different way of working.

It would require the majority of reputable suppliers to provide the HSP values of their relevant materials, something they either do not know or are reluctant to supply. This has to change.

Many of the large companies talk about digitalisation, often with only the vaguest idea of what that means. Here we have a prime case for digitalisation. By having sets of numbers describing the solubility properties of key formulation ingredients and high-powered tools such as SFBox with which to work out the consequences, we could all move to a much better digital methodology.

As I write this, it's only a dream. Behind the scenes a number of us are working to encourage a "push" strategy from the suppliers. This book is one way to encourage the customers to create the necessary "pull".

4.10 SF of aqueous systems

SF theory works perfectly well with charged particles and polymers in water. Within HSPiP there is the ability to simulate neutral di-blocks and surfactants in water and look at the host-guest solubilization interactions. But these are not the concern of this book. SFBox can also be fed all sorts of charge-related parameters and does an excellent job at calculating many fascinating aspects of charged particle systems in water.

It makes no sense to use the current SFBox-FE to try to include charged systems. A front end focussed on charged systems might be an option for the future. But my current judgement is that simple water-based systems can be modelled adequately via DLVO if due caution is taken with ζ values and that I'm not sufficiently skilled or motivated to tackle the more complex charged systems.

4.11 Appendix: An SF master class

With some minor re-phrasing, here is SF the master class provided by Prof Leermakers. For those who are interested, it will explain some of the behind-the-scenes workings of SFBox.

We will work in lattice units l . For the sake of the argument you may take this to be $l = 0.3$ nm if you like. Typically I will not mention l when I use unit lengths, but it is understood that all linear lengths are in units l

We have a container with length L and thus with volume $V = L^3$ (Formally better $V = (Ll)^3$, but l is unity so is ignored)

Our particles are assumed to be cubes. We will use p as shorthand for particle. The linear size is D . The volume will be $V_p = D^3$. The area $A_p = 6D^2$. Although in reality particles are spheres, and we use radius R as an input, the refinements of adding factors around π are not worthwhile. In all calculations in the program, $D=2R$.

In our container there are n_p particles. The volume fraction of particles $\Phi = n_p V_p / V$ which, of course, must be smaller than unity.

In the remainder of the volume we have polymers, which to avoid confusion with p for particles we'll call poly, and we have solvent molecules. Both of these are composed of segments with volume $v=l^3$. In lattice units they thus have a volume of unity.

The polymer has a degree of polymerisation N , the radius of gyration of the chain is assumed to be $R_g = \sqrt{N/6}$, hence we assume ideal Gaussian size. The volume that the chain segments occupied on the lattice is also equal to N (molecular volume), again because the volume of a segment is unity. Note that a coil assumes a much larger volume, because it coils through space. This volume may be estimated as $R^3 \sim N^{3/2}$, though this fact is not used below.

The solvent is a monomer and thus has a molar volume of 1.

The number of polymer chains is n_{poly} and hence the volume fraction of polymer $\phi_{\text{poly}} = n_{\text{poly}} N / "V"$. Again this is a number smaller than unity. The scare quotes around "V" will be explained in a moment.

The number of solvent molecules is n_s and hence the volume fraction of solvent $\phi_s = n_s / "V"$. Note that we will require $\phi_s = 1 - \Phi - \phi_{\text{poly}}$.

Now let us assume that the particle size D is much larger than the coil size R_g . This allow us to work in SFBox in the planar geometry. We focus on the volume that is available for the polymer and the solvent, so we consider only the volume $V_f = V - n_p V_p$. That's why we had the scare quotes around "V" previously. In the modelling volume we only have polymers and solvent. The particles are outside the volume we consider. The polymer volume fraction used in the calculations is given by $\phi_{\text{poly}} = n_{\text{poly}} N / V_f$ and that of the solvent is $\phi_s = n_s / V_f$.

In SFBox a key parameter (historically called Nlayers, though many find that name to be confusing) is the number of lattice elements, NL. Now let us estimate what NL should be. This SFBox volume sits next to the area of the particles and thus the volume per unit area, which is equal to $NL = V_f / n_p A_p = (V - n_p V_p) / n_p A_p$. So per unit surface there are exactly NL of lattice sites available in the system. When the volume fraction of particles goes to zero we thus notice that NL diverges. So typically we ensure that $NL < N$. For homopolymers it is often sufficient to limit NL even more, e.g. to $NL \sim 6R_g$. For di-blocks and combs, the segment sticking out may not be so coiled, and in this version we use $NL \sim 10R_g$ for computational safety.

We number these lattice elements by $z = 1 \dots NL$, where usually the surface site is located at $z = 0$ (for interaction calculations there is a second surface at $z = NL + 1$). Of special interest to us is the polymer density profile which is $\phi_{\text{poly}}(z)$ and that of the solvent is $\phi_s(z) = 1 - \phi_{\text{poly}}(z)$. In our system the overall amount of polymer segment equals $\theta_{\text{poly}} = \sum_z \phi_{\text{poly}}(z)$ and thus the amount of solvent segments $\theta_s = NL - \theta_{\text{poly}}$. Now it is easy to see that the number of polymer segments (a key input to the calculation) is given by $\theta_{\text{poly}} = \phi_{\text{poly}} NL$.

In a closed system (canonical ensemble) the θ_{poly} is an input. Some polymer gets adsorbed (trains, loops and tails), the remainder floats freely in the solution. An estimate for the adsorbed amount is the excess amount $\theta_{\text{ex}} = \theta_{\text{poly}} - \phi_b NL$ where

ϕ_b is the bulk volume fraction of polymer. Often this is the volume fraction that exists near the last layer, i.e. $\phi_b = \phi_{\text{poly}}(\text{NL})$.

In order for the user to provide the volume fraction of added particles and of added polymer, the calculations are “canonical” which means that the “polymer” input to SFBox is θ_{poly} .

However to create interaction curves the calculations must be “grand canonical” which take ϕ_b as the input. The trick is to first do a canonical calculation to estimate ϕ_b then to use this as the input to the grand canonical calculations where free energies are calculated stepwise as NL are reduced one by one.

The calculated energies are normalized by subtracting the value at the largest distance, making all values far out ~ 0 . If there is a barrier to particle association, the value will increase. To relate these values to the familiar intuitions from DLVO theory, they are calculated on a “per-area” basis (A_{layer}), which in turn depends on the input particle size. Values are in terms of kT .

5 Rheology of particle systems

Many of our planned formulations turn out to be unworkable because their rheology doesn't allow us to pump, coat or otherwise handle the formulation as we had hoped. This chapter links the fundamentals of particles to the outcome of our measurements of, say, viscosity, and offering, wherever possible, ways out of formulation disasters.

For those who want a crash course in rheology, I can recommend the (free) *Abbott Guide to Rheology* on my Practical Rheology website, <https://www.stevenabbott.co.uk/practical-rheology/the-book.php>. This gives a broad overview of the main issues and is linked to a set of apps to bring the ideas alive. I also recommend Thomas Mezger's *Rheology Handbook*, which is an expert's guide that takes you step-by-step through each of the ways to measure the many possible rheological parameters. The book encapsulates one of my key problems with rheology: there are so many different ways to do approximately the same thing, with no obvious reason why one method, or one way of plotting the results of that method, is what you need to solve a problem. As a rough approximation, my guide tells you what and why, Mezger tells you in detail how.

In this chapter I am going to assume that you have some idea of the basic methods of rotational and oscillatory rheometry because I want to focus on what the techniques are telling us, and how we can influence the outcomes, rather than on the techniques themselves.

For particle systems there are two domains of interest:

1. The low shear domain where we probably are interested in yield stress effects, i.e. the inability of the system to flow unless some minimum stress is applied.
2. The high shear domain where most formulation processes take place and where the shear can usually decrease viscosity but can sometimes give a catastrophic increase.

Our problem with both is that the standard theories show that everything depends on volume fraction, ϕ , but not on particle radius, r . This goes strongly against our intuitions and experience that smaller particles are generally more troublesome than larger ones. To get around this problem we need to add an interaction energy term, E , which then feeds back into the rheology. After decades of papers on particle rheology, we still have only rudimentary ways to incorporate E into our calculations. In aqueous systems, DLVO gives us a chance to estimate E . In solvent systems using steric hindrance DLVO via the χ parameter might help. But in my experience we just go with the HSP values of the particles, solvents and polymers to get an idea of whether the particles prefer to clump with themselves (high E) or not (low E).

In what follows, we will start with the simplistic ideas, then add the E-related terms.

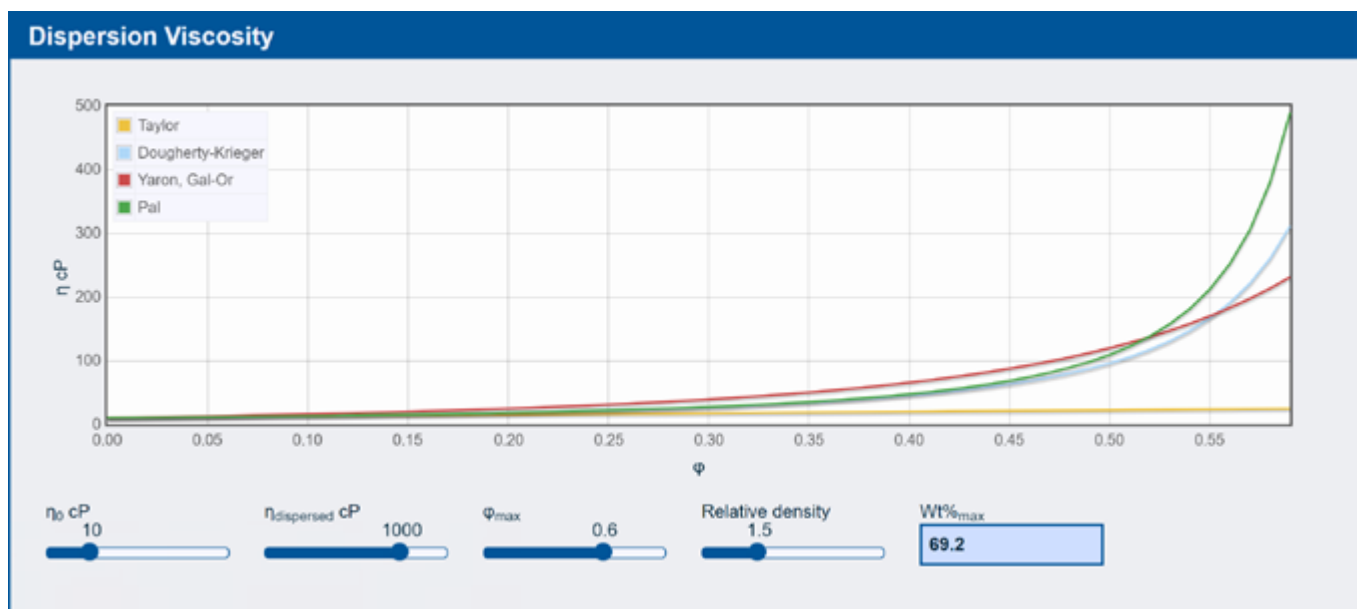
5.1 Low Shear and Yield Stress

The viscosity of any spherical particle system doesn't change much up to, say, 30% then it starts to rise catastrophically to some semi-infinite value at the maximum packing fraction, ϕ_m . All the equations, therefore rely on your estimate of ϕ_m and that can be anything from 0.55 to 0.63, with 0.58 and 0.61 being popular values. The differences between equations seem to me to be smaller than the differences between a value of 0.58 and 0.61.

The equation from Taylor is always mentioned as it is some sort of ideal, but is irrelevant. The most common equation for describing the effect of ϕ is Dougherty-Krieger. Others, such as the one from Pal (or Yaron, Gal-Or that is liked in emulsion particle formulations) may or may not be superior in some ways, but again it hardly matters for our real-world formulations of imperfect particles:

$$\text{Equ. 5-1} \quad \eta = \eta_0 \left(1 - \frac{\phi}{\phi_m} \right)^{-2.5\phi_m}$$

The important point is that as ϕ approaches ϕ_m the viscosity goes to infinity, as the app demonstrates:



App 5-1 <https://www.stevenabbott.co.uk/practical--rheology/Low-Shear-Particles.php>

There is only one thing of interest here which is generally overlooked. The calculations are all about viscosity relative to the solvent. If the viscosity at 50% is, say, 50cP and if you double the solvent viscosity from 5 to 10cP the viscosity at 50% also doubles to 100cP. This is obvious when pointed out, but most

formulators would assume that a 5cP increase in the base solvent would simply add 5cP to the higher viscosities at large volume fractions. In a rather difficult formulation crisis where we could not solve the problem of a high viscosity by reducing the particle volume fraction, we found that a small tweak to the base solvent to reduce it by a few (apparently insignificant) cP fixed the problem immediately.

We all know that this idealistic picture of low shear viscosities is not realistic. To get a feel for what really matters, I have invoked a theory called YODEL²⁵ which is simple enough to be usable and complex enough to be adequate. This allows us to introduce the association energy, E. Although YODEL is about yield stress, described in more detail shortly, it is also giving you an idea of what the low shear viscosity is - in simple terms, a high yield stress will give a large measured low shear viscosity.

My simplification of YODEL says that:

$$\text{Equ. 5-2} \quad \sigma_y = k \frac{E}{r} \phi \frac{(\phi - \phi_c)^2}{\phi_m (\phi_m - \phi)}$$

E is an interaction energy in units of kT, where, as DLVO theory tells us, this might be something like 3 or 4 for a lightly flocculating system²⁶. Well-dispersed systems in the 2-3 range give lower yield stresses than systems with bad dispersants in the 8-10 range. The particle radius is r, with smaller particles, other things being equal, giving a higher yield strength as our instincts tell us. The constant k is my fudge value to provide plausible results. A screenshot of the relevant app follows shortly because it is incorporated into the high-shear app thanks to the commonality of many of the key inputs and calculations.

Although the effects of the interaction energy and radius are important (proportional to E and 1/r respectively) the biggest effect is the rapid rise with ϕ as it starts to approach ϕ_m . So paying attention to small changes in ϕ can be more rewarding than worrying too much about E or the meaning of r in a system with a range of particle sizes. If ϕ_m is a given then E becomes important.

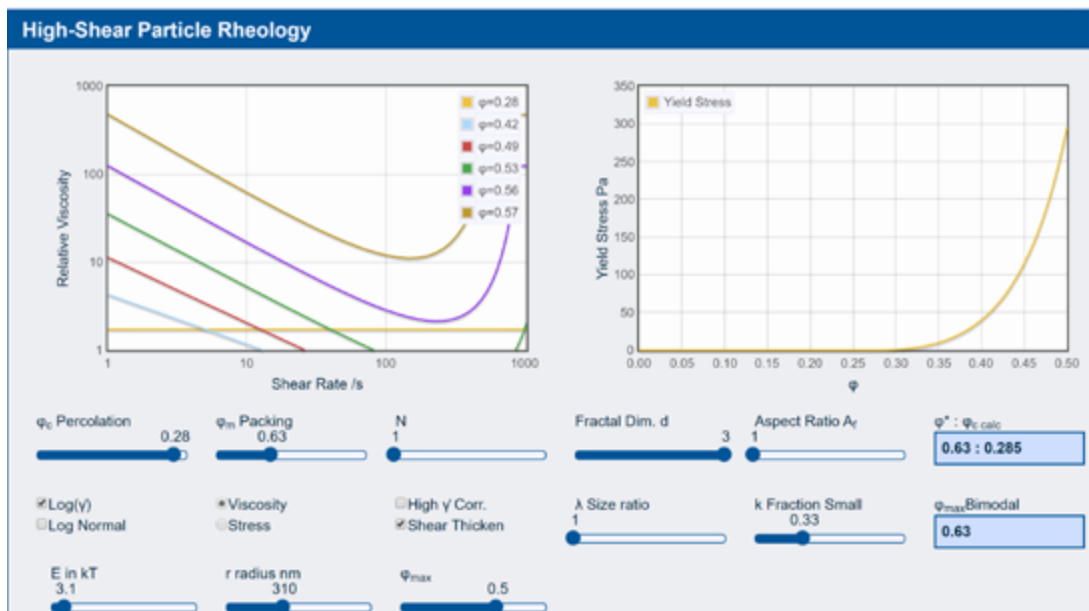
A specific example of world-scale importance is the use of "super-plasticizers" in cement to reduce E so the cement has the maximum flowability with the least amount of water. This is important because just a 1% reduction in water use has a huge impact on the mega construction projects around the world that consume vast amounts of cement. I had always imagined cement to be a rather boring product; reading other work by Flatt and Bowen soon dispels that idea; indeed,

25 Robert J Flatt and Paul Bowden, *Yodel: A Yield Stress Model for Suspensions*, J. Am. Ceram. Soc., 90, 1038–1044, 2007

26 Alert readers will say that E is -3kT for a flocculating system. I just find negative signs in equations confusion.

a lot of the cutting edge rheological work on particulate system comes from the cement industry.

All these ideas are captured in the app containing YODEL. Because many of the inputs are required for high shear particle viscosity, the one app contains both models:

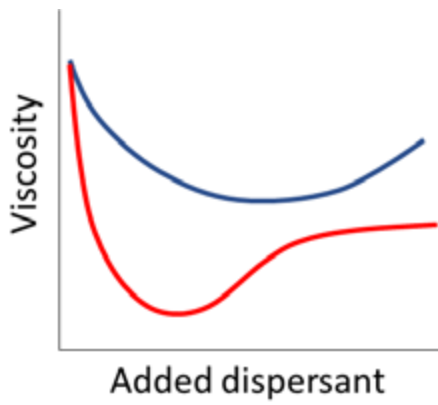


App 5-2 <https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php>

The settings in the app have an aspect ratio of 1, i.e. spherical particles. They are not so small, 310nm and they have a small attraction, 3kT. At $\phi=50\%$ the yield stress is 300Pa. If you make the particles modestly elliptical, aspect ratio =3, take the radius down to 100nm and double E to 6kT then the yield stress is 10x higher at 3000Pa.

You can see now why I am less interested in the simplistic low shear curves. They tell us nothing much other than the base solvent trick. It needs a model like YODEL to give us an idea of how hard we should try to keep our particles bigger, rounder and better isolated from each other.

An alternative, or maybe a complementary way to look at (especially) small particle interactions is discussed later when we have the necessary tools to implement that methodology.



Those who need to choose a dispersant with no way (such as HSP) to make a rational choice, can run a set of ODC (Optimal Dispersant Concentration) tests where the viscosity of a dispersion is measured while drops of a given dispersant are added. In the diagram, the blue curve is of a rather ineffective dispersant where you need to add rather a lot and

get a small decrease in viscosity; i.e. the E value is relatively high. The red curve is for a superior dispersant where adding relatively little provides a rather large reduction in viscosity. Why does the viscosity increase after a minimum? Because these dispersants are medium-sized molecules and can start contributing their own viscosity and/or start bridging or depletion flocculation via mechanisms which, in the past, we are probably too busy to worry about.

If you can combine some good HSP knowledge and have a good idea of the surface area of your particles then you can reduce the guesswork by choosing a well-matched particle/dispersant/solvent combination and adding the amount of dispersant required to cover the surface. With SF theory you can make a "tail rich" dispersant system and gain greatly in stability and ease of formulation.

To judge the right amount of dispersant, if you know the surface area via BET or dibutyl phthalate oil measurements discussed in the next chapter you may or may not have a relevant number. If the measure is of the true surface then it's directly useful. If it includes lots of hidden pores then the value is of little help.

If you know the area average radius $r = D[3,2]/2$ then if you know the total number of particles, n , the surface area is $4n\pi r^2$. The total number comes from the known mass/density to give volume and dividing by the volume, $4/3\pi r^3$, of each particle. If the previous surface area is likely to be too large because of pores, this idealised surface area assuming perfect spheres is likely to be too small.

But having some knowledge of the surface area and testing it against reality tends to lead to insights that are unavailable via blind formulation. If one type of particle works well with one estimate and a different one with the other estimate, that is revealing something interesting about the structures of the surfaces. Now we know the surface area of the particle, we need to know the "head" area per molecule of the dispersant. From the known weight of dispersant and its MWt we then know the relevant total surface area. The catch is that suppliers of dispersants are reticent about what they are supplying and we often don't even know the MWt. Here we can use common sense. In the bad old days, a typical dispersant would be a surfactant with MWt of 400-600 and a head area of 40-60Å². A modern hyperdispersant has multiple head and tail units, but the individual heads and tails are likely to be of a similar length, and the joining units won't be too significant. So guessing 600 and 50Å² will be a good start. Although the calculations are surprisingly tricky, the app does the calculations for us so we can see what's going on:



App 5-3 <https://www.stevenabbott.co.uk/practical-coatings/ODC.php>

As you can see, the app calculates surface areas in multiple ways. You get the actual surface areas for the given weights of particle and dispersant. You also get particle m^2/g to check against the BET/Oil value. For the dispersant you get m^2/mg (hopefully you compare mg of dispersant to g of particle!) and also the inverse which is mg/m^2 , a value sometimes provided by suppliers to tell you how much to add with respect to your known particle surface area. As you can see, a guess of $\text{MWt}=600$ and a head area of 50\AA^2 gives the value of $2\text{mg}/\text{m}^2$, which is a figure I found for one manufacturer's hyperdispersant.

The app shows that the given weight of dispersant provides $\sim 100\%$ coverage of the particle. Because of the multiple assumptions behind the calculation we can be sure that this isn't the exact amount required. But it gives you a good idea of where to start your optimization work.

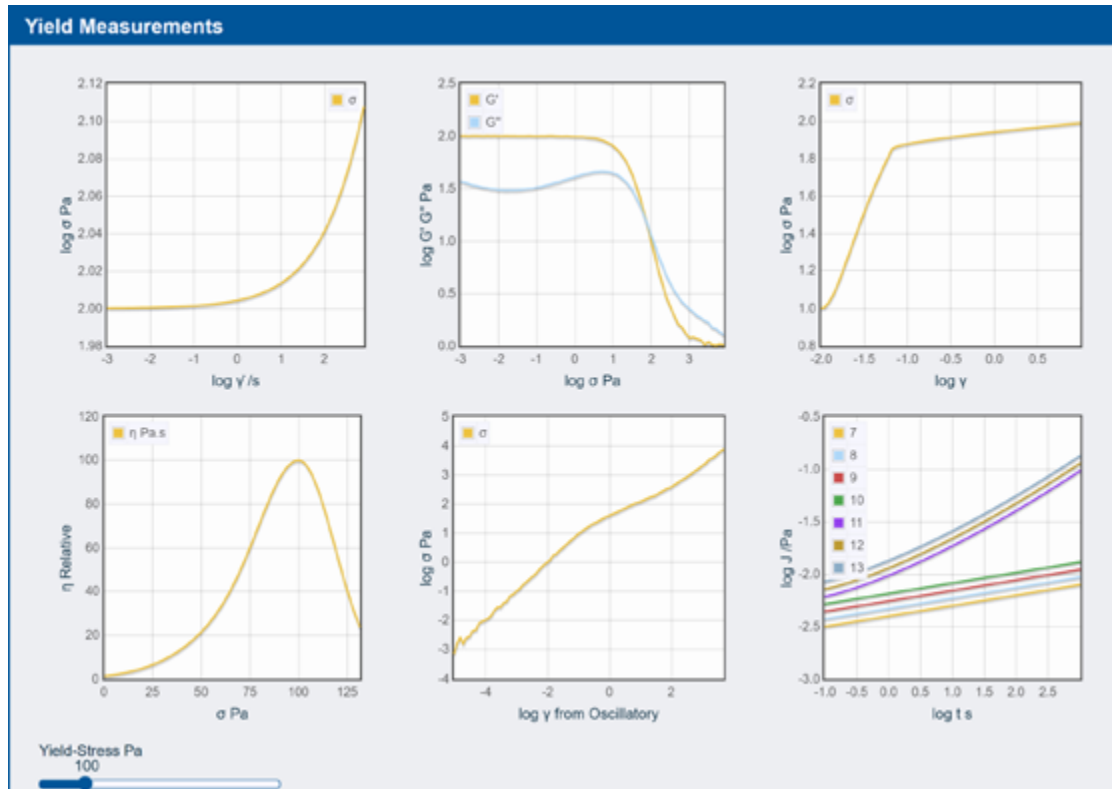
5.2 Yield stress

Intuitively we know what yield stress is. A blob of mayonnaise will stay as a blob until spread with a knife. An expensive-feel hand cream will stay as a "rich" blob on the hand yet flow smoothly when the hands are rubbed together. They are each effectively a solid until some stress is applied, then they flow like a liquid. Although both examples are from classic emulsions, not covered in this book, the same phenomena, for good or bad, exist in any high-concentration particle system. If you are lucky, yield stress is irrelevant and you can ignore this section. If it is significant for you or your end product then the problem is not so much "how is yield stress measured?" but "which of the 6 yield stress measures should I use?". Yes, everyone seems to announce, for no obvious reason, that they used *this* standard way to measure their yield stress objectively, even though others might measure a similar system using a very different technique. It is very confusing.

Fortunately a wonderful paper²⁷ from Prof Daniel Bonn's group at U Amsterdam describes many of the different ways, making it easy for me to write an app that shows what you would see if you measured the same yield stress, σ_y , (as an input to the app) in six different ways. Whenever I have to read a yield stress paper I go back to my app to find the nearest approximation so I can make

²⁷ Maureen Dinkgreve, José Paredes, Morton M. Denn, Daniel Bonn, *On different ways of measuring "the" yield stress*, Journal of Non-Newtonian Fluid Mechanics 238 (2016) 233–241.

sense of the paper and, perhaps, compare it with measurements on similar systems which have used a different method.



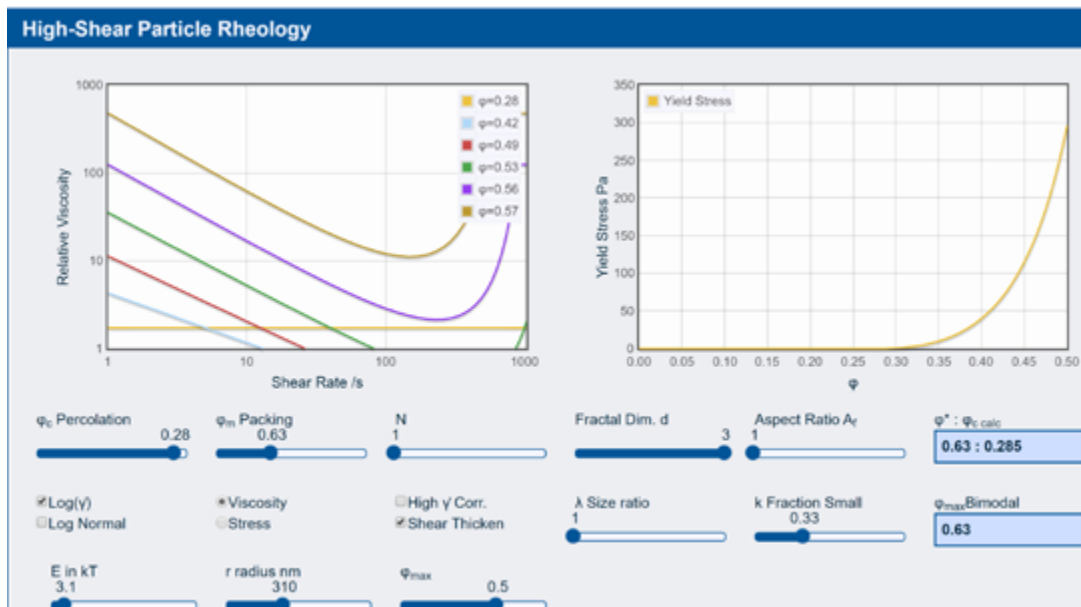
App 5-4 <https://www.stevenabbott.co.uk/practical-rheology/Yield-Measurement.php>

Those who want to understand the methods can read all about it in my Guide to Rheology. I just needed to alert you that the problem exists and to assure you that everyone is confused about this (which is why Prof Bonn's team produced the paper).

Having rationally minimized your yield stress and low shear viscosity (by having the smallest possible aspect ratio and the best possible dispersant and/or solvent compatibility) to be able to handle the formulation at low shear rates, it is time to look at what happens at high shear.

5.3 High shear rheology

By repeating the same image used to show YODEL we can get an instant guide to some key features of high shear particle rheology:



App 5-5 <https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php>

1. Below ϕ_c , the percolation limit, there is no shear thinning. The particles are not forming a continuous network so there is nothing that shear can do to help.
2. Although the lines are regular steps in the graph, note that the viscosity is a log plot and that the steps of ϕ get smaller, so tiny changes in concentration, when you are near the packing limit, have a massive effect on viscosity.
3. If you change the size ratio λ to, say, 3 and set the fraction of small particles, k to 0.333 then you get the maximum packing fraction ϕ_m rise from 0.61 to 0.67. This makes a huge difference to viscosities at $\phi > 0.5$. Even though this is an ideal bi-modal distribution, as we all know, having a proportion of small, medium and "standard" sized particles allows us to increase ϕ_m by a modest amount with a significant reduction in high ϕ viscosities.
4. Shear is very effective at reducing viscosity - typically by 2-3 orders of magnitude, which makes it possible to push the packing fraction for that little extra bit of performance from a formulation.
5. In this specific case, there is a catastrophic problem at very high shear - the viscosity heads off to infinity. This phenomenon is *not* well-understood and even for the same particle system, small changes in formulation can flip the behaviour from continued shear thinning to catastrophic thickening.

There are two other factors with a large effect on the rheology.

1. If you change the Aspect Ratio then ϕ_c goes down, so low shear viscosity goes up, so the high shear viscosities are also higher. Spherical particles are much easier to deal with!
2. If you change the Fractal Dimension to less than 3, you imply that there is some self-association within the dispersion. If you increase N , the average number of particles that are associating to 2 or 3 then you find that the viscosities increase strongly. Although this double effect captures the reality,

it is a very indirect way to see the effects of self-association. Don't blame me. For a long time, this was the best model I could find that could reproduce some of the effects of poor dispersion. The next section shows another app that somewhat improves on the situation, especially for very small particles.

5.4 Self-associating particles

It is a common experience to have a formulation containing a relatively small (say 1%) fraction of some specialist, small particles and a large (say 30%) fraction of large particles, but where the rheology is dominated by the small particles because they self-associate into loose networks.

A common example is battery formulations that have lots of big particles doing the main anode/cathode duties, with a small fraction of carbon black added to provide the necessary conductivity. Many of the practical problems in high-speed, high-quality production come from the high viscosities created by those small particles.

As we know from HSP, it would not be so hard to find a solvent in which those small particles were especially happy, so their tendency to self-associate would be minimized. However, the formulation has many other components that require optimal solvents so the result is a sub-optimal self-association. Using the language of DLVO, the trick is to provide a polymeric shell that produces steric repulsion. Given a basic DLVO curve, what is the effect of, say, the thickness of the protective shell on the low-shear viscosity, and how readily does shear reduce the viscosity.

By adopting a micro-rheological view of this²⁸, thinking through how the extent of clumping is affected by shear, it is possible to produce a credible model of the viscosity effects from the small fraction of small, interacting particles and add them to the standard high-shear rheology of the previous section. We add together the "structural" viscosity with the "hydrodynamic" (conventional) viscosity.

The specific implementation of the micro-rheology comes from Prasher's team at Lawrence Berkeley labs²⁹ and was based on battery formulations, though the approach is perfectly general.

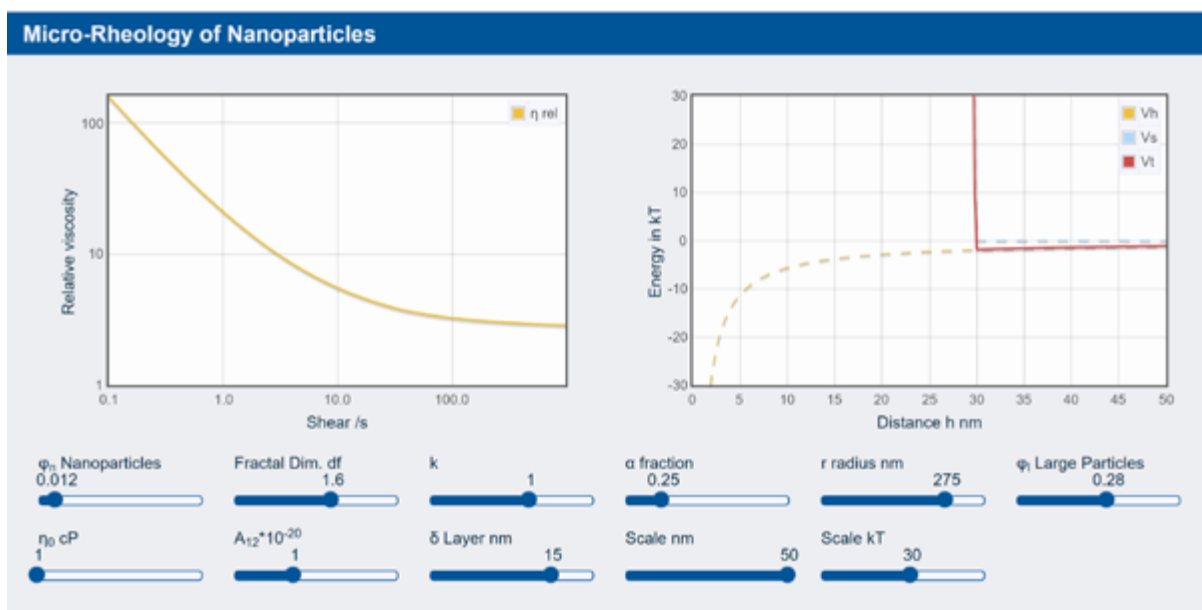
From the shape of the DLVO curve near its minimum (1st and 2nd derivatives, dependent on the shell thickness δ and the Hamaker constant A_{12}) it is possible to extract 3 parameters used in the calculations. F_c is the force holding the particles together at distance h_c , and a factor k_0 tells us the steepness of the

28 A. A. Potanin, R. De Rooij, D. Van den Ende, and J. Mellema, *Microrheological modeling of weakly aggregated dispersions*, J. Chem. Phys. 102, 5845-5853 (1995)

29 Fuduo Ma, Yanbao Fu, Vince Battaglia, Ravi Prasher, *Microrheological modeling of lithium ion battery anode slurry*, Journal of Power Sources 438 (2019) 226994

potential curve. From these plus a few other factors such as α , the probability of collisions producing structure and d_f a fractal dimension, it is possible to calculate a ϕ_a at a given shear rate which tells us the effective volume fraction of these small particles. When particles clump, they encompass a large amount of solvent so their effective volume fraction is much larger than the actual fraction. As we know from classical particle rheology, the larger the volume fraction, the higher the viscosity.

So the calculations consist of finding the ϕ_a at a given shear rate and via the other parameters working out the relative "structural" viscosity. Then the conventional "hydrodynamic" viscosity is added depending on the total volume fraction of small and large particles.



App 5-6 <https://www.stevenabbott.co.uk/practical-rheology/Micro-Rheology.php>

The reader will rightly protest that there are too many hard-to-know factors such as α or d_f or the precise shape of the DLVO. My experience using the app in real-world situations is that the details are not that important and that, instead, the core ideas of reducing A_{12} (maybe via HSP matching as discussed in the Solubility chapter) and increasing δ (smart dispersants) take on an urgency that is absent if you don't have a model that allows you to explore the options.

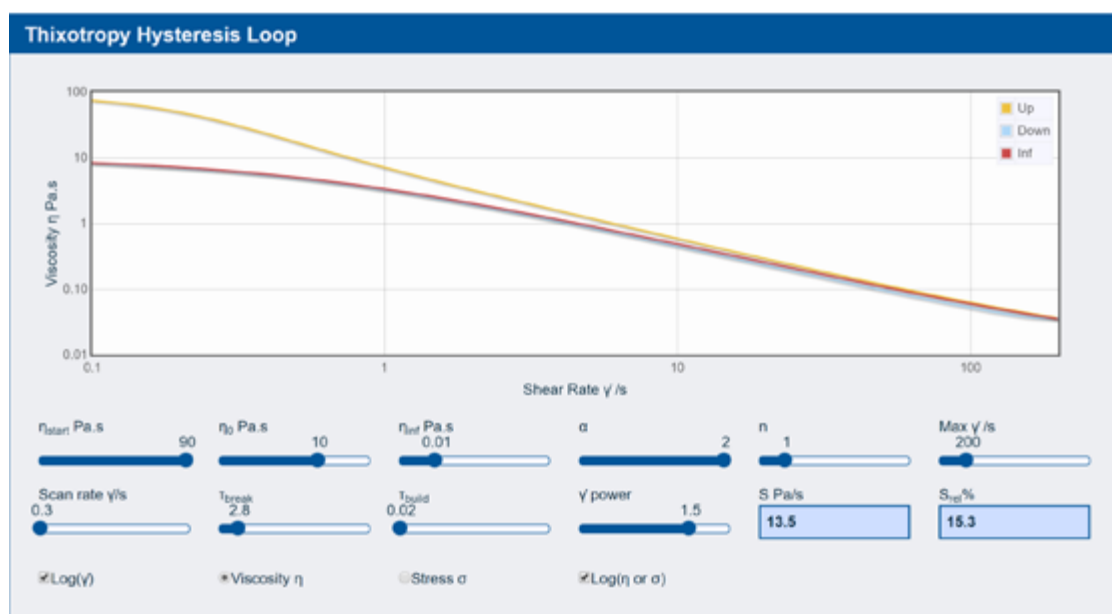
5.5 Thixotropy

We know that a thixotropic dispersion not only shear thins, but if you keep measuring at a constant shear, the viscosity continues to decrease. We need to be able to formulate either to avoid thixotropy or to use it (as in some types of paint) as a way to provide multiple functionalities in different parts of the processes with different shears and time-scales.

The problem is that because at any moment a measured value of viscosity depends on every other moment, it is hard to get meaningful numbers with which to understand what your formulation changes are doing.

Of all the possible ways to measure thixotropy, I've found only one which gives a good balance of insights for the least amount of work. I wrote another (<https://www.stevenabbott.co.uk/practical-rheology/Thixotropic-Recovery.php>) because some people like to use their oscillatory rheometers to look at thixotropy, but I can't see that the formulator gains much insight from it.

The good method is just a standard shear rate sweep to higher shear rates ("up") then back again ("down"). From the differences between the two curves it is possible to get some useful numbers:



App 5-7 <https://www.stevenabbott.co.uk/practical-rheology/Thixotropy.php>

The up curve is, by definition providing shear rates not encountered previously by the formulation. You have no idea how much of the decrease is due to normal shear thinning and how much is due to thixotropy. But the down curve allows you to disentangle the effects. At the lowest shear rate you are likely to see the true shift in viscosity after "infinite" shear, though you may need to do a couple more sweeps, or do the sweep at a slower scan rate (so the formulation spends more time at higher shear) to confirm that the lower curve is the true one.

Using the app, you set the lower curve's shape via a standard Cross equation with parameters η_0 , η_{inf} , n and α (the Practical Rheology site has a Cross equation app if you want to know more). Now you fit the upper curve by changing the three parameters that affect thixotropy:

1. τ_{break} which is the timescale for breaking the structure, which depends on
2. γ_{power} which is how rapidly τ_{break} decreases with shear rate, then

3. τ_{build} which is the rate (assumed a constant) at which the structure reforms.

What is often missing from discussions of thixotropy are the timescales for the structure breaking and making along with the size of the effect controlled by γ power. With this hysteresis app approach you get a deep understanding of these factors.

If you want a quick comparison of the "amount" of thixotropy then you use the Thixotropic Area, S , or its more meaningful S_{rel} which are measures of the area between the different curves (larger area means more thixotropy) made more meaningful as a % of the area under the up curve.

Thixotropy can come from many sources within a formulation. If we assume that it all comes from the particles themselves we can put numbers to the obvious fact that smaller particles will give higher thixotropy for at least two reasons:

1. Let's say that with each particle-particle collision there is a chance of the particles showing some sort of temporary adhesion. How does the probability of such collisions scale with particle radius? It turns out that the diffusion coefficient, D , of a particle scales as $1/r$ (with kT , π and viscosity involved). The collision rate between particles scales as Dn^2 where n is the number of particles in a given volume. The r 's in D and collision rate cancel out, so collision rate goes as n^2 . But the number of particles scales as $1/r^3$: if I have a volume of 1000 units then I need 1000 particles of 1 unit and 1 particle of 10 units. As a result, the collision rate goes as $1/r^6$.
2. If we assume an excellent dispersant and add enough to fully cover the particles with radius r , then if we happen to break the particles down to $r/2$, we will have half the required dispersant, so thixotropy will increase. This is because we reduce the surface area per particle by r^2 and increase the required number of particles by r^3 giving the final r^1 dependence.

5.6 That's not a long chapter!

Readers might, at this point, feel cheated. What a short chapter! But again it doesn't seem to be my fault. There are 1000s of papers on particle rheology and only a tiny fraction are of any practical use. Just extracting these few apps was a torture. One aim of my books is to either encourage or shame academics into producing theories and tools that make my own versions look foolish. Imagine what would happen if (to invent a name) The Association of Particle Rheologists reviewed the useful content they have created over the past decades of meetings, found that it is near zero and then did something about it. I'm not picking specifically on rheologists. It's just that I'm so exasperated by this short chapter that I have to say something.

6 Understanding the surface

A simple calculation (see <https://www.stevenabbott.co.uk/practical-coatings/surfvol.php>) tells us that any particle with a radius of greater than 8nm has (far) less than 5% of its atoms at the surface. Yet it is this tiny fraction of our particle that causes us the most problems. If the particles are smooth then we can use a mix of HSP and DLVO to understand the surface interactions during each step of the formulation and transformation into final product. If the particles are rough then there are extra frictional issues that can cause processing problems. If the particles are porous then we are either making use of the porosity or the pores add an extra confusing layer of complexity to our measurements of "surface" properties.

What I find odd about the particle world is that some measures which may in themselves be valid (but, as we shall see, are frequently invalid) are used totally inappropriately in an attempt to understand formulation issues. This chapter is my attempt to alert you to the strengths and weaknesses of the standard surface tests and, especially, to point towards what sort of test is appropriate relative to any given formulation issue.

6.1 Fancy techniques

If you really know what you are doing with an AFM tracing out the surface, or an XPS or TOF-SIMS analysis of the surface elements, or have specific questions about, say, distribution of elements within a particle that can be answered via EDS SEM then go right ahead. If, instead, you are rather vague about what the acronyms mean then you are probably happier without knowing. In general these techniques are a lot of effort for limited information, often on an unrepresentative subset (say 5 particles out of 5 zillion) and with only a vague link to genuine formulation issues.

In this chapter I want to focus on techniques that, at least in principle, are looking at the broad distribution of particles and providing general information about them. Even if we wanted to use the fancy techniques, the chances that the particles are sufficiently pure (rather than some formulated mix from a supplier or our own process) to make such analyses worthwhile are slim.

There are plenty of times when I have used fancy techniques to address interesting pure science questions about particle-matrix interactions, but these are big question issues ("How can we investigate the effect of polymer relaxation on particle adhesion?") rather than issues about getting better adhesion between a complex resin and a complex particle formulation.

6.2 Surface energy

It is odd that surface energy is revered as *the* method for characterising a particle surface. The one thing we know about surface energy is that the

slightest contamination can change it significantly, i.e. a monolayer on the surface can totally change how it interacts with test probes such as water. So when we use our particles in just about any formulation, the first thing that happens is that our pristine surface is no longer pristine, so measurements of its surface energy have ceased to be relevant.

We also know that most surface energies lie in the range from 30-50 mN/m, not even a factor of 2 difference across a vast range of chemical functionality. Yet people write papers getting excited about a change from 41 to 43 mN/m.

When I mention my scorn for surface energy, the immediate response is that I'm ignoring the fact that surface energy is important for wetting. There are multiple replies to this:

1. For most of the formulation steps our particles are already wetted so the wetting process is unimportant. When we mill them or disperse them or mix them with a spatula, within a very short time the surfaces are wetted because they have no choice - they are physically surrounded by the liquid and *everything* [if we exclude complex porous shapes] is wetted when surrounded by a liquid. The energy we supply in any form of mixing is much larger than can be resisted by super-weak surface energy.
2. It is rather rare for us to expect the particles to spontaneously wet as per the classic Washburn tubes discussed shortly. We force particles into a liquid phase by applying plenty of shear, so the niceties of "wetting" are irrelevant.
3. Everyone likes to measure wetting in terms of advancing contact angle (pure liquid onto pristine surface), which is mostly irrelevant, and almost no one measures the far more relevant receding contact angle which reflects a truer picture of how the liquid interacts with the surface. "It's much harder to measure receding contact angle" is an excuse often used. While true, it is also irrelevant, because if you are genuinely interested in wetting (and most of the time we are *not* interested) it is the problem of spontaneous de-wetting that is more serious because we can always enforce wetting by shear or removal of air.

To those who say that surface energy is important for adhesion of the particles to the matrix around them, I have to repeat that surface energies vary by a factor of 2 and adhesion by factors of 1000s and that a quick read of my "popular science" book, *Sticking Together* from the Royal Society of Chemistry will explain how to really get adhesion via intermingling, entanglement and dissipation. The book's YouTube channel also has some helpful videos.

6.2.1 Washburn tubes

Take a capillary tube, fill it with your powder, dip it in the liquid of interest and measure its mass m over time t . The well-known Washburn formula which

depends on contact angle, θ , the surface tension, density and viscosity of the liquid, σ , ρ and η , and a constant c gives us:

$$\text{Equ. 6-1} \quad \frac{m^2}{t} = \frac{c\rho^2\sigma\cos(\theta)}{\eta}$$

For a test liquid we know σ , ρ and η , so from a plot of m^2 versus t we can find $c.\cos(\theta)$. To determine c we do an experiment with a liquid with an assumed zero contact angle (e.g. hexane). If we assume that c is a constant across all test liquids we can work out the contact angle for other liquids. That is not controversial, though whether these angles mean anything on a powder *is* controversial. The bigger problem is that the contact angle, via some popular formula, is converted into a surface energy. Whether an idealised, bulk surface energy on a flat surface (calculated from contact angle via, e.g., the Owens-Wendt formula) has any significant relationship to that of a micro-rough or even micro-porous surface is contentious, especially as there are competing schools of interpretation of contact angles on controlled-roughness macro surfaces.

The fact that we are trying to get a mostly-irrelevant number (surface energy) via a dubious route (Owens-Wendt formula) is made worse by the assumptions behind c . This is supposed to account for the particle radius and packing. If every tube is packed perfectly, and if the particles are fully wetted by, say, hexane, then this correction for c is OK. But packing these tubes is an exercise in frustration. Indeed, one of the supposed reasons for measuring surface energy is that high surface energy particles are difficult to pack properly.

I am assured by experts in the field that the best thing to do with a set of Washburn tubes is to lock it in a drawer and throw away the key. The test is the maximum of frustrating work for the minimum benefit.

6.2.2 Simple IGC

I've written a book on Inverse Gas Chromatography (or, as I prefer to think of it, Interfacial Gas Chromatography) IGC (<https://www.stevenabbott.co.uk/practical-chromatography/the-book.php>) and although later in this chapter we will return to why IGC can be used to gain unique insights into particle surfaces, the one thing for which it is useless is measuring surface energies of relevance to formulators.

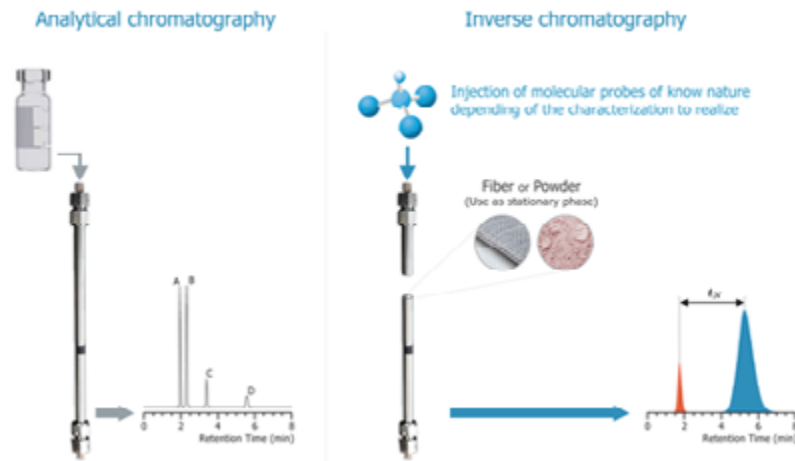


Figure 6-1 How IGC compares to regular GC

First, what is IGC? We know what standard GC is, a standard column where we try to find out things about the molecules we inject. IGC is the inverse because the column material is the unknown and we find the retention time of different known probe molecules such as heptane:

One of the standard techniques is to measure the retention times of hexane, heptane, octane ... and via a standard technique called Dorris & Gray a surface energy value can be extracted. The problem with this is that many relatively dull materials give surface energies of 100s of J/m² when we know that in reality they are in the typical 40-50 range. The naive acceptance of these meaningless numbers pervades the IGC community. The reason that most of the numbers are meaningless artefacts is obvious, but widely ignored.

It turns out that the measured retention time depends on the exponential of the interaction energy. This means that if 99.9% of the sites have an interaction energy of 15J/mole and 0.1% have an energy of 40J/mole, the retention times are so skewed by the 0.1% (they are so long) that the surface energy, calculated on the assumption that all sites are equal, becomes absurdly large.

Maybe that 0.1% of "high energy" sites is excitingly interesting and important? Usually not. Suppose we have 99.9% planar surface and 0.1% of holes into which the probe molecules can fit, where the surface energy inside the hole is no different from the planar surface. The probe experiences interactions on all sides rather than on one side, so its interaction energy is necessarily higher. This means that those measured high surface energies might be an artefact of 0.1% of very dull holes in the surface.

It turns out that IGC *can* give novel insights into surfaces, unavailable by other techniques and we will come to those later. Just measuring surface energies with IGC is doubly useless; first because the technique is usually misleading, second because surface energies are usually unhelpful.

I've been assured that the IGC-measured surface energies of carbon nanotubes and graphene are genuinely large. That's fine, but the least of my problems when trying to work with these materials is knowing their surface energy.

6.3 Surface areas

We can measure surface areas indirectly from measurements of particle size or directly from techniques that respond to interactions with the surface. Both approaches have major limitations so need to be treated with caution. For some simpler materials, the measurement techniques will work well, and that's good. I have chosen to emphasise the problems because most of the time our samples are complex enough that we need to be alert to the misinformation potential of any of these techniques. I also have to do this because manufacturers, understandably, prefer to focus on the splendour of their technique rather than on the many problems that arise because particles and their distributions are complex.

6.3.1 Via particle size

If we have measured the area-weighted radius of our particles, then, as we did in the Optimal Dispersion Concentration app, we get our surface area as $4\pi r^2$. So how do we get these particle size distributions and their radii? As discussed in the first chapter, we can get size from light scattering, centrifugation or electron microscopes. Each has many problems as the EU report mentioned in that chapter emphasises. The least bad seems to me to be the centrifugal method, but each method has its own strengths for specific types of particles.

6.3.2 Via "direct" area measurements

In the world of pigments for paints and coatings, there exist ASTM tests for surface area. You take a known weight of pigment and add drops of oil (e.g. linseed oil or dibutyl phthalate) till the surface becomes saturated. Although these values are used directly through other rule-of-thumb procedures, with some effort the specific surface area in m^2/g can be calculated. There are many problems with this method, not the least of which is knowing when the pigment is saturated. The fact that one ASTM test is called the "spatula rubout" method gives an indication that this is not a refined test. Indeed, the pigment industry really uses them as an indication in 5 steps from ultra-low via medium to ultra-high.

The standard alternative is the BET test where nitrogen is absorbed onto the particles at 77°K and via pressure (manometric) or mass (gravimetric) the amount of nitrogen absorbed at different pressures can be fitted to a theoretical curve (BET) from which a surface area can be derived. For those who want to know more, my IGC book describes BET in detail because it is relevant to an IGC-based alternative to BET.

Although there are plenty of issues with the BET algorithm, with the right equipment and analysis a reliable BET surface area can be measured. The question is how relevant this is to the formulator. A nanoporous particle will have a vast surface area and this might be relevant to the particles' interactions with solvents, but a hyperdispersant will see only the "real" surface area which will be much smaller. The EU report on particle sizing mentioned in the first chapter is clear about this - BET is pretty good for simpler particles and more problematic for more complex ones.

6.3.3 Use with caution

If there were a simple, objective measure of size and/or surface area, we would all be using it. The problem is not the lack of methods, it's the fact that particles are complex, with strange shapes, surface structures and pores.

The smart formulator will combine whatever information is readily to hand to form a judgement. If a particle sizer gives a large size and BET implies a larger surface area and, by implication, a small size then if prior knowledge suggests that the particles are porous, a quick calculation will give an idea of the porosity. If it is known that the particle is not porous then either the surface itself is amazingly rough or there's a flaw in one or both of the measurements. Maybe it's an unexpected aspect ratio problem. A very quick SEM will confirm or refute this. Maybe there is unexpected clumping in the size measurement, a hint that the dispersant or solvent isn't as good as expected. A re-measure of the particle size in different solvents might reveal that one solvent is especially good at keeping the particles apart; indeed that's what the centrifugal method for measuring HSP is all about. So a mismatch of surface areas might lead to an improved understanding of the solubility properties of the particles. In science, a contradiction between methods can be an opportunity to a better understanding or, sadly, a realisation that the user needs more training on how to prepare samples for a measurement technique.

6.4 Subtle surface properties

Having disparaged naive IGC-based surface energies, I want to recommend two ways that IGC can help understand surfaces in a way that other techniques cannot.

6.4.1 The 5-fold way

The "IGC 5-fold way" to surface understanding is made possible by newer IGC machines, using smart columns, clever injectors and autosamplers, run by smart algorithms.

The starting point for this is the near-universal frustration of particle suppliers and users where Batch A particles work wonderfully and Batch B don't.

Whatever fancy techniques are thrown at the batches, they are both in spec, even though they perform very differently.

There have been plenty of examples where IGC was able to pick out differences between different batches that were judged identical by other tests. However, to the extent that these IGC tests relied on "surface energy" they were subject to the uncertainties of those measurements. In any case, a test that relies on a set of linear alkane probes is not going to be able to pick out subtle changes of chemistry at the surface.

The 5-fold way is a set of measurements that give 5 different measures of the surface properties. While I have theoretical doubts about any one of these measures, it is clear that looking at surfaces with 5 different measures is more likely to lead to root cause understanding of what is changing. The smart thing about the 5-fold way is that each additional measurement requires data from only one or two extra probe molecules - the data build on themselves. Here they are:

1. The core measurements are the familiar linear alkanes and the extraction of a "surface energy" that is, for all its faults, a base reference from which the other values are calculated
2. The use of a few polar probes, with corrections for their molecular size (i.e. providing them with an equivalent size to a linear alkane) shows whether there are significant polar interactions, shown by a larger retention time than that expected from the probe's equivalent size.
3. The use of a few acid and base probes, following a similar logic, allows some insights into specific base or acid functionalities on the surface.
4. A few branched and cyclic alkane probes might show a lower retention time than expected from their equivalent size. This tells us that their shape is not allowing them to gain access to more complex sites, which gives an indirect measure of the surface roughness or porosity.
5. If a cyclic probe shows a greater retention time than a branched probe and, especially, if the retention time is greater than expected from its equivalent alkane number then we know that there is some specific interaction, beyond mere surface energy, that is encouraging the probe to linger longer. For example, a thin polymer coating might interact more strongly with a cyclic probe than a linear equivalent (in HSP terms, the δD for cyclic molecules is higher).

Whatever reservations there might be about the logic behind each of these 5 measurements (discussed in the IGC book), it is clear that the intelligent use of, say, 5 linear alkane probes, 2 polar, 1 acid, 1 base, 2 branched, 2 cyclic for a total of 13 measurements is not a lot of extra work (setting up the sample in the column is far more work) and provides a lot of potential insights. My practical experience is that the 5-fold way regularly shows up differences between seemingly identical batches, and because it can distinguish between roughness/

porosity effects, polar, acid/base and "other" interactions, it leads to helpful hypotheses of the root cause of the differences and, therefore, potential cures.

In the IGC book I add a 6th experiment, generating the Absorption Energy Distribution Function (AEDF) for a few probes as a complementary method to gain access to subtle surface understanding. While I admit that it is not yet as routine as the 5-fold way, the fact that it happens to generate a BET-style surface area is a welcome bonus.

6.4.2 Surface chemistry via IGC?

If you make a relatively thick (10-100nm) layer of some relatively mobile material (an oligomer or a polymer above its T_g) on an inert carrier particle and measure the retention time of solvent probes that cover a good part of HSP space then it is possible to fit the data in a way that provides the HSP of that material. This is described in detail in the IGC book and there's an app page showing the principle: <https://www.stevenabbott.co.uk/practical-chromatography/hsp.php>.

By a loose analogy, the chemistry of, say, the dispersants around a particle might be obtained in this fashion. It would be very exciting to get such detailed data, which would be so much more informative than the approximations provided by the 5-fold way, or, worse, classic surface energy values.

However, there are at least two problems with this idea. The first is that there does not seem to be a theoretical approach that can separate the classic IGC surface interactions from the more chemical interactions within the bulk of the dispersant or other covering.

Although I can imagine ways around this problem, I am not aware that there is significant effort going into this. This is partly because the IGC world is largely stuck in the past with poor measures derived from poor techniques on poor machines. But I think the second problem is more fundamental, because it affects not just this specific idea, but the whole idea of analyzing the surface of a powder through the use of, effectively, infinitely small doses of probe molecules in an otherwise hyper-dry atmosphere of nitrogen or argon.

6.5 Surfaces via HSP

When we measure the HSP of a particle we are neutral about which aspects of the surface are being measured. The measure might be of pure particle or pure dispersant, or maybe it is a complex mix. With our other knowledge of the particle we can make a more refined assessment of what the values mean.

The key point is that via this one measurement we have a rather good idea of how the particles will interact with solvents, polymers, other additives and, even, other particles. If I had to choose only one surface measurement technique

(allow me to know the particle size independently), this is what I would choose because it provides the most insight for the least amount of work.

6.6 Surfaces in real formulations

When we formulate particles we want them to be well-dispersed, compatible with their matrix, and maybe show strong adhesion to the matrix. Each of these properties is a function of the complex environment - solvents, polymers, other formulation ingredients. The properties depend on association, intermingling, entanglement and in any formulation they will be a mix of thermodynamics (if time allows) and kinetics (if not). All of this is far away from Washburn tubes or from AFMs, XPS, SEMs, IGCs.

So we have to ask what value we are getting from our investigations into the surface properties of isolated particles in an alien environment.

Certainly we need to know the particle size because that has such a fundamental effect on so many properties. Here my personal preference is to get the information from a centrifugal technique because, for all its acknowledged faults (for example, the EU particle sizing document pointed out a lot of variance of values that could be put down to variable training and equipment), it is dealing with significant concentrations, and the effect of the solvent environment comes "free" with each measurement because a bad environment will give fast sedimentation times via flocculation. Because a centrifuge handles multiple samples, it is quick to get multiple answers. For solvent-based systems, if you can get particle sizes *and* HSP values from the same technique, what's not to like? For aqueous systems, the effects of pH, ionic strengths, "minor contaminants" (such as surfactants with the wrong charge) can all be found at the same time.

It is possible that in your specific system, subtle surface effects have a significant effect on the final properties. The fact that IGC can distinguish between good and bad batches is evidence that this can happen. So the IGC 5-fold way would be a good approach. What's important is for there to be a good reason why, say, surface nanoroughness might be important, leading to a decision to follow the IGC route.

What is unacceptable is to say "surface energy is important so we must measure surface energy" without a clear reason for thinking that surface energy is important and a clear knowledge that technique X will measure the surface energy correctly.

6.6.1 The reality of surfaces

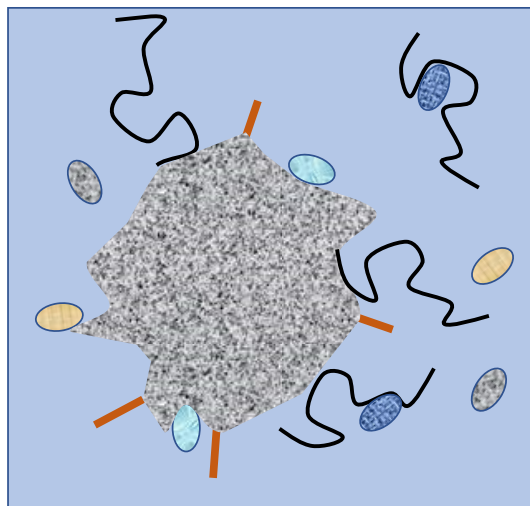
It's very nice to be talking about particles sizes, and BET surface areas and the niceties of surface interactions with small-molecule probes. We *have* to make

assumptions and simplifications and there *is* value in knowing the measured values.



The problem is not with the measurements but with the implicit assumption that we can think of the particle, as in this image. In addition to the oversimplification of the particle itself, the problem with this image is that the real world is nowhere to be seen.

When we think about particle surfaces, we should think about them in their context.



Instead we need to have a more realistic notion of what our particle is like, and the environment in which it is formulated. It is much more like this:

- An uneven shape
- Functional groups (orange bars)
- Polymers attached and free
- Other formulation components interacting with the particle and with other elements
- An environment of solvent or other polymers

If we keep this vision in mind when we measure "surface" properties we are more likely to measure what matters for the whole formulation. This is why I like to know the HSP of the particles. On their own the HSP values tell me nothing about the actual surface (pure particle, dispersant covered etc.), but at least I will have a good idea of how the particle "system" might interact with all those other things I need to put into the formulation.

I've tried to be fair to other surface measures, but if I need to formulate with some particles, I need to know how they will interact with the real world. The wealth of surface knowledge derived from the HSP infrastructure of human observation, centrifugal measurements and NMR relaxation seems to me to be maximum information for minimum effort. My reading of countless papers on particle surfaces is that researchers have often managed to create minimum information for maximum effort.

I know which approach I prefer.

7 Formulating in water

If you ever want to spend a scientific day wandering around in circles getting nowhere, then select an X and look for the "X effect in water". You will find hundreds of papers, each of which gives, finally and completely, the "real" explanation for the effect. The first third of the paper will explain why everyone else has got it wrong. The next third will be some re-hash of some old data. And the final third will be a confused description of their explanation that leaves you as puzzled as when you started. X might be the hydrophobic effect, the Hofmeister series, salting out (or, then again, salting in) of proteins, amazing polymers that become less soluble at a specific temperature - or many more.

This has been going on for a century, and is only now starting to show some sanity as converging lines of reasoning expose how much nonsense has been written about effects in water. One delightful read³⁰ describes how the author, having written a rather straightforward paper on how many water molecule were strongly bound to specific ions, found that he had stumbled into the world of strong opinions around the Hofmeister series (salts tending to have systematic effects depending, for example, on their size), mostly based on confused language and incoherent logic. Other trains of thought (see my Solubility Science chapter on the topic) have converged on some cases of specific interactions and others where the effect relies mostly on the obvious fact that if a molecule is somewhere then other molecules cannot be in that space - the excluded volume effect. And very often it's because a few kJ/mole of enthalpy one way and a few kJ/mole of entropy the other way happen to balance out to give the X effect.

My job in this chapter is to extract from this unhappy mess a few principles that can help you to formulate better in a confusing environment. We can start with a repeat discussion of DLVO and the zeta potential.

7.1 DLVO and zeta

The idea that charged particles will repel each other in water (and not in other solvents because the dielectric constant is too low) was formulated at a time where there was little chance of directly measuring the effects being discussed. Because the ideas were based on sound theory, it is claimed that they have been validated many times. If you put a, say, silica particle onto an AFM tip and measure the forces as it approaches a silica surface, you can get data that seems to validate the DLVO story.

But we can say four things against this story:

30 Andreas A. Zavitsas, *Some opinions of an innocent bystander regarding the Hofmeister series*, Current Opinion in Colloid & Interface Science 23 (2016) 72–81

1. These effects are studied and validated in situations that are of little interest to formulators. For good scientific reasons, the particles are simple, monodisperse, with a readily-measured and reliable zeta potential which will typically change in a rational way as the pH is changed - from negative to positive via an isoelectric point as the pH goes from high to low. Almost never do we have such nice circumstances. We have complicated particles in the presence of many other molecules, each of which might have some effect, good or bad, on the charge repulsion - even if we aren't dealing with the extra complexities of polymers with steric (or "electrosteric" in this case) stabilization, and bridging or depletion flocculation.
2. Even in pure studies, we find that, say, Li^+ behaves differently from Cs^+ (or F^- is different from SCN^-) even when they are at the same ionic strength which is what controls the effective repulsion between ions in the idealised theory. These are the Hofmeister effects. Worse, if I have a negatively charged particle and add a little bit of extra Na^+ , nothing happens. If I add a tiny bit of a long-chain quaternary ammonium salt (a surfactant like CTAB), then the system crashes out. The opposite can happen with a positively charged particle and the addition of an anionic surfactant such as SLS.
3. As discussed in the second chapter, the zeta potential is a scientific fairy story. We all have a general idea of what it means - *the* charge on the particle. But given that around any particle is a cloud of ions, counterions and water molecules which are more or less bound to the ions and to the particle, the reality is a mess. The zeta potential is what is measured when we measure a zeta potential. Because the behaviour of the cloud of ions is a mix of kinetics and thermodynamics, different zeta techniques measuring at different frequencies will give different values for the potential. The zeta potential is a circular and convenient definition that often works well in simple systems and becomes increasingly unreliable in real-world systems.
4. As mentioned in the earlier discussion on DLVO, there are so many intrinsic flaws in it that, as Ninham, a friend of both D and O said, in the context of AFM measurements, "*Indeed if anyone claims agreement with DLVO theory, their measurements are wrong.*"

The fairy story in the third point is that if we test the stability of some particles as we increase the ionic strength, their stability decreases partly as predicted by DLVO (because the extra ions decrease the particle-particle repulsions) and partly because there is a systematic effect whereby the size of the ion has an extra impact on the stability. You can spend days (as, sadly, I have done) reading papers which explain why *these* particles crash out faster with larger counterions and why *those* particles crash out faster with smaller counterions. Although they will dress up the explanations with terms like Hofmeister series, and the inevitable kaotropes and cosmotropes (if you don't know what they are, stay happily ignorant of them as they don't help), it's pretty simple: if your charged species on the particle has an especially strong specific attraction for, say, large counter-ions then the assumptions behind DLVO break down and the effective charge of the particle goes down.

But that's not how formulations work. We don't say "Oh, let's add 1M NaCl, NaBr and NaI to see what happens". We say "I need to add some sodium citrate to create a separate functionality within my formulation, what will happen to the particle stability?" The combined wisdom of 1000's of papers over decades of work on charge stability of particles will say "On the one hand and on the other".

As mentioned before, but worth repeating, the zeta potential story at least emerges with one golden rule: "Never measure a zeta potential, measure a series of them". I don't just mean "Measure the pH effect and identify the isoelectric point", useful though that might be if your formulation is likely to be (for other reasons) in the critical zone around the isoelectric point. I mean "You know the sorts of things that have to be added to your system for other reasons. Just try a few of them, at plausible concentrations, to see if for any reason they show a significant shift in zeta values."

This isn't random science. We know that zeta potential is not a well-defined scientific measure, we know that it can easily be wrecked by small amounts of things like cationic surfactants or trivalent ions. We have some plausible ideas of what the surface might be like, we have some plausible ideas of how our various ingredients might interact, so let's check if we have any significant (and therefore unwelcome) possibilities of sudden changes. If a few such experiments show that zeta remains unchanged, rejoice because this is going to be an easy system to work with. If they show that the system is very jumpy, well, it's good to know it sooner rather than later, and from the data there will be some plausible ideas of what to do about the problems.

This advice is especially important for those who want to do fancy formulations within the biological domain. It can sound very impressive to say that some pure nanoparticles in a test tube show some activity which might plausibly be anti-cancer. But it often needs only nano- or micro-moles of the wrong protein or DNA to fully wrap around the surface of the particles, changing their stability or functional properties. Biological molecules are complex mixes of anions, cations and hydrophobic zones and real biology contains complex mixtures of these molecules - any one of which might show a nano-molar liking for the particle and comprehensibly mess up whatever properties seemed interesting in an idealised formulation.

7.2 Balancing a formulation

The fairy stories are especially unhelpful because real-world formulations require us to add molecules that in general make our formulations unhappier, yet are required for other reasons. A good example of this is the need to add low volatility glycols to aqueous inkjet formulations. No one likes adding these - they're extra chemicals with the usual safety restrictions, and they have the potential to mess up the printed image over time (because they hang around, being low volatility). But they *have* to be added otherwise the inkjet head blocks

up as the water evaporates between prints. In my Solubility Science book I include a diagram from a conference paper from AgfaLabs³¹ which describes the desired behaviours of 5 different components in the formulation. Here is my own version.

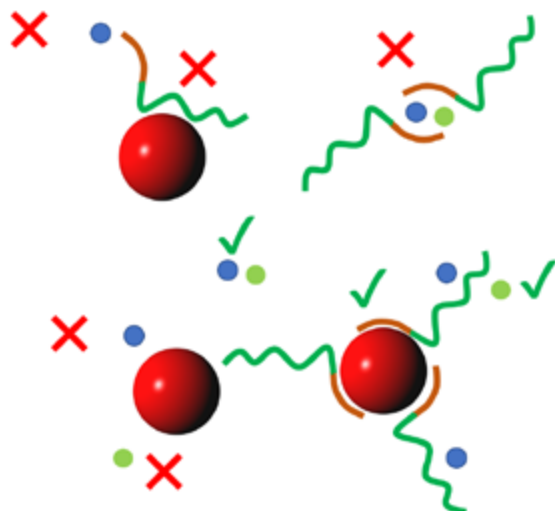


Figure 7-1 Juggling an aqueous inkjet formulation

We have the nanoparticle pigments which like neither the water (blue dots) nor the glycol (green dots) (bottom left), though they like each other (middle). We have a di-block dispersant which we hope doesn't like to self-associate in solution (top right), and where the redder block loves the particles and the greener block loves the water and glycol (bottom right) and where the reverse state (top left) is also not liked.

Because inkjet formulators don't, in general, like adding salts to their aqueous solutions, the particles cannot be stabilized via charges, so stability relies more on the hydrophilic part of the di-block being something like a polyethyleneoxide (PEO). As the pigments they use are not likely to be as hydrophobic as, say, polyethylene, a simple long alkane chain as the other part of the di-block won't be strongly attracted. So how about a polypropyleneoxide (PPO) chain? That would probably work OK in a pure water environment. But because ethylene glycol additives are now generally frowned on as unsafe, more typical would be functionalised propylene glycol equivalents. Now we start to see the problems that arise. The added propylene glycols may be rather too attracted to the PPO block, reducing the attraction to the pigment.

This neatly takes us back to SF theory and the various χ parameters. We can be happy that the χ of the PEO is adequate (low) with respect to water, but with the extra glycols added to the water, the χ of the PPO might be inadequate (not high enough) with respect to the water/propylene glycol blend. Even worse, the χ_s of the PPO block will not be in the required <-1 range.

³¹ Bart Wuytens of AgfaLabs, Measurement of HSP as a commercial service, found on <https://www.hansensolubility.com/conference/papers.php>

What is interesting is that the original AgfaLabs work on these complex systems was based on HSP measurements. While it is normal to measure the HSP sphere of a polymer, it is tricky, but possible, to measure the double sphere of a di-block polymer, allowing the informal intuitive analysis carried out in the conference paper. With the newly-enabled SFBox-FE these issues can be modelled in more detail.

7.3 Neutral molecules matter too

The world of proteins has many of the same stability problems as the world of particles - which is no surprise because I cannot see a fundamental difference between them. The effects of ionic strength as well as the effects of specific ions preoccupy protein research every bit as much as particle research, with endless mythologies about the causes and effects. Admittedly the situation is more complex because the proteins do not just "clump" or "not clump", they can also fold and unfold. But there are enough similarities that the comparison is valid.

What is more interesting is that everyone knows in the protein world that neutral molecules can have effects every bit as powerful as ions. The protein world has a superabundance of papers on the effects of urea, trimethylamine N-oxide (TMAO or TMNO), glycerol or sugars. The papers invoke all sorts of fancy ideas about water structure and are equally unhelpful and contradictory. The effects are readily understood in terms of the fundamental statistical thermodynamics of Kirkwood-Buff theory.

Given that formulators must add plenty of other things to their particles, including neutral molecules to provide other functionalities, it would seem obvious that there would be an equally vast and unhelpful literature on the effects of rockstar molecules such as urea or glucose on nanoparticles.

Despite my best efforts I can find no papers that have been interested in testing out the effect of these neutral molecules. Maybe I've not known where to look, maybe people smarter than me know that it's not a topic worth investigating. Or maybe no one has thought to try.

7.4 Another very short chapter

And that is all I can find to say about aqueous formulations. A bit of obvious DLVO using the app, a bit of scepticism around zeta potentials and the effects of any required charged species, some caution around χ parameters when using di-block non-charged stabilizers and that's it.

I had hoped/expected to do something really interesting by transferring the insights of Kirkwood-Buff theory from large proteins etc. to nanoparticles. While it is clear that this is achievable, in this first edition of the book it hasn't been achieved.

8 Simulating Particle Solubility Effects

This was going to be my grand finale chapter. Instead of just talking about particle solubility, I'd show how to simulate it.

Alas, that's not how it's worked out. When I started writing the book, I wasn't familiar with any nanoparticle simulation studies but assumed that's just because I hadn't looked very hard. As the book went on, I started to get worried that this chapter might not work out too well, but I still assumed that this was simply because I'd not looked hard enough. Well, finally, I have to admit that the situation really is dire. The number of convincing simulations of nanoparticle interactions with solvents is, depending on your criteria, somewhere approaching zero.

The reason is simple. Any simulation complex enough to hold the possibility of being realistic requires many days of mindless numerics on a supercomputer cluster, and even then you have to put in plenty of assumptions that others will question. Any calculation that can be done with a few non-linear equations involving a few pages of maths are so full of simplifying assumptions as to be unconvincing to most of us as well as being (mostly) incomprehensible. Intermediate "coarse grained" simulations are a lot of work that convince only those who have put in the assumptions that, say, a nanoparticle is a perfect sphere interacting with perfect spheres of solvents. I'm not including SF theory in this chapter because it is in a different simplistic domain of calculations from the ones that interest me here.

The studies that *are* carried out, and they are relatively rare, reflect the interests of the research team. If the idea of "solubility" of particles has never occurred to the team, it is unlikely that the study will think of comparing good and bad solvents. If, on the contrary, the team has no choice but to think of differences between good and bad solvents, then these appear in the calculations. Or the team might be intrigued between the differences of a particle in (essentially) liquid argon (in the calculations this is called Lennard-Jones fluid to avoid embarrassment) and a real solvent like n-decane - though the number of people who use n-decane as a solvent is very small - because, it turns out, having a lot of connected atoms allows an interesting comparison with the simple spheres of a Lennard-Jones fluid.

Because it is relatively easy to do calculations of idealised charged particles in idealised water with an idealised low ionic strength, and because it has been confirmed experimentally that such charged particles follow DLVO when they are larger than, say, 50nm there are plenty of papers addressing the fine details of such idealised systems, without the worries of the rest of us about real ions. As soon as you add, say, tetrabutylammonium salts to the system then "DLVO is fine" cease to apply, and there are still all those small nanoparticle systems to worry about.

Exploring some of the kinetic (rather than thermodynamic) aspects of simulations took me into gelation and so the chapter ends with some insights into models for gelation effects. These models offer some hope - some types of particle gelation are remarkably gentle and easily reversed, so we don't have to instantly despair if our formulations happen to gel over time.

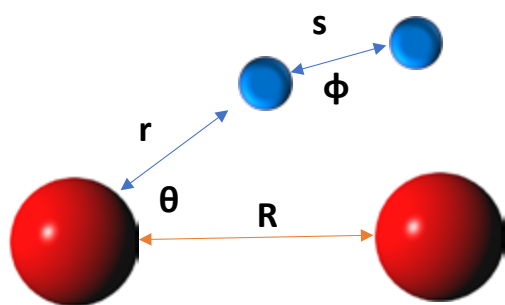
8.1 Particles in a box

The most obvious way to understand what might be going on when you have nanoparticles and solvent molecules is to put a pile of them into a virtual box and let them sort themselves out into some stable equilibrium via repetitive calculations. Whether you do this via Monte Carlo or Molecular Dynamics methods makes no difference to the fundamental issue which is the massive amount of computation, even with today's resources, required to do just one simulation.

The problem is that just to simulate a couple of nanoparticles with any degree of realism you might need 10,000 solvent molecules inside the box with them. And for any simulation of multiple nanoparticles with the potential to aggregate you might need a few thousand of them and a few hundred thousand solvent molecules. And your problems have only just begun.

The calculations need to know the exact forces between any two particles or solvent molecules at any instant in time so that their positions relative to each other can be adjusted ready for the next cycle of the calculation. The safe default assumption is to give everything a "hard sphere" interaction, with modest attraction at a distance and strong repulsion when in (near) contact. These are often Lennard-Jones potentials with a σ/r^{12} repulsion and ϵ/r^6 attraction. There are plenty of alternatives. I love the phrase "hard core Yukawa" for an alternative potential, and if you prefer there are formulae for penetrable sphere potentials that aren't as abrupt. For any degree of realism, the molecules need to interact via complex force fields depending on distance and angle. But then we have no plausible idea for the force fields between particles and solvents. If we did then we'd probably be able to work things out via simple solubility theory.

8.2 A few integrations



We have two particles a distance R apart and we have a solvent molecule at a distance r and angle θ and in general we have solvent molecules at distance s and angle ϕ from each other. Assuming (inevitably) Lennard-Jones potentials between everything (though with differing values for the repulsion and attraction parameters σ and ϵ) then via a series of integrations over r , R , s , θ and ϕ we can do a

credible job of calculating the overall energies, especially with respect to R where we want to know the extent of the attraction or repulsion.

The appeal of this approach is that we can get a feel for what might be going on without waiting days for a more realistic calculation to take place. The obvious downside is that it is far too crude to capture the complexities of what is going on.

8.3 A sliding scale

We can save ourselves a lot of work with MD/MC calculations if we go "coarse grained", i.e. we throw away a lot of the details of molecules by just saying they are point blobs with spherical properties.

And we can mathematize the trajectories of particles, to save a lot of random jumping around over femtosecond timescales using Langevin and Brownian dynamics.

As we continue this move away from molecules and particles jiggling around, we gain speed, so we can investigate more options, plus more understanding of the forces that are in play. At the same time we lose whatever realism had been placed into the original force fields and interaction potentials.

It's a tough and unrewarding space in which to model. People like me criticise whatever you do and the results after decades of difficult work are rather unimpressive.

So why am I even writing this chapter? For all the many difficulties, some key messages come through via whatever means are used for calculations.

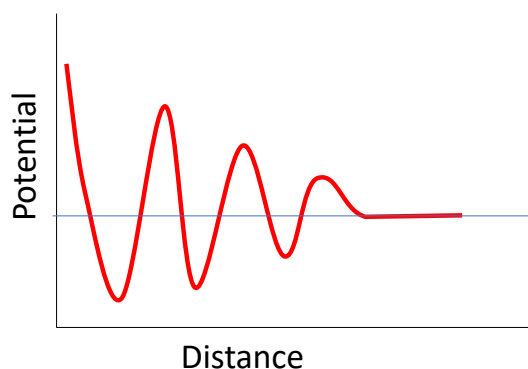
8.4 The potential of mean force

The classic curves showing DLVO forces are smooth because the core assumptions behind them don't recognise the existence of molecules with finite sizes and molecular interactions. DLVO experts are well-aware that this is an approximation, which often works very well for the academic-style systems that they investigate. The problem is that this omission of the role of molecules has become the default for all discussions on particle interactions in solvents. When you automatically dismiss the word "solubility" because these are "dispersions" we are deprived of a key formulation tool - our knowledge of molecular understanding.

What's funny about the situation links strongly to the Shimizu proof that nanoparticles are in solution, not dispersion. As he shows, acts like measuring particle interactions via neutron scattering or measuring particle size via ultracentrifugation intrinsically base themselves on single phase solubility science. Similarly, the act of doing MD/MC simulations, or any of the variants

down to the crude potential calculations, necessarily assumes that we are applying solubility thinking.

However we do the calculations, or, indeed, if we do subtle measurements with AFMs, we find that when particles come together, the potential contains a bunch of maxima and minima which, by no coincidence, coincide with the shape/size of the solvent.



If, for simplicity, we imagine solvent shells round particles, then we have a first shell strongly held, then a second shell, less strongly held and so on. As the particles come together, there are zones with nice stable shells between the particles, with it getting harder to push past shells where the shells are closer to the particles, or we have zones where you can't fit solvent molecules neatly, so they are forced out, giving a net attraction.

Whenever we do any calculation/simulation with any rational link to theory, we get such curves. If the solvent molecules are deeply unhappy next to the particle, these peaks and troughs might be very small. If the solvent has strong intermolecular interactions (e.g. hydrogen bonds) with the particle, then the peaks and troughs are large

Common sense suggests, and high quality simulations show that ordered solvent shells are a myth. Instead they are *statistical* solvent shells, meaning that on average you might have many more solvent molecules around the edge of the particle than expected by chance. So we have average or "mean" solvent shells. So the forces exerted between particles are mean forces. Classically this is shown as the radial distribution function which shows peaks and troughs coinciding with integer (peaks) and half-integer (troughs) values of solvent size.

The graph, repeated below on the left is, therefore, showing a "potential of mean force", PMF. In that specific example, the PMF is ultimately stabilizing - the solvent shell is obviously rather (statistically) strong. There are plenty of alternative possibilities:

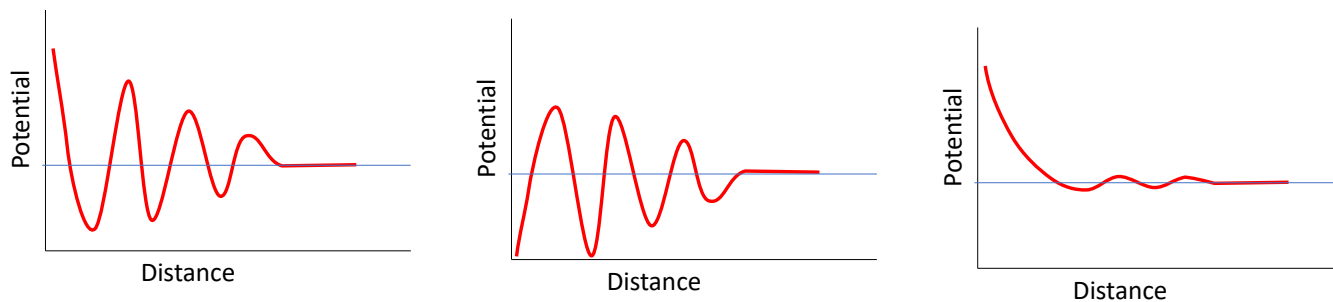


Figure 8-1 There are many possible PMFs.

In the second one, the solvent does a good job at providing a barrier, but eventually is overwhelmed by the particle-particle forces. In the third one, the solvent effect is very small and in many classic DLVO cases, the solvent effects don't show at all.

The point of introducing the PMF is to provide a visual metaphor for the particle community. If you have strong-enough solvent effects then the PMF can be ultimately completely protective, just as seems to be the case for the silicon nanoparticles mentioned earlier that are soluble in DMSO or NMP.

8.5 How stable is stable?

All but the bravest of simulations assumes that the particles are at "infinite dilution". There's a sort of contradiction here - if the particles are infinitely dilute then no two particles will ever see each other so there's no problem of stability to be solved. What the term is meant to convey is that when we examine particle-particle interactions in solvents, there are no other particles within significant distance to perturb things, so we can focus on the two particles and the many solvent molecules.

If you continue with this infinite dilution idea, you can calculate, as discussed earlier, the Stability Ratio, W . If we assume that the rate at which two particles will collide and associate is controlled just by diffusion, with a certainty of sticking together you get a collision rate of R_d but if there is some form of barrier which reduces the rate of collision we have R_b and $W = R_d/R_b$ with a larger value of W being desirable for stability. Calculating the two R values is surprisingly tricky and my DLVO app only manages to do it for pure charge stabilization via some mathematics full of incomprehensible (to me) functions.

Fortunately, the calculations are often wrong for cases of smaller barriers because it is frequently observed that such systems are more stable than DLVO would suggest - which is good news.

There are at least two reasons for this; the first is thermodynamic, the second is kinetic.

The thermodynamic reason comes from the potentials of mean force that involve real solvent molecules getting in the way of the classic, continuum theory. Bigger molecules with stronger sets of interactions will cause larger deviations in the PMF.

There are two kinetic reasons. The first is linked to the thermodynamic reason: if there are bigger molecules with stronger sets of interactions, even if their thermodynamic barrier is not so great, it takes time and energy to move the molecules out of the way. As the molecules jostle out from the gap between particles, they dissipate energy and decrease the chances of the particles overcoming the local barrier. A few heroic examples of molecular dynamics calculations of nanoparticle systems demonstrate this effect quite convincingly.

The second kinetic reason applies even to the continuum model of charge stabilization. If we measure "the" zeta potential, it turns out that the value obtained depends on the speed of the measurement technique. At slow speeds, the various ion clouds can cancel each other out to a large extent, leading to a smaller zeta potential. At higher speeds, local higher charge densities can persist during the measurement cycle, giving a higher zeta. I don't know about D&L, but V&O were well aware of this and in their publications urged caution about over-reliance on "the" zeta potential (in addition to the fact that zeta is only a pragmatic approximation for the "true" potential). But the DLVO we all know generally glides over these subtleties, so we can end up being (pleasantly) surprised at the stability of particles that should, naively, be unstable.

8.5.1 25 kT

It is commonly said that if the classic DLVO barrier is 25 kT or more then the particles are stable. As mentioned before, the unit kT is the amount of energy available through random thermal motions. The time taken for particles to go over a barrier of height B depends on e^B , so 25 kT gives a very low rate of barrier crossing.

Let's suppose that for a given system, 25 kT gives an effective stable life span of 1 year, which is approximately 9000 hours. Each 1 kT less is a factor of $e^{-1} = 0.368$ faster crossing over the barrier, so by the time we are at 20 kT the system will be stable only for 5 days, and by 17 kT only 3 hours. A kT is rather little, so a solvent change to boost a system from 17 to 18 kT does not have to be major to take something that will crash out over a long lunch to something that is stable for a working day. It might take only a tweak to SF theory to get the polymer/solvent system to deliver 2 extra kT and give us 24hr stability.

The fact that many particle systems are not stable indefinitely is a concern if, like me, you use thermodynamic solubility arguments. We know that things with "real" solubility are stable indefinitely so maybe systems with a finite lifetime before they crash out cannot use solubility thermodynamics. However common sense intervenes. It's normal and respectable to do thermodynamics on a system in a timescale where it is effectively stable. Experiments over minutes on a system stable for hours are as acceptable as experiments over years in a system "only" stable for centuries.

We also have the reverse question of whether crashing out is permanent or reversible. Experience says that if I crash some particles out of a bad solvent, then it is harder to get them back into a formulation than if they'd come out of a good solvent - the residual solvent, dispersants and polymers around the particles from a good solvent can be enough to allow things to be salvaged. This is not an effect limited to nanoparticles. I've had plenty of ordinary polymer formulations which are difficult to redissolve if I've messed them up via a bad solvent choice.

This admittedly vague train of thought finally brought me to gels.

8.6 Gels

This is the final section of the book. Whether it really belongs here is debatable, but there is at least some logic to it because a common sign of "crashing out" is the formation of a gel. Is this a real crash? It might be, or it might, as explored here, be a surprisingly gentle and reversible effect even though the gel itself is as solid as a real crashed-out formula. The reasons for the combination of gentleness and solidity are rather interesting.

Once you start having significant concentrations of particles, systems can swap from sols to gels. Well, that's the common terminology based on the idea that low particle concentrations aren't "solutions", but because they look and act like solutions they were given a different name, i.e. sols.

By being relaxed about calling them solutions, we can then ask what's special about gels. But first we have to remove an ambiguity. Many nanoparticles, such as silica, are prepared by the "sol-gel" method where there is a chemical reaction that transforms a real solution (e.g. tetraethyl orthosilicate) into a gel of concentrated nanoparticles. As the science of that transform is easily found on the web and is endlessly discussed, it is not included here.

The reason we are discussing gels here is that while to some people gels are formulation forms with positive attributes, to others the fact that their particle system gels is yet again a proof that particles are "dispersions" so will crash out in the form of failed gels. My worse-ever crash was on a production trial where our nanoparticle formulation failed within a couple of seconds of turning on the

pump to send it to the production machine. Its encounter with some very low levels of silicones within a boring filter we'd added to remove large particles was enough to turn it into an intractable gel.

The persistent aim of this book is to give us a language with which to formulate rationally away from dangerous regimes within particle/solvent/polymer space. A rational solubility-focussed language makes it easier to navigate around the complexities of gels.

A reasonable definition of a gel is a state in which the particles form a space-filling network that can support a stress. Like many definitions it's neither complete nor specific. We can have polymer solutions that are so viscous that it's hard to say whether they are a gel or not. At the other extreme, when there's enough of a space-filling network we have a glass. And in any case, the existence or non-existence of the network depends on history - it's rarely unique. So we have, as in the previous section, a curious mix of thermodynamics and kinetics.

If we start with mythical hard particles that have no interactions other than bouncing off each other when they make contact, then we can easily define the onset of a gel-like state because we have a simple space-filling network. This point is the percolation threshold, discussed in the first chapter. For spherical particles, we need 28% by volume, for rod-like particles it's less than 1% and for clay-like platelets it's a few percent. This behaviour is nicely captured in the high-shear rheology app.

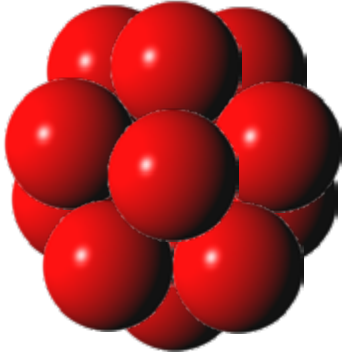
If we have highly-charged particles, such as Laponite clays, in a low ionic strength environment then although we get what looks like gel formation, it's generally considered to be a "Wigner glass" where the name originates from Wigner's original idea that electrons can form a crystal enforced by their mutual repulsions. As this form of particle gelation seems to be relatively uncommon, we'll discuss it no further.

8.6.1 From ergodic to icosahedral via isostatic

I couldn't miss the chance to put the words "ergodic", "icosahedral" and "isostatic" into a heading. It's partly for fun and partly because the three words capture some deep truths about gelation, and tend to be found in the better papers on the subject.

Ergodic simply means that the particles can, over time, reach every part of the 3D space in their container, i.e. they are free to move without global constraint. Of course in the very short term they (literally) hit constraints when they bash into other particles, but in the seconds/minutes/hours time-scales of gelation, they are clearly ergodic.

Isostatic means that the average net force on any particle is zero and by Maxwell's definition this means that the 3 degrees of freedom (x,y,z of 3D space) of any particle are balanced by forces in 3 directions. If your gel is isostatic then it's going to be rather rigid because there's not much opportunity for particles to optimise their energy by moving. Conversely, if you have any other form of gelation (e.g. isotropic, forces in any mix of directions) it is not as rigid.



We need the word icosahedral for a reason that surprised me greatly. The isostatic rule implies 3 connections per particle. It turns out that icosahedral packing provides a compact structure, made from 12 spherical particles, each touching in 3 places. The icosahedron is made up of pentagons and if you carefully look at gelling particles you find plenty of 5-membered structures, with plenty of icosahedra if the gelation process is gentle enough to attain such good equilibrium packing.

These "LFS", locally favoured structures, are hugely significant in two ways. First, they are amplifiers of small effects. Papers on gelation often marvel at how small the interaction energies can be (just a few kT) while providing solid structures with rheological G' values significantly higher than G'' , i.e. elastic networks³². Indeed, gelation is sometimes defined as the point when G' exceeds G'' . The key phrase is "cooperative bonding" - the existence of one very temporary bond can, by chance, be matched by another and the structure takes on a strength beyond the individual contributions. Second, this amplification leads directly to percolation across the sample, which provides the gel structure. This second effect is often stated but I had never understood it. The diagrams are my summary of how these structures generate percolation.

Before looking at them, we have to clear up a confusion. Right at the start of the book we saw that there is a percolation threshold of ~28% before there can be a continuous chain from one side of a system to the other. With classic hard sphere particles we see this as the onset of significant viscosity as soon as 28% is exceeded. This type of percolation is "isotropic" - it's got no structure to it, it just happens. For gelation, percolation is very different. We'll first look at a gelled structure then discuss the different type of percolation.

³² A simple guide to G' and G'' is on my Practical Rheology site: <https://www.stevenabbott.co.uk/practical-rheology/G-Values.php>

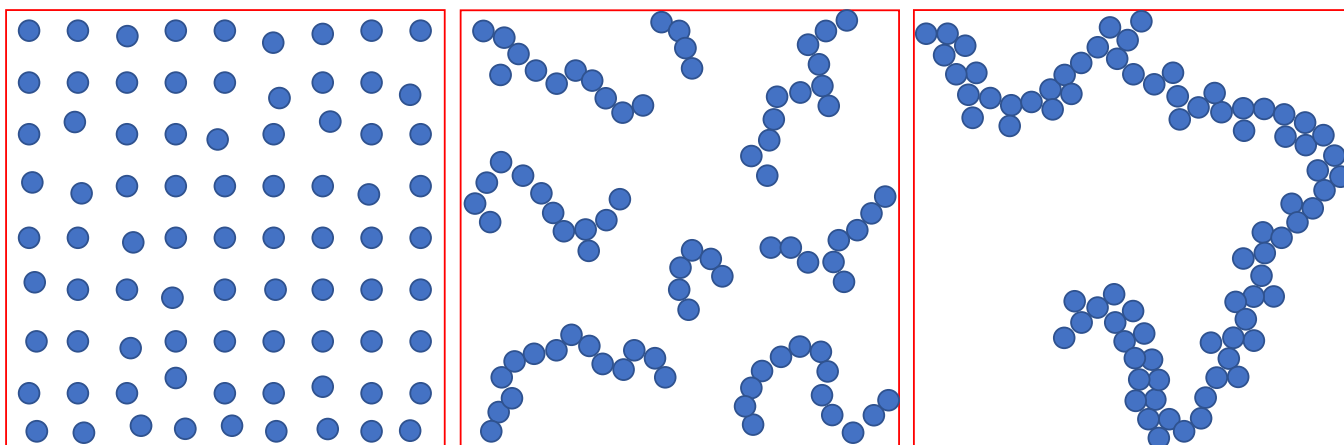


Figure 8-2 Left, 15% ergodic particles don't span the container. If they join up loosely (middle) then they still do not span. But as clusters (right) they can now efficiently span the system, which implies gelation. There are the same number of particles in each box.

The examples in the figure all have $\sim 15\%$ spherical particles so cannot show isotropic percolation. In the left-hand image, the particles are more-or-less randomly placed and are enjoying full ergodicity. In the middle image, at some moment in time there happen to be some pseudo-chains, presumably helped by some small particle-particle attraction, but they are of no significance. This system is still ergodic because those chains can appear and disappear anywhere. The system on the right has taken particles away from random positions and joined them up into sub-clusters. These sub-clusters are joined to each other and cross from one wall to another - so this system is non-ergodic and gelled. For the gels studied carefully around their gelling points, you cannot get this number of particles into this coherent a structure without cooperative bonding - individual particle-particle interactions are too weak.

We can be confident about the science of all this precisely because the physics of gelation is generic, i.e. it does not depend on subtle details. We cannot predict how our *specific* system will gel because we don't have the necessary fine details, but we can describe the *general* process with precision. This means that computer simulations are relevant and meaningful, and link very nicely to the recent sophisticated measurements of the gelling process using techniques like XPCS (x-ray photon correlation spectroscopy) that can give the spatial and temporal resolution required to be able see what is happening.

8.6.2 A different percolation

When the same word means different things to different communities, things get confusing. I wanted to write a simple app based on Adhesive Hard Sphere (AHS) theory that captured the effects of the local favoured structure, LFS. There is a relatively simple equation from which to calculate the percolation threshold. For non-interacting spheres this turns out to be $\phi \sim 0.75$, even though we know that the percolation threshold is 0.28. It turns out that the definition I'm used to is *when there is any single path across the bulk*. The other definition is

when each particle has a path to every other particle, hence the higher value for non-interacting spheres.

Once the definitions are cleared out of the way, things become interesting. As mentioned above, these gelation processes can be analysed by relatively basic physics. The AHS approach³³, which is very basic, says that the particles have no interaction until they are, effectively, touching, in which case they have an interaction which can be characterised by a pseudo-temperature τ (this is called the Baxter temperature) which is large (say >10) for weakly interacting particles and small (say <0.1) for particles with significant interactions. If you try to implement some basic rheological analyses on real-world systems using the simple AHS approach, the results are respectably good.

There are two things which interest me about this approach. First, for a given τ there is a ϕ where percolation (in terms of everything being connected to everything) sets in given by:

$$\tau = \frac{1 - 2\phi + 19\phi^2}{12(1 - \phi)^2}$$

Equ. 8-1

When $\tau=0.1$, this percolation volume fraction ϕ_{perc} , is 0.1, for $\tau=1$ it is 0.44.

Once we know ϕ_{perc} we can calculate the mean cluster size, S , at lower ϕ values. This is given by:

$$S = \frac{1}{\left(1 + \frac{\phi}{\phi_{perc}}\right)^2}$$

Equ. 8-2

For $\tau=0.24$, $\phi_{perc}=0.22$ and at $\phi=0.2$, S is 13 and at $\phi=0.21$, $S=50$. As you see, when ϕ reaches ϕ_{perc} the cluster size goes to infinite, i.e. everything is touching everything.

We are almost there in our journey through AHS theory. The cluster size and the viscosity increase near the percolation limit can, in many systems, be significantly higher than suggested by the simple theory. This is because the theory assumes that individual clusters can slip past each other - they are "lubricated". Once they start to be "rigid" then things get more viscous more quickly. The trick is to apply a probability that the individual clustered particles are themselves actively connected. When this probability is 0 we have the equation above, when it is closer to 1 then *if* we have significant clustering *then*

³³ I'm following the rheology paper of Gang Wang, Andrew M. Fiore, and James W. Swan, *On the viscosity of adhesive hard sphere dispersions: Critical scaling and the role of rigid contacts*, J. Rheol. 63, 229–245, 2019

the viscosity is much higher. I'd thought of applying all this, but it's not really suitable and, in any case, the general point is that modest particle-particle interactions can lead to large increases in viscosity through cooperative binding into significant-sized assemblies.

What about the time-scales for all this? The analysis above indicates, and experiments show³⁴, that there are really two sorts of time-scales. The first is the build up of the local clusters, i.e. up to ϕ_{perc} which takes place with a time dependence of $t^{0.57}$. This curious dependence is actually t^{1/d_f} . The value of d_f might be imagined as being 2 because it involves an area, leading to $t^{0.5}$. But clusters are fractals with d_f being the fractal dimension, typically 1.75, giving the $t^{0.57}$ value. This is because the hydrodynamic radius of a cluster, R_H , depends on this fractal and the diffusion coefficient, which drives assembly, goes, as is well known, as $D = k_B T / (6\pi\eta R_H)$.

During this first time-scale you might not notice much because the viscosity increase is not very large. Indeed, you might be fooled into thinking that your system is not going to gel. But once percolation has set in, the time-scale for motion of the clusters goes as $\sim t^1$, the big clusters are getting jammed with each other so it's harder for them to grow into larger clusters. The gel is definitely there but it will take a long time to reach its end state.

8.6.3 How they did it

If you are going to do careful measurements of gel formation, you need some well-controlled particles and you need to turn on the gelation mechanism at a precise time. The quoted papers used two variants of the same trick, which is to use a known borderline solvent for the protecting groups, then suddenly change the temperature so the solvent becomes bad.

In one case, you use octadecyl groups in decalin, which switch from "good" to "bad" at 33°C and gives strong gelation at temperatures near 5°C. The particle-particle attraction changes from 0 in the neutral solvent to $\sim 3kT$ near the gel temperature. This is very gentle gelation, though the gel itself is definitely solid!

In the other case the protective polymer is PNIPAM which is happily soluble in its solvent, water, below $\sim 30^\circ\text{C}$ then becomes unhappy (there are furious rows, discussed in my Solubility Science book, about whether it becomes "insoluble", but basically the answer is that it doesn't) as you go over that temperature. To reduce any risk of charge-charge effects you add some salt and for good measure you tweak the density by using an $\text{H}_2\text{O}/\text{D}_2\text{O}$ mix so there is no sedimentation.

³⁴ I found this paper especially clear: Jae Hyung Cho, Roberto Cerbino, and Irmgard Bischofberger, *Emergence of Multiscale Dynamics in Colloidal Gels*, Phys. Rev. Lett. 124, 088005, 2020

Although I tend to get frustrated by the many academic papers that study idealised systems that often provide little of interest to formulators, I relish papers like these which use fancy tricks in idealised systems to work out what is going on within complex systems such as particle gelation.

8.7 So what?

This is it, the end of the last chapter of the book. Where's the satisfying conclusion, where is the masterly summary of all that has gone before?

I admit it, I've failed. This is not the book I had hoped to write.

If the failure was simply because I was too stupid or hadn't tried hard enough, then I'd not publish this version. One can always learn more and one can always try a bit harder. But I think my failure also reflects a failure of 70+ years of rather poor attempts to understand particle behaviour. As far as I can tell, the reason I have not added more apps providing more insight is that there aren't the adequate theories out there to appify.

I had wrongly almost reached that conclusion some months previously, but by persisting, the world of SF theory opened up. Having had the luxury of email encounters with some of the brightest minds in modern SF theory and who are tuned in to modern science trends, it seems to me that there really is nothing else out there that I am missing, unless it's hints of theories that take SF as a starting point and go further.

The answer to the So What? is that a few decades of solubility-based HSP applied to particles is emerging from the shadows to be seen as a valid, pragmatic approach for particle formulation, encouraged by the fact that solubility thinking is now sanctioned via Shimzu and Matubiyasi. Modern centrifuges and NMR techniques are making this impressive approach even more powerful. The fact that HSP are linked nicely to χ parameters makes it astonishing to me that the parallel thread of SF theory, based around χ parameters, has been around, unknown to me and underutilized by the wider community, for a similar number of decades.

Despite the book's many faults, and with chapters that come to abrupt ends, the particle formulation community now has access to a common core of apps covering a broad range of humble but necessary issues along with access to the power of solubility thinking via the dual engines of HSP and SF theory. Although the apps are, for historical reasons, scattered around different parts of my website, I've brought them together into Practical Particles for ease of access.

You might reasonably demand that I produce a better book. I certainly aspire to add a large section on Kirkwood-Buff theory which I think will become a complementary third branch. But maybe that will never happen.

Fixing errors and adding extra apps is a given - if you let me know what's needed and if there's a good-enough theory to be implemented, I'll do it.

But I can make a demand on the readers. If you are an academic with a theory too complicated to be used or a computational methodology too complicated for a laptop, that's very nice but not too helpful. Can you really not do better? Can you really not produce something that the formulation community can use? We fund you because you say that "particles are important for formulations". We don't demand that *every* funded program should result in a user-friendly outcome. But we *do* say that adding to the 10's of 1000's of papers that no one will ever use is not an obviously good use of precious resources.

If, on the other hand, I've missed your important, insightful and usable work then let me know how to describe and appify it. I will apologise to readers of the previous edition and give full credit for your achievements.

So here's a positive note to end on. By tradition, my books come with a guarantee. Here's this book's guarantee: if you can show me how to improve what's already in the book or add new material that makes it more useful for the formulation community, then I will do my utmost to include your improvements for a revised edition, with, of course, full acknowledgement. Yes, it's embarrassing to be proven to be wrong or ignorant, but compared to the excitement of learning something cool and being able to pass it on to others, that's OK by me!

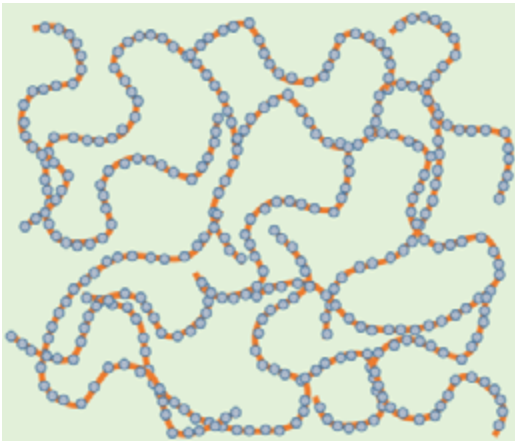
9 Capillary Suspensions

This chapter is an accidental extension of the previous chapter's discussions on non-ergodic, percolating structures. It is accidental because I *should* have known about capillary suspensions (which have been formally known since 2011) when I wrote the book, but only learned of their existence when talking about something entirely different to some of the team behind the ideas.

These suspensions get a whole (short) chapter to themselves because I don't want to rewrite a big chunk of a previous chapter and because they are fascinating in their own right - with spin-offs in things like 3D printing and foams.

9.1 Just a drop of oil (or water)

As we've seen before, a suspension of, say, 10% hydrophobic particles in water isn't anything special. Over time, without dispersants, those particles will agglomerate and settle out, but with some modest shear, they're just a boring suspension with a viscosity little changed from the bulk fluid. The particles are attracted by van der Waals forces which are relatively modest. The same thing applies to hydrophilic particles in oil.



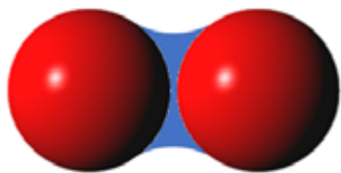
Now add 0.5% of an oil to the aqueous suspension, or 0.5% of water to the oil suspension and the system becomes a gel with a large yield stress, τ_y , and plenty of G' and G'' . Since the original paper from Koos & Willenbacher in 2011³⁵ at KIT this gel has been called a capillary suspension. The yield stress is larger than the original suspension because the capillary forces holding the particles together are orders of magnitude larger.

If, as happened to me in the past, the effect is unintended, the result is an inexplicable disaster. How could such a small amount of water or oil "contaminant" cause such a large change in properties? So understanding the rules of capillary suspensions can help us avoid such problems in the future. Indeed, the original paper notes that the *phenomenon* was well known in, for example, PVC plastisols, but "without a clear understanding" of what was going on.

If, however, you want to produce gels with modest amounts of particles and relatively benign additives, then capillary suspensions open a wide range of opportunities for foods (low fat chocolate spread), novel nanocomposites,

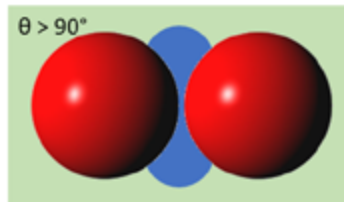
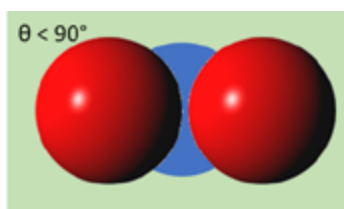
porous ceramics (including 3D printed ones), printable electronics and, as we shall see, foams.

9.2 The basic rules



Most of us are familiar with capillary bridges. Just as a liquid can rise inside a capillary tube driven by capillary attraction, a small amount of a liquid between two particles or fibres will pull them together. Very wet and very dry hair is free flowing, but a bit of water or oil (but not silicone oil which does not wet the hair) causes the hair to stick together via capillary bridges. And on a humid day a powder can stick together via capillary condensation that in turn creates capillary bridges. These are 3-phase systems, though we tend to ignore the fact that air (the white background in the image) is the 3rd phase. For capillary suspensions, the 3rd phase (green in the image below) is the bulk, non-wetting, fluid instead of the air.

To make a reliable capillary suspension, too much or too little of anything is a bad thing. If you are using a hydrophobic particle in water, don't go too wild, because dispersing the particle will be too much of a challenge. The opposite is true for hydrophilic particles in oil - don't go too hydrophilic with the particle or too hydrophobic with the bulk fluid phase.



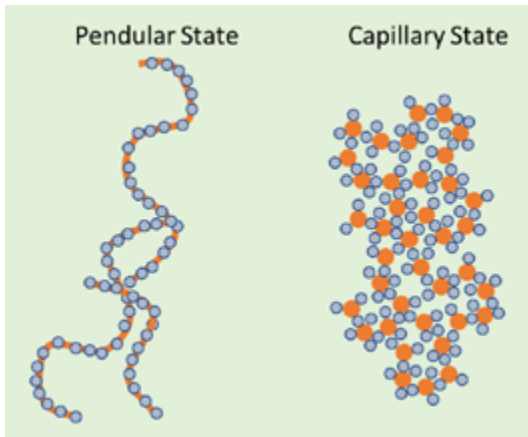
So when you add your secondary fluid, you don't want the contact angle, θ , with the particles to be too high or too low. With $50^\circ < \theta < 85^\circ$, as in the top image where the background is the bulk fluid, there's a reasonable balance, giving a "pendular" state which is what most of us want most of the time. Given that measuring θ is hard and that determined experimentalists can measure capillary suspensions with θ as low as 45° , we don't have clear rules, but you get the general idea.

If the particles wet too well ($\theta < 45^\circ$), or if you add too much secondary fluid then you tend to get a "funicular" state which is rather fluid and uninteresting.

What about the opposite case, the lower of the two images, when they wet badly, with $\theta > 90^\circ$? A well-known formula (I'm not providing it, for reasons the footnote shortly explains) shows that the capillary force scales as $\cos\theta$, which would imply that a non-wetting secondary fluid would *not* form an interesting capillary suspension as $\cos\theta$ becomes negative above 90° . However, the applicability of this formula is debatable even for pendular bridges³⁶ and in any case these "capillary state" capillary suspensions (I don't understand why they chose the

36 "[The equation] was validated only up to 40° in the original publication. Thus, for pendular state suspensions with contact angles larger than 40° , the use of this equation is not thoroughly warranted." Frank Bossler and Erin Koos, *Structure of Particle Networks in Capillary Suspensions with Wetting and Nonwetting Fluids*, Langmuir 2016, 32, 1489–1501

confusing term "capillary" for this form of capillary suspension) don't form via strings of bridges but by particles clustering around a blob of secondary fluid.



Why does the capillary state form when (except for a minor surface tension attraction) there is a net repulsion from the particle? The *thermodynamic* answer is that it doesn't - the stable form is phase separated. The *kinetic* answer is that the small drops are shielded from the thermodynamic state by their contact with the particles. It is a local minimum where separation from the particles to join up with another drop is energetically unfavourable.

The original literature (not surprising when there was so much new and complicated) tended to suggest that capillary state suspensions were stronger than pendular ones. But now it is accepted that a pendular formulation is more likely to lead to a higher yield stress, partly because of the net attractive capillary bridges and partly because, as discussed later, it is trickier to get a reliable capillary state formulation.

Because the capillary bridges are formed when the secondary fluid is incompatible with the bulk fluid, you want a large interfacial tension, Γ , between them, typically > 20 mN/m. If, for example, you add some surfactant to the system, reducing Γ , the suspension falls apart.

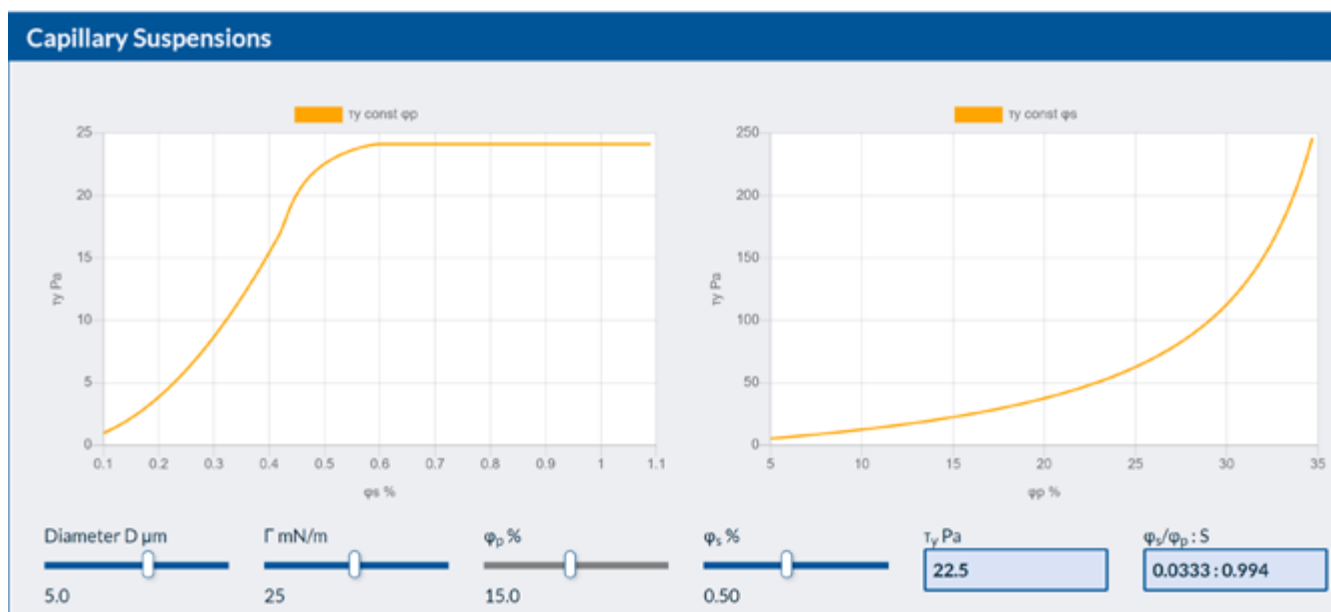
However, for easy, reliable formation of a suspension, you want the secondary fluid to be present as small drops, comparable in size with the particles, especially for the capillary state where the particles need to cluster around drops of comparable size. In general, the larger Γ , the harder it is to get the secondary phase to form small drops, a topic to be discussed later.

And because, for the same volume fraction of particles, you get more capillary bridges if the particles are smaller, the strength of the suspension (e.g. measured by yield stress) increases with smaller radius R .

The more particles, the more bridges, so yield stress increases with ϕ_p the volume fraction of particles. Once you are above $\phi_p=28\%$ then you are in the domain of conventional isotropic percolation theory and we start to lose interest in capillary suspensions.

However, increasing the fraction, ϕ_s , of the secondary fluid quickly becomes futile - once you've got a tiny capillary bridge between most of the particles, extra fluid achieves nothing and you go more and more into the funicular state.

Frustratingly, although there is a vast literature on all the relevant individual effects³⁷, there is no standard way to bring together all the factors in an app. Happily, a set of correlations from Orpe's team at CSIR, Pune³⁸, makes it possible to create an app that captures, surprisingly well, the essence of what's going on:



App 9-1 <https://www.stevenabbott.co.uk/practical-rheology/Capillary-Suspensions.php>

Just looking at the sliders and the boxes, a 15% dispersion of 5 μm particles in a fluid where the interfacial tension with the secondary fluid is 25 mN/m becomes a gel with a 22.5 Pa yield stress, τ_y with only 0.5% of that fluid, a volume ratio ϕ_s/ϕ_p of 0.05 and a secondary fluid fraction (compared to total fluid) of 0.994.

Looking at the graphs, you see that for the fixed ϕ_p , increasing ϕ_s above 0.5% produces no increase in yield stress. But for the fixed ϕ_s , there's a continuing, and accelerating increase in yield stress as ϕ_p is increased. In academic literature these graphs are usually plotted with a log scale which makes them linear. Because we formulate in a linear world, I prefer to see the real trend.

As discussed above, increasing Γ and decreasing $R = D/2$ produces an increase in yields stress, i.e. $\tau_y \sim \Gamma/R$.

The equations used are shown on the app page but aren't as important as the core take-home messages about the trends (if not the absolute values) in the effects. The value of ϕ_s at which no further increase in strength is obtained is

37 Sebastian Bindgen, Jens Allard and Erin Koos, *The behavior of capillary suspensions at diverse length scales: From single capillary bridges to bulk*, Current Opinion in Colloid & Interface Science 2022, 58:101557

38 Sameer Huprikar, Saurabh Usgaonkar, Ashish K. Lele, Ashish V. Orpe, *Microstructure and yielding of capillary force induced gel*, Rheologica Acta (2020) 59:291–306

fixed in the app at 0.5%. The literature suggests that the limit value depends on the relative sizes of the particles and secondary drops, but as most of us have little idea of our drop size, and because practical formulations tend to be in the 0.3-1% range, just treat the 0.5% as "illustrative" rather than definitive.

9.3 It's not so easy

So far it all sounds very easy. But like so many things, there's a sort of reverse guarantee. If you *don't* follow the above ideas, you won't get a stable capillary suspension. But even if you *do* follow the rules, you may still have trouble. And there's a good reason why that is so.

Suppose I have my suspension of particles that are, because that's how all this works, unhappy in the bulk fluid. It doesn't take much for them to decide to agglomerate and, as we all know, agglomerated particles can sometimes be frustratingly hard to re-disperse. If you choose to use great dispersants then the outer portion of the particles will be so comfortable in the fluid that they may not wish to attract the secondary fluid.

But let's go with our nice suspension. Now add the incompatible secondary fluid. This gives the problem that this incompatible fluid might just want to float around as big blobs, unaware of the opportunity to create capillary bridges or, if this is to be a capillary state, in the thermodynamic global minimum rather than the kinetic local minimum.

No problem, we just add lots of shear to the system to break up the secondary fluid. But high shear can readily break up nascent capillary suspensions so the strategy might not work. It turns out that pendular suspensions tend to fall apart with high shear (their floppy, more linear networks are ripped apart) while the polyhedral (we'll get to this later) capillary state is more robust to shear, and can, in any case, only be created if the shear can create small secondary drops.

OK, so let's do it the other way round. Disperse, under high shear, the secondary fluid in the bulk then, at lower shear, add the particles. Although there are potential flaws with this idea, it turns out to be generally preferred in practice. However, if you want a really solid system it might get increasingly difficult to add the extra particles for a larger ϕ_p .

If you pre-disperse and are slow to add the particles, the benefits disappear. You can potentially fix this by making the bulk fluid viscous to slow down droplet agglomeration, giving more time to add the particles. But viscous bulk fluids bring their own problems and the physics of dispersion shows that if the viscosity ratio $\eta_{\text{secondary}}/\eta_{\text{bulk}} < 0.1$ it gets harder to disperse, as described in <https://www.stevenabbott.co.uk/practical-surfactants/critcap.php>. If you happen to want to disperse a viscous second phase in a less viscous bulk phase then you have more severe problems because the same app shows that conventional high

shear mixing cannot create a good dispersion if $\eta_{\text{secondary}}/\eta_{\text{bulk}} > 4$. As the app shows, the easiest dispersion is with matched viscosities.

There are more complications. So far we have assumed nice uniform, solid spheres with wetting behaviour and contact angles that are under good control. Many of the systems we would choose to use might be non-spherical and/or porous. Now defining contact angles becomes more difficult because, as is well known, wetting depends on roughness, with large differences between advancing and receding angles. Moreover, the definition of φ_s becomes tricky if some of the secondary fluid is absorbed into a porous particle.

With expert teams such as those in the Koos lab in U Leuven, it is possible to spend a year looking in detail at just one system to see what's going on during methodical changes to the fluids, the hydrophobicity, the radius and the roughness of the particles. With measured contact angles, confocal microscopes to look at fluorescent oils/water³⁹ revealing the capillary bridges and distinguishing between pendular and capillary systems, then sophisticated rheology of the resulting formulations, it is possible to see how all this comes together in terms of the properties of the formulation.

Most of us don't have the time and resources, so we need to absorb the key lessons from all that work and try to apply it to our own far-from-perfect setups. So let's sort out some tips and tricks.

9.4 Tips and Tricks

There is a logical way to approach the challenge of creating capillary suspensions. Do it step by step.

9.4.1 Understanding the secondary fluid

If you can't readily disperse your secondary fluid in the first, then you won't make good capillary suspensions. The opposite is also true, if it is easy to add to the bulk then it might be too compatible. So do a set of simple tests:

- Add 0.1%, 0.2% ... to the bulk fluid. If the small additions disappear into solution, the fluids are probably too compatible.
- Is the viscosity of the secondary fluid 5x or 10x higher than the bulk? If so then, as mentioned above, standard dispersers won't be able to form fine drops. Either give up or use a disperser that provides extension rather than shear.

³⁹ You have to match the refractive index of your fluids to your particles to make them invisible, add just the right amount of fluorescent dye to the secondary phase then have a great confocal microscope to see the capillary bridges for pendular or drops for capillary suspensions. You then change the focal plane and reconstruct a full 3D image. Without all this, much about these suspensions would have remained unclear. The images in the papers are awesome.

- Can you observe some nice scattering from a quick wizz with a disperser? That's a good start. Does the secondary phase separate out in a few seconds? Life is going to be tough.
- Can you tweak the properties of the secondary phase in a rational manner? If secondary phase water is too incompatible with your oil, then a water/glycerol mix or water/propylene glycol mix might get you into a sweet spot. If the secondary phase is, say, decyl decanoate, you can tune its incompatibility with the water phase via the vast variety of x-yl y-anoate oils, where x and y can be anything from 1 to 20. Or you just make a blend across a range of mineral oils.
- Similarly, can you tweak the properties of the bulk phase via similar tricks?

9.4.2 Understanding the particles

Now some quick tests with the particles.

- Are they highly porous? Probably not a great thing to use if you aren't already expert. The secondary fluid will have many complex interactions, maybe inside the pores and certainly in terms of complex advancing and receding contact angles.
- Do the particles have a large aspect ratio - highly elliptical or flake-like? These, as explained in <https://www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php> will show earlier onset of high viscosity and yield stress so it is important to do ϕ_p tests over your desired range without any secondary phase to get a feel for when these effects kick in. By playing with the "fractal dimension" and "N" sliders in the app, you'll get a feel for what small amounts of secondary fluid might do if you are already in a fairly viscous zone.
- Do a set of tests at ϕ_p of 0.05, 0.1, 0.15 ... in your bulk fluid. If the resulting dispersions are stable indefinitely, then the particle is too compatible. If the particles phase separate rapidly, creating practical capillary suspensions is going to be hard.
- Think very carefully about dispersants for your particles. If they are solidly locked on then the dispersion properties are controlled by whatever portion of the dispersant is somewhat compatible with your bulk fluid. But, of course, it mustn't be too compatible or else the system won't work because the wetting angle will be too small. If the dispersant is a typical surfactant then it might greatly reduce the interfacial tension, Γ , between the two fluids and, as theory suggests and experiment confirms, the system will fail. In any case, any free polymeric dispersant might cause depletion flocculation, giving a more complex system. In general, therefore, sort out your particle with some sort of tuned, pre-reacted surface treatment to avoid several complications.
- Tweaking the bulk phase can help. If the particles cream or settle too quickly, density matching might help and, as with tuning to the secondary

phase, the bulk hydrophobicity/hydrophilicity can be altered rationally - though tuning for both the secondary phase and the particles sounds complicated.

9.4.3 Creating the suspension

With the pre-knowledge from the experiments on the individual components you should:

- Pre-disperse whichever is the more difficult (secondary fluid or particle) and add the other phase to it.
- Even better, if you can manage it, is to pre-disperse each using, say, a stirrer for one and a disperser for the other, then combining them to their final composition. This requires each to be made at a higher concentration, but with your exploratory experiments it shouldn't be too hard to make each at, say, 2x the concentration and create a 1:1 blend.
- If you have a choice, it is better to build up ϕ_p than ϕ_s because the increase in viscosity and yield stress from higher particle concentrations is dramatic.
- *Do all this with air-free mixers/dispersers!* As has been often observed, it is easy to get air into these suspensions and hard to get it out again. Air gets in the way of any rational measurements of rheology and/or visual inspection of the suspension. This advice isn't so hard for production-scale equipment because in-line circulating dispersers are naturally air-free. Bench-top air-free mixers/dispersers are less common, but they do exist. If you are serious about capillary suspensions then sorting out the bench-top air-free mixing should be a priority before starting. For example, vacuum dispersers allow you to mix with relatively little air present - as long as you don't have volatile solvents. Although it is easy to get recirculating inline dispersers for pilot scale and production, bench-top ones are somewhat less common.
- If you can't manage to get the suspension perfect via dispersers, you can always take an OK suspension and try ball milling to improve its properties. This is common in the academic literature but is not ideal if you want to scale up the process.
- *Alternatively, work out how to get maximum air into the mix!* Now you are creating a capillary foam, discussed below.

9.4.4 What can go wrong?

You are always fighting against these effects:

- The secondary fluid re-agglomerates into big blobs
- The secondary fluid forms a stable emulsion protected by the particles - i.e. it forms a Pickering emulsion

- Alternatively, the secondary fluid is dispersed into drops much finer than the particles so they merely "decorate" the particle rather than form a capillary bridge.
- There's insufficient shear for the particles to break up enough to then reform via capillary bridges or ...
- ... there's too much shear forcing the capillary bridges to fracture, spraying out ultra-fine secondary fluid which ends up just decorating the particles.

9.4.5 How do you know what you've got?

If 0.5% of secondary fluid reliably take your 10% particle suspension to a stable, viscous suspension with the short- and long-term properties you like, well, the chances are that you have a great capillary suspension. Whether it is pendular or capillary hardly matters. A few images under an optical microscope will be a good record to have.

My reading of the literature is that it is so hard to really work out where the secondary fluid is and precisely what it's doing, that attempting to diagnose a failed formulation via direct observation is an inefficient approach. You have probably had warnings from the preliminary experiments that one or more aspects of your system is borderline, so take the failure as proof that you are at just the wrong side of borderline and tweak things with, say, the particle surface treatment or the solubility/miscibility of the secondary phase to shift things in the right direction.

9.5 Making porous solids

With 25% of 1 μ m particles in an aqueous suspension and 1% of an oil you've just made a near solid suspension. If you could just get rid of the water, while maintaining the current structure (you'd want the oil to stay in place longer, to keep the bridges), you would have a somewhat delicate porous solid - a "green" solid ready for making properly strong via sintering.

Similarly, you can have 1% of water with the 25% particles in an oil suspension, so by evaporating the oil while keeping the water bridges, you will have a reasonable green solid.

With some sintering to form solid inter-particle bridges you now have a highly porous solid. Because these suspensions can be readily extruded, you can even 3D print highly porous structures.

There are some big "ifs" in the previous paragraphs. Removing the bulk water (for oil bridged) or oil (for water bridged) will, in general, have a devastating effect on the structure. Most naive attempts will be inglorious failures.

The trick is to allow the bridges to phase change from liquid to solid. Then the porous structure has higher green strength while the bulk phase is removed and during subsequent sintering.

Many possibilities are open to you. For example:

- Use a wax as the bridge material, form the suspension above its MPt, then evaporate the water at a lower temperature.
- Use a UV-curable oil (an acrylate) as the bridge material and do a UV exposure once the solid is formed.
- Use a conventional radical-cure polymer and raise the temperature to both create the initiators for the reaction and to help evaporate the bulk phase
- Use a sugar solution for the aqueous bridge. As the water evaporates, the bridge becomes solid. Yes, this works well in the literature.
- Use a hydrogel solution for the aqueous bridge. Depending on the polymer the bridge becomes solid at higher or lower temperature.
- Include, say, 25% nanoparticles in the secondary phase⁴⁰. As the secondary liquid evaporates the bridges become solid via the nanoparticles which can sinter nicely into the whole ceramic structure.
- Include a wax in an oil bulk phase. As some of the oil evaporates, the wax becomes less soluble so on cooling to below its MPt there is some bulk strength as well as bridge strength - this can help the green strength for handling before the sintering process.

With yield stresses in the region of 2000 Pa not only can you cast complex 3D structures, you can also 3D print (DIW, Direct Ink Write) them⁴¹ with a suitable extrusion nozzle in the few 100 µm diameter range.

In practice, the shrinkage on drying then sintering can impose significant stresses on the system so it needs some smart tricks to deal with them. In the cited papers, for example, they print the oil bulk phase material onto a polyethylene oxide, PEO, film surface. When the print is completed, the PEO film is wetted to become semi-liquid, allowing the structure to shrink in the x-y plane as it dries.

9.6 Percolated polyhedra

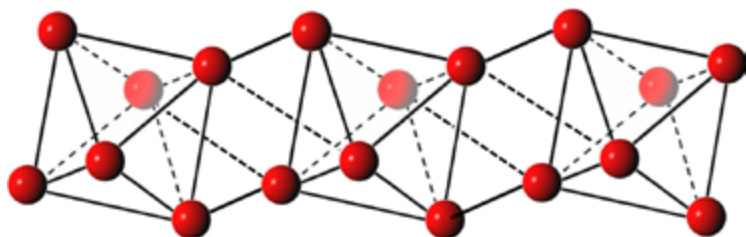
Feel free to ignore this section. The strange title is intended as a challenge in the spirit of the section *From ergodic to icosahedral via isostatic* in the previous chapter. The reason for this section is that most diagrams produced by most

40 Moritz Weiß, Patrick Sälzler, Norbert Willenbacher, Erin Koos, *3D-Printed lightweight ceramics using capillary suspensions with incorporated nanoparticles*, J. European Ceramic Society, 40, (2020), 3140–3147

41 Johannes Maurath, Norbert Willenbacher, *3D printing of open-porous cellular ceramics with high specific strength*, J. European Ceramic Society 37 (2017) 4833–4842

of us showing capillary suspensions are bogus. They don't capture the reason these structures are so strong for such little material.

In the early days of capillary suspensions it wasn't at all clear what was going on. With a combination of theory and experiment a nice picture emerged⁴². First of all, these suspensions might not be thermodynamically stable - phase separation is the preferred state. For the *capillary* state this thermodynamic instability is clear because the contact angle $> 90^\circ$ means a repulsive wetting force with only a small surface tension force remaining. In practice many *pendular* state suspensions seem to be readily transformed into other structures so their fate is also kinetic. They are *kinetically* stable because the secondary fluid is shielded from the primary fluid by contact with the particles. This immediately tells us that things can (and, in practice, do) go horribly wrong when trying to make these suspensions: if the secondary fluid or the particles self-aggregate before the suspension is formed, it can be hard to force things into the required small-drop state.



If things go right, then the secondary drops like to get shielded by modest numbers of particles. It turns out that an efficient way to do this is by building up tetrahedral and octahedral chains.

The diagram tries to capture the idea with 3 octahedra, though obviously the spacings of the particles are not representative of the real case where the particles are in near contact via the bridging fluid. Nor, of course, are real structures this perfect: they are *statistical* octahedra

Fortunately, building up large 3D *clumps* is entropically unfavourable - it is the *chains* that are the best balance of shielding and entropy. These polyhedral chains then create a percolated structure, using, as before, the definition of percolated as *when each particle has a path to every other particle*, rather than the more usual isotropic percolation which is *when there is any single path between particles across the bulk*. Isotropic percolation of spheres requires 28% particles while the non-ergodic (reminder - this means that any given particle is *not* free to be at any place) polyhedral percolation kicks in even at 5% particles if the polyhedral chains (initially tetrahedral then more octahedral) are able to form via the correct balance of capillary linkages.

Real structures aren't the neat chains shown in the diagram. But equally they are not a bunch of beads on a chain which is the diagram used earlier simply

42 Erin Koos and Norbert Willenbacher, *Particle configurations and gelation in capillary suspensions*, *Soft Matter*, 2012, 8, 3988

to capture the general idea of the secondary fluid creating capillary bridges between neighbouring particles. The idea of these non-ergodic, polyhedral, percolated structures is unfamiliar to many in the particle community and a wider familiarity will surely lead to a better grasp of what is going on in many of our formulations.

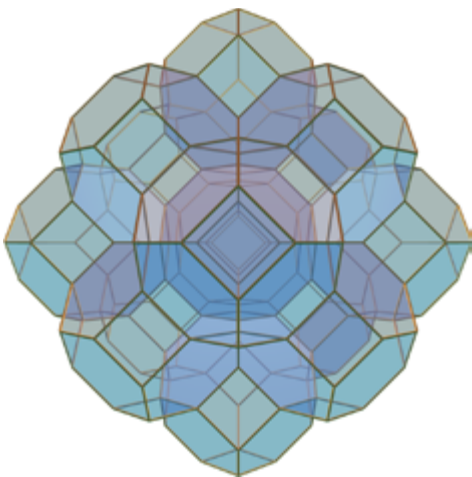
Finally, an alternative way for the high contact angle secondary phase to reach some sort of stability is for each drop to be protected from the bulk fluid by a coating of particles, i.e. it forms a Pickering emulsion. Those trying to make capillary state suspensions often find themselves making uninteresting Pickering emulsions instead⁴³.

9.7 Capillary foams

Any time you get air into a capillary suspension, getting it out is a real nuisance. So what if you deliberately put a *lot* of air into a capillary suspension?

If you put, say, 65% air then you have a Kugelschaum, a "spherical foam", i.e. one where the individual bubbles remain as bubbles. In normal fluids a Kugelschaum is unstable, because the bubbles easily rise to the top. In a capillary suspension with a large yield stress, even a Kugelschaum is stable.

What is more interesting is when you have, say, 80% air. Now the bubbles are pressed against each other and you have a conventional foam (Polyhedral foam).

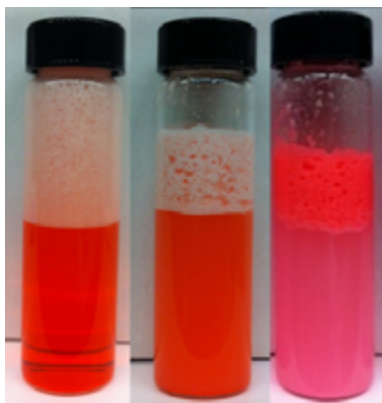


These are surprisingly hard to visualise and this 3D view (you can explore it interactively in the app at <https://www.stevenabbott.co.uk/practical-surfactants/Foam-3D.php>) of a portion (15 tetrakaidecahedral bubbles) of an idealised Kelvin foam shows how you have thin cell walls, some hexagonal, some square, relatively thick borders between the walls and nodes where the borders connect.

In normal foams, the borders are a problem because they are like drainpipes through which the foam fluid can flow, causing the bubbles to "drain", starving the walls of fluid so they get thin enough to burst. With capillary foams, the viscosity is so high that drainage is not a problem.

43 Frank Bossler, Lydia Weyrauch, Robert Schmidt, Erin Koos, *Influence of mixing conditions on the rheological properties and structure of capillary suspensions*, Colloids and Surfaces A: Physicochem. Eng. Aspects 518 (2017) 85–97

More importantly, the cell walls are stabilized by the capillary bridges between the particles. For those familiar with Pickering foams, created via a large fraction of particles, the idea of a foam wall made tough via particles is natural. They also know that it is rather hard to create Pickering foams as the required balance of properties is tricky to organise. As shown by Behrens and Meredith from Georgia Inst. Tech.^{44,45}, assuming you can create a capillary suspension it is relatively easy to generate a capillary foam, provided you can deliver sufficient mixing energy to get the air into the suspension.



One delightful attribute of these foams when in a bulk aqueous phase is that the oil used to form the capillary bridges can contain a dye. Because of the strong partitioning to the air interface, the oil and particles are extracted from the bulk fluid which loses much of its original colour, while the foam itself is brightly coloured, as you see in this image courtesy of Behrens and Meredith. The tube on the left is a conventional foam, in the middle it is a Pickering foam; the capillary foam is on the right.

Classically, air-in-oil foams are hard to create, but starting with a capillary suspension, they are no more difficult than an aqueous foam. This creates significant opportunities for new types of foam.

The practical problem in creating these foams is that foam creation requires air to get into the formulation and then to be broken into smaller bubbles. As mentioned above, it is famously hard to get low viscosity fluids to mix efficiently into high viscosity fluids ($\eta_{\text{secondary}}/\eta_{\text{bulk}} < 0.001$). If for capillary foams you use powerful dispersers, there is the strong possibility of destroying the capillary suspension itself, as noted earlier.

A logical alternative is to generate the bubbles from inside the formulation via the same mechanism that gives champagne its bubbles: use a CAF (Compressed Air Foam) - though for champagne this is a Compressed CO₂ Foam. Dissolving air under high pressure into the capillary suspension, followed by rapid release, will generate many small bubbles (there are plenty of seeds available thanks to the particles) with a high expansion ratio giving a low density foam.

44 Yi Zhang, Michael C. Allen, Ruiyang Zhao, Dimitri D. Deheyn, Sven H. Behrens, and J. Carson Meredith, *Capillary Foams: Stabilization and Functionalization of Porous Liquids and Solids*, Langmuir 2015, 31, 2669–2676

45 Omotola Okesanjo, J. Carson Meredith, and Sven H. Behrens, *Structure–Property Relationship in Capillary Foams*, Langmuir 2021, 37, 10510–10520

9.8 Summary

Capillary suspensions can, as many of us have found, be a nuisance - 0.5% of water in oil or oil in water can ruin your formulation. So knowing about them is important for debugging unwelcome surprises.

On the positive side a new set of opportunities is open to us to create imaginative semi-solid formulations with small amounts of particles via a process that is, at least in principle, relatively simple.

The academic work has, of necessity, required the sorts of analytical equipment and access to computational resources that most of us won't have. But after a decade of such work we can see that these suspensions are part of a particle suspension space that includes classic higher-solid suspensions, Pickering emulsions, looser suspensions with larger amounts of secondary fluids, normal flocculated systems and plain phase separated failures. The key to getting capillary suspensions is to avoid extremes of incompatibility or compatibility between the bulk phase and the particles and secondary fluid, and to avoid the need for extremes of dispersion energy. It is those extremes which take us into one of those undesired regions.

As we start to use these for ourselves, we need to focus on modest incompatibility that allows dispersion of the particles and secondary fluid without too much effort, giving the suspension the chance to form without too much drama, and without the need for extra steps such as bead milling. For this, small, smooth, spherical particles provide the fewest surprises.

- Small. Because of the Γ/R dependence, smaller particles more easily give stronger suspensions at smaller ϕ_p values, reducing complications.
- Smooth. Because smooth particles have a simpler wetting behaviour and fewer of the particle-particle interactions that could move them into unwanted zones of particle suspension space.
- Spherical. Because spherical particles have a larger safe zone, while particles with high aspect ratios can readily agglomerate for other reasons.

The fact that these suspensions have provided access to a new type of foam opens up new types of collaboration between the foam world and the particle world. While Pickering foams have their uses, capillary foams offer many more opportunities. The ability to bring extra ingredients into the foam itself, as shown delightfully by the coloured foams, is a significant bonus feature.

And, finally, an apology. This book has often despaired about the poor quality of particle science and the theories that have traditionally been applied. Capillary suspensions and foams have been developed by smart people using good theories that are backed up by high-quality data. I *should* have known about them, but I didn't, so I offer this chapter as an apology for my ignorance.