NANOCOATINGS: PRINCIPLES and PRACTICE

From Research to Production

Steven Abbott
Nigel Holmes

NANOCOATINGS: PRINCIPLES *and* PRACTICE

HOW TO ORDER THIS BOOK

BY PHONE: 877-500-4337 or 717-290-1660, 9AM–5PM Eastern Time

BY FAX: 717-509-6100

BY MAIL: Order Department
DEStech Publications, Inc.
439 North Duke Street
Lancaster, PA 17602, U.S.A.

BY CREDIT CARD: American Express, VISA, MasterCard, Discover

BY WWW SITE: http://www.destechpub.com

NANOCOATINGS: PRINCIPLES *and* PRACTICE

From Research to Production

Steven Abbott, Ph.D.
Nigel Holmes, Ph.D.

DES*tech* Publications, Inc.

Nanocoatings: Principles and Practice

DEStech Publications, Inc.
439 North Duke Street
Lancaster, Pennsylvania 17602 U.S.A.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Main entry under title:
Nanocoatings: Principles and Practice—From Research to Production

A DEStech Publications book
Bibliography: p.
Includes index p. 323

Library of Congress Control Number: 2013939360
ISBN No. 978-1-60595-090-7

Table of Contents

Preface

ALTHOUGH the promise of nano is great and the market for nano products is large, the reality is that a miniscule fraction of nanohype has made it into sellable products [1]. Sunscreens and wood coatings are admirable examples of the promise of nano, yet they are relatively humble applications.

A bright PhD student might generate an exciting new nanoapplication during a year's work in an academic lab. Getting that application to market will then typically take ten times as long, though the majority of ideas never make it beyond a few publications in the vast nanoliterature.

This book could be filled with wonderful speculative nanopossibilities. Instead it concentrates on those aspects of nanocoatings that have a realistic chance of being produced in thousands to millions of square meters per year. If you want to read about fascinating ideas that might result in square centimeters of product, this is not the book for you.

The aim of this book, then, is to boost the chances of getting functionalized nanocoatings to large markets. Although sales and marketing issues are vital, they are not the authors' focus. Instead, the basic notion is that even if the market is there, new products generally fail because of some minor technical detail in the path from raw materials through to the final stage of production. This detail might (sadly) invalidate the whole process. Or (more usually) it delays the product launch or increases costs sufficiently for the whole opportunity to be missed.

To those who focus on production issues, it is often a matter of astonishment that the lab people worked with a nanoparticle supplier who simply can't scale up from grams to tons. To those who focus on lab steps, it is often a matter of astonishment that the production people

can't take the beautiful lab-proven formulation and crank it out at high yield and low cost.

This book is an attempt to allow everyone in the chain from lab to production to see what the issues are at all stages of the chain. Using the above examples, once the production people understand why a batch of nanoparticles might suddenly become an unrecoverable gel, and the lab people understand why a minor problem in a 20 μm coating becomes a major problem in a 10 μm coating, the whole team can work more effectively to avoid the problems or to find creative solutions if the problems appear.

To achieve this goal, every chapter has to be more-or-less comprehensible to everyone on the team. Yet, each chapter must be deep enough to provide the relevant specialists with the means to master the complexities of that part of the process. Experience shows that a lot of the 'depth' provided in the literature is profound but worthless. An exquisite and accurate formula might be provided from which the reader can take no actionable items. It may be intellectually satisfying to know that a property of crucial importance can be calculated via a triple integral over phase space, but if there is no practical way for that integral to be realized, the project is not going to be much advanced by that knowledge.

Hence we have grounded the book in equations and tools that members of the team can use in spreadsheets (downloadable from the book's website) or in software that is either free (on-line apps) or affordable. Even if other team members don't fully understand the formulae and models, the fact that they can be discussed live in a team meeting is important. Playing "what ifs" is a vital part of a team's work. If a 20% improvement in product properties comes from a 50% reduction in coating thickness but produces a 100% increase in pinholes, the whole team needs to see why, other things being equal, the reduction in thickness has such a big impact on pinholes. If that 20% improvement is a "nice to have" then maybe it can be abandoned. If it is a "must have" then the other factors affecting pinholes will have to be attended to. These issues aren't somebody else's problem—they are a matter for the whole team, even if eventually the resolution of that issue becomes the responsibility of one member.

Readers should be aware that one of us (Abbott) has helped create two of the commercial software packages used in this book, so there is a risk of commercial hype. All he can say in defense is that, if these tools weren't of use, then either he wouldn't be using them himself (but he is) or he would improve them so they became more useful (and he has).

At times a note of world-weariness will enter into the book. Both authors have been there, done it and have the T-shirts and the battle scars. They have both struggled with formulations in the lab and have spent many long days and nights on production lines nurturing new products through to (sometimes!) successful launches. Many of the common mistakes highlighted in the book are included because the authors have made those same mistakes and have met others who have made them, too. By clearly bringing out the downsides, they hope that your team will either say "Hey, we forgot about that—it's a show stopper, let's try something else" or "Hey, we forgot about that, but we have a brilliant way around it that our competitors won't figure out, let's go for it". Stopping a program for a good reason is no shame—indeed, it should be celebrated—as the alternative is usually a program that drags on fruitlessly. For an academic lab, the typical misjudgements about nanotechnology might lead to a few months' delay in finishing a PhD or writing a paper—painful but not usually serious. For an industrial lab, a mistake can be the difference between sales of $10m and sales of $0, with knock-on consequences for careers and employment.

Driving the authors' enthusiasm for writing the book is that wonderful feeling that comes from knowing that a lot of people are spending lots of money to buy a product that a team has worked hard to produce. There is a lot of competition out there; there are people with more money or better access to the market or a potentially better idea. For those who go along with nanohype and commit the usual mistakes, there is almost a guarantee that the competitors' advantages will win through. For those who work as a team and understand in advance the cross-functional difficulties of bringing a nano-based product to market, the chances of beating the competition have become much greater.

The word "team" is used extensively through this book. Such usage is not intended as mere pandering to modern sensibilities. Both authors have previously worked in environments where there was a high wall between research and production. Products were thrown over the wall and blame for failure to produce (or, if we add the Marketing wall, sell) the material was thrown back. When organizations have become serious about working as a team, life becomes productive and more pleasant. All sides began to appreciate how a "trivial" issue may not be so trivial after all. Solving that issue early on is not so hard. Coming as a complete surprise late in the project can often cause a huge, expensive delay.

References

The authors have made their best efforts to keep abreast of the latest developments in the topics under discussion, but freely admit that this is not easy. The number of scientific papers being published worldwide on nanotechnology is increasing yearly. Add to this company technical presentations and marketing pronouncements and the size of the task becomes apparent. We have exercised our best efforts to separate wheat from chaff. In chapters involving many well-known, standard equations, no attempt is made to provide references to them.

It is also useful to delve into the past (deeply in some cases). There are many useful papers pertaining to nanotechnology published before the term was invented and it is useful to bring this older and still pertinent work to light. Perhaps these papers should be dated BF (Before Feynman).

What the authors suspect, but cannot prove, is that there is much proprietary knowledge that is not present in the public domain. It is a little frustrating as it is this hard-won knowledge that is the core purpose of this book. Where such knowledge has been freely provided, the authors are doubly grateful.

A Note on Units

It is tedious for the authors, and for the readers in > 99% of the countries of the world, to have two sets of units—metric and US, so only metric units will be used. It is unhelpful to use uncommon, but "official" SI units when 99% of users choose "common" metric units. It is confusing to switch between different units for different parts of the book. Therefore, temperatures will only be in °C (not °F or °K), viscosities will only be in cP (not Pa.s), surface tensions will be in dyne/cm (not N/m), densities will be in g/cc (not kg/m^3), and thin coatings will be in μm (microns) or nm (nanometers), not in mils or meters.

One of the big barriers to trying out any formula comes from the need to convert from common units into SI for the calculation and back to common units for the answer. It is surprisingly easy to be out by many orders of magnitude through some simple error in unit conversion. The spreadsheets provided with the book all provide inputs and outputs in common units, with explicit conversion to SI units for the calculations.

Common Abbreviations

To avoid tedious repetition each time they appear, here are the abbreviations we felt necessary to use throughout the book:

Physics

DLVO	Theory to explain some aspects of particle (in)stability. Named after Derjaguin, Landau, Verwey and Overbeek.
kT	Boltzmann energy, i.e. Boltzmann constant, *k* times *T*, absolute temperature
RI	Refractive Index
RT	The same thing as kT but with different units via the universal gas constant, *R*

Polymers

EVOH	Copolymer of Ethylene and PVOH—polyvinylalcohol
PDMS	PolyDiMethylSiloxane
PE	Polyethylene
PET	Polyethylene Terephthalate—Mylar, Melinex, etc.
PLA	Poly Lactic Acid
Tg	Glass transition temperature of a polymer—the onset of main-chain segmental motion

General

CNT	Carbon NanoTubes
HSP	Hansen Solubility Parameters
MEK	Methyl Ethyl Ketone solvent
RER	Relative Evaporation Rate
SHE	Safety Health & Environment. Catch-all phrase to encompass interlocking aspects of how we work responsibly.
THF	Tetra Hydro Furan solvent
UFP	Ultra-Fine Particles
UV	Ultra-Violet radiation

Choice of Topics

There is only so much space available in a book. The authors had to select what seemed to them to be important within a range of generally "wet" nanocoatings (thereby excluding vacuum coating) and with a bias towards coating onto surfaces other than metals (thereby exclud-

ing the vast topic of specialized corrosion inhibitors). The criterion was the answer to the question: "What do we wish we'd known in advance rather than discovering the hard way?"

Writing the book and, especially, creating the spreadsheets has provided a wealth of resources that are proving useful in the authors' daily life within an excellent coating company (Holmes) and as a consultant (Abbott).

Figures

All figures, except when explicitly acknowledged, are created by the authors, including screen shots from HSPiP and TopCoat. Most graphs (except for calculations too complex to be included) are taken from the spreadsheet provided with the book—this means that readers can confirm the calculations themselves, then play with the input parameters to obtain a better understanding of what is going on.

Downloading Active EXCEL Spreadsheets for Formulas and Equations

As noted on page x, many equations and calculations in this book are available on active EXCEL spreadsheets. These spreadsheets can be downloaded from the book's page on the publisher's web site, which is www.destechpub.com.

Thanks

The authors are most grateful to the individuals and groups that provided the Case Studies throughout the book.

Thanks, too, to the wonderful support from our excellent editor, Joseph Eckenrode.

Finally, thanks to Dr Sem Seaborne for specialist advice on the vast subject of nanosafety and to our non-scientist readers, Anne Abbott, Fran Holmes and Francesca Lacey who were able to recommend so many ways to make the text more readable.

Reference

1. David M. Berube. (2006) *Nano-Hype: The Truth Behind the Nanotechnology Buzz.* Amherst, NY: Prometheus Books.

CHAPTER 1

Nano or Not?

MAKING a functionalized nanocoating seems straightforward to those who have not tried it themselves. The key step is to create some brilliant new nanoparticle which has merely to be formulated and coated, ready to be sold to an astonished and grateful market. Countless researchers in countless labs are focusing on these brilliant new nanoparticles. The rest is "mere production". In reality, the outcome of this approach is generally a few grams or milliliters of an unstable formulation which produces a few small samples showing an interesting effect, followed by a few papers, then silence.

This book is about reducing the chances that those bright ideas and hard work will end in oblivion. This chapter aims to help in that task by pointing out all the reasons why a nano project is likely to fail.

1.1. QUESTIONS AND CHAPTERS

A series of questions provides a structure for thinking through the key issues. If you don't have a good answer (or a good way to approach the answer) to one or more of the questions then it may well be a good idea to stay away from nano.

It is important that everyone involved in the project team helps to think through each of the questions. Although different members will take different roles in answering the questions, everyone has to understand why the answers are either positive or negative.

Each question forms the basis for a subsequent chapter. Everyone in the team should have at least a passing knowledge of the contents of each chapter so that they have some familiarity with the sorts of issues

their expert colleague is facing. Sometimes the solution to a problem in one area is best found in another area and a "dumb" question from a non-expert can often turn out to be the key to solving a problem.

1.2. WHY GO NANO?

For a few inglorious years, the answer to the question was obvious: nano is good, nano is exciting, and we've got to have nano in order to survive. Then panic set it: nano is evil; the grey goo is going to devour us all. The hype in either direction is irrelevant. You should only choose nano if you really need what only it can provide because in general working with nanoparticles is difficult!

If you need the invisibility of nano (from particles that are smaller than the wavelength of light), then "molecular" is even smaller than nano and can do a pretty good job. The hard-coated plastics on touch screens can be made much harder by adding silica particles, but clever molecular design can achieve much the same result, without the risk of the nanoparticles clumping to a size large enough to cause light scattering which is then visible to the customer's eagle-eyed quality control inspectors.

Or if you need the surface area of nano (surface-area/volume scales as 1/radius), again, why not go "molecular"? The large surface areas of nanoantimicrobials can make them potent, but molecules can be even more potent.

Maybe you need special electronic/optical/magnetic properties that set in towards the very low end of the nanodomain. Whether you reach this stage by creating small nanoparticles or via large molecular assemblies will affect your definition of this being nano or not.

Controlled shapes in the nanodomain can offer interesting possibilities. Some shapes/sizes of particles are good at killing microorganisms. The high aspect ratio (length/width) of carbon nanotubes (CNT) gives distinctive advantages in creating percolation networks for conductivity. Thin, flat clay particles can provide excellent barrier properties. There are good alternative antimicrobials, conducting materials and barriers—is it worth the trouble not only of getting small particles but also getting them into the right shape?

A final blunt but important point is that it is usually much cheaper not to go nano. In summary, it is not a good idea to choose the nano route to a product purely from the viewpoint of intellectual vanity or fashion. Chapter 2 discusses the science behind choosing nano.

1.3. HOW TO FIND THE RIGHT NANOADDITIVE?

Some companies that profess to provide nanoadditives don't really know what they are doing. This is not a criticism, it is a fact. Similarly, some companies that profess to have a need for a specific nanoadditive don't know what they really want. It is comparatively easy for a supplier to divert a few resources away from their old, tired, but reliable product development efforts to get a few quick samples of nanoproducts with potential properties that will excite a few potential customers. It is similarly easy for a potential user to make a lab sample sufficiently good to excite nano suppliers and one's own senior management as to its potential. The problem is that the scale-up from those supplier's lab samples is a process with relatively few established rules, and the scale-up effort by the customer is full of a similar level of uncertainties. The usual outcome is mutual frustration and disappointment.

The key advantages of nano—small size and large surface area—are also the key problems for suppliers and users. Suppliers generally don't know exactly what they have and why it works, because the information from an electron microscope image of 1μm × 1μm is not sufficient to characterize a 1 kg batch, let alone a 1 ton batch. Users generally have no idea how to go about understanding why sample X forms a perfect homogenous, stable mix and near-identical sample Y forms a gelatinous mass after a week.

In many cases, the properties of nano-X are dominated not by X but by the stuff around X, put there (deliberately or accidentally) by the supplier of the nano-X. For example, if nano-X is made inside a microemulsion then it will automatically come with a layer of surfactant. This surfactant must be perfect for creating a microemulsion and might be worthless for your application. A surfactant shell that might be perfect for your application might equally be worthless for making the necessary microemulsion.

Similarly, if the particles are made via a plasma or flame process, it is almost inevitable that they will exist as agglomerates. It usually takes a lot of energy and clever dispersants to break up those agglomerates. Whether that energy has resulted in changes to the surface molecules of X and whether the dispersant that covers X is also unchanged are both, generally, unknown.

If the stabilizing shell is relatively thick then it is easy to show that a large percentage of nano-X is, in fact, stabilizer. Although the stabilizer is necessary to keep the particles apart, many of the final properties

of the product will depend not on X but on the stabilizer, which can all-too-often compromise the performance. When extra functionality is required (e.g., particles with holes in the middle), the number of commercially viable routes becomes much smaller.

Chapter 3 discusses the science of finding the right nanoadditive, made by the right method, from the right supplier.

1.4. HOW TO CREATE STABLE NANOFORMULATIONS?

Nanoformulations are an uneasy truce between kinetics and thermodynamics. Sometimes, but rarely, we can be certain that thermodynamics guarantees that, with sufficient time, the formulation will become a lump of agglomerated particles. We therefore focus all efforts on ensuring a large kinetic barrier to that agglomeration. Sometimes we know the opposite—that if we can get the particles nicely separated, they will stay that way till the end of time; therefore we focus all our efforts on finding a dispersion technique that puts in sufficient energy/shear to accomplish that aim at the smallest cost in time, equipment and energy. Generally, though, there is insufficient information to know where the balance lies between kinetics and thermodynamics, or how close the formulation is to tipping one way or the other.

This uncertainty is not through lack of general theories about stability. DLVO theory (which is an obligatory part of any book on nanoformulations) provides a highly reasonable framework for understanding stability. Unfortunately for most systems, most of the time, the basic parameters required to apply the theory simply are not available for any reasonable expenditure of time and effort and so there are many cases where the theories simply don't apply. Not many of us, for example, can be confident that we meet the Smoluchowski criteria for calculating zeta potential from experiments, and many aspects of polymer stabilization of particles are entirely missing from DLVO. Fortunately, there exist some useful practical methods for characterizing nanoformulations in ways that can avoid many classic formulation problems.

For some systems the "solvent" is a UV-curable lacquer. For high production speeds, the UV-curing has to be rapid, so the system must be as reactive as possible. Therefore, acrylates are used rather than slower (more stable) methacrylates. It is well-known that the functional properties of particles can change significantly as their radius decreases—for example, high curvatures can cause surface atoms to be in unusual states of stress. So it is not uncommon to find that dispersions of

nanoparticles can cause an acrylate system to gel during storage. Adding stabilizers to intercept stray radicals is an obvious way to increase the storage time, but at the cost of reaction speed. Therefore, finding rational ways to understand the interactions between particles and curing systems is important.

The method used to disperse the particles has a profound effect on the overall system. The best system for a given project is one where no dispersion effort (other than a quick stir) is required. This, in effect, shifts the problem up one level to the supplier of the nanoingredient. If this policy also means that the team never has to handle dry nanopowders, this also shifts a lot of the nanosafety issues (discussed in detail in Chapter 9) up to the supplier. The supplier has the choice of producing powders which then have to be dispersed, or producing the nanoparticles pre-dispersed in media likely to be wanted by customers. Both methods have advantages and disadvantages which the users have to understand in terms of their own formulation needs.

Chapter 4 discusses various approaches to obtaining stable systems along with the compromises such approaches can entail.

1.5. HOW TO FIND SUITABLE SOLVENTS OR SOLVENT BLENDS?

One key aspect of obtaining a stable nanoformulation involves finding the right solvent system. The past decades have produced a welcome improvement in the safety of solvent-based formulations. Many once-common solvents are now unusable in a production environment for reasons of Safety, Health and Environment (SHE). In the past, if a certain solvent was not appropriate, then usually another could be found to replace it without too much difficulty. Now with the choice of solvents so greatly restricted, the chances of finding *the* perfect solvent are small. To gain the right balance of cost, solubility, vapor pressure, etc., a mixture of solvents is generally required. In many cases the mixture has to be one of both good and bad solvents, so that during the drying process the properties of the coating can be made to change in the desired manner.

Given a choice of 10 safe solvents, simply testing all possible 50:50 combinations would require 45 experiments. This is already an impractical experimental task. What usually happens in practice is that formulators stick with what they know and hope for the best. In the context of all the other uncertainties, this is a sub-optimal strategy.

Clearly the team needs a rational tool for thinking through the complexities of solvent choice. An "ideal" lab blend is not so ideal if it cannot be taken through to production. If the production department requires a specific solvent blend, it is better for the nanoparticle supplier to know this requirement near the start of the project rather than towards the end. It is all too common to find that the one solvent in which the nanoparticle is not happy is the one specified by Production.

Aqueous systems require different concepts of "solvents"; these generally comprise mixtures of water, alcohols and surfactants to ensure coating quality and compatibility. Nevertheless, the behaviour of these systems can still be understood in terms of the techniques used for solvent based systems. Chapter 5 offers practical approaches to obtaining the best solvent systems.

1.6. HOW TO SOLVE THE PROBLEMS OF COATING AND DRYING?

Most labs will have a few metering rods ("Mayer bars") used for creating surprisingly acceptable coatings. A few minutes' drying on the bench or in a (solvent safe) lab oven will complete the task. For thin coatings, a spin coater can do an excellent job in the lab. The production team will already have machines that coat (or print) and dry with a few key techniques. The trick is to know how to scale between the lab and production, all the time preserving the interesting properties built in to the formulation. There are many opportunities for misunderstanding between lab and production. What the lab might regard as a minor detail (e.g., that their new product has to be coated 5 μm thick and the machine normally coats 50 μm thick) can turn out to be an absolute showstopper. Conventional slot coating of a 50 μm wet thickness is simple; coating 5 μm wet with the same set up is essentially impossible, so an alternative technique has to be found.

Drying a high-solids aqueous coating slowly on the bench can give perfection; drying it quickly at production speeds can guarantee total mud-cracking.

Printers can often provide "almost good enough results". The team needs to know whether the results are at the limits of the technique (which they often are) or whether the issues can be fixed by further effort.

A coating created via a "drop" process (e.g., ink jet or spray) might form a perfect homogeneous coating at one (average) thickness and be

a disaster at a slightly lower thickness. The rules for understanding this are rather simple and can be described in a couple of simple spreadsheet models. The ideas are also relevant to pinhole formation in conventional coatings and to the formation of "necklace" defects in fiber coating.

Chapter 6 concentrates on a few key processes for coating/printing and drying/curing nanoformulations and provides models for knowing how close the process is to the border between perfection and disaster. This is highly important because lost production time costs money and failed trials can cost careers.

1.7. HOW TO GO NANO IN THE THIRD DIMENSION?

Sometimes it helps to create a third nano dimension—creating surface structures on coatings which may themselves contain nanomaterials. This is now done routinely using (UV) nanoreplication. Creating extra functionalities with large particles within the coating is not an option, so any particulate additives have to be nanosized.

There is no shortage of potential structures that can be produced on large scales. As is so often the case it is the little details that get in the way of profitable products. Bio-inspired structures have provided a strong boost to the world of nanostructuring—suggesting structures that might not have otherwise been thought of. By exploring some classic examples of bio and non-bio nanostructured surfaces and, in particular, the physics behind why many of the applications have not been a great commercial success, the team will be able to think through whether this is a route they want to follow. Chapter 7 provides a guide to the issues involved.

1.8. HAVE THE NANOADDITIVES DONE THEIR JOB?

Because many of the problems of nanoadditives are due to particle-particle interactions, the lower the concentration of particles the fewer problems will occur. This also has the obvious benefit (to the user but not the supplier) of reducing the cost of these often-expensive components.

So how does one determine an optimal level? Sometimes there is a simple formula to provide the answer. If the requirement is a high refractive index (RI) layer created by the addition of a high RI nanoparticle, there is little choice but to use the value suggested by a simple weighted average percentage of the RI of the medium and the particles.

Usually the relationship between particles and performance is less direct. If the particles are toughening a whole layer, they need to be evenly distributed (how do you prove that?) at an optimal spacing (what is that?) with strong links into the matrix (how do you prove they are strong enough?). Or, if you are relying on a surface effect, you have to prove that most (but what excess is sufficient?) particles are in the top few nm (but how few?).

These problems don't matter if, when you carry out your functional tests, you have all the properties you require at an affordable, stable level of nanoadditives. But they are crucial when, as is usually the case, intelligent trade-offs are required. So a team must have members capable of posing the right questions and finding cost effective and time effective ways of answering them. This is why Chapter 8 is so long; understanding the complex interactions of the nanoadditives requires a broad set of intellectual and analytical tools.

When it comes to matters such as adhesion (of the formulation to substrate and of the formulation to the nanoadditives), measurement methodologies alone are insufficient. A clear understanding of the science of what is required for good adhesion is important. Unfortunately, there are many myths about adhesion science, so it is necessary to explore in detail why surface energy is relatively unimportant, why chemical adhesion isn't necessarily a good thing and why the words "intermingling" and "entanglement" should be a key part of any discussion when strong adhesion is required.

Because nanoadditives are frequently used to provide exceptional hardness, the science of hardness and its relation to adhesion is discussed in some detail. Many hardness tests provide little information from much work. A few key tests can allow a much deeper understanding of whether the nanoadditive is doing a good job.

Chapter 8 discusses methodologies for getting to the answers to these problems, even for those without labs full of expensive analytical equipment.

1.9. ARE THE NANOADDITIVES SAFE?

It is a fact of life that people (and that means all of us) don't react rationally to risks. A scientist might know beyond reasonable doubt that a system containing a nanoadditive is perfectly safe, but that is of little value if staff or customers become irrationally fearful of that nanoadditive. Because of the huge emotions aroused by nanoadditives, the team

has no choice but to accept that the burden of proof is with them to demonstrate the safety every step of the way. Scientific logic is necessary but not sufficient. Convincing stakeholders that nanosafety is assured is slow, often tedious, sometimes restricting, but always a necessary process. What, after all, could be safer than totally inert, fire-proof, "natural" (you could dig it out of the ground) asbestos? Any costs incurred by taking the nanosafety issue seriously right from the start (and they can be considerable) are miniscule compared to the costs if the "harmless" system being used turns out to be the new asbestos.

Team-wide thinking required for other aspects of a successful nanoproject is equally valuable for addressing the safety concerns. If nine-out-of-10 steps in the safety chain are OK but one step poses an unacceptable hazard, then by thinking through the issues far in advance the chances are that one step can be removed by clever re-thinking of the process. If discovered during the big scale-up towards launch of the product, it could derail the whole project.

Fortunately, modern approaches to nanosafety fit very naturally with the quality systems that any responsible manufacturer must have. Just as external quality audits stop companies from falling into bad habits, so external nanosafety audits provide a vital independent check to help satisfy both staff and customers that things are being done correctly. Ultimately, we are responsible for the health and safety of our colleagues, employees and customers.

Chapter 9 provides an extensive background on the debate around nanosafety and offers a number of proposals for achieving the necessary consensus of co-workers, customers and regulatory authorities that the products and processes are safe.

1.10. A GREAT FUTURE FOR NANOCOATINGS!

This chapter ends in a positive statement rather than a question. Of course nanocoatings will provide opportunities for profitable growth. Nano *does* deliver things that can't be delivered by other means. The aim of this book is to increase the chances that your team will reap the benefits of nano. If the questions help you to avoid a nanoproject that is for one or more reasons doomed to failure, then that is doubly good. It avoids an expensive mistake and it frees up resources so you can focus on other nanoprojects with a higher chance of success.

CHAPTER 2

Why Do We Need Nano?

As discussed in Chapter 1, it only makes sense to use nanoadditives if the coating requires the properties that only nano can bring. Rather than attempt to provide an exhaustive survey of all possible properties of nanomaterials, the aim is to provide a survey of knowledge that is fundamental to the understanding of why these materials behave as they do. The smallness of nano is a source of many problems as well as a source of many positive attributes, so it is especially important to understand the specific issues of small particles. The emphasis is on generality, so the highly specific benefits of materials such as carbon nanotubes (CNT) or graphene (which are discussed in huge detail in other books) will not be included. A lot of the disappointment with wonder materials such as CNT comes not from their lack of marvellous properties, but from the fact that the general principles of nanomaterials can make it hard to realize those properties in a practical system. The aim of this chapter, therefore, is for the whole team to grasp the implications, positive and negative, of a decision to work with nanomaterials.

2.1. SIZE-RELATED CALCULATIONS

A lot of profound insights come from some simple equations. Even though they are simple, it is remarkably easy to get them wrong by many orders of magnitude, so the Spreadsheet does them for you—taking inputs in common units, doing all calculations in SI units and converting results back to common units.

All calculations will be based on radius (r) rather than diameter.

The choice is arbitrary but many simple errors arise from accidentally switching from one to the other.

As we all know from elementary geometry:

$$\text{Surface Area} = 4\pi r^2 \tag{2.1}$$

$$\text{Volume} = \frac{4}{3}\pi r^3 \tag{2.2}$$

A few important ideas come from these equations.

2.1.1. Surface Area/Volume

$$\frac{\text{Surface Area}}{\text{Volume}} = \frac{4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{3}{r} \tag{2.3}$$

So as the radius decreases, the relative surface area increases. Halving the radius doubles the Surface Area/Volume ratio. This has profound effects on stability—smaller particles tend to be less stable because they have relatively more surface area by which they can clump together. Whether a higher surface area ratio helps in the final product depends on the specifics of the application.

Given that atoms or molecules at the surface of a nanoparticle can be particularly relevant to catalysis or stability, it is interesting to estimate the % atoms at the surface. Using an estimate of the radius s of the surface species (e.g., 0.1 nm for an atom, 0.3 nm for a small molecule), the number at the surface is Surface Area/($4\pi s^2$) and the total number is Volume/($4/3\pi s^3$), then:

$$\%\ \text{Atoms at Surface} = 100\ \frac{4\pi r^2}{\frac{4}{3}\pi r^3} \cdot \frac{\frac{4}{3}\pi s^3}{4\pi s^2} = \frac{100s}{r} \tag{2.4}$$

For s = 0.4 nm this gives 20% for 2 nm particles and 2% for 20 nm particles. So anything which requires surface atoms/molecules for functionality is effectively wasting 98% of the molecules in a 20 nm particle, Figure 2.1.

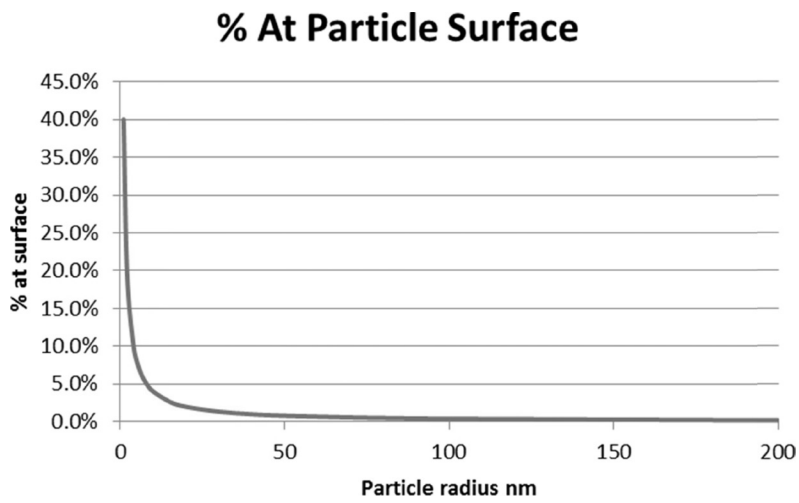


FIGURE 2.1. Percent of atoms/molecules at the particle surface.

2.1.2. Distance Between Particles

As the formulation becomes loaded with particles, the average distance between them decreases; for a given Vol % loading, there are many smaller particles. Therefore, the inter-particle distance becomes smaller—in fact, it is proportional to *r*. The dependence on Vol % is a bit ugly:

$$\text{Inter Particle Distance} = 2r\left[\left(\frac{1}{3\pi\ \text{Vol}\,\%} + \frac{5}{6}\right)^{0.5} - 1\right] \qquad (2.5)$$

The Distance spreadsheet performs the calculation for you. The viscosity of a particle solution increases with Vol % and the same spreadsheet provides an estimate of that increase via:

$$\text{Viscosity Increase} = \frac{1}{\left(1 - \dfrac{\text{Vol}\,\%}{\text{Close Packed}\,\%}\right)^2} \qquad (2.6)$$

The interparticle distance is of great importance in terms of particle stability. Using DLVO theory (discussed in Chapter 4) it is possible to calculate a distance where there is maximum repulsion due to charge effects. It is generally said that if the value of this maximum is > 20 kT

(20 times larger than the Boltzmann energy) then the particles will be stable. Unfortunately it is not quite that simple. If the average inter-particle distance is less than the point of maximum charge repulsion then the barrier ceases to be effective because the particles are pushed below that limit by packing forces stronger than Brownian motion. Because the inter-particle distance is proportional to r, a stable 30% dispersion at, say, $r = 40$ nm might become unstable for $r = 20$ nm because the inter-particle distance has halved. Given that the height of the DLVO barrier also tends to fall with radius (the DLVO spreadsheet lets you confirm this for yourself), the effect of going smaller can be catastrophic. In other words, an innocuous and simple formula on particle distance can illustrate profound consequences in terms of stability.

When steric stabilization is included in the DLVO analysis then the barrier tends to a high value and the above issues are less of a problem—except when DLVO does not apply. This situation occurs when there is too little polymer (bridging failure) or too much polymer (depletion failure).

2.1.3. Dominance of Large Particles

The volume, and therefore mass, of one particle of $r = 100$ nm is 1000× greater than a particle of $r = 10$ nm. So if a dispersion technique creates 1000 of the desired 10 nm particles and has just a 0.1% "contamination" by a single 100 nm particle, this single particle will make up 50% by weight of the total particles. This trivial calculation should alert all users to statements of "purity" by nanoparticle suppliers. An unscrupulous supplier could in this extreme case claim > 99% purity based on particle counts, yet still ship product that was only 50% pure in terms of weight of the desired 10 nm particles.

Users should understand that it is exceedingly difficult to provide high mass % purity of nanoparticles. If a supplier quotes > 90% mass % purity then that is an exceptional accomplishment. If the supplier quotes > 90% size % purity this could, in extreme cases (though most suppliers are not so unscrupulous), mean that their product is mostly worthless for your application.

The Distribution spreadsheet lets you play with these issues. The curves are generated via a so-called Gamma distribution (creating a Gaussian), and the best way to get different curves is to change the values until you get a feeling for which parameter affects which part of the curve. The Y-axis is in arbitrary units.

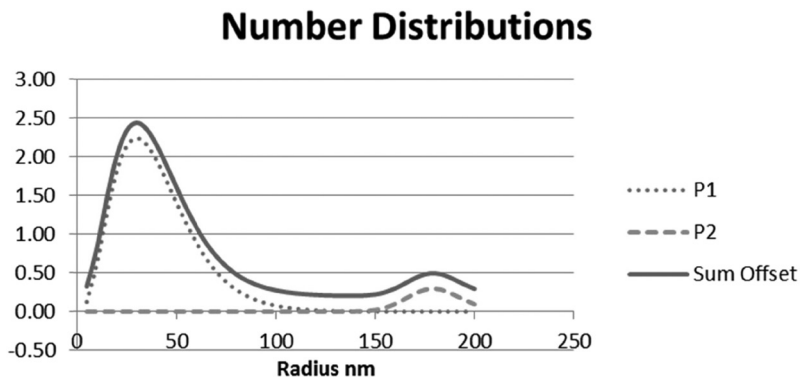


FIGURE 2.2. *The Number Distribution of simulated particle data.*

For example, here is a mix of a large number of small particles (P1) with a relatively small addition of larger particles (P2). The Sum is offset for clarity, Figure 2.2.

The cumulative number distribution emphasizes the large number of smaller particles and the larger particles barely register as a blip towards the end, Figure 2.3.

On the other hand, the cumulative mass distribution tells a completely different story—the small particles play very little part in the distribution, Figure 2.4.

The Cumulative Area distribution (not shown here) is included in the spreadsheet—its shape is somewhere in between the number and mass curves.

Size distributions are important when considering the ratio of surface area to volume. The spreadsheet calculates the number averaged Surface Area/Volume ratio.

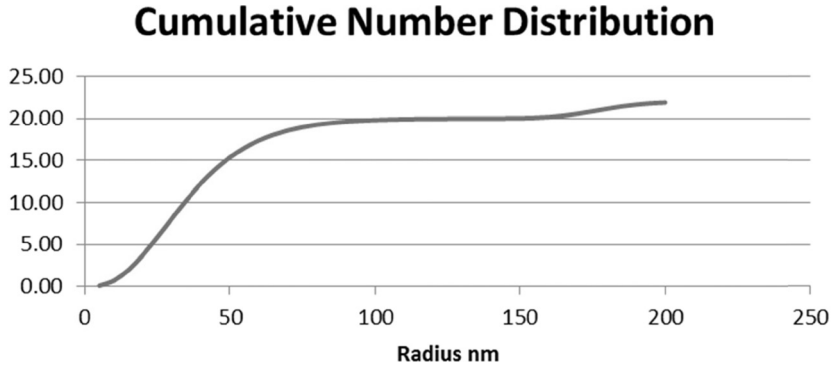


FIGURE 2.3. *The Cumulative Number Distribution of the same data.*

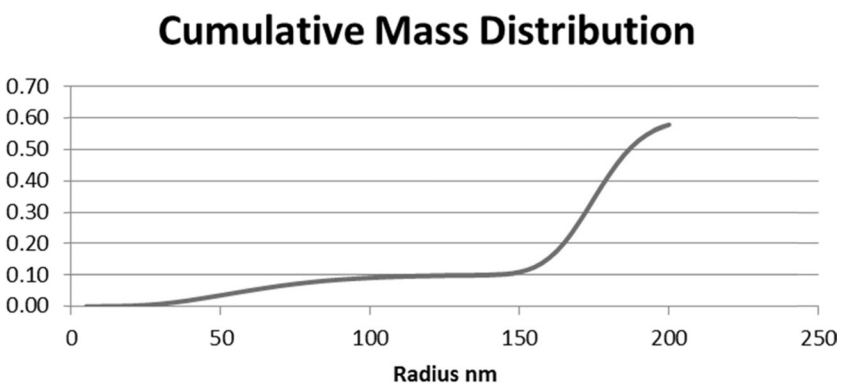


FIGURE 2.4. The Cumulative Mass Distribution of the same data.

Tip: Always check whether graphs of purity are by number or mass distribution. If they are by number distribution, then simulate the distribution in the Distribution spreadsheet to see if the product is good enough for your application.

The spreadsheet calculates quite a few numbers that you see in various datasheets and academic papers. It is important to realize that there is a large variety of these numbers and it is very easy for confusion to reign if you are unfamiliar with them.

The first are the "mean" values. There are many possible means, indicated by a nomenclature such as $D[1,0]$, the Number mean $D[3,2]$, the volume/surface (or Sauter) mean and $D[4,3]$, which is the mean diameter over volume (DeBroukere). Note the use of Diameter in these quoted numbers—a shift from the use of Radius adopted as the standard in other formulae. A mean value $D[x,y]$ is defined as:

$$D[x,y] = \frac{\Sigma D_i^x \cdot n_i}{\Sigma D_i^y \cdot n_i} \tag{2.7}$$

In the above example the means shift steadily to higher values: $D[1,0]$ = 104 nm, $D[3,2]$ = 265 nm and $D[4,3]$ = 322 nm. Note that these numbers are sometimes shown with an averaging bar along the top as, for example, $\overline{D}43$, which gets confusing when thinking about the common $D50$ described next.

There are also various Median values—Number, Area and Volume. The Median Volume is the diameter where half the particles have less volume and half have more volume. This is the standard D50 value that

is often quoted. Again, you need to play with your distribution to see if the *D*50 really gives you what you want in terms of number (or, indeed, area) distribution.

2.1.4. Dominance of Shells

Most nanoparticles have a "shell" around them containing, for example, a dispersing agent. Simple geometry leads to a surprising conclusion: a nanoparticle may contain a rather small amount of the nanomaterial. For a particle of nanomaterial of radius *r*, with a shell of size *s*, this can be calculated from the ratio of their volumes based on $4/3\pi(\text{radius})^3$ where radius is either *r* or *r* + *s*:

$$\text{Vol \% Nano} = 100\frac{\frac{4}{3}\pi r^3}{\frac{4}{3}\pi(r+s)^3} = 100\left(\frac{r^3}{(r+s)^3}\right) \tag{2.8}$$

or, for a particle of total radius *r* with a shell of size *s* the radii become *r* and *r* – *s*:

$$\text{Vol \% Nano} = 100\frac{\frac{4}{3}\pi(r-s)^3}{\frac{4}{3}\pi r^3} = 100\left(\frac{(r-s)^3}{r^3}\right) \tag{2.9}$$

For a 20 nm radius particle with a 6 nm shell, the first case gives % Nano = 46% and the second gives % Nano = 31%. In other words, if you bought a "20 nm nanoparticle", over half of it would be the shell. Although this is a deliberately extreme example, many formulations have been a disappointment to formulators because they were unaware that the properties were so greatly diluted by the shell.

The shell is often less dense than the core of the particle, so these volume-based calculations are pessimistic in mass terms. The Shell % spreadsheet also calculates the mass ratio. In the above examples, for relative densities of 2.5:1 the % Nano values are 68% and 57%, respectively, Figure 2.5.

The graph shows the %Vol and %Mass curves for *r* = 20 nm and densities of particle and shell being 2.5 and 1.

The effect of the shell also has ramifications for the refractive index

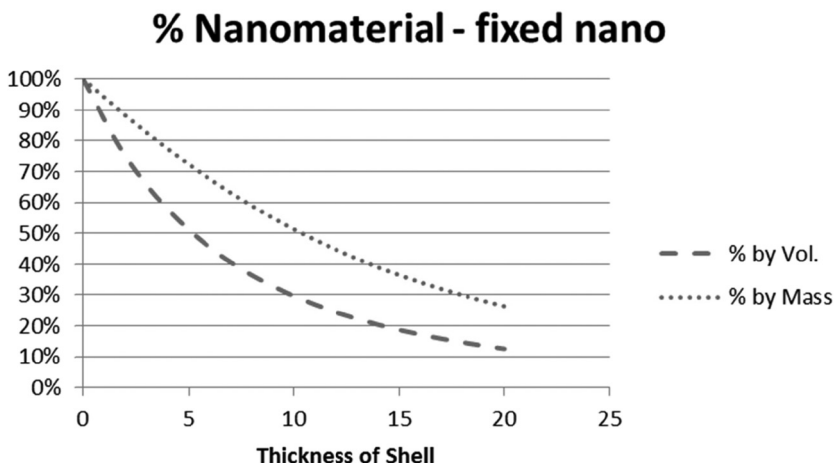


FIGURE 2.5. The % nanomaterial for a fixed size of core with varying shell thickness, shown by Volume and by Mass.

of the particle as a whole. The next section shows how to calculate the average refractive index of any system from the volume fraction of the individual components. If the size of the core and of the shell are known, then instead of using the volume % of the "nanoparticle" in the formulation, the corrected volume % (multiplied by %Vol of the core) should be used if the shell is presumed to be the same RI as the general medium. If the RI of the shell is also known and is significantly different from the general medium, then a 3-term RI average can readily be calculated. The authors learned the hard way that corrections for this core effect can sometimes be significant. Turning a problem into a solution, if you are confident of the RI of the core material and can accept the approximation that the shell is similar to the rest of the matrix, then by measuring the RI of the whole system it is possible to back-calculate the size of the shell if the supplier can't or won't tell you what it really is.

2.2. INVISIBILITY

To provide a tablet screen with extra functionality (such as hardness or conductivity) via solid additives (silica, conductive oxides, etc.), the particles have to be effectively invisible at normal wavelengths of light.

There are only two ways to make this happen.

First, if the refractive index (RI) of the particles is identical (at all relevant wavelengths) to that of the matrix, the photons have no way

of distinguishing the two phases. With some care, even large silica particles can be made invisible in a matrix, as it is not too difficult to find materials of RI ~ 1.46. The greater the difference between refractive indices of the particle and the matrix, the greater will be the scattering and consequently the greater the visibility of the particles. This applies not only to particles with high RI such as ZnO (RI = 2), but also to air-filled particles used for anti-reflection coatings with a remarkably low RI ~ 1.12—though the difference in RI between 1.12 and 1.46 is not as severe, so it is less of a challenge to hide these low RI particles [1].

Second, if the particles are much smaller than the wavelength of light then the photons "see" an homogenous coating with a refractive index that is a (weighted) average of the matrix and the nanoparticles. There are a number of ways to calculate the average refractive index; for example, via volume fraction or Lorentz-Lorenz. Both methods are available in the spreadsheet. The volume fraction method gives:

$$\mathrm{RI}_{av} = \mathrm{RI}_{matrix}\ \mathrm{Vol\ Fraction}_{matrix} + \mathrm{RI}_{particle}\ \mathrm{Vol\ Fraction}_{particle} \quad (2.10)$$

In some cases the important factor is that the RI is average; in other cases such as anti-reflection coatings, the precise average value is of great importance.

To complete the picture, all that is needed is a simple formula relating particle size distribution and RI difference so the user can specify to the nanoparticle formulator exactly what is required. Unfortunately, no simple formula exists. The user, typically, will ask for "as small as possible", which brings many problems in terms of dispersability (high relative surface area) and, generally, cost. The supplier may hope that the user will be able to find a matrix of sufficiently high RI that will mask some of the problems from the larger-radius fraction, but this also brings problems to the user.

There is also a mismatch between what the supplier sees and what the user's customer sees. To most casual observers a "good enough" blend of size and refractive index is indistinguishable from perfection. Yet experienced eyes can pick up even small amounts of scatter and that can make the difference between being accepted or rejected as the supplier of the top protective film on 1 million tablet computers.

To those who have access to full solutions of Maxwell's equations, modelling can provide a basis for discussions between supplier and user. For the rest of us, two simple approaches have to suffice.

2.2.1. Rayleigh Scattering

In the ideal case of an isolated nanoparticle much smaller than the wavelength of light, the Rayleigh scattering formula can be used—the same one that explains why the sky is blue.

Light of (modified) wavelength λ is scattered to an angle θ at intensity I_θ relative to the incoming intensity I_0 by particles of radius r and RI n:

$$\frac{I_\theta}{I_0} = 64\frac{\pi^4 r^6}{(8R^2\lambda^4)\left(\frac{n^2-1}{n^2+1}\right)(1+\cos^2\theta)} \tag{2.11}$$

The R^2 term is the "inverse square" falloff of intensity with distance R.

The reason for "modified" wavelength is that for scattering in a medium with refractive index m and real wavelength λ_0, λ is given by:

$$\lambda = \frac{\lambda_0}{m} \tag{2.12}$$

The Rayleigh formula encapsulates the intuitions that large r and

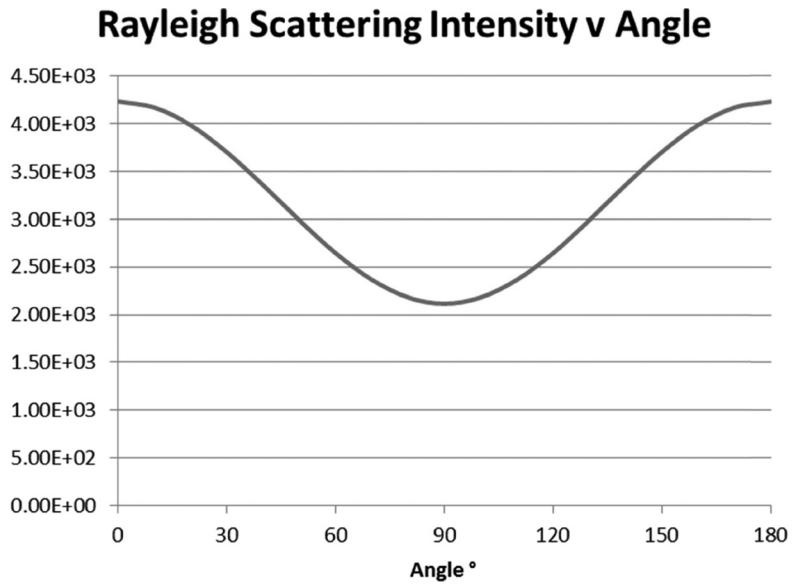


FIGURE 2.6. *A Rayleigh scattering plot from a mixture of two sizes of particles.*

large *n* (or n-m) are bad. The angular part of the equation which seems, at first sight, to be uninteresting governs the all-important visibility of the scattered portion of the light. The trained eye catches the small fractions of light coming off at higher angles, Figure 2.6.

In principle, the Rayleigh spreadsheet could perform a fully integrated calculation related to the particle distribution curve. For simplicity, however, two particle sizes, and their relative ratios, are entered and the weighted average of the scattering is calculated. For a pure 20 nm particle, the scattering intensity increases by a factor of 10 if 10% of the number of 200 nm particles are present. This at least gives a feel for the scale of the effect of the presence of larger particles.

Although playing with the formula in the spreadsheet is instructive, the basic assumption behind Rayleigh scattering is that it is coming from individual particles. This is clearly inappropriate once the inter-particle distance approaches the wavelength of light, so a more sophisticated theory is required. Nevertheless, the key insights from Rayleigh scattering are useful, providing the particles are small.

2.2.2. Mie Scattering

As particles get bigger, the assumptions behind Rayleigh scattering cease to be valid and Mie scattering takes over—indeed, Rayleigh scattering is equivalent to Mie scattering for small particles. A full model of Mie scattering is too complex to discuss here and is largely irrelevant because, again, the assumptions behind Mie scattering don't reflect the realities of typical nanodispersions with small inter-particle distances. As with Rayleigh scattering, it is still useful to have some feel for the effects, so the spreadsheet contains a simplified "van de Hulst" version which captures the essence of Mie scattering without going to a level of complexity which in reality delivers little extra benefit for a lot of extra work [2]:

$$\text{Scatter} = 2 - \frac{4}{p}\sin(p) + \frac{4}{p^2}(1 - \cos(p)) \tag{2.13}$$

where

$$p - \frac{4\pi r\left(\frac{n}{m} - 1\right)}{\lambda} \tag{2.14}$$

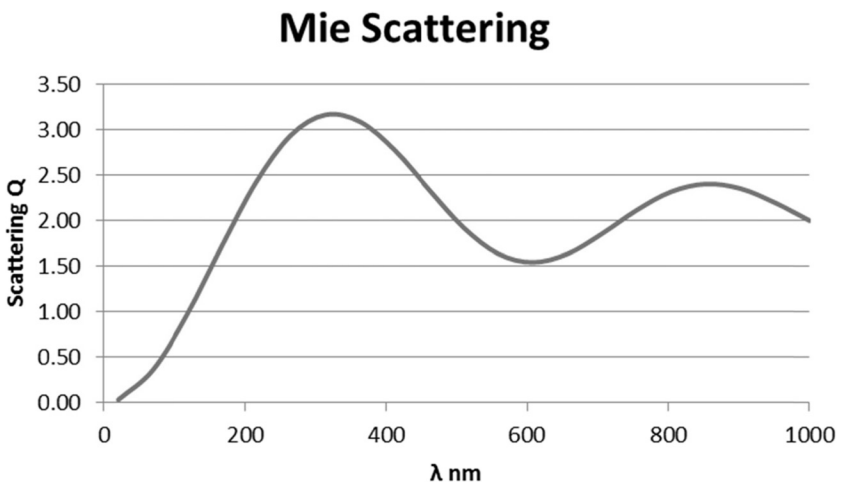


FIGURE 2.7. A Mie scattering plot using the van de Hulst approximation.

n is the RI of the particle, m the RI of the medium and λ is the wavelength of the light, Figure 2.7.

The van de Hulst plot is good enough to show that in most circumstances Mie scattering sets in only at sizes that should be irrelevant for most nanoapplications.

2.2.3. Other Shapes

The above scattering theories apply only to spheres. For highly anisotropic materials such as carbon nanotubes, graphene or nanoclays, there is no simple theory that can offer much help. The team can either work with experts who have access to full scattering code, or use their intuitions about size and refractive index based on Rayleigh and Mie scattering.

2.2.4. Common Sense Optics

If the nanoparticles are very small (< ~20 nm, depending on their RI), there is no scattering. If they are "big" (> ~100 nm, again depending on their RI), then there is gross scattering. If there is a blue color in scattered light and an orange color in transmitted light then the particles are in the Rayleigh regime. These common sense rules are surprisingly powerful for practical formulators. If an additive turns a clear formulation blue, then things are heading in the wrong direction as the hitherto small particles must be getting larger and therefore providing stronger

Rayleigh scattering. If an additive gives a hint of blue to a somewhat cloudy formulation, then things are heading in the right direction, as the hint of blue means that the particle sizes are decreasing.

Although these discussions apply to liquid and solid formulations, there is less need for speculation with the liquid formulations, as modern light-scattering measurement tools can tell you much of what you need to know.

2.2.5. Quantum Optics

At the other end of the optical scale, quantum optics gives strong reasons for going nano [3]. The optical properties of materials are a given from the scale of meters down to 10's of nm. At the low nm scale (below the "exciton Bohr radius") the rules change. Most of us have a memory of the quantum physics of a "particle in a box". The quantum levels are dictated by the integral number of wavelengths that can fit into the box. The smaller the box, the harder it is to fit integral wavelengths, so the bigger the gap between them; in other words, the smaller the quantum box, the bluer the emitted light.

The beauty and the challenge of quantum dots arises from the fact that the color is controlled by size. The good news is that, instead of having to tune the chemistry of the material to change the color, the same material produces different colors through change of radius. The bad news is that the radius effect is so important that it requires exquisite control of manufacture to provide the desired color.

In a later Case Study on quantum dots, the complexities of real life start to impact. The electronic properties of the dots are modulated by the dispersants. For example, the charge characteristics of phosphonate groups used for attaching the dispersants shift the wavelength of the dots from their "pure" but unusable (because of clumping) state.

An equivalent way to think about these quantum phenomena is in terms of plasmon absorption which, by definition, is relevant to conducting metal particles. The key here is that a broad-band absorber/reflector such as silver or gold becomes specifically absorbing, and therefore colored, below a critical particle size. Because the color is strongly dependent on particle size and these small particles are non-scattering, it is possible to create stunning colors. It is humbling that Roman nanotechnologists in the 3rd century could create gold and silver nanoparticles in this size range (though slightly larger, as scattering effects are significant) when making the Lycurgus cup.

2.3. LOTS OF SURFACE AREA

The interior of most particles is not adding much value. Most of what a particle can offer takes place in the outer few nm, unless, that is, the requirement is for bulk padding. Although some nanoparticles are sufficiently cheap that they can be used as padding, in most cases those inner atoms are an expensive waste of money. If the shell is particularly expensive, then a cheap core material with a few nm of shell is a good idea scientifically, though the cost of the process for creating the core-shell particles might be higher than the savings from using less of the expensive material.

If we are paying for surface area it is important to know why it is wanted.

2.3.1. Sintering

It is possible to "melt" particles together at temperatures much below their melting point. This sintering process is vital for nanoapplications such as glass-like hardcoats and printed conductors. Small particles sinter faster for multiple reasons that are inter-related.

- The free energy of a surface is higher than that of the bulk, so reducing the surface area is the thermodynamic driving force. The greater ratio of surface area to volume of smaller particles provides a greater relative driving force.
- The curvature of smaller particles creates a higher free-energy and driving force.
- Liquid present during sintering experiences very high capillary forces ($1/r$), which impose a strong pressure gradient that forces the particles together.

Various formulae exist to model the different forms of sintering (surface driven, diffusion driven, viscous, liquid-assisted . . .), typically showing how the % width of the bridge between particles (i.e., Width/r) varies with respect to time, t, radius r and a constant K, according to a formula of the form:

$$\%\,\text{Bridge}^{x} = t\left(\frac{k}{r^{y}}\right) \tag{2.15}$$

Typical values for the exponents are x~3 and y~2. What these formu-

lae mean is that growth of the bridge can be relatively fast at the start but slows down considerably (t^{-x}) with time and that the radius has a significant ($y = 1$) to highly-significant ($y = 3$) effect, with smaller r being faster. The slow-down comes because the driving forces (e.g., radius of curvature) decrease the bigger the bridge becomes.

Unfortunately, the general need to have some stabilizing shell around a nanoparticle means that for smaller particles there is a greater relative proportion of shell and, therefore, a greater potential barrier to perfect sintering. Many of the failures to produce high-quality printed conductors from nanoparticle inks can be traced to this basic fact. Announcements of inks that sinter at low temperatures to give high conductivities are more common than products that one can actually purchase. When the shells are organic molecules and the particles are ceramics, high temperatures can make the shells disappear as carbon dioxide and water. For those who require low temperature sintering, such a simple solution is not available. Tricks that allow the shell to detach from the particles (e.g., oxidizing thiols attached to gold nanoparticles) then become necessary [4].

It is well-known that perfect spheres cannot pack with a density higher than ~74%. Typical "random packing" tends to provide void levels of 40%, significantly higher than the 26% of perfect packing. This creates a problem because sintering can transform a coating from lots of small voids to a few large ones. These large voids can cause problems such as acting as scattering centers for light. Packing densities can be higher if two different particle sizes are used—the smaller particles can fit into the voids of the larger ones. Fine-tuning a distribution for optimal packing is an interesting trade-off between formulation complexity and product performance.

The fact that smaller particles sinter faster is loosely connected (because surface free energies are involved) to the fact that smaller particles melt at a lower temperature. For example, the melting point of gold falls by ~200°C for 5 nm particles—an interesting fact but perhaps not too useful for most nanocoatings.

2.3.2. High Surface Adsorption/Desorption

Because of the $1/r$ dependence of relative surface area, it makes sense to have small particles when high adsorption is required, or when the particles need to deliver a material from the surface. Unfortunately, things aren't this simple. As it gets harder to stop smaller particles from clumping, the surface has to be ever-better protected by stabiliz-

ers from doing so, reducing the adsorptive and desorptive capacity. It might, therefore, often be better to accept larger particles engineered to provide surface area via porosity. Whether clumped small particles can be considered as highly porous large particles becomes a matter of philosophy rather than science. What matters is whether the system provides the required adsorptive/desorptive capacity as measured by techniques such as BET.

2.3.3. Changing Chemistry

Gold nanoparticles are simply small particles of ultra-inert gold—until they reach a critical size, when their chemistry changes from dull to interesting. As this happens in the domain of the few nm, the question arises of how to preserve them as individual particles while keeping their interesting properties. Any reasonably sized protective shell will dwarf the particle in both size and volume.

The sudden flip of properties at these small dimensions is reminiscent of the change of energy levels when materials such as CdSe become quantum dots or gold particles show plasmon absorption. It is also related to the s/r dependency of the ratio of surface to bulk atoms. As this is, therefore, a general phenomenon, there is no doubt that a lot of interesting coating chemistry with such small particles remains to be explored. Almost as much ingenuity will have to be expended in keeping the particles small and active as will be needed in creating such small particles in the first place.

2.3.4. Fighting Microbes

Some antimicrobial materials such as silver work best in nanoform because they present a high surface-to-volume ratio that allows insoluble materials (such as silver) to become sufficiently soluble to diffuse out and attack the microbes. The advantages and disadvantages of this diffusion are discussed in detail in the section on antimicrobial formulations.

There is another way in which nanoparticles and nanostructures can show antimicrobial activity. It has been claimed that microbes seem to be easily upset by contact with surfaces containing nanosized features. In principle, this means that an antimicrobial surface can be created without the need for specific antimicrobial chemicals. At the time of writing, this idea does not seem to have resulted in practical products. For example, even in ideal lab conditions, controlled hydrophilic/pho-

bic nanostructures of lines or squares showed variable patterns of activity to *E. coli* and *S. aureus* [5].

The ability of a material which in itself is regarded as safe to have potent antimicrobial activity via surface shape seems to be good news for formulators. Two drawbacks spring to mind. First, a pristine nanosurface may well stop microbes from growing, only for the effect to be neutralized by tiny amounts of surface contamination from, say, grease or oil. Thus, the antimicrobial effect might require regular, vigorous cleaning, in which case those doing the cleaning can achieve the same antimicrobial effect with a common surfactant cleaning spray formulation—some of the cationic surfactants are potent antimicrobials.

The second drawback also applies to active nanoparticles and arises from the simple word "antimicrobial". There are numerous reports which cite the latest wonder product (often "natural") as showing potent antimicrobial activity. Translated, this means "kills living cells", which then carries the possibility of "killing the cells of humans, pets, aquatic organisms, bees, etc." One man's "antimicrobial" is another man's "toxin". Those who aim to fight disease bias the discussion towards "antimicrobial"; those who fight for a safer planet bias the discussion towards "toxin" (and indeed might simply want to imply that all nanoparticles are toxins, as discussed in Chapter 9).

Given the rising realization that the "cleanliness hypothesis" is quite a good explanation of outbreaks of many human ailments, it is arguable that the world needs much less general antimicrobial activity, while the rise of dangerous bacteria such as MRSA makes it arguable that we need more specific strategies for fighting the real killers.

It is not at all clear whether antimicrobial nanocoatings would be "bad" because they kill large numbers of harmless (or perhaps beneficial) microorganisms via, say, nanosilver, or "good" because they kill modest numbers of lethal microorganisms. It is also not clear if nanocoatings can provide routes to such desirable specificity. This all feeds into the issues of nanosafety discussed in Chapter 9. If nanoparticles can be potent antimicrobials, it follows that they have the possibility of harming humans directly (attacking specific cells) or indirectly (by making the environment too "clean"). Sorting out that balance promises to be a herculean task.

2.3.5. Magnetism in the Nanodomain

Just as there are significant optical changes at a critical size with

quantum dots and chemistry changes with metals like gold, super-small particles can take on novel magnetic properties.

As independent magnetic entities (rather than parts of a larger magnetic region), a magnetic nanodomain can have properties not present in bulk (or micron-sized) magnetic particles, such as spontaneous magnetic order at zero applied magnetic field. It is also possible for superparamagnetism to occur, where the particles have no net magnetism but behave as strong magnets in the presence of a magnetic field. Such nanomagnets have great promise in many areas such as MRI, magnet-controlled drug delivery and the specialized area of magnetic recording media. Their potential in nanocoatings seems to be largely untapped except, of course, for the now ancient industry of magnetic tape recording. The extreme pyrophoric nature of typical nanometals such as iron or cobalt (a consequence of the high surface-to-volume ratio) make the early steps of producing formulations problematical, and the need for protective surface layers such as gold or graphene imposes further constraints. Life is easier if the metal oxides have the desired magnetic properties.

2.3.6. Printable Electronics

The hype of roll-to-roll electronics, displays, solar cells and so forth has matched nanohype in many ways; but behind the hype there is, at last, some solid prospect of real-world products. The need for high-conductivity printable electrodes has led to much progress in understanding the sintering of nanoparticle metals (e.g., silver) as discussed above. Controlled nanodomains are a requirement for some products. For example, organic photovoltaics require nanosized phase-separated particles of electron donors and acceptors. The classic way to achieve this is via fullerenes (arguably we can call them nanoparticles) and nanodomains of conducting polymers (though this is stretching the definition), and rational ways to achieve this are discussed in a subsequent chapter.

Helping to drive this process is the sense that conventional vacuum-coated conductors (typically, conducting oxides) have reached a price/performance limit. For example, thicker ITO (indium tin oxide) is more conducting but also more expensive, yellower and more brittle than is desirable. “Grid” structures with locally high conductivity from opaque printed stripes interconnected by a low conductivity (cheaper, less yellow, more flexible) ITO is sometimes suggested as a compromise.

Printable CNT conductors that rely on their high anisotropy (rather

than surface area) to form a percolation network (see below) have been promised for many years. The properties of graphene may so eclipse those of CNT that the impetus to solve the many problems in getting CNT coatings to full functionality may vanish.

In many cases of printed electronics, protective layers around nanoparticles, required to make them processable, are literally a barrier to full functionality. In many cases the requirement is for large (μm) crystal domains rather than individual particles. So, the usability of nanoparticles is sometimes questionable. A specific example is that the large crystals of classic silver inks often provide a higher net conductivity than a network of nanosilver particles each surrounded by a dispersant molecule.

Nevertheless, it seems highly likely that printable electronics will require a wide variety of processable nanoparticles and the techniques described in this book will be highly relevant to the greatest challenge of all—producing printable electronics on the kilometer scale.

2.4. STAYING SUSPENDED

Everyone knows that large particles (for a given density) fall out of solution faster than smaller ones. This is not as obvious as it might seem. Although smaller particles have a lower mass, they also have a smaller cross-section, so the viscous drag on them as they fall is lower. The reason that larger particles fall faster is that mass depends on r^3 but the viscous drag depends only on r. The balance of these effects gives a terminal velocity v, given by Stokes' law:

$$v = \frac{2}{9}\frac{(\rho_p - \rho_f)}{\mu} g r^2 \tag{2.16}$$

where μ is the viscosity of the fluid, g is gravity and ρ_p and ρ_f are the densities of the particle and the fluid. The Settling Time spreadsheet carries out the calculations and allows selection of two different particle sizes so their relative velocities and settling times can be shown.

The inclusion of g is a reminder that in a centrifuge, the settling velocity can be much higher. For example, a 100 mm tube spinning at 1000 rpm gives a relative g-force of 112. The Settling Time spreadsheet includes the ability to do calculations at higher g-forces.

If a centrifuge is set up with an optical sensor to measure how the opacity of a suspension changes with time, it is easy to calculate the par-

ticle size for simple dispersions and, with more powerful formulae, to determine a lot about particle size distribution. Such a technique has an advantage over light-scattering measurements because it can be performed on un-diluted suspensions, thereby eliminating worries that the dilution necessary for the light-scattering has also changed the particle sizes [6].

Brownian motion offers some resistance to settling. The Stokes-Einstein equation provides the diffusion coefficient, D, for a particle of radius r in a fluid of viscosity μ being buffeted with Boltzmann energy kT that provides the Brownian motion:

$$D = \frac{kT}{6\pi r\mu} \tag{2.17}$$

The factor of 6 is debatable; some prefer a factor of 4.

The diffusion coefficient is not of much help on its own. However, the distance d, travelled in time t is given by:

$$d = \sqrt{2Dt} \tag{2.18}$$

The reason d varies as the square root of t is because progress is made via a random walk. This makes it impossible to calculate an average velocity which can be compared to the sedimentation velocity calculated from the Stokes' equation.

Instead, d can be calculated from the time taken for an average particle to fall to the bottom of a test tube, which is given by the height of the test tube divided by the Stokes' velocity. If the calculated d is similar to the height of the test tube, then sedimentation will be essentially counteracted by Brownian motion. If the calculated d is very much smaller, then the Brownian motion is essentially irrelevant.

From the Settling Time spreadsheet it can be found, for example, that for a particle with relative density ($\rho_p - \rho_f$) of 1 g/cc present in water with viscosity 1 cP, a particle of radius 34 nm will diffuse 5 mm in the time taken (45 days) for the particle to fall by 10 mm. This means that particles of 68 nm diameter will not have fully settled after the 45 days.

Another way to look at Brownian motion and settling is via the comparison (using ρ in the sense of the differential density):

$$kT \sim \frac{4}{3}\pi r^3 \rho g h \tag{2.19}$$

If the Brownian motion, kT is greater than the gravitational component (which necessarily includes a "height" component, h), then the particle is indefinitely stable.

Yet another way (suggested by Prof. Coleman at Trinity College Dublin, used with permission) is to think of diffusion as a thermodynamic force. The particle is exactly balanced when this force equals the gravitational pull and the result is an estimate of the height above the bottom of the tube where the concentration will have fallen exponentially to $1/e$:

$$\text{Height} = \frac{3}{4} \cdot \frac{kT}{\pi \rho g r^3} \tag{2.20}$$

A log-log plot shows that for a density of 2 g/cc, particles greater than ~20 nm radius (which have a $1/e$ value ~10 mm) will not be much helped by Brownian motion, Figure 2.8.

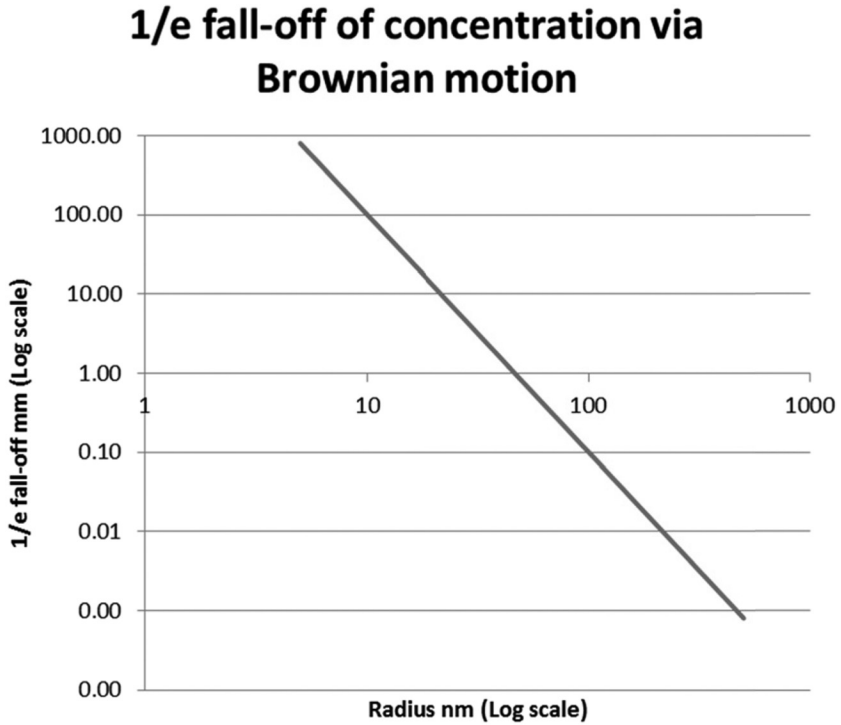


FIGURE 2.8. *Fall-off of concentration from Brownian motion depending on particle radius.*

The calculations are included to give a feel for the sorts of effects that are relevant down in the nano domain.

For example, if a ZnO particle really has the bulk density of 5.3 then a 34 μm particle will tend to fall out of suspension in 10 days, so the Brownian motion has taken it only 2.4 mm. This is a significantly smaller relative effect. Even in this case, the square root of time dependency means that the relative stability isn't quite as bad as that large increase in density might have suggested, so the particles won't completely settle out. Using the 1/*e* argument, the increased density means that only 10 nm-radius particles will be stable rather than the 20 nm particles when the density is 2.

There are many assumptions behind these calculations, not the least of which is the assumption that the particles are spheres. Those who have particles of different shapes will need to use the "Spherical Equivalence Principle" to find an "equivalent radius" r and a "shape correction factor" K' suitable to their specific shape [7]. The spreadsheet includes an example which uses calculations for a spheroid falling with its most streamlined orientation; i.e., as a prolate spheroid. The spreadsheet calculates two versions of the equivalent radius and K' is derived from the aspect ratio of the spheroid. These values can then be plugged into the other formulae to see how, for example, a CNT might behave. A tube with diameter 2 nm and length 1 μm has a K' of ~50 and a Stokes equivalent radius of 3 nm (i.e., it falls with the drag equivalent of a 3 nm sphere) and a volumetric equivalent radius of 10 nm (from which its gravitational and Brownian motions can be estimated).

Going back to the velocity equation, this applies only to very dilute suspensions. As suspensions get more concentrated the particles interact with themselves. The generally accepted rule of thumb (to call it a law or formula would be to give it too much credence) is that there is a power-law dependency (with 5 being a typical number) on the free volume (i.e. 1- volume fraction φ) so that the actual velocity v is related to the undisturbed velocity v_0 by:

$$v = v_0(1-\varphi)^5 \tag{2.21}$$

2.5. THE RIGHT SHAPE AND SIZE

The final reason for going nano may not require nano in all directions. Sometimes micro in one direction is a positive requirement. This

section looks at why being nano in one or two directions but micro in the other direction(s) can be a positive advantage.

2.5.1. Percolation Theory

A network of particles touching each other is one way to achieve high strength for hardcoats, and is a necessity for achieving conductivity when using conducting additives.

At ~70 %Vol spheres become close-packed. For most purposes, this is far too much particle and even if it is technically possible, it might be unaffordable. Often, to achieve the same results an interconnected network that delivers most of the desired properties can be obtained at a significantly lower % addition. The onset of such a network is the *percolation threshold*. The phrase comes from the analogy of water percolating through soil where the focus is on individual (spherical) pores. At a low density of pores there is no pathway that allows the water to flow from pore to pore. At a level of pores where it is possible for water to percolate through the soil via at least one interconnected path, the pores have reached the percolation threshold.

In Figure 2.9, the top array of particles don't show any connected path from top to bottom. The lower array has hit the percolation threshold because, by chance, there is one path all the way through.

For spheres, percolation is reached at ~28% volume fraction and this number is independent of sphere radius. 28% is not too hard to for-

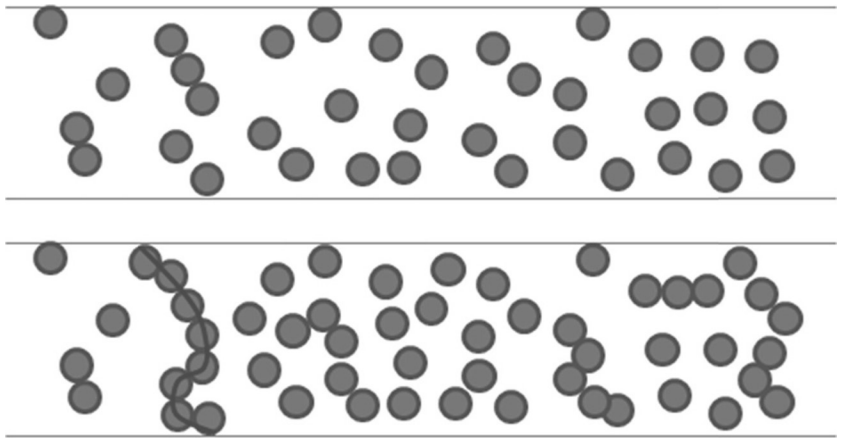


FIGURE 2.9. *A non-percolated (top) and just percolated (bottom) dispersion of particles in a coating.*

mulate with some nanoparticles (though a considerable challenge for many others). For producing hard-coats it can be argued that significant extra strength should be obtained by being just above the percolation threshold, and adding 30% silica might not appear at all difficult or costly. However, a 30% volume concentration of silica translates to a weight concentration of > 60%; this not easy to achieve industrially. Most silica nanoparticles dispersions are provided at between 30–50% concentration by weight. Attempts to produce dispersions containing greater amounts of silica generally suffer from viscosity and stability problems. It is possible to circumvent this problem by skillful formulating ability, but this is discussed later in the book.

To achieve electrical conduction through a nanocoating, a 28% volume concentration of particles is usually far too high to be practical because of the issues discussed in the previous paragraph and also because of the generally high cost of typical transparent conductors.

Fortunately, the laws of percolation theory (which are surprisingly complex) show an easy way to obtain percolation at a much lower concentration. If a particle (assuming for simplicity just two major dimensions such as length a and cylindrical diameter *b*) has a high value of the *aspect ratio* (AR), *a*/*b*, this leads to a low value of the percolation threshold. If the aspect ratio is 10 then 8.4% is required for percolation; if the aspect ratio is 100 then 0.6% is enough. These values are obtained via the Padé approximation for calculating the threshold [8]:

$$\text{Threshold} = \frac{S.AR + AR^2}{h + f.AR + g.AR^{1.5} + c.AR^2 + d.AR^3} \qquad (2.22)$$

where *c*, *d*, *f*, *g*, *h* and *s* are Padé constants shown in the Percolation spreadsheet, Figure 2.10.

For those who require percolation within their coating, conversations with nanoparticle suppliers should be focused as much on aspect ratio as on size. The contrast between conducting coatings with carbon black and with carbon nanotubes is striking. It needs 30% carbon black to be conducting and << 1% with CNT [9]. One practical nanocoating that utilizes this phenomenon is the ClearOhm nanowire silver that can achieve conductivities in the ITO range with ~1% of nanowires and, therefore, a transparency/color equal to or better than ITO [10].

Being at the threshold merely ensures continuity; it does not ensure that the desired property (such as conductivity) is high. To achieve this

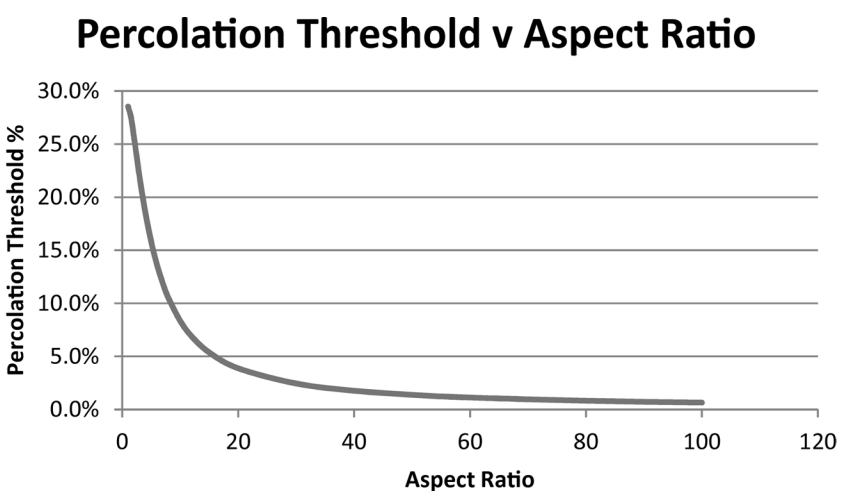


FIGURE 2.10. *Variation of Percolation Threshold with Aspect Ratio of the particles.*

it is necessary to go beyond the threshold. Using conductivity as a typical example:

$$\text{Conductivity} = (\phi - \phi_c)^t \tag{2.23}$$

Here ϕ is the volume fraction, ϕ_c is the percolation threshold and t is a power dependence that is generally between 1 and 2, depending on the system.

Although the theoretical percolation threshold is independent of particle size, because it is a statistical phenomenon it is possible to be beneath the threshold and yet still percolate, or to be above the threshold and not percolate. The zone of uncertainty is larger with larger particles. Because of this uncertainty, and because the theory deals with particles of a single size, there is no substitute for experimentation to find the optimum balance between the issues of adding higher levels of particles and the need for strong percolation to deliver the desired properties.

For those who have no choice but to accept a spherical particle, there is an elegant trick to achieve cost-effective percolation. Ensure that the bulk of the sphere is something low cost, low density and reliable like silica or carbon, then ensure an outer shell comprising the percolating material.

Given that conductivity near the percolation threshold depends mostly on the strength of the percolation network rather than the bulk conducting properties, by increasing ϕ to take advantage of the $(\phi - \phi_c)^t$

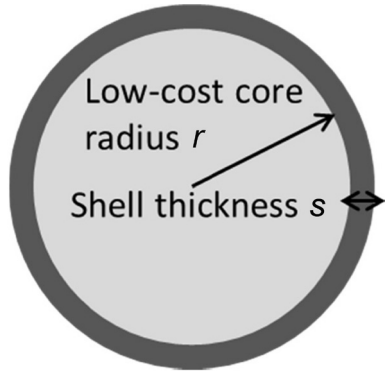


FIGURE 2.11. A core-shell particle of radius r and shell thickness s.

dependency, surprisingly good performance can be obtained cost effectively. Note, too, from the "shell" calculations that a relatively thin shell can contain a relatively large amount of material, so the sacrifices from the bulk core might not be as great as first appears. If the thickness, *s*, of the shell is 2 nm and the radius, *r*, of the core is 20 nm, then 25% of the volume is the shell, so there will be lots of conductivity with a 75% reduction in expensive material. If 1 nm is sufficient to gain conductivity then there is an 86% reduction in expensive material.

2.5.2. Barrier Properties

Because inorganic nanoparticles have essentially zero permeation rates for water, oxygen, flavors, etc., including them within a polymer film (extruded, printed, coated . . .) would seem a great way to provide barrier properties. Unfortunately, most particles are essentially useless in this respect. The particles may be beautifully distributed within the film while still leaving a simple path along which the permeant can travel. The reality is that only high aspect-ratio clays (and, perhaps, graphene) can provide significant barrier properties in practical systems.

Even for clays, the improvements to barrier properties can be minimal. Following the approach of Bharadwaj they have to get two proper-

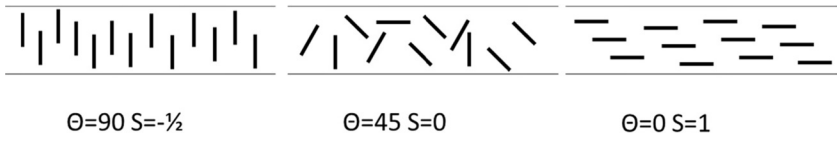


FIGURE 2.12. The "S" value for orientation of clay particles in a coating.

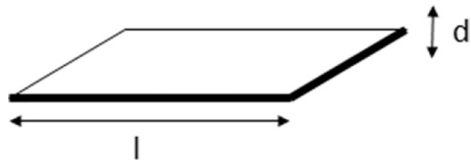


FIGURE 2.13. *Definition of Aspect Ratio = l/d.*

ties right at the same time [11]. The first is the *S* parameter that depends on the orientation θ of the particles, Figure 2.12.

$$S = \frac{1}{2}(3\cos^2\theta - 1) \tag{2.24}$$

When the clay particles are oriented at 90° to the substrate, $S = -1/2$. When they are random, effectively at 45°, $S = 0$. When they are perfectly aligned with the substrate, $S = 1$.

The second property is the Aspect ratio, *AR*, defined this time as the ratio of the length *l* of the edge of the nominally square particle to its (nano) thickness *d,* Figure 2.13.

$$AR = \frac{l}{d} \tag{2.25}$$

From these it is possible to calculate the dependence of the relative permeability P_r (permeability with particles divided by permeability without particles) on the fraction φ of particle. There are many ways to calculate P_r; the version by Bharadwaj seems to capture most elements of the situation in a single formula which depends on the orientation parameter *S*:

$$P_r = \frac{1-\varphi}{1+0.5AR\varphi\frac{2}{3}\left(S+\frac{1}{2}\right)} \tag{2.26}$$

The Clay Barriers spreadsheet allows you to play with these variables.

Given a typical high value of AR such as 50 it is possible to plot the dependence of permeability upon orientation parameter *S*, Figure 2.14.

The need to ensure a good, parallel alignment of the particles ($S = 1$) is clear.

Assuming that $S = 1$, the effect of AR is also clear—high values are crucial for good barrier properties, Figure 2.15.

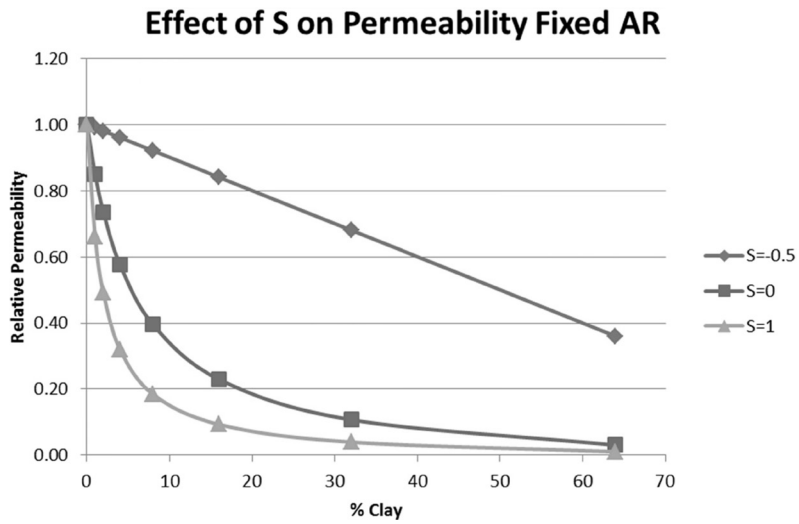


FIGURE 2.14. Effect of the orientation parameter S on Relative Permeability for a fixed Aspect Ratio and varying % Clay.

It is especially obvious that spherical particles (AR = 1) are of little value for providing good barrier properties.

Two case studies in later chapters show that these apparently simple principles are far from simple when it comes to creating real-world products with excellent barrier properties. One key issue is that clays are thought of as nice, cheap fillers. This contradicts the need for them to be totally colorless, free from bits of "dirt", and with a stable func-

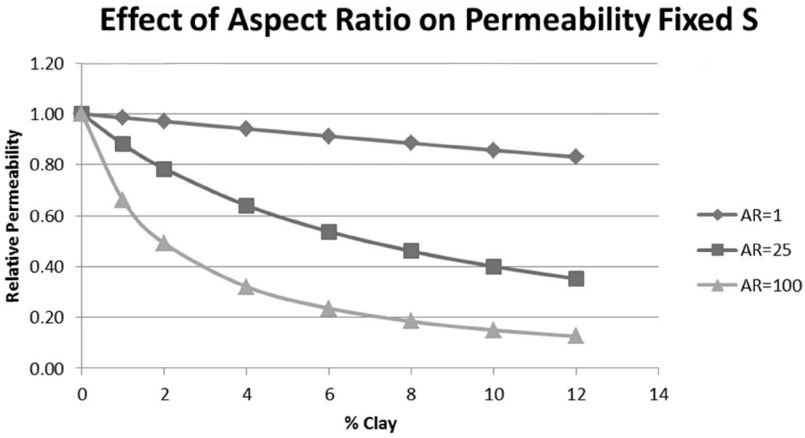


FIGURE 2.15. Effect of the Aspect Ratio on Relative Permeability for a fixed S parameter and varying % Clay.

tionality that allows every individual clay layer to separate (exfoliate) and become surrounded by the polymer in which it is dispersed.

In the above example, if the average clay particle is composed of 2 clay layers then AR is halved from 100 to 50 and at 6% loading the permeability increases from 13% to 24%. This makes it clear how important it is to get the correct formulation and processing to deliver 100% exfoliation of the clays.

And as we know from the section on scattering, relatively small percentages of relatively large particles ("dirt") can lead to a high level of visual scatter that would make a product unsuitable for packaging or bottling applications.

2.6. CONCLUSION

A few key formulae govern the world of nanoparticle behavior. By having them in spreadsheet format, the team can explore "what if" scenarios so that the problems and the promise can be made clear. Then the decision whether or not to go down the nano route is owned by everyone.

2.7. REFERENCES

1. Jonathan Moghal, Johannes Kobler *et al.* (2012) High-performance, single-layer antireflective optical coatings comprising mesoporous silica nanoparticles. *ACS Applied Materials and Interfaces* 4:854–859.
2. H.C. van de Hulst. (1957) *Light scattering by small particles.* New York: John Wiley & Sons.
3. Victor I. Klimov. (2010) *Nanocrystal Quantum Dots (2nd Edition)*. Boca Ratonk FL: CRC Press.
4. Michael J. Coutts, Michael B. Cortie, *et al.* (2009) Rapid and controllable sintering of gold nanoparticle inks at room temperature using a chemical agent. *Journal of Physical Chemistry C* 113:1325–1328.
5. Andras Komaromya, Reinhard, I. Boysen, *et al.* (2008) Influence of surface nanostructure on the extent of colonization and cell viability of *E. coli* and *S. aureus*. *Biomedical Applications of Micro- and Nanoengineering IV and Complex Systems,* edited by Dan V. Nicolau, Guy Metcalfe. Proc. of SPIE, 7270:727006-1 to 727006-8.
6. D Lerche. (2002) Dispersion stability and particle characterization by sedimentation kinetics in a centrifugal field. *Journal of Dispersion Science and Technology* 23:699–709.
7. Claus Bernhardt. (1994) *Particle Size Analysis: Classification and Sedimentation Methods.* London: Chapman & Hall. Chapter 3.1.3, pp. 32–47.
8. E.J. Garboczi, K.A. Snyder, *et al.* (1995) Geometrical percolation threshold of overlapping ellipsoids. *Physical Review E.* 52:819–828.
9. J.K.W. Sandler, J.E. Kirk, *et al.* (2003) Ultra-low electrical percolation threshold in carbon-nanotube-epoxy composites. *Polymer* 44:5893–5899.
10. http://www.cambrios.com Accessed June 2012.
11. R.K. Bharadwaj. (2001) Modeling the barrier properties of polymer-layered silicate nanocomposites. *Macromolecules* 34:9189–9192.

CHAPTER 3

Finding the Right Nanoadditive

3.1. INTRODUCTION

THE properties of a given type of nanoparticle depend strongly on how they are made. Understanding the strengths and weaknesses of each type of manufacturing route is important to all aspects of a nanocoating project. The lab team might get wonderful performance from a small, specialist supplier who can fine-tune the properties. The production team might be able to source the "same" particle at low cost from a large-scale manufacturer, only to find that the functionality is significantly different.

It is not necessarily the case that the particle you require can be manufactured identically using different methods. However, in terms of the amount of material required, some methods will be more appropriate than others. Do not choose a manufacturer who is producing for a test-tube scale if the requirements are for tons of material. The team needs to carefully construct the project plan before starting work.

The means of manufacture are wide and varied, ranging from gentle precipitation to controlled explosions. There is no best method, although some processes are better suited to certain products than others. There are many nanoparticle manufacturers worldwide and a wide selection can be found via dedicated internet sites such as http://nanoparticles.org [1].

As with all work involving the use of nanoparticles, the watchword is consistency. To an extent, it does not matter whether the manufacturer is large or small, provided that they can demonstrate consistency of product from their favored manufacturing method so the user can feel more confident in the chosen raw materials. This requirement is, of

course, more difficult to meet for a start-up company; therefore, the customer may have to become rather more closely involved with the supplier than is normally the case. This chapter may therefore be of some practical assistance. This final point should not be regarded as a disadvantage; indeed, a good case can be made for more proactive supplier/customer relationships in this field. If a supplier is proving incapable of providing consistent material then the best option might be to bring the process in-house, though this is not a step for the faint-hearted.

3.2. MANUFACTURING METHODS

Basically there are two major manufacturing strategies; these are generally referred to as "bottom-up" and "top-down". The former includes techniques where the particle is built from the molecular level by combination; the latter refers to methods whereby larger particles are processed to smaller sizes by techniques such as grinding. Both approaches have their advantages and disadvantages. This separation of techniques is convenient and easy to remember, but it is a simplification of reality. In fact, for the production of nanomaterials on an industrial scale, the two classes are often combined into a unified whole. The primary particle may be produced from a vapor phase process which produces (loose) agglomerates which are then reduced in size and stabilized using a milling process. Nanoparticle dispersions are often produced by this step-up/step-down procedure.

It must be remembered that there is no perfect way to make a nanoparticle. In the descriptions that follow, strengths and limitations are described as objectively as possible. Users should never underestimate how hard it is to make and supply nanoparticles.

It is not the intention of this chapter to present a fully comprehensive review of all the nanostructure manufacturing methods; rather, a brief overview will be provided with the bulk of the chapter being dedicated to issues that are of particular interest to the nanoparticles user. It should be noted that within the methods shown below, there are many variants on the basic process. A very useful review of manufacturing methods has been presented in Zhong's recent book on nanoscience and nanomaterials [2].

3.3. BOTTOM-UP METHODS—VAPOR PHASE

The common rationale of these varied manufacturing methods is that

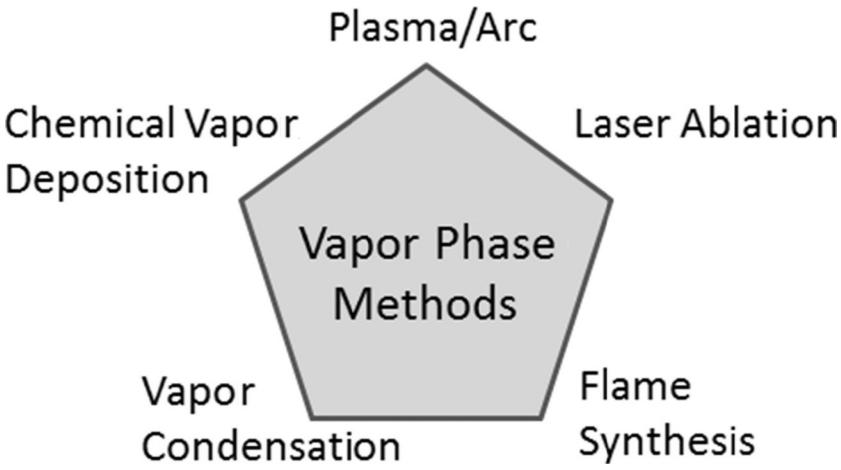


FIGURE 3.1. *Vapor Phase nanoparticle synthesis methods.*

it is simpler to build the required particle from a molecular source than it is to try and physically grind down a micron-sized starting material.

Vapor phase synthesis is the most commonly practiced method of producing metal and metal oxide nanoparticles. The energy source and the physical nature of the fuel used to produce the nanoparticles can be varied in many ways to produce a variety of different techniques that sit under this banner, Figure 3.1.

3.3.1. Vapor Condensation

Vapor condensation is perhaps the simplest of the methods and other techniques derive from it in many ways. The process is generally dual stage: generation of a supersaturated vapor is followed by condensation of the vapor to produce the nanoparticles. It should be noted that of all of the vapor phase production methods discussed, the most important factors affecting particle growth are the partial pressure of the gas, gas flow rate and temperature.

Physical vapor condensation (PVC) is a method that is particularly suited to the production of metal nanoparticles, especially gold [3] and silver [4]. The bulk metal is vaporized in an inert gas atmosphere to produce a supersaturated vapor, which, at the correct concentration, will condense to form the primary particles. On quenching the mixture with a cold inert gas, the primary particles begin to grow in size due to coagulation and coalescence.

It is important to remember that, unlike liquid phase synthesis, there is no opportunity to stabilize the particle; as a result of this, agglomeration is bound to occur. It is therefore necessary to operate the process using conditions that produce a loosely agglomerated final product.

Chemical vapor condensation (CVC) has certain advantages over the PVC process, being able to operate successfully using low vapor pressure materials. There are a wide range of raw materials that are commercially available which are particularly suited to this process—metal hydrides, halides and organometallic compounds are typical examples. Process operating procedures are similar to the PVC process, but the required temperatures are generally lower. Particle growth follows a similar mechanism, but yields are generally greater.

The nucleation process associated with CVC is basically similar to that found with the PVC process. By judicious choice of reactant gases and conditions, it is possible to produce metal alloy nanoparticles by this route. Using Iron and Cobalt Carbonyl as starting materials, nanoparticles comprising an Iron/Cobalt core surrounded by a passivating layer of oxide have been produced [5]. It should be remembered that most metals rapidly form a protective oxide layer on their surface (passivation). The growth of this layer can be controlled during the CVC process by addition of oxygen in small quantities as part of the reactant gas.

3.3.2. Chemical Vapor Deposition

Chemical vapor deposition (CVD) is a powerful and versatile manufacturing technique and has been used successfully to produce a wide variety of nanoparticles of varied sizes and shapes. It has been used to produce wires, tubes and rods at the nanoscale. It is also capable of producing coatings and patterned three-dimensional arrays. A useful review of the technique has been provided by Choy [6].

The CVD and CVC processes are rather similar; most of the procedures that apply to one apply to the other. The operational control of temperature, pressure and flow rate is critical to both processes. A difference between the two processes is that CVD only involves heterogeneous reactions between the substrate and the reactant gas. It is perfectly possible to view either technique as a variation of the other.

There are many variants within the method but all have a common factor. Particle nucleation takes place between the reactant gas and a heated substrate; it is the heterogeneous interactions at the interface that

helps to determine the shape and nature of the product. Several mechanisms have been proposed to explain the structural growth that takes place during the process, for example: Oxide Assisted Growth (OAG), Vapor/Solid (VS) and Vapor/Liquid/Solid (VLS).

Metal catalyzed CVD is probably the most established and understood CVD process. The metal forms a eutectic alloy with the source material and then acts as a source for particle growth. The vapor generated from the target diffuses into the metal droplets. The size distribution of the metal droplet determines the diameter of the growing structure. It is believed that the VLS mechanism predominates in this technique, which is a particularly effective means of growing nanowires.

The *modus operandi* of the metal-catalyzed CVD process is simple and powerful but there is a disadvantage attached (literally) to the process. The resulting nanostructures are contaminated with metal impurities which may compromise properties that the product requires. Therefore, a purification step is required in the process which increases overall costs and reduces yields.

It is possible to operate a catalyst-free process which overcomes the contamination at a stroke. However, the mechanisms involved in these kinds of process are not as well understood as for the metal catalyzed route; it is quite possible that self-catalysis plays a role in some of these methods. This lack of understanding is a barrier to industrial scale-up. An exception to this is metal/organic CVD (MOCVD), which is capable of producing large scale coatings on a variety of substrates. The fact that the process utilizes organometallic compounds of relatively low thermal stability means that the reactant vapor of metal and metal oxide nanoparticles can be generated at low temperature, a great advantage for an industrial process.

There are many more variants on the basic CVD process. For further information on these processes, their strengths and weaknesses, the reader is referred to Zhong's review [2].

3.3.3. Flame Synthesis

Flame or combustion synthesis is a process that has the advantage of being suitable for large scale production at relatively low cost. The particles are created within a flame, the combustion reaction providing the necessary heat for nucleation and particle growth. The oxidative environment makes this a very good way of producing metal oxide nanoparticles.

TABLE 3.1. Some Products from Flame Synthesis.

Material	Size, nm	Specific Surface Area, m^2g^{-1}	Phase Identification
Al_2O_3	10–15	~100	Mixture of γ- and δ-Al_2O_3
CeO_2	10–15	80–100	Cubic CeO_2
ZnO	8–15	60–90	Mainly tetragonal
TiO_2	25	80–100	Mainly Anatase and trace of Rutile
Doped TiO_2	30	90–100	Mainly Rutile and traces of Anatase

The ability of the process to operate on the industrial scale has boosted the availability of nanoparticle materials significantly. The large-scale availability of fumed silicas such as the Aerosil has led the way, encouraging others to make the leap from academic to industrial quantities [7]. It should be noted, however, that products such as Aerosil are largely agglomerated. A discussion on the reduction of these agglomerates is presented below. Interestingly, the agglomeration means that when such particles are purchased in large sacks, they are not technically classed as bags of nanoparticles—defusing many of the nanosafety issues during shipping and handling.

An example of fine-tuning particles via this method is the pilot-scale flame synthesis reactor being used by Johnson Matthey [8]. Typical products produced from this reactor are shown in Table 3.1. The "size" parameter is the primary particle size, which is very different from the agglomerated size.

The flame used within the process is capable of being configured in different ways and a variety of different gases, both inert and combustible, may be used. These factors all play a role as to the size and shape of the particle produced. The gas flow rate also plays a significant role in flame configuration and hence the nature of the final product.

It would be very useful if metallic and metallic alloy particles could be produced using this method, given its merits as a truly industrial scale process. However, there is a problem that needs to be considered; the flame produced is usually oxidative, and this is a serious disadvantage if we need non-oxidized products. The problem can be overcome by using a non-oxidative flame or a reducing atmosphere. In the former case, a chloride flame may be used, with the metal chloride being burned in the presence of sodium. This produces the metal particle which is coated with a layer of sodium chloride which is readily removable. The latter case can be achieved by either limiting the input of oxygen or by using

a gas stream rich in a reductant such as hydrogen or carbon monoxide. This approach has been used to produce pure iron nanoparticles from ferrous and ferric oxide [9].

3.3.4. Plasma/Arc-discharge

The plasma or arc-discharge process utilizes the evaporation of the source material through an electrical arc in the presence of a background gas, thereby producing a plasma. The gas used can play either an active role in the reactions, or can be simply the source of the plasma. This is a fundamental difference to the vapor deposition processes discussed earlier.

The electrode assembly may comprise a solid metal ingot as the anode and source material; alternatively, it might comprise a solid dispersed in graphite, which is then formed into the anode block. The arrangement depends on the particle being produced. The cathode is usually made from carbon or tungsten; its distance from the anode being of the order of several millimeters. It must be remembered that the anode is consumed during the process, altering the distance of the electrode gap. If unchecked, this can destabilize the arc; therefore, active control of gap distance must be maintained during the process.

Depending on what process is being carried out, the time of discharge can vary from several seconds to several hours. The arc operates at several thousand degrees centigrade, so active cooling of the electrodes is required. This huge temperature gradient between the arc and the cooling system is believed to be one of the driving forces for particle nucleation in the process.

The process is capable of producing a wide variety of nanomaterials and is particularly associated with the production of carbon nanotubes. Core-shell nanoparticles may also be produced using this method. With careful attention to the problems of large surface areas of pyrophoric materials, pure metal nanoparticles (such as iron) can also be produced on a large scale, making this route generally preferable to the flame route with chlorine and sodium.

An interesting variation on the standard manufacturing technique is the possibility of running the process with the electrodes submerged in a liquid environment. The electrodes may be submerged in deionized water, and when the arc is struck the source materials are vaporized as bubbles within the water, which also confines particle growth. In this respect, there are similarities with the laser ablation technique dis-

cussed in the next section. This modification to the technique is claimed to make the process less expensive, more environmentally friendly and more versatile. A variety of materials and structures has been produced using this modification, which appears capable of being used for consistent industrial scale production.

3.3.5. Laser Ablation

Laser ablation is a further way of generating a gaseous nanoparticle precursor. The factors critical to production control are: laser operating wavelength, pulse duration, pulse energy, energy per unit area and repetition rate; related to these operating parameters is the angle of attack of the laser. Pulsed laser ablation can be carried out in vacuum, gas or liquid environments; the latter is of interest, as it offers the possibility of directly synthesizing colloidal nanoparticles. The technique also has particular use in nanopatterning.

The choice of laser is relatively wide: Nd-YAG, Ti-sapphire and ArF excimer have all been used for this process. Conventional laser ablation uses nanosecond pulses where interactions between the laser and all three phases of matter may take place. It is also possible to carry out femtosecond ablation; the short pulse duration means that there is insufficient time for the substrate target to become heated with consequential melting. The irradiated target is converted directly to vapor. This is ex-

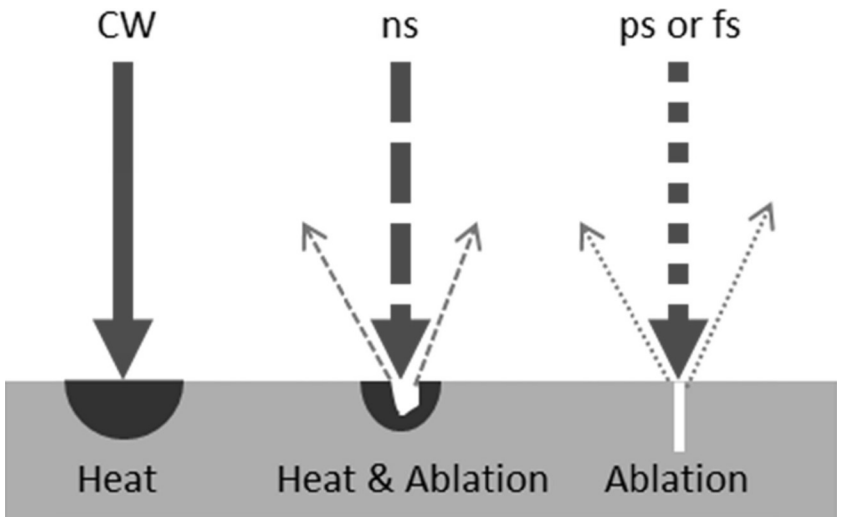


FIGURE 3.2. Laser ablation modes.

tremely useful in nanopatterning, where the occurrence of rims around the drilled areas of the substrate can create problems. The figure shows the differences between continuous wave (CW), nanosecond (ns) and superfast (ps or fs) pulsing.

The femtosecond process is much cleaner than the picosecond and nanosecond pulse methods and this probably contributes greatly to the generation of smaller and more tightly distributed particles. The difference in quality and cleanliness of the surface ablation, along with the theoretical background, is well described by Chichkov and co-workers [10].

As mentioned, laser ablation can take place in a variety of media. The possibility of carrying out the reaction in water allows for extra control over particle properties. While superficially similar to the gas phase process, there is a significant difference. The plasma plume, produced when the laser beam impinges on the target, compresses the surrounding liquid, producing a shock wave. Interactions between the plasma plume and the shock wave lead to increased temperature, pressure and density at the interface between the plume and the liquid. In such an environment reactions can take place within the plasma, at the liquid/plasma interface and within the liquid. Therefore:

- It becomes possible to produce colloidal nanoparticles.
- Surface modification of the particles becomes easier.
- The liquid phase helps to produce and stabilize metastable structures.
- The liquid phase allows a more rapid rate of cooling, resulting in smaller and denser particles.

A wide variety of materials have been manufactured using this technique; these include metallic colloids, diamond-like carbon and core-shell particles.

The nature of the liquid phase can determine what particles are produced during synthesis; an example of this is shown in Table 3.2.

TABLE 3.2. Different Co Forms from Laser Ablation in Different Solvents.

Target	Ablation in Hexane	Ablation in Water
Co	Co	Co_3O_4
Co_3O_4	Co	Co_3O_4
CoO	CoO	Co_3O_4

TABLE 3.3. Effects of Surfactant Concentration on Silver Particle Size.

Molarity of Solution	Particle Size (nm)
0.05	12 ± 5
0.01	15 ± 8
0.003	16 ± 4

Similarly, the presence of additives in the liquid phase can affect particle size growth. The experimental results for silver nanoparticles produced in different concentrations of sodium dodecyl sulfate are shown in Table 3.3—though admittedly the effects (compared to the standard deviation) are not large [11].

3.4. BOTTOM-UP METHODS—LIQUID PHASE

Liquid phase synthesis provides an alternative bottom-up approach to vapor deposition and similarly comprises a wide range of techniques. The approach has the advantage over vapor deposition in that no expensive energy source is needed. The reactions can largely be carried out at low temperatures and at atmospheric pressure, making the process cost- and energy efficient. A further advantage is that the method allows for greater control of the reaction stoichiometry; this enables the synthesis of ceramic materials of unusual composition. The liquid phase reaction also allows the easy introduction of stabilizing agents, which can be used to modify the surface properties of the nanoparticle.

The disadvantage of the wet route is that, in addition to the nanoparticles pre-cursor, other additives (e.g., pH modifiers) need to be present to ensure the success of the reaction. This results in contaminants that need to be removed from the final product. As a result of this, the products generated by liquid phase synthesis are generally not as pure as those obtained from vapor deposition methods. However, this has not prevented the commercialization of processes that provide a means of industrial scale production, Figure 3.3.

3.4.1. Precipitation

Chemical precipitation is a simple and convenient means of producing metal oxide and ceramic nanoparticles with controlled composition. The starting materials, which are usually metal oxides, salts, or organometallic compounds, are dissolved in solvents, which can include

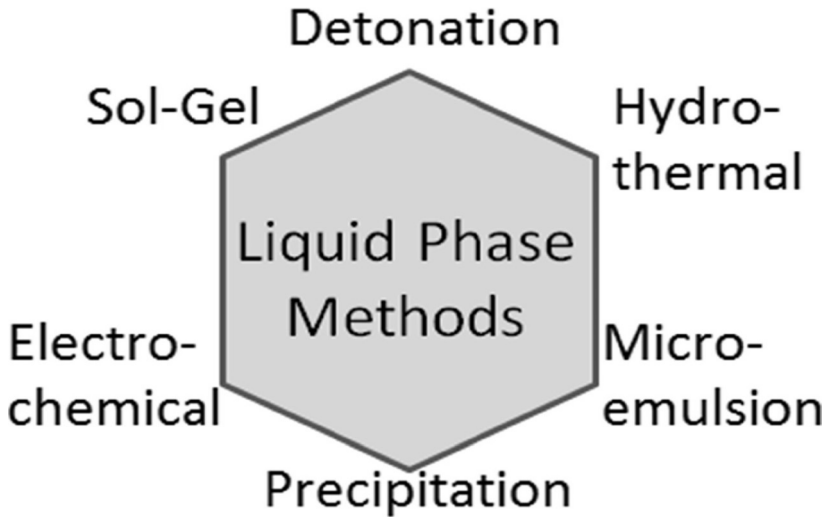


FIGURE 3.3. *Liquid Phase nanoparticle synthesis methods.*

water, alcohols and acids. The prepared solutions are then mixed and stirred vigorously; the product is then allowed to precipitate from solution. This is the simplified reaction scheme. It is not unusual for other solvents to be added to aid precipitation. To obtain larger and purer particles, it is not unknown for the collection of the precipitated material to last for several days. Naturally, the purer the starting materials, the purer will be the product.

A general review of the methods employed is provided by Zhong and co-authors in the text referenced earlier [2]. It is however useful to consider the theory of precipitation in rather more detail and this is provided in the following case study.

3.4.1.1. Case Study: Barium Sulfate

To consider the theory of the precipitation process in a little more depth, this study reveals some subtle physical chemistry taking place in what appears to be quite a mundane reaction.

Making an insoluble particle such as barium sulfate seems to be easy: react barium chloride and potassium sulfate with good mixing and the highly insoluble sulfate will simply fall out of solution. Intuitively, if the mixing is very rapid and there is no opportunity for the particles to pick up further material (i.e., they are removed from the system as soon as they are formed), particles of arbitrarily small size can be produced.

The key barrier to this scheme is the impossibility of crystals forming without a seed. If some 1 nm seeds were to appear in vast numbers then crystals of, say, 10 nm size could rapidly form. But where do 1 nm seeds come from? They come from 0.5 nm seeds which in turn . . .

If, instead, there are only ever a few seeds (e.g., some random dirt in the process) then, in principle, the whole precipitation would end up as a few large lumps centered on those seeds. This central dependency on seed formation is the key reason why nanoparticles via precipitation are not as common as might be imagined.

In this simple introduction to precipitation, a single particle will be considered—Barium Sulfate. Some standard approximations reveal that a seed must be 5 molecules or ~1 nm in size. For small particles it helps to have a large supersaturation ratio, S, which depends on the concentrations of the anion and cation and the solubility product, K_{sp} which is [Anion].[Cation] at saturation (assuming a 1:1 Anion:Cation ratio):

$$s = \sqrt{\frac{\text{Anion.Cation}}{K_{sp}}} \tag{3.1}$$

This high ratio helps drive the formation of seeds. Higher temperatures T also help seeds form (provided they don't decrease S). Smaller molar volumes MVol make it easier for the seeds to form, and finally the surface energy σ should be as small as possible because the energy needed to create a seed goes as σ/radius, which is why the initial small particles are so hard to produce. Combining these terms, with Boltzmann's constant k and a base rate B_0 (typically 10^{36} $(m^3s)^{-1}$), the theoretical rate of seed formation, B is given by:

$$B = B_0 \exp\left[-\frac{16\pi\sigma^3 \ \text{MVol}^2}{3(kT)^3 \ln(S)^s} \right] \tag{3.2}$$

It is common practice to replace this complex formula by a simpler approximation, but before doing that, the influence of σ needs to be re-emphasized. Dispersion agents added to precipitation reactions can have a big influence on particle size. In addition to stopping agglomeration of particles, it seems reasonable to assume that the agents can also decrease the surface energy and therefore increase the rate of seed formation. The speed at which the dispersing agent can migrate to the fresh

surface is a key parameter; thus, smaller molecules in solution (rather than as micelle agglomerations) are generally more suitable.

The simplifying approximation gives B as:

$$B = k_s(c - c_{sat})^b \tag{3.3}$$

where c is the concentration of ions, c_{sat} is the saturated concentration and b is somewhere between 5 and 15. In other words, high rates of seed formation are created when there is a big difference between the concentration of the ions and the saturated concentration.

The other key equation tells us the growth rate G of the particles once they have formed. The simplified version is similar, with the exponent g typically being 2:

$$G = k_g(c - c_{sat})^g \tag{3.4}$$

When $(c - c_{sat})$ is large, B overwhelms G, so particles are small, so making nanoparticles of barium sulfate is relatively easy by precipitation. Making nanoparticles from a saturated solution of sucrose is impossible because G far outstrips B to produce sugar crystals.

Experts in precipitation know that the above summary covers a multitude of complications. But it is sufficient for the practical nanoparticle user to see the sorts of issues involved.

A paper on barium sulfate particles brings these ideas to life [12]. As the feed-rate increases, the supersaturation value increases and particle sizes decrease from ~250 nm down to ~100 nm. A more explicit experiment based on supersaturation value shows a decrease from ~200 nm down to ~80 nm as the supersaturation increases from 2,500 to 25,000. The effect of a dispersing agent in this particular case was merely to inhibit agglomeration to ~4 μm. The agent itself had no significant effect on the primary particle size. As the paper points out, dispersing agents can also act as growth inhibitors, but again in this example there was no evidence of such an effect. Perhaps, as discussed above, the agent was too slow to migrate to the fresh surface.

The above paper used a conventional stirred tank reactor. It was natural to suspect that, without perfect mixing, particles might intercept fresh reagent and therefore grow larger, rather than allowing the fresh reagent to create fresh seeds. A paper which puts the above ideas into rigorous simulation was able to match theory and experiment to show that the better the mixing (using a high-pressure T-mixer) the smaller

the particles, down to ~40 nm [13]. The paper also showed that these particles took about 300 μs to form.

With nanoparticles there are no free lunches. Smaller particles require higher powers.

An alternative route to barium sulfate nanoparticles via precipitation relies on thermodynamics rather than raw power. To understand that route, we need to understand microemulsions, discussed below.

3.4.2. Sol-gel Synthesis

Sol-gel is a term that covers a vast range of chemistries and materials [14]. In this context, it means that a soluble form of a chemical such as silica (e.g., an alkoxide) is carefully reacted so that clusters of the (now insoluble) particle are formed. The "gel" part of the terminology is probably not helpful—if a gel has been formed, the chances are that the whole system is a network, not what is required for particles. The archetypal sol-gel nanoparticle process is from Stöber; a solution of TEOS (tetraethyl orthosilicate) in alcohol is mixed with a solution of ammonia and controlled amounts of water. The orthosilicate is hydrolyzed to form a growing network of silica particles. The impressive thing about the Stöber method is that it naturally creates particles of a very controlled size, though with an open/porous structure which is either something required or something to be fixed by subsequent densification processes (such as heat). In contrast, the same sort of reaction using acetic acid (plus some water) tends to give a less uniform particle size, though the particles tend to be denser.

There is nothing magical about the Stöber process—the considerations about seeds and rates of growth discussed in the barium sulfate case study are still relevant because the nascent silica particles in this "bad" (for them) alcohol solvent cannot grow into a broad polymer network. It is possible to control the process by adding seeds—with the advantage that the seeds might be a "core" onto which the sol builds a "shell". Such core-shell capabilities are an important aspect of sol-gel. For example, given that TEOS is relatively cheap and Stöber-like processes can generate large amounts of low-cost, uniform seeds, these can be the basis of a further reaction through the addition of alkoxides (etc.) of more exotic elements, thereby producing particles with the external characteristics of the exotic elements and the basic cost of a silica particle.

Because the process is essentially happy in an alcohol environment, and because the acids and bases used (e.g., ammonia and acetic acid)

are relatively volatile, it is possible to produce particles in a wide variety of final forms—from powders to solutions in alcohols to solutions in acrylates.

Similarly, the mild, controlled environments (the precipitation is much slower than in the barium sulfate case) make it easy to add extra functionalities such as dispersant chemicals locked onto the outside of the particles. The process can also be combined with *in situ* polymerization reactions so that the particles and the polymer are co-produced, thereby eliminating the need to disperse the particles in a subsequent step into the (reluctant) polymer. A typical example of a more sophisticated sol-gel system is the Ormocer® from Fraunhofer [15].

A variation on the classical sol-gel approach has been published by Schönstedt and co-workers [16]. A sol-gel formulation based on TEOS was combined with pyrogenic silica (Aerosil) to produce coatings that are claimed to show good abrasion resistance. It was found (unsurprisingly) that stabilization of the silica within the reactive medium varied with the type of Aerosil employed. Milling of the silica was also found to be necessary and this is discussed later in this chapter.

There are industrial processes that, although they cannot be classically described as sol gel systems, do bear similarities. It is important to consider these methods, as they include one of the most successful processes for manufacturing nanoparticle dispersions on an industrial scale. This was pioneered by the German company NanoResins, which is now part of Evonik.

Descriptions of the NanoResins process can be found in the patent literature [17]. A silica sol is passed through an acidic ion exchange resin to reduce the pH of the sol to 2–3. At these conditions, a variety of trimethoxysilanes may be added to produce a stabilizing shell around the nanoparticle. Analytical studies have shown that for every 100 nm^2 of particle, approximately 1 to 100 stabilizing groups are attached.

At this point in the patent it is claimed that the treated silica particles may be extracted into an organic solvent, the water removed and the organic phase dried. Evaporation of the solvent leaves (it is claimed) clean, free flowing silica that can be dispersed in the solvent of choice.

More interesting, however, is the possibility of adding an excess of solvent such as isopropanol to the mix and using the azeotrope produced to remove the water, leaving the silica dispersion in the remaining isopropanol. At this point, it is possible to add radiation-curable acrylate monomers or oligomers to the mix and, after removal of the remaining isopropanol, a silica nanoparticle dispersion in a radiation curable acry-

late remains. It is possible to produce dispersions of consistent quality, which may contain up to 50% silica by weight, by this process. Also of importance is the fact that the process can be operated at the ton scale.

A wide variety of silica dispersions in different acrylate monomers and oligomers are available under the commercial name of Nanocryl™. The manufacturing method has also been extended to produce silica dispersions in UV cationically-curable resins known under the commercial name of Nanopox™.

A somewhat similar industrial process has also been patented by Advanced Nanotechnology Ltd. of Australia to produce Zinc Oxide nanoparticle dispersions in radiation-curable acrylates [18].

The take-home message is that if your particle can be produced from something like an alkoxide, then there are sophisticated possibilities available if they are necessary for a specific application. Versatility comes at a price: sol-gel reactions are complex and can head off into many different directions—the addition of a well-intentioned modifier can easily send the reaction into an undesired pathway.

Interestingly, the sols can be made in oils, and the acid/base reactions take place in an aqueous emulsion environment, as discussed in the next section.

3.4.3. The Microemulsion Method

The microemulsion method comprises a group of synthetic systems that are carried out in a thermodynamically stable colloid comprising two immiscible fluids. An amphiphilic surfactant is necessary in the reaction scheme to form micelles (or reverse micelles) which contain the reactants. It is an unfortunate historical accident that "microemulsions" means "nanoemulsions" (because they are intrinsically nanosized) and "nanoemulsions" has a specialist meaning that happens to be largely irrelevant to the manufacture of high-quality nanoparticles—though they can be used if desired.

The method has the possibility of finely controlling particle size, distribution and shape. In essence, the method is very simple:

- take a solution of a metal salt in water
- add a few % of the correct surfactant
- add an excess of an "oil" (to use surfactant nomenclature) such as heptane
- mix gently

this spontaneously creates a uniform dispersion of nanosized aqueous drops in the oil. Now:

- add an oil-soluble reactant (e.g., a reducing agent) with gentle stirring

or

- add a similar aqueous microemulsion where the water drops contain a suitable reactant

The reactions take place within the drops without significantly changing their size, so this gives a dispersion of perfect nanoparticles, which are automatically surrounded by their own dispersing agent (the surfactant) ready for further processing.

This all sounds too easy to be true, and yet it can really work as described. The microemulsion forms spontaneously driven by the laws of thermodynamics and the size of the drops is governed by the simple fact that the more surfactant that is present, the bigger the interfacial area and, therefore, the bigger the drops.

The difficulties exist mainly in finding the right combination of surfactant and oil.

Ignoring, for the moment, the issues of specific interactions between the metal salts and the surfactants—i.e., making the assumption that the salt can be modelled as if it were equivalent to a standard saline (NaCl) solution—then a simple equation exists which allows the formulator to design the appropriate system. The equation and its practical implementation is based on work by Salager, Sabatini, Aubry, Acosta and others, and a full description of it (along with software to implement it) is available on the Abbott website [19]. First, choose an oil by its Effective Alkane Carbon Number (EACN). For linear alkanes this number is the actual carbon number; 6 for hexane, 10 for decane, etc. For other oils, the number can be found experimentally. For example, toluene has an EACN of 1 and the EACN of cyclohexane is 3. Now define a Hydrophilic Lipophilic Distance (HLD):

$$\mathrm{HLD} = f(\mathrm{Salinity}) - 0.17EACN - a.\Delta T + Cc \tag{3.5}$$

where f(Salinity) is a function of the salinity which is ln(Salinity) for ionic surfactants and 0.13 × Salinity for non-ionics, ΔT is the difference

in temperature from the reference value of 25°C, a is 0.01 for ionics, –0.06 for ethoxylates, and 0 for many sugar-based surfactants, and Cc is a "characteristic curvature" for a given surfactant or surfactant blend.

When HLD = 0 the "optimal surfactant" blend is obtained where the interfacial energy is very small and there is the highest solubility of oil in water and water in oil. This "optimal" situation has drops of "infinite" radius (i.e., the curvature of the drops is effectively 0) and is therefore not of direct use for creating nanoparticles. However, it is a key navigation marker in surfactant formulation space. If the salinity, oil and surfactant (blend) are tuned to this point, then the formulation is very close to the desired water-in-oil blend. Being very far away from this point provides essentially no indication of how to adjust temperature, oil, salinity or surfactant to reach the desired point.

This optimal point is a so-called Winsor Type III. By going to a small positive value of HLD (by decreasing the EACN of the oil, increasing salinity or changing the surfactant blend to a slightly higher *Cc* value) you are then in the correct Type II domain, where essentially all the added surfactant is at the oil/water interface and where it is now possible to control the drop size by fine-tuning the relative levels of surfactant and water, as discussed below.

It is also possible to fine-tune the HLD by adding alcohols; it is common to see nanoparticle recipes that include alcohols. Alcohols add further complexity to an already complex formulation space. In general, it is better to formulate rationally with EACN blends or surfactant blends rather than add the complexity of alcohols.

3.4.3.1. Blending Rules

It is a fact that is frequently observed but seldom explained that single surfactants are used less than surfactant blends. Within the context of HLD theory the reason is obvious: it is highly unlikely that any single surfactant will have exactly the right *Cc* to give the correct HLD value. In fact, many common surfactants are very far from being appropriate as they tend towards high positive and negative extremes of *Cc* values.

Given an unsuitably low *Cc* surfactant, blending with an unsuitably high *Cc* surfactant is the only way to get a good formulation. The blending rule is one of molar weighted average where Wt*x* and MWt*x* are the weight and molecular weight respectively:

$$Cc_{blend} = \frac{\frac{\mathrm{Wt}_1 Cc_1}{\mathrm{MWt}_1} + \frac{\mathrm{Wt}_2 CC_2}{\mathrm{MWt}_2}}{\frac{\mathrm{Wt}_1}{\mathrm{MWt}_1} + \frac{\mathrm{Wt}_2}{\mathrm{MWt}_2}} \tag{3.6}$$

For a mixture of oils, the data show that the blending rule is a simple weighted average:

$$\mathrm{EACN}_{blend} = \frac{\mathrm{Wt}_1 \mathrm{EACN}_1 + \mathrm{Wt}_2 \mathrm{EACN}_2}{\mathrm{MWt}_1 + \mathrm{MWt}_2} \tag{3.7}$$

The Salinity used in the HLD equation always refers to NaCl. For other salts of charge Z the effective salinity (in g/100 ml water) is calculated from the real salinity *S* as:

$$S_{effective} = \frac{58S}{\mathrm{MWt}_{salt}} \frac{2}{1+z} \tag{3.8}$$

where 58 is the MWt of NaCl.

3.4.3.2. Droplet Size

Near the optimal surfactant formulation it can be assumed that all the surfactant is at the oil/water interface rather than as a solution in the oil (for w/o emulsions) or as micelles in the water (for o/w emulsions). The droplet size can then be calculated in one of two ways. The first turns out to be a useful rule of thumb:

$$r = A + BR \tag{3.9}$$

where *r* is the radius in nm, *A* and *B* are constants, typically 1.5 and 0.175 or 1.1 and 0.22 and *R* is the molar concentration ratio of Water/Surfactant.

It is not always clear what "*r*" means—is it the radius of the water droplet or the radius of water + surfactant; is it the Stokes or the hydrodynamic diameter? Given that a typical surfactant tail is 1nm long, for small drops this ambiguity is significant.

The second way is to say that:

$$r = \frac{3V}{S} \tag{3.10}$$

where V is the volume of water and S is the surface area of the surfactant. The factor 3 comes from the fact that $V = 4/3\pi r^3$ and $S = 4\pi r^2$.

The Microemulsions spreadsheet performs these calculations from primary inputs of weights of water and surfactant.

3.4.3.3. Droplet Concentration

The droplet size depends only on the ratio of water to surfactant concentrations, so in principle you can obtain any desired droplet concentration simply by increasing the total concentrations of water and surfactant while maintaining the required ratio. However, the HLD-NAC theory shows that there is a limit to the ability of the system to absorb enough water and surfactant. That is why there are so many disappointments for those who try to formulate practical microemulsion nanoparticles.

3.4.3.4. The Perfect "Oil"

At modest salinities, oils such as heptane are reasonable for formulating microemulsions with the surfactant AOT (which happens to be a common surfactant used in this sort of work). But for nanoparticles in a different carrier it quickly becomes clear that AOT simply will not give the required results. To make the whole process rational, it is necessary to know the EACN of the carrier, then find a surfactant blend to match it. Any manufacturer who invests in systems to make such measurements will be able to offer particles in a wider range of carriers.

3.4.3.5. The Perfect Surfactant Blend

The HLD part of the equation, combined with a list of surfactant *Cc* values makes it possible to find a blend with the right basic properties. But the user probably prefers to have the least possible surfactant, so surfactant efficiency is important. The NAC part of the theory explains that the length of the surfactant tail and the so-called xsi parameter control the overall efficiency. One particularly efficient surfactant class is

the "extended surfactants", with a mixed ethylene oxide and propylene oxide spacer that seems to like to reside in the interfacial region and therefore gives significantly higher efficiencies.

3.4.3.6. Back to Barium Sulfate

In the previous section, barium sulfate made by traditional stirring gave particle sizes in the 100nm range—only high-powered T-mixers could give sufficient mixing to generate 40 nm particles. A microemulsion based on cyclohexane as the oil and a C13EO4 surfactant was obtained using 15% surfactant and 4% (separately) of concentrated potassium sulfate and barium chloride solutions [20]. Using the same traditional reactor, particles of ~6nm were obtained irrespective of reaction conditions, provided a 1:1 ratio of reactants was used. With so much surfactant present, the particles were highly resistant to agglomeration, even after prolonged storage.

The microemulsion theory described above allows the initial droplet sizes to be tuned in the 5–50 nm range, so in principle it is straightforward to generate particles in this range with no requirement for high energies, and with little risk of agglomeration. The downside, of course, is the task of separating the particles from the oil and excess surfactant.

3.4.4. Electrochemical Generation

When a cation such as Ag+ arrives at the cathode of an electrochemical cell, it becomes Ag metal. Under most circumstances, the metal just plates out onto the cathode. Given the right circumstances, the atoms start to form clusters that can move away from the cathode, creating a dispersion of nanosilver particles.

A typical example comes from the work of Prof López-Quintela at the University of Santiago de Compostela, where the silver was supplied as a sacrificial anode, the solvent was acetonitrile and the electrolyte to provide the required conductivity was tetrabutylammonium bromide [21]. With an aluminium cathode the result was a coating of silver. With a platinum electrode no deposition was obtained and nanosilver dispersions were obtained with sizes that could be controlled by, especially, current density. Compared to other techniques it seems particularly suited for creating ultra-small particles. As discussed in Chapter 2, such silver clusters show interesting optical properties as they can be produced small enough to show plasmon resonance, and for very small

particles, essentially atomic clusters, the chemistry becomes very different from that of normal silver.

Because they are obtained in a solvent, it is possible (after removal of the electrolyte) to transfer the particles into other media. For example, they can be put into acrylate systems which, on cross-linking, become antimicrobial abrasion-resistant coatings, provided of course that all volatile solvent has been removed prior to curing.

A totally different electrochemical process is the Kirkendall effect, whereby rather ordinary nanoparticles can be etched into strange shapes—cubes, double cubes, particles with holes, nanotubes (from solid wires) and so forth. It is generally explained as being the result of a mismatch of diffusion rates at the interface of two materials [22]. Take, for example, a solid particle of material A and react it at the surface with material B, often electrochemically. In principle, the whole particle can be converted to AB by diffusion of B deeper and deeper into the core. Thanks to the mismatch, A might diffuse to meet B faster than B can diffuse to meet A, so the result is a hollow particle. Depending on geometries, the hollow might be a complex shape, and if there are ways for A to escape via pores, these can become magnified, producing a very complex shape.

That simplistic explanation glosses over many complexities. For a potential user, the important thing may not so much be an in-depth understanding of the effect than the knowledge that particles with extraordinary shapes can be created. While the Kirkendall processes are described in this electrochemical section for the sake of convenience, any other interfacial process which involves differential diffusion at boundaries can produce fascinating and extraordinary shapes for those who need them.

3.4.5. Supercritical Hydrothermal Processes

At the supercritical point of water (374°C, 22.1 MPa) many of the properties of water are finely balanced. Below that point its solubility behavior approaches that of conventional water, while above it it behaves like a non-polar liquid simultaneously having a low pH and a relatively high concentration of H^+ and OH^- ions. Therefore the solubility and reactivity behavior can be tuned exquisitely by adjusting temperature and pressure [23].

In particular, metal salts are rapidly hydrolyzed to form their hydroxides which, at these extreme conditions, are immediately transformed into their oxides by removal of a water molecule (dehydration).

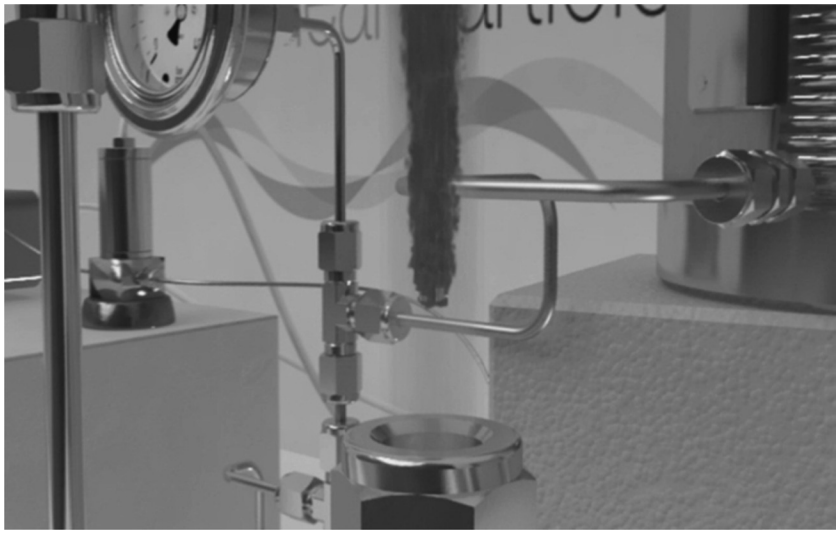


FIGURE 3.4. *Production of iron oxide nanoparticles in a supercritical hydrothermal process.*

The supercritical water hydrothermal synthesis (scWHS) technique makes it possible to carry out many interesting reactions that can create nanoparticles. Control of size comes from control of mixing and the temperatures/pressures around the critical point.

The process can be made continuous with obvious advantages in terms of throughput. Because the generated particles are already dispersed, there is no need for a subsequent dispersion step, Figure 3.4.

An example of the process is the one developed by Promethean. An image from the Promethean website shows cold water containing a metal salt (in this case, iron nitrate) colliding with a stream of supercritical water coming from above and instantly (thanks to efficient mixing from the countercurrent) producing a stream of iron oxide nanoparticles, typically in the 50 nm region [24]. Because the process is based on water the products are most naturally available as aqueous dispersions, though post-processing routes to dispersions in other solvents are clearly possible.

3.4.6. The Detonation Process

For an exciting means of nanoparticle production, there are few methods that challenge the detonation process that is used to produce nanodiamonds [25]. The method was originally devised by Zababakhin

and co-workers in the Soviet Union during the early 1960s. An oxygen-deficient mixture of TNT and RDK is detonated in a closed chamber and diamonds with a diameter of about 5 nm are produced at the front of the detonation wave. The yield from the process is strongly dependant on the operating conditions; active cooling (water, carbon dioxide, etc.) of the detonation chamber is employed, and the greater the cooling capacity the greater the yield of diamond. The process also produces soot comprising various carbonaceous materials and it is from this that the diamonds must be extracted, generally by heating with strong acids in an autoclave.

A further point to note is that nanodiamonds agglomerate very readily. To prevent this from producing unusable material, dispersion techniques in polar solvents must be applied as part of the production process. PlasmaChem in Germany have carried out much interesting work in this field and a useful review of this and associated technique can be found on their website [26].

3.5. TOP DOWN METHODS

The processes described in this section are predicated on the basis that it is very useful to start an industrial manufacturing process with a readily available raw material that requires no unusual processing conditions. ‘Hit a large particle, shatter it and create a smaller particle’ is the simple description of this method. Of course, the truth is a little more subtle than this blunt force trauma explanation. A useful review of the subject is that of De Castro and Mitchell [27].

It is also worth noting that much of top-down manufacturing is complementary with dispersion technology. Indeed, milling and grinding steps are themselves used to reduce the size of nanoparticles produced by bottom-up methods. An example of this is the NanoCentral project reported in Chapter 8.

3.5.1. Dry Milling

Processing of materials to nanoparticle size usually takes place within a mill, of which there are many designs. The fundamental principles of size reduction derive from the energy transfer during impacts with the milling media; compaction of the particles between the milling media leads to elastic and plastic deformation of the particles. Metallic particles may undergo cold welding at this stage. The final stage of

compaction involves the fracture, further deformation and fragmentation of the particles.

The issue of contamination must be confronted when milling techniques are employed. This can arise from the expected sources such as impure starting materials and from wear of the grinding media. A more subtle source of contamination arises from the atmosphere under which the milling is carried out; this is not simply a problem of "dirty air" but also includes the potential problem of oxidation. This is an issue that must be taken seriously if reactive materials, such as pure metals, are being processed.

As might be expected, simple mathematical descriptions of the energy requirements of the size reduction produced during the milling process are difficult to achieve. It is probably useful to present some of the factors that are in action during the milling process to provide the reader with a better understanding of the complexity of the situation.

A key question is: Why does it take so much energy to break something apart? The basic theory of Griffiths, modified by Irwin, shows what the problems are [28,29].

For brittle materials, mechanical fracture is related to stress cracking. This insight into the mechanics of particle fracture was presented by A.A. Griffiths in 1921. The diagram below shows the idealized situation, where a is the length of the flaw and σ is the applied fracture stress, Figure 3.5.

Griffiths showed that the stress σ_f at fracture was proportional to the reciprocal of the square root of the crack length such that their product is a constant C:

$$\sigma_f \sqrt{a} = C \tag{3.11}$$

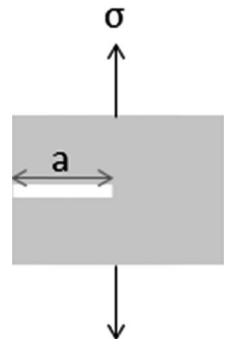


FIGURE 3.5. *Griffiths cracking definitions.*

Crack growth depends on the creation of two new surfaces; this enables the constant to be defined in terms of the surface energy (γ) and the Young's modulus (E) of the material:

$$C = \sqrt{\frac{2E\gamma}{\pi}} \tag{3.12}$$

From this relationship the fracture stress can be calculated:

$$\sigma_f = \sqrt{\frac{2E\gamma}{\pi a}} \tag{3.13}$$

While providing a good degree of agreement for brittle materials, the Griffiths analysis was less successful for ductile materials. Studies by Irwin and co-workers at the USNRL revealed that the problem arose because plasticity at the tip of the crack had not been considered. Additional energy is required to overcome this plasticity for crack growth to continue; this led to the observation that the energy required for crack propagation to occur comprised two components:

- The release of stored elastic strain energy as the crack grows; this is the thermodynamic driving force of fracture.
- The dissipation energy which provides the thermodynamic resistance to fracture; this included the plastic dissipation energy, surface energy and any other sources of dissipative energy for the material.

With these additions taken into account, the Griffiths equation is modified to become the following relationship, where γ is the surface energy and G_p is the plastic dissipation per unit area of crack growth and $G = 2\gamma + G_p$:

$$\sigma_f = \sqrt{\frac{EG}{\pi a}} \tag{3.14}$$

Although the initial Irwin approach is a useful analysis of the problem, it assumes that the plastic deformation zone at the crack tip is small in relation to the crack size; this is not necessarily the case. Irwin attempted to overcome this problem by instituting the crack extension resistance curve (R curve) approach. Unfortunately, the R curve depends on the geometry of the sample and is not easy to calculate.

Other approaches to the problem exist, but also suffer from the similar difficulties [30].

It is also an open question as to whether the Griffiths/Irwin approach in analyzing the problem of particle milling is useful, although this is claimed to be the case by De Castro & Mitchell [26]. The fact that any particle without a crack ($a = 0$) requires infinite energy to break it is a rather embarrassing feature of the theory, though those who know the difficulty of grinding particles might think that this explains why it is so difficult. The theory makes more sense in the case of dis-agglomeration of clusters of primary nanoparticles.

It should be noted that as the size of the particles decreases, the tendency to agglomerate will increase and a limiting size will be attained. The major factors contributing to the grind limit have been defined by Harris [31] as follows:

- Increasing fracture resistance
- Increasing cohesion as particle size is reduced
- Excessive clearance between impacting surfaces
- Surface roughness of the grinding medium
- Bridging of large particles to protect smaller particles during grinding
- Increasing viscosity as particle size decreases.
- Decreasing internal friction of the slurry as particle size decreases.

An alternative approach to the problem of predicting particle size reduction has been proposed by Hukki and others; this is an empirical approach that builds upon much earlier work [32].These earlier attempts are worth mentioning as they provide a useful initial analysis of the problem.

In 1867, Rittinger proposed that the energy (E) that is required to reduce a particle in size from a value of x to a value y is directly proportional to the area of new surface created:

$$E = C_R\left(\frac{1}{y} - \frac{1}{x}\right) \tag{3.15}$$

The approach ignores the energy absorbed by elastic deformation, which may be greater than that required for the creation of a new surface. To remedy this shortcoming, Kick and Kirpiĉev proposed an alternative expression:

$$E = C_K \ln\left(\frac{x}{y}\right) \tag{3.16}$$

Unfortunately, it was found that this expression suffers from scaling problems and it was left to Bond to propose a further formula:

$$E = C_B(y^{-0.5} - x^{-0.5}) \tag{3.17}$$

A problem with all of these approaches is that no account is taken of particle size distribution, or of interparticle attraction and plastic deformation. Nevertheless, the expressions are useful in predicting behavior in certain size ranges. Interestingly, the approach of Rittinger appears to best serve the situation where the particle size is < 1 μm. The Bond equation appears to work best for intermediate sized particles, while the Kick-Kirpiĉev relationship works best for coarse particles.

Hukki and others attempted to find a general equation for which the three earlier models would emerge as special cases. The following differential equation was proposed:

$$\frac{dE}{dx} = -Cx^{-n} \tag{3.18}$$

For the Rittinger case $N = 2$ and $C = C_R$
For the Bond case $N = 1.5$ and $C = C_B$
For the Kick-Kirpiĉev case $N = 1$ and $C = C_K$

However, from experimental data Hukki came to the conclusion that n is not a constant but a variable; the form of the equation becomes:

$$\frac{dE}{dx} = -Cx^{-f(x)} \tag{3.19}$$

There have been further attempts to refine the Hukki approach by treating the generation of new ground surface as a problem of fractal geometry leading to expressions such as that proposed by Thomas & Filippov, where C is of the order of 100–200 and $f(x)$ is of the order of $1/3[11 + (\ln x)/2]$. Of course, these relations are empirical and to date have not been determined from first principles [33].

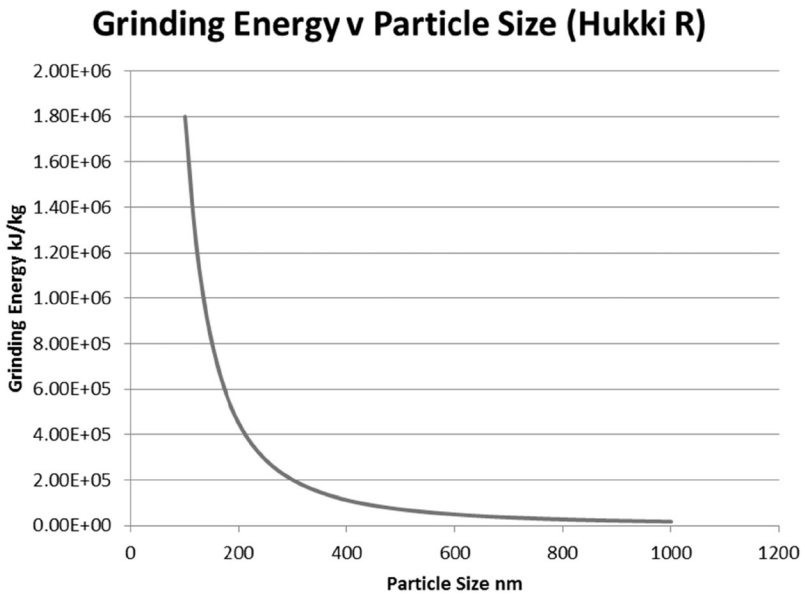


FIGURE 3.6. *Grinding energy v particle size via the Hukki (Rittinger) method.*

The Milling spreadsheet allows you to specify any two particle sizes and calculate the energy required go from one to the other. It also performs a simplified numerical implementation of the Hukki (Rittinger) equation, with $N = 2$ to show how most of the energy is required for the last few nm reduction in size. The spreadsheet plots in both a linear and log scale for the energy; the linear scale shows how dramatic the effect is, Figure 3.6.

The take-home message is that milling is unlikely to get particles down much below 100 nm unless the particles are full of cracks down to that level or are agglomerates of small primary particles, in which case the Griffith laws can provide considerable assistance to the process, especially for particles where plastic deformations are small.

3.5.2. Milling in Liquid Media

It is certainly the case within the surface coatings industry that most milling and grinding procedures are carried out in liquid media. There is much debate as to how different milling techniques compare in their ability to produce nanoparticles within a fluid medium. Although this is a vast subject, the work of two groups seems to provide a lot of practical insight.

As mentioned earlier, a study carried out by Schönstedt and coworkers used additions of Aerosil to a sol-gel type medium to try to produce abrasion-resistant coatings [15]. The agglomerated particles of silica need to be reduced in size, so as part of the study the dispersing efficiency of various types of milling equipment was investigated. Two types of equipment were used; a Stirred Media Mill (Bühler PML2) and a High Pressure Homogenizer (APV Gaulin Lab 60-15TBS). A variety of different Aerosil grades were used for the trials and a number of different stabilizing agents were also used. The dispersing (stabilizing) agents are 3-glycidyloxypropyltrimethoxysilane (Glymo), 3-glycidyloxypropyltriethoxysilane (Glyeo) & hexamethyldisilazane (HMDS). A summary of the dispersion results is presented in Table 3.4.

Aerosil materials, as purchased, may have particles in the size range of microns; but even accepting the fact that the methods described have not been optimized, these values are rather disappointing in terms of getting down to "real" nano sizes. It is also indicated in the paper that the particle size distributions are rather broad, although only one actual set of values is illustrated.

Further insights into the problem come from a wide-ranging study of the efficiency of different dispersing devices for producing silica and alumina nanoparticle dispersions carried out by Schilde and co-workers [34, 35]. The equipment investigated during the study is listed in Table 3.5. The materials were Aeroxide Alu C & Aerosil 200 V from Evonik,

TABLE 3.4. Equipment, Stabilizer and Final Particle Size in the Schönstedt Grinding Experiments.

Aerosil Grade	Machine	Wt. Solids (%)	Agglomerate Size d_{90} (nm)	Modifying Agent
OX50	Homogenizer	15	331	Glymo
200	Homogenizer	15	318	Glymo
300	Homogenizer	15	296	Glymo
300	Homogenizer	15	305	Glyeo
R812S	Homogenizer	5	241	HMDS
R812S	Homogenizer	10	230	HMDS
R812S	Homogenizer	15	279	HMDS
R812S	Homogenizer	18	226	HMDS
OX50	Media Mill	15	2330	Glymo
OX50	Media Mill	15	4300	No additive
300	Media Mill	15	411	Glymo
R812S	Media Mill	15	364	HMDS

TABLE 3.5. Equipment Used in the Schilde Grinding Experiments.

Equipment
Dissolver: (Dispermat CA60: VMA-Getzmann GmbH)
3-Roll mill: (Exakt 80 SE: Exact)
Kneader: (HKD-T0.6: IKA)
Stirred media mill: (Labstar, Netzsch Feinfabrik GmbH) & (PML-V/H: Drais)
Disc mill: (Romaco FrymaKomura)
Ultasonic Homogenizer: (UP200S: Dr. Hielscher)

each with starting sizes in the 200 nm range (and this is important to note) with primary particle sizes of 12–13 nm, which puts the final best results of > 100 nm into some sort of context.

The studies show that there is no "best" process. For example, the 3-roll mill is wonderful in terms of energy efficiency and in terms of quickly breaking down the large particle clumps. The stirred media mill breaks particles down to a significantly smaller size than the 3-roll mill (which quickly runs out of ability to decrease size further) but requires much more energy to do so and is especially inefficient at breaking down very large clumps (where the 3-roll mill excels). The ultrasonic Homogenizer has very high *specific* energy efficiency in terms of how well the energy from the tip is able to break down particles whilst the overall efficiency is very low because of the losses of going from electrical power to ultrasonic power. This complex web of effects can be disentangled with a few key concepts.

For the 3-roll mill, the disperser, the kneader and the ultrasonic Homogenizer, there is only one force available to rip the particles apart. This is viscous shear stress—where the top of the particle experiences a different fluid velocity from the bottom of the particle. The absolute difference is proportional to the particle size and the induced stress intensity is inversely proportional to the size, so the net effect is that for all these techniques the stress intensity is independent of particle diameter. For the mill and the kneader, the intensity depends only on shear rate γ and viscosity μ:

$$\text{Stress Intensity} = 2.5\gamma\mu \tag{3.20}$$

and for the disperser and Homogenizer it depends on density ρ, kinematic viscosity v and specific power input ε:

$$\text{Stress Intensity} \propto \rho(\varepsilon v)^{0.5} \tag{3.21}$$

For any given set of processing parameters, grinding will take place until the stress intensity is less than the minimum required to break apart the particle. After that, extra time and energy achieves nothing—as seen in the graphs below.

For the stirred media mill, the shear stress is imparted by collisions with the grinding media. For a mill with media diameter d and density ρ, tip speed v and particle size x the situation is very different:

$$\text{Stress Intensity} \propto \frac{d^3 v^2 \rho}{x^3} \tag{3.22}$$

This means that the mill is very inefficient with large particles, so large particles (for which other mills are good) can often be untouched. As the particles get smaller, the stress intensity increases, so given enough time and energy the technique can result in a finer overall grind.

Another key concept for understanding the process is N, the number of stress events; i.e., the number of times a particle experiences the stress intensity. A simple approximation is:

$$N = \frac{\gamma t}{4\pi} \tag{3.23}$$

where N is proportional to shear rate γ and time. This value can be adjusted for a specific situation. For example, for a 3-roll mill N is reduced by a factor that describes the volume of liquid actually in the roll gap divided by the total volume of the mill.

Experiments with a 3-roll mill show a rapid reduction in particle size for small values of N—each encounter with the nip in the roll is largely successful in reducing the size. For further reduction in size more and more work is required, along with much higher shear stresses. This is consistent with the notion that smaller particles have a higher intrinsic strength, consistent with their higher surface area to volume ratio.

There are many more complexities. Although the 3-roll mill can be very efficient, this only happens if there are high concentrations of particles. And although the silica and alumina particles are similar in many ways, the tendency of silica to form highly viscous networks means that during grinding more energy is required to overcome the viscous drag.

The whole can be summarized by Schilde's grinding equation, which seems to do a more efficient job of fitting data than the alternatives:

$$x(t) = x_0 + \frac{(x_{end} - x_0)t}{t + K_t} \tag{3.24}$$

This says that the particle size x at time t depends on the initial size x_0, the smallest size limited by the physics of the process x_{end} and on K_t, which is a function of N, volumes, velocities and concentrations. The Grinding spreadsheet shows the equation fitting real data for four different grinding techniques, Figure 3.7.

Although these are only fitted equations (for the kneader, x_0 is, mysteriously, much less than the original particle size of 200 nm) they show the general trend: that the stirred media (because of its dependence on x^{-3}) is slow at first but keeps on going, while the other techniques are relatively rapid at the start but achieve nothing from the investment of extra time. With further work (not included in the spreadsheet), Schilde was able to find equations for x_{end} and K_t based on known input parameters, making the equation a predictive tool.

What emerges from this excellent work is a sense of disappointment. Recall that the primary particle size for the alumina and silica is 12–13 nm; so the reductions from ~200 nm to ~100 nm are not all that impressive. The particle size distributions at the end of these processes are, according to the papers, not very sharp, so their cumulative *mass* distributions will be heavily biased towards high-end particles. Of course these experiments are deliberately simple—there is no attempt to add

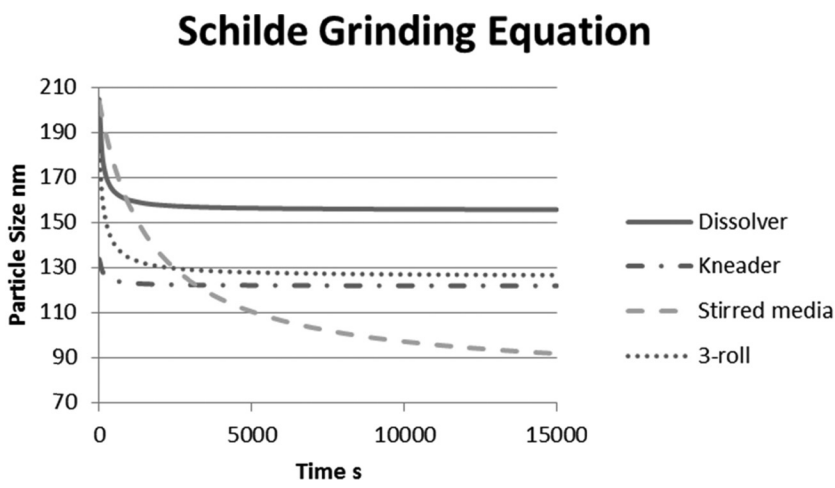


FIGURE 3.7. *The Schilde grinding equation for different methodologies.*

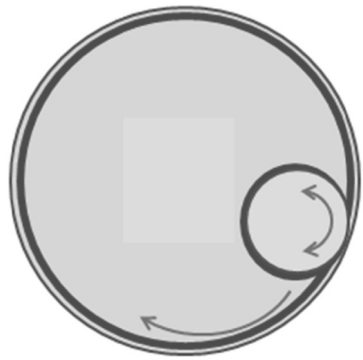


FIGURE 3.8. The ConCor™ grinding process.

dispersants which might fight against the tendency of smaller particles to be harder to break up. The ZnO case study in Chapter 8 reinforces this point; even in the hands of a team with expert knowledge of dispersers and dispersants, particle sizes could not be brought down to the 20 nm size typical of ZnO prepared via a liquid process. So, although it is possible to make low cost particles with inherent sizes of 12 nm, realizing those sizes in a practical formulation via milling/grinding may not be possible, or might require very energy-intensive post-processing. A higher-cost process which creates the right particle in the right solvent (or UV curable material) may result in overall lower cost.

An interesting dispersion and grinding idea has recently been presented by Primary Dispersions. The technique is based on a shearing force that is generated by two rotating surfaces; the direction of the rotation can be varied to produce co-rotational or con-rotatory motion. Depending on the relative rotational direction, direct or indirect stressing of the particles can be achieved. Direct stressing of the particles is achieved with a co-rotary motion, the particles impacting on each other on their passage through the gap. Indirect stressing is achieved by stressing the fluid by contra-rotation; this produces a shearing force on the particles. The process is known as the ConCor™ process and the technology has been reviewed by Cash for OCCA [36], Figure 3.8.

3.6. CONCLUSION

There is no perfect way to make nanoparticles. The point of this chapter is that the team now has some idea of the trade-offs in price/performance/capabilities between many different techniques. The au-

thors' bias has always been towards systems that deliver nanoparticles in dispersion without the need to grind/disperse. As the Top Down section shows, milling is not a good way to get small particles, even if the low cost of the original clumped particles is attractive. Your team will have different needs, and so may reach a very different conclusion.

3.7. REFERENCES

1. http://nanoparticles.org/ Accessed August 2012.
2. Wei-Hong Zhong, Bin Li, *et al.* (2012) *Nanoscience and Nanomaterials: Synthesis, Manufacturing and Industry Impacts.* Lancaster, PA: DEStech Publications.
3. Kwang-min Lee, Sung-tae Park, *et al.* (2005) Nanogold synthesis by inert gas condensation for immuno-chemistry probes. *Journal of Alloys & Compounds* 390: 297–300.
4. M. Raffi, A.K. Rumaiz, *et al.* (2007) Studies of the growth parameters for silver nanoparticles synthesis by inert gas condensation. *Journal of Materials Research* 22: 3378–3384.
5. C.J.Choi , X.L. Dong, *et al.* (2001) Characterization of Fe and Co nanoparticles synthesized by chemical vapor deposition. *Scripta Materialia* 44: 2225–2229.
6. Kwang Leong Choy. (2003) Chemical vapor deposition of coatings. *Progress in Material Science* 48: 57–170.
7. http://www.aerosil.com/lpa-contentdb/page/literature/show?lang=en. Accessed August 2012.
8. Bénédict Thiébaute. (2011) Final analysis: Flame spray pyrolysis: A unique facility for the production of nanopowders. *Platinum Metals Review* 55: 149.
9. B. Kumfer, M. Shinoda, *et al.* Gas-phase flame synthesis and properties of magnetic iron oxide nanoparticles with reduced oxidation state. *Journal of Aerosol Science* 41: 257–265.
10. B.N. Chichkov, C. Momma, S. Nolte, F. von Alvensleben and A. Tünnermann. (1996) Femtosecond, picosecond and nanosecond laser ablation of solids. *Applied Physics A: Materials Science & Processing* 63:109–115.
11. Fumitaka Mafuné, Jun-ya Kohno, *et al.* (2000) Formation and size control of silver nanoparticles by laser ablation in aqueous solution. *The Journal of Physical Chemistry* B 104: 9111–9117.
12. Aneta Petrova, Werner Hintz and Jürgen Tomas. (2008) Investigation of the precipitation of barium sulfate nanoparticles. *Chemical Engineering & Technology* 31: 604–608.
13. Ying Ying, Guangwen Chen, *et al.* (2008) A high throughput methodology for continuous preparation of monodispersed nanocrystals in microfluidic reactors. *Chemical Engineering Journal* 135: 209–215.
14. http://www.solgel.com/articles/June00/phalip/introsolgel.htm. Accessed August 2012.
15. http://www.ormocer.de/EN/. Accessed August 2012.
16. B. Schoenstedt, A. Kwade, *et al.* (2011) Influence of pyrogenic particles on the micromechanical behaviour of thin Sol-gel Layers. *Langmuir* 27: 8396–8403.
17. *Nanoresins.* (2011) Patent application number 2011040031.
18. Advanced Nanotechnology Ltd. (2006), PCT/AU2006/000454.
19. www.stevenabbott.co.uk/HLD-NAC.html. Accessed August 2012.
20. Dendy Adityawarman. (2007) *Precipitation of barium sulfate nanoparticles in microemulsion: Experiments and modelling.* PhD Thesis, Otto-von-Guericke-Universität Magdeburg.
21. L. Rodríguez-Sánchez, M.C. Blanco, and M.A. López-Quintela. (2000) Electrochemical Synthesis of Silver Nanoparticles. *Journal of Physical Chemistry B* 104: 9683–9688.
22. Hong Jin Fan, Ulrich Gçsele, and Margit Zacharias. (2007) Formation of nanotubes and hollow nanoparticles based on Kirkendall and diffusion processes: A review. *Small* 3: 1660–1671.

23. Edward Lester, Paul Blood, *et al.* (2006) Reaction engineering: the supercritical water hydrothermal synthesis of nano-particles. *The Journal of Supercritical Fluids* 37: 209–214.
24. http://www.prometheanparticles.co.uk/index-2.html. Accessed August 2012.
25. K. Iakoubovskiia and M.V. Baidakova. (2000) Structure and defects of detonation synthesis nanodiamond. *Diamond and Related Materials* 9: 861–865.
26. http://www.plasmachem.com/overview-powders.html#diamond. Accessed August 2012.
27. C.L. De Castro and B. Mitchell. (2002) Nanoparticles from mechanical attrition. *In: Synthesis, Functionalization and Surface Treatment of Nanoparticles.* Marie-Isabelle Baraton (Editor). Valencia, CA: American Scientific Publications.
28. A.A. Griffith. (1921)The phenomena of rupture and flow in solids. *Philosophical Transactions of the Royal Society of London, A* 221:163–198.
29. G. Irwin. (1957) Analysis of stresses and strains near the end of a crack traversing a plate. *Journal of Applied Mechanics* 24: 361–364.
30. J.R. Rice. (1968) A path independent integral and the approximate analysis of strain concentration by notches and cracks. *Journal of Applied Mechanics* 35: 379–386.
31. C.C. Harris. (1967) Minerals beneficiation—size reduction—time relationships of batch grinding. *Transactions of the AIME.*
32. R.T. Hukki. (1961) Proposal for a Solomonic settlement between the theories of von Rittinger, Kick and Bond. *Transactions of AIME* 220: 403–408.
33. Thomas and L.O. Fillipov. (1999) Fractures, fractals and breakage energy of mineral particles. *International Journal of Mineral Processing* 57: 285–301.
34. Schilde, Ingo Kampen, *et al.* (2010) Dispersion kinetics of nano-sized particles for different dispersing machines. *Chemical Engineering Science* 65: 3518–3527.
35. Carsten Schilde, Arno Kwade, *et al.* (2011) Efficiency of different dispersing devices for dispersing nanosized silica and alumina. *Powder Technology* 207: 353–361.
36. http://primarydispersions.com/?page_id=490. Accessed August 2012.

Creating Stable Nanoformulations

THE stability of particles during production and during the lifetime of the product are both important. The first aspect means that the nanoparticles don't clump together in a hopeless mess. The second aspect is more concerned with the integrity of the nanoparticle-to-matrix interactions over time, with strong interactions generally being desirable. This latter aspect also has ramifications for the concerns about nanosafety.

There are many fine, theoretical approaches to thinking about these stabilities. For those who have time and resources, approaches such as DLVO can provide useful insights into what is going on. For most of us these approaches produce little or no practical insight, especially when comparing (for example) different formulations of the "same" product, from one or more suppliers. Even if the suppliers were willing to reveal all their technical information about their stabilizing shell (and many of them, rightly, want to keep this as trade secret), they themselves may know rather little about why some of their formulations are more stable than others, and if they do they might not be able to explain it in terms of DLVO.

This chapter, therefore, is going to focus on relatively simple, pragmatic methods for thinking through stability issues. Classically, the discussion would center on the three DLVO factors:

- stability via charge repulsion and then to a discussion of zeta potential
- steric stabilization via polymer chains
- dispersion stability via Brownian motion

There is a fourth factor which is less frequently discussed yet is of great practical importance—both directly and in terms of its effects on the steric stabilization:

- solubility effects

This chapter will cover the DLVO factors as briskly as possible, spending much more time on the fourth factor, solubility, which is of far more practical importance to the nanocoating formulator. There is, of necessity, some overlap with Chapter 5, where the choice of solvents and solvent blends is covered in more detail.

4.1. DLVO (DERJAGUIN, LANDAU, VERWEY AND OVERBEEK)

Despite the fact that most of us find that we cannot apply DLVO to the real-world formulations we deal with, it is obligatory to include some discussion of the subject. The theory is not particularly difficult—the problem is that assembling all the correct formulae with the correct parameters and, most difficult of all, the correct units is too hard for most of us. The authors, therefore, are most grateful that Dr. Robert Lee of Particle Sciences provided a working model (in Wolfram Mathematica) which we were able to translate into Excel and include in the Models spreadsheet. As an example of the problems, even with that help it took time to track an error of 6 orders of magnitude in the Steric term that arose from the fact that most definitions of molar volume are in cc/mole and the formula required it in the unfamiliar units of m^3/mole. Because the spreadsheet has all the formulae and provides all the inputs and outputs in "normal" units, there is no need to take up time in this chapter explaining things in detail.

The theory wants to describe the "potential" of a system of particles depending on the distance between them. A high positive potential means a strong barrier against association, while a large negative potential means a strong attraction. For DLVO discussions, entropic effects are put to one side, so in terms of stability the assumption is made that potential is equivalent to free energy. The total potential V_T is composed of three terms: Hamaker (H) representing van der Waals interactions; Debye (D) representing charge interactions and Steric (S) representing interactions via polymer chains:

$$V_T = V_H + V_D + V_S \tag{4.1}$$

In each case, the terms are expressed in terms of h, the distance between the particles. The values are expressed in units of kT (Boltzmann energy). There is general agreement that if the barrier is greater than, say, 20 kT then the particles will not have a chance to overcome that barrier. As the barrier decreases to just a few kT, then thermal motion can kick two particles over the barrier and once they are stuck they are stuck! Note the important point discussed in the section on inter-particle distance in Chapter 2: high volume fractions of small particles can impose an inter-particle distance less than the maximum height of the barrier, so "stable" dispersions can become unstable at higher volume loadings.

The most basic of the terms is V_H, the dispersion term. This arises from van der Waals (VdW—it seems more usual to capitalize the V in the acronym) interactions—mutually induced interactions between every atom of every molecule. VdW interactions are weak, but because they apply to every atom in a system, their sum can add to a large number.

These are so fundamental that a specific VdW worksheet is included in the spreadsheet as a reminder of the balance of attraction (that goes as 1/distance^6) and repulsion (1/distance^{12}) that is typically used to understand molecular VdW interactions. The graph is a typical example of what would happen if two atoms (such as argon) approached each other, Figure 4.1.

The classic Lennard-Jones 6-12 VdW forces, with the repulsive

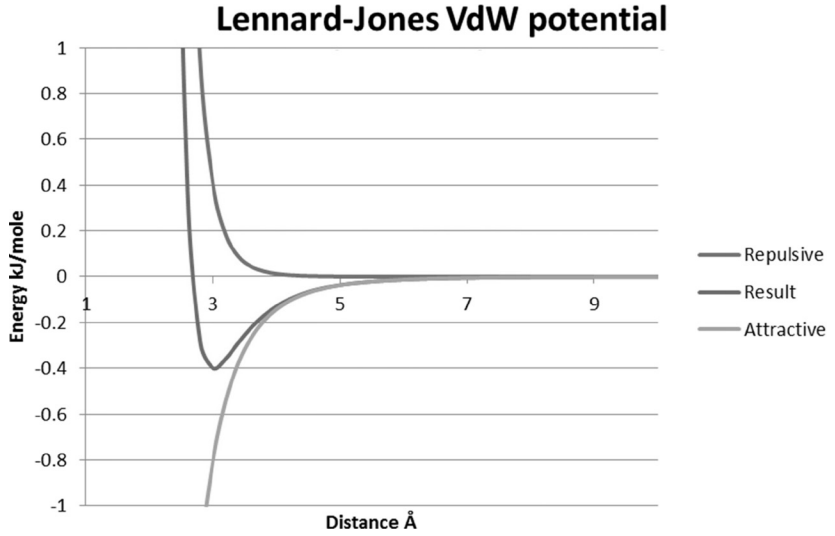


FIGURE 4.1. *The classic Lennard-Jones van der Waals potential for atom-to-atom interactions.*

force eventually winning out, results in a stable distance of 3Å. This view of VdW gives the impression that the net result is always a safe, stable distance between "particles" which can, therefore, never clump closer together.

Although the same VdW forces are in play for nanoparticles, they do not behave like the atoms, because at close distances the net effect is an attractive force. This means that nanoparticles are always fighting the VdW tendency to clump together. For nanoparticles of radius r separated by distance h the V_H term becomes:

$$V_H = -\frac{A_{12} r}{12hkT} \tag{4.2}$$

where A_{12} is the Hamaker constant and kT is the Boltzmann temperature. With such a simple formula what can go wrong? The problem is A_{12}. This represents the average interaction of a particle with its surroundings. There are plenty of rough guides to what it might be in various circumstances (and some examples are included in the spreadsheet) but without considerable work (usually not worth the formulator's trouble) it is not going to be known for a given real-world formulation. A note of caution—when looking up Hamaker constants, check whether they are quoted in terms of 10^{-20}J or 10^{-21}J; the values in the spreadsheet are in terms of 10^{-20}J. Some papers describe them in terms of zJ which turns out to be zeptoJoules, or 10^{-21}J.

Because V_H is always attractive (negative), particles (as opposed to atoms or molecules) will always wish to clump together unless there is another barrier. If that barrier is greater than ~20 kT it is generally assumed that there is not enough Brownian motion to force the particles over that barrier into the VdW trap.

The next term is V_D, the Coulombic effect based on the Debye parameters. It is this term in standard DLVO discussions which is likely to provide the 20 kT barrier:

$$V_D = \frac{2\pi e_0 \varepsilon r \varphi^2 \ln(1 + e^{-(h/k^{-1})})}{k_B T} \tag{4.3}$$

k^{-1} is the Debye length given by:

$$k^{-1} = \sqrt{\frac{e_0 \varepsilon k_B T}{2N_A\, e^2 I}} \tag{4.4}$$

N_A is Avogadro's number, e_0 is permittivity of free space, e is the electron charge, ε is the dielectric constant, and I is the ionic strength. The electric potential φ can be approximated by the zeta potential ζ (discussed in detail in the next section).

What V_D is telling us is how the charges across space (φ) interact through the medium (the e, e_0 and ε terms) while being shielded by the other charges from surrounding ions. The Debye length k^{-1} represents the distance across which the pure charges can reach before disappearing in the fog of other ions.

The third term, V_S, is the steric effect. For the moment, the formula is presented with a minimal explanation of each of the terms, with a fuller discussion later. When V_s is positive (repulsive), it usually overwhelms V_H so there is no worry about a 20 kT barrier—steric stabilization tends to be all or nothing:

$$V_S = \frac{30N_A 4\pi r\Gamma^2(0.5-\chi)\left(1-\frac{h}{2\delta}\right)}{\rho^2 \mathrm{MVol}} \tag{4.5}$$

Here, Γ is the absorbed weight, δ is the layer thickness (the effective length of the polymer chain), ρ is the density and MVol is the molar volume (MWt/Density) of the solvent. The formula is saying that there is more steric interaction the more absorbed material there is Γ, the larger the radius (because there's a bigger chance of interaction if the particle is less curved) and the larger the polymer chain δ. These numbers are, in principle, calculable.

The problem is the Flory-Huggins parameter χ (chi). Mathematically, when it is less than 0.5 the steric repulsion provides a firm barrier against association (V_S is positive). When it is more than 0.5 the particles (mathematically) are positively attracted to each other (V_S is negative), and at precisely 0.5 the polymer acts as though it is not there ($V_S = 0$). If you know χ then you know your steric repulsion.

But what is χ? It is a basic parameter in polymer physics which describes the mutual interaction of a polymer and the surrounding solvent. When $\chi < 0.5$ the polymer chain prefers to extend itself away from the particle into the solvent, providing the ideal circumstances for the entropic stabilization effect (this is the most common explanation for steric stabilization) to take place. When $\chi > 0.5$ the polymer chain prefers to curl up on itself rather than extend into the solvent. So if another polymer chain arrives from another particle, the chains are happier to-

gether and can exclude the solvent that was in between. When $\chi = 0.5$ (the so-called theta solvent state) the polymer and the solvent are exactly balanced—with no preference either way, so no steric effects arise.

Though most of the time, most of us have no idea what χ actually is, we shall see later that χ can be estimated readily for a given polymer/solvent combination. This means that the powerful steric forces can be harnessed to prevent or, indeed, to promote, particle-particle interaction.

In terms of DLVO, the key thing to understand is the science behind the mathematics. The vital take home message from this discussion is that to have reliable steric stabilization, a good solvent is required to ensure that the chains are extended and therefore provide mutual repulsion.

With the DLVO spreadsheet it is a simple matter to see how these three effects work out. Each individual term is calculated and plotted over a given distance along with the total, Figure 4.2.

In this particular example, V_S is dominant, a situation typical of solvent-based systems where V_D effects are largely irrelevant. It is easy to construct the opposite system, where stability (in water) depends on large potentials on the particles. Given that formulators have essentially no control over V_H, DLVO theory boils down to: *when in water, use charges; when in solvents, have a large polymer and a good solvent.*

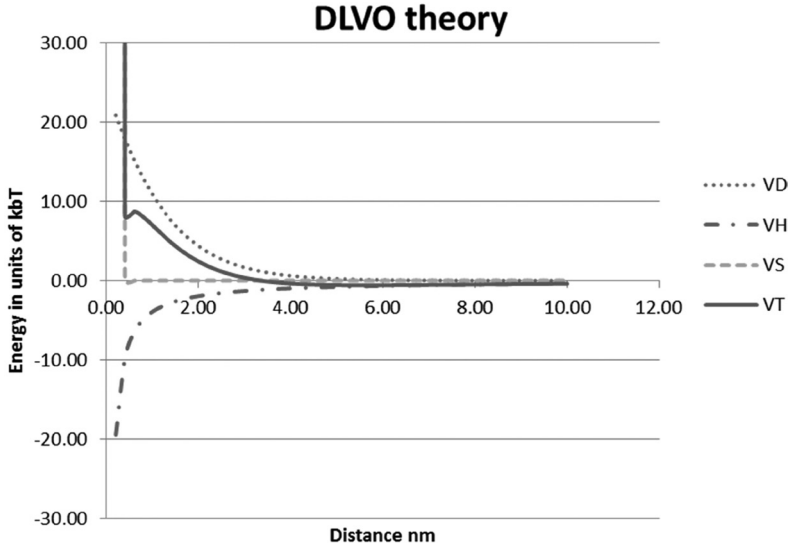


FIGURE 4.2. *The three DLVO terms plus their total.*

The model itself, therefore, adds very little to what a smart formulator would have already done, save, perhaps for getting to understand χ as will be discussed below.

4.1.1. Zeta Potentials

Knowing the effective overall charge of the nanoparticles in the chosen environment is critical for understanding formulation effects on the VD term. If the particles are effectively neutral then addition of charged additives will have relatively minor effects. If the particles are strongly charged (and, by implication, if stability comes from electrostatic repulsion), then accidental addition of an inappropriate ion of the opposite charge can be catastrophic.

Knowing the exact charge at the surface of the particle is generally seen as impractical. Instead we use a surrogate potential which is the charge at the effective surface of the particle. By "effective surface" it is meant the surface which moves through the liquid medium. Outside this surface is, of course, the neutral bulk solution. Inside this surface is the complex mix of ions and liquid called the double layer or Stern layer, which acts as if it were part of the particle as a whole. The thickness of the double layer is typically a few nm (and comes from the Debye length discussed above). The zeta potential is assumed to exist at the "slipping plane" between the Stern layer and the bulk solution, Figure 4.3.

In general, it is assumed that the charge at the effective surface is of the same sign as the "true" charge of the particle, whatever that means. This assumption is convenient, as it allows us to rationalize the charges that are typically found. Silicas normally show a negative charge (they have a negative zeta potential), and alumina has a positive charge in acidic environments and a negative charge in basic environments. Otherwise, neutral particles covered with cationic surfactants generally have a positive zeta potential and those covered with anionic surfactants have a negative zeta potential.

However, it doesn't take much to tip the balance. The potential at the effective surface depends on a complex balance and the judicious addition of a, say, trivalent salt can upset that balance and switch the sign. Impurities can also interact in that critical environment, neutralizing or reversing the potential. Thus the zeta potential has a reputation of being nice and simple when it gives the expected result, but slippery and complex when things don't work out as predicted.

One issue is that the calculations behind the zeta potential assume

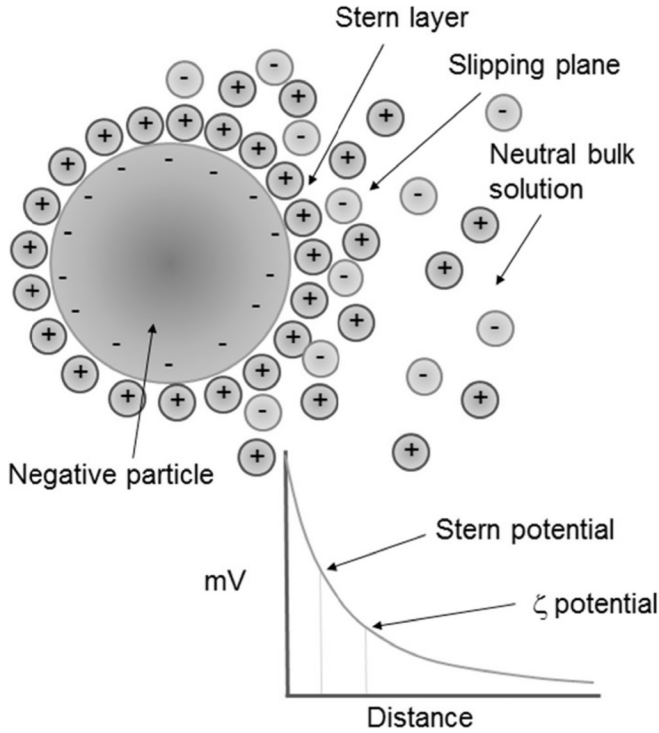


FIGURE 4.3. *The origin of the zeta potential of a charged particle in an ionic solution.*

that the Smoluchowski criteria apply (roughly speaking, that the double layer is much thinner than the particle radius, and that surface conductivity effects are neglected). This assumption is generally sound for typical solid nanoparticles with modest levels of (charged) dispersant around them. It is inappropriate for "soft" nanoparticles and for cases where the dispersant polymer is deliberately made large to provide steric stabilization.

So it makes little sense to worry too much about the subtleties of zeta potential. Instead the big trends are important. In your specific application, does the zeta potential change dramatically over exactly the pH range that is of interest to you (a danger sign) or is it relatively unaffected in that range (a good sign)? If small quantities of likely contaminants (e.g., surfactants) are added, does the zeta potential give a big change (danger sign) or stay nearly constant (a good sign)?

In other words, there is no point in determining a *single* zeta potential for your particle. Instead, the zeta potentials must be measured over a plausible sweep of likely formulation variants. In the past, measuring

zeta potential was so painful that this approach was not possible. These days the measurements are quick, simple and reliable (and come "free" with your particle sizer), so there is no excuse for not making lots of measurements to get a feel for what your particle is likely to do.

Modern analytical machines make it unnecessary to discuss the theory of measurement. Particles of a given size with a larger charge will migrate faster towards the appropriate electrode than those with a smaller charge. So from (typically) electrophoretic measurements characterized by (typically) laser-based size/velocity methods, the size and sign of the zeta potential is readily known. There is no cut-off value for the zeta potential that divides "charged" from "uncharged" particles, but particles below 20 mV and above –20 mV would generally be seen as being of small charge and those above 50 mV or below –50 mV would generally be seen as having a large charge. The usual rule of thumb is that particles of > 30 mV or < –30 mV are "charged" and will resist flocculation via charge repulsion, while those with smaller (absolute) charges will need to rely on polymeric stabilization. As hinted above, this rule of thumb is often misleading, especially for "soft" particles.

In protic environments (in particular, in water), the zeta potential of a particle changes with pH. A particle coated with amines would typically show a small potential at high pH and a large, positive potential at low pH. A carboxylic acid will show the reverse trend. Amphoterics such as silica or alumina can show strong zeta potentials at low and high pHs, and small potentials at neutral pH. The isoelectric points—the pH at which the potential changes—can be indicative of what would be expected of a particle. Typical examples are indicative only—measurements can differ wildly and many of these "oxides" are really "hydroxides" or "hydrates" and values can vary, Table 4.1.

TABLE 4.1. The Isoelectric Point of Some Typical Nanoparticles.

Particle	Isoelectric pH
SiO_2	2
Al_2O_3	5
Fe_2O_3	6
TiO_2	6
SnO	7
ZnO	9
$CaCO_3$	11
MgO	12

Not surprisingly, the reality can be quite different. As discussed above, it takes only a small percentage of impurity or stabilizer to be at the surface of the particle to change its isoelectric point dramatically. At the very least, even without zeta potential measurements, the pH of the nanoparticle formulation should be known and any changes from that pH should be treated with caution.

The effects of the counter-ions that balance the ions of interest are somewhat complex. It is possible to rationalize some of these effects via Hofmeister series. This is the idea that ions tend to interact with complex systems in an orderly fashion, such that for anions:

$$F^- \sim SO_4^{2-} > HPO_4^{2-} > CH_3CO_2^- > Cl^-$$
$$> NO_3^- > Br^- > ClO_3^- > I^- > ClO_4^- > SCN^-$$

and for cations:

$$NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+}$$

Though this may be true, such changes are usually not all that relevant to a practical formulator. To a positive zeta potential particle that has a Cl- counter-ion, adding extra, say, NaBr might do little harm (ignoring more subtle ionic-strength effects). But a large alkylsulfate counter-ion (such as lauryl sulfate from the common surfactant SLS) might prove catastrophic—an effect way beyond any Hofmeister thinking. Similarly, adding K^+ ions to a negative zeta potential particle surrounded by Na^+ counterions might have a small effect, but a long-chain tetraalkylammonium counter-ion (such as cetyl-trimethyl ammonium from the common surfactant CTAB) might prove catastrophic. Or to put in another way, small amounts of the wrong surfactant (often present for innocuous reasons) can cause chaos. The large chain of the surfactant probably prefers to be away from the water and so attaches itself, and its charge, preferentially to the surface, though as with so much to do with zeta potentials, other effects might be at play because (as stated earlier) the zeta potential represents the net result of a complex inner double layer.

The point about surfactants is worth emphasizing. They are present in many different additives and are sometimes added for other reasons (such as fixing wetting problems during coating) so unexpected things can happen for no apparent reason.

4.1.2. Steric Stabilization

If two uncharged particles come together, then a strong van der Waals attractive force sets in when they are very close and they clump together. If the particles are surrounded by a shell of polymer chains then the particles themselves cannot get close enough together to be able to clump. This leaves the question of why there is not clumping via the polymer-polymer interactions; after all, van der Waals is not specific about what atoms/molecules are in close proximity. Indeed, if the polymer is nicely wrapped around the particle then the clumping takes place in exactly the same fashion; stabilization does not occur simply because a polymer is attached. The trick is that the polymer chains need to be sticking out like bristles of a hedgehog. The argument, then, is that as the particles come together the polymer chains start to overlap and decrease each other's degree of freedom—in other words, decrease their entropy. This increases the free energy of the system so the particles prefer to stay apart (entropic stabilization). That this explanation is not totally convincing is shown by the large literature attempting to find alternatives or refinements. The debate would be interesting if it were possible to provide convincing and useful formulae that could be applied to optimize practical formulations. This does not seem to be possible, so instead the discussion needs to focus on the four things that really matter and which we can do something about:

- attachment strength to the particle;
- % coverage of the surface by the polymer;
- MWt of the polymer;
- degree of extension of the chain.

As discussed earlier, it is the fourth factor (related to the Flory-Huggins χ parameter) which will take us deep into solubility theory which, conveniently, allows further practical refinement of dispersion and compatibility.

4.1.3. Attachment and Steric Stabilization

If the polymer chains are not solidly attached to the particle, the chains fall off and stabilization is lost. Attachment can be via chemical bonds, ionic bonds or via general compatibility. For the bonding routes, standard chemical intuition is your guide. For general compatibility, a

typical trick is to use a di-block polymer, with one block that is compatible with the particle and another block that likes to stick out into the medium. This raises the question, answered below, about how to practically determine the relative compatibilities and incompatibilities.

Multiple attachment points are more secure, and the popular comb stabilizers rely on a long backbone polymer that is compatible with the particle plus many "teeth" sticking out from the backbone to provide the steric repulsion. There are two problems with this approach. The first is that getting the complex polymer to wrap itself around the particle in the required manner can take time—and if the aim is to add a dispersant during particle formation, this process might be too slow to take effect. The second problem relates to the need for a high surface coverage; if the density of teeth along the comb is not high enough then there is insufficient protection.

An important warning about "good" solvents is necessary for systems not stabilized by chemical attachment. A good solvent very much helps to disperse particles, but a problem arises if the dispersant is so happy in the solvent that it detaches from the particle. One of the tricks in formulating such systems is to get a "good enough" solvent that keeps the particles dispersed without causing catastrophic failure by removing the dispersant.

4.1.4. Surface Density and Steric Stabilization

At very low surface densities of the chains, it is not too hard for the two particles to find ways to come together and clump. The V_S term in DLVO shows a squared dependence on the surface coverage—so doubling the coverage quadruples the stabilization. Therefore it is important to add plenty of stabilizer, though presumably it is possible to have too much of a good thing, as a close-packed coverage is indistinguishable from a solid particle, so the steric protection vanishes.

4.1.5. MWt and Steric Stabilization

It is an astonishing fact that within a high temperature melt, high MWt polyethylene is not soluble in high MWt deutero-polyethylene. There is a very small enthalpic difference between the two polymers which imposes an enthalpic penalty of mixing. Normally we expect entropic effects to more than compensate—which they do for low MWt versions of the experiment. At high MWt, there is insignificant entro-

pic contribution, so they remain mutually insoluble. The point of this story is that for a process relying on entropic penalties, high MWt is a requirement. In addition, if the chain is too small, the particles might get close enough for van der Waals forces to take over. Rules of thumb (confirmed by DLVO) suggest that polymers in the 1–5 nm range are required for stabilization. Assuming that a typical polymer bond is 0.15 nm this implies (under the most favorable circumstances of a fully-extended chain, an over-optimistic assumption) something between an 8-mer and a 40-mer. For a hydrocarbon chain this means a MWt between 100 and 500, or for a polystyrene chain, 400 to 2000. Longer chains give greater stability (other things being equal) but they seriously dilute the nanoparticle. As discussed in Chapter 2, a 5 nm chain on a particle with 20 nm radius means that only ~30% of the volume of the (now 30 nm) particle is particle, and 70% is polymer.

These guidelines are necessarily vague. For a stabilizer attached by a single chemical or ionic bond, the MWt can be as described, while for a di-block maybe twice the MWt is required; so there is plenty of opportunity for the particle-compatible part of the di-block to adhere to the particle. The extra complication is that the estimate is for the best-case scenario where all of the chain is sticking out straight. As soon as the chain starts to coil in on itself, much of the stabilization is lost, so a higher MWt is required to compensate.

This last point brings us once more to the solubility considerations and χ.

4.1.6. Solubility and Steric Stabilization

A polymer placed in a poor solvent curls up on itself and remains undissolved. In a so-called "theta" solvent there is (by definition) no positive or negative net interaction between the polymer and the solvent, so it expands to form a "statistical coil" which is characterized by a standard radius of gyration. In a good solvent where there are positive interactions between polymer and solvent, the chain can extend towards its maximum length. In experimental terms, this shows as straightforward insolubility for a poor solvent, a modest viscosity for a theta solvent and a high viscosity for a good solvent, because the expanded chains can readily overlap and tangle. Polymer physicists are familiar with the Mark-Houwink equation, which states that the intrinsic viscosity (which is a theoretical viscosity at zero concentration) depends on $K.MWt^{a}$, where K is a constant for a given solvent/polymer

pair and a is 0.5 for a theta solvent (i.e., increasing MWt doesn't have a big effect) and nearly 2 for a good solvent, where doubling the MWt quadruples the viscosity. Measuring the Mark-Houwink relationship for the polymer used as a dispersant is not practical. Instead, a simple way to characterize polymer-solvent interactions is to measure the viscosity at a given % polymer, looking for high viscosities as evidence for good compatibility. The match between polymer physics and a chemist's intuition is, in this case, very good.

Because steric stabilization requires that the polymer chains stick out from the particle, it follows that this can only happen in a good solvent. The implications for the coating chemist are profound. A supplier may say that they have a "stabilized nanoparticle" and in their test systems the particle might be infinitely stable. The formulator might take that "stabilized nanoparticle" and place it in a different solvent regime and get instant flocculation.

The trick, therefore, is to know what are good or bad solvents. While this is just about possible for a standard list of solvents (acetone, MEK, ethanol, toluene . . .), in practice solvent blends tend to be used. So how might we anticipate whether a particle that is nicely stable in MEK would also be stable or unstable in MEK:ethanol or MEK:toluene blends? The most practical answer known to the authors is to use Hansen Solubility Parameters, which is why they shortly follow in some detail.

Along the way we even find a way to calculate the Flory-Huggins χ parameter to allow us to calculate the V_S term in the DLVO equation—or simply use our common sense and choose a system with a low χ parameter.

4.1.7. Beyond DLVO

The predictions of steric stabilization are clear: only long polymer chains can provide stability. Yet everyone knows that short-chain dispersants can work remarkably well. This goes completely against DLVO theory. So what theoretical frameworks can be used instead?

In fact there are probably more exceptions to DLVO than systems that follow it. Typical examples are "hydration layers" which can keep some particles apart. A good example is where Li^+ salts provide greater stability than K^+ salts because a strong layer of water is held around the lithium ions, stopping further approach of the particles.

Another important exception concerns steric effects when both low

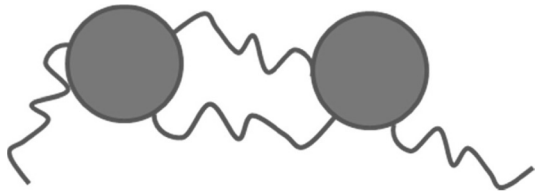


FIGURE 4.4. *Low levels of polymer can cause bridging flocculation.*

levels and high levels of polymer "stabilizer" can make matters worse. At low levels the polymer can find itself partly attached to one particle and partly to another, rather than forming a protective barrier around a particle—causing *bridging* flocculation, Figure 4.4.

At high levels, the polymer can behave effectively as a spherical particle that does not like being trapped between the particles and moves out of their way. This (using naïve language to explain a complex effect) creates a low concentration of polymer; therefore, the solvent seeks to move to areas of higher polymer concentration (i.e., osmotic pressure) and the particles come together to fill the gap created by the movement of the solvent. This is *depletion* flocculation. Because the root cause is the dislike of being trapped, smaller polymer chains are less likely to cause flocculation than large ones, as their radii of gyration are smaller. At very high levels of polymer the effect disappears, as the polymer molecules can combine to push the particles apart once more, Figure 4.5.

Surfactants should not, according to DLVO theory, provide any significant stabilization—they can destroy the zeta potential effects, yet they are not large enough to provide significant steric stabilization. In fact, in water surfactants can act catastrophically—a single layer

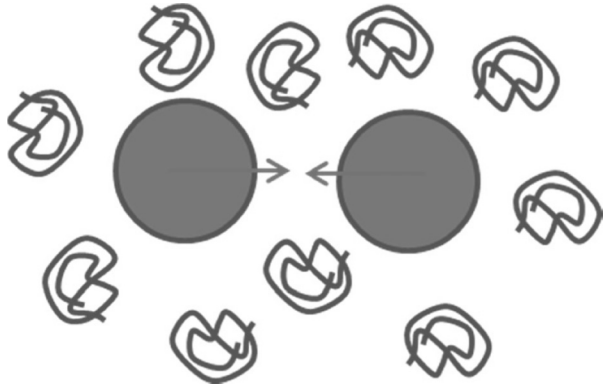


FIGURE 4.5. *The polymer chains are forced out of the gap between particles, causing a net osmotic force bringing the particles together causing depletion flocculation.*

around the particle creates a hydrophobic environment which, again simply stated, prefers to be out of the way of the water. This results in the particles crashing out just as oil separates from water. At higher levels of surfactant a double-layer can be formed (with the charged end pointing towards the water,) thereby providing charge stabilization once more. Note that if the particle is positively charged, the surfactant will be negatively charged, so the double layer will show the opposite zeta potential to the original particle.

Yet we know that many nanoparticles are perfectly happy in solvents even with small-molecule dispersants/stabilizers. None of the above effects seem to be relevant to the sorts of stabilities that are so common. What other theories can be applied?

In Chapter 2, the effect of Brownian motion on the ability of particles to stay suspended was discussed, along with some simple calculations that compared Brownian energies to gravitational forces. For typical particles in typical solvents, for $r \leq 20$ nm the particles should be relatively stable against falling out of the solvent. This is still not enough. What other approach can be adopted?

The answer seems to be solubility theory. As discussed in detail in the solubility theory section, there are good thermodynamic reasons for believing that nanoparticles are soluble and that the Hansen Solubility Parameter framework can be legitimately applied to it. Although this is (at the time of writing) rather controversial, the fact that DLVO utterly fails to explain why short-chain stabilizers provide stability in specific solvents means that an alternative approach is required—and as 40 years of experience show that solubility thinking works adequately, maybe the ideas shouldn't be so controversial after all.

The investment in understanding how solubility parameters apply to steric stabilization pays off because the same rules apply to the stabilization (solubilization?) of naked nanoparticles such as CNT and also apply to nanoparticle-matrix interactions where the matrix is another polymer or (for example) an acrylate blend. It opens a rich world of practical understanding that DLVO simply cannot provide.

4.2. HANSEN SOLUBILITY PARAMETERS (HSP)

As large parts of this chapter (and other parts of the book) are based on HSP, some background justification is required [1].

As mentioned above, beautiful theories such as DLVO can provide endless intellectual stimulation, yet precious little practical help for the

sorts of complex formulations required in commercial coatings. There is a big difference between understanding a pure nanoparticle in a scientifically clean environment and an affordable nanoparticle present in a blend of commercial components, each of which might itself be a blend of other components (think acrylates, surfactants, plasticizers, etc.).

The practical formulator tends to turn to simple tools. "Hydrophilic/Lipophilic" or "Polar/Non-Polar" are terms often encountered in discussions of formulations. They certainly capture some aspects of molecular environments, yet are non-numeric and surprisingly vague. "X is lipophilic and was therefore dissolved in ethanol" [a real sentence found in a real academic paper] is a sentence which makes some sense to the authors who were disappointed that X was insoluble in water, but many of us would take exception to the idea of ethanol being "lipophilic".

There have been heroic attempts to better characterize the complexities of solvents, solid chemicals and polymers more numerically. Kamlett-Taft parameters are a noble example.

Modern computer power makes it feasible to understand compatibility at a more fundamental level. The COSMO-RS theory as embodied in COSMOtherm is an excellent example of a blend of quantum mechanical and theoretical rigor with practical implementation [2]. It places each molecule in a virtual dipolar sea made up from the solvent and can produce solubility results of very high accuracy.

The early nanoparticle community (otherwise known as paint and ink formulators) were initially beguiled by Hildebrand's Solubility Parameters (SP). This technique elegantly cut through the complexities of thermodynamics to show that a single number, SP, could characterize the ability of, say, a polymer to dissolve in a solvent. Unfortunately, the theory was based on the assumption that the solvents and polymers had no significant polar or hydrogen-bonding attributes, which rather limited the applicability.

Charles Hansen was not the first to think of overcoming that crushing limitation of Hildebrand's SP by breaking it into multiple components. He was, however, the first to create a coherent set of parameters for solvents, chemicals, polymers and pigments, which suddenly allowed the paint and ink nanoformulation industry to create robust, complex formulations based on just three SP which came to be called (but not by Hansen) Hansen Solubility Parameters (HSP). More than 40 years on from their creation, HSP are still proving themselves capable of providing practical insights to the latest nanoformulations: carbon nanotubes, graphene, quantum dots, nanoclays, organic photovoltaics and more.

The successes of this old technique in these new areas, and the fact that it can be applied to so many aspects of nanocoatings provides, the authors believe, more than sufficient justification for using the technique extensively throughout the book.

4.2.1. 3 Numbers

HSP are three numbers:

- δD is the Dispersion parameter that describes the polarizability of a molecule. Small, simple molecules like methane have low δD values; large aromatic molecules such as naphthalene—or molecules with lots of chlorine atoms, such as chloroform—have higher δD values. δD is strongly related to the refractive index of a molecule. It can also be thought of in terms of the van der Waals forces described in the DLVO section.
- δP is the Polar parameter that describes the polarity. Dull molecules such as methane have small or zero δP values; acetonitrile has a very high δP value because of the large charge difference across a small straight molecule. δP is strongly related to the dipole moment.
- δH is the Hydrogen-bonding parameter. Methanol has a high δH as it is an obvious H-bond donor and acceptor. Chloroform is a reasonably strong H-donor, so it has a reasonably high δH and the C=O group in acetone is a reasonably strong H-bond acceptor, thus providing a reasonably high δH.

Although the specific values aren't immediately obvious to a standard chemist, the general trends conform to chemists' intuitions. That is one of the many strengths of HSP—the numbers make sense. Here are some typical examples, Table 4.2.

By tradition, the HSP for any material are shown as triplets in square brackets in the order δD, δP, δH. Ethanol, for example, is [15.8, 8.8, 19.4] which means a moderate δD (15.8), a reasonably high δP (8.8) and a very high δH (19.4).

What are the numbers for? Everyone knows that "Like Dissolves Like". The key insight of HSP is that the definition of "Like" is a simple equation. Given two materials with HSP [δD_1, δP_1, δH_1] and [δD_2, δP_2, δH_2] the HSP Distance is given by:

$$\text{Distance} = \sqrt{4(\delta D_1 - \delta D_2)^2 + (\delta P_1 - \delta P_2)^2 + (\delta H_1 - \delta H_2)^2} \quad (4.6)$$

TABLE 4.2. The Hansen Solubility Parameters of Some Well-known Solvents.

Solvent	*δD*	*δP*	*δH*
Acetone	15.5	10.4	7
Acetonitrile	15.3	18	6.1
Benzene	18.4	0	2
Cyclohexane	16.8	0	0.2
Dimethyl Sulfoxide (DMSO)	18.4	16.4	10.2
Ethanol	15.8	8.8	19.4
Ethyl Acetate	15.8	5.3	7.2
Hexane	14.9	0	0
Methanol	14.7	12.3	22.3
N-Methyl-2-Pyrrolidone (NMP)	18	12.3	7.2
Methylene Chloride	17	7.3	7.1
N,N-Dimethyl Formamide (DMF)	17.4	13.7	11.3
Tetrahydrofuran (THF)	16.8	5.7	8
Toluene	18	1.4	2
Water	15.5	16	42.3

Apart from the factor of 4 before the *δD* terms, that Distance is simply the 3D equivalent of the familiar Pythagoras distance (square root of the sum of the squares) in 2D. When the Distance is small, the two materials are very alike; when the Distance is large, the two materials are not alike. This is the core of what makes HSP so useful—an objective, numeric measure of (dis)similarity.

The definitions of "small" and "large" Distance (where 0 is ideal) depend on the system; for a large molecular weight crystalline polymer which is not soluble in many solvents, even a relatively small Distance can lead to low solubility. For a small molecular weight amorphous polymer which is soluble in many solvents, it requires a large Distance for insolubility. This rather vague definition will be made more precise shortly.

The parameters are grounded in thermodynamics. The sum of the squares of the parameters is related to the enthalpy of vaporization of a molecule, and because of the strong links to refractive index (*δD*) and dipole moment (*δP*) there are tough constraints on their values. Nonetheless, it is an unfortunate fact that there is currently no objective way to assign reliable values to a molecule. Instead, for small molecules users rely on a core set of ~1000 values that have withstood the test of time, from which automated methods have been derived for predicting

HSP values from the chemical structure. For polymers and nanoparticles there is a simple way to measure the HSP using a suite of 20–40 solvents that span HSP space. This so-called Sphere technique is what gives HSP its great practical strength.

4.2.2. The HSP Sphere

Place a small amount of polymer or nanoparticle in each of, say, 20 test tubes and add a small amount of a different solvent to each tube. Shake, then judge how "happy" the sample is in each tube. Give a score of "1" to samples that are happy and "0" to those that are unhappy. Now find the portion of 3D HSP space where all the "1" solvents are inside and all the "0" ones are outside. This defines a sphere that characterizes that material. The center of the sphere marks the HSP of the material and the radius marks its limits of solubility, compatibility or, in general, "happiness". Other solvents, polymers, particles that are inside that sphere will, in general, be compatible with the material and those outside that sphere will be incompatible. In one simple experiment a wealth of new information, immediately applicable to real-world formulation, has become available.

As this is a scientific book, the word "happy" looks rather out of place, although it is a term frequently used within the surface coating industry. Yet it is used deliberately. Look at the picture of some CNT placed into four different solvents, sonicated and centrifuged [3], Figure 4.6.

Even a non-expert on CNT knows that they are not at all happy in the solvent on the right, are rather unhappy in the next tube, are moderately happy in the next and entirely happy (assuming, as is the case, that the picture looks the same after many weeks) in the tube on the left. The

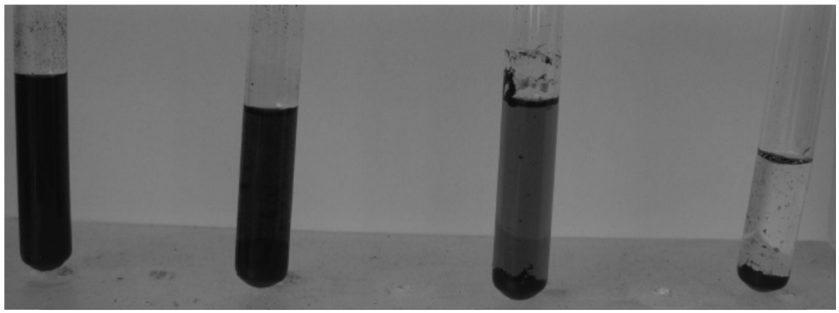


FIGURE 4.6. *The difference between "happy" and "unhappy" CNT in four different solvents.*

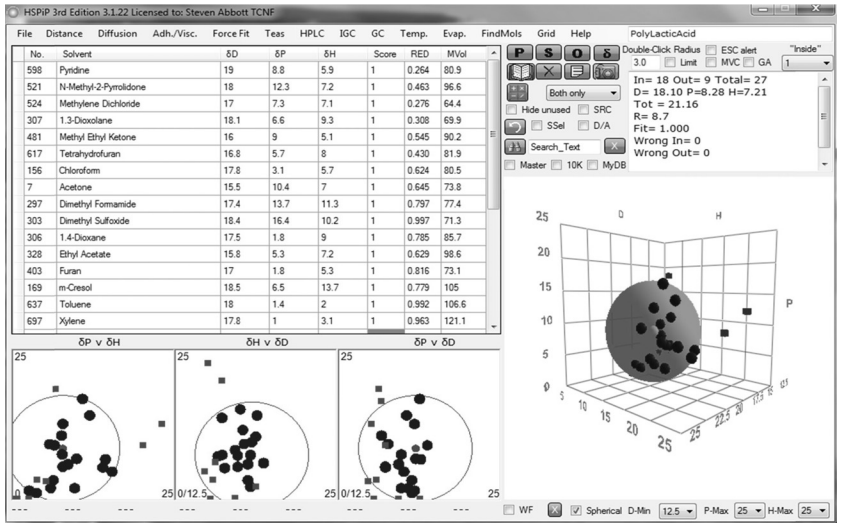


FIGURE 4.7. *The HSP Sphere fitted to the happy/unhappy PLA data from 27 solvents.*

author of the CNT study (discussed in more detail later) had originally tried to measure the solubility of the CNT in each solvent. It quickly became apparent that an entire PhD project could be taken up merely obtaining good solubility values for a range of solvents. By going for the less scientific but more practical test of general happiness, many deep insights into CNT behavior were readily attained.

As a simple demonstration of the Sphere technique, here is the result of fitting the solubility data [4] of Poly(Lactic Acid) (PLA) in 27 solvents, as discussed in detail elsewhere by Abbott [5], Figure 4.7.

Of the 27 solvents, 18 were "happy" (shown inside the sphere) and 9 were "unhappy" (shown outside the sphere), giving unique values for the center (a thermodynamic value independent of the definition) and the radius (which depends on the definition of happy/unhappy). As discussed below, from this rather simple characterization of the properties of PLA it is possible to reach some profound conclusions about success and failure of dispersions of nanoparticles within the polymer.

For convenience, HSP Sphere fits will be shown in this book using output from the software package called Hansen Solubility Parameters in Practice, HSPiP [6]. Readers should be aware that Abbott is a co-author of the software. Because the key HSP data and algorithms are in the public domain, and to ensure that readers do not have to rely on an external software package, the spreadsheet package includes a Sphere

fit routine that gives the same results as HSPiP—plus some other necessary HSP-oriented calculations.

A different approach to measuring happiness for the HSP Sphere of a nanoparticle is provided in the work of the Azema group at the Ecole des Mines d'Alès [7]. They determined the HSP of fumed silica—raw and octadecylamine functionalized by two techniques. The first was the simple binary division into "It sediments" and "It doesn't sediment". This gave adequate results, but by using numeric data from centrifugation they were able to get higher-quality values. Using the methodology discussed below, by reducing the HSP mismatch between the silica and their chosen polymer, polypropylene, they were able to greatly improve the quality of the dispersion.

4.2.3. Applying Solubility Parameters—Solvent Blends

In the discussion on steric stabilization the question was raised as to how to go from a few known solvent compatibilities (e.g., with acetone, MEK, ethanol, toluene, etc.) to compatibility with solvent blends.

If the compatibility of the test material with a reasonable set of pure solvents is known, then the resulting data can be placed into a Sphere calculation. It is important to remember that the selection of solvents is made with respect to their distribution in HSP space, not for their utility within a practical formulation. The center of the sphere characterizes the HSP of the polymer dispersant (assuming that the coverage is sufficiently good that the solvents are effectively seeing just the dispersant). The radius of the sphere characterizes how far away from that center it is possible to go without causing flocculation.

The Sphere data can be applied directly to any other solvent; just calculate the HSP Distance from the center and compare it to the radius. If it is less, then the solvent will be OK. To normalize distances, the RED (Relative Energy Difference) number is calculated as Distance/Radius. By definition, a solvent with RED < 1 is good and one with RED > 1 is bad.

The same logic applies to solvent blends. The HSP of a 50:50 MEK: toluene blend is a 50:50 blend of their δD, δP, δH values. It is simple, therefore, to calculate the HSP of each blend and therefore the Distance from the target and the RED, where a value > 1 predicts flocculation. The HSP Solvent Blends spreadsheet does the calculations for you. In this specific example the target is [17, 9.7, 3] with a radius of 5. This target has been set to show an interesting and important point. The target

TABLE 4.3. The HSP of Blends of Two Solvents.

	Solvent	***δD***	***δP***	***δH***	**Distance**	**RED**
S1	Methyl Ethyl Ketone (MEK)	16	9	5.1	3.07	0.61
S2	Toluene	18	1.4	2	6.97	1.39
% S1	**% S2**					
100	0	16	9	5.1	3.07	0.61
90	10	16.2	8.24	4.79	2.41	0.48
80	20	16.4	7.48	4.48	1.98	0.40
70	30	16.6	6.72	4.17	1.91	0.38
60	40	16.8	5.96	3.86	2.25	0.45
50	50	17	5.2	3.55	2.85	0.57
40	60	17.2	4.44	3.24	3.59	0.72
30	70	17.4	3.68	2.93	4.39	0.88
20	80	17.6	2.92	2.62	5.23	1.05
10	90	17.8	2.16	2.31	6.09	1.22
0	100	18	1.4	2	6.97	1.39

polymer is readily soluble in MEK and insoluble in toluene. When the "bad" toluene is added to the "good" MEK, the resulting blend is closer to the target so that 70:30 MEK: toluene would prove to be significantly better than pure MEK. Only at 40:60 does the solubility get worse and 20:80 would give flocculation, Table 4.3.

It is an important consequence of HSP thinking that it is possible to find two bad solvents which together create a good solvent or, as in the above example, a bad solvent can improve the solubility of a good solvent. An example commonly observed in coatings of UV acrylates is that some—which are neither soluble in IPA or Toluene—are readily soluble in a mixture of the solvents. This is very liberating for formulators who rarely find that any single solvent has the right balance of properties and are therefore forced to use solvent blends. In the chapter on coating, it is shown that the changes of solvent blend during drying can be used to tune the solubility/flocculation behavior throughout the drying process.

So useful is the concept of HSP of solvent blends that a "grid" method of measuring the HSP sphere is now being used. First exemplified by Machui and colleagues at U. Erlangen in the context of organic photovoltaics (see below), the idea of taking a few solvents and scanning an HSP space via solvent mixtures has proven to provide a fast (it is particularly suited to robotics) and accurate (because it can be fine-tuned) method for determining the HSP sphere [8].

4.2.4. Applying Solubility Parameters—Getting Good Solvent Dispersions

As mentioned at the start of this chapter, there are two inter-related aspects of the stability of nanoadditives. In this section the focus is on nanodispersions. In the following section, the focus is on the interactions between the nanoadditives and the matrix (polymer).

Whether CNT are dispersed or dissolved in solvents is a matter of some debate and the issue is raised in the discussion on solubility theory in Chapter 5. In either case, it is clear that HSP can help in understanding which solvents are good or bad. The Namur CNT Case Study is just one example of such insights.

4.2.4.1. Case Study: CNT

Dr. Simon Detriche, working in the Mekhalif group at U. Namur, Belgium, wanted to better understand the behavior of different classes of CNT—SWNT (Single Walled Nano Tubes), MWNT (Multi Walled Nano Tubes), FNT (Functionalized NanoTubes) [3]. At first it seemed scientifically sensible to measure the absolute amount of CNT held in dispersion/solution but it became apparent that there was no simple way to measure what is an ill-defined concept. Detriche therefore shifted to a quick, simple, reproducible protocol of placing a fixed amount (5mg) of CNT into a tube containing 5 ml solvent, sonicating for 2 minutes, centrifuging at 3500 rpm for 30 minutes then scoring the resulting tubes visually on a 1–6 scale, where 1 was “excellent dispersion/solution” and 6 was “no dispersion/solution”. Tests showed that sonicating for longer or centrifugation under somewhat different conditions made no significant difference to the scores.

Even the best solvents seemed to be incapable of fully dispersing a sample. Careful tests showed that the non-dispersible portions could not be dispersed even in fresh solvent. Microscope analysis showed why—these CNT were clusters stuck to lumps of metal catalyst.

Further work showed that solvents could select CNT by size: the non-dispersed CNT were much larger in size than the dispersed portions; in other words, classic entropic effects that apply to polymer solubility seem to apply to CNT.

Following the usual procedure of entering the scores, a reasonable fit could be obtained for those solvents that scored a “1” for very happy, Figure 4.8.

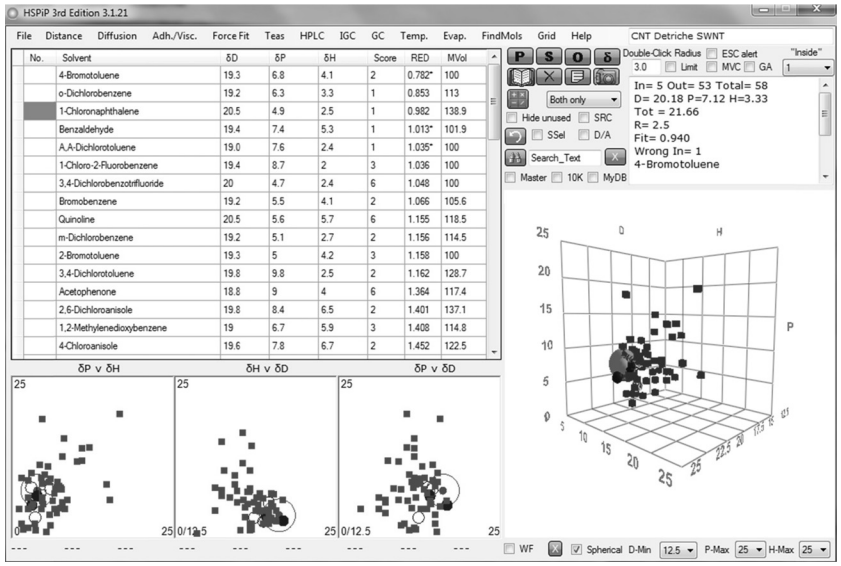


FIGURE 4.8. *The fit to SWNT using the most restrictive definition of "happy" yields a very small sphere.*

This led to a value for the HSP ~ [20, 7, 3]. If the solvents that scored a 2 are included, Figure 4.9.

The HSP value changes only slightly to ~ [20, 7, 4] with an increase of radius, which is what would be expected from a more relaxed definition of a good solvent.

The fact that the Coleman group, using a very similar procedure, found values closer to [18, 9, 8] raises the question "what is a CNT?" [9]. Detriche believes that his SWNT were (despite the catalyst clusters) of very high quality, having been carefully annealed for many hours at high temperatures. Whether the batch of CNT used by Coleman is somehow different is an open question. The differences are not small. An "excellent" solvent for Coleman (NMP) is a "poor" solvent for Detriche. Good (i.e., pure and annealed) MWNT are not so different in HSP terms from SWNT, which makes sense.

The real insight came from analysis of FNT created from MWNT. The method used by Detriche is one common to the CNT world: heating in strong nitric/sulfuric acid. This oxidizes the CNT. The HSP "Sphere" for FNT is a mess—there isn't a reasonable sphere at all—as is more clearly shown using a wire-frame view, Figure 4.10.

When HSP Spheres fail, this may be due to simplifications in the theory; or it may be a sign that the samples are impure. What could be

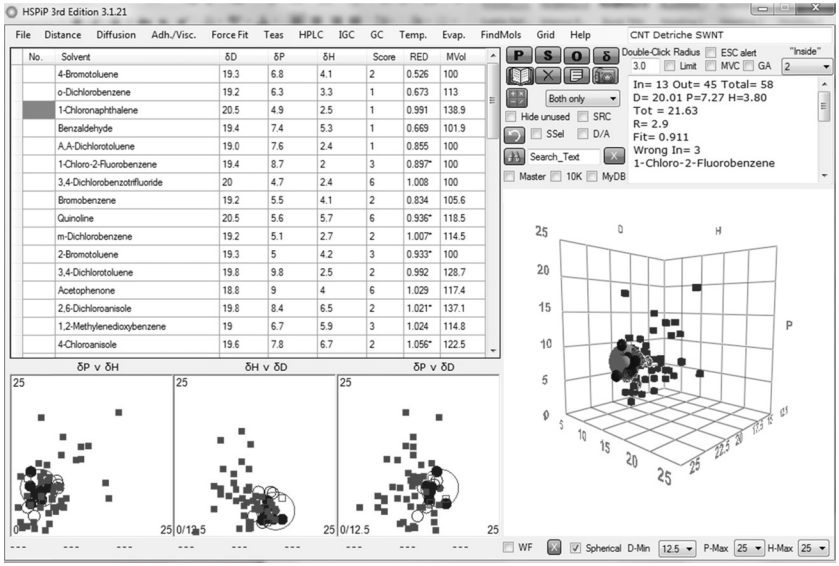


FIGURE 4.9. *The fit to SWNT with a more relaxed definition of "happy" gives the same HSP and a larger radius.*

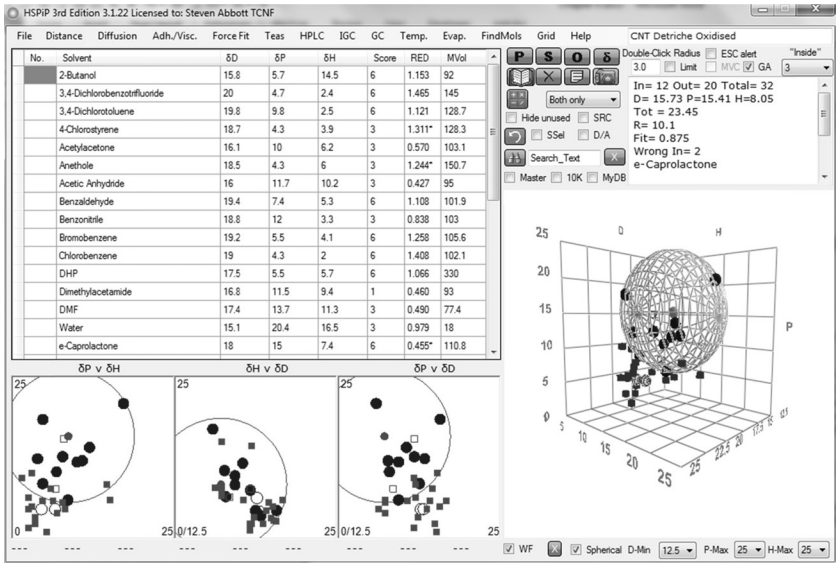


FIGURE 4.10. *The fit to Functionalized CNT makes no sense.*

the "impurity" in FNT? Detriche's conclusion is that the impurity is super-perfect MWNT. By this he means that a really perfect MWNT is not attacked by the nitric/sulfuric mixture. The MWNT with minor defects are the ones that are oxidized. Indeed, using a Double Sphere fit, it looks as though the hypothesis is at least plausible. One Sphere contains the un-reacted MWNT and the other seems reasonable for the highly functionalized nanotubes, Figure 4.11.

4.2.4.2. Case study: Quantum Dots

LED lights are typically monochrome and users for lighting applications want them to be white. Combining red, green and blue LEDs can give white light. Alternatively, a blue LED can be coated with a phosphor blend that emits white light. Many phosphors are specific chemical compounds that emit light at a wavelength depending on the chemical structure.

Quantum dots are optically versatile materials whose behavior arises not from chemical structure, but from dot size. As discussed in Chapter 2, the electrons get trapped, like the classic quantum "particle in a box", and can only emit at the wavelength controlled by the size of the

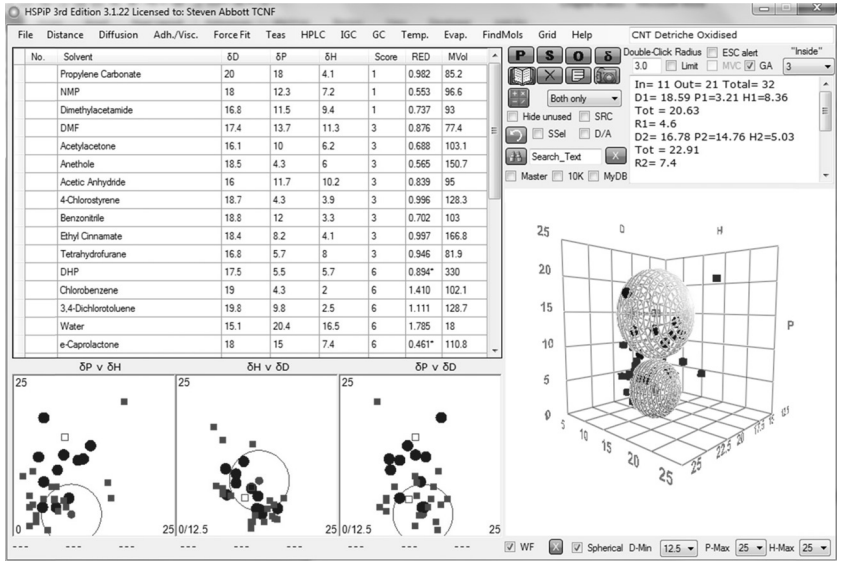


FIGURE 4.11. *Functionalized CNT fitted to two spheres, one of which is in the right region for MWNT.*

box. Thus, a single chemical structure such as CdSe can emit all wavelengths from blue to red as the size of the dot changes over a range of a few nm. In effect, the size of the dot is the size of the box.

To be well-dispersed, quantum dots require some form of stabilizing shell around them. With their delicate quantum electronic nature, the energy levels of the dots can be tuned by the shell. The Rosenthal group at Vanderbilt University set out to obtain both dispersibility and tunability through the use of substituted phosphonic acids [10]. The electronic effects are discussed in their paper. This case study involves the team's efforts to understand the dispersibility through the use of HSP. Each phosphonic acid variant was subjected to the Sphere test and the data were sent to Abbott and Hansen for analysis. The calculated values for the butyl phosphonate dots [17.1, 4.2, 1.5] seem reasonable for a somewhat polar butyl chain, Figure 4.12.

The problem is that when one goes to the phenyl phosphonate, the values are remarkably similar. The fit of [17, 1, 5] (not shown) has no obvious reason for a lower δP and a higher δH. The fit for the hexadecylphosphonic acid version [16, 5, 1] shows some expected decrease in δD for the hexadecyl chain, though the δP and δH might have been expected to decrease. And, surprisingly, the 2-Carboxyethylphosphonic acid fit [16.4, 4.8, 3.2] shows no evidence for the expected higher δP and δH. Even worse, one of the fits was of poor quality. Here is the dodecylphosphonate, Figure 4.13.

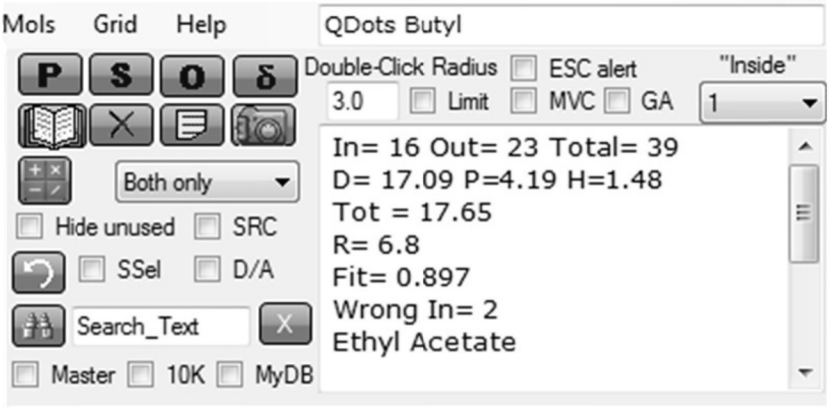
This makes no sense in the context of the butyl, the hexadecyl and others in that homologous series.

The worries about the small changes between the different substituents assume that the CdSe surface is entirely covered by a shell of substituted phosphonic acids, with the chains sticking out into the solvent, so the HSP should be that of the chains. But what if some of the CdSe or the phosphonate group is accessible to the solvent—how much would that contribute to the HSP? Those issues are still unresolved.

The key part of this case study is that for the dodecyl samples the Sphere was unconvincing. This could mean that HSP does not apply to these particular dots, which would be worrying. An alternative explanation was that this batch was contaminated by excess phosphonate. The sample that gave poor results was checked using Rutherford Backscattering and was indeed found to contain excess phosphonate. The take-home message is that a relatively crude macro technique—shaking up quantum dots with solvents—was able to reveal problems with ultra-small nanoparticles.

The issue of dispersing the quantum dots into polymers in order to make practical phosphor coatings is discussed in the next section.

There is also an interesting observation to be considered concerning the behavior described above. As discussed earlier, these stabilizers



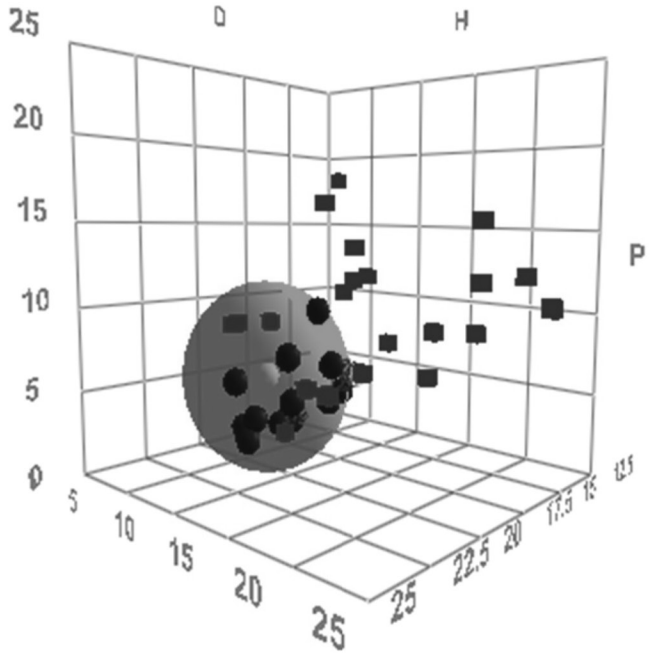


FIGURE 4.12. *The fit of the butylphosphonate quantum dot showing how even quantum dots can behave reasonably in terms of HSP.*

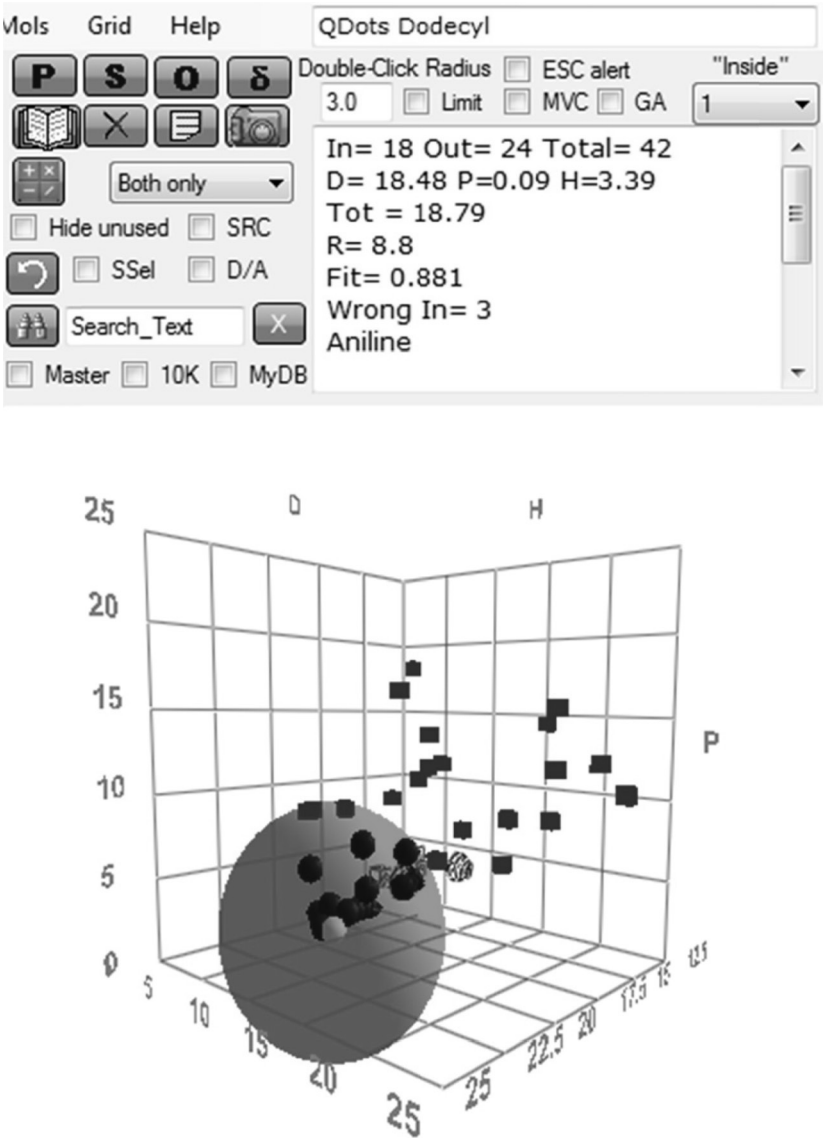


FIGURE 4.13. *Although the other 8 substituted dots had reasonable values, the dodecylphosphonate made no sense.*

fall far short of being long-chain polymers, so DLVO theory cannot be applied (or if it were applied it would predict flocculation). This is a reminder of the need for (and justification of) the use of solubility thinking applied to nanoparticles.

4.2.5. Applying Solubility Parameters—The χ Parameter

The DLVO theory of steric stabilization requires knowledge of the χ parameter. This can be measured (e.g., by light scattering) for the stabilizing polymer but most of us don't have the time, experience or patience to do the measurements, especially within the context of messy real-world formulations. This is where HSP are so useful. It can be shown that the χ parameter is related to the HSP Distance via:

$$\chi = \frac{\text{MVol.Distance}^2}{4RT} \tag{4.7}$$

where MVol is the molar volume of the solvent and RT is the usual gas constant times absolute temperature. The formula is included in the Solvent Blends and Chi spreadsheet. The factor of 4 is only approximate because both sets of theories (DLVO and HSP) are making assumptions about polymer-solvent interactions. So although DLVO makes it clear that dramatic things happen precisely at $\chi = 0.5$, the reality is that the aim of a formulator is to get χ down to a reasonably low value to give a large margin of error to cope with the differences between theoretical models and practical reality. Recall that the HSP of a messy real-world stabilizer can be measured by the Sphere technique and the HSP of a complex solvent blend can readily be calculated. To achieve a low value of χ in practice means obtaining a small Distance.

4.2.6. Applying Solubility Parameters—Getting Good Polymer Dispersions

Getting good stand-alone dispersions of nanoparticles is often an important step for further formulations. Now it is time to look at ways to ensure that the particles are easily dispersed, and remain dispersed, in a polymer matrix. As mentioned earlier, work on silica and polypropylene showed that reducing HSP distance greatly aided dispersion [7]. Another way to explore this question is via the search for improvements to the properties of the "green" polymer polylactic acid, PLA.

PLA is produced from lactic acid which in turn is produced via fermentation of biomaterials. It is a polyester which might act as a replacement for synthetically sourced PET (polyethylene terephthalate), the polyester most often used in soft drink bottles. PLA is, for example, crystal clear like PET and can be quite readily molded into bottle

shapes. Unfortunately, PLA is not as tough as PET and cannot withstand the normal handling shocks of PET bottles. There have been numerous attempts to increase the toughness of PLA. One method, which could also improve its barrier properties, is to incorporate nanoclays. As the clays are "natural", PLA can maintain its green status.

The nanolayers that make up clay are held together by Na^+ ions. To separate the layers so that they can be dispersed into non-aqueous environments ("exfoliated") requires replacement of the Na^+ ions by larger cations that are too big to fit nicely between two layers. The most common exfoliants that are also approved for food use are tetralkylammonium groups, often based on tallows—long-chain alkyl groups in the C14-18 range, which can be hydrogenated or non-hydrogenated.

A typical example is Cloisite 15A, which has two methyl groups and two hydrogenated tallows. How likely is it that a clay which is exfoliated via Cloisite 15A will be compatible with PLA? Language such as "hydrophilic/lipophilic" doesn't help. The Cloisite exfoliants have both hydrophilic (ammonium) and lipophilic (alkyl chain) components. If it is assumed that the hydrophilic parts will be closely attached to the clay then the two tallows will be exposed to the PLA which will see, therefore, a largely hydrophobic environment. Given that PLA is sometimes called a hydrophobic polymer (it is insoluble in water), that could be taken to mean a perfect compatibility. A chemist's intuition would be uncomfortable with that argument, as a polyester isn't very much like an alkyl chain, even if both are considered to be "hydrophobic". This vague sort of debate is rather common in the PLA/nanoclay literature and does not seem to have advanced the subject greatly. Indeed, many of the papers express frustration that the nanoclays provide little or no benefit to the properties of PLA.

A numerical approach seems a rather more helpful way of examining the issue. First, then, what are the HSP values of PLA? As noted above, the HSP for PLA are ~[18, 8, 7]. From data scattered through the literature, it is possible to provisionally assign a value of [17, 3, 3] to Cloisite 15A. The HSP Distance between them is ~8, which is generally seen as a large distance. In other words, the chances are small that the PLA and nanoclay will be compatible.

Now try a different Cloisite (10A) with one tallow and one benzyl group. At ~[18, 4, 5] the HSP distance (5.8) from PLA is significantly smaller. Finally, Cloisite 30B with one tallow and two $–CH_2CH_2OH$ groups has rather less δD (no aromatic), but considerably more δP and δH and at [17, 9, 8] has a distance of 3.7. This should be nicely com-

patible with PLA. Because PLA is very susceptible to hydrolysis and transesterification, during processing the OH groups of the Cloisite might link into PLA chains to provide some extra compatibility (positive) while degrading the PLA MWt (negative).

What is the best Cloisite to use with PLA? At the time of writing it is impossible to disentangle the literature sufficiently. The tallow nanoclays seem to be rather unhelpful. The one with the benzyl group is, on the basis of the approximate data available, the "least bad", unless it turns out that the $–CH_2CH_2OH$ functionality improves things by creating strong bonds with the PLA rather than causing the molecular weight to plummet unhelpfully. Some careful measurements and, perhaps, an ester-substituted tallow might provide the best combination. The nanoclay Case Study is only a slight digression from PLA. It involves PET which, in solubility parameter terms, is not so very different. It is a salutary lesson in the need for clear thinking about what a nanoadditive is supposed to do. But first, a return to quantum dots.

4.2.6.1. Case Study: Quantum Dot Polymer Dispersions

Returning to the Rosenthal group's work on quantum dots, in a separate paper on phosphonic- and phosphine oxide-stabilized CdSe nanocrystals it was shown that solubility parameter analysis could help to map from solvent dispersibility onto polymer dispersibility [11]. Although the paper does not contain a full analysis of all the materials, it seems safe to assume that HSP would provide helpful insights both to the solvent dispersibility of the quantum dots (toluene was the closest match in their list of solvents) and to the quantum dot compatibility with the polymers. The epoxies at ~[18, 12, 10] are very far from the quantum dots, and so did not produce good encapsulation. Not surprisingly, therefore, they obtained the best encapsulation with a polymer BP-PFCB that is likely (by reasonable estimates) to be much closer in HSP, though no sphere measurement of its value was made.

4.2.6.2. Case Study: Nanoclays in PET

The aim of the project led by Dr. Tim Gough at the University of Bradford was to improve the barrier properties of blow-molded PET bottles using nanoclays. This was a major study that used the group's access to excellent processing and analytical equipment to ask searching questions in an attempt to get unequivocal answers.

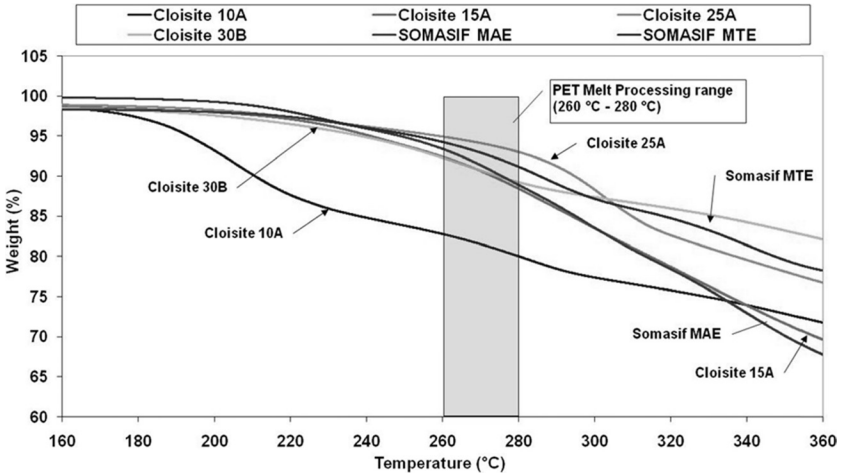


FIGURE 4.14. *TGA weight-loss plots for various clays showing, especially, that Cloisite 10A degrades at too low a temperature for the melt processing of PET.*

The first question was: PET must be processed at ~260–280°C, will any of the commercially available clays survive such temperatures? The way to answer the question was via TGA (Thermo Gravimetric Analysis), Figure 4.14.

Immediately Cloisite 10A, the most likely material (in HSP terms) to be compatible with the PET, is ruled out— it degrades even below the 260–280°C temperature range.

The 30B has –OH groups that (as discussed above) can react with the ester functions and therefore must be rejected.

The choice is therefore restricted to the following materials:

- Cloisite 15A (which is unlikely to be compatible)
- Cloisite 25A, which replaces one tallow group with an ethylhexyl group
- The two Somasif clays, where MAE is similar to Cloisite15A, which might explain their near-identical plots, and is therefore unsuitable and the MTE which contains methyl trioctyl ammonium groups and is similar to the 25A but still not ideal.

Tests on the Tg and MPt of the PET showed essentially no effect of even 20% addition of the clays—showing little effective interaction.

There were some interesting effects on the temperature of crystallization from cooling—or, to put it another way, the clays acted as seeds for crystallization, in principle reducing the cycle time of the molding

process by a factor of 2 or 3. This may not be a primary aim of adding the clays, but it is a useful property to have been gained.

Measuring the rheological behavior over a wide range of temperatures and shear rates required a lot of work. A summary, that does not do the work full justice, is that there was another modest, but helpful, improvement via a reduction in high-shear viscosity by a factor of 2 for moderate loadings. This could potentially reduce molding cycle times.

In order to study mechanical effects, the team produced high-quality films via extrusion. These were then measured on a tensile tester. Interestingly, a few % loading of the MTE was sufficient to double the tensile modulus, Figure 4.15.

Why the MTE should be significantly better than the MAE or 25A is not at all obvious, given their similarities in structures and, so far, similar effects on other properties.

The main aim of the project was to achieve better barrier properties. The hope was that a highly-oriented dispersion would provide good barrier properties thanks to a high tortuosity (see Chapter 2), Figure 4.16.

Measurement of the barrier properties is not an easy task when the material (PET) is already quite a good barrier. The team was able to

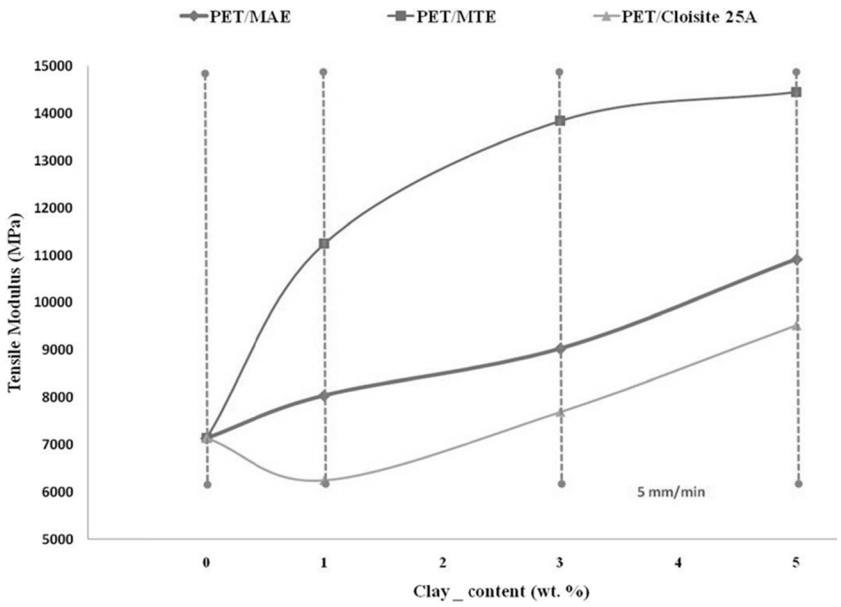


FIGURE 4.15. *The effect of % clay content on tensile modulus for different clays.*

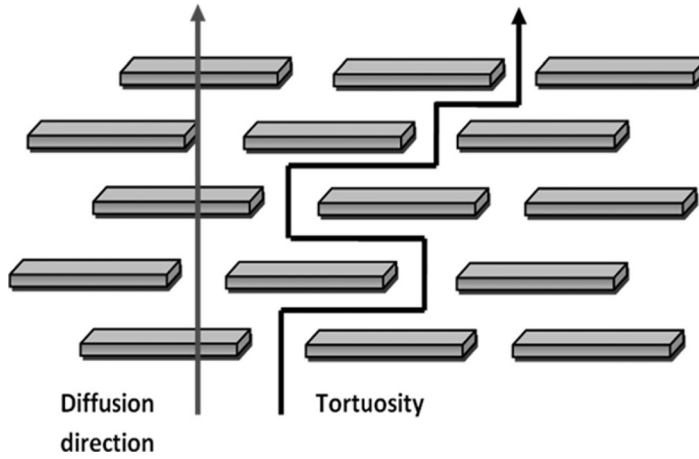


FIGURE 4.16. How tortuosity increases barrier properties.

gain access to a sophisticated MOCON OX-TRAN test instrument to measure the oxygen permeability. The good news was that a factor of 2 improvement could be obtained with the 25A, Figure 4.17.

Arguably, the differences between the rather similar clays are not too significant (though the data themselves were shown to be highly

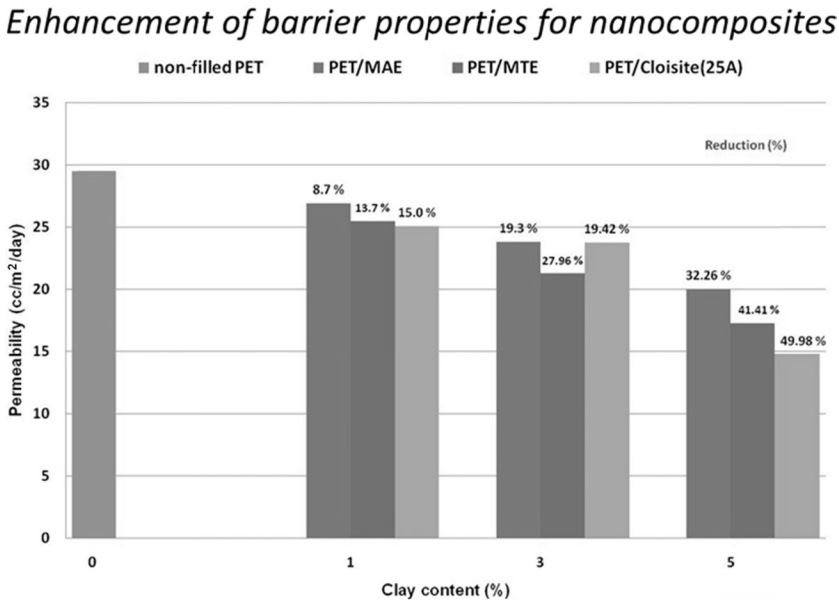


FIGURE 4.17. Reduction in permeability with increasing concentrations of three different clays.

repeatable) at the 1 or 3% level. At 5% loading the 25A is definitely superior, which is odd because it happens to have a significantly smaller aspect ratio (length/thickness), whereas tortuosity and barrier properties should be better for higher aspect ratios.

The final question is whether the 5% loading of the clay gives acceptable optical quality. A glance at the raw extruded resin seems to suggest not, Figure 4.18.

Fortunately, further experiments with real films and bottles showed that the coloration was not quite as unacceptable as the images might suggest.

In summary, this case study shows that what seems to be simple ("just throw in some nanoclays to improve barrier properties") requires a great team with access to a wide range of excellent equipment, asking tough questions and recognizing that everything we do in life is a compromise.

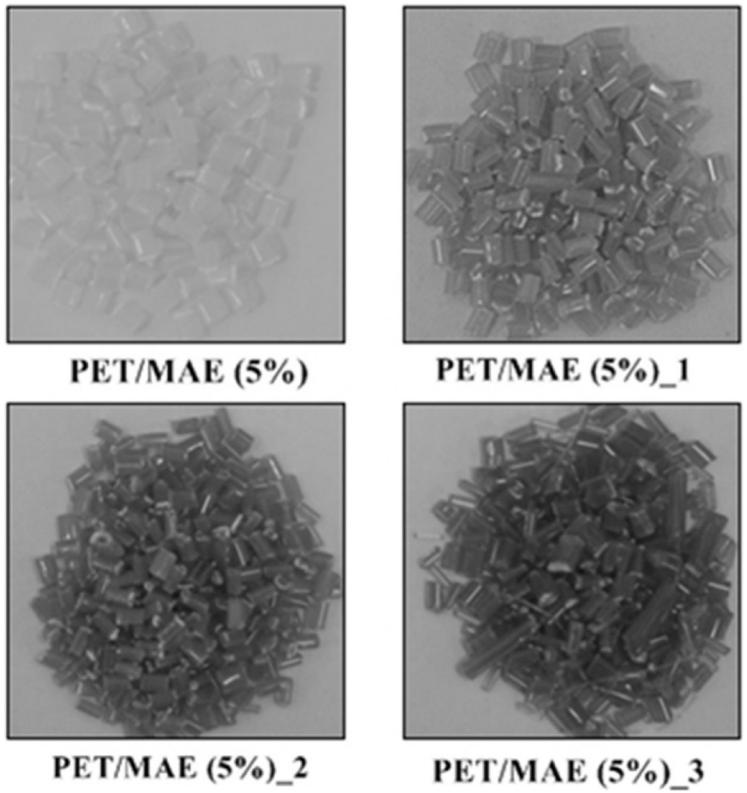


FIGURE 4.18. *The effect of clay concentration on the color of extruded PET resin.*

The authors' opinion is that if the clays were functionalized with groups that were naturally more tuned to the PET (without the drawbacks which caused the team to remove two more promising clays from the tests), the probability of more significant improvements in the overall properties would be increased. Whether it is in the interest of the clay manufacturers to provide additional functionalities is a question the authors cannot answer.

4.2.7. Applying Solubility Theory—Stable Dispersions in UV Formulations

There are many attractive possibilities for nanoparticles dispersed in UV-curable resins. A typical resin is HDDA (1,6 Hexanediol Diacrylate), which is popular because it is di-functional, has low viscosity and reacts rapidly on curing. It has often been noticed that some nanoparticle dispersions are happy in HDDA and some are unhappy—which is not a surprise. The worst situation is when a "happy" dispersion turns unhappy sometime after arrival at the customer's premises. Why might a seemingly stable dispersion go bad?

Part of the answer to the question is linked to the fact that we want UV systems to cure quickly, so adding a lot of radical stabilizer to the formulation is not helpful, though adding none at all is usually a recipe for disaster. One stabilizer is potent and provided free with each pot, provided enough ullage is included: oxygen. If anything about the nanoparticle consumes oxygen then the formulation can suddenly go solid.

Very small nanoparticles (usually a good thing) can be a source of problems. Their high curvature means that the surface atoms are in an unusually high-energy state; this might help trigger some self-polymerization.

The other part of the answer comes when the acrylate is considered as a solvent. If it is a poor solvent then for χ parameter reasons any steric stabilization might be compromised so that the formulation might flocculate—though this should be a rapid and obvious process.

There is one possible cause that is rarely discussed. Another reason that HDDA is much used is that it is a generally all-round good solvent. In terms of steric stabilization and in terms of solubility theory for particles, that's a good thing. However, as discussed in terms of real solvents, if the dispersant happens to be too soluble in the HDDA and if it is not strongly attached to the particle then it might, over time, remove itself from the particle—and eventually stabilization is lost.

4.2.8. Dispersions Summary

What applies specifically to nanoclays in PLA and PET and to quantum dots in other polymers applies generally to nanoparticle dispersions within polymers. Compatibility between the dispersant shell and the polymer is necessary to provide easy dispersion of the nanoparticles with little tendency to clump together during further processing. To the authors' knowledge, no better technique for thinking about "compatibility" exists than the HSP approach. Given that it is easy to measure the HSP of the particles and, (if they are not already in the literature) the HSP of the polymers, formulating for compatibility is relatively straightforward and far more productive than trial and error or reliance on ill-defined terms such as "polar" and "non-polar". So compatibility is a necessary condition for nanoparticle dispersion. The fact that it might be necessary but not sufficient to gain high performance requires a major digression.

4.3. AN IMPORTANT DIGRESSION ABOUT ADHESION

This section deals with the use of nanoparticles within a polymer matrix in order to give it extra toughness in the context of, for example, a hardcoat. A key insight from the previous section is that good HSP compatibility is vital for good dispersion. With compatibility it is possible to create a high loading of a nanoparticle such as silica. As we shall see, a high loading is necessary but not sufficient to guarantee toughness. For that, the particle must be adhered to the overall matrix. There appears to be much confusion about which aspects of the science of adhesion are relevant to nanocoatings, so this section attempts to resolve this issue.

4.3.1. Adhesion Via Surface Energy

Place two surfaces into intimate contact and an astonishing degree of adhesion can be obtained. The gecko uses a set of tricks to allow its super-compliant feet to be in nanocontact with a rough wall. The van der Waals forces between the feet and the wall are more than sufficient to support the weight of the gecko. No "glue" or other special effect is required to explain gecko adhesion. Scientific studies with artificial gecko feet show that the difference in adhesion between a low-surface-energy silicone and a high-surface-energy polyester is minimal [12].

FIGURE 4.19. *Abbott loses 30 kg using pure surface energy.*

A simple demonstration of the strength of perfect contact between two smooth surfaces is shown in the photo. Abbott is attempting to support his 80kg from a half-ton hoist using two pieces of rubber simply pushed together. After reducing his weight to 30 kg, the tape broke. Simple calculations show that an area of rubber about 1/4 the size would be sufficient to hold 80 kg.

There are two points to these two stories. The first is that obtaining a strong adhesive force via perfect surface contact is not a problem. The second is that this sort of adhesion is useless if any form of permanence is required. The gecko merely flexes its foot for the adhesion to vanish by crack propagation along the interface. Abbott could pull the two sheets of rubber apart with no effort by picking at one of the corners. In both cases the overall adhesive force against an overall pull is strong, but the work of adhesion is weak.

The work of adhesion of a bond created by surface energy is a pitiful 0.05 J/m² (i.e., the same as the surface energy itself, 50 dyne/cm). Real adhesion needs values 2000× larger. Nanoparticles in a nicely compatible matrix can be well-dispersed and show excellent surface adhesion—yet the matrix/particle bond can be separated with as little effort as the gecko's feet or the rubber pads.

4.3.2. Strong Adhesion Without Chemical Bonds

The formulator's instinct for increasing adhesion is to create a chemical bond between the particle and the matrix, but this appealing strategy has some serious pitfalls. To describe these, it is necessary to look at a surprisingly little-known source of very strong adhesion—adhesion so good that the work needed to break the bond increases the more vigorously the bond is tested. To understand this adhesion it is necessary to understand the difference between nails and screws.

If two polymers are nicely compatible (as judged by HSP), it is possible (via heat or solvent) to get them to intermingle. If the chains are of modest length, this interface gives a 20× increase compared to pure surface energy, i.e. 1 J/m². As the interface is forced apart, the polymer chains simply slide past each other. The extra work of adhesion comes from the friction forces between the chains. There is an obvious analogy with planks of wood held together by nails. Wool's group at the University of Delaware took this analogy seriously and wrote a paper (one of the few in the scientific literature to acknowledge the assistance of a carpenter) based on an analysis of the number of nails, their length and their friction coefficients [13]. A formula allowed them to nicely match their experimental results. As a beautiful example of the universality of science, this same formula applies exactly to intermingled polymer chains and the value of 1 J/m² can be plausibly estimated from nail theory. This "nail adhesion" between polymers is a welcome boost to adhesion, but just as it is unwise to rely on nails for strength (screws are superior), relying solely on nail adhesion for polymers is not a good idea.

This idea of nail adhesion has been investigated at the molecular level using the well-known JKR theory (discussed in more detail in Chapter 8) and test procedure (pressing a sphere onto a surface) and measuring the "adhesion hysteresis" which shows the difference between standard work of adhesion (surface energy) and extra effects such as intermingling. If two polystyrene surfaces are pressed together,

below their Tg, there is very little adhesion hysteresis. If the surfaces are cross-linked (separately) using UV light under nitrogen, there is no adhesion hysteresis. If the surfaces are opened up by UV under oxygen then there is significant adhesion hysteresis with a 10× increase in work of adhesion due to intermingling [14]. Using Sum Frequency Generation spectroscopy (a form of IR that shows only molecules at surfaces and interfaces), it can be shown that this intermingling effect correlates with increased adhesion [15].

Adhesion scientists typically model this sort of adhesion via variations on the basic de Gennes formula which says that work of adhesion, G, is given by [16]:

$$G = \Sigma NU \tag{4.8}$$

where Σ is the density of polymer chains across the interface, N is the number of bonds along the chain (typically 8 if the chains are overlapping by 1 nm), and U is the energy per bond that is required for it to slip past another chain—a sort of friction energy—which is taken to be ~ 3.5 kJ/mole, i.e., 1/100th of a typical C–C bond energy.

Actually, U is an energy per molecule and so needs to be divided by Avogadro's number (6.02E23) to make the calculation work. Σ is typically 1/nm². But why 1? Because a close-packed molecular array such as a Langmuir-Blodgett film typically needs 0.5 nm^2/molecule equivalent to $\Sigma = 2/nm^2$, and a typical polymer is not going to be as tightly packed as that. The unit cell of crystalline polyethylene has a cross-sectional area of 0.36 nm^2 allowing $\Sigma \sim 3/nm^2$ so that really is a sort of upper limit. Using the de Gennes formula adhesion values in the region of 1 J/m² are attained.

For stronger adhesion, let the polymer chains intermingle more efficiently. In particular, ensure that the chains are long enough that they can become entangled. What happens when the bond is challenged? It can no longer slide out like a nail and, more importantly, the load is spread over many other polymer chains as each is entangled with the others. Just as a sharp tug on a tangled ball of string will make the tangle even worse, the faster the bond is pulled apart, the stronger the resistance and the higher the work of adhesion.

The key to strong practical adhesion is entanglement, so how is entanglement ensured? Following the approach of Wool, the answer is surprisingly simple: ensure that the polymer chains are long and flexible enough to be able to cross the interface at least three times [17]. Figure 4.20 shows why.

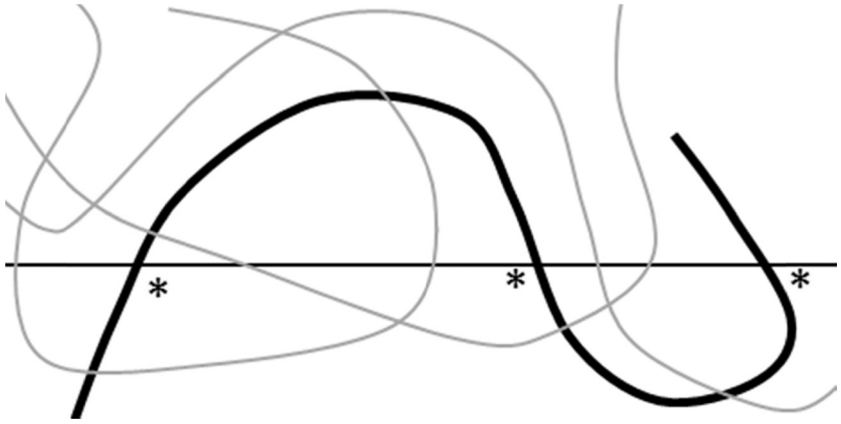


FIGURE 4.20. *The definition of entanglement: a polymer chain crosses a plane at least three times.*

With just two crossings (marked with a *), plucking a chain on one side allows the two dangling strands to be drawn out as with nail adhesion. With three crossings, pulling on a strand causes the chain to trap a different strand which in turn traps another, creating the load-spreading self-reinforcement that makes entanglement so effective.

For a few standard polymers, M_c, the critical entanglement molecular weight, is known. Unfortunately, for most relevant polymers there seems to be no good way to estimate M_c, so it must be measured. Viscosity dependence on molecular weight in the melt or in solution allow such measurements for those who have the time and the equipment. Without such measurements, the only recourse is to assume that "longer is better", up to the limit where "longer" becomes impractical in terms of using the polymer in your coating. Because longer polymers are more easily entangled, their viscosity in the coating solution can increase too rapidly.

Another way to ensure "entanglement" which turns out to be identical in terms of adhesion is to create cross-links between polymer chains—where each cross-link is equivalent to an entanglement point. It is interesting that the adhesion from cross-link entanglement is not significantly different from real entanglement. For those who might be skeptical about the equivalence of entanglements and cross-links it is interesting to note that the first two chapters of Magomedov's book *Structure and Properties of Crosslinked Polymers* discuss the cluster and fractal science of cross-links entirely via standard thermoplastics such as polystyrene or polyethylene [18].The important factor is en-

tanglement density, irrespective of whether that is chemical or physical. In other words, chemical cross-linking does not magically guarantee super-strong adhesion.

The work of adhesion in this entangled case is given by the formula ascribed to Lake & Thomas [19]:

$$G = \Sigma NU \tag{4.9}$$

Yes, the formula is deliberately written in the same form as the de Gennes formula and has a similar meaning. Σ is our Link Density (number of entanglements or number of cross-links per nm^2), N is the number of bonds along the chain between links and U is the Polymer bond strength of (typically) 350 kJ/mole (the standard value for a C–C or C–O bond), again scaled by Avogadro's number. What Lake & Thomas found is that you have to stretch every bond between a link to its limit before one of them will break, so the adhesion strength is amplified greatly. This is probably the single most important fact in practical adhesion science. Even though Σ is smaller than in the de Gennes case (0.2–0.5 rather than 1), U is 100× larger and N can easily be 10 or more for a well-entangled polymer. So G can easily be 50× larger, giving adhesion in the 100 J/m² range, more than enough for really good adhesion between surfaces.

4.3.3. Implications of Entanglement on Nanoparticle Dispersants

Take two dispersions of nanoparticles, each dispersed using the same type of polymer which happens to be nicely anchored to the particle. Using the HSP test, both are nicely compatible with your polymer matrix. Experiments show that they form excellent dispersions and fine coatings. Now test their toughness. One will be significantly less tough than the other. The difference? A short polymer chain on one will mean only nail adhesion to the matrix and therefore a relatively easy splitting at the dispersant/matrix interface.

Should you always demand long-chain dispersants for your particles? For adhesion purposes, yes. But it is important to remember the "shell" calculations from Chapter 2. A large polymer chain on the outside of the particle will constitute a large percentage of the overall particle. The price of high entanglement adhesion is a reduced load of nanoparticle.

These are the trade-offs when adhesion is produced via intermingling of dispersant and matrix. It seems intuitively obvious that a much bet-

ter approach is to use chemical bonds. Surely these will give all the required adhesion and none of the tradeoffs. It turns out not to be that easy—and entanglement is part of the explanation behind the problems often found with chemical adhesion.

4.3.4. Why Chemical Adhesion Isn't So Simple

A typical van der Waals attraction that provides the adhesion behind surface energy is 1 kJ/mole. A typical chemical bond is 350 kJ/mole. It seems obvious that chemical bonds win every time—that is, until a few questions are asked. The first question is about the surface density of the chemical bonds. The radius of gyration of a medium-sized polymer is 1nm. To make calculations simpler, imagine it as a 2nm square. A bond to the surface will occupy, perhaps, a square of 0.1nm sides, so the density of the chemical bonds is $(0.1/2)^2 = 1/400$. Suddenly the 300× advantage in bond strength looks less compelling.

Crucially, there is no amplification via N in the Lake & Thomas formula. If adhesion depends only on that chemical bond, $N = 1$ and a large amount of adhesion is lost.

The obvious solution to this issue is to increase the functionality of the polymer so there is more than one bond/molecule. Now there are two issues. The first is that as the groups start to react, the polymer motion becomes more constrained so it becomes harder for other groups to react with the surface. The hoped-for large increase in surface density of links might not occur.

The second problem with the increased functionality is to do with entanglement—which is why entanglement had to be described before getting on to chemical adhesion. The analysis and data are due, once again, to the Wool group.

As the polymer chain reacts more and more with the surface, its adhesion to the surface increases. If a typical chain has two chemical attachments, the stresses during the breaking of an adhesive bond can become shared just as in entanglement and cross-linking; in other words, the N value increases beyond 1. At the same time, its chains are more and more associated with the surface *and therefore less and less associated with the polymer chains in the rest of the matrix*. At a critical point, the polymer at the surface is no longer entangled with the rest of the polymer matrix. The result is that what might previously have been an adhesive failure between matrix and particle becomes a cohesive failure between matrix and polymer. When a stress is applied to the system, a

crack can propagate not at the particle/matrix interface but at the particle + polymer/matrix interface.

This effect is shown starkly in a paper by Wool on adhesion of a polymer to an aluminum surface [20]. As the % functionality rises to 1% the peel energy goes from 100 to 250 J/m^2, then at 1.2% it dips to 50 J/m^2—less than the peel with no chemical bond.

As the % functionality increases, the adhesion to the aluminum increases, until the polymer is so attached to the aluminum that it is no longer entangled with the rest of the polymer, so the failure shifts to the polymer-polymer interface.

Many of us in the (nano)coatings industry have found such effects—too much of a good thing seems to be a bad thing—without a coherent way to explain it. Although the entanglement theory does not explain all such effects, it must surely help address many such mysteries.

Going back to simple chemical attachment, there is another question to be asked. Against the forces of a sharp crack propagating, how strong is a chemical bond? The answer seems to be "not very". If the whole of the crack energy is focused along a narrow interface of chemical bonds, breaking those bonds is not so hard. Once again, the contrast with entanglement is instructive—with entanglement, the energy readily gets dissipated over many polymer strands, so there is less snapping of polymer chains and a greater resilience to crack propagation.

What we want is entanglement without too many chemical bonds and without too long a polymer chain. The above logic makes this seem impossible. It is here that the equivalence of entanglement and cross-linking becomes vital. If the end functionality of each dispersant can bond to the surrounding polymer (while the other end is securely bound to the particle), then although each individual bond is not especially helpful, the system has constructed the equivalent of a cross-linked network—with the particle at the center like a gigantic pseudo-dendrimer. It seems reasonable to suggest that this is the secret of a tough, crack-resistant, particle-filled system.

Although not directly relevant to the nanocoating theme, it is helpful to think how effective many pressure sensitive adhesives (PSAs) can be. Here the fracture energy is absorbed not by strength but by weakness. The PSA undergoes plastic deformation and even cavitation, which absorbs the crack energy very effectively. It can quickly flow back to recover from the attempt to separate the two surfaces. In fact, simple theory shows that up to a fairly generous limit, the weaker the PSA (e.g. in shear tests) the stronger it is in peel tests!

As we shall see in Chapter 8, such PSA-style resistance is not helpful for hardcoats. It is not clear under what circumstances it might be helpful, but the information is included on the grounds that someone will be able to do something clever with it.

4.3.5. Summary of Adhesion of Nanoparticles within Matrices

This has been a very long but necessary diversion into adhesion theory. Many nanoformulations provide disappointing properties because, although the basic problems of compatibility and anti-flocculation have been solved via solubility thinking, there still is not sufficient intermingling to give good nail adhesion—i.e., in the de Gennes formula, N is too small. But even if N is large, this sort of adhesion is still relatively poor; only strong entanglement and/or cross-linking and/or chemical adhesion with multiple attachment points can produce the desired strength. Even here, if (relatively) short chains are used, the N multiplier is small. Finally, too much chemical adhesion shifts the adhesion problem from particle-to-matrix to particle + matrix-to-matrix allowing cohesive failure.

Happily, these adhesion concepts are simple, powerful and largely proven. If applied consistently to nanoparticles there would be a steep increase in the impact of nanoparticles within matrices because the particle-to-matrix adhesion would be 10–1000× greater.

Strong adhesion of the particles to the matrix may also be regarded as an extra level of protection in terms of nanosafety. Strong adhesion makes it less likely that contact with, handling and disposal of the coating will allow significant release of the primary nanoparticles. Some experimental evidence to support this view is discussed in Chapter 9.

4.4. CONTROLLED NON-COMPATIBILITY

There is a beautiful example where non-compatibility between nanoparticles and matrix is a requirement. For OPV (Organic Photo-Voltaics) that use the classic combination of a photoactive polymer and a fullerene, it is vital that the fullerene exists in a separate phase.

In the lab it has been relatively easy to obtain this phase separation. The two phases are thermodynamically mismatched and warming a sample to a modest temperature such as 60°C for an hour is sufficient for the phases to separate and the device to become an efficient solar cell. In production it is not practical to have such long annealing times,

given that one of the advantages claimed for OPV is that, in principle, they can be manufactured by the kilometer on roll-to-roll equipment.

The ideal would be for the phase separation to happen while the solvent is evaporating. If, however, a single solvent has been used which happens to be good for both the fullerene and the polymer, there is no reason for the phase separation to take place.

A rational way to deal with this issue is to use a solvent blend that is adequate for the two components, with a volatile component which, when it evaporates, reduces the solubility of the fullerene while keeping the polymer soluble. This will encourage phase separation. The way to achieve this is clear: measure the HSP of the two components then find solvents that behave in the desired manner. Two groups have achieved at least part of this approach.

The Nguyen team at UCSB determined the HSP of two of their key molecules. For DPP(TBFu)$_2$, an electron donor, they were [19.3, 4.8, 6.3] and for the fullerene PC71BM, an electron acceptor, they were [20.2, 5.4, 4.5] [21]. Based on these values the group was able to find an optimum solvent blend that delivered increased power conversion efficiencies.

The Brabec group at the University of Erlangen were similarly able to characterize some of their key components: two polymers, P3HT at [18.5, 5.3, 5.3] and PCPDTBT at [19.6, 3.6, 8.8] and a fullerene PCBM at [20.4, 3.5, 7.2] [22]. Recognizing some difficulties and uncertainties in the values obtained by the standard sphere technique, the team developed the "grid" method and re-measured P3HT at [18.5, 4.6, 1.4] and PCBM at [19.7, 5.2, 6.0] [8]. Armed with these values the team was able to formulate solar cells with a more practical (in terms of SHE and cost issues) solvent blend while at the same time achieving the controlled phase separation without the further annealing step [23].

4.5. CONCLUSION

DLVO theory quickly reaches its limits, and even with extensions to the theory such as the ideas of bridging and depletion flocculation, there isn't a practical toolkit for the formulator to use in day-to-day work. By using a robust solubility theory such as HSP, it turns out to be possible to formulate nanoparticles for improved dispersion in solvent and for improved dispersion and adhesion within a polymer matrix. The solubility ideas feed naturally into a practical theory of adhesion where the aim is to achieve as much entanglement as possible, via mutual compat-

ibility and a cross-linked network. Many of these ideas will be put to good use in Chapter 6, where the production team is faced with difficult choices about the best solvents to use for coating and printing.

4.6. REFERENCES

1. Charles M. Hansen. (2007) *Hansen Solubility Parameters. A User's Handbook (2nd Edition).* Boca Raton, FL: CRC Press.
2. Andreas Klamt. (2005) *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design.* Amsterdam: Elsevier.
3. Simon Detriche. (2010) *Chimie de surface des nanotubes de carbone, vers une meilleure dispersion et solubilisation.* PhD Thesis, Facultés Universitaires Notre-Dame de la Paix, Namur.
4. Abhishek Agrawal, Amit D. Saran, *et al.* (2004) Constrained nonlinear optimization for solubility parameters of poly(lacticacid) and poly(glycolic acid)—validation and comparison. *Polymer* 45: 8603–8612.
5. Steven Abbott. (2010) Chemical compatibility of poly(lactic acid): A practical framework using Hansen Solubility Parameters. *In: Poly(Lactic Acid): Synthesis, Structures, Properties, Processing and Applications.* Rafael Auras, Loong-Tak Lim *et al.* 2010. New York: John Wiley and Sons, pp. 83–95.
6. Steven Abbott, Charles M. Hansen and Hiroshi Yamamoto. *Hansen Solubility Parameters in Practice: Software, Datasets, eBook.* www.hansen-solubility.com (Accessed Augst 2012).
7. Arnaud Lafaurie, Nathalie Azema, *et al.* (2009) Stability parameters for mineral suspensions: Improving the dispersion of fillers in thermoplastics. *Powder Technology* 192: 92–98.
8. Florian Machui, Stefan Langner, *et al.* (2012) Determination of the P3HT:PCBM solubility parameters via a binary solvent gradient method: Impact of solubility on the photovoltaic performance. *Solar Energy Materials & Solar Cells* 100:138–146.
9. Zhane D. Bergin, Zhenyu Sun, *et al.* (2009) Multicomponent solubility parameters for single-walled carbon nanotube solvent mixtures. *ACS Nano.* 3: 2340–2350.
10. Michael A. Schreuder, James R. McBride, *et al.* (2009) Control of surface state emission via phosphonic acid modulation in ultrasmall CdSe nanocrystals: The role of ligand electronegativity. *Journal of Physical Chemistry C* 113:8169–8176.
11. Michael A. Schreuder, Jonathan D. Gosnell, *et al.* (2008) Encapsulated white-light CdSe nanocrystals as nanophosphors for solid-state lighting. *Journal of Materials Chemistry* 18:970–975.
12. Kellar Autumn, Metin Sitti, *et al.* (2002) Evidence for van der Waals adhesion in gecko setae. *PNAS* 99:12252–12256.
13. Richard P. Wool, David M. Bailey and Anthony D. Friend. (1996) The nail solution: adhesion at interfaces. *Journal of Adhesion Science and Technology* 10:305–325.
14. Nobuo Maeda, Nianhuan Chen, *et al.* (2002) Adhesion and friction mechanisms of polymer-on-polymer surfaces. *Science* 297:379–382.
15. Betul Buehler. (2006) *Molecular Adhesion And Friction At Elastomer/Polymer Interfaces.* PhD Thesis. University of Akron.
16. M. Deruelle, L. Léger, and M. Tirrell. (1995) Adhesion at the solid-elastomer interface: Influence of the interfacial chains. *Macromolecules* 28:7419–7428.
17. Richard P.Wool. (2006) Adhesion at polymer–polymer interfaces: A rigidity percolation approach. *Comptes Rendus Chimie* 9:25–44.
18. Gasan M. Magomedov, Georgii V. Kozlov and Gennady E. Zaikov. (2011) *Structure and Properties of Crosslinked Polymers.* Smithers Rapra Technology.
19. G.J. Lake and A.G. Thomas. (1967) The strength of highly elastic materials. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 300:108–119.

20. Richard P. Wool and Shana P. Bunkera. (2007) Polymer-solid interface connectivity and adhesion: Design of a bio-based pressure sensitive adhesive. *The Journal of Adhesion* 83:907–926.
21. Bright Walker, Arnold Tamayo, *et al.* (2011) A systematic approach to solvent selection based on cohesive energy densities in a molecular bulk heterojunction system. *Advanced Energy Matererials* 1:221–229.
22. Florian Machui, Steven Abbott, *et al.* (2011) Determination of solubility parameters for organic semiconductor formulations. *Macromolecular Chemistry and Physics* 212:2159–2165.
23. Florian Machui, private communication.

CHAPTER 5

Finding the Perfect Solvent

FINDING a perfect solvent for a coating is no easy task. Having no volatile solvent at all—as in UV coatings—is desirable in many ways, yet is highly restrictive. Water-based coatings have many advantages in terms of SHE, solvent cost and overall green image, yet it takes a lot of energy to evaporate the water and many nanoparticle systems cannot be made compatible with water.

If real solvents have to be used, then—as happens all too often—the one solvent that is low cost and effective becomes banned for some good SHE reason. There is no steady stream of new solvents to replace the old ones, partly because it can cost $1 million to perform all of the necessary safety and environmental testing. It is therefore wise to accept that there will generally be no single solvent that can combine solubility with all the other desirable properties such as cost, odor, environmental compliance and so forth.

This means that the team has to find solvent blends that are acceptable both to the researchers and to the production team. Finding blends is difficult without a rational process: some solvent blends are so unusual that it is highly unlikely that they would be found by either intuition or chance.

This chapter uses Hansen Solubility Parameters (HSP) as a way to find rational solvent blends [1]. It is not the only way. Other solubility tools such as COSMOtherm provide powerful methodologies [2]. HSP are chosen for several reasons: they seem to strike a balance between simplicity and rigor, and because they have a long and successful history of use in precisely this way. The authors know much about them, having used the technique successfully for many years for (nano) coatings.

As the basics of HSP have been described in Chapter 4, this chapter provides some essential solubility theory, including recent developments in the theory of nanoparticle solubility.

5.1. THE MINIMUM NECESSARY SOLUBILITY THEORY

Most formulators, including the authors, don't find thermodynamic theory to be much fun. The aim of this section is to provide the absolute minimum theory that provides the absolute maximum of benefits to the formulator. To put it another way, the authors wish they had known this theory years ago, as they would have been able to tackle many tricky issues starting from a more solid basis. The section starts with a question: *What is the ideal solvent for a system?* Most of us would answer "one that is good, cheap, safe . . ." In this section the word "ideal" will be introduced and used in a strict scientific sense because it is such an important starting point for any sensible discussion of solubility.

By definition, an *ideal solvent* is one where the solute doesn't notice the presence of the solvent and vice versa. To put this into thermodynamic language, the mutual activity coefficients at all concentrations are 1. This definition of "ideal" is important, as it introduces the notion of *ideal solubility*, a concept that should be a routine part of any discussion of solubility, yet one that seems (from the authors' own experience) to be virtually unknown. The authors are highly embarrassed to admit that they spent most of their working lives ignorant of the concept.

There are some examples of solvents that are better than ideal—where specific solvent-solute interactions positively encourage solubility. In these cases the activity coefficients are less than one, but generally not so much less that it makes a great difference to solubility. Of course, adding an amine solvent to an acid solute can make a huge difference to solubility. Arguably, however, this is simply shifting the definition of solubility from the acid to the acid salt. Cases of "significantly better than ideal" solubility are relatively rare.

The norm is for solvents to be worse than ideal—their activity coefficients are > 1 and often >> 1. The aim of rational solvent development is to bring the activity coefficient nearer to ideal.

The point of raising the issue of ideal solubility is that for crystalline solids it can be calculated from first principles and, even more importantly, can be estimated via a simple formula. This means that if there is a requirement for X% of this solid in a formulation and the ideal solu-

bility at your chosen temperature is less than X% then it is unlikely that it can be dissolved to this level.

A clear negative is always helpful in science. The negative can stop a lot of wasted effort trying to identify the perfect solvent that doesn't exist. The project team has to decide whether 0.5X would be adequate or whether the temperature could be raised or whether the project should shift to a similar molecule with a much higher ideal solubility.

A particularly clear exposition of solubility theory is that of Meekes' group and is the one used here [3]:

$$\ln(x) = \frac{\Delta_F}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) + \frac{\Delta Cp}{R}\left[\frac{T_m}{T} - \ln\left(\frac{T_m}{T}\right) - 1\right] \tag{5.1}$$

The solubility, x, is defined as mole fraction. It depends on Δ_F, the enthalpy of fusion, the gas constant R, the melting point T_m, the desired temperature T, and on ΔCp, the change in heat capacity between the solid and the (virtual) liquid at the given temperature. The formula captures the intuition that solubility is greater for solids with lower MPt, lower enthalpy of fusion and (less intuitively) a large ΔCp.

Although the enthalpy of fusion can be readily measured using a DSC, it is very hard to determine ΔCp, and most people are happy if they just know T_m. Fortunately, Yalkowski has shown that the uncertainties in the full equation are sufficiently large that it is almost as good to use a much simplified equation [4]:

$$\ln(x) = -0.023(T_m - T) \tag{5.2}$$

In other words, solubility has an exponential dependence on melting point. High melting point solids give lower solubilities than lower melting point solids. This has the advantage of being intuitively obvious.

A graphical view of these equations brings out their essence. For those who like straight lines, the van't Hoff plot of ln(x) versus 1/T is the one to use, Figure 5.1.

The same equation can be plotted with linear (x) and T, Figure 5.2.

The Yalkowsky approximation can be plotted against the van't Hoff straight line, Figure 5.3.

And the effects of ΔCp can also be compared to the van't Hoff line, Figure 5.4.

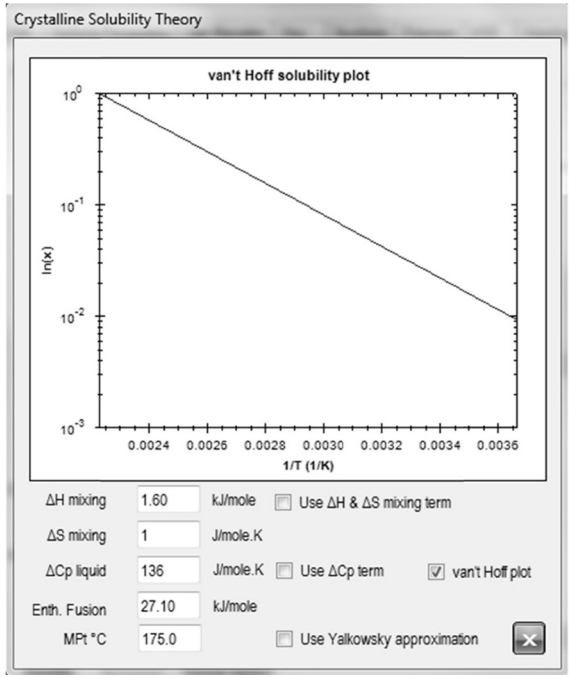


FIGURE 5.1. *The classic van't Hoff plot of solubility v 1/T.*

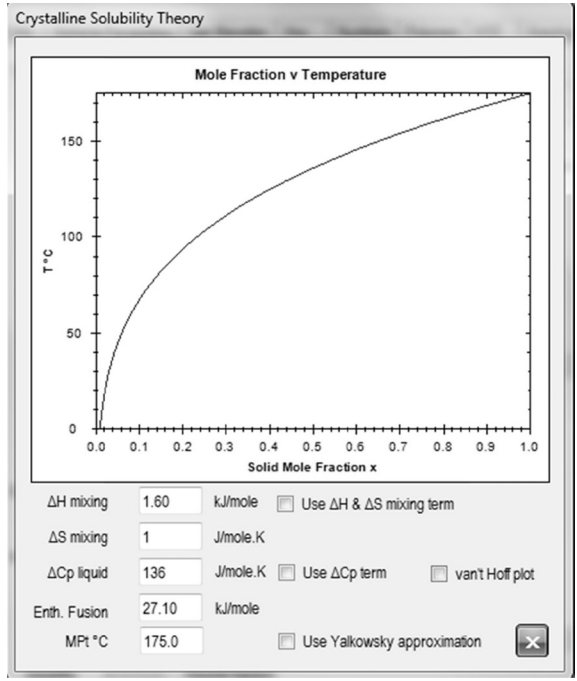


FIGURE 5.2. *Dependence of solubility in mole fraction on temperature.*

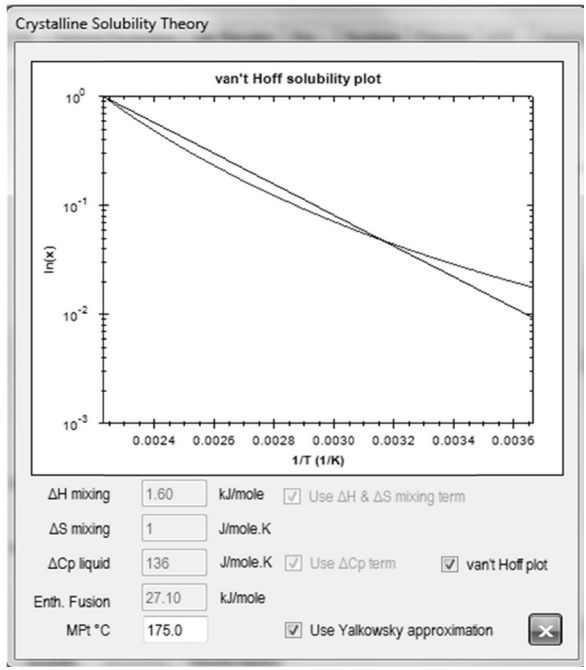


FIGURE 5.3. The Yalkowsky approximation compared to the linear van't Hoff line.

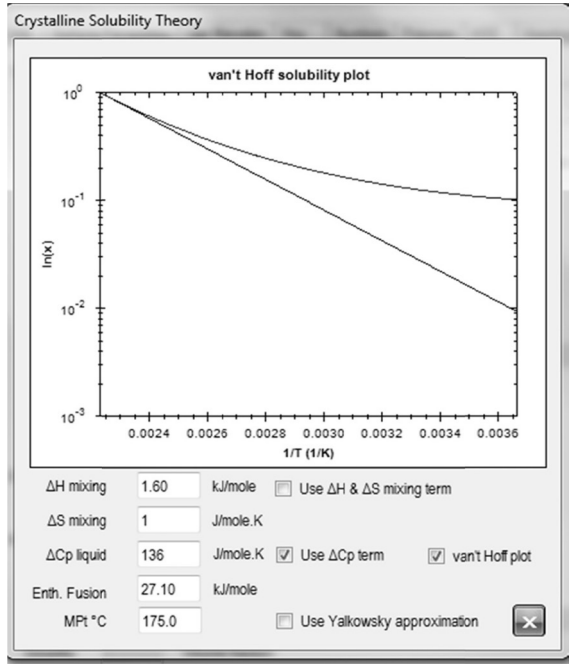


FIGURE 5.4. The large effects of changes of heat capacity ΔCp on solubility compared to the linear van't Hoff line.

What this shows is that when the heat capacity of the liquid form is higher than the solid form (which is a common occurrence) this significantly increases solubility at lower temperatures.

To repeat, if the ideal solubility is not high enough for the application then there is little chance of success—it requires a "super-ideal" solvent which may be hard to find. In general, though, the problem is that most solvents are far from ideal. To extend the equation further:

$$\ln(x) = \frac{\Delta_F}{R}\left(\frac{1}{T_m} - \frac{1}{T}\right) + \frac{\Delta Cp}{R}\left[\frac{T_m}{T} - \ln\left(\frac{T_m}{T}\right) - 1\right] - \frac{\Delta H_{mix} - T\Delta S_{mix}}{R}\frac{(1-x)^2}{T} \tag{5.3}$$

This is the same equation as before with an added non-ideal term, where ΔH_{mix} and $T\Delta S_{mix}$ are the enthalpies and entropies of mixing. For many cases the entropy of mixing is small enough to be irrelevant, so if the enthalpy of mixing is known the real solubility is also known. Unfortunately there is no simple way to calculate ΔH_{mix} for most practical systems; it is therefore put into the form of an activity coefficient and the trick then is to find a method of estimating its value. One well-known method is UNIFAC, but for complex real-world systems it seems that an estimate of activity coefficient based on HSP Distance is a reasonable approximation.

The point of this section is that for real-world teams who have no great interest in solubility theory there is a simple take-home message: solubility depends on melting point, which is easily measured, and activity coefficient, which can be estimated adequately with real-world tools such as HSP. So the solubility of crystalline materials can be thought through rationally rather than by trial and error.

Before getting to the "solubility" of nanoparticles it is important to look at the thermodynamics of polymer solubility, as their behavior is far from simple.

5.2. POLYMER SOLUBILITY

Instead of looking at the solubility, the convention is to plot the change in free energy, ΔG, with respect to the volume fractions of the solvent φ_1 and the polymer φ_2. In other cases we could mostly rely on enthalpic effects, but as mentioned in Chapter 4, the fact that deuterated polyethylene is not soluble in the protonated form demonstrates that entropic terms are significant for polymers. The entropic term depends

on the ratio of molar volumes, x, which can be approximated as the ratio of molecular weights $x = M_1/M_2$, giving an overall dependency of:

$$\Delta G_1 = RT[\ln(\varphi_1) + \varphi_2\left(1 - \frac{1}{x}\right) + \chi\varphi_2^2] \tag{5.4}$$

where R the gas constant and T the temperature in °K.

The χ is the Flory-Huggins chi-parameter which was discussed in Chapter 4. It can be replaced by an equivalent term based on Hansen Solubility Parameter Distance and the molar volume V_1 of the solvent along with RT:

$$\Delta G_1 = RT[\ln(\varphi_1) + \varphi_2\left(1 - \frac{1}{x}\right) + \frac{V_1}{RT}\text{Distance}^2\varphi_2^2] \tag{5.5}$$

For now just think of this replacement term as a measure of "like dissolves like". When Distance is small, this term is small, and given that it is a positive term which increases ΔG, the larger it is the lower the solubility.

The point of these equations is that they allow us to look at what happens as we try to increase the volume fraction of a polymer (i.e., get more into solution) and how it depends on relative molecular weights. At a Distance just below 8 for a typical solvent and polymer (which corresponds to the important value of 0.5 for the χ parameter), the ΔG_1 hovers just below 0—the polymer is not convincingly soluble in the solvent; in polymer science language, the solvent is a theta solvent, Figure 5.5.

Go slightly greater in Distance and the polymer positively dislikes the solvent, Figure 5.6.

Go slightly less and the polymer is happy in the solvent, Figure 5.7.

This is important for formulators who need to dissolve polymers in their solvent formulations.

What is not so obvious, or well-known, is that (as discussed in Chapter 4) this critical behavior is vital for those relying on polymer shells to provide steric stabilization for particles. A relatively small change in solvent can drive a polymer into a deeply unhappy state, curled up on itself and therefore positively welcoming more polymer chains from other particles—causing the particles to crash out of solution. In the state of "theta" balance (when χ = 0.5 or Distance = 8) the polymer

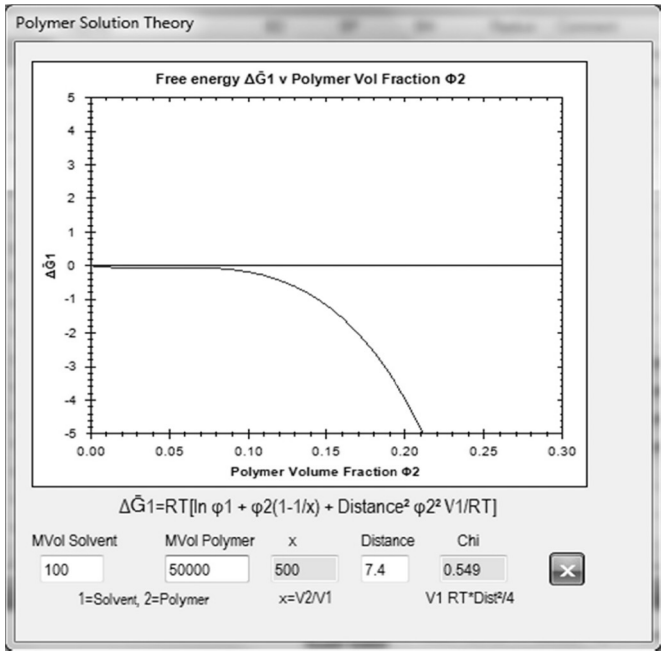


FIGURE 5.5. Free energy of polymer solution for increasing mole fraction of polymer. A borderline situation.

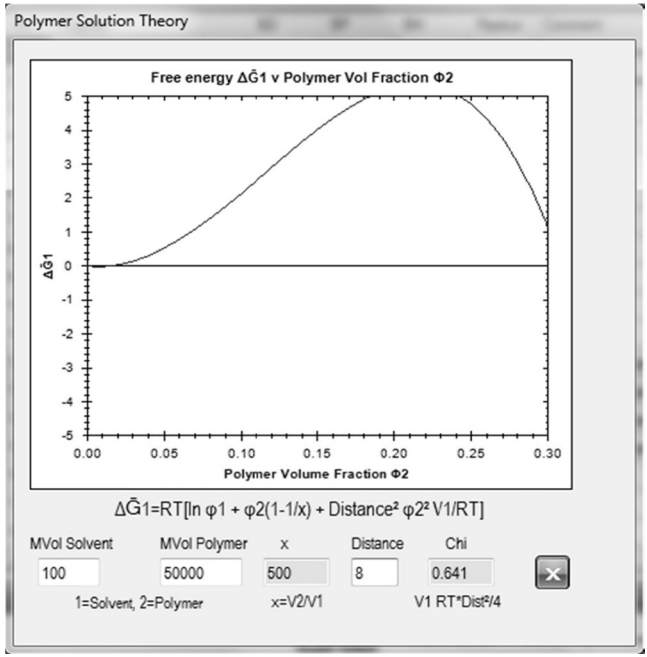


FIGURE 5.6. A slight increase in solvent-polymer distance leads to a large positive ΔG and, therefore, low solubility.

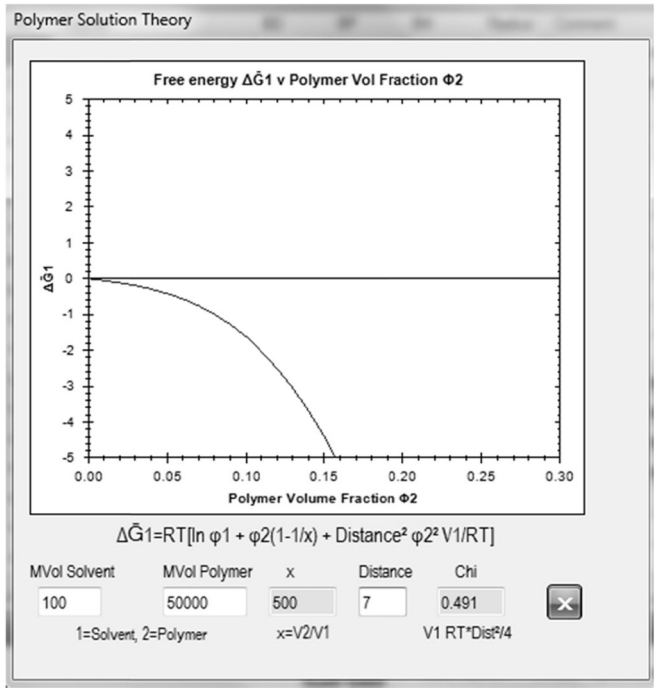


FIGURE 5.7. *A slight reduction in the solvent distance yields good solubility.*

is strictly neutral; then in the state with a smaller distance the chain is happy to extend and steric stabilization can take place. So, for DLVO steric stabilization, a relatively small change in solvent can lead to a catastrophic failure of a dispersion.

That steric stabilization can be unreliable is a well-known fact. That the flip between states can depend on a tiny thermodynamic balance derived from a rather simple equation is much less well-known.

5.3. DISSOLVING NANOPARTICLES

The *next* section is on the contentious subject of nanoparticle solubility. *This* section is a minor diversion into nanoparticles dissolving in the sense that the molecules making up the particle disappear into the solution and the particle gets smaller and smaller.

The idea is that surface energy forces increase as $1/r$, so for very small particles there is an extra driving force for the molecules in the particle to dissolve. This is basically the Gibbs-Thomson effect and is usually described in terms of the Ostwald-Freundlich equation. The

mention of Ostwald links this to *Ostwald ripening*—that big particles get bigger at the expense of smaller particles. Naïve users of Ostwald-Freundlich hope that it can make insoluble pharmaceuticals nicely soluble by making them small enough. The problem (other, as we shall see, than the fact that the effects are usually small) is that Ostwald ripening often overcomes solubilization, so all that effort to make super-small particles ends up by making larger, ripened, particles by the time the pharmaceutical reaches the patient. The Ostwald spreadsheet contains both the Ostwald-Freundlich equation and the Ostwald ripening equation to allow users to benefit (hopefully) from the first without suffering from the second.

Ostwald-Freundlich:

$$\ln \frac{S}{S_0} = 2 \frac{\gamma}{r} \frac{V_m}{RT} \tag{5.6}$$

Here the solubility S of a particle compared to its bulk solubility S_0 depends on the particle's surface energy γ, its radius r, the molecule's molar volume, V_m, and the usual RT term; thus, large surface energies and small radii lead to larger solubility. The practical effect of this is much debated. For ordinary surface energies, say 60 dyne/cm, the radius has to be very small in order to see any significant solubility effect, as is seen in this log plot from the Ostwald spreadsheet, Figure 5.8.

Only when the radius is < 10 nm does anything interesting happen. For inorganics which may have large surface energies (assuming

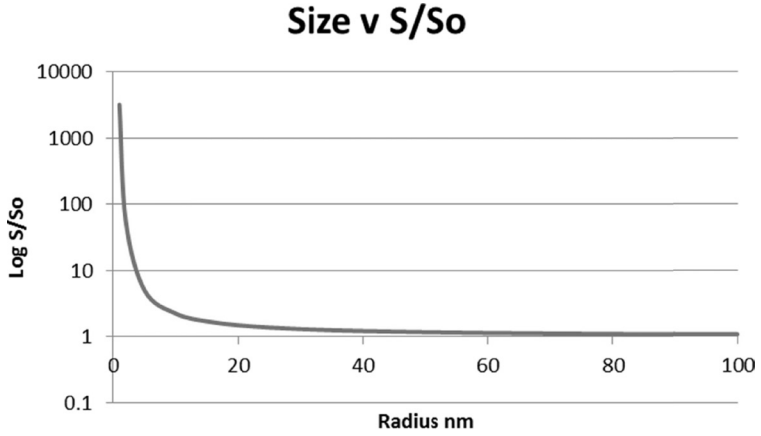


FIGURE 5.8. *The Ostwald-Freundlich dependence of relative solubility on particle radius.*

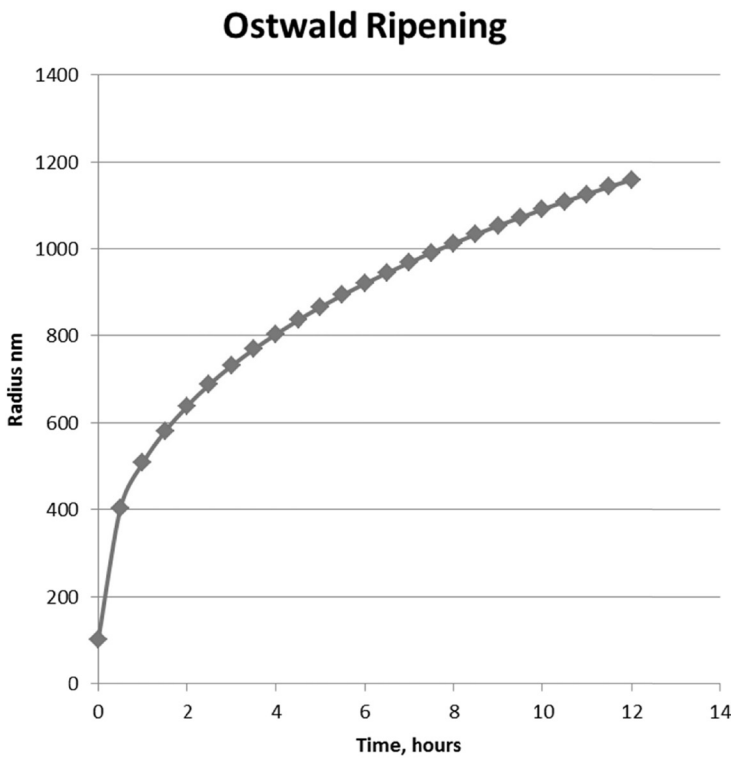


FIGURE 5.9. *Ostwald ripening—the growth of average particle size with time.*

they aren't neutralized by stabilizers), the theoretical effects are large, though a factor of 1000 increase in solubility of a particle with a natural solubility of 10^{-12} may not be very exciting.

Ostwald ripening occurs when small particles spontaneously get smaller and large ones get larger. The effect is given by:

$$\langle r_t \rangle^3 - \langle r_0 \rangle^3 = 8\frac{\gamma DcV_m}{RT}t \tag{5.7}$$

where r_t is the radius at time t, r_0 is the original radius and the brackets mean the average, γ, V_m and RT are defined as above, D is the diffusion coefficient of the molecule through the solution and c is the solubility of the molecule in the solution. With some plausible parameters (see the spreadsheet), a typical ripening scenario is found, Figure 5.9.

Purists will note that the behavior around $t = 0$ is not well-described by the simplified formula, which is used for illustrative purposes only.

The summary of this short section is that hopes of using Ostwald-Freundlich to actually dissolve the molecules of an insoluble nanoparticle are largely illusory. The effects are small and the chances of Ostwald ripening are large. For those who need ultra-small particles and who aren't aware of these effects, it is a warning that carefully prepared particles might disappear through dissolution or, worse, grow to large sizes for no apparent reason.

Having completed this diversion into nanoparticles that erode by dissolution it is time to return to the main theme of solubility of the entity as a whole.

5.4. NANOPARTICLE SOLUBILITY

Some would object to the title of this section. The idea that a nanoparticle could be "soluble" seems laughable. Surely nanoparticles are merely "dispersed". A glance at the literature suggests that most nanoparticle authors are more comfortable using the word "dispersed", even those where "solubility" is in the title of the paper.

A thought experiment shows that matters aren't so clear-cut. Everyone agrees that DNA is a polymer for which the idea of "solubility" is perfectly meaningful. It is possible to take a DNA strand and stretch it to lengths of hundreds of nm. It remains, of course, soluble. Now, what is the difference in principle between this long strand of DNA, with its highly structured double helix, and a CNT of similar dimensions? The fact that the CNT is considerably more rigid doesn't provide any obvious reason why it should not be defined as soluble.

To take another analogy, typical ribosomes have a radius of 10–15 nm and are highly structured. People also talk casually about ribosome solutions, presumably because they are made of biopolymers and no one questions that polymers are soluble. So why should a 10 nm nanoparticle not be called soluble?

When Coleman and co-workers explored the behavior of CNT within solvents that varied from bad to good, they could find no better word than "solubility" to describe what they saw in the best solvents [5].

Finally, when a nanoparticle is stabilized by a floppy shell of a polymer which on its own is soluble, is the ensemble "soluble" or not?

One possible distinction between "particle" and "solute" is behavior under gravity. Materials in solution don't fall to the bottom of the test tube; particles do. Except that, as discussed in Chapter 2, small particles don't—they stay suspended indefinitely via, at the very least, Brown-

ian motion. Artificial gravity can resolve that issue: spin the particles in a centrifuge and the high G forces bring them to the bottom of the tube. However, it is commonplace in biology to use an ultracentrifuge to bring large molecules such as DNA out of solution and onto the bottom of the tube.

The authors' reason for being relaxed about the concept of nanoparticles being soluble is that applying solubility thinking seems to deliver many practical formulation benefits. There is no technical doubt that "like disperses like" because exactly the technique that can pinpoint the solubility properties of a polymer (along with very "solubility related" ideas such as the χ parameter) can be used to get the solubility properties of nanoparticles such as CNT, quantum dots, ink pigments, carbon black, surface-treated silicas, etc. Whether properties are "solubility" or "dispersibility" related is less important than the fact that the measured values allow formulators to use techniques that work well for classic solubility issues.

As discussed in the Case Study of quantum dots in Chapter 4, their solubility or dispersibility cannot be explained by DLVO because the "stabilizers" are far too small to be relevant in terms of steric stabilization.

By allowing such solubility thinking, the benefits aren't merely pragmatic, they lead to specific predictions. In the CNT case study of Chapter 4, it was shown that large CNT are less soluble than small CNT for (probably) exactly the same entropic reasons that large polymers are less soluble than small ones (a high "*x*" value in the polymer solubility equation).

Two recent papers by the Coleman group put this solubility thinking on firmer ground [6, 7]. They analyze 0D (normal solutes), 1D (polymers and CNT) and 2D (graphene, etc.) solubility via classic "lattice theory" and conclude that the standard solubility parameter approach applies equally to 1D and 2D materials, albeit with factors of 2/3 and 1/3 in front of the standard Distance formula. Crucially, their approach makes it clear that the rigidity of the nanoparticles in no way affects the theory compared to the flexibility of polymers. Indeed, their approach, arguably for the first time, shows that solubility parameters may legitimately be applied to polymers (i.e., 1D molecules)—something that had hitherto only been assumed, albeit with considerable success.

This approach also looks at the oft-neglected entropic terms, reaching the conclusion that almost any solvent orientation at the surface will be catastrophic in free energy terms because the effect is multiplied by

the ratio of molar volumes of the nanoparticle (huge) and the solvent (small). This is a work in progress as some of their data can be fitted only if the effective size of the nanoparticle is significantly smaller than its actual size. The point, however, is that their work is showing the possibility of a respectable theory of solubility for nanoparticles which will give insights into not only approaches such as solubility parameters, but also the effects of particle size on solubility.

5.5. AQUEOUS SOLUBILITY

In theory, water is just another solvent, so everything discussed above applies to water. Yet we all know that water behaves unusually. The problem, to use a term that arouses strong passions, is the "hydrophobic effect". In the absence of strong hydrogen-bonding groups such as alcohols or amines, the solubility of a chemical in water is strongly controlled by its molar volume—the larger the volume, the less the solubility. Whether this effect is enthalpic or entropic, clustering or non-clustering, the effect is real. The single simplest way, for example, to estimate Log(*P*), the octanol/water partition coefficient is via molar volume.

The other peculiarity of water is its very low molar volume. It is generally little known that octanol in the octanol/water partition coefficient experiment contains 27.5% water when expressed as mole percent. In weight percent it is a more intuitive 5%. A very small amount of water can provide a surprisingly large mole percentage. The small molar volume affects entropic effects in solubility theory and also allows water to diffuse quickly into coatings, even when its solubility in the coating is low.

In terms of almost everything discussed in this chapter the rule with water is "if it works it works if it doesn't it doesn't". This is, admittedly, most unhelpful, even if it is, in practice, mostly true. An attempt by academic groups with considerable intellectual fire power to predict the aqueous solubility of a collection of chemicals resulted in a rueful conclusion that considerable effort had resulted in remarkably little progress [8]. Water is a very difficult solvent to understand.

The main route connecting water to other solvents is via mixes with alcohols. At high-alcohol ratios, the water acts like a rational co-solvent; at low-alcohol ratios, water is dominant. Unfortunately the transition isn't linear. This is shown, for example, in a plot of surface tension versus water content, which is highly non-linear because the lower surface tension alcohol is preferentially concentrated at the

surface, leading to a lower value than expected from a linear blending rule, Figure 5.10.

This plot is a reminder that one of the key reasons for adding alcohol to a water-based coating is to decrease the surface tension to allow wetting of the substrate. The fact that the relationship is non-linear in the right direction means that relatively small amounts of ethanol lead to significant reductions in surface tension.

The alternative way to reduce surface tension is via an added surfactant. This can lead to disappointment if the surfactant is not "fast" acting. Some very efficient surfactants (with a low CMC, Critical Micelle Concentration) are also large molecules which can take a long time to reach the surface. Even for typical modest-sized surfactants above their CMC, the time to reach a low surface tension is in the order of 100 ms [9]. So although in static tests they reduce the surface tension dramatically, at the millisecond time-scale that matters during a coating process (as the liquid meets the substrate) the surfactant might migrate far too slowly to the fresh surface to make much difference. Alcohols have no such problems, as they are present in relatively large concentrations and are small molecules with large diffusion coefficients. Alcohols also disappear during the drying process whereas surfactants can remain on the surface and create downstream problems.

It is obvious to any sensible formulator that a few percentages of i-propanol in an aqueous coating will not be flammable in any meaningful way; thus, using the alcohol to reduce surface tension or, perhaps,

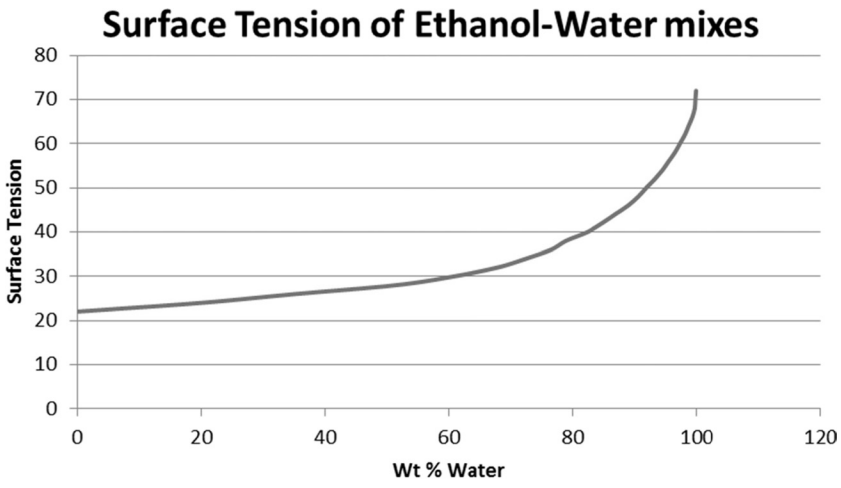


FIGURE 5.10. *The (non-linear) dependence of surface tension of ethanol-water mixes.*

solubilize something that is marginally water-soluble is sensible while preserving the safety advantages of an aqueous coating. Sadly, what is obvious happens to be wrong. The flash-point of i-propanol is ~12°C; that of 20:80 i-propanol:water is 29°C while even 10% i-Propanol has a flash point of 41°C. This makes the 10% solution (depending on local regulations) "flammable" and, therefore, subject to all the rigors of coating from traditional solvents.

Returning to real surfactants, it is worth repeating the warning in the discussion of zeta potential and DLVO: small amounts of the wrong surfactant can cause catastrophic failure of a nanodispersion. Adding some sodium lauryl sulfate to a particle with a positive zeta potential or some cetyl tetrammonium bromide to a particle with a negative zeta potential can cause instant crashing out of the particles.

5.6. MEASURING THE HSP OF A NANOPARTICLE

The basics of HSP have already been explained in Chapter 4. To summarize, each solvent, polymer and nanoparticle can be described in terms of 3 solubility parameters: δD, δP and δH, representing the Dispersion, Polar and Hydrogen bonding components. Because "like dissolves like", the way to gain solubility/compatibility is to reduce the Distance between, say, solvent and nanoparticle; given two materials with HSP [δD_1, $\delta P1$, δH_1] and [δD_2, δP_2, δH_2] the HSP Distance is given by:

$$Distance = \sqrt{4(\delta D_1 - \delta D_2)^2 + (\delta P_1 - \delta P_2)^2 + (\delta H_1 - \delta H_2)^2} \qquad (5.8)$$

Because the distance formula creates a sphere in HSP space (without the factor of 4 in front of the δD the shape is an ellipse and, more importantly, data fits are worse) it becomes possible to define the HSP of a polymer or nanoparticle via a sphere, the center of which represents a perfect matching solvent (Distance=0). The radius of the sphere defines the difference between "good" and "bad" or "happy" and "unhappy" solvents. Examples in Chapter 4 of sphere fits of a classic polymer (PLA), a classic "naked" nanoparticle (CNT) and a set of "stabilized" quantum dots (CdSe) show the versatility of this approach.

The HSP of many chemicals, solvents and common polymers are known (or can readily be estimated), so that the Distance formula can be applied across a wide range of chemicals, solvents, polymers and nanoparticles to maximize compatibility/solubility.

5.7. FINDING THE PERFECT SOLVENT

Given a target such as a polymer or nanoparticle, a table of distances between target and a range of solvents makes it easy to select those solvents that are close and have reasonable properties such as cost, safety or RER (Relative Evaporation Rate). Using the HSPiP software, such a table is readily prepared. As a specific example, suppose that the "Target" values for a nanoparticle are [17, 7, 7]. Using a table of reasonably common solvents and sorting them by distance from the target the following is observed, Figure 5.11.

The practical formulator will spot that, although methylene chloride is a near-perfect match and will give great solubility, chlorinated solvents are generally seen as unacceptable. Dimethyl Isosorbide (DMI) is in the list as it is a relatively new "green" solvent. Its high BPt, low RER and, perhaps, high cost will exclude it from common formulations. THF is next—a really useful solvent rendered generally useless by safety concerns. It is interesting to note that DMI contains a sort of double-THF structure which helps explain its similarity in HSP values. The Dibasic Esters (DBE) are relatively cheap and benign solvents often used for cleaning (they would be ideal to use after a coating run) but are generally too involatile to be the main solvent for a coating applica-

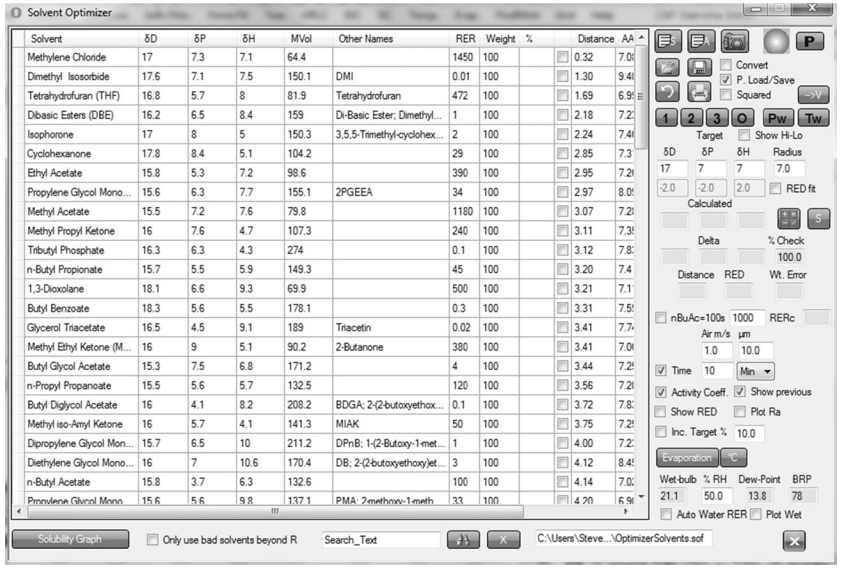


FIGURE 5.11. *Typical solvents listed by their HSP distance from the Target.*

tion. Isophorone is much used in printing, so it is at least a possibility as is cyclohexanone, though both are rather pungent. Butyl acetate has an odor of bananas, which is pleasant at first but often deemed unacceptable on a coating line that has to handle it day-in and day-out (not to mention the factory's neighbors!). The PGEEA (Propylene Glycol Ether Ethyl Acetate) is the first solvent in this list which has a reasonable profile. Maybe it is the right balance of cost, safety etc. and can be used on its own. But suppose it is not quite right for some reason. Does one go down the list getting further and further away (certainly rejecting Tributyl Phosphate!) or is there another approach?

5.8. RATIONAL SOLVENT BLENDS

As discussed at the start, and as the example above shows, in general it turns out to be impossible to find the perfect single solvent, so it is necessary to find a blend that will do the job. Finding this via trial and error is a thankless task, so why not employ a rational method?

One more important principle is required. An X:Y blend (always expressed as volume %) of two solvents gives an HSP which is an X:Y average of each individual parameter. To take a specific example, a 50:50 blend of [18, 10, 10] and [16, 4, 4] gives a solvent of [17, 7, 7]. If the two individual components are acceptable (cost, safety, etc.) while being too far from the target to provide good solubility, then a blend which is close to the target is likely to be acceptable.

Let's apply this principle to the example above. Suppose, for example, that there is a very good reason why 50% PMA (Propylene glycol monomethylether acetate) should be used. An automatic check through all the available solvents finds that the single best solvent that can be combined with PMA is cyclohexanone. The combined HSP of the two is closer (0.8) to the target than either PMA (4.2) or cyclohexanone (2.85). The result is a solvent blend with less odor than pure cyclohexanone plus an improved solubility, Figure 5.12.

The consequences of this decision are just as important during the drying process. Using a simplified drying model (a more complex one is available in the TopCoat suite described in Chapter 6), it is possible to follow what happens during drying. Because the two solvents above have similar RER values (33 and 29), nothing much of interest happens during the drying as the composition stays similar to its original. Using, instead, a rather volatile blend can illustrate the sorts of issues that are important during the drying of a solvent blend, Figure 5.13.

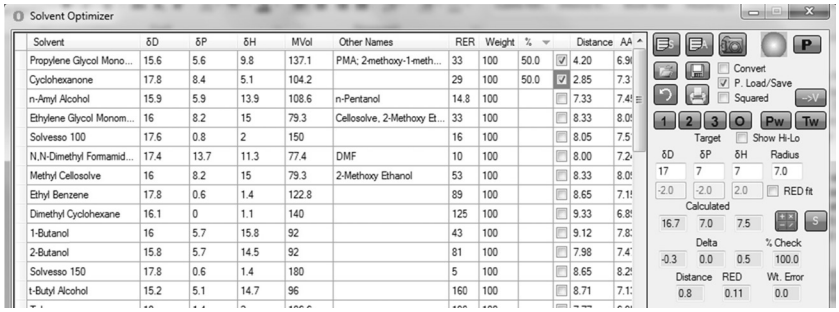


FIGURE 5.12. *A close match to the target using a 50:50 solvent blend.*

The 1,3-Dioxolane (lower curve) is about twice as volatile as the ketone (upper curve). The 50:50 blend starts with a near-perfect match to the target. The distance (shown as Ra in the table) remains remarkably close to 0 for the first third of the evaporation (shown in arbitrary units of 0–30). As the Dioxolane disappears, the blend becomes closer to the ketone, so the Ra approaches 3 towards the end of the drying. This may well be adequate to keep the whole system together during the drying process.

Suppose, however, that our Target was [16, 7.6, 4.7], Figure 5.14.

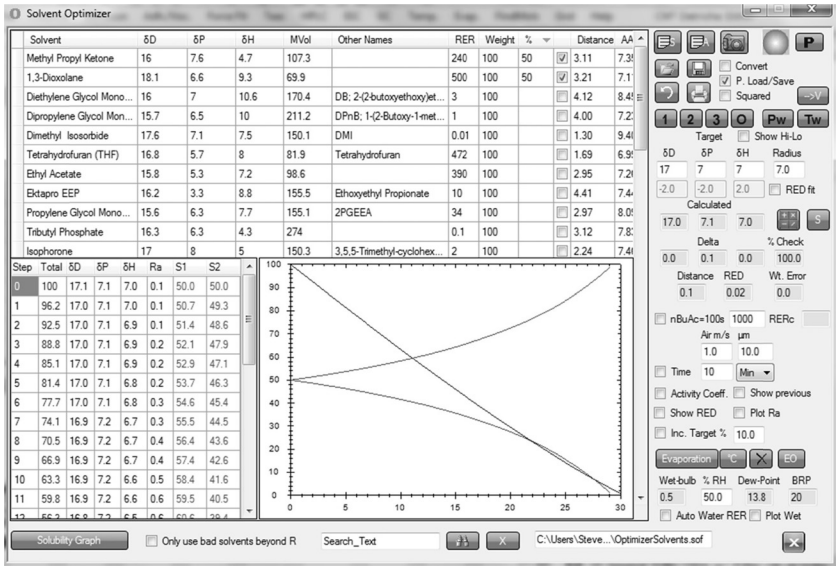


FIGURE 5.13. *HSP distance increases with time as the better solvent evaporates faster.*

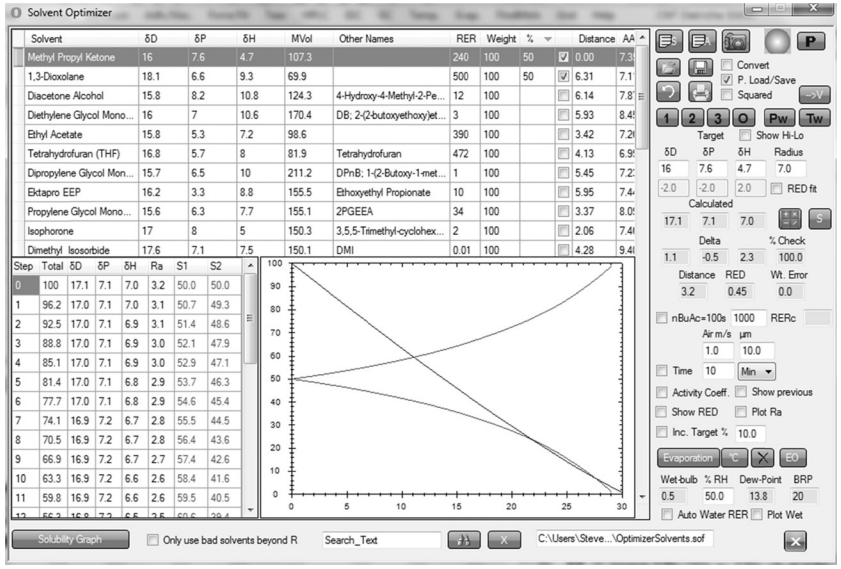


FIGURE 5.14. *Solubility improves with time as the less-compatible solvent evaporates.*

Now the drying starts with a Ra value of 3—perhaps good enough to keep everything in solution when there is plenty of solvent. By the end, the solvent is mostly the ketone which (because this is an artificial example) is a perfect match for the Target so the Ra ~0. This means that as the drying takes place, the key material will have the maximum probability of staying compatible with the solvent. If the material is a polymer rather than a nanoparticle, it will probably dry as a nice, gloss, stress-free coating. Gloss, because it has no reason to crash out and produce an uneven coating. Stress-free because residual stress is proportional to the percentage solvent remaining in the coating once it becomes unable to move and as a good solvent that point will be reached late in the drying process.

There is no need to show another screen-shot of the reverse idea. If the Target is a perfect match for the Dioxolane then what starts as a reasonable solvent blend rapidly develops a large distance from the Target which may well crash out early. If you are making an organic photovoltaic and you require a rapid phase separation of the conducting polymer, then this sort of deliberate crashing-out of a component early in the process is fundamental to success. For most applications, however, this is a situation to be avoided.

The use of the phrase “crash out” is deliberate. The section on poly-

mer solubility shows how there really is a critical point where a polymer (in the formulation or as a nanoparticle stabilizer) can flip from compatible to non-compatible over a small solubility range.

These examples have encapsulated a great deal in a small space. What they say is that it is possible to tune solvent blends rationally, not only to give adequate solubility at the start of the process (something which in itself can be quite difficult) but also because of the ability to tune the relative solubilities of components during the drying so they stay in, or crash out of, the blend at a time and speed of your own choosing.

In Chapter 6 it will be pointed out that in a high-efficiency drier on a typical modern coating/printing machine "high-boilers"—i.e., solvents that are typically seen as being too slow to evaporate—pose little problem. On the lab bench a solvent such as N-methyl pyrrolidone seems impossibly slow to evaporate, yet as Chapter 6 will show, it can disappear far more quickly than, say, water. The key is not so much the boiling point but the enthalpy of vaporization. So it is important for the team to include high-boilers in its discussions. Almost by definition, these solvents show strong interactions with many polymers and nanoparticles (if they didn't, they wouldn't show such strong self-associations) and can be a vital part of any formulation, especially when their evaporation is slow with respect to lighter solvents. This ensures that in the difficult final phase of drying, the formulation can still remain open and mobile—thereby (as will be discussed in Chapter 6) speeding up the overall drying process by not becoming diffusion-limited too early.

Although the above examples are created for illustrative purposes, in the real-world such techniques do work. Because they work so well, it is interesting to ask why they are not used more frequently. The answer seems partly to be that people are generally afraid of solubility theory (it involves thermodynamics). The authors' experience suggests that another part of the answer is that formulators are simply expected to know what to use and therefore don't need to turn to theories or computer models. Overturning well-established or entrenched practice requires a serious battle with inertia. Finally, there is an element of job protection. If Joe the Formulator relies on magic to attain results then job security is assured because the magic would go if Joe went. If anyone can press a few buttons and obtain a superior solvent blend, then the magic is gone. Joe naturally fights against any attempt to demystify the magic. It is an unfortunate fact of life that this is a common reason why there are so many unsatisfactory formulations taken into production.

5.9. TRADE-OFFS

For simplicity, the previous section dealt with just one Target. In a real formulation there will be multiple polymers, chemicals and nanoparticles with conflicting requirements for solubility. These conflicts make the above logic even more compelling for real-world use.

The first question that needs to be asked is which of the components is the most difficult to dissolve to a level that creates an acceptable coating. For example, coating from 99% solvent because a key additive is soluble only up to 1% is deeply unsatisfactory. The most difficult component must then get allotted priority as the Target for the solvent blend. Assuming that the other components are adequately soluble in this blend (or the blend can be tweaked to reach a compromise), the formulation can be further developed to accommodate the needs during drying. Is it really, really important that the polymer stays in solution as long as possible? Then the solvent blend should ensure that the least volatile component is a good match for the polymer.

The 3D nature of HSP makes it important to be able to visualize some of the problems, Figure 5.15.

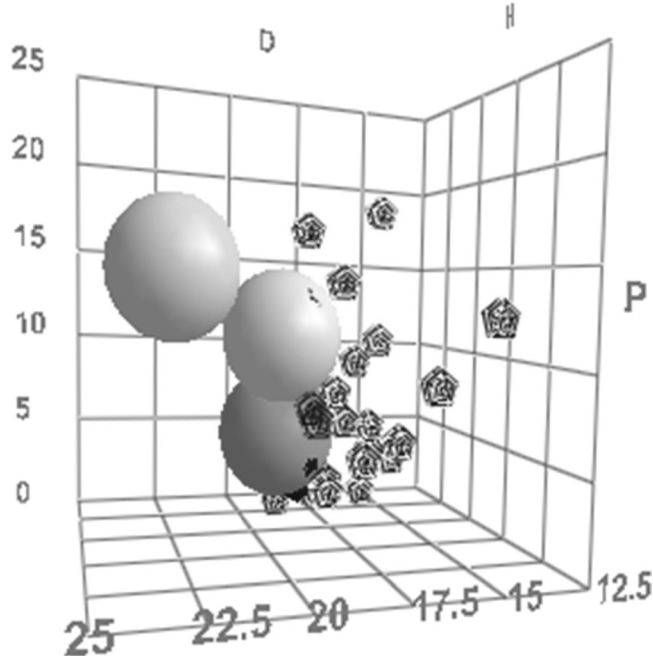


FIGURE 5.15. *Three components in different parts of HSP space pose a challenge.*

In this example, three important components (e.g., a polymer, a chemical and a nanoparticle) occupy three very different parts of HSP space. A single solvent that is ideal for either of the two extremes will be worthless for the other extreme, so a solvent closer to the middle sphere is going to be preferred. Using a mixture of solvents, it is possible to start with a blend that is near the middle sphere which, over time, might move towards either of the two extreme spheres. The team can rationally plan the solvent strategy based on which of the extreme spheres is more important to keep in solution/dispersion during the drying process.

The Production Department must be present during these discussions because they have their own important needs. Even if the lab formulation is scientifically perfect, if a key solvent is too malodorous, has the wrong flash-point, is too expensive or degrades the catalytic burners, then this "perfect" formulation won't make it into production and further work must be carried out.

The examples in the previous section were restricted, for simplicity, to two components. The same logic applies to blends with more solvents so the players have more options.

There is an interesting wild-card to play during such discussions. Everyone tends to focus on solvents that in themselves are typically good. Instinctively the team will reject solvents known to be useless. The trick is to remember that two bad solvents can produce a good solvent. If these solvents are exactly opposite each other in HSP space then a 50:50 blend will be a perfect match to the target. This surprising idea was spotted by Hansen when he was developing his theory. It seemed to him to offer a chance of refuting his own theories. If a theoretical good match from two bad solvents turned out to be bad, then he would have been forced to abandon the theory. When he performed the experiment with an epoxy resin and nitromethane and butanol, each a useless solvent, he obtained a crystal-clear solution. Interestingly it had been known for many years in the paint industry that blends of toluene and isopropanol were capable of dissolving resins that were insoluble in the individual solvents. Until HSP there was no scientific explanation of this phenomenon; it just worked.

So taking the original solvent list and the original target, invoking an option that creates solvent blends only from "bad" solvents, and asking the software to find a blend of 3 such solvents, the following formulation was suggested, Figure 5.16.

It is undoubtedly not a combination that anyone would have come up

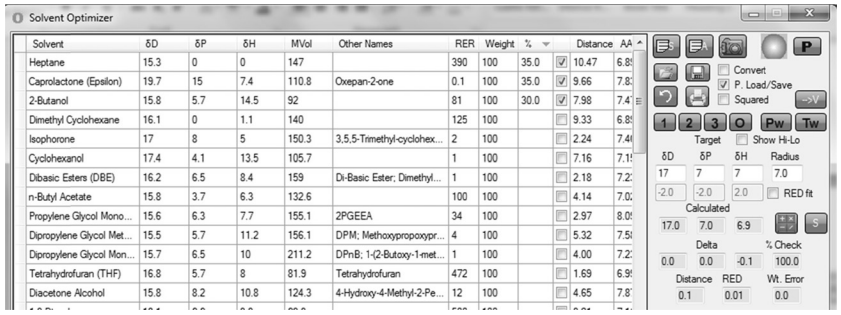


FIGURE 5.16. *A blend of bad solvents that combines to be a close match to the Target.*

with via any normal route and there are reasons (e.g., the excessively low volatility of the caprolactone) why this combination would be unlikely to make it to production. Yet the impetus to fresh thinking might send the team towards a combination of bad solvents that could actually deliver the desired trade-offs. The reason the option to use only bad solvents was added to the software is because a major chemicals company asked for it. When it was provided, they confirmed that it had proven to be of great use in creating fresh ways to approach solubility issues.

5.10. NOT SOLVENTS

Nanocoatings based on UV lacquers generally are "solvent free" or, if they first go through a solvent-coating stage, they then exist for a significant time as liquid UV systems in their own right. It takes just a small shift of mind-set to start calling the UV lacquers "solvents", allowing the above logic to work perfectly in terms of arranging the functionalities to obtain solubility of the nanoparticles, polymers and other chemicals.

Typical acrylates have HSP values in the [15.5, 7, 6] range. If nanoparticles are stabilized by acrylate-containing shells, then the HSP match will be excellent, so dispersion should be straightforward. Typical vinyl ethers are in the [16, 7, 8] range so they require particles in the same HSP space as acrylates. If the nanoparticles are measured to be in, say, the [17.5, 10, 10] range, the HSP distance is around 6, which is starting to be a bit distant. This is closer to typical urethanes. Similar considerations can be given to UV-curable epoxide and oxetane systems.

The logic is straightforward: determine the HSP of the nanoparticles via the standard Sphere technique. Work out the HSP of the UV mono-

mers either by looking them up in tables of known materials or via the automated technique (using SMILES or Molfiles) in HSPiP. Similarly, find the HSP of polymers or other additives such as photoinitiators. Where possible, try to shift the work to your suppliers. It helps reduce your workload, and providing HSP information should be a routine part of the service offered by suppliers. The suppliers themselves will benefit as it helps them to formulate better and to reduce the number of irate calls from customers who have found themselves with a pot of viscous goo when everything crashes out.

It is, in any case, probably a good idea to have a discussion with your nanoparticle supplier about solubility theory, χ parameters, steric stabilization science (DLVO) and HSP. If they already know all there is to know, the chances are high that they will be able to help fast-track your development process. If these things are a mystery to them then you have to decide whether to educate them or seek an alternative supplier.

5.11. CONCLUSION

A little theory goes a long way to help the team through the maze of concerns about solubility. For crystalline additives, ideal solubility theory helps to define whether the required level is theoretically feasible. For polymers, the notion of a rapid flip between stable and unstable is a powerful insight into why systems might crash out when the χ parameter approaches 0.5, and the ability to calculate the χ parameter via HSP is of great benefit. Applying solubility thinking to nanoparticles seems to be a rich source of formulation knowhow, even if one chooses not to buy in to the concept that the particles are actually "soluble". HSP works across crystalline solids, polymers and nanoparticles and can be used to identify single solvents or, more likely, blends of solvents that deliver not only the initial solubility but also control it during drying—either to keep things in solution or (as in the case of organic photovoltaics) to separate out one component early on. No one claims that HSP are a perfect scientific tool. In each case there might be scientifically superior, specific ways to think through solubility issues for those who have the time and energy to devote to them. The authors, for example, love to use COSMOtherm for detailed calculations of solubilities of crystalline solids. Yet for a hard-pressed team having to resolve multiple, conflicting, issues ranging from lab to production, this integrated approach via HSP seems to be the best all-round tool for the job.

5.12. REFERENCES

1. Charles M. Hansen. (2007) *Hansen Solubility Parameters. A User's Handbook* (2nd Edition). Boca Raton: CRC Press.
2. Andreas Klamt. (2005) *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design.* Amsterdam: Elsevier.
3. J.Th.H. van Eupena, R.Westheim, *et al.* (2009) The solubility behaviour and thermodynamic relations of the three forms of Venlafaxine free base. *International Journal of Pharmaceutics* 368:146–153.
4. Samuel H. Yalkowsky and Min Wu. (2010) Estimation of the ideal solubility (Crystal–Liquid Fugacity Ratio) of organic compounds. *Journal of Pharmaceutical Sciences* 99:1100–1106.
5. Shane D. Bergin, Valeria Nicolosi, *et al.* Towards solutions of single-walled carbon nanotubes in common solvents. *Advanced Materials* 20:1876–1881.
6. J. Marguerite Hughes, Damian Aherne and Jonathan N. Coleman. (2012) Generalizing solubility parameter theory to apply to one- and two-dimensional solutes and to incorporate dipolar interactions. *Journal of Applied Polymer Science* DOI: 10.1002/APP.38051.
7. J. Marguerite Hughes, Damian Aherne, *et al.* (2012) Using solution thermodynamics to describe the dispersion of rod-like solutes: application to dispersions of carbon nanotubes in organic solvents. *Nanotechnology* 23.
8. Anton J. Hopfinger, Emilio Xavier Esposito, *et al.* (2009) Findings of the challenge to predict aqueous solubility. *Journal of Chemical Information and Modelling* 49:1–5.
9. Xi Yuan Hua and Milton J. Rosen. (1988) Dynamic surface tension of aqueous surfactant solutions: I. Basic Parameters. *Journal of Colloid and Interface Science* 124:652–659.

CHAPTER 6

Coating, Printing, Drying

THIS chapter contains very little that is specifically related to nano. So why has it been included? The authors' experience is that the basics of high-quality 21st century coating and printing are unknown to many teams who are trying to introduce exciting new nanocoatings. There have been whole conferences devoted to roll-to-roll production of, say, printed electronics where the basic assumption of most talks was that the hard part was the creation of the nanoformulation, following which kilometers of fine product would be produced by "standard coating techniques". This is naivety of heroic proportions.

As a general rule, all coatings are easy to produce, except for the one you happen to be working on. Or—to put it another way—everyone thinks that everyone else is having a great time coating/printing *their* products, while *our* product just seems to have a special feature that makes it a particular challenge.

The truth is that everyone finds their coating to be right at the borderline of achievability. This is because if the coating is particularly easy then everyone thinks that it should go faster or thinner or cheaper, and the boundaries get pushed till the product becomes borderline coatable. This tendency to push is automatic and hardly noticed. Who would produce a product at 5 m/min when it can be produced at 50 m/min? So the speed goes up till, for some reason generally unknown, the process seems to become unmanageable above 45 m/min and everyone settles on 44 m/min as "good enough", though with a sense of shame that it is not 50.

The aim of this chapter is to allow the team to make rational choices of the relatively few coating and printing processes that should be ap-

plied to a 21st century nanoproduct. This aim instantly rules out most old-fashioned coating methods. Those readers who routinely use roll coating techniques may be shocked to find them dismissed out of hand.

Because all coatings have a tendency to form defects, an equally important aspect of this chapter is the understanding of what factors can amplify or deter defects. Sadly, given the trend for coatings to get thinner, the laws of physics are rather clear that thinner coatings are much more prone to defects than thicker ones. The relevant laws are rather simple to understand, which makes it regrettable that they are very little known. Many coating teams have scratched their heads when a minor reduction in coating thickness has resulted in a major increase in coating faults. Understanding why this happens is the first step to coming up with rational strategies to solve the problems.

Throughout this chapter, extensive use will be made of TopCoat [1]. This is commercial software and once again you are alerted to the fact that Abbott is a co-author. If the authors knew of a better way to simulate these coating methods then they would be using it. For those who have access to state-of-the-art computational fluid dynamics (CFD) software, a powerful computational platform and the time to get to know how to use the code, it might be possible to get deeper insights into these coating issues. The authors greatly admire what CFD can do in the hands of a good research team, and CFD has been used extensively in providing and validating the TopCoat models. It seems implausible, however, that most practical coaters could use CFD to solve a coating problem at 3:00 in the morning, which (as the authors know because they've been there) is a time when a quick check on a good-enough simulation can be the difference between success and failure.

6.1. A FUNDAMENTAL CHOICE

There are two ways to coat: metered or unmetered. With metered coating the (wet) coatweight is set by a simple formula:

$$\text{Coatweight} = \frac{\text{litres/min}}{\text{metres/min}} \tag{6.1}$$

or, more succinctly:

$$\text{Coatweight} = \frac{\text{Feed}}{\text{Speed}} \tag{6.2}$$

In other words, if you know that you are pumping a certain volume per unit time which is all being transferred onto the web, the one thing you don't have to worry about is average coatweight. It is a strange fact that many people who would never coat without knowing their machine speed find it acceptable to coat without a flowmeter. This is probably the single biggest mistake in the whole of coating. For those who want success in coating the basic rule is: never, ever, do metered coating without a suitable flowmeter.

Now that the coatweight is guaranteed, the "only" thing left to worry about is coat quality.

By contrast, those who choose unmetered coating have a quadruply hard life. First, they have to tune the system to get the right coatweight. Second, they have to be able to measure the coatweight online to ensure they have the right value. Third, they have to tune the system to get the right quality. Fourth, they have to either recycle or discard excess coating fluid. Recycling implies re-filtering, compensating for lost solvent and generally coping with the fact that what they coat now is not what they coated some time ago because it contains a mixture of fresh and recycled solution. Discarding is usually unacceptable because of waste and cost.

To put it bluntly, only desperation should cause a team to choose an unmetered coating method.

6.2. PROVING A NEGATIVE

It can be very helpful to know that something is impossible. It may be regrettable that it is impossible, but at least no time is wasted attempting the impossible. Here are three cases of impossibility to help the team from trying coating methods that stand zero chance of success.

6.2.1. Cannot Go Slow Enough

An outstandingly good metered method is curtain coating. The metered fluid is pumped through a slot nozzle, falls in an elegant curtain onto the substrate and instantly creates a perfect coating. Apart from some issues of controlling the very edge of the curtain (for which some elegant solutions exist) and issues of air-flow at the curtain/substrate interface, there seems no reason why it should not be the ideal technique that everyone uses. What can possibly be wrong with it?

This is where coating science can save a team from committing a

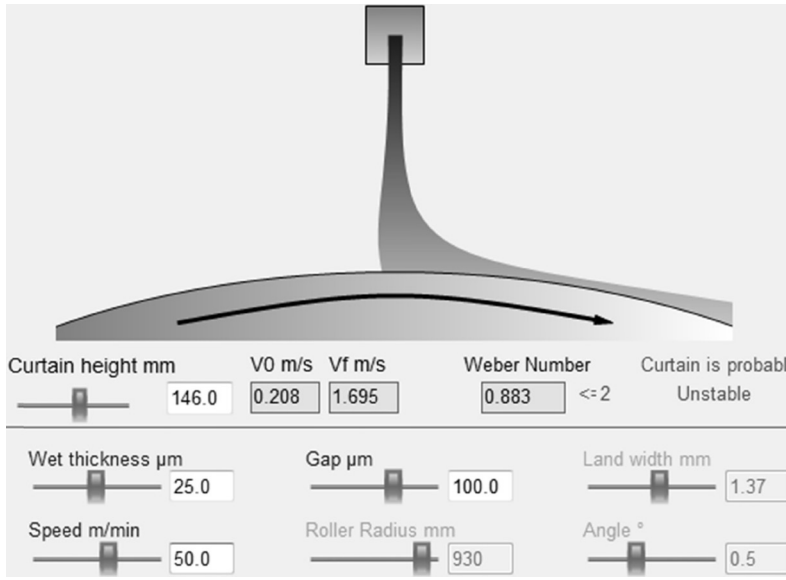


FIGURE 6.1. *A curtain coater trying (and failing) to coat 25 µm at 50 m/min.*

dreadful mistake. Curtain coating is, indeed, a fabulous coating technique that is used around the world every day. Sadly, it is only applicable to coatings that are some combination of thick and fast.

A typical scenario is shown in Figure 6.1.

The user wants a 25 µm coating at 50 m/min. The picture shows an apparently happy curtain flowing out of a slot, delivering exactly the required flow rate. The Weber number, *We*, unfortunately, is < 2, so a warning appears that the "Curtain is probably Unstable". The Weber number is one of many useful dimensionless numbers that aid coating scientists. It represents the balance between inertia and surface tension and is defined as:

$$We = \frac{\rho Q V}{\sigma} \tag{6.3}$$

where ρ is the density, Q is the flux (proportional to coating thickness and coating speed), V is the velocity of the curtain (controlled mostly by gravity) and σ is the surface tension. In this particular example, We turns out to be significantly lower than 2. What is significant about 2σ? The only force keeping the curtain stable is surface tension. This force is 2σ because the force acts on both sides of the curtain. Any chance instability in the curtain can only be stabilized by 2σ. If 2σ is too small

but the curtain is falling fast enough that the instability has no chance to move up the curtain, then the curtain remains stable. This balance of stability and velocity is captured in We and when it is less than 2 the instability wins (inertia is too small), travels up the curtain and rips it apart—just as a slow flow of water from a tap (faucet) breaks into drops rather than forming a steady stream [2].

Going back to this example, the model shows that the curtain becomes stable only at speeds greater than 110m/min. If you are lucky enough to be able to handle such speeds (e.g., have sufficient drying capacity) then curtain coating looks viable. But think about the development process. You will not be able to do small experiments at low speeds. Everything, start-up, coat quality assessment/measurement, drying, handling has to be done at speeds in excess of 110 m/min. It is a brave team that will go straight to those speeds.

And just when things seem OK, it will turn out that the coating needs to be half the thickness. Now the minimum speed at which you can do any coating at all is 220 m/min. At those speeds some of the subtleties such as air-flow dragged along by the web start to become significant factors in making life difficult.

So although curtain coating looks and is beguilingly simple, it is rarely the technique of choice unless and until a stable product requiring high volume production at high speeds is required.

6.2.2. Cannot Go Fast Enough

Another fine negative case can be found within the world of roll coating. Rolls can move either in the forward or reverse directions. In the simplified figures the applicator roll in the bath picks up the coating fluid and the web around the top roller either moves in the same or reverse direction, Figure 6.2.

As a gross simplification of some complex fluid dynamic issues, it might be expected that the forward mode would be generally more useful than the reverse mode for a roll coating process. That it is not comes down to the fact that as the coating fluid comes out of the coating head it must split—either staying on the substrate or returning on the applicator roller. The diagram of what is going on in the forward coating gap shows an applicator roller moving at speed U_{app} and the web moving at speed U_{web}. The split takes place at the curved meniscus on the right of the diagram, Figure 6.3.

The splitting point is a meniscus which is stabilized only by surface

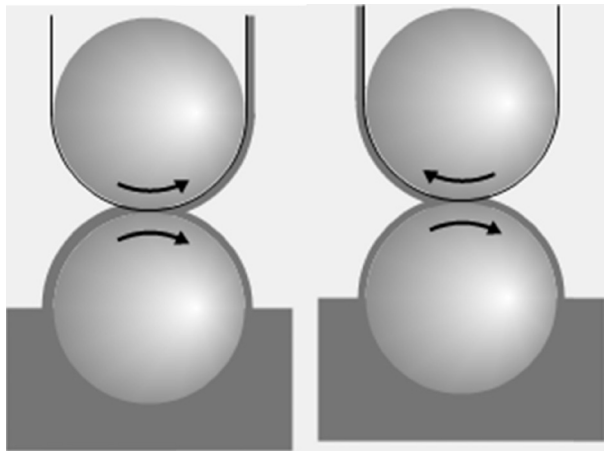


FIGURE 6.2. Forward (left) and Reverse (right) coating modes.

tension forces and the curvature of the meniscus. The pressure field, indicated in the diagram, is very delicate near the splitting meniscus and slight changes to it can upset the meniscus. Above a critical combination of speed and viscosity, the surface tension forces in the one direction of curvature are overwhelmed, so stability is created by creating curvature in a second dimension at right angles to the first—i.e., across

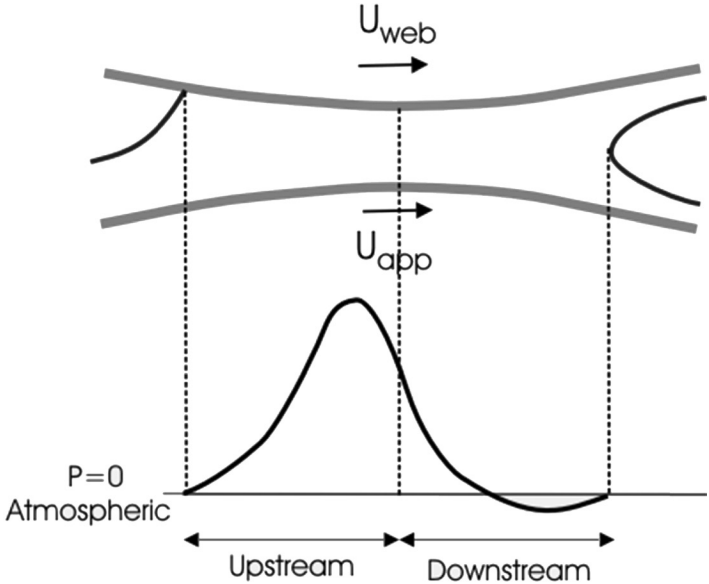


FIGURE 6.3. The pressure distribution and meniscus positions in forward roll coating.

the coating. This creates a sinusoidal profile in the coating meniscus and the result is a beautiful set of parallel lines in the coating, running in the machine direction. This "ribbing" or "record grooving" (for those old enough to know what a record groove is) is a fundamental aspect of forward roll coating.

Whereas problems occurred at slow speeds for curtain coating, for forward roll coating problems increase as speed increases. There is a critical capillary number (itself a function of roll-to-roll Gap and roller Radius) above which ribbing is inevitable [3]:

$$\text{Critical Capillary Number} = \frac{U\eta}{\sigma} > 4\left(\frac{\text{Gap}}{\text{Radius}}\right)^{0.65} + 50000\left(\frac{\text{Gap}}{\text{Radius}}\right)^{4} \tag{6.4}$$

where U is the web velocity, η is the viscosity and σ is the surface tension. The capillary number is another much-used dimensionless constant in coating science and captures the ratio of viscous forces ($U\eta$) to surface tension forces (σ). Unfortunately for forward roll coating, for most practical coating scenarios this critical capillary number is rather small, so ribbing is the norm rather than the exception, Figure 6.4.

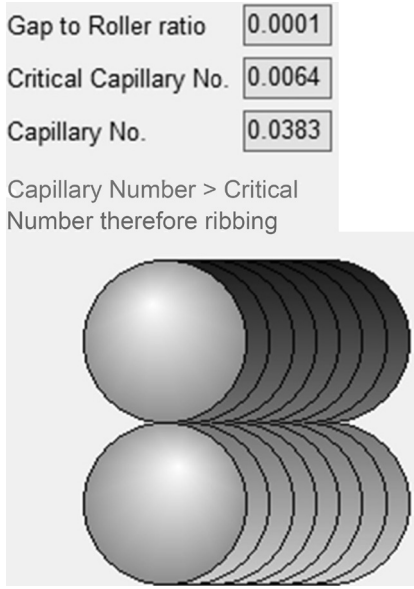


FIGURE 6.4. *Ribbing in forward mode when the capillary number exceeds the critical value.*

In the example shown in the ribbing modeller, the roller radius is 200 mm, the gap is 50 μm, the surface tension is 40 dyne/cm and the viscosity is 20 cP. The ribbing appears at all speeds above 1.6 m/min.

Roughly speaking, it is not possible to forward coat most products at commercially sensible speeds. This applies not only to simple 2-roll coaters. Any forward roll configuration in a multi-roll system is likely to produce ribbing which will cascade through the stack of rollers to the final coating.

Reverse roll configurations don't (generally) suffer from ribbing, and hence are much more popular. The sorts of defects that ruin reverse roll configurations are less easily described with simple physics, and so aren't discussed here. Those limitations combined with the ribbing problems of forward roll make it a nightmare to find the right combination of roll speeds to give the correct coatweight and the correct coat quality. Simple roll coaters can seldom achieve the correct combination of speed and quality. Complex roll coaters can achieve both but at the cost of greater complexity, greater lack of understanding of what influences what, greater openness to operator "tweaking" that sends the system out of control and to the obvious problems of extra sources of defects from all those open rollers that can pick up dirt, dry out, fail to be cleaned properly and so forth. Although there is no knock-out proof that roll coating should never be used, this section (plus the fact that non-metered methods are intrinsically unsatisfactory) at least reminds readers why it should be a choice of desperation. The authors' experience is that the cost of an alternative coating head (such as slot, discussed extensively below) is far less than the cost of months of fruitless trials with a roll-coating head, though if Production is happy with the roll coater speed and performance it would be unwise to change. Just don't budget for large improvements in coating efficiency.

6.2.3. Cannot Go Thin Enough

Knife or blade coating seems a very attractive method as it is so cheap and simple. An excess of coating fluid is applied and a carefully-positioned sharp knife blade scrapes off the excess, resulting in a nicely metered coating.

The diagram shows a sophisticated knife setup along with a pressure curve (above) which indicates that everything is nicely under control. Seen like this, the only objection to knife coating seems to be that it shares all the other problems of unmetered coating.

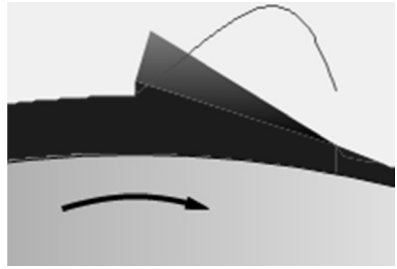


FIGURE 6.5. *A knife coating operation and its pressure profile.*

The problems arise when a thin coating is required. Although there are sophisticated formulae to describe knife coating, in essence the rule is that:

$$\text{Coatweight} = \frac{\text{Gap}}{2} \tag{6.5}$$

So for a 25 μm coating, the knife needs to be held in a controlled manner 50 μm away. If one side is 52 μm and the other is 48 μm then the coatweights will be respectively 26 and 24 μm, an error of ± 4% which is relatively serious. Controlling a gap to 2 μm across a 1.5 m coating is actually quite hard, so already knife isn't looking so easy. Add the complications of hydrodynamic forces trying to bend the knife away from the roller and things are starting to look difficult.

Now go to a 12 μm wet coatweight. This requires a gap of 24 μm, so a ± 2 μm error in knife gap translates to a ± 8% coating error.

The problem is that knife coating relies on absolute gaps for control and on a real-world coater, absolute gaps are very hard to set and maintain down to the accuracies required of precision knife coating.

Because metered coating is the only wise choice, and because curtain coating is so limited, the obvious technique to use is classic slot coating. As has been found by many people, slot coating is a wonderful process till you have to coat thin; then it becomes a nightmare. The point is that it is *impossible* for conventional slot coating to produce thin coatings without exquisite engineering and even then there is a limit. This at first sounds surprising. The coatweight is not controlled by anything other than flow rate and web speed. So how can slot fail to deliver the correct coatweight? It cannot—the coatweight will always be perfect on average. The only problem with slot coating is coat quality: and for thin coatings it is impossible to achieve good quality.

There are two reasons for this. The first is similar to the problem of knife coating—it is impossible to sustain a stable bead in the slot coater if the gap is greater than about twice the coatweight. The second is an unfortunate consequence of Poiseuille's law of flow.

Poiseuille states that the flow Q of a liquid through a gap depends linearly on the pressure drop P across the gap (that's intuitive) and inversely on the length L along which it flows (intuitive) and the viscosity μ (also intuitive). It also depends inversely on the cube of the gap, h:

$$Q = \frac{kQL}{\mu h^3} \tag{6.6}$$

If the gap is halved, the flow decreases by a factor of 8 or, for the same flow, the pressure must increase by a factor of 8. When the gap is large, small changes in the gap mean only modest changes in the flow. When the gap is, say, 10 µm, a 1 µm reduction in gap means a pressure change of $(10/9)^3$ an approximately 37% increase. This big pressure dependence has two unfortunate consequences, Figure 6.6.

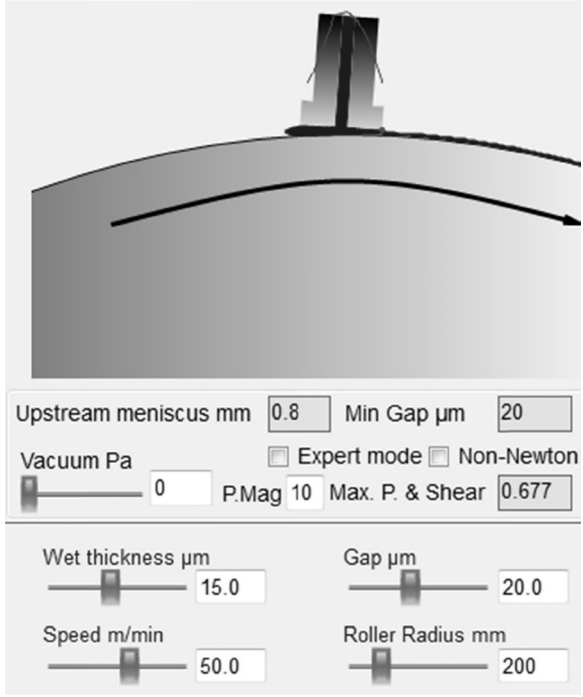


FIGURE 6.6. *Slot coating 15 µm wet—all seems to be OK.*

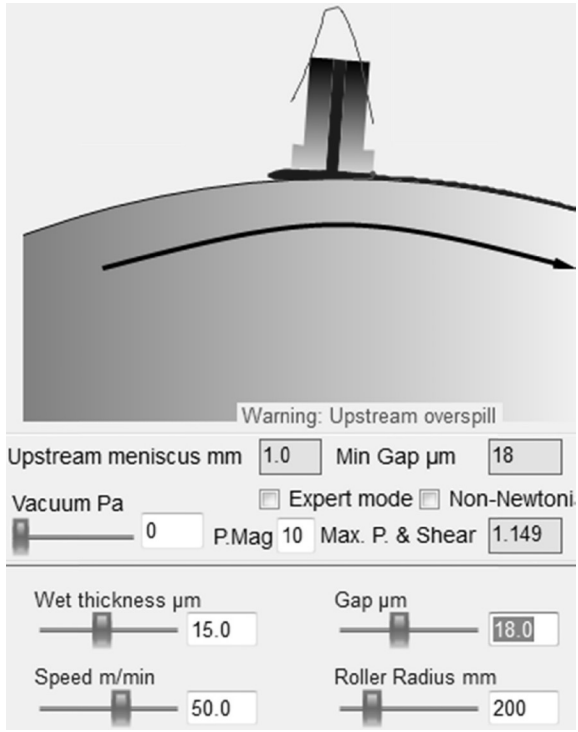


FIGURE 6.7. *A slight decrease in slot gap gives upstream overspill.*

A typical slot coating is shown in Figue 6.6. 15 µm wet is being coated, the gap is 20 µm and with this viscosity and speed the back pressure (Max. P. in the image) is 0.65 bar—a perfect slot coating setup.

Reduce the gap to 18 µm and observe what happens, Figure 6.7.

The pressure shoots up to 1.15 bar and the liquid squirts out of the back of the slot, producing a poor-quality coating. Increasing the gap to 25 µm produces the opposite problem, Figure 6.8.

The pressure becomes too low and the fluid fails to form a proper flow pattern within the slot, causing a characteristic pattern of high and low coatweights sometimes called "cascade".

Even if the gap is controlled between either of these extremes, small differences in gap across the web mean large pressure differences across the web, so that the coating on average will be pushed from high- to low-pressure. A 1 µm-larger gap on one side has no first-order effect on coatweight [which is (liters/minute)/(meters/minute)]; instead there is a second-order effect leading to a cross-web flow producing a thicker coating on that side. Precision controls of the cross-web gap can poten-

tially fix this, but substrates to be coated are often not accurate to this level; therefore, the points of high and low coatweight can shift around thanks to small changes in substrate thickness.

Some teams pin their hopes on adding a vacuum box to help stabilize the coating when the gap is too large. Except for low speeds and/or viscosities, the pressures within a typical thin slot setup mean that applying an upstream vacuum (which is only a reduction in pressure by 100–1000 Pa, i.e., 0.1–1% of an atmosphere) is largely irrelevant.

There is no way out of the Poiseuille dilemma with conventional slot coating. Exquisite engineering can take you to impressively thin coatings but at a high cost, a high risk of damaging the slot lips and the certainty of not being able to go that little bit thinner as demanded by the customer. At the extreme limit is the LCD industry who, thanks to their billion-dollar-plus budgets, are able to create very thin coatings on super-flat glass using super-precise slot coaters. Because even under these circumstances it is possible for the coating lips to touch the glass, they are often made from ceramic rather than steel so as to be able to survive the encounter.

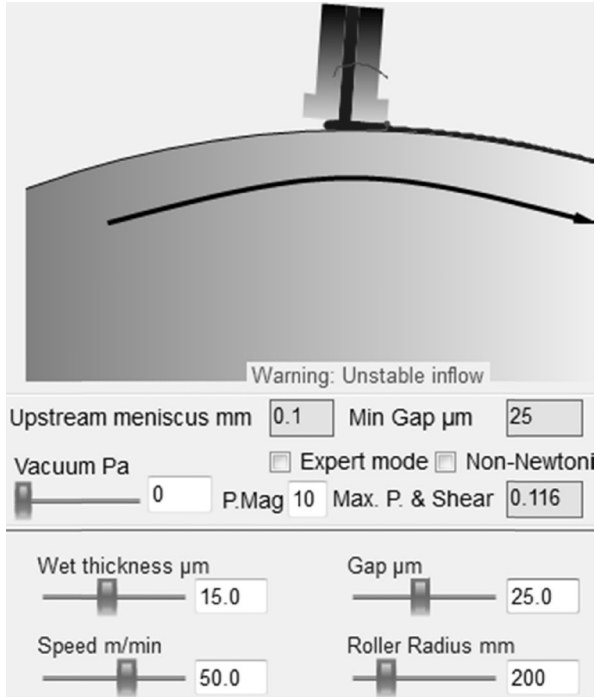


FIGURE 6.8. *A slight increase in slot gap gives an unstable inflow.*

Because slot coating is such a beguilingly simple technique and works so magnificently over such a wide range of coatings, teams think that they are doing something wrong when they fail to produce thin coatings on their slot coater. The only thing they are doing wrong is fighting the laws of physics, which is always a losing proposition.

But all is not lost. We will shortly meet two variants of slot coating that can readily deliver thin coatings.

6.3. GOING THIN

It seems to be generally the case that high-tech applications require thin coatings. Conventional slot and knife have just been shown to be essentially impossible. Multi-roll coating is such an ugly process that it too should be ruled out. So what *can* be used?

6.3.1. Gravure

The answer for many is gravure coating. A finely-engraved cylinder picks up excess liquid and a perfect doctor blade scrapes off the excess, leaving an exactly metered amount of liquid in the cells. The cells then contact the substrate within the coating nip and, as they separate, a fixed percentage of the contents of the cells is transferred to the substrate. The diagram is for the kiss configuration in the forward direction. There are many variants with and without impression rolls, transfer rolls, with enclosed doctor blades, even pressurized doctor blades, Figure 6.9.

In this simple mode, there is a high risk that the individual dots of coating transferred via this process will fail to join up to a perfect uniform coating, creating the well-known cell pattern that mars many attempts at gravure coating. With the reverse setup (web moving in the opposite direction to the one shown), this patterning process is avoided (along with any tendency towards the ribbing defect). The cells feed a meniscus and the substrate removes fluid from this meniscus with no danger of a cell pattern showing in the product. In reverse mode, the relative speed of the web can be adjusted to alter the final coatweight—another advantage of this technique.

There is surprisingly little science in gravure. There is no obvious way to specify the design of the cells in a gravure cylinder from the basic inputs of required thickness and fluid properties. A guess that 50% of the volume will be transferred is usually seen to be reasonable, and with luck, some on-machine tweaking of impression roll pressures, doctor

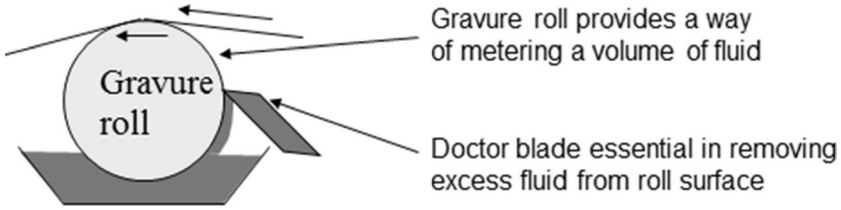


FIGURE 6.9. *One of many gravure setups.*

blade pressures, and relative roller speeds can compensate for cells that do not have quite the right combination of volume and % release.

To produce thin coatings is not trivial. Generally, 3-roll setups are required with high precision, which is particularly difficult when a rubber roller is involved—precision engineering of rubber is a challenge. With a high-quality setup, coatings of a few μm can be delivered reliably—making them the obvious default option when all else fails.

If gravure is relatively low in science, it is (or should be) very high in engineering. Creating a high-quality cylinder is difficult. Finding the right balance of doctor blade assembly so that the scraping off of excess fluid is perfect without undue wear on blade or cylinder, is a major engineering challenge. Because gravure coating has been established for decades, these engineering problems have been confronted and largely been solved. A good gravure head with a good doctor blade assembly and a good cylinder can produce excellent thin coatings in a manner that is simply impossible for the other techniques discussed so far.

One phenomenon that is a problem is that during a run the cells tend to fill up with solids from the coating via a variety of mechanisms. As the cells fill up, the volume delivered goes down. The severity of the effect depends on subtle interactions between the formulation, the doctor blade and the transfer mechanism—plus the length of the coating run.

For those who require printed, as opposed to coated, products, gravure's advantages are even more striking. Many printed electronic applications require "patch coating" with uncoated areas in both machine direction and transverse direction. Gravure is excellent for this.

Yet the fact that there is no real understanding of how much fluid comes out of the cell, when coupled with the facts that the coating is being recycled (even in enclosed chambers), and that blades and cylinders can wear, makes coating specialists search for a process under more natural control. The fact that so many different gravure configurations exist is possibly an indication that the technique is not perfect.

What might a more perfect process be? This brings us to the question of how to defeat slot coating's Poiseuille problem.

6.3.2. Rubber-backed Slot

The core problem of slot is that the downstream edge of the slot lips needs to be close to the substrate for thin coatings (the gap cannot be greater than 2× wet thickness). This produces the Poiseuille problem that small changes in gap lead to large changes in pressure. If excess pressure could be somehow absorbed, then the Poiseuille problem disappears and slot merely needs high-quality engineering to achieve its aim.

One attempt to achieve this is by using a rubber roller instead of a steel one. Any excess pressure can be absorbed by deformation of the rubber through the substrate. For normal, low, pressures the rubber will hardly be deformed, so the downstream and upstream areas will behave conventionally, and slot theory will work perfectly. In the areas of local high pressure, the rubber can deform. If it deforms too much then the pressure will fall catastrophically, so the deformation itself will be reduced and the whole system can quickly reach an equilibrium condition.

Producing a model of this process is a difficult technical challenge. The version in TopCoat seems to do an adequate job of predicting the outcomes of changing key parameters including, importantly, the modulus (or, crudely, the Shore hardness) of the rubber, Figure 6.10.

Here a nominal 10 μm gap produces a reasonable distortion of the 60 Shore rubber, and the previously-impossible 15 μm coating seems to be no problem at all.

It is said that "a major corporation" has used rubber-back slot "to considerable advantage" in certain "critical thin coating applications." It would be good to know more, but no publications on their successes have been forthcoming. Obviously if the technique is a source of competitive advantage, then there is no reason why any details should be published.

The authors' own experience is both triumphant and farcical. Triumphant because with a relatively unsophisticated setup they were able to create wonderfully thin coatings that were impossible to produce using the same equipment and a steel roller; farcical because at every revolution of the roller, the number 23 appeared on the coating. This came about because someone in the supplier's shipping department had help-

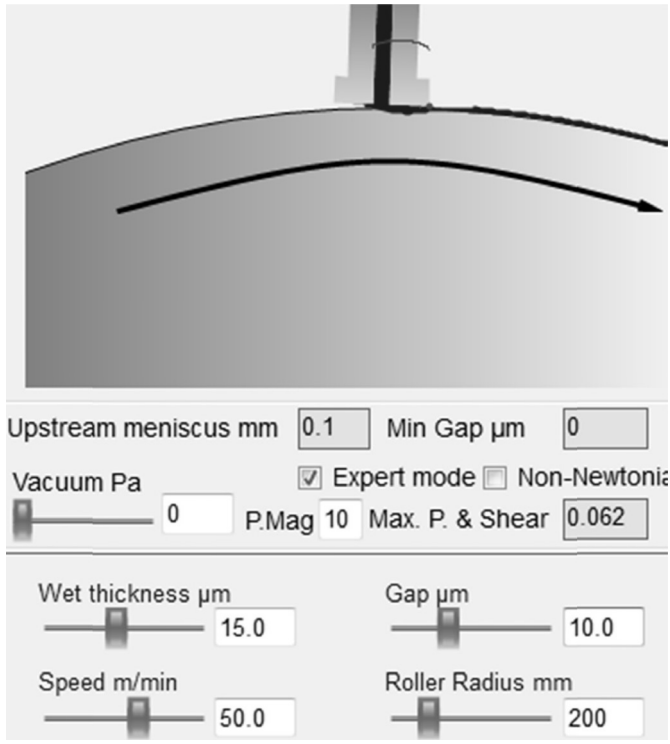


FIGURE 6.10. *Rubber-backed slot coating absorbs the high pressure, making it easy to produce thin coatings.*

fully written the code number 23 onto the protective sheet of polyester wrapping the roller for transport. The pressure of the marker pen had transferred through to the rubber and deformed it sufficiently for the impression to appear on the coating. In addition, all attempts to clean the rubber from any slight dirt created further subtle defects in the surface which appeared in the final coating.

Rubber-backed slot coating, then, seems to be a wonderful technique for those who can obtain a perfectly engineered, perfectly resilient, perfectly cleanable rubber. There is also the problem that rubbers can undergo work hardening. It is therefore necessary to consider the potential lifetime of a rubber roller. Presumably the "major corporation" managed to obtain such rollers or had a large stock of identical rollers to be used for each coating run.

Although the principle is sound, there are still practical challenges to be met. Fortunately, an alternative technique exists which has proven to be game-changing.

6.3.3. Tensioned-web Slot

The key problem with thin slot coating is the backing roller. If it is steel it is too rigid; if it is rubber it seems to be too difficult to maintain its perfection. So why not throw away the backing roller?

To most people this sounds wrong; slot is all about precision; placing a slot near a wobbly web does not sound like a recipe for success. But as with much coating technology intuitions are wrong; tensioned-web slot works astonishingly well for thin coatings and an obscure law of web physics makes it possible.

Here is the same 15 μm coating, Figure 6.11.

The web is pushed by 5 mm into a span of web 200 mm long (the Y-axis is greatly exaggerated in the image; in reality, the effect is not so dramatic). The tension is a typical 100 N/m. With this relatively crude visual model, it is hard to see that the web is deformed under the zones that would be of high pressure and, as with the rubber-backed slot, the deformations are auto-correcting. A slightly rounded upstream lip (not drawn in the simple graphic) solves the issue of the web touching the lip during exit.

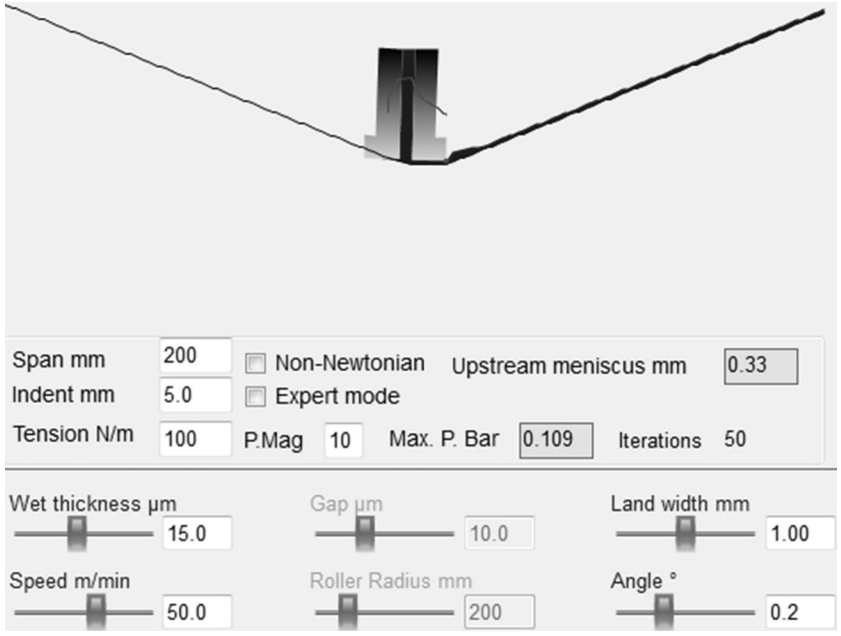


FIGURE 6.11. *Tensioned-web slot coating lets the web deform to absorb high-pressure regions.*

The first academic paper on the technique appeared in 1998 and nothing more was heard about it in public [4]. Following enthusiastic adoption of the technique in (especially) Japan, Taiwan and South Korea, there is now a substantial amount of literature confirming that the simple model in TopCoat is adequate for understanding the various trade-offs of speed, viscosity and thickness [5]. It is said that the current state-of-the-art is high quality coating of 1 μm wet thickness but "only" up to a width of 1m.

Anyone familiar with a real web will know that the absolute tension is hard to control with precision. Fortunately, the self-correcting ability of the technique means that the coating above works well at tensions between 50 and 150 N/m. In other words, absolute tension is of little consequence. The key problem faced by everyone is that the cross-web tension must be uniform. A slack edge means lower pressure, a net flow of coating from high to low pressure and, therefore, poor cross-web uniformity. The obvious way to fix this is by an adjustable roller that compensates for a slack edge. Those familiar with web handling are immediately alarmed by this. In most circumstances, changing the position of a roller causes wrinkles in the web. Happily, in this case the physics of webs shows that the motion required to compensate for the tension is in a plane that has no effect on wrinkles. What is fascinating is that, for a technique which is delivering coatings in the μm range, the required movement to correct for tensions is in the mm range. So relatively simple engineering can deliver absolutely high coating quality.

6.3.4. Lane and Patch Coating

Another strength of slot coating is that it can readily be engineered to deliver lane coating—i.e., bands of coating in the machine direction. This is of particular importance for printed electronic devices such as organic solar cells. A blockage of appropriate width is placed between the slot lips at appropriate points (via more or less sophisticated methodologies) and, with a little theory and/or trial-and-error, it is possible to create repeatable lanes.

Patch coating, with interruptions in the cross-web direction, is possible with slot but suffers from the obvious problems of how to stop the flow instantly (the pressure within the slot die tends to lead to a dribble of coating) and how to restart with a perfectly clean edge. It is said that both are possible given ingenuity and good engineering. An excellent

thesis describing the issues of the start-up part of patch coating also contains references to key patents that tend to cluster around the basic ideas of (a) keeping the pump going continuously (via a bypass) and (b) providing a controlled push at the start and pull at the end to avoid a starved start-up and an extended stop-down [6]. That the details of slot-based patch coating are not readily available in the open literature is, perhaps, another case of competitive advantage being preserved in-house.

6.3.5. Dual-layer Coating

Putting down multiple layers in a single pass was routine in the photographic industry. By focusing on a single chemistry, a relatively modest range of viscosities and tightly controlled surface tensions, it was possible, after years of development, to put down tens of layers in a single pass. For almost every other industry, a single coating was difficult enough, so attempting even two layers was beyond the range of most coaters, Figure 6.12.

Dual slot coating is, in principle, not too difficult to implement. The internal structure of the die allows two feeds. The downstream feed (the top layer in the final coating) flows onto the fluid provided by the upstream feed. The relative thickness of the two coatings depends only on their relative flow rates.

Before there were good models of the technique, dual-layer slot coating was confusing to those brave enough to try it. Under some circumstances it worked marvellously, and then some small change would flip it to instability. New models now exist, and it is possible to create reliable simulations of the process [7].

A simple rule emerges from the simulation. In the example below, the top layer is 50% the thickness of the bottom layer and the coating is stable, Figure 6.13.

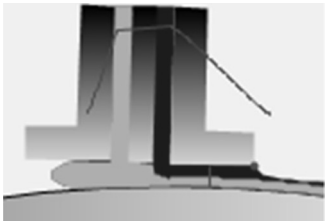


FIGURE 6.12. *The basics of two-layer slot coating—the downstream fluid coats onto the fluid from the upstream feed.*

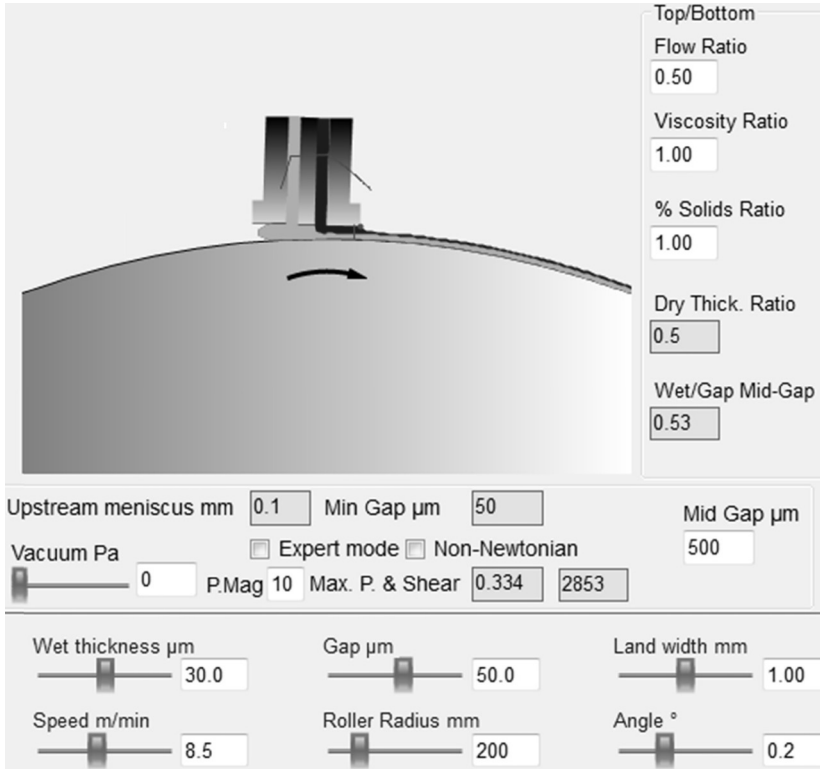


FIGURE 6.13. *Dual-layer slot coating working well.*

A small increase in the relative thickness of the top layer from 50% to 60% flips the system into instability, Figure 6.14.

What is happening is visualized in the graphic. The "mid-gap" (i.e. the space between the two flows) is "invaded" by the top layer, which moves to the input zone of the bottom layer, rendering the whole system unstable.

It is not the flow ratio itself which is causing this problem. What is happening is that the "1/3" rule has been violated: the thickness of the bottom layer is < 1/3 of the total gap between slot and base. The invasion can also happen if the flow ratio is returned to 50% and the overall thickness is reduced. In the case of the 60% flow ratio, if the overall gap is reduced by 10 µm, then the mid-gap invasion doesn't happen and the coating becomes stable.

Why 1/3? Because that's how the fluid dynamics work out within this complex system. Without a grasp of the 1/3 rule it is hard to spot

why a small change in flow ratio or gap can flip the coating from stable to unstable. Now that the rule is understood, it is possible to rationally adjust relative % solids and coatweights (in addition to the general adjustments of slot coating) to obtain reliable, in-control dual coatings. The technique works for tensioned-web slot as well, and so can produce thin coatings [8].

6.3.6. Spray Coating

There is not a lot to say about spray coating. Its advantages are obvious. It is a metered and non-contact system. It can, in principle, do lane and patch coating and, if the definition of "spray" is extended, it can include inkjet and nozzlejet printing techniques. It is a technique that has been used for many years within the paint industry and a wide range of engineering expertise is available. The disadvantages are also obvious. It is exceedingly difficult to create a uniform spray across the full width of a web, and it is exceedingly difficult to ensure that each uniform drop that hits the surface will fuse perfectly with each other drop and yield a coating free of imperfections. As discussed in Chapter 9, the fact that

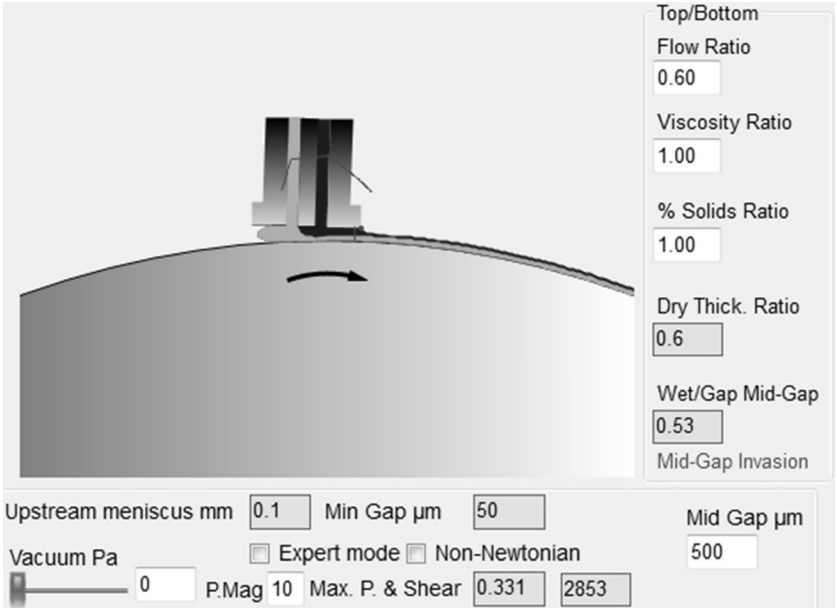


FIGURE 6.14. *A small change to the setup causes the top layer to invade the mid-gap region, producing an unstable system and poor quality coating.*

these are nanoparticle coatings in the form of aerosols means that there are also potentially more health and safety issues to be considered.

For relatively thick nanocoatings applied to 3D shapes, such as polycarbonate car headlamps or parts of mobile (cell) phones and also for large-volume coatings onto wooden flooring, the spray technique is of proven capability, albeit with many issues of recycling overspray and reducing the problems of gravitational flow.

For relatively thin precision coatings the upsides of a successful implementation are so compelling that it may be just a matter of time before the downsides are solved via clever science or engineering. To the authors' knowledge, that happy state has not been reached.

6.4. PRINTING

Like coating, printing breaks down into metered and unmetered techniques. To make this section manageable, only key aspects—often not generally known or understood—are touched upon.

6.4.1. Inkjet

If hype were reality, the whole of printed electronics would now be performed via inkjet. Despite the fact that the limitations of inkjet have always been clear, the dream of being able to place each drop where it was wanted when it was wanted persuaded far too much money to be wasted on far too many hopeless schemes.

A technology that is wonderful for placing drops fairly carefully in the right place for a convincing graphical image does not necessarily translate to one that can place drops absolutely accurately every time. A partially blocked nozzle in a graphics printer can still provide adequate results; such a nozzle for electronics is a disaster. Fatally, nozzles can be blocked by the slightest error in formulation or handling of the head. Getting a tiny drop of liquid from an inkjet nozzle to fly in the right direction through space is a delicate balancing act of surface tension and viscosity, with the ghost of Rayleigh ever present to create an instability in any "neck" that forms, thereby creating a satellite drop that can head off in a random direction.

The limitations on surface tension and, especially, viscosity (anything over ~30 cP is difficult) are a massive constraint. A drop will spread with frightening speed once it contacts a surface. Many examples of "printed electronics" rely on a photolithographic step to place

hydrophobic barriers around the places where the drop will land, thereby confining it. If photolithography can be used in one step, why not use it in other steps; why not just make conventional electronics instead?

Adding to the woes of inkjet is the infamous coffee ring effect. Again, the surface tension and viscosity constraints of inkjet make the drops particularly susceptible to this phenomenon, whereby the edge of the drop evaporates faster, gets pinned in place and thence attracts more liquid, which brings more solid material, so that the majority of the coating ends up at the edge of the printed feature [9].

Because inkjet nozzles are so small, pigmented systems must, of necessity, use nanosized ingredients. It is not possible to inkjet-print the large silver flakes that create adequate circuit tracks for conducting significant amounts of current. Therefore, inkjet has had to embrace nanosilver (or nanocopper, although this has proved more difficult to produce at an industrial scale because of the strong tendency to form copper oxides and because of sintering problems). As discussed in Chapter 2, there are reasons why nanosilver can give comparatively fast sintering to provide good conduction, though this is countered by the need to ensure that the particles are wonderfully dispersed in order to pose zero risk of blocking the nozzle. The net result is a lot of unhelpful dispersant molecules between silver nanoparticles, rendering the hoped-for advantages illusory—at least as judged by the gap between promises of nanosilver inks and successful delivery of inks that perform routinely well in production when sintered at temperatures low enough to be compatible with common substrates such as PET or PEN.

With Herculean efforts, many of these problems are being overcome. The prize of being able to put each drop in the right place at the right time is one worth fighting for. The point of this section is to alert the incautious that inkjet for printed electronics has not lived up to even a fraction of the hype that has been pumped out over the past decades.

6.4.2. Offset Printing

The basis of offset, a deliberately tenuous oil and water mix placed onto a plate, seems to offer scant promise of success for serious nanoprinting. One of the key issues about film splitting during separation also affects flexo printing which, having more promise, is discussed in more depth.

6.4.3. Flexo Printing

Traditionally, flexo printing is often thought of as a relatively simple technique that works exceptionally well for cardboard boxes and packaging films. Because it hardly differs in principle from a home-made rubber stamping technique or potato printing, it seems to have received scant attention from those concerned with elegance and precision. A glance with a normal magnifying glass at anything that has been flexo printed shows that it has some serious flaws. A letter T, for example, is often found with a thin outline more or less following the T. Small dots sometimes have rings around them while "donut dots" with holes in the middle are often observed. Solid areas are often full of lighter areas, so much so that in order to print a solid area some printers resort to printing a 95% dot coverage, finding that this results in a solid print with fewer defects.

The reasons for these defects can be visualized via a simple model written by Abbott for MacDermid Printing Solutions, who are experts in the manufacture of flexo plates [10]. A core idea in the modeller was provided by Dr. Nik Kapur of the University of Leeds.

Before addressing those defects, another piece of physics needs to be explained. Because a flexo plate is rubber, it might be assumed that it can readily conform to any surface. This ignores a fundamental formula for the compression modulus of a material which shows that, for a perfect rubber, the compression modulus is infinite and for a normal rubber the compression modulus is merely very large. The reason is that the compression modulus E_c (or, as it more normally referred to, the bulk modulus K) depends on the (low) tensile modulus E and the Poisson Ratio, ν, which is the amount by which a movement in one direction translates into a movement in the orthogonal direction:

$$E_c = \frac{E}{3(1-2\nu)} \tag{6.7}$$

For a perfect rubber, ν = 0.5 so E_c becomes infinite. For a reasonable flexo rubber, ν ~ 0.49 so E is 9× higher than expected, making the rubber effectively non-compressible.

For a plate to be conformal, a structure that allows true compression has to be included in the system, typically a foam layer between the plate and the system that holds it onto the press.

A stylized illustration of flexo now explains the standard faults. It is

FIGURE 6.15. *An idealized flexo dot just about to print against the substrate at the top.*

assumed that the rubber dot has been covered with a thin layer of ink via the standard anilox (gravure-like) roll feed system. The dot is close to, but not quite touching, the substrate at the top, Figure 6.15.

Once it contacts the substrate, two things happen: the dot expands laterally, and the ink gets squashed, expands and starts to run down the side of the dot, storing up trouble for later cycles, Figure 6.16.

Once the dot starts to separate from the substrate, the ink has a dilemma. It knows theoretically that 50% of it must end up on the substrate and 50% remain on the rubber. There are no physics that can significantly alter this ratio (surface effects, for example, play no part in the splitting of the ink). One way for this split to occur is for a meniscus to be created at the outer surface of the ink and then run between the sepa-

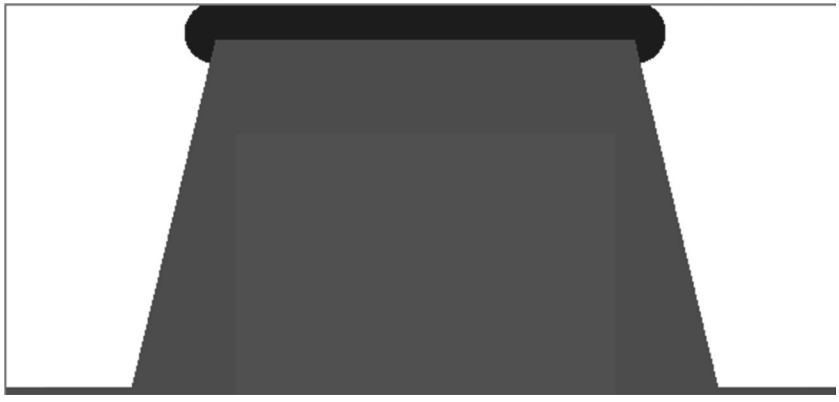


FIGURE 6.16. *The ink and dot spread under contact pressure.*

FIGURE 6.17. The only way to induce separation is to create voids in the fluid.

rating surfaces to produce a clean break. A moment's thought shows that this is not possible, given the small ratio of ink thickness (typically 3–4μm) to dot width (50μm up to 1m). Given that the ink must split, its only choice is to create internal splits—cavitation. Bubbles of air or solvent spontaneously appear as the internal pressure falls to sub-ambient. That is indicated in the animation, Figure 6.17.

After final separation, both surfaces show the remains of the cavities that were formed. For small dots, this can be a nice clean donut. For large solids, this is a messy random mix of defects that show up as a strongly inhomogeneous "solid", Figure 6.18.

FIGURE 6.18. On separation, the coated dot is uneven because of the remnants of the voids.

FIGURE 6.19. *Ink accumulates along the edge of the dot, allowing, for example, ghost images outside straight lines.*

After a few prints, the ink has built up significantly along the side of the dot and it can be readily imagined that the edge ink can print as a halo. This is particularly likely along straight lines because the tendency to buckle under pressure is higher for a straight line than for a curved dot, Figure 6.19.

For all its obvious limitations, the modeller provides insights that should guide the choice of whether to use flexo. The buckling around straight lines and the tendency to cavitate are inherent to the system and either have to be accommodated or avoided by clever science.

These truths are in addition to the more well-known features of flexo: that dot-gain by lateral expansion is to be expected; that the system seems to be suitable only for inks with viscosities (after shear) of a few cP; that the ink deposit can only be a few μm thick.

The generally poor resolution can be addressed via harder dots. This goes against the instincts of those who want the rubber to be compliant, but by using a proper approach to compliance (such as a foam) this problem can be overcome. Traditionally, the height of a typical dot is large (100 μm or more), making it very hard to control very thin dots. There is no obvious reason for such a large height. A smart team should be able to get good results with much shorter dots.

By working within its limitations and intelligently pushing the boundaries based on the laws of physics, flexo can be an awesome combination of simplicity and functionality.

6.4.4. Gravure

There really is little to be said scientifically about gravure printing. As discussed in gravure coating, the rules governing the transfer of fluid from the cells to the substrate are not fully understood and most printers come to their own conclusions about the optimal shape and size of cell for their own coating. The fact that a single test cylinder can have numerous patterns placed on it makes it comparatively easy to home in on the right combination. It is not an elegant approach, but it works.

The key issue that can be addressed by pure science is whether the printed pattern will join up to create a uniform film.

Assuming that the doctor blade has not been too perfect, the printed dots at least start off with a minimum layer of connecting fluid. This allows the use of a formula that is both remarkably simple and remarkably powerful. Imagine that the coating has a sinusoidal variation in thickness. The wavelength (divided by 2π) is λ and the average thickness of the coating (not the height of the sine wave) is h. That a real gravure coating is anything but sinusoidal is less of a problem than it seems. As we will see, the levelling time has a strong dependence on λ and via a Fourier analysis it can be seen that features sharper than a sine wave are equivalent to features with shorter wavelengths which level far quicker than the basic sine wave.

The time for a coating of viscosity μ and surface tension σ to level by $1/e$ is given by the following formula which is implemented in the Levelling spreadsheet:

$$t = \frac{3\mu\lambda^4}{\sigma h^3} \tag{6.8}$$

To those who are unfamiliar with this expression, a common instance of its predictive power is in painting a door. A cheap paintbrush with coarse bristles gives a larger value of λ and the only way to get the paint to level before it dries is to apply a thicker coating (increasing h). This has the unfortunate side-effect of the paint running down the door. A wise painter spends a little more on a high-quality brush with very fine bristles, leading to a 16× reduction in the time taken to level if the bristles are twice as fine. This means that h can be reduced by a factor of 2 (giving an 8× increase in levelling time) to reduce runback whilst still having a factor of 2 overall improvement in the quality of the painted door.

The good news is that the fine gravure patterns used in high-quality printing (small λ) have a profound effect in reducing the time for the cells to level out. The bad news is that gravure is typically used for thin coatings and a small h means not only a slower levelling speed but also faster evaporation of the solvent, which increases μ and therefore reduces the chances of levelling. It is highly likely that this h^3 effect is the biggest barrier to successful gravure printing (and to direct gravure coating). Given that it comes from a fundamental law of physics (the cubic factor comes from Poiseuille), the chances of overcoming it are small.

If the dots are isolated (i.e., there is no thin connecting film between them), then the problems are even more severe. Dots spread via surface tension; simple "Tanner theory" (discussed below) shows that the speed of spreading depends on θ^3, where θ is the contact angle of the drop with the surface. For a very thin dot (small h), θ is (by simple geometry) very small, so the spreading speed is very low. Given the dual issues of preferential drying at the perimeter and pinning (coffee ring effect), it is highly unlikely that isolated dots will be able to join up in time.

6.4.5. Screen Printing

Screen printing is so often treated with scorn (it is, after all, used for printing T-shirts and decorating cakes) that its formidable strengths are forgotten. It is a technique that struggles to reliably print features below 50μm in lateral dimensions; it is a technique that involves large areas of messy ink being scraped to and fro by squeegees and flood bars. Finally, it is a technique which generally cannot print thicknesses much less than, say, 15 μm of ink.

What it has going for it is that unlike inkjet, flexo and gravure, it can handle high viscosities and can cope with large loadings of relatively large pigments and flakes. It is no coincidence that the vast majority of solar cells which require silver tracks with high conductivity (for efficient current extraction) but narrow width (to avoid blocking light) on the top surface are produced via screen printing.

The perceived low-tech nature of screen printing has led to much unhelpful mythology about how it actually works. This is entirely unnecessary as the mechanism is simple and, astonishingly, rather impervious to variables in the process. A print slowly produced on a hand bench is essentially the same (thickness and resolution) as one produced on a high-speed automated unit.

Another simple modeller produced by Abbott for MacDermid Au-

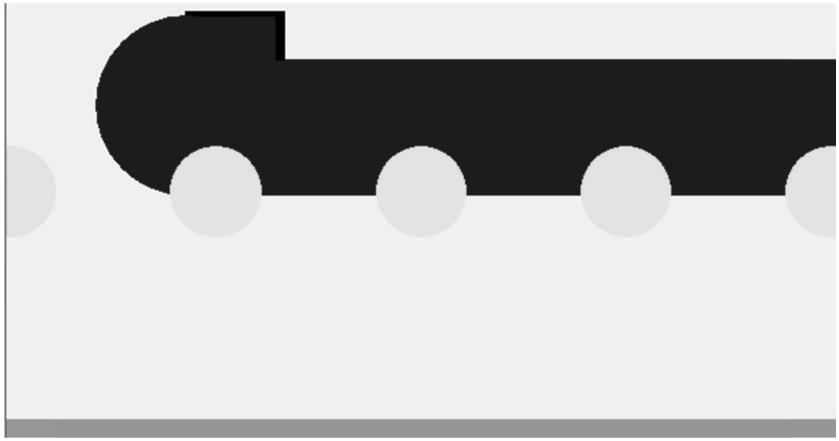


FIGURE 6.20. The flood bar partially fills the screen print mesh.

totype using insights from Prof. Phil Gaskell and Dr Nik Kapur at the University of Leeds, captures the essence of screen printing. A full technical paper describing the science is available [11]. In particular, the approach shows that the much-discussed problem of "how does ink get out of the mesh?" is neatly solved by the fact that it doesn't. The opposite is the case, the mesh gets out of the ink via a process no more mysterious than a spoon coming out of a jar of honey. For brevity, only the basic process in an open mesh is shown here. Those who wish to learn more can refer to the free eBook (Abbott was one of the authors) "How to be a Great Screen Printer" downloadable from MacDermid Autotype. A free set of software modellers is also available from MacDermid Autotype [12].

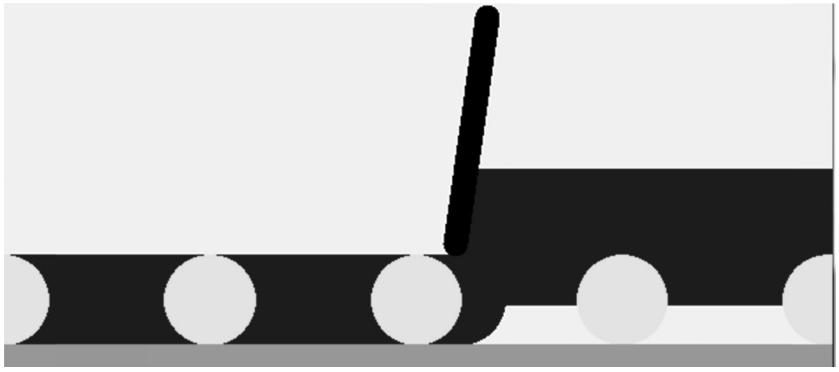


FIGURE 6.21. The squeegee forces the mesh against the substrate, fills the mesh with ink and scrapes off any excess: 3 actions in one.

FIGURE 6.22. *The squeegee is far away, allowing the mesh to come out of the ink—bringing some ink with it.*

The process is a cycle. First the screen is lightly covered with ink using the flood bar, Figure 6.20.

Second, the squeegee comes along and does three things: press the mesh against the substrate; fill the mesh with ink; scrape off any excess ink, Figure 6.21.

Once the squeegee has disappeared, the mesh can start to rise. This, incidentally, proves that the squeegee has nothing to do with the ink transfer—it is only when it is out of the way that the printing takes place, Figure 6.22.

As the mesh comes out of the ink (there is nothing to stop it coming out—just as there is nothing to stop a spoon coming out of a jar of honey) the liquid must flow around it in the most energy efficient manner, which creates a meniscus, Figure 6.23.

As it rises further, the liquid flows round to form a liquid bridge underneath each mesh fiber, Figure 6.24.

Finally, the liquid bridge breaks (creating a drop that is just visible in the animation) and the print is complete. Note that the mesh is covered by a thin layer of ink. It is easy to show experimentally that this is ~30% of the original volume of ink. This basic fact seems to have been essentially unknown to all but a few screen printers, despite it being as obvious as a spoon being covered by honey when it comes out of the jar.

What is at first surprising is that the percentage of ink remaining on the mesh does not depend on either the viscosity of the ink or the speed

FIGURE 6.23. *The ink naturally forms these menisci and liquid bridges.*

FIGURE 6.24. The liquid bridge snaps leaving a larger ink deposit directly beneath the mesh, and ~30% of the original ink wrapped around the mesh fiber.

of the mesh rising from the ink. The reason for this is that above a critical capillary number (speed × viscosity/surface tension) the amount of fluid picked up by a spoon from honey or a mesh from ink remains constant. For screen printing, the capillary numbers are nearly always much larger than this critical value.

Thus, the ink deposit from screen printing is Ink-in-the-mesh-at-the-start–Ink-on-the-mesh-at-the-end. As the starting volume can be readily calculated by basic geometry and because the latter value is about 30% of the original, it is straightforward to calculate the ink deposit from any mesh.

As it happens, the 3 general classes of mesh (polyester, liquid-crystal polymer, stainless steel) all converge on a starting volume equivalent to 25–35 μm of ink, so the ink deposit from most meshes is in the range of 18–25 μm. This makes screen printing hopeless for thin layers, just as gravure is hopeless for the sorts of thick, high viscosity layers that screen can deliver naturally.

Because the alternatives to screen printing are not without their problems, it is interesting to note that an entire flexible solar cell can be screen printed on a mass scale [13]. The key was to recognize that conventional solvents were simply unusable for the very thin layers—such as a 30 nm ZnO electron transport layer—because of the large amount of solvent required and the inevitable problems of "drying in" during processing. Instead, a non-volatile solvent was used, making it very easy to process. The trick was to then turn the non-volatile solvent (3,6-Dimethyl-heptan-3-ylphenylcarbonate) into a volatile one via thermal degradation. Admittedly this was a rather slow process, but the point is that all 4 active layers of a solar cell (30 nm ZnO electron transport, 90 nm conducting polymer, fullerene, ZnO, 250 nm PEDOT:PSS conductor, 6 μm Ag conductor) could be screen printed on relatively simple equipment with relatively high yields.

6.5. DRYING

A lot of misconceptions about drying of coatings can be dispelled via a simple question. Assuming you have long, flowing locks, if you have just washed your hair, which is the better way to dry it: in front of a hot fire or in a warm breeze on a summer's day? The answer, for those who are unsure, is the latter. Drying in front of a hot fire is a slow and painful process. The problem is that the hair and the head both get very hot and the drying is frustratingly slow. A few minutes in a warm breeze flowing through one's locks is all it takes for the hair to be pleasantly dried.

The reason for the difference is that heat alone dries nothing. Heat just makes things hot. What dries things is removal of the molecules. Hot molecules, of course, move faster than cold ones. The trouble is that the speeds with which even hot molecules move are far too slow for effective drying. What is needed to remove molecules is a stream of air containing few of those molecules. More specifically, that stream of air has to reach down close to the surface to be dried. Unfortunately the "no slip boundary condition", which is a fundamental law of fluid flow, tells us that the velocity of air at the surface of a coating is precisely zero—increasing slowly away from the surface. This isn't good enough to remove lots of solvent molecules quickly. What is needed is turbulence. With turbulent flow, fresh air can reach right down into the (comparatively) still zone, pluck solvent molecules away and carry them to their final destination.

So for good drying, the most important element is a good flow of turbulent air; hence the need for the breeze in the summer afternoon. Because evaporation of solvents absorbs heat, the result of all that flowing air is a cooler layer of coating. At some point, the cooling effect reduces the temperature to a level where there is a balance between evaporation and cooling. Clearly, evaporation needs a source of heat, so the temperature of the incoming air is important. You cannot dry a coating without a good source of heat; drying your hair in a cold winter's wind is slow and unpleasant.

Drying therefore requires both airflow and heat. In the lab, a typical coating placed into an oven will take minutes to dry because the airflow is so low. That same coating, in air of the same temperature, will dry in seconds within a modern drying oven which supplies air at just the right degree of turbulence for efficient drying, Figure 6.25.

In this example, a 15 μm coating of 20% solids in water enters an oven of 2 m length at 80°C. The heat from the oven at first raises the

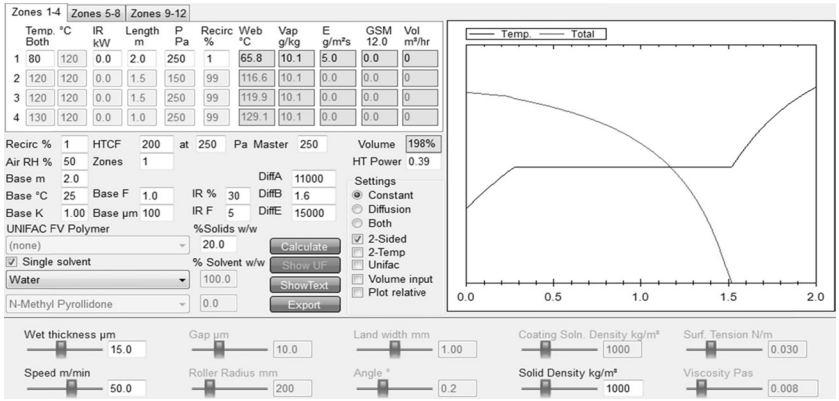


FIGURE 6.25. A 15 µm water-based coating drying in 1.5 m in an 80°C oven.

temperature of the web and the coating (lower curve) from its initial 25°C to 39°C, at which point it remains constant as the water evaporates (top curve, falling to zero at 1.5 m). As soon as the water has disappeared, the web temperature rises towards 80°, though in this example it fails to reach it before it exits the oven with a nicely dried coating.

The domain where the temperature remained constant is called, universally, the "constant rate" zone. It represents the balance of heat coming in from the hot air and the cooling via evaporation. If the solvent is changed from water to one with a similar but somewhat higher boiling point (MIBK), then the drying is much faster, Figure 6.26.

The web had no time to reach a constant temperature because evaporation was over so quickly. The reason that a solvent with a higher boil-

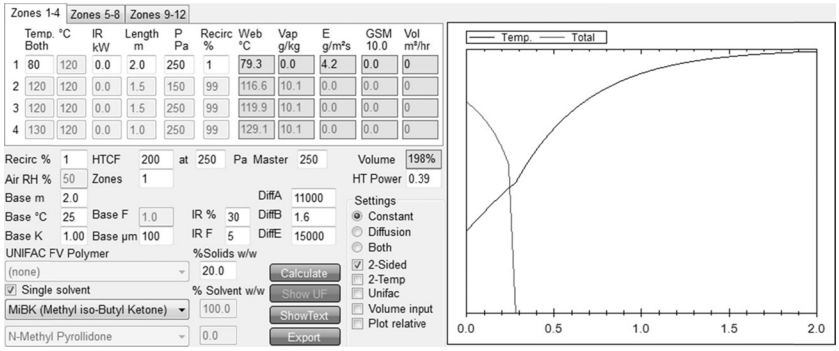


FIGURE 6.26. A higher boiling solvent dries faster because it has a lower enthalpy of vapourization.

ing point evaporates faster is that its enthalpy of vaporization is considerably lower than that of water, so less heat is required to evaporate the same amount of solvent. People often need reminding that heat is not the same as temperature.

In Chapter 5, there was much discussion about rational solvent blends being used to ensure that different components of the coating would remain in solution or, alternatively, crash out of solution depending on the requirements of the coating. There is a general fear of using high boiling solvents, which is entirely unnecessary. A 50:50 mix of MIBK and NMP (BPt 202°) dries much faster than the original coating based on water (which needed 1.5m to dry) and takes only an extra 0.2 m in this oven compared to MIBK, Figure 6.27.

Basic drying theory, with its emphasis on air flow and enthalpy of vaporization rather than temperature and BPt, is highly liberating. In general, most coatings of modest thickness (typical of nanocoatings) with reasonable solvents, even high boilers, are dry within a meter of entering the oven.

Unfortunately, there is a caveat. At some point, the drying polymer coating starts to block the exit of the solvent. The molecules have to diffuse through the polymer before they can escape. At this point, the coating enters the "falling rate" or "diffusion-limited" zone, where the temperature rises and no amount of air flow can increase the rate of drying. Now the only thing that matters is diffusion rate and time. As the drying process nears completion, the diffusion rate decreases further, making it especially hard to remove the last few percentages of solvent. There is only one way to increase the diffusion rate and that is to increase the temperature.

So now we have the opposite advice. For drying in the diffusion-

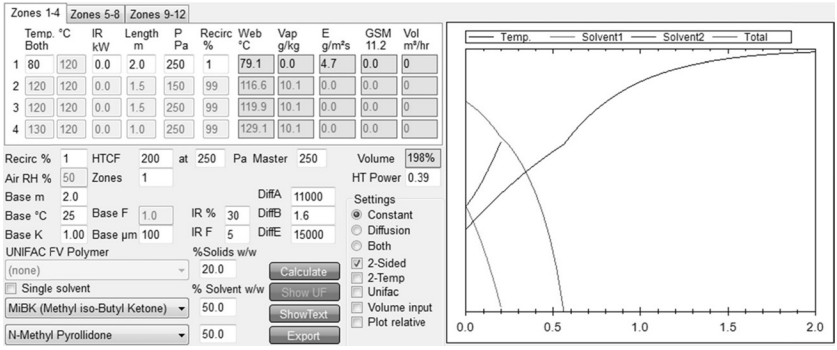


FIGURE 6.27. *Even a very high boiler such as NMP can dry faster than water.*

limited zone, air flow is (mostly) irrelevant and temperature is what matters.

This switch between drying modes is vital for any coating team to grasp. If the coating is in the constant mode, air flow is what matters most. When the coating is in the diffusion limited mode, high-powered fans blowing lots of air are a waste of energy; all that is required is an efficient source of heat and a modest air flow to remove the residual level of solvent. The authors have known machines that were wrong in both directions: machines using temperature instead of air flow and machines using air flow instead of temperature.

Because it is so important to know when you should be providing air and when you should be providing temperature, it is good to know that it is trivially easy to find out where you are in your process, providing the ovens have simple web temperature gauges built in. Failing that, it is often possible, provided side-access is available, to point a temperature "gun" at a web and get a good indication of its temperature. The trick is to see where in the oven the web temperature suddenly rises from below the oven temperature. That is the point where drying shifts from constant to diffusion-limited. It is the single most important measurement you can make in a drying process. If that point is within 0.1m of entering an oven (typical of coatings with solvents such as MEK) then the airflow in that oven can be reduced considerably (subject to LEL safety requirements). If that point is towards the end of Zone 1, then Zone 2 can be run hot with low air flows in order to drive out residual solvent. If the point is towards the end of your final zone then you have serious problems, as it is likely that there will be no chance of driving out the residual solvent.

The cooling effect in the constant zone allows another important degree of freedom. Many coatings have an upper temperature limit beyond which damage sets in. There is, therefore, great fear of setting an oven temperature beyond that limit. If, however, the coating is comfortably in the constant zone, then the oven temperature can be increased beyond the limit, giving much faster drying with no risk of overheating the coating or the web. Many coating lines have been able to double their speed (which had previously been limited by drying) using the fact that the web in a zone considered to be too hot never even gets close to the critical temperature.

There are many more things that could be included in a discussion on drying. The aim in this summary is to bring out the crucial points that are missing so often in discussions of drying. Even without a drying

modeller, the team can make rapid progress by working out when to use air, when to use temperature and when it is safe to raise the temperature without damaging the coating or the web. A simple measurement of web temperature helps the whole team understand where they are in the balance between constant-rate and diffusion-limited drying, so they can apply more air, more temperature or can simply run much faster than they thought. Losing the fear of high boilers makes it possible to keep coatings "open" for longer. As many of these solvents (such as NMP) have very attractive solvency properties for difficult polymers and nanoparticles, a reluctance to use them because of their perceived non-volatility can be swept away.

A final important point that needs to be made is that of emissions; it is usual these days to place catalytic burners on the end of coating lines to incinerate the solvent that is being removed. Care must be taken not to choose a solvent that, whatever its other benefits, will interfere with the operation of the burner. It is also necessary to determine how much solvent is being discharged to atmosphere if a burner is not used.

Having access both to the simplified evaporation modeller in HSPiP in order to understand the balance of solubility and to the more realistic modeller in TopCoat allows the whole team to progress much faster in overcoming the necessary compromises of solvent formulations. The fact that Abbott has a commercial interest in both products does not alter the fact that the previous sentence is true.

6.6. DEFECTS

Two types of defect have already been discussed.

6.6.1. Ribbing

Ribbing appears in forward roll coating when the capillary number exceeds a critical threshold. Ribbing can also appear in slot coating when the pressure gradient at the downstream lip goes negative. Many people performing slot coating have been puzzled by regularly spaced lines appearing on the web with no apparent cause. The fact that they are regularly spaced proves that they are nothing to do with dirt or air bubbles—they are a physical outcome of the wrong pressure gradient, compensated by a sinusoidal fluctuation in the cross-web bead. The fix is gratifyingly simple—move the slot lips very slightly towards the coating and the ribbing disappears. If you are called in the middle of the night to solve a "serious

coating line problem" and merely tweak the gap smaller to make them go away, your reputation as a coating expert will be assured.

6.6.2. Irregularities

The flow-out of irregularities has also been discussed in the context of gravure. It is worth repeating the formula (in the Levelling spreadsheet) because it is so relevant to many coating issues:

$$t = \frac{3\mu\lambda^4}{\sigma h^3} \tag{6.9}$$

If the time to level out, t, is small, then defects will disappear before they reach the drier. The golden rule of defects is to make them small wavelength (λ). Broad coating lines never level, whereas quite severe sharp defects can magically heal before the oven. The physics that makes life difficult for everyone is the h^3 dependence. Marketing, or customers, almost never ask that the coating should be made thicker—they generally want it thinner. Halving the coatweight increases t by a factor of 8, so a defect that could heal itself before the oven may now have no chance to do so.

6.6.3. Pinholes

The malign influence of h appears in another key type of defect. If a small speck of dirt lands on a coating then (with luck) it might be small enough so that it won't interfere with the product performance. Very often, however, a small speck of contamination leads to a large hole in the coating. The universally-observed phenomenon is that a coating which seemed to be relatively easy to coat without pinholes suddenly becomes full of pinholes when the coating thickness is reduced to meet some new product requirement. Why would a 10% reduction in coatweight lead to a 1000% increase in pinholes when nothing else (e.g., cleanliness standards) has changed?

The laws of surface tension say that a hole of diameter d in a coating of thickness h which has a contact angle with the substrate of θ will self-heal if:

$$\frac{h}{d} > 2(1-\cos\theta) \tag{6.10}$$

The formula is included in the Pinhole and Fibers spreadsheet. It isn't immediately obvious what the formula means. The right-hand side makes sense. When the contact angle approaches 0° for perfect wetting, cosθ approaches 1, so the right-hand side approaches 0—in other words, all holes will self-heal via wetting (given enough time). To understand the left-hand side, imagine a very thick layer with a very small hole punched through it—h/d is large and it is obvious that that will self-heal. Now imagine the opposite—a very thin layer with a huge hole punched in it—h/d is small and there is no way that that it can self-heal.

The surface-tension driven motion has a velocity v that depends on the contact angle θ (in radians), the surface tension σ and the viscosity μ as:

$$v = \frac{\theta^3 \sigma}{\mu} \tag{6.11}$$

If a coating has been run for some months or years, the chances are that pinholes have been cured by a cleanliness regime that gives particles of a diameter d that are just at the limit where h/d is less than the critical value—and v is high enough that the hole heals before meeting the oven. Why go to extra (expensive) work to get a higher standard of cleanliness when the pinhole count is essentially zero? When h is reduced by, say, 10%, what was previously a stable system above the critical value now is unstable (h/d is now less than the critical value) and pinholes appear. The only cure is to increase the standard of cleanliness or to tweak the substrate or the formulation so that $\cos\theta$ is sufficiently high. If the coating fluid completely wets the substrate ($\cos\theta = 1$) then this form of pinholing cannot take place and the critical dependence on h disappears.

Note, however, that the measurement of a contact angle over a "long" time (such as a few seconds) might be very deceptive if surfactants are used. At these long timescales, the surfactants have time to migrate to the surface and reduce the surface tension. But in the short time-scales of coating, it is very possible that few surfactant molecules have reached the surface during the critical moments involved in pinhole formation, so the actual contact angle might be much higher than measured. For those who cannot measure these time-dependent "dynamic surface tensions", the best advice is to add, if possible, a small amount of a short-chain alcohol (which reaches the surface very

quickly) or to choose a smaller surfactant molecule in preference to a larger one—though size is not a totally reliable guide to dynamic behavior of surfactants.

While on the topic of pinholes, it is worth noting that the most common analysis of their cause is "oil contamination". This analysis is nearly always wrong. The most common causes of pinholes are the dirt-induced holes discussed above and air bubbles in the coating solution. At one point, Abbott worked for a company that was (rightly) paranoid about pinholes. It was almost a sackable offense to think of bringing a can of a well-known oil spray anywhere near the coating line. During one severe outbreak, Abbott convinced the management to let him do a very light spray with this oil onto the coating to see how quickly or slowly things recovered. To his horror, his nervous "light" spray was a blast that covered the entire web and spilled out into the coating enclosure. While starting to contemplate a life of unemployment, Abbott watched the coating recover completely in about 10 seconds. It turned out that this coating was not particularly sensitive to oil contamination. Further investigation revealed the true cause of the pinholes. It was dirt carried in to the coating machine by people like Abbott trying to find the cause of the pinholes. The biggest source of contamination in many machines is from people trying to be "helpful". Every visit to the coating head to tweak some setting is another chance to introduce dust. Far better to have set up a process which does not need tweaking.

Because a clean machine means fewer pinholes and better product, there is a great temptation to carry out a thorough cleaning just before a quality-critical run. Another hard-won truth about coating is that a machine is at its dirtiest just after such cleaning. Scrambling around the machine to do the cleaning puts lots of particles into the atmosphere. Once this truth is grasped, then the team won't panic when the pinhole count goes high when they think (because they've cleaned the machine) that it should be low. If the cleaning has been thorough, the quality will quickly be restored.

Finally, on the subject of cleanliness, just because a machine is in a clean room doesn't mean that it is clean (though if it is not in a clean room it won't be.). Cleanliness is partly to do with cleanroom air flows, but mostly to do with disciplines that everyone abides by. The time a junior operator criticized (senior) Abbott for not following a cleanroom procedure was a key moment in establishing a cleanroom culture for that operation. Abbott was sent out and had to be re-trained before be-

ing allowed to enter. This sent a strong message that these disciplines were being taken seriously.

Because many people think that pinholes must be caused by oil contamination there is strong resistance to alternative explanations. Air bubbles are, for no rational reason, seldom included in the list of possible causes. A simple calculation shows that if a coating solution has equilibrated at temperature T in a relatively cold mixing department, then brought into a warmer coating department with a temperature of T+5°, 126 bubbles of 500 µm diameter can potentially appear in each ml of liquid. They only "potentially" appear; bubble formation requires a seed. It just needs a slight bit of contamination in some pipework or within a coating head to create the seed and, therefore, a stream of bubbles to appear as pinholes, Figure 6.28.

The cure for such air-bubble defects is to ensure that the coating solution is produced and stored at a temperature slightly higher than the coating process—there is then no danger of dissolved air coming out of solution.

The authors' experience over a wide range of coatings on many different machines is that air-induced and dust-induced pinholes are far more common than those caused by oil contamination. However oil contamination is not unknown; Abbott was once called to an outbreak of pinholes that showed a weird pattern towards one side of the web.

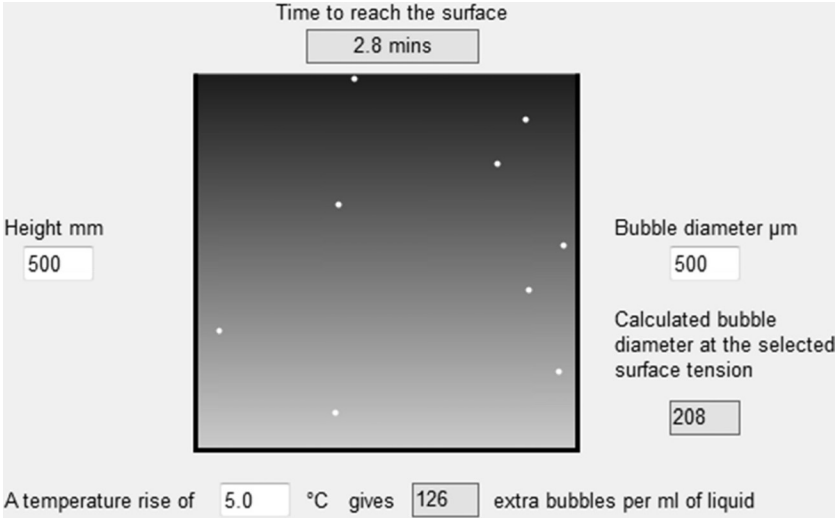


FIGURE 6.28. *Bubbles formed when the temperature of a coating solution is raised by 5°C. These can cause coating defects.*

This turned out to be from a very fine spray of oil drops coming from a contaminated compressed-air line—the first, and last, time he had seen pinholes caused by oil.

6.6.4. Fiber-coating Defects

When coating onto a fiber, a further instability can occur which is similar to the pinholing issue of the previous section [14]. The Pinholes and Fibers spreadsheet performs two calculations. The first assumes that the fiber of a given radius r is moving at speed U through a bath of liquid with a given viscosity μ and surface tension σ, creating a capillary number $Ca = U\mu/\sigma$. The fiber picks up a thickness of liquid h_0 given by:

$$h_0 = 1.34r\frac{Ca^{0.666}}{(1-1.34Ca^{0.666})} \tag{6.12}$$

Unfortunately, the initially-uniform coating around the fiber spontaneously breaks up via the Rayleigh instability, giving beads of spacing:

$$\text{Rayleigh Spacing} = 2\pi\sqrt{2}r \tag{6.13}$$

This occurs in a timescale:

$$t = \frac{12\mu r^4}{\sigma h_0^3} \tag{6.14}$$

In other words, if the coating can be cured/dried in a time less than t, the Rayleigh instability will not have had time to manifest itself.

There is a critical thickness, h_c, below which no instability sets in, if the coating has density ρ and g is gravity:

$$h_c = \frac{1.68r^3}{\sigma/\rho g} \tag{6.15}$$

These instabilities are well-known to those companies who coat fibers all day for a living. They are a complete mystery to those who encounter them for the first time, as it is not at all obvious that a coating on a fiber should in any way be unstable.

6.6.5. What's in a Name?

This discussion about causes of defects leads to a key observation from real life. A typical scenario is that the Production Department phones the technical team to alert them to an outbreak of "repellencies". Instantly, the technical people start asking questions about sources of oil contamination, demand that a sample of the coating fluid be brought to the Analytical Department to test for oil, etc., etc. In the end, the cause will most likely be air bubbles. The point is that the word used to describe the defect, in this case "repellencies", presupposes a root cause (contamination causing the coating to repel from the oil) and well-meaning people rush off in the direction implied by this word.

As was noted long ago by three experts in their excellent book on coating defects, there should be a law in every coating company forbidding the use of names for coating defects until the defect has been properly analyzed (e.g., via a microscope) to identify its root cause [15]. A typical example is that what is called "ribbing" in this chapter has at least 14 names used within the industry: Ribbing; Barring; Coating Lines; Comb lines; Corduroy; MD lines; Phonographing; Puckers; Rake lines; Railroad Tracks; Rib Pattern. Even to say the coating has "holes" can be misleading. To one person a "hole" might mean a perfectly round defect with a raised rim, characteristic of air. To another person a "hole" might mean a "starry night" pinprick. To another, a "hole" might be any shape and size. It might turn out that there isn't a "hole" at all. To the casual observer a light spot in a coating might look like a hole but might actually be a clear bit of polymer gel. So those who ran around trying to find how the different "holes" were formed would completely miss the fact that the problem was one of gel contamination.

Today, microscopes with digital cameras are not at all expensive. Insist, therefore, that at every outbreak of a (relevant) coating defect, a sample is quickly put under the microscope and the image, *including indications of the absolute size*, is included in the email alerting people to the fact that there is a problem. The point about size is important. People can look at an image of a hole and imagine it to be the 100 μm defect that they've seen before, unaware that it is either low magnification (so is a 1 mm defect) or high magnification (a 10 μm defect) which might have totally different causes. Although a single image might still mislead investigators, it is a great advance over a phone call saying "We've got repellencies, come and help".

6.6.6. Mud Cracking

There is one type of drying defect that is very specific to nanoparticle coatings, especially those coated from water. The coating starts to break apart, looking just like mud cracks. This phenomenon has proved puzzling and there have been many explanations. The most convincing one comes from Routhe and is best described with a diagram, where r is the radius of the particles [16], Figure 6.29.

To the right is the wet coating containing the nanoparticles. To the left, the dry particles form a nice coating. The problem arises in between. Just where the particles have a minimum amount of water there is a meniscus of water of radius, r, equivalent to that of the particles. This gives a surface-tension-induced negative pressure proportional to $1/r$. At the wet zone, the pressure is atmospheric, so there is a strong pressure gradient attempting to push water through the drying particles. This flow is frustrated by the close-packed small particles, so the net result is a strong tension across the particles, resulting in cracks.

In this idealized scenario there is not much that can be done to solve the problem. Drying slower gives more time for the pressures to equalize, but has the downside of reducing production rates. Doubling the radius of the particles halves the cracking force, but defeats the specific requirement for small particles. The driving force depends on the surface tension, γ, as $10\ \gamma/r$, which is why mudcracking is more common in water-based than solvent-based systems. Adding surfactants to water-

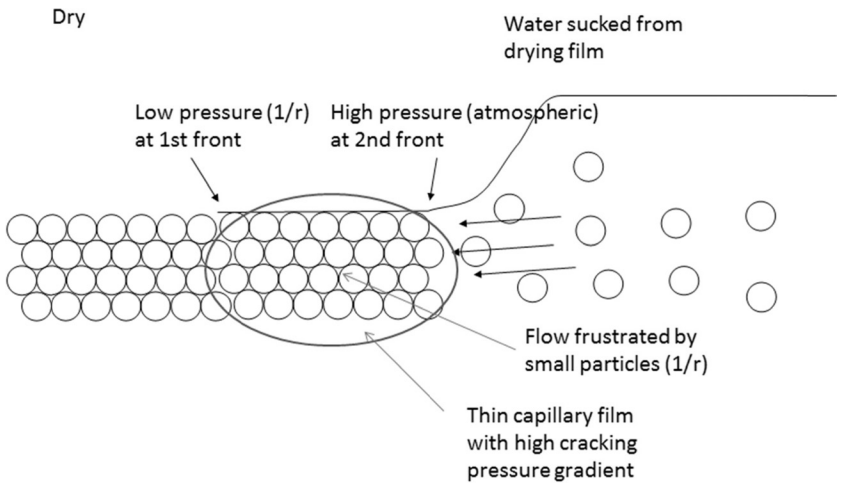


FIGURE 6.29. The origin of mudcracking in nanoparticle coatings.

based systems can in principle help, but these raise other problems. If the shear modulus of the particles is less than 35 γ/r then the particles will sinter rather than crack. So to flip a problem into an opportunity, this capillary-driven pressure gradient can be used to speed up sintering of small, relatively soft, particles.

There is anecdotal evidence that mudcracking is less severe with non-spherical particles. It might also be the case that a small amount of high aspect-ratio particles mixed with the spherical particles will be able to absorb some of the strain and resist the mudcracking.

6.6.7. Slump

Slump is a problem of the growth of a printed feature. The physics is the same as that discussed in the pinhole section, where growth was a good thing leading to healing. Slump is generally a bad thing as it takes a feature of size x immediately after printing and changes it to a size $x + \delta$ where, invariably, δ happens to take your printed feature just beyond the specification limit. As is so often the case with printing problems, there is much vagueness about the causes and cures of slump. One popular but erroneous idea is that it is driven by gravity. In fact, the physics are simple—slump is driven by surface tension—though it is surprisingly complicated to do calculations on specific systems.

The key insight is Tanner's law that the velocity v with which a drop (in this case a spherical cap, typical of a drop of liquid on a surface) of surface tension σ, viscosity μ and contact angle θ, spreads in the same manner as discussed in the pinhole case [17].

$$v = \frac{\theta^3 \sigma}{\mu} \tag{6.16}$$

The problem with applying the formula is that as the drop spreads the contact angle decreases (because the drop is elongated), so a complex integral is required to solve the equation. For real printed dots the assumption that the feature is a spherical cap is also false, so more complex integrals are required. McHale provided algorithms for solving these equations and Abbott has written a modeller App (available from his website) that implements the equations for practical printing cases [18]. A typical example is a line 10 μm high with a nominal width of 50 μm but with an initial 10 μm "wedge" either side of the line, giving a total width of 70 μm at $t = 0$. The liquid has a viscosity of 50 cP,

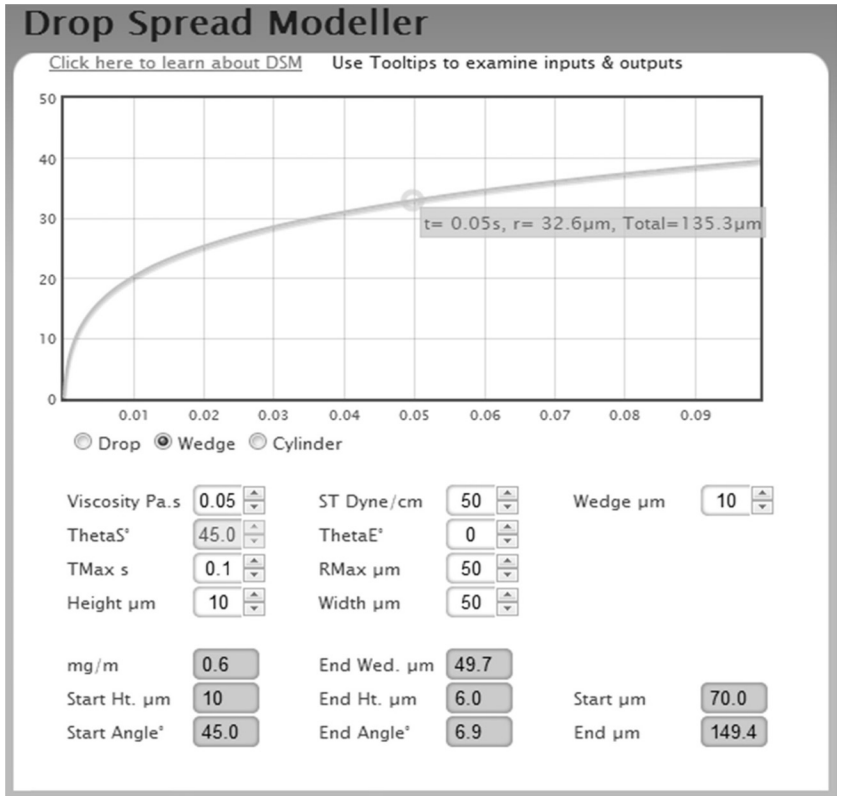


FIGURE 6.30. *A printed line can expand its width rapidly. Here a 70 µm line grew to 135 µm in 50 msec.*

a surface tension of 50 dyne/cm. After 100 ms the line has a total width of ~150 µm, with most of the growth having taken place in the first 10 ms, Figure 6.30.

Because of the dominance of the θ^3 term, increasing the viscosity even by a factor of 10 makes surprisingly little difference. In the above example, the final width after 100 ms is 110 µm. Playing with surface tension (which can only change by a factor of 2 at most) makes even less difference. Reducing the height from 10 µm to 5 µm has an effect equivalent to a 10-fold increase in viscosity. This is due to simple geometry—the initial contact angle is smaller so the velocity is much smaller.

The problem here is that 100 ms isn't enough time for the growth to stop via evaporation of solvent or UV curing. There is only one root-cause cure. The more exact form of the Tanner equation is:

$$v = \frac{\theta(\theta^2 - \theta_e^2)\sigma}{\mu} \tag{6.17}$$

where θ_e is the equilibrium contact angle of the liquid with the substrate. When the drop expands sufficiently so that $\theta = \theta_e$, then the drop stops growing. Hence there is a large expenditure of resource within printed electronics to tune the ink and substrate so the equilibrium contact angle is high enough to stop the slump. Because people "know" (wrongly) that adhesion depends on surface energy there is much worry about the loss of adhesion that will result from this strategy. The discussion in Chapter 5 on the mythology of surface energy and adhesion frees formulators from the irrational fear of increasing contact angle to control slump.

6.7. CASE STUDY: PRINTABLE GAS BARRIERS

Chapter 4 used a nanoclay case study to examine some of the issues of creating polymer/clay compatibility in order to gain barrier properties within a film. This study on the SunBar™ printed barrier (courtesy of Sun Chemical) shows how the combination of printing technology and nanoclays can deliver barrier properties on top of a film [19].

The starting point is the need for extra oxygen barrier performance within low-cost packaging applications. It often comes as a surprise to those who know little about packaging film that it is already a sophisticated product. Multi-layer films are the norm in packaging. A typical example is PE/EVOH/PE where the two outer layers of polyethylene provide excellent barrier properties against water and hydrophilic constituents in the package while the polyethylenevinylalcohol middle layer provides an excellent oxygen barrier (it has a large HSP Distance from oxygen) as well as a barrier to hydrophobic constituents. A mere 5μm of EVOH is all it needs to transform a packaging film from unsatisfactory (PE alone) to acceptable.

The starting point of this project was to gain the barrier properties of a multi-layer film without the complexities (and recycling issues) such films require. Therefore a food-safe, environmentally friendly barrier was required that could simply be printed.

As discussed in Chapter 2, a high aspect-ratio particle is required in order to gain the tortuosity for a good barrier. Clays can naturally provide this if they are fully exfoliated. The systems in Chapter 4 were intended for inclusion in non-aqueous systems, and so had to be ex-

foliated with quaternary ammonium salts. In this project the aim was water-based inks, so sodium bentonite, which is food safe, was used. The bentonite was dispersed in a water-soluble polymer of undisclosed nature but presumably something safe and low-cost such as EVOH. This appears simple in principle, though making a practical flexo ink is not at all simple. Skipping over the complexities, the net result, when printed onto PET (similar results were obtained on other polymers), is a dramatic improvement in oxygen barrier properties, Figure 6.31.

At 50%RH the oxygen transmission rate hardly registers on the scale. Not surprisingly, at higher humidities the barrier properties are somewhat compromised, but not as much as a coating of EVOH.

Such an improvement required considerable optimization work. For example, clays and other particles of different aspect ratios were studied to ensure the maximum barrier properties for the least percentage of clays, Figure 6.32.

Careful studies of temperature/barrier properties proved that the effect really was one of tortuosity and not something to do with diffusion through the polymer matrix, Figure 6.33.

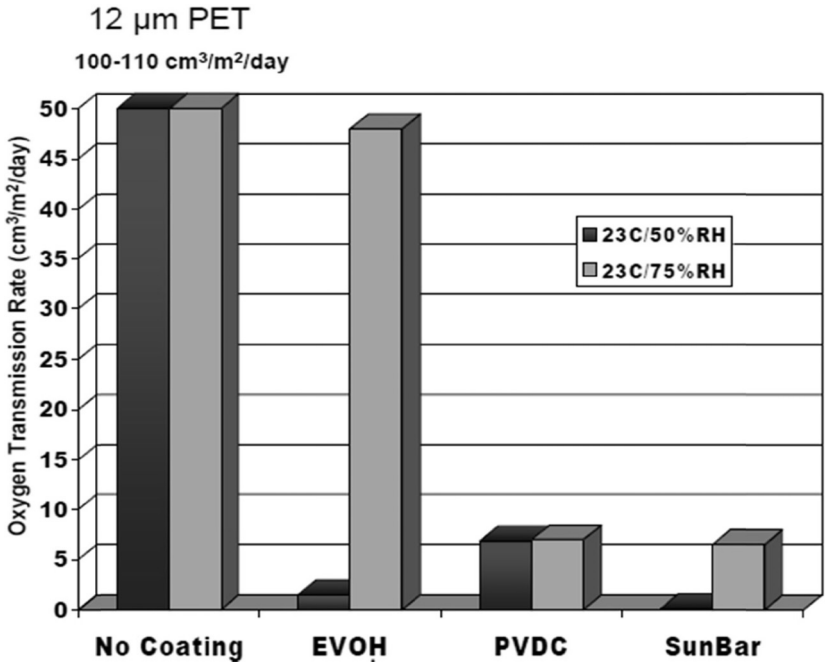


FIGURE 6.31. Comparison of oxygen permeability and humidity effects for various barrier coatings on PET including the nanoclay SunBar.

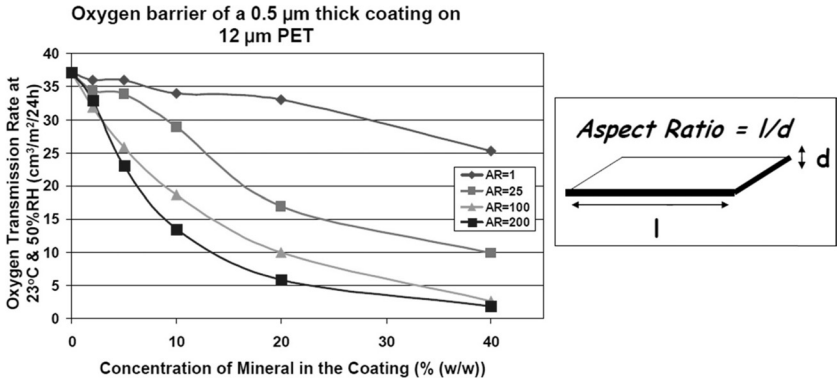


FIGURE 6.32. *High aspect-ratio particles achieve low oxygen permeability at lower overall loadings.*

As discussed in Chapter 2, the quality of the barrier depends strongly on the degree of orientation. For these printed barriers, tests showed that orientation was in near-perfect alignment with the substrate, providing the maximum barrier for the minimum clay.

Why this emphasis on minimum clay? Affordable clays contain impurities which can affect the color and clarity of the final product. Keeping the level of clay addition to a minimum reduces the chances of the final product being visually unacceptable.

The nature of the barrier means that there is no special reason why it should be damaged by flexing. This was confirmed via a set of "Gelbo Flexes"—very strong twists like those of an old-fashioned sweet wrapper, Figure 6.34.

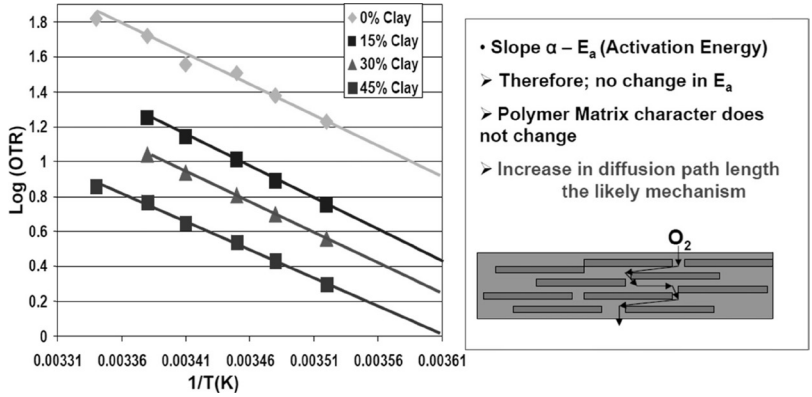


FIGURE 6.33. *A temperature analysis shows that the barrier effect is due to the tortuosity of the clay.*

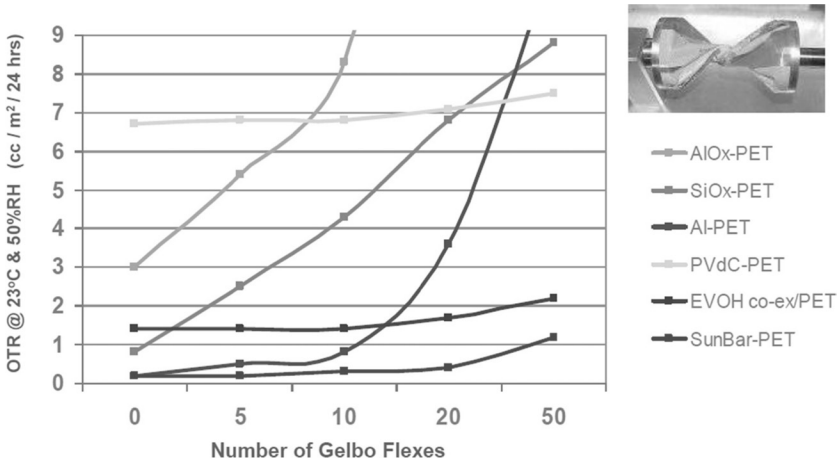


FIGURE 6.34. *The nanoclay barrier retains its properties during flex tests.*

The barrier was much more robust than inorganic AlO*x*, SiO*x* or aluminized film.

The team had one more key job to do. Although the clay they used is food grade, it is required to exist as individual nanoparticles within the coating. The question then arises whether these nanoparticles can pose a health hazard. The answer is a clear "no". One key test carried out (anticipating a discussion in Chapter 9) involved abrading the coating. No individual nanoparticles could be found in the dust—only particles embedded within the matrix and therefore of sizes far exceeding any definition of "nano".

In summary, a focused approach on a particular target (water-based, environmentally friendly, printable barriers) can, through the use of good nanoscience and good formulation, produce a satisfactory product, though one that is not yet perfect (humidity resistance, color and scattering are all being improved via further development). Although it sounds like a relatively simple application of nanocoating science, the development required an enormous amount of work for a well-motivated group with access to excellent test and measurement equipment. It confirms the rule that many things are simple till you have to bring them into production at an affordable but profitable price.

6.8. CONCLUSION

Coating and printing are not as easy as they seem to those who don't

actually have to do them. By definition we always push to the limits in speed, thinness, solvent suitability and so forth because if a coating is easy to produce we know we could make more money by going faster, thinner, using cheaper solvent, etc.

The aim of this chapter is to pass on the key science that will help the team to understand how close they are to any particular limit. There is no shame in hitting a limit if it is known why it is there. The whole team can rally around the science of the limit and decide whether there is an alternative route or whether to accept that no more can be done. What should be a cause of shame is spending time and resources trying to fix a problem that is beyond some fundamental limit and cannot be fixed. Trying to defeat the laws of physics is a singularly unproductive way of spending one's working life.

6.9. REFERENCES

1. TopCoat. www.rheologic.co.uk Accessed June 2012.
2. S.P. Lin. (1981) Stability of a viscous liquid curtain. *Journal of Fluid Mechanics* 104: 111–118.
3. M.D. Savage. (1991) Meniscus instability and ribbing. *Industrial Coating Research* 2: 47–58.
4. J.Q. Feng. (1998) Computational analysis of slot coating on a tensioned web. *AIChE Journal* 44: 2137–2143.
5. Jaewook Nam and Marcio S. Carvalho. (2011) Flow visualization and operating limits of tensioned-web-over slot die coating process. *Chemical Engineering and Processing* 50: 471–477.
6. Konstantinos Apostolou (2012) Slot Coating Start-Up. PhD Thesis. 2007. University of Minnesota.
7. Jaewook Nam and Marcio S. Carvalho. (2009) Mid-gap invasion in two-layer slot coating. *Journal of Fluid Mechanics*, 631: 397–417.
8. Jaewook Nam and Marcio S.Carvalho. (2010) Two-layer tensioned-web-over-slot die coating: Effect of operating conditions on coating window. *Chemical Engineering Science*. 65: 4065–4079.
9. Robert D. Deegan, Olgica Bakajin, *et al.* (1997) Capillary flow as the cause of ring stains from dried liquid drops. *Nature*, 389: 827–829.
10. Steven Abbott. Available on email request steven@stevenabbott.co.uk.
11. N. Kapur, S.J. Abbott, *et al.* (2012) Screen-pinting: A mathematical model for predicting liquid transfer. *Transactions on Components, Packaging and Manufacturing Technology,* (2012), in press.
12. MacDermid Autotype. www.macdermidautotype.com.
13. Frederik C.Krebs, Mikkel Jørgensen, *et al.* (2009) A complete process for production of flexible large area polymer solar cells entirely using screenprinting—First public demonstration. *Solar Energy Materials & Solar Cells* 93:4, 22–441.
14. David Quéré. (1999) Fluid coating on a fiber. *Annual Review of Fluid Mechanics* 31: 347–384.
15. Edgar B. Gutoff, Edward D. Cohen and Gerald I. Kheboian. (2006) *Coating And Drying Defects: Troubleshooting Operating Problems,* 2nd Edition. New York: John Wiley & Sons.
16. Wai Peng Lee and Alexander F. Routh . (2004) Why do drying films crack? *Langmuir* 20: 9885-8.

17. L.H. Tanner. (1979) The spreading of silicone oil drops on horizontal surfaces. *Journal of Physics D: Applied Physics* 12: 1473–1484.
18. G. McHale, M.I. Newton, *et al.* (1995) The spreading of small viscous stripes of oil. *Journal of Physics D: Applied Physics* 28: 1925–1929.
19. Peter Brownrigg. (2010) Lifting the barrier to new technologies. *European Coating Journal* 12: 97–100.

CHAPTER 7

3D Nanocoatings

ANOTHER definition of nanocoatings is one where the surface is nanostructured; such structures can provide interesting properties. Often the coatings into which the structures are formed contain nanoparticles for reasons discussed in other chapters. So we have nano and double-nano.

This chapter covers the essential technologies and includes quite a lot of physics relevant to the application—emphasizing the point that, for any successful nanocoating, chemists must understand what the physicists say and vice versa. It is simply not possible to bring such complex products to market without cross-functional comprehension.

7.1. STRUCTURING A SURFACE

Although there are elegant self-structuring methods such as opalization from close-packed spheres, this chapter will focus on proven techniques that are scalable into mass production. This means in practice that a master structure is provided (typically in the form of a metal or polymer "shim") and the structure is replicated via contact with the master.

The replication can be done by a combination of heat and pressure—old-fashioned thermal embossing or injection molding. The drawbacks of these techniques are severe, so they will not be discussed in detail. The first problem is that there are relatively few polymers (e.g., PMMA and PVC) with the right thermal and flow properties to allow these techniques to be used. It is relatively rare that the surface properties of these polymers are right for other aspects of the product; therefore the choice is between good replication and good surface properties.

The second problem is that high pressures and high temperatures are not generally compatible with subtle and complex nanoproducts.

The third problem is one of cycle time. If the master is kept in contact with the replica while the ensemble is cooled below the T_g of the polymer, then the replica will be of high quality, though the cycle time will be slow. If, on the other hand, the replica is separated quickly to give a fast cycle time, then the chances are that that the structure will relax and lose fidelity.

Fourth, for high aspect-ratio structures (large depth/width ratio), it is very hard to get a highly viscous polymer to flow right to the bottom of the master structure.

Another way to impart a structure is to thermally polymerize a monomer in contact with the master. This is a popular method for making PDMS stamps. It is not practical for any high-throughput method alone, though the PDMS replicas can often be used as excellent masters for subsequent high-throughput techniques.

It seems, therefore, that only one general-purpose technique has sufficient power and flexibility to be viable for mass replication. This is UV embossing. The technique requires only moderate pressures and temperatures while viscosities are low enough for flow into high aspect-ratio structures. Importantly, there are a great number of UV lacquers available, providing the ability to tune the bulk properties and the surface properties in a way that is not possible with thermal embossing. Adding particles is typically not helpful for nanoreplication unless those particles are nanosized. It then becomes possible to create some potent double-nano formulations.

It should be noted that, in principle, similar techniques might be employed using electron beam curing. However, the cost of an e-beam curing line is not insubstantial; so although e-beam offers certain advantages over UV, the authors are unaware of e-beam embossing being carried out commercially.

Between them, the authors have many years of experience with production-scale UV embossing. The principles are well-known within the specialist circle of UV embossers, so these will be discussed in this chapter. Individual implementations of these principles encompass a great deal of confidential knowledge which, of course, will not be revealed here.

7.1.1. UV Embossing

By whatever appropriate means, a thin coating of a UV curable lac-

quer is placed on the surface of a suitable substrate such as PET. The thickness is generally not of great significance, other than in the sense that thinner coatings are, as discussed in Chapter 6, more difficult to produce. The thickness must be large enough to allow the structure to be filled completely and small enough to be readily curable. The coating rarely has to exactly match the structure, leaving (in theory) no excess material. As we will see, this is essentially impossible to achieve; therefore it is generally not worth attempting.

The viscosity of the lacquer is important. In general (and, it turns out, always in your own specific domain), UV monomers and oligomers with desirable properties tend to have viscosities that make them hard to coat and to emboss. Fortunately, their viscosities tend to decrease quickly with temperature, so it is common to apply them at a temperature above ambient. With only modest pressure, the liquid resin comes into contact with the master and flows into the structure.

In the most common situation, the master is a nickel shim, often treated in some proprietary fashion with a release coating. This means that the UV curing must take place through the transparent substrate. As many such substrates are either UV absorbing (PET) or have UV absorbers added for stability (PMMA), it is important to tune the wavelength of the photoinitiator to the wavelengths that pass through the substrate.

At one time, there was no viable alternative to mercury-based lamps, with their attendant broad spectrum (much of it wasted by absorption in the substrate) and large heat output. Rapid advances in UV LEDs have resulted in viable LED arrays, where the power (in total much less than the mercury systems) is concentrated in a relatively narrow wavelength band (compensating to a significant extent for the reduction in total power). This makes it even more necessary to find the right photoinitiator to absorb the LED photons.

Of increasing importance are systems where the master structure (often as a PDMS replica) is transparent and light is delivered through the structure, typically wrapped around a quartz glass cylinder. In some cases, the structure might be etched into the glass. These systems were largely impractical before LED lamps became available, which partly explains their current resurgence. Naturally, this allows the replicas to be made on non-transparent substrates such as metals or silicon for solar and other applications [1].

By whichever way the light reaches the lacquer, the idea is to fully cure it before separation from the roller. Experience shows that failure

to fully cure is a disaster. Not only do you fail to have nice replicas but the (usually expensive and precious) master structure fills up with semi-cured resin that is essentially impossible to clean without damaging the structure.

Most common UV lacquers are sensitive to inhibition by oxygen. One nice feature of UV embossing is that, by definition, oxygen is excluded during the curing process. So lacquers that require super-high powers to cure in ambient conditions can be cured rapidly even with relatively low-powered LED lamps. Curiously, this makes it somewhat difficult to formulate in the lab because there is no need (or ability) to cure the formulation against the master structure. Curing in the open air is not sensible because the right combination of power, photoinitiator and resin system for air curing is not usually relevant to the real curing system. Therefore, lab formulators have to either have access to an inert-gas (nitrogen or carbon-dioxide) blanket on their curing machine or have to create a laminate with a suitable non-adhering film.

One problem that everyone has to wrestle with is what to do during start-up and shut-down of the process. It is likely that at some point of the process there is wet lacquer that does not have the benefit of oxygen-free curing with relatively low UV power. This means that it is surprisingly easy to have uncured lacquer reaching downstream rollers. The experience of the horrendous clean-ups that follow such mistakes is enough to convince a team of the necessity for techniques and procedures to be put in place to avoid repeats of such mistakes. It is strongly recommended that senior managers are asked to join in the clean-up. This helps free-up funds for providing root-cause cures to the problem.

Temperature control during the curing process is vital for four reasons. First, like any process, being "in control" is important. Second, the lamps and the curing reaction can send the system soaring to excessively high temperatures. Third, controlled viscosity is vital for accurate filling of the structure and, as mentioned above, many oligomers and monomers show a strong dependence of viscosity on temperature. Fourth, the quality of the cure is very strongly temperature-dependent. This needs explanation for those who are unfamiliar with this fact.

It seems at first that the most important requirement for a good cure is lots of photons. To a certain extent this is true, but the key point to note is that the curing is generally intended to create a cross-linked matrix. By definition, this makes it hard for components within the matrix to move around. So what happens is that above a certain level of cure, an active radical seeking out a fresh molecule of monomer finds it dif-

ficult to move the extra nm to react with remaining monomer; therefore chain termination takes place instead. The only way to increase the efficiency of the reaction is to increase the mobility of the components, and the only way to achieve this is to increase the temperature.

Those who doubt this should try curing the same acrylate lacquer at room temperature and at 60°C, then run an IR (ATR) spectrum to look at residual monomer. There is usually a depressingly large amount of unreacted monomer in the sample cured at room temperature. Of course, this depends on many aspects of the formulation, such as the functionality of the oligomers used or the percentage of "rubbery" monomers. The general rule remains—curing at higher temperatures is generally a good idea and rarely a bad one.

Getting good release from the master depends above all on obtaining a good cure. Once that is established, release may be good enough without further effort. If release remains borderline then one approach, as hinted above, is to pre-treat the master with a release coating. A typical example from the literature is 1H, 1H, 2H, 2H, perfluorodecyltrichlorosilane, which can be reacted onto an alumina surface (used for antireflection coatings) to produce a surprisingly robust, low surface energy monolayer [2].

Another approach is to build release into the lacquer. If the product requires silicone-style properties, then UV curable silicones give, not surprisingly, good release without further tricks (though this requires considerable formulation skill to work well). For general-purpose acrylates, the only way to get release is to provide an additive which has time to rise to the surface of the coating before reaching the embossing nip. Beyond this point, the additive gets cured into the system. Suppliers of UV curing materials generally have such surface-active additives as part of their range. A small percentage in a coating can give large benefits in terms of release. The trick is to ensure it reaches the surface of the coating before curing—otherwise it adds no value to the process—and that it all becomes entirely locked in to the cross-linked system. Any low-energy materials not locked into (or onto) the coating will tend to be lost with (at best) loss of performance and (at worst) gross contamination of other surfaces, causing coating or printing problems in subsequent steps.

It was stated earlier that it is impossible to create a replica of perfect thickness with no excess lacquer between the replica and the web. In other words, to go from an initial thickness h_1 to a finite thickness h_2 is relatively easy but going to zero thickness ($h_2 = 0$) is impossible, Figure 7.1.

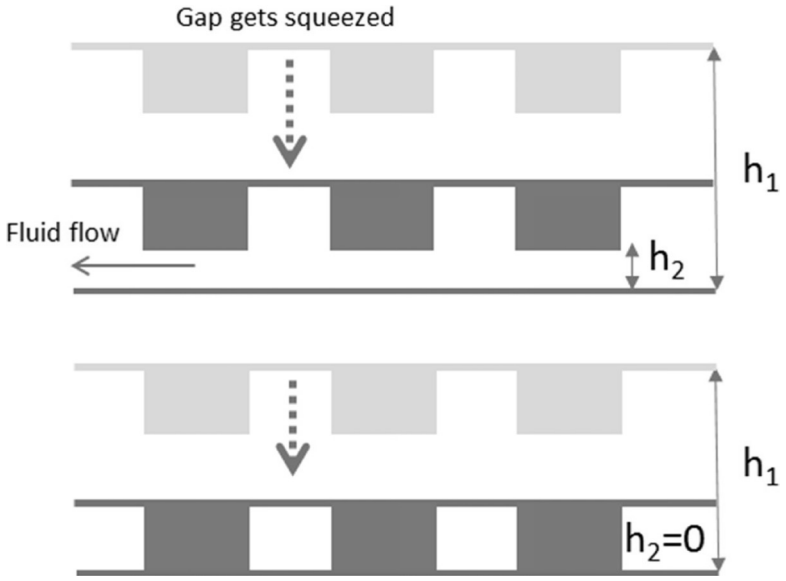


FIGURE 7.1. *The problem of squeezing out the residue liquid.*

Those who are familiar with the Poiseuille problem in Chapter 6 will immediately grasp the issue of the dependence on h^3—though the relevant effect is generally called the Stefan equation (named after MJ Stefan, not the J Stefan of Stefan-Boltzmann) [3]:

$$F = \mu \frac{3\pi R^4}{4h^3} \frac{dh}{dt} \tag{7.1}$$

This states that the force, F, needed to squeeze two discs of radius R together depends on the viscosity μ, the gap h, and the speed of closing, dh/dt. This means that as the master and the substrate get very close together, in order to squeeze out the last remnants of the liquid, the required force heads towards infinity. The integrated version of the equation says the same thing. The time, t, taken to go from initial thickness h_1 to the desired thickness h_2 is given by:

$$t = \frac{3\pi R^4 \mu}{16 f} \left(\frac{1}{h_2^2} - \frac{1}{h_1^2} \right) \tag{7.2}$$

Again, as h_2 approaches zero, t tends towards infinity. The Stefan spreadsheet shows this in action, Figure 7.2.

The spreadsheet allows some crude estimates based on the applicable "*R*" (depending on the contact width of the embosser) along with the time available (from speed and contact width).

A residual layer of 200 nm is generally considered quite impressively small, and can only be achieved in structures that do not have large flat areas (i.e., a large *R* in the equation) and with lacquers of relatively low viscosity. Even then, the process requires either a large force *F* or a slow speed.

The sad consequence of Stefan's law is that many schemes intended to make interesting isolated structures ($h_2 = 0$) on top of a release coating hit the problem that the residual layer ($h_2 > 0$) will link together the individual desired structures, invalidating the whole idea.

But it isn't all bad news. Earlier it was stated that the coating thickness wasn't too critical for the overall functionality. Using the default values in the spreadsheet, going from 10 µm to 140 nm the required time is 23.4 ms. If, instead, the starting thickness had been 15 µm, the required time is still 23.4 ms—the extra 5 µm of coating takes only a few extra µsec to flow.

A final consideration is the issue of the flow of the UV lacquer into the master structure. Done badly, air bubbles can easily become trapped. Once again, viscosity control via formulation and temperature is important. An analysis of this issue is beyond the scope of this chapter, but an excellent summary is available from Scriven's Minnesota group [4].

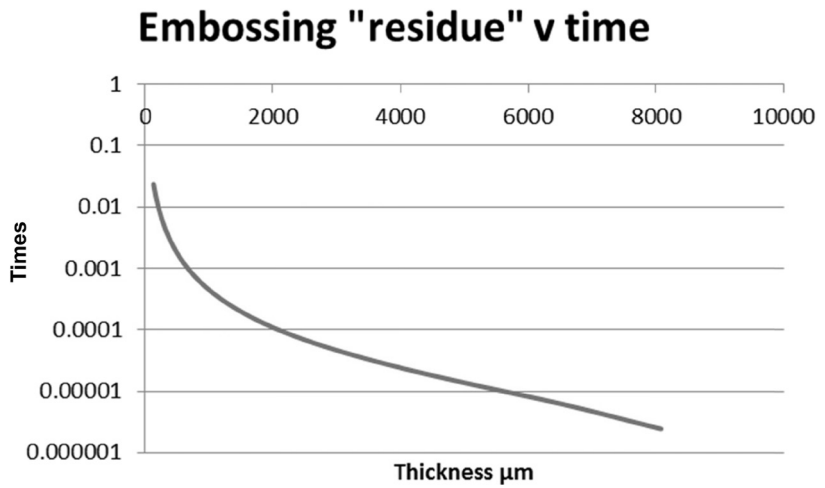


FIGURE 7.2. *The embossing residue versus time showing the rapid increase in time as the residue approaches zero.*

7.2. STRUCTURES WORTH EMBOSSING

Because nature is very good at manipulating nanoscale features, a number of interesting UV-embossed nanostructures have been developed based on such features [5].

Four bio-inspired applications will be discussed as an illustration of what can be done. Although each application has, from time to time, received near-hysterical adulation, the number of commercial products based on them is, at best, highly limited. Beware; just because something is bio-inspired and nano does not mean that it is profitable.

7.2.1. Gecko Effects

One of the many things the world does not need is a "gecko tape" adhesive. Just because a lizard can walk up walls does not mean that it has a technology that is particularly useful for adhesives. The adhesive strength of a gecko is somewhat lower than a Post-it note. If you want a gecko adhesive, go buy a Post-it. This is a rather harsh comment on all the enthusiasm generated by the "gecko effect"; it is not intended as a comment on the truly excellent science carried out by the pioneers in gecko-effect research. It is that research which demonstrated that the gecko *adhesion* (we will come to other smart aspects of the gecko effect in a moment) is nothing more than pure surface energy. The reason the gecko can grip is that its feet have a hierarchy of compliance; the nanotips (~20 nm) at the end of that hierarchy are in intimate contact with the nooks and crannies of the surface [6]. An ordinary pad of rubber ~100 mm × 100 mm in smooth contact with a sheet of glass is more than enough to support the weight of a human. The only reason Abbott did not demonstrate this live during a lecture in Reno was his inability to get sufficient health & safety clearance from the venue. The photo in Chapter 4 shows his lab-based attempt to support his own weight with two simple rubber pads and no extra adhesive. In both cases (gecko and rubber pad) the work of adhesion is ~40 mJ/m^2. A typical Post-it can give 100 mJ/m^2 through extra effects beyond surface energy.

It is this hierarchy of structures that makes human attempt to mimic the true gecko effect largely futile. We can manage the nanostructures at the tips of their feet, but we cannot connect these nanostructures to coarser microstructures, nor can we connect these microstructures to millistructures. We simply don't have access to the technologies that geckos have at their disposal.

The reason geckos show this low level of adhesion is that evolution is smart. At a high level of adhesion the gecko would be able to stick but not walk. The gecko has to be able to free one of its four legs with a flick of its ankle (or whatever the gecko equivalent of an ankle happens to be) so that a crack propagates along the rather weak interface. When it places that leg on fresh surface the next leg must be ready to free itself—and so forth.

So the gecko adhesion is nothing special. The ability to turn the adhesion on and off with a small motion is clever. To find a useful application for this we need one more bio-invention. It turns out that tree frogs invented the idea behind a car tire tread long before humans did. Tree frogs can clamber around wet leaves without slipping. It was long postulated that the mucous within the nanostructures on their feet was some special adhesive which allowed them to stay stuck to the wet surroundings. Via an ingenious technique, Federle and colleagues showed that the viscosity of this mucous was identical to that of water—so it wasn't an adhesive [7]. The mucous turned out to be an irrelevance. The nanostructures existed, as in car tires, to allow the water to flow out so that the rest of the foot could make contact with the surface and obtain the standard 40 mJ/m² work of adhesion.

Consider robotic surgery: if you want to drive a robot along the inner surface of someone's abdomen during keyhole surgery, how would you obtain the necessary grip combined with freedom of movement? A nanostructure that has sufficient compliance to make good contact with the abdomen wall, and a tire tread pattern able to allow the biological fluids to get out of the way seems to do the job very well. Careful analysis by the team of Prof. Neville at the University of Leeds revealed that a structure shown in the diagram was the best compromise between adhesion against vertical gravitational tug and resistance to slipping on the curve of the abdominal wall [8], Figure 7.3.

Assuming such a bio-robot existed and was a commercial success, the volume of nanostructured material required could be made in about 10 minutes on a modern UV embossing system. There don't seem to be many formulation challenges. The surface energy effect is hardly altered by choice of lacquer. UV acrylates are much used in biomedical applications so there are no significant toxicology challenges. The example was included because the authors had some involvement in the early phases of the bio-robot project at the University of Leeds and because the whole gecko effect story from hype to reality is fascinating.

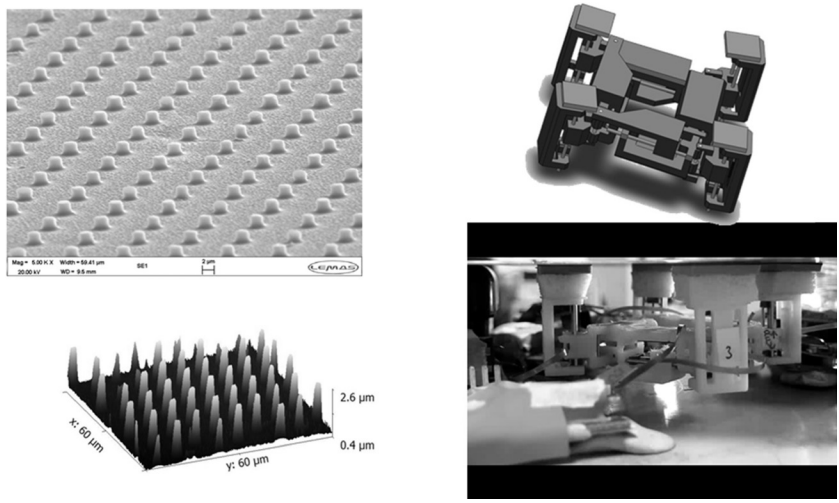


FIGURE 7.3. Some of the basics of the University of Leeds bio-robot. Clockwise from the top left: the optimal structure of the pads; a robotic walker with 4 pads that are also conformal; a lab prototype adhering to a test surface; white light interferometer plot of the optimal structure.

7.2.2. Self Cleaning

Another bandwagon started rolling when the Lotus Effect was announced [9]. The name is very clever, as the Lotus is a symbol of purity partly because the leaves stay clean in the dirtiest of waters. The fact that cabbage leaves show the same effect could have caused it to be named the Cabbage Effect, which would have made it seem far less appealing.

At the time of its discovery it appeared to be a truly remarkable effect. Science was quickly able to catch up and show that the physics were quite simple and (with the benefit of hindsight) obvious. The work of Queré's group provided the key formulae [10]. These describe the actual contact angle, θ^* in terms of a contact angle θ of the liquid on a smooth substrate, the fraction of solid at the surface, φ and roughness, r. φ varies from 1 for a smooth surface down to ~0.1 for a very spiky surface. The metric r is defined as the distance travelled along the surface of horizontal length 1, so a smooth surface has $r = 1$ and the rougher the surface the longer the distance travelled. For example, a 50:50 sawtooth of depth = width has $\varphi = 0.5$ and $r = 3$.

A roughened structure can have more or less wetting depending on the static contact angle. For this particular structure, for a static contact

angle ≤ 78° the effect is called Hemi-wicking—the structure is filled and the measured contact angle is 53°, Figure 7.4.

Hemi-wicking wins over the next effect, Wenzel wetting, when:

$$\cos\theta < \frac{1-\varphi}{r-\varphi} \tag{7.3}$$

and the Hemi-wicking contact angle is given by:

$$\cos\theta^* = \varphi\cos\theta + 1 - \varphi \tag{7.4}$$

A small increase to a static angle of 80° takes the system into the Wenzel wetting region, where the structure outside the drop is not filled and the decrease in contact angle is not so large, Figure 7.5.

The Wenzel contact angle is given by:

$$\cos\theta^* = r\cos\theta \tag{7.5}$$

A roughened hydrophobic (contact angle > 90°) structure can look impressively non-wetting. The static contact angle of 95° is amplified to 105°, still in the Wenzel domain, Figure 7.6.

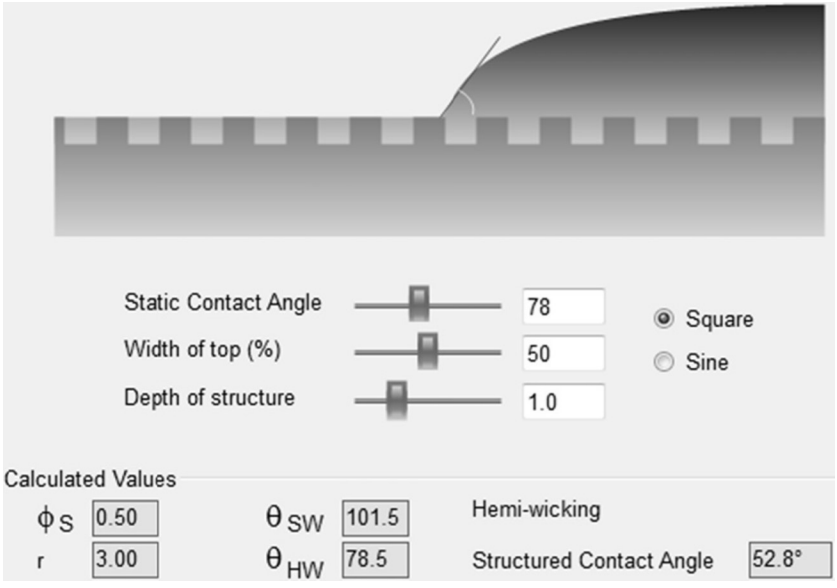


FIGURE 7.4. *Hemi-wicking fills the whole structure.*

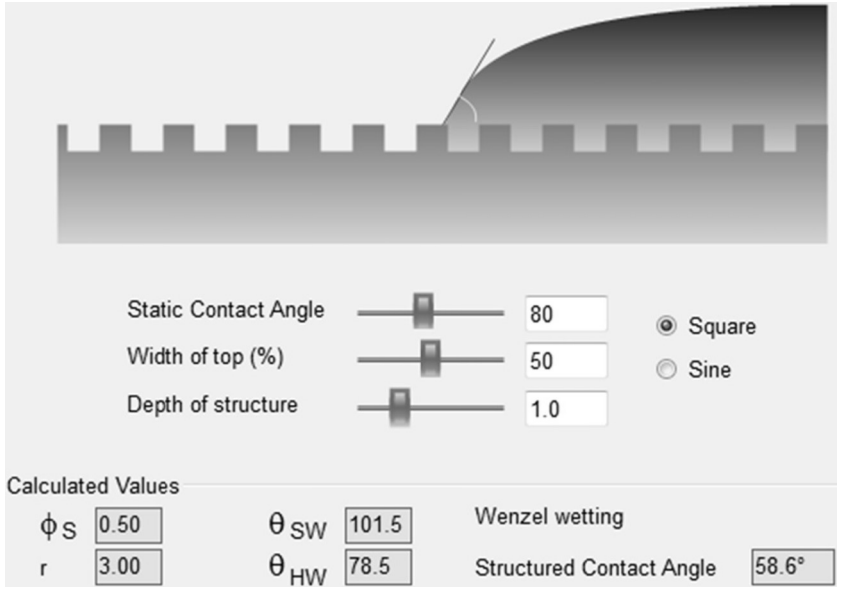


FIGURE 7.5. *Wenzel wetting means that the structure outside the drop is no longer filled.*

At the critical point when:

$$\cos\theta < \frac{\varphi - 1}{r - \varphi} \tag{7.6}$$

the system becomes super-hydrophobic and the contact angle becomes:

$$\cos\theta^* = -1 + \varphi(\cos\theta + 1) \tag{7.7}$$

So with a final push in the right direction (a slight increase in static contact angle and an increase in the depth of the structure), a high-aspect ratio roughened surface with a very hydrophobic surface coating can cause a super-hydrophobic effect, whereby drops of water run off the surface (taking dirt with them) and will bounce if dropped onto the surface from a modest height. Although the calculated value is 126°, in reality the drop would behave like a round ball, Figure 7.7.

To achieve a super-hydrophobic foil in a single pass is a difficult challenge. The structure must be high aspect-ratio, which is not particularly difficult to replicate but very hard to originate. The dimensions don't seem to be critical—1000 nm structures work as well as 200 nm structures—provided the aspect ratio is high and the percentage of the

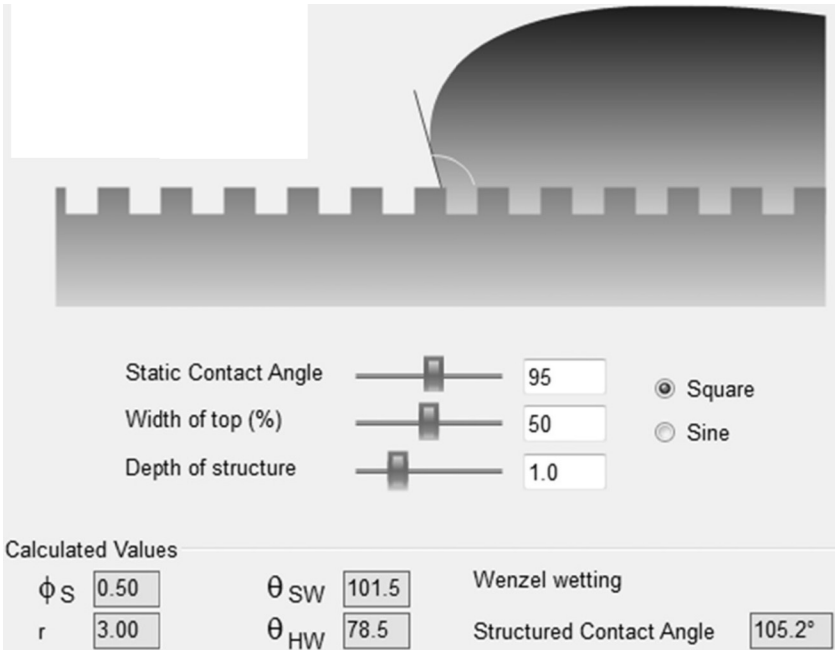


FIGURE 7.6. Wenzel wetting when the contact angle starts greater than 90°.

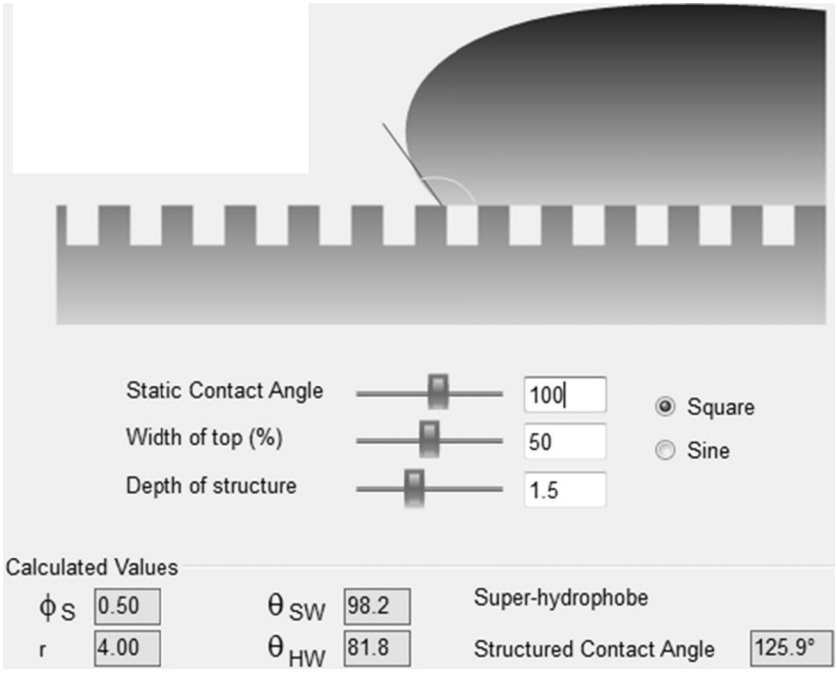


FIGURE 7.7. The flip into super-hydrophobic mode when no liquid enters the structure.

top surface is low. The replicated surface requires an unusually high static contact angle. Achieving that in an acrylate-based system requires the use of fluoroacrylates, but it is achievable. How useful is the product?

Unfortunately, the product is almost entirely useless. The slightest contamination from almost any oil changes the surface contact angle from > 100° to, perhaps, 95°, flipping the system into the non-self-cleaning Wenzel state where the water trapped below the drop prevents the drop from freely moving around. Similarly, if a drop of non-distilled water is allowed to dry on the surface, the salts in the water are sufficient to change the surface energy dramatically, rendering that part of the surface non-super-hydrophobic. Lotus leaves have exactly the same problem but they can solve it by continually creating a fresh surface—not an option for the usual replicated product.

Recently, a variation on the Lotus theme has been announced: the Pitcher Plant Effect, where a lubricant is released into a structured surface, mimicking the effect found in the Pitcher Plant [11]. Though technically ingenious, creating such structures on self-wetting structures on a semi-industrial scale has yet to be attempted. The best that has been achieved to date has been the production of test surfaces on 10 × 10 cm aluminum plates. The coatings were tested to determine their efficacy in preventing ice build-up, and the results were positive [12].

Exceptions are to be found in single-use products, for example in pharma applications, where even a small amount of wetting of a surface might be sufficient to remove or damage a low concentration of some precious biomolecule. A 96-well plate with superhydrophobic properties might be very interesting.

For those who require greater robustness in a super-hydrophobic surface, the answer is paper. Ordinary tissue paper plasma treated with the right fluoro-material shows remarkable, robust super-hydrophobicity [13]. The fibers in the tissue paper are naturally of a very high aspect ratio, so the effect is far more difficult to lose with a minor contaminant. While super-hydrophobic tissues may be an interesting gimmick, super-hydrophobic fabrics are of very real benefit and the plasma process onto fabrics and leather is proving to be a commercial success.

Another way to stay "clean" is to avoid fouling of surfaces by marine organisms. The skins of marine animals such as sharks and pilot whales contain multiple levels of structures. At the coarse level they are able to deliver efficient hydrodynamics (the "sharkskin effect"). At the nanolevel they can provide effective anti-fouling effects [14]. With

the correct nanoparticles that can strike the balance between toxicity to the marine organisms without pollution of the marine environment (see the discussion on antimicrobials in Chapter 8), it is possible to imagine potent surfaces that could be integrated into marine craft.

Finally, surface nanostructures can significantly reduce the crystallization of minerals (such as calcium carbonate) onto surfaces with potential uses (as in the quoted paper) in the oil industry [15].

7.2.3. Anti-reflection

Moths tend to fly at dusk when there is very little light. They need big eyes in order to be able to see. When light goes from air to a medium such as the surface of the eye with a refractive index ~1.5, 4% of the light is reflected. Those big moth eyes would therefore look like car headlights to their predators. Moths needed to go into stealth mode. If moths had had access to vacuum sputtering machines they may well have chosen to put down multilayers of differing refractive index, just as humans do to make their spectacles anti-reflective. Because they didn't have vacuum technology they had to be smarter than humans (who only caught onto the moth's technology two decades ago). They realized that if the surface of their eye was nano-rough, the incoming light wave would see a slow, steady increase of refractive index and, therefore, hardly reflect at all [16], Figure 7.8.

Using either the full Maxwell equation or some less computationally intensive approximations, it can be shown that a sinusoidal structure with a period and depth ~230 nm makes a very effective anti-reflection coating [17]. Going to <200 nm, while still keeping a similar depth removes any tendency to show diffraction colors at oblique angles. UV

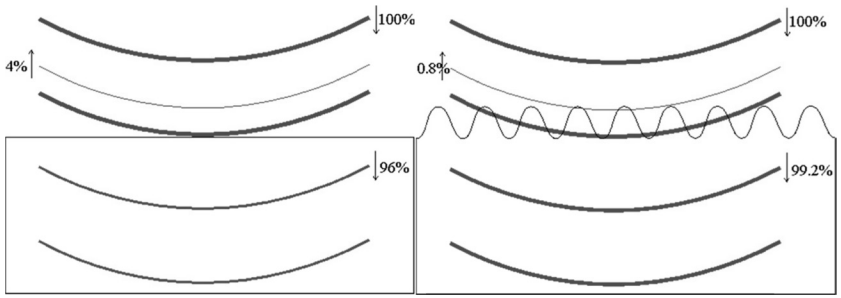


FIGURE 7.8. *The wavefront hitting a planar surface sees a sudden change of RI, so 4% light is reflected. The nanostructure surface on the right is smaller than the wavefront; therefore, it sees only a gentle increase of RI, so the reflectivity is <1%.*

replication can create such structures by the kilometer, so the prospect of vast quantities of motheye anti-reflection coatings has appealed to research teams around the world.

To be able to replicate large areas for modern LCD screens, there is a need for large scale master structures. These can be produced in approximately two ways. The first uses holographic interference to create the structure in a photoresist. While most holographic labs can create uniform beams over 150 × 150 mm, it is extremely challenging when the need is for masters that are 1.5 × 1.5 m.

The second method uses a technology applied to aluminum windows—anodic oxidation. By applying a current in an acidic bath to an aluminum roller, a structure in the range of 100–200 nm (semi-random as there is some interesting self-ordering in these structures) is created with aspect ratios easily exceeding 1. Such a roller can then be used with the UV embossing process to generate the kilometer of motheye [2]. Although this is easy to write, the implementation on an industrial scale is a severe technical challenge, and one small mistake by an operator can destroy a very expensive roller.

Because the human eye is amazingly sensitive to optical effects, whichever method is used, the resulting film has to be essentially perfect in quality over its whole area. A single defect anywhere on the master surface (holographic or anodized) renders the whole production process non-viable and a fresh roller has to be made at considerable expense.

Another problem with motheye anti-reflection is that any grease from the user's fingers can fill the structure, rendering it useless. A quick wipe with a cloth is usually not sufficient to remove the defect because the grease tends to get to the bottom of the high aspect-ratio structure. Instead, the combination of a microfiber cloth and a mild cleaning agent such as an alcohol is required to restore the surface.

To reduce the tendency to soil, additives such as fluoroacrylates can reduce the surface energy, with the unintended consequence of providing a lotus effect in addition to a motheye effect. If the replicating material also contains nanoparticles then the whole is double-bio and double-nano. Holmes has created such a material.

7.2.4. Photonic Crystals

It is well-known that the invention of sophisticated applications for photonic crystals pre-dates mankind by about 500 million years. Photonic crystals are regular arrays of sophisticated regular structures, of-

ten in complex 3D shapes with features below 200nm, that can diffract, refract, reflect or waveguide light in a fantastic variety of ways. Plants, insect, birds, worms have all created sophisticated structures that produce startling colors (think of the Morpho butterfly) along with many more subtle effects, such as a green color when viewed without a polarizer and a blue/yellow distinction when viewed with a polarizer [18]. By contrast, artificial holograms or diffraction gratings that can be replicated by UV embossing appear relatively crude.

The double challenge for would-be replicators of such effects is to find effective, large-scale methods for creating master structures containing any of the many photonic crystal structures and then to find applications that require mass-replication. It is relatively easy to take a small area of, say, butterfly wing and create, then replicate, the structure. It is another matter to do so over large areas with a structure that is of technological value. When in doubt, proposing a sensor is generally sound advice. Because many of these structures show profound spectral shifts with changes in the refractive index of the surrounding medium, it can be readily imagined how a sensor might be created. As with most sensor ideas, the tough part is thinking why this particular sensor mode is advantageous compared to the vast number of other proposed sensors.

It seems inconceivable that the rich store of ideas from nature on photonic crystals will not eventually result in some blockbuster applications. Some clues are provided by the successes of paints and inks containing color-shifting or pearlescent effects, which find applications varying from anti-fraud effects on bank notes right through to automobile paints. Earlier it was noted that Stefan's law makes it impossible to obtain replicas where there is no residual layer. This is an inconvenience for this type of application because it makes it hard to impart a structure onto a continuous film and then to strip it from the film in convenient μm-sized platelets.

7.2.5. Anti-glare

Moving from the bio-inspired and, often, unrealistic, to the mundane but commercially important, anti-glare structures provide the option for double-nano coatings.

The human eye is very sensitive to specular ("mirror-like") reflection and a controlled level of scatter at the surface can be sufficient to weaken the intensity of the reflection as its light is sent off in other

directions. Such a surface can be highly desirable for touchscreens and touchpads, though there is a fierce debate between those who prefer the sharpness and clarity of a truly gloss surface and those who prefer to mask the obvious reflections.

Creating an anti-glare master surface is a combination of art and science. The science is sufficient to specify the sorts of spatial frequencies (few microns) and amplitudes (sub-micron) required for a good anti-glare. The art comes from the fact that only the experienced human eye can judge which samples of a selection of scientifically good anti-glare surfaces are broadly acceptable to customers; this is by no means trivial.

Once the surface has been created, replicating it via UV embossing is straightforward. The real challenge is to make the surface tough enough for the real world. As discussed in Chapter 5, the right nanoparticles locked into the structure via the right reactive dispersants can produce a level of hardness (as defined by a nanoindenter, discussed in Chapter 8) that is difficult to achieve without nanoadditives. The additives can also be used to fix the problem of iridescence (also in Chapter 8) at the interface with the substrate.

7.2.6. Polarizers

LCDs require polarizers. Classical polarizers waste 50% of the light through absorption of the "wrong" polarization. An alternative polarizer that reflects the "wrong" polarization so that it can be recycled has a dramatic effect on the power consumption of an LCD. One way to achieve this is via a "wire grid polarizer". A conducting grid in one orientation acts as an efficient reflector for light polarized in that direction and has no effect on the perpendicular polarization. The principle was known by Hertz in the 1880s for radio waves. For LCD applications "all" that is needed is kilometers of films with a fine grid structure replicated into the surface which can then be metallized within the grid. The problem is that for visible light the grid has to be ~150 nm in spacing with wires that are ~60 nm wide and ~150 nm deep [19]. Origination of grids of these dimensions on the scale required for modern LCDs is a difficult challenge. Replicating such structures at such fine detail requires serious optimization of viscosity and release characteristics of the UV lacquer. Teams with deep pockets, a good patent portfolio and access to the full chain of master development will gain access to a vast market, dwarfing almost every other possibility covered by this book.

7.3. MAKING THE MASTER ROLLERS

It is worthwhile to say a little about the process of making the master rollers. The key to success is having control over every step in the production chain. Without such control the process will fail. For those with the internal resources the whole chain can be brought in house. Others might deliberately choose to outsource many of the steps to ensure they have flexibility. If, for example, you have holographic mastering in house, that is of no use if you require femtosecond laser engraving. The outsourcing route requires a determination of each of the parties in the chain to be honest about issues. It may be satisfying to point the finger of blame at someone else in the chain, but if this doesn't fix the problem then everyone loses out.

In general, creating a sophisticated structure is time-consuming and expensive. If the structure is created onto the embossing roller then the slightest mistake in production means that the whole roller has to be rejected and the expensive process repeated. If the master structure can be easily copied into many submasters which can then be attached to (or made into) a roller, then damage to a submaster is merely an inconvenience as another one can be brought out of stock and put into production.

The chain of production via submasters follows this route:

- Create the master structure by whatever method is appropriate: holographic exposure into photoresist; step-and-repeat lithography into photoresist; direct laser writing into photoresist; femtosecond laser writing onto a suitable surface; diamond etching into copper; etc.
- Place the master (which may first have to be made conductive by spraying with silver or sputtering with gold) into a tank of nickel salts, apply some current and grow a nickel master shim. Carefully pull them apart.
- Passify the nickel surface and grow a copy of the master.
- Repeat, either with the master or the sub-master or both. Repeat with further generations.
- Attach the nickel replicas to your roller by any available means—as a sheet or as a cylinder.
- Alternatively, cast PDMS (or equivalent) replicas from a suitable nickel replica and use those in either conventional or quartz cylinder mode.

Very conveniently, this process from photoresist to multiple nickel

copies was developed and refined by the CD and DVD industry, because that is how their injection molding inserts are made. Refinements to this process include step-and-repeat nickel growing to create a larger master from smaller originals and various forms of laser cutting and/or laser welding to create masters of suitable shape and size.

Realistically, it is very unlikely that any but the largest organization will possess all of these techniques in-house. Clearly, it is vital that all of the parties involved in the project are involved in a regular dialogue, not only concerning their individual parts of the process, but also as to how their procedures might affect downstream processes. If this is borne in mind then the chances of success are reasonable; if not, then failure is guaranteed.

7.4. CONCLUSION

UV nanoreplication is a powerful technique. There is enough knowledge and manufacturing capability in place to be able to create a fascinating array of wonderful master structures at industrially relevant scales. The basics of a replication machine are well-known. Suppliers of UV resins offer many different combinations of materials to be able to satisfy the demands of formulators. The physics of what is required in the nanodomain is generally well-known or, at least accessible. There are hundreds of academic labs generating small-scale nanoreplication ideas and publishing interesting academic papers on effects that cover a few square mm.

What seems to be lacking is the combination of a killer-app and large-scale mastering capability to deliver it.

Truly large-scale motheye anti-reflection masters (via holography or via the alumina route) are becoming available. The lure of wire-grid polarizers is immense. The authors' crystal ball then becomes rather cloudy, other than to say that evolution has been in operation for eons and nanotechnologists have only entered the game a relatively few years ago. The advances made in these few years have been impressive, but there is a lot of catching up to do. There is still plenty of room at the bottom [20].

7.5. REFERENCES

1. H. Hauser, B. Michl, *et al.* (2012) Honeycomb texturing of silicon via nanoimprint lithography for solar cell applications. *IEEE Journal of Photovoltaics* 2:114–122.

2. G. Hubbard, M.E. Nasir, *et al.* (2012) Angle dependent optical properties of polymer films with a biomimetic anti-reflecting surface replicated from cylindrical and tapered nanoporous alumina. *Nanotechnology* 23:155302 (8pp).
3. M.J. Stefan. (1874) Parallel platten rheometer. *Akad Wiss. Math.-Natur.,Wien* 2: 713–735.
4. M. Torigoe, W.J. Suszynski, *et al.* (2004) Study of "embossing" by casting, curing and peeling on a patterned roll. *12th International Coating Science And Technology Symposium,* Rochester, New York.
5. S.J. Abbott and P.H. Gaskell. (2007) Mass production of bio-inspired structured surfaces. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 221:1181–1191.
6. Kellar Autumn, Metin Sitti, *et al.* (2002) Evidence for van der Waals adhesion in gecko setae. PNAS 99:12252–12256.
7. W. Federle, W.J.P. Barnes, *et al.* (2006) Wet but not slippery: boundary friction in tree frog adhesive toe pads. *Journal of the Royal Society Interface* 10:689–697.
8. G.W. Taylor, A. Neville, *et al.* (2010) Wet adhesion for a miniature mobile intra-abdominal device based on biomimetic principles. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 224:1473–1485.
9. W. Barthlott and C. Neinhuis. (1997) Purity of the sacred lotus, or escape from contamination in biological surfaces. *Planta* 202:1–8.
10. J. Bico, U. Thiele, and D. Quéré. (2002) Wetting of textured surfaces. *Colloids Surfaces A, Physicochem. Engineering. Aspects* 206:41–46.
11. Tak-Sing Wong, Sung Hoon Kang, *et al.* (2011) Bioinspired self-repairing slippery surfaces with pressure-stable omniphobicity. *Nature* 477:443–447.
12. Tak-Sing Wong, Sung Hoon Kang, *et al.* (2012) Liquid-infused nanostructured surfaces with extreme anti-ice & anti-frost performance. *ACS Nano* June 2012.
13. www,p2i.com. Accessed June 2012.
14. C. Baum, W. Meyer, *et al.* Average nanorough skin surface of the pilot whale (Globicephala melas, Delphinidae): considerations on the self-cleaning abilities based on nanoroughness. *Marine Biology,* 140, (2002), 653–657.
15. W C Cheong, A Neville, et. al. Using nature to provide solutions to calcareous scale deposition. *SPE International Oilfield Scale Conference,* 28-29 May 2008, Aberdeen, UK. Downloadable from http://www.stevenabbott.co.uk/Nanostructure-Formation.html, accessed September 2012.
16. S.J. Wilson and M.C. Hutley. (1982) The optical-properties of moth eye antireflection surfaces. *Optica Acta* 29:993–1009.
17. A. Gombert. (1998) PhD Thesis. Université Louis Pasteur, Strasbourg.
18. Pete Vukusic and J. Roy Sambles. (2003) Photonic structures in biology. *Nature* 424:852–855.
19. D.P. Hansen and J.E. Gunther (1998) *Polarizer Apparatus for Producing a Generally Polarized Beam of Light.* U.S. Patent 6,108,131.
20. R.P. Feynman. (1960) There's plenty of room at the bottom. *Engineering and Science* pp. 22–36.

CHAPTER 8

Have the Nanoadditives Done Their Job?

IT takes a lot of work and cost to get nanoparticles into coatings. The question in this chapter is whether it is worth the effort. Have those nanoparticles delivered knock-out performance that makes the effort worthwhile? Or are the results (and this is often the case) only modest improvements that are not really worth the hassle?

In some cases the success of the nanoparticles is obvious and unambiguous. In others it is much harder to prove.

8.1. OPTICS

Optics is an exact and exacting science. There is usually no place to hide.

If the coating is intended for a display application or for a decorative finish on high-quality flooring, then clarity is of utmost importance. As discussed in Chapter 2, light scattering arises from a mixture of refractive index mismatch and particle size. The larger the refractive index difference the smaller the particles must be to avoid haze. This means that the primary particle size must be small, that dispersion in the formulation should be excellent and that effects during drying/curing should work together to avoid clumping. Although haze and gloss meters can unambiguously indicate failure, success is rather harder to quantify.

Ultimately, the human eye is the judge of whether the haze is acceptable to the potential customers viewing the final product. In the highly demanding world of films for laptops and tablets, it is possible to devise one's own in-house viewing test for such low levels of haze,

but there is no substitute for spending 30 minutes at the customer's test facility where highly-trained staff gaze daily at samples distinguishing the subtle differences between good and bad products. If possible, bring back a stack of "pass" and "fail" materials and immediately show them, unlabelled, to senior managers. They will be totally incapable of seeing any difference and so might show a little more sympathy towards the challenges of developing and manufacturing products to high optical quality standards.

Viewing conditions are critical and must be strictly controlled. Samples need to be viewed in transmission via diffuse backlight, linear backlight (i.e., a fluorescent tube) and also spot backlight. They also need to be viewed in reflection via diffuse frontlight, linear frontlight (i.e., reflection from a fluorescent tube) and spot frontlight. Viewing conditions need to be changed from close-up to distant. Each viewing condition reveals different types of defects, though only after careful training. Some people seem to be better than others at spotting defects. Abbott's ability to understand the causes of some optical defects was severely hampered by his inability to spot many of them. Holmes, on the other hand, could spot most defects instantly. Once spotted, however, the defect becomes etched on the retina and memory. On-line inspection systems have the same issues—they need dark/bright field, low/high angle, low/high wavelength and lots (months) of training.

8.2. REFRACTIVE INDEX

If the nanoparticles have been added in order to obtain a refractive index of 1.623 and the RI of the basic coating is 1.456 then a measured value of 1.623 is definitive proof that the nanoparticles have done their job. Getting there should be easy. Using either the volume fraction approach:

$$RI = \sum_i VolFraction_i . RI_i \tag{8.1}$$

or the Lorentz-Lorenz approach (a more complex formula that often arrives at the same result—see the RI spreadsheets), the refractive index should be predictable from the known quantities and parameters of the ingredients. Failure to reach the theoretical value is either a failure of formulation or failure of the assumptions behind the theories. Either way, it is vital to measure the RI of the coating.

The most powerful way to do this is via ellipsometry. A polarized

light beam is reflected from the coating and a measurement of the light reflected in the p- and s- polarizations is made. If everything about the coating (substrate, other layers, thickness of the layer of interest) is known, then the refractive index can be calculated directly from the ratio of p- and s-intensities via a process of data fitting. If, as is usual, the thickness is not known, then more data (e.g., at different angles) is required and the fitting algorithm has to work somewhat harder.

In practice, the whole process is simplified by coating onto a known, standardized substrate such as aluminum or silicon, so it is relatively easy to disentangle the effects of thickness and refractive index. Modern ellipsometers can do the measurements automatically over a wide wavelength range, making it far easier to unambiguously determine both thickness and refractive index, both real (pure refraction) and imaginary (absorption).

An alternative method is via surface plasmon resonance, where the refractive index of the coating influences the resonance signal via evanescent waves. It is possible using this method to obtain refractive index and film thickness results quickly and simultaneously from coatings on flexible substrates.

Ellipsometers and surface plasmon resonance machines are not common in typical coating labs, although it is not uncommon to hire time on one in a university lab. An alternative technique can lead to "good enough" measurements using just a UV-Vis spectrometer and the principles of optical interference. When light is reflected from the lower surface of a coating, it can interfere with the light reflected from the top surface. When the optical path is one quarter of the wavelength of light, $\lambda/4$, then the reflected light will be exactly $\lambda/2$ out of phase and produce zero reflected intensity. The simple technique follows directly from this basic fact of optics.

A coating of approximate known thickness in the range of a few 100nm is made onto an optically thick sheet of glass. The example below assumes that the RI of glass is 1.500 and that the spectra are calibrated against the uncoated glass. For convenience, the reflection spectrum is shown but for those who only have access to the more-usual transmission spectrum, the reflection spectrum can be calculated by subtraction from a reference.

The aim is to check whether the RI is the desired 1.623. Using any convenient thin-film modelling software (the examples shown are from an App available from the Abbott website), a simulated spectrum is created. There are multiple peaks because the thickness is large compared

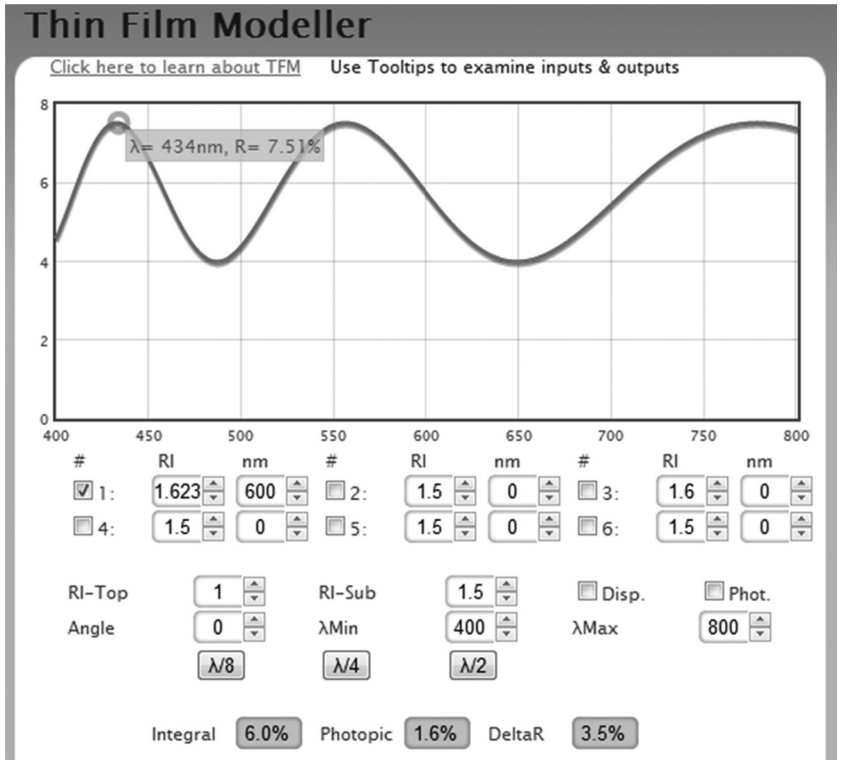


FIGURE 8.1. *The reflectance spectrum from a thin coating of RI = 1.623 and thickness 600 nm.*

to $\lambda/4$, the thickness required for a reflectance minimum. On the plots, the mouse readout shows the key values, Figure 8.1.

Here, it is assumed that the RI (n in the software) is 1.623 and that the thickness is 600 nm. If in fact the thickness was 700 nm the curve would look like, Figure 8.2

Now the λ max value is 414 nm but the maximum %R is unchanged. To put it another way, the effect of errors in the estimated coating thickness is that the number and positions of the peaks in the spectrum change but, importantly, their maxima and minima are unchanged. Going back to 600 nm, the plot now shows what happens if the RI is 1.61, Figure 8.3.

The position of the peaks has changed slightly (the first maximum has gone from 434 nm to 430 nm) but as we now know, peak positions also depend on thickness, so they are not much help. The minima of both curves are unchanged at 4% because, independently of the thickness or the refractive index, the perfect interference minimum (thick-

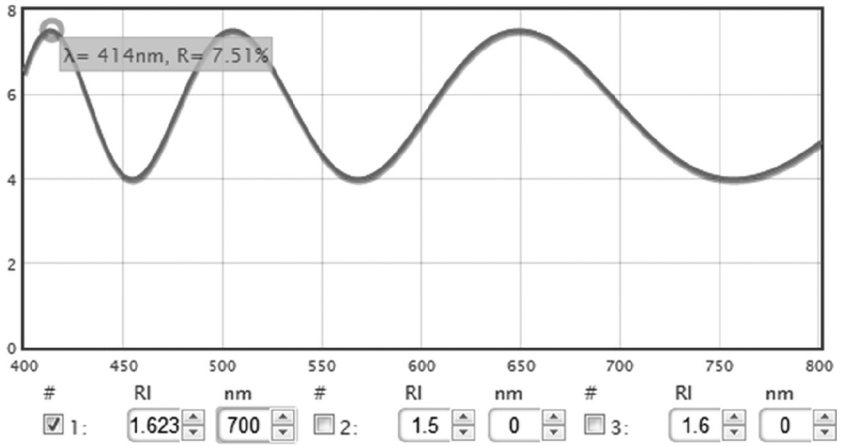


FIGURE 8.2. *The same spectrum but with thickness 700 nm.*

ness = $\lambda/4$) gives the reflectance equivalent to the RI of glass (hence the need to have calibrated against uncoated glass). So peak positions and minima are no direct help. However, the maximum reflectivity (thickness = $\lambda/2$) has changed from 7.51% to 7.14%. If the real data show a few peaks (so the thickness is about right) and minima of 4% (so the calibration is OK) and maxima of 7.28%, then it is easy to find that the RI is 1.616.

There are some complications due to "dispersion" effects (i.e., the refractive indexes change with wavelength) but these can be taken into account via options in the thin film software.

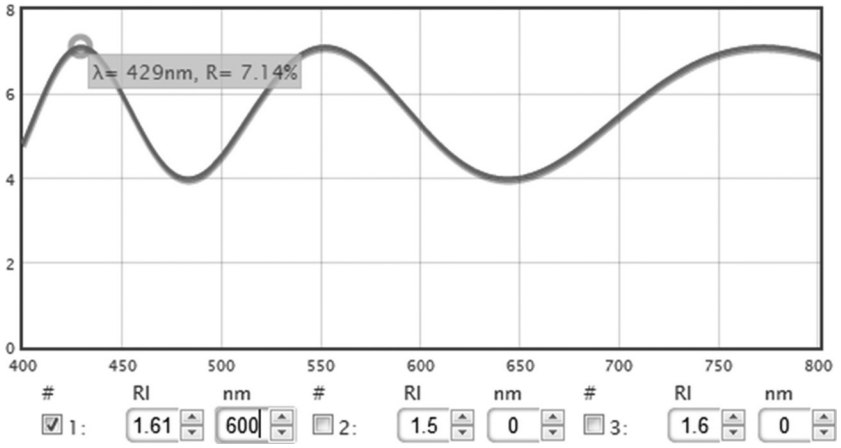


FIGURE 8.3. *A change of RI has a bigger effect on the spectrum.*

The above techniques work not just for the high RI samples. Coatings with low RI made from air-filled nanomaterials such as silica which are used for anti-reflection coatings can be examined in exactly the same way. For these systems, the measurement challenge is to ensure that the air pockets don't accidentally get permanently filled with liquid during processing. For example, if they are dispersed in a UV-curable system, any monomer that seeps into the pores will increase the overall RI and decrease the anti-reflection effect. Although making porous nanoparticles is relatively easy, making those pores totally closed, batch after batch, is a tougher challenge. One approach being applied successfully to single-layer anti-reflection coatings is to ensure that the pores, although open, are covered with highly hydrophobic groups so the liquid does not enter the pores [1].

Given that nanocoatings with specific refractive indexes are most often required for anti-reflection and anti-iridescence, it is fortunate that the same spectrometer which measures the effectiveness of the final product can be used to get "good enough" estimates of test formulations with very little effort.

8.2.1. Anti-iridescence

The same spectrometer is also helpful for that most subtle and annoying of defects in hardcoats—iridescence. In normal light, a coating might look perfect; yet, in some types of light (e.g., fluorescent or LED lights with sharp phosphor emission lines), the coating is covered with iridescent hues. The cause of the iridescence is the same as that of the colors of an oil slick—interference from a refractive index mismatch between the coating (oil) and substrate (water). For oil (RI~1.4) on water (RI~1.3) the interference is intense because of the large mismatch of refractive index and, therefore, reflectivity and also because the thin oil layer gives perhaps just one interference peak. For coating iridescence the RI difference is often smaller (though for coatings onto PET the differences can be similar: 1.68 to 1.53), so the interference is of low intensity, distributed over more interference peaks (if the coating is thicker). Using the same setup as in the RI measurement example, but changing the RI of the coating to 1.51, gives a peak difference in reflection of just 0.3% compared to the 2.4% in the case of oil and water. In white light, the iridescence is spread across the whole spectrum and is not easily visible, whereas in lights with just three sharp emission peaks (red, green, blue) the iridescence is concentrated in narrow bands and therefore more visible.

The cure for iridescence is either a perfect match of RI between coating (e.g., tuned by nanoparticles) and substrate, or an extra interference coating between the layers that cancels out the effect.

To formulate for such a degree of cure is difficult without an objective measure of improvement in iridescence. The (trained) human eye is the ultimate means of assessing iridescence but rather poor at providing numerical values of improvement. The best way to get objective measures of progress towards lower iridescence is via a spectrometer. Iridescence shows as regular fluctuations (as discussed earlier) in the transmission or reflection spectrum of the coated product and the amplitude of the fluctuations provides the objective measure—the smaller the amplitude, the lower the visual effect is likely to be. Complications come from interference peaks created by the substrate itself, so it is important to get to know the optical characteristics of the substrate, either by using the same measurement conditions or by using an uncoated substrate as the reference beam in the spectrometer.

If nanoparticles can be added to the coating to give a perfect RI match then the iridescence will disappear. If that is not possible, then a thin film modeller will readily guide the formulator to a cure via an intermediate layer. For example, a 2.5% fluctuation in reflectivity from a mismatch of a 500 nm coating RI = 1.5 on a substrate RI = 1.6 can be halved by an 80 nm layer of RI = 1.55.

8.3. THE RIGHT THICKNESS

Assuming that the nanocoatings are relatively thin, then measuring thickness is surprisingly difficult. The classical method of measuring a cross-section with a microscope works adequately down to ~1–2 μm; it is, however, tedious, error-prone and not a great way to carry out measurements over a large area.

If the coating has some obvious absorption or fluorescent signal (it is possible to add minute amounts of fluorescent marker), it is possible to create suitable off-line and on-line probes. For organic electronics there are usually very strong absorption bands, so a fiber-optic spectrometer working in reflection or transmission (both have their strengths and weaknesses) can do an excellent job of on-line measurement.

If there is any significant degree of scatter, then optical methods quickly become problematical. How, for example, to distinguish between the haze caused by a thicker sample and one with slightly bigger particles?

In reality, most practical thickness measurements in this range rely on the coating being sufficiently transparent and non-scattering for interference effects to be readily measurable via a spectrometer. The thicker the coating, the more interference peaks and troughs are produced over a given wavelength range. There is a simple formula (used in the Thickness spreadsheet) for calculating "thickness" (*nd*) from the number of peaks, *k* between wavelengths λ_1 and λ_1:

$$nd = k\frac{\lambda_1\lambda_2}{2(\lambda_2 - \lambda_1)} \tag{8.2}$$

The "thickness" is shown as nd because it is the optical thickness (or optical path)—physical thickness *d* times refractive index *n*. To know the physical thickness *d*, the refractive index *n* has to be known via the technique discussed earlier. For many applications, *nd* is the desired

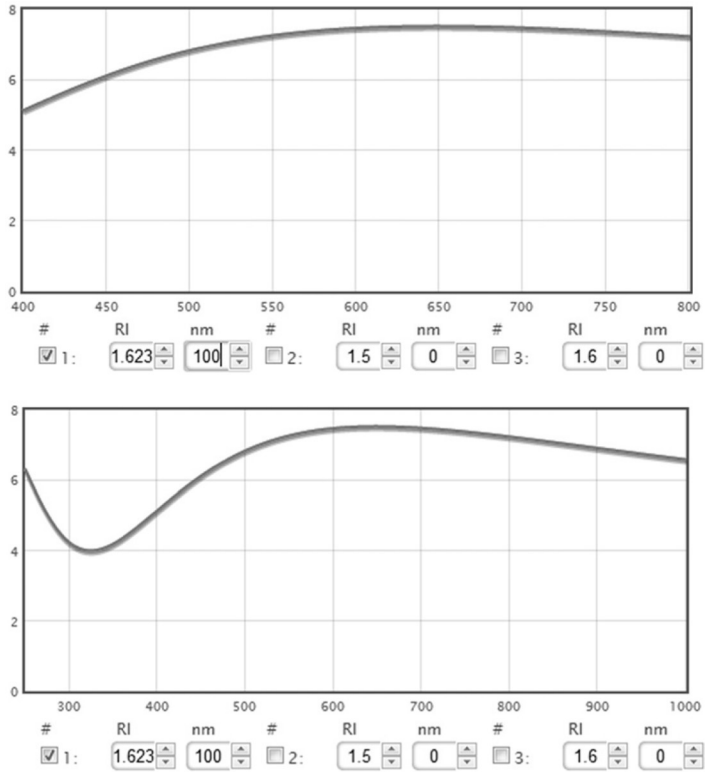


FIGURE 8.4. *Thin coatings need a wider spectral range. 400–800 nm (top image) is too small, so ideally a 250–1000 nm range (bottom image) should be used.*

quantity, so during coating small changes in n can be accommodated by corresponding changes in d to give the correct number, k, of interference peaks.

In reality it can be difficult to count k directly, especially with the noise of on-line measurements. If there are multiple layers then there are two sets of interference peaks superimposed, making direct counting even harder. Fortunately, simple Fourier transform techniques make it possible to filter the data for noise and to pick out the signature(s) of the layer(s).

Of course, specialist thin-film thickness measurement systems are available for off-line and on-line analysis. The larger the wavelength range of the system, the more accurate and versatile it can be. For thick coating, the peaks can become too numerous in the visible range but can be readily counted in the 1000–2000 nm range.

For thin coatings there aren't enough peaks in the 400–800 nm range (left image), so going to a range starting at a lower wavelength (250–1000 nm, right image) is necessary. Of course, absorption phenomena start to complicate things in the sub-300 nm range, but generally these effects can be accommodated in the calculations, Figure 8.4.

8.4. ANTIMICROBIAL BEHAVIOUR

Proving that the final product, such as a coated surface, is antimicrobial is a lot harder than proving that the nanoadditive is antimicrobial. The additive can be quickly tested in standard bio-assays on multiple microorganisms. With respect to proving that the product is OK, there are multiple problems that have to be addressed.

The first problem is complex. Because antimicrobials are viewed with much suspicion by users and regulatory agencies, making sure they are locked firmly into the coating seems a good idea. It is not too difficult to devise tests to show that the antimicrobial remains nicely in the coating after extended challenges with time, temperature and humidity. The downside of this is that the antimicrobials will have no chance to go out and kill some microbes.

The other extreme is to have designed a wonderfully open structure so that the antimicrobials pour out and overwhelm the microbes. Given the proven efficacy of the antimicrobial and the proven degree of pouring out, it will not be hard to prove overall antimicrobial efficacy in the product. The obvious downside is that the antimicrobial material can be transferred in large quantities to humans and the stock in the coat-

ing can be depleted rapidly until there is no longer any antimicrobial left. A surface that once reliably killed all known germs then becomes worthless.

For regulatory and life-time purposes, the bias of the formulation should be towards the locked-in position. How, then, can one demonstrate that the nanoparticles are coming out at just the low, safe rate that also kills the microbes? That is a key question for the team to answer.

The next problem is somewhat related to the first. If the nanoparticles are reasonably locked in and immobile, then a small layer of dirt (or biofilm) will ensure that they cannot migrate to the surface and cannot, therefore, kill any microbes. Users have to be aware that regular cleaning is required in order to allow the nanoparticles to attack the microbes. Arguably, therefore, the user should rely just on the cleaning and avoid the double concerns about nanoparticles and antimicrobials. If, on the other hand, the particles are relatively mobile, they will be able to penetrate the dirt and remain active. Now the alternative concerns about spread of the antimicrobial and loss of efficacy again become the focus.

The third problem is that there are no good tests for proving antimicrobial efficacy. This might sound surprising, but think of the problems. What are the test circumstances? A pristine technical surface specifically dosed with a known amount of a known microbe? A real-life surface containing unknown quantities of unknown microbes? A simulated real-life surface somehow dosed with whatever microbe is known to thrive in that environment and which also poses a risk to human health? This third test sounds to be the most relevant, but it also sounds far too complex. The second test is probably irrelevant—what possible pass/fail criteria could be adopted? So the first test, although probably irrelevant for real-world use, becomes the default.

If this first test is to be used, then how much of which microbes should be put onto the surface? How long should they be left to die? Would they have died anyway on that naked surface largely free of nutrients? If the test microbes are placed using a nice nutrient broth, will a nice thick layer of broth allow the microbes at the top to thrive while the slow-moving antimicrobial kills a few at the bottom of the pile?

The fourth and final problem is that in some markets, such as USA, making claims of antimicrobial efficacy is not something to be done lightly. The cost of proving to the FDA that the antimicrobial system is efficacious and safe probably far exceeds any possible profit from the product. There are good reasons why these nations have a jaundiced

view of antimicrobials. There are multiple risks. Users might rely on them (rather than old-fashioned cleanliness) when they aren't, in fact, efficacious. Or they might promote the evolution of strains that are resistant to the antimicrobial, reducing the chances of being able to fight them when really needed. Finally, putting it very bluntly, most of the time we don't need antimicrobials. Humans have been living happily with surfaces swarming with bacteria, the vast majority of which do us no harm at all and many of which are positively beneficial to us ("friendly bacteria"). Why should we be so keen on killing them?

In summary, make sure that the whole team understands that there is a big difference between having a really clever nano antimicrobial and having a really successful antimicrobial product. Don't even think of starting on such a project till the whole team understands where they want the balance of antimicrobial to be (locked or flooded), how they are going to test for the right degree of locking in, how they are going to test for antimicrobial activity and how they are going to answer the questions of regulatory agencies who really don't want large amounts of chemicals that kill microbes being used in an uncontrolled manner in homes or hospitals.

8.5. THROUGH-COAT DISTRIBUTION

Sometimes it is vital that the particles are evenly distributed through the coating; sometimes it is best if the particles are concentrated at a surface. For example, if the whole coating has to be conductive then the particles must run through the whole coating layer, but if surface conductivity (anti-static) is needed, any particles within the coating are wasted. Similarly, for slip coatings, anything not at the surface is a waste.

Measuring such a distribution is generally hard. Taking cross-sections and checking under an SEM is splendid for fundamental work and for validating other techniques while not being generally practical for the lab formulator or the production QC team. Pure surface techniques such as SIMS and ESCA are, in the authors' experience, more talked about than used; the ratio of cost and complexity to valuable data is generally not satisfactory.

Given that most particles have distinctive vibrational spectra then ATR (Attenuated Total Reflectance) spectroscopy and/or confocal raman can be used to get some idea of what is going on, with the obvious limitations that ATR really only gives some sort of average over the top

~1 μm of coating (depending on the crystal used—Ge is ~0.66 μm, Si/ZnSe is ~ 0.85 μm and ZnSe is ~2 μm) while confocal raman, which can take images throughout the coating, again only creates averages over ~1 μm.

With clever tricks and more expense, the vertical resolution can be improved. For example, in ATR the effective penetration depth of a beam of the correct polarization at the highest possible angle with the highest refractive index crystal is 0.32 μm. With an automated way to adjust the angle, the spectrum over effective depth ranges from 0.32 μm to 1.6 μm can be gathered, giving many clues about the constancy, or otherwise, of the formulation. Similarly, clever optics and a low-wavelength laser can improve the vertical resolution of confocal Raman. For the very brave who have pico-second or femto-second IR lasers at their disposal, Sum Frequency Generation spectroscopy can tell you exactly what is present right at the surface.

If the particles have a very different refractive index, then the interference techniques discussed for thickness measurement can spot the difference between an homogenous and layered system—the latter giving multiple interference peaks.

8.6. SURFACE ENERGY

For some applications such as anti-fingerprinting and slip, it is necessary to achieve a controlled low surface energy. Measuring the surface energy is therefore necessary, even if there is no necessary correlation with the desired effect (friction depends on degrees of inter-surface contact more than it does on surface energy). The contact angle is also used as a surrogate measure of anti-fingerprinting, though again there is no obvious reason for this. On classical anti-reflection coatings, some components of finger grease can migrate through a high contact angle coating and react with one of the underlying layers, changing the RI and so reducing the AR performance. This means that the only real test is with real fingerprints left on for real times. Similarly, for surfaces that need to be printed, a surface energy sufficient to allow ink spreading is required, even if there is no necessary correlation with the subsequent adhesion of the ink. The insistence on “no necessary correlation” is explained below.

Generally, a simple water contact angle test will suffice. This actually tells you nothing direct about surface energy (because the angle also depends on a balance of the solid/liquid, liquid/air and solid/air en-

ergies—2 unknowns with one datapoint); it is merely a convenient test where higher contact angles are generally associated with lower surface energies. With automated machines now commonplace, the days of fiddling with a manual goniometer should be over and such measurements are routine. Given such machines, it is not much extra work to get some slightly deeper insights into the surface energy by measuring the contact angle with two liquids (e.g., bromonaphthalene and water) or with three (add in ethylene glycol or formamide). The software provided with the machine can then calculate values via Owens-Wendt (geometric mean), Wu (harmonic mean) or with three solvents Lewis Acid/Base methods. The Surface Energy Calculator App (from the Abbott website) also lets you do the calculations, Figure 8.5.

In that example, the total surface energy is 62.6 dyne/cm or 68.9 dyne/cm depending on the model, and the dispersive component is ~47.8 with a disagreement about the polar component, which is either 14.9 or 21.1. For the acid-base model, the dispersion value is called the Lifschitz -van der Waals interaction. Although it contains dispersive components, it is calculated rather differently and does contain dipole contributions (among others). None of these numbers should be taken too seriously. It is, however, important to be *consistent*: do not mix

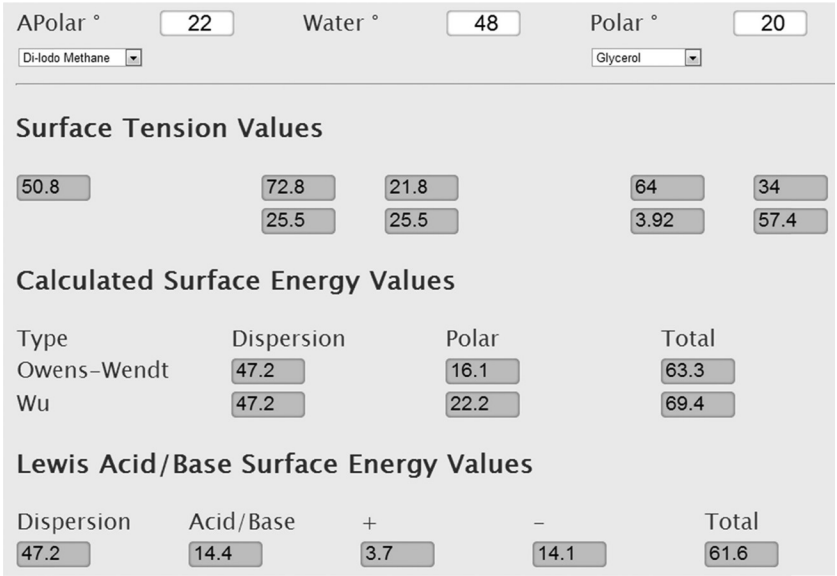


FIGURE 8.5. *Calculating surface energies from contact angle measurements with 2 or 3 solvents.*

values from different models when presenting results. The assumptions behind the calculations are remote from the messy reality of practical surfaces. The point of doing the measurements is that *changes* in these numbers away from what you know to be suitable material can reveal chemical insights into what might be causing these changes.

As discussed below, having, say, Owens-Wendt numbers for the substrate and the coating can provide insights into the problem of adhesion failure—pinpointing, in principle, whether failure is adhesive or cohesive.

A deeper insight into the surface—and more relevant measure of surface energy—comes from a more complicated piece of equipment. As the equipment is highly relevant to adhesion (though not for surface energy reasons), it is worth a little digression to explain it. Another free App (ZAAM, the Zeng-Abbott Adhesion Modeller, from the Abbott website and also the Practical Adhesion guide, which introduces the concepts in step-by-step manner) allows you to make more sense of the theory by playing with the various parameters.

Imagine a sphere of radius R and modulus E (actually an adjusted modulus E^*) being pressed against your surface. Imagine, further, that there is absolutely no surface energy interaction between the two. As the sphere is pressed harder and harder with force F, the contact width, a, increases in a manner described by Hertz in the 19th century:

$$a^3 = \frac{3FR}{4E^*} \tag{8.3}$$

With the right equipment, such as an SFA (Surface Force Apparatus), it is easy to measure F and a, and for this hypothetical Hertzian contact the results would look like the following, with perfectly superimposed loading (F increasing) and unloading (F decreasing) curves, Figure 8.6.

In reality, the sphere is attracted to the surface via surface energy forces, so the contact width, a, is now a function both of the Hertzian term and of the surface energy γ. The theory for this case was developed by Johnson, Kendall and Roberts, and so is called JKR theory. There are alternatives such as DMT, Tabor, and Maugis-Dugdale, but in general JKR is suitable for the sorts of issues in this book. The JKR equation (which reduces to Hertz when $\gamma = 0$) is:

$$a^3 = \frac{3R}{4E^*}\left[F + \sqrt{6\gamma\pi RF + (6\gamma\pi R)^2}\right] \tag{8.4}$$

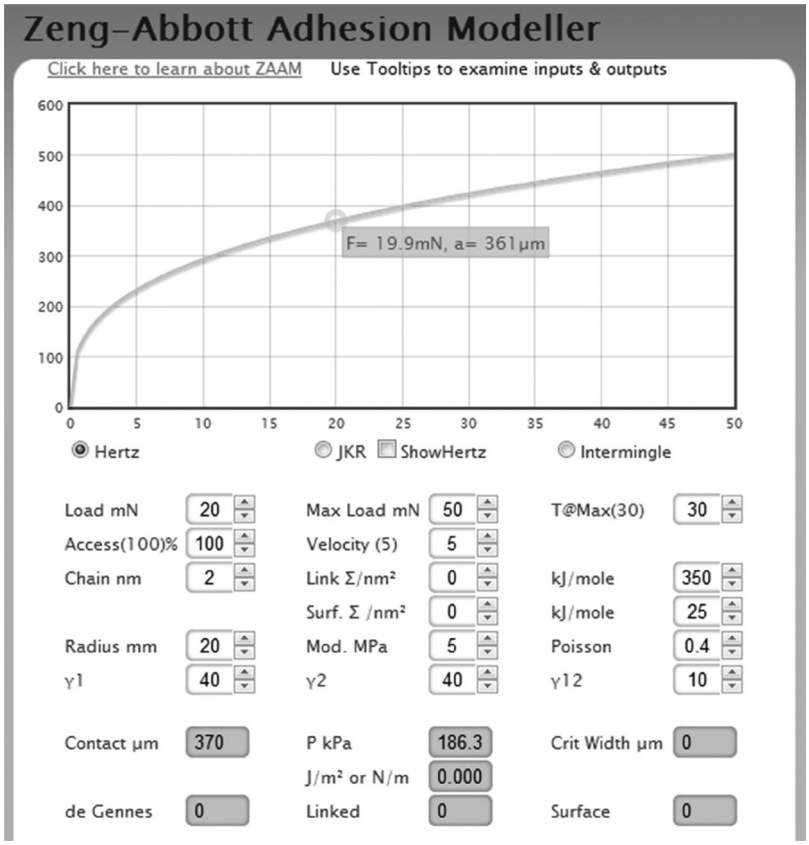


FIGURE 8.6. *A Hertzian load/unload curve. At a force of 19.9 mN the contact width is 361 µm.*

So even at zero force there is a measurable contact width, a, and it needs a negative value of F to remove the sphere from the surface. The loading and unloading curves are again superimposed, shown, in this plot along with the Hertzian plot for comparison, so the surface energy effect is clear, Figure 8.7.

So instead of a zero contact width at zero load there is a width of ~400 µm.

Again the experiment essentially provides one datapoint for two unknowns. In this example it is assumed that γ_1 (the sphere) is known to be 40 dyne/cm and the unknown γ_2 is calculated to be 40 dyne/cm, assuming (but actually a value that is impossible to know) a surface/surface energy γ_{12} of 10 dyne/cm. As this is a modeller, the supposed output (γ_2) is actually an input, but the general idea should be clear.

This seems a lot of work to get a value which is hardly more valid than one obtained by contact angle measurements. The reason for showing the technique is that often the plots show a difference between the loading and unloading curves, Figure 8.8.

The lower curve to the right of the plot is the loading curve which is identical to the JKR curve in the previous plot because none of the surface energy parameters have been changed. The fact that on the return (unloading) the curve has shifted to much larger contact widths (usually called "contact hysteresis") indicates that adhesion beyond mere surface energy effects has started to take place. In this particular example, there is some modest chain intermingling between the two surfaces, leading to ~10× increase in Work of Adhesion or Peel-strength. Note that this sort of hysteresis is not related to contact angle hysteresis in classic sessile drop measurements—where the changes are just a few mJ/m², not the hundreds of mJ/m² found in typical JKR experiments.

The authors' view is that SFA measurements will become increasingly common for those who have a need for this kind of surface data. As mentioned above, there is no necessary link between low surface energy and anti-fingerprinting or slip. The reason is that complex phenomena such as these may involve, for example, some intermingling of the molecules on the tip of the finger and the molecules on the surface of the coating. Such intermingling could very easily produce a 10-fold increase in surface force and, therefore, a big diminution in desirable properties. SFA systems often allow lateral movement to measure fric-

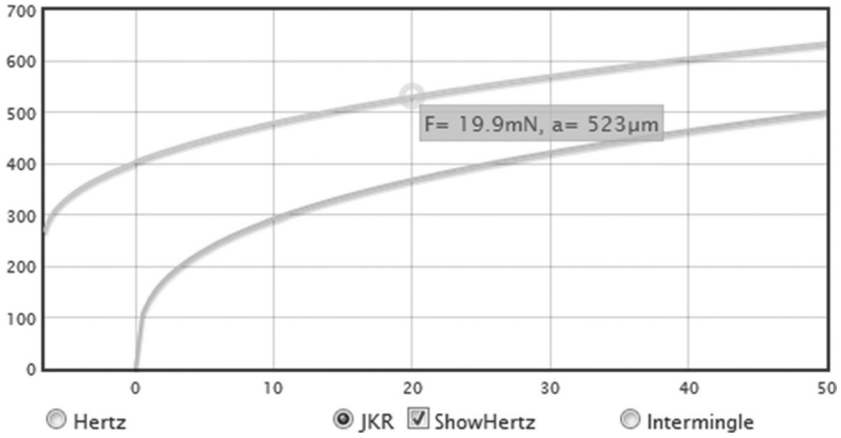


FIGURE 8.7. The Hertzian loading curve (lower) and JKR loading/unloading curve. The extra attraction means that at 19.9 mN the contact width is 523 µm.

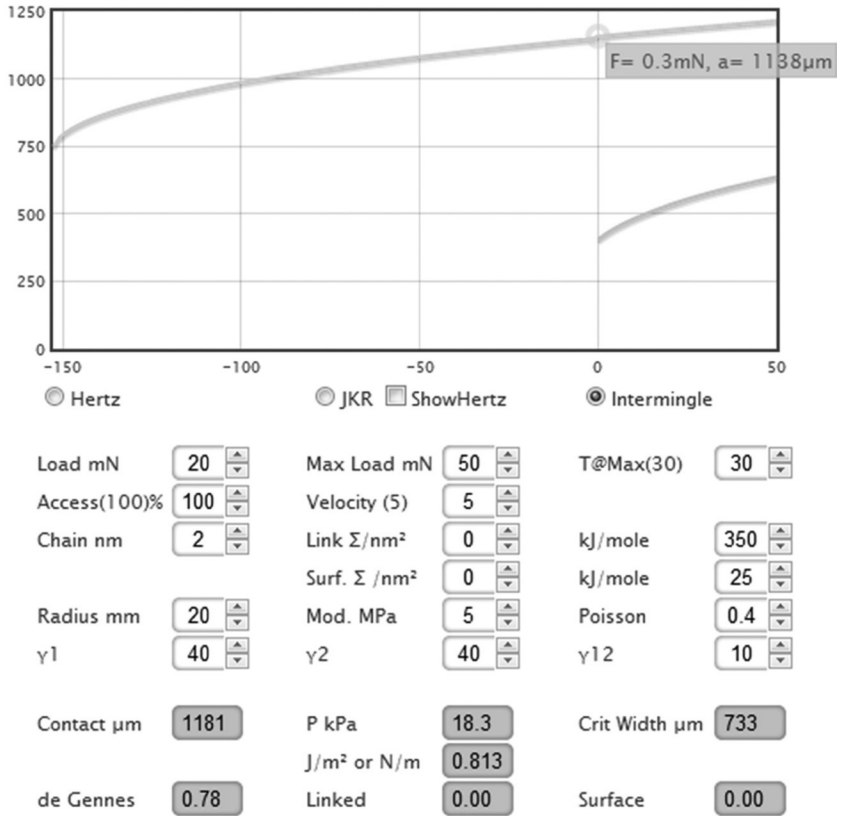


FIGURE 8.8. *A JKR curve with "hysteresis"; i.e., an unloading curve showing more adhesion than the loading curve. At 0 load instead of 400 µm, contact width is 1138 µm.*

tion in the same setup and, not surprisingly, find that friction depends on time/pressure during vertical contact.

Similarly, for printing, once there is sufficient wetting to allow the printed liquid to flow on the surface, the surface energy effect is of minor importance for overall adhesion. The intermingling and, as discussed in the adhesion section in Chapter 4, entanglement are much more important factors in determining adhesion. The point of showing the SFA here is that when checking the surface for its adhesion capabilities, any change in the degree of contact hysteresis indicates changes in the capability of the polymers and particles at the surface to interact with the printing ink (or anything else that is intended to stick to the surface). This is far more insightful than any values that might come from following changes in surface energy.

Whereas Owens-Wendt or Wu give numbers that have no direct relevance to adhesion, the JKR contact hysteresis provides lots of important information. A more sophisticated analysis might test the effect of contact time and/or maximum load on the JKR hysteresis. If, for example, it took a longer contact time or higher load to reach the same level of adhesion, that would indicate that the polymer chains at the surface were less free to intermingle. If, on the other hand, even at higher contact time or higher load the hysteresis was not increasing, that indicates that the polymer chains are accessible (they don't need extra time) and that there are less of them. This might occur if, for example, the level of adhesion promoter on the nanoparticles was less than expected or the particles could not properly reach the surface.

Going from an SFA plot to a root-cause understanding of a problem is not at all direct. The point is that it is providing important clues that other techniques simply cannot provide.

Sadly, at the time of writing, SFA systems don't seem to be readily available as standard pieces of test equipment. It is to be hoped that this situation will change when the predictive and explicative power of SFA is better appreciated.

8.7. STAIN AND SOLVENT RESISTANCE

Among the most successful nanocoatings are those used for wood floors. The nanoparticles undoubtedly contribute to the hardness of the coating. It is likely that they also help resist staining. To those who have never thought about it, creating a stain-resistant coating doesn't sound too difficult. Anyone who has carried out the infamous mustard test knows just how difficult it can be both to produce a meaningful test and, in the case of mustard, to pass it.

Unfortunately people spill things; drinks, foods, cleaners and almost anything else you might care to mention. It is easy to be dismissive of commonplace items such as tomato ketchup and mustard (although food technologists are probably spitting with rage at this point); but they are highly complicated mixtures of chemicals. For instance, most commercial mustards contain turmeric which comprises amongst its other ingredients the phenolic compound curcumin, Figure 8.9, (pictured in its enol form), essential oils and stabilizers. That combination makes the material suddenly appear a lot less innocuous.

Such mixtures can show extremely good penetrating (solubilizing) power for many coatings, and as the standard test methods involve leav-

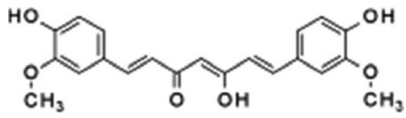


FIGURE 8.9. *Curcumin.*

ing the contaminant on the coating at an elevated temperature (e.g., 50°C) for up to 24 hours before any removal is attempted, it becomes rapidly apparent that these challenges to the coating are not to be taken lightly.

An example encountered by the authors was an extremely promising formulation containing nanoparticle zinc oxide (see the Case Study below) that showed excellent properties in all respects, except that it was critically stained by tomato ketchup. The initial cause of the problem was penetration of the coating and attack on the zinc oxide by the acetic acid present in the ketchup; this led to the subsequent staining of the coating by the red coloring matter (lycopene and other carotenoids). These examples are mentioned because it is important to distinguish between resistance to pure colorants such as curcumin and lycopene and the real-world stains that contain oils (for curcumin) or acids (for lycopene) that greatly complicate the problems of stain resistance. The HSP idea of bad solvents being able to combine to create good solvents is highly relevant in this context.

A typical set of stain resistance tests might include 20 or more commercial materials, including foodstuffs and commercial cleaning agents. The formulator should always be aware that a potential customer may well "spring a surprise" and ask for resistance testing with a material that has not been investigated previously. This can range from the mundane to the frightening. On one occasion a question was asked of Holmes as to whether a product possessed mustard gas resistance!

Solvent resistance tests are carried out on a regular basis with a wide range of solvents on a wide range of materials. Typically, a solvent-soaked wipe is held in contact with the coating under a watchglass for an extended period.

If the tester is unwary, unpleasant surprises may await. Imagine, for example, a silica nanoparticle formulation, in a radiation-curable acrylate monomer, coated onto polyester film. The nanoparticles will add a high level of solvent resistance to the coating, especially if the dispersing agents are able to co-react with the acrylates. Solvent and stain resistance tests would be passed successfully. Now coat the same formulation onto polycarbonate. The coating itself is identical so it should pass the test, but this is decidedly not the case. A solvent such

as MEK (methyl ethyl ketone) will not damage the coating, but unless the coating is totally impervious to solvent (which is unlikely) then as soon as the MEK reaches the polycarbonate interface it will attack it, leading to severe blistering. The question of what is happening at the interface is interesting. Some UV monomers are reasonable "solvents" for polycarbonate and can therefore "bite" into the surface during coating. This can be seen in cross-sections of many polycarbonate coatings. The question now is whether the monomers that have mixed with the polycarbonate become sufficiently cross-linked as to protect the polycarbonate or, perhaps, plasticize the polymer so that it is more susceptible to attack. Polyester film is essentially impervious to such solvents (though as primer coatings are generally involved it is an interesting question as to what is happening in the few nm of intermingled coating and primer), which is why the coating on polyester succeeds while that on polycarbonate fails.

The point of this story (one that is well-known to those who put hardcoats onto polycarbonate) is a reminder that when solvent resistance testing is carried out, the *whole* of the coating construction is challenged. The substrate, primer, coating and supercoat; everything that is present in the construction is a potential target for attack by the test solvent.

8.8. CONDUCTIVITY

There is very little to be said about measuring bulk conductivity through the coating or, above a certain level, the surface conductivity. Those concerned with through conductivity (solar cells, printed electronics) have the necessary equipment for measuring current/voltage curves. Those who need surface conductivity for touch panels, solar cells, etc. can also easily measure whether the surface resistivity is 10, 100 or even 1000 Ω/sq. Getting good through conductivity or high surface conductivity is difficult; these comments are simply about proving that the required values have been obtained.

The opposite is true for anti-static performance. It is not too hard to achieve some sort of anti-static performance. The difficulty is in proving it, to a level that will provide anti-static protection under the desired down-stream conditions.

The smallest of the problems is understanding what the measurement provides. For resistance, R, in Ω, all that is required is a known voltage V and a measured current I so that $R = V/I$. We know, however, that

the current depends on the size of the conductor; the current that flows increases linearly with the width, W, of the current path and decreases linearly with the distance, D, between the test electrodes. To define a *resistivity*, ρ, requires a nominal unit width and unit distance. So $\rho = (V/D)/(I/W)$. This still has units of Ω, because the length units cancel out, but is labelled (erroneously, according to purists) Ω/sq, where the "sq" implies a square cm or a square m or whatever unit is of interest. The unit is not important because if D goes from 1 cm to 1 m (reducing the current because of a longer path), W also goes from 1 cm to 1 m (increasing the current by the same amount because of the wider path).

The next problem is the impossibility of measuring pure surface conductivity. Electrons do not flow along an ideal, infinitesimally thin layer of coating (except, perhaps, in graphene); there are always complications through surface effects. The way out of this dilemma is to work with electric fields which are strongly geometry dependent. The geometry of concentric rings of radius r_i (inner) and r_o (outer) happens to be particularly convenient for analysis and if the resistance R (i.e., a delicate measure of Voltage/Current) is obtained for a concentric electrode then the surface resistivity is given by $\rho = R.2\pi/\ln(r_o/r_i)$. In other words, you buy a standard tester with a known concentric ring geometry and the delicate electronics necessary to measure the tiny currents that flow when ρ is in the range of 10^{10} to 10^{14} Ω/sq.

The real problem is relating surface resistivity to anti-static performance. When the resistivities are below, say, 10^9 the problem goes away—under most circumstances static will not be an issue. Frustratingly, it is possible to find two coatings each with the same measured resistivity of, say, 10^{10}, measured at the same temperature and humidity, with very different tendencies to shrug off static. The phrase "shrug off" is deliberately vague. When handling polyester films, for example, some anti-static treatments seem to allow the sheets to be handled (e.g., during conversion) with no problem and others end up with a stack of sheets locked together by static. The fact that there is often no relationship between surface resistivity and anti-static performance is well-established, though too little known [2].

The missing link is the way the anti-static coating encourages or discourages triboelectric charging. This suggests that a better test of anti-static behavior is a device that charges the surface by rubbing and measures the speed of discharge. Such a device would be rather tricky to devise, so in reality a charge decay anti-static meter induces the sorts of charges (kV and pA) that are typical of static and monitors the decay.

Note that this requires corona charging rather than the polite 25–1000 V offered in many charge decay machines. The authors' experience is that for marginal surface resistivities, such a test is more informative than surface resistivity, with the obvious caveat that the test itself is subject to the vagaries of handling high voltages and super-low currents. Continuing the frustration, sometimes a good charge dissipation performance with a negative charge is not matched by the performance with a positive charge.

The hope is that nanoparticle anti-statics will provide reliable low surface resistivities, largely independent of humidity. It is also hoped that they will be less susceptible to triboelectric charging, with better overall charge dissipation characteristics. If the particles are present at the surface they might reduce triboelectric charging by offering a nano-rough surface with less contact.

The point of this section is to emphasize the need for tests beyond surface resistivity, such as charge decay devices and practical tests (such as going through a conversion machine) that mimic what the customer will do with the product. Even more importantly, because surface effects can be so subtle, a robust, agreed QC test protocol that relates to end-use performance has to be established. If the lab have done a good job in obtaining a reliable distribution of the right particles at the surface—with low(ish) surface resistivity, with low triboelectric charging and low susceptibility to effects such as Coulomb trapping—then maybe a standard ring electrode surface resistivity test will be enough to ensure that no major production problem exists. If, however, the lab has produced a coating that is more typical of anti-static coatings in general, then the QC lab will need some tests beyond surface resistivity such as a charge decay test run with both positive and negative charges.

Finally, there needs to be agreement as to what to do if the QC test fails. The lab team needs to have some heuristics (lower line speed, hotter drying, etc.) that offer some hope of bringing the coating quickly back into specification. The production team needs to understand that with anti-static performance it is a universal phenomenon that finding such heuristics is difficult.

8.9. SCRATCH RESISTANCE

Like many in the (nano) coatings industry, the authors have scars from the battle to produce scratch-resistant coatings. The scars are partly to do with the challenge of producing scratch resistance. Most of the

scars, however, are a result of the need to prove internally and externally that sufficient scratch resistance has been obtained.

It would be possible to fill this section with discussions about various scratch test procedures. This, arguably, is of little real-world value. Instead, the focus is on why one common method is worthless and why, in the view of the authors, only one method makes scientific and practical sense, especially for those who are relying on nanoparticles to produce the required scratch resistance.

Before discussing the bad and the good, it is worth touching on the splendid technique of ensuring that the surface has a nice low coefficient of friction. In many practical situations, the fact that objects will slip rather than gouge into the surface reduces scratching significantly. There are a number of well-known additives that will migrate to the surface during production of the coating and give a low surface energy, typically long-chain fluoro or siloxane groups. The trick is to ensure that they come to the surface (within the matrix they add no value) and are locked to the surface via suitable chemical bonds. As these functionalities are rather soft, as a thick layer they would be easily damaged. So the thinnest possible coherent surface layer is required—provided it can stay in place for the lifetime of the product. The presence (or absence) of such materials on the film surface can be readily determined by contact angle measurements. Whether they provide "real" scratch resistance is something that can only be decided by your customer. Whether they also provide anti-fingerprinting is a question addressed above.

8.9.1. Pencil Hardness

The pencil hardness test is beguiling in its simplicity and is widely used across the surface coatings industry. For the unscrupulous, it can be readily manipulated to give almost any number desired. Pencils of increasing hardness (2B, B, H, 2H, etc.) are drawn across the surface with a controlled pressure and the surface inspected for a visible scratch. If there is no scratch with a 2B, the B pencil is tried. If that is OK the H pencil is tried. Eventually, say at 3H, scratches are found, so the coating is labelled as 2H.

There are a number of ways to massage the technique to give the desired answer:

- Forget to sharpen the pencil for each stroke—a blunt pencil doesn't scratch so much.

- Go easy on the pressure.
- Carefully choose the vendor of pencils; a 3H pencil from one supplier might be much softer than a 3H from another.
- Take off your glasses, turn down the lights, or make sure there is a bright light in a strategic position—this makes it much harder to see the scratches and can easily gain you an extra pencil or two.
- For coatings designed for plastics, do the tests with the coating itself on glass. This can be worth 3 or more pencils. When this is done, be careful to word the claim correctly. To say "This hardcoated film is 6H" is lying. To say "The coating passes a 6H test" is perfectly true, even if it happens to be irrelevant to the hardcoated plastic film. The reason for this trick is that the softness of the underlying polymer makes pencil breakthrough inevitable. A 6H for a coating on glass might be a 3H on PET and B on PC.
- Of course, standards exist that describe a method for carrying out the test. These include ASTM D3363, ISO 15184, etc. This should prevent outright "cheating", but does not address the fundamental problems described above and below. It is also worth noting that it is still quite unusual for the test method to be quoted alongside the pencil hardness values in technical and marketing presentations. It is the authors' opinion that such presentations of data are "not even wrong".
- Finally it is worth pointing out that the pencils used for these tests were never designed for scientific testing; the pencil manufacturers are quite clear on this point. The whole range of pencils was designed for use by artists and draftsmen, not as a means of scratch indentation.

If the above list reflects a certain world-weary cynicism, scientific objectivity reaches the same conclusion—that the test is effectively worthless. Holmes and colleagues sacrificed some weeks of tedious lab time to carry out a "process capability" study on the pencil hardness test using more than 1000 individual pencil hardness measurements. This involved a set of blind tests on a set of films by a set of trained, skilled testers genuinely trying to see whether it was possible to distinguish reliably between, say, a 2H and 3H coated film. The procedure made it possible to disentangle the elements of the test and the tester.

To give some idea of the effort involved, here is just one of the datasets gathered by the team, Figure 8.10.

The basics are all there. Three operators (1–3) tested ten different

GAUGE R & R DATA SHEET
LONG METHOD - 3 x 3 x 10

Operator		Sample											Average
		A	B	C	D	E	F	G	H	I	J		
1	Trial 1	3.000	4.000	3.000	5.000	5.000	4.000	4.000	1.000	5.000	2.000		3.600
1	Trial 2	3.000	3.000	3.000	5.000	5.000	4.000	5.000	2.000	3.000	4.000		3.700
1	Trial 3	5.000	4.000	4.000	5.000	5.000	4.000	4.000	4.000	4.000	5.000		4.400
1	Average	3.667	3.667	3.333	5.000	5.000	4.000	4.333	2.333	4.000	3.667	Xa	3.900
1	Range	2.000	1.000	1.000	0.000	0.000	0.000	1.000	3.000	2.000	3.000	Ra	1.300
2	Trial 1	3.000	2.000	4.000	3.000	1.000	2.000	2.000	3.000	1.000	4.000		2.500
2	Trial 2	3.000	4.000	4.000	2.000	3.000	4.000	3.000	2.000	4.000	5.000		3.400
2	Trial 3	3.000	3.000	2.000	4.000	1.000	4.000	5.000	2.000	3.000	5.000		3.200
2	Average	3.000	3.000	3.333	3.000	1.667	3.333	3.333	2.333	2.667	4.667	Xb	3.033
2	Range	0.000	2.000	2.000	2.000	2.000	2.000	3.000	1.000	3.000	1.000	Rb	1.800
3	Trial 1	3.000	3.000	2.000	4.000	1.000	4.000	5.000	2.000	3.000	5.000		3.200
3	Trial 2	3.000	2.000	3.000	2.000	2.000	3.000	5.000	1.000	1.000	3.000		2.500
3	Trial 3	1.000	1.000	4.000	1.000	1.000	2.000	2.000	3.000	2.000	3.000		2.000
3	Average	2.333	2.000	3.000	2.333	1.333	3.000	4.000	2.000	2.000	3.667	Xc	2.567
3	Range	2.000	2.000	2.000	3.000	1.000	2.000	3.000	2.000	2.000	2.000	Rc	2.100
	Part Ave	3.000	2.889	3.222	3.444	2.667	3.444	3.889	2.222	2.889	4.000		3.167
	Xa	3.900										Rp	1.778
	Xb	3.033											
	Xc	2.567											
Rbar		= [Ra+Rb+Rc] / #Ops										Rbar	1.733
Xdiff		= MaxXbar - MinXbar										Xdiff	1.333
UCLr		= Rbar * D4										UCLr	4.472
LCLr		= Rbar * D3										LCLr	0.000

Measurement Unit analysis		
Repeatab	EV =	5.29
Reproduc	AV =	3.47
R & R	R & R =	6.32
Part Varia	PV =	2.88
Total Vari	TV =	6.95

% Total Variation	
	0
% EV =	76.09
% AV =	49.92
% R&R =	91.00
% PV =	41.45

FIGURE 8.10. *A small fraction of the pencil hardness data used to calculate process (in) capability.*

films (A–J), each three times (Trial 1–3) with 5 attempted scratches, scoring the number of scratches from 0 (totally unscratched) to 5 (a scratch every time). To the untrained eye the variability looks significant. Applying the technique of Gauge R&R (Repeatability & Reproducibility), it is possible to provide a score for the % variation for Repeatability (76%), Reproducibility (50%), and Part Variability (41%). The total R&R variability is 91%. In industry, an "in control" process/test has an R&R of <10, a process that's OK is in the range 10–30%. So a score of 91% shows how unfit for purpose this test is.

The same data can be plotted as a graph. It looks as though Operator 3 is more forgiving than Operator 1. So Marketing could request Operator 3 for their own product and Operator 1 for the competitor's, Figure 8.11.

And, to repeat, this is a test done by three highly trained people using the same samples, the same standard procedure, the same viewing conditions and same batch of pencils from a single supplier.

More damning evidence is provided by a test carried out by a single operator using a series of pencil leads of nominally the same hardness from a single manufacturer. All of the leads were taken from a new unused box of pencils.

It is readily apparent that there is no reproducibility in scratch performance of the 5 individual 2H leads that were tested. For example, lead 1 will show 0 scratches out of a possible 5 tests about 25% of the time and 3 scratches out of a possible 5 < 10% of the time; lead 5 will show 0 scratches < 20% of the time and 3 scratches 50% of the time!

This experimental result is obtained from 300 individual pencil hardness tests carried out under identical conditions, Figure 8.12.

In other words, a test used by a large industry (hard coated films for switches and touch screens) was proven to be unfit for the purpose.

One further outcome from this study was the results of the same hard coat, produced under identical conditions on three different substrates, plus the value for the same hard coat coated (under lab conditions) onto glass. The results emphasize why the test says rather little about the hard coat, Table 8.1.

The question as to why PET and PMMA behave so differently to glass and polycarbonate has no simple answer. At a purely empirical and (it has to be said) practical level, it is probably sufficient to say that the softer the substrate, the less physical support it provides for the hardcoat when this is indented by the pencil. One factor that can be

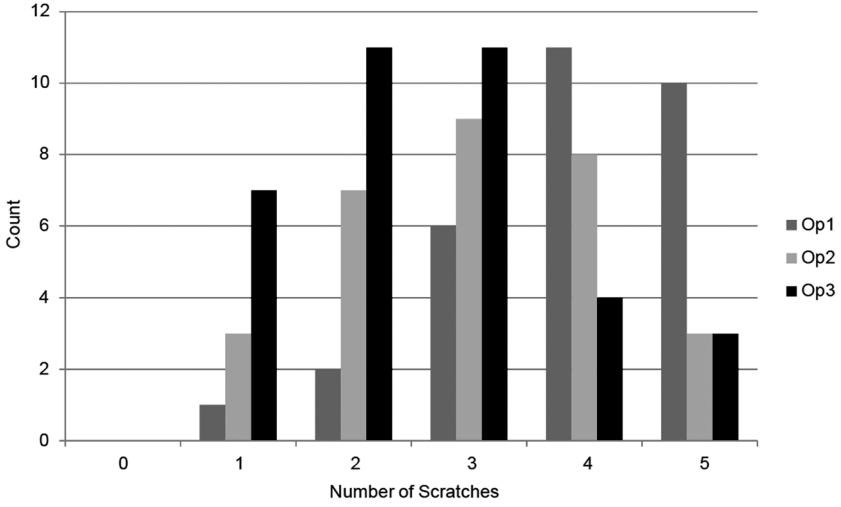


FIGURE 8.11. *Person-to-person variability of pencil hardness tests.*

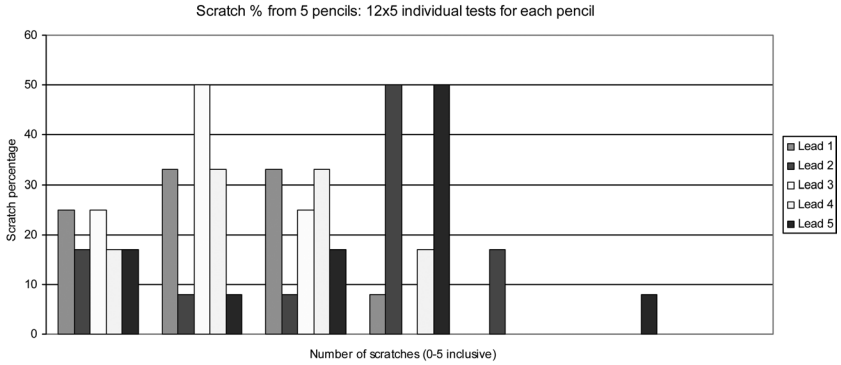


FIGURE 8.12. *One person—different leads.*

readily used by the formulator is the dependence of scratch resistance on the coating thickness; quite simply, for hard films on compliant substrates thicker coatings are generally better.

Unfortunately, a full treatment of the effects of substrates on scratch resistance is beyond the scope of this book. The fracture mechanics at the interface of the coating and substrate cannot be described in simple mathematical terms. Factors that affect the overall performance include the coating thickness, elasticity mismatch between the coating and substrate, the corresponding Poisson ratio mismatch, and the interfacial friction [3].

Although many people agree in private that the test is worthless, its simplicity, the absence of a viable simple alternative and the fact that marketing departments can ensure that the softest possible set of pencils is used by the technical staff ensures that the test is still used worldwide.

Useful data on abrasion resistance can be obtained using Taber abrasion testing and similar techniques; but these do not fully replicate the interaction of stylus driven instruments on a coated surface. In the authors' opinion there is only one scientific means of obtaining scratch

TABLE 8.1. Pencil Hardness of the Same Coating on Different Substrates.

Substrate	Pencil Hardness
Polycarbonate	2B
PET	2H
PMMA	3H
Glass	6H

resistance data that helps to optimize the formulation and that is by nanoindentation.

8.9.2. Scientific Insight Into Scratch Resistance

The terms "pencil hardness" and "hard coat" help to prejudge the issue of what makes a good product: hardness. Given that silica or alumina are both very hard, it seems obvious that adding plenty of silica or alumina nanoparticles will increase hardness and a much-improved coating will result. The experience of many disillusioned researchers is that the nanoparticles can often make astonishingly little difference.

Those who abandon the pencil hardness test naturally seek a more scientific test. Because hard materials have a high modulus, it seems logical that measuring the modulus and formulating for as high a value as possible is the way forward.

There is a wonderfully simple way to measure the modulus of a hard coat—provided it can be produced as a stand-alone film by, for example, coating it onto a substrate with a release coating.

Cut out a strip of length L and thickness H. The width is not important—it should be wide enough that edge defects from cutting out the strip are not significant. Now clamp the strip horizontally and measure the deflection D of the strip under its own weight; the larger the deflection, the lower its modulus.

For a cantilever (supported at one end) the modulus is calculated from the classic beam bending formula (in the Modulus Spreadsheet) where ρ is the density and g is gravity:

$$\text{Modulus} = \frac{1.3L^4\rho g}{H^2 D} \tag{8.5}$$

If the experiment is done with the sample supported at both ends, the same absolute deflection in the middle requires a smaller modulus, by a factor of 8, because to get the same sag with something supported at both ends the material must be weaker.

The formula shows that the deflection goes as L^4 so it is relatively easy to "tune" the measurement by increasing or decreasing L so that the deflection is readily measurable.

The main practical difficulty comes from the amount of curl in the sample. Sometimes a modest annealing process can remove the curl.

Otherwise the test can be done curl up and curl down and the deflection can be assumed to be the average of the two values.

For scratch resistant coatings intended for rigid substrates, such as metals or glass, the modulus can be made arbitrarily high through high loadings of (sol gel) nanoparticles, large cross-link densities and high-temperature annealing. Such brute-force tactics generally fail for hardcoats on flexible substrates and there seems to be no correlation between scratch resistance and modulus once a certain minimum modulus (~1 GPa) has been reached.

A simple shift of nomenclature away from hardness directs our minds to the real source of scratch resistance. What is a scratch? It is a permanent deformation in the coating. What is modulus? The result of a stress-strain experiment that assumes Hookian elastic behavior. In other words, a scratch is a plastic deformation and a modulus (in the regions where they are conveniently measured) is an elastic deformation. Put like this, it is clear why there is no necessary relationship between modulus and scratch resistance.

Of course, high modulus materials will tend to require a large strain before they move into the plastic domain, so modulus is not irrelevant to scratch resistance—low modulus materials tend to be useless for hardcoats. The point of the shift of nomenclature is to shift the experimental emphasis away from the elastic domain into the plastic domain. To understand scratch resistance we need a quick way to measure resistance to plastic deformation.

Purists will note that at high enough frequencies (or low enough temperatures), everything is elastic and at low enough frequencies (or high enough temperatures), everything is plastic. This is captured in the Deborah number ("The mountains flowed before the Lord", Judges 5:5), which is the ratio of the response time of the system to the timescale of the measurement. In the discussions that follow, the time-scales are seconds and for typical hardcoats the response time (above a yield threshold) is also in seconds; in other words the system is borderline in terms of Deborah number so some systems will yield and others won't. A microsecond nanoindenter test would allow most hardcoats to escape scratch-free, while an hour-long test would be a severe challenge to most hardcoats.

Modern nanoindenters make it (relatively) easy to obtain the key data. A stylus of a known shape is pressed with known force and known rate of increase of force into the coating. For a purely elastic material, the stylus will form a modest indentation, and the reverse curve

(withdrawing the stylus) will be an exact mirror. For a plastic material, the stylus will give a deeper indentation than allowed by the modulus, and on withdrawal of the probe the forces will return to zero before the probe has returned to its starting position.

Although the theory isn't particularly difficult, instead of the usual Oliver and Parr approach [4], for clarity the approach of Hainsworth *et al.* is used in the Nanoindentation spreadsheet to describe what happens [5]. Given a reduced modulus E^* and a hardness H (to be discussed below) then the depth of impression, δ, for a pressure P during loading (the loading curve) is given by:

$$\delta^2 = \frac{P}{E^*}\left(\varphi\sqrt{\frac{E^*}{H}} + \psi\sqrt{\frac{H}{E^*}}\right) \tag{8.6}$$

The two factors φ and ψ in the equation are fitted parameters which seem to work well in a standard nanoindenter fitted with a Berkovich indenter (a sharp three-edged corner with a specific angle). The φ term is associated with the plastic effects and the ψ term with the elastic effects. The unloading curve in the spreadsheet is a simplification—it sets the φ term to zero and shifts the curve to match the end of the loading curve. On a real piece of nanoindenter equipment the factors would be deduced from the fitting software. E^* depends on the true modulus, E; the Poisson ratio ν (which expresses how much the material expands in another direction due to pressure in the orthogonal direction, and assumed to be 0.3 unless data are available) and the modulus and Poisson ratio E_i and ν_i for the indenter, Figure 8.13.

$$\frac{1}{E^*} = \frac{1-\nu^2}{E} + \frac{1-\nu_i^2}{E_i} \tag{8.7}$$

The spreadsheet uses the input values of E and H to show what the experimental results would look like. It also calculates two additional values discussed below. In the figure, E^* is 4.4 GPa (E = 4 GPa) and H is 0.5 GPa, typical of a good hardcoat. In a real experiment, the equipment's software would calculate four pieces of information:

- Modulus
- Plastic depth
- Elastic recovery
- Hardness

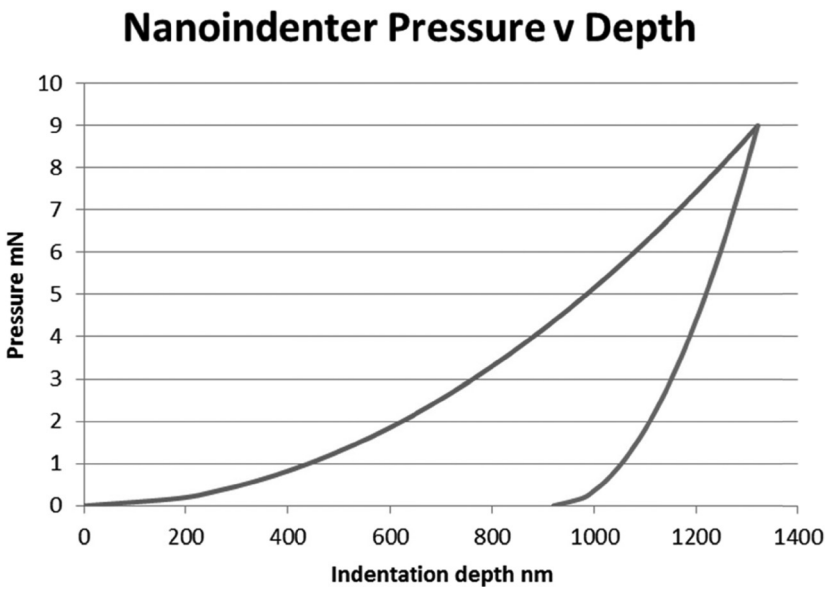


FIGURE 8.13. *A typical nanoindentation plot using simulated data.*

As discussed above, the modulus is the least informative in terms of scratch resistance. The three other parameters are inter-related. If the indenter has moved in by 2000 nm and according to the modulus it should have entered only by 1000 nm, then the plastic depth is 1000 nm and the elastic recovery is 50%. The plastic depth is an absolute number depending on the test, and so is not usable in comparisons. The elastic recovery is useful—the higher the value, the more likely it is that the hardcoat is resistant to scratches.

If discussions on scratch resistance should focus on a single number, then that should be hardness as defined by the nanoindenter community rather than the common understanding that "hardness = modulus". The definition in terms of the nanoindenter test is simple:

$$\text{Hardness} = \frac{\text{Maximum Load}}{\text{Contact Area}} \tag{8.8}$$

This is often called the Meyer Hardness and is closely related to the Vickers Hardness (used for expressing the performance of metals and ceramics), which is based on a single measurement of depth of penetration for a given load and a given indenter.

At first, it is not obvious what this Hardness means, but the idea turns

out to be quite simple. The further the nanoindenter probe enters into the sample, the wider the area in contact with the sample. As shown in the top of the figure, for two samples with zero plastic deformation (i.e., full elastic recovery), the one with the smaller modulus will have a bigger indentation and a larger contact area. In this extreme case, Hardness is essentially the same as Modulus. Now (lower part of the figure) take two samples with the same modulus, with one having a small plastic deformation and the other having a large plastic deformation. For the latter, the indenter will go deeper and the contact area will be larger, so the Hardness is smaller, Figure 8.14.

The Nanoindentation spreadsheet calculates the contact area for each load using a typical Berkovich formula:

$$\text{Contact Area} = 24.5\delta^2 + 1500\delta \tag{8.9}$$

then back-calculates the Hardness. Because of the simplifications in the calculations, the input and output Hardness don't match, especially when Hardness exceeds 10% of Modulus.

The ideal situation for the formulator (and you get both numbers au-

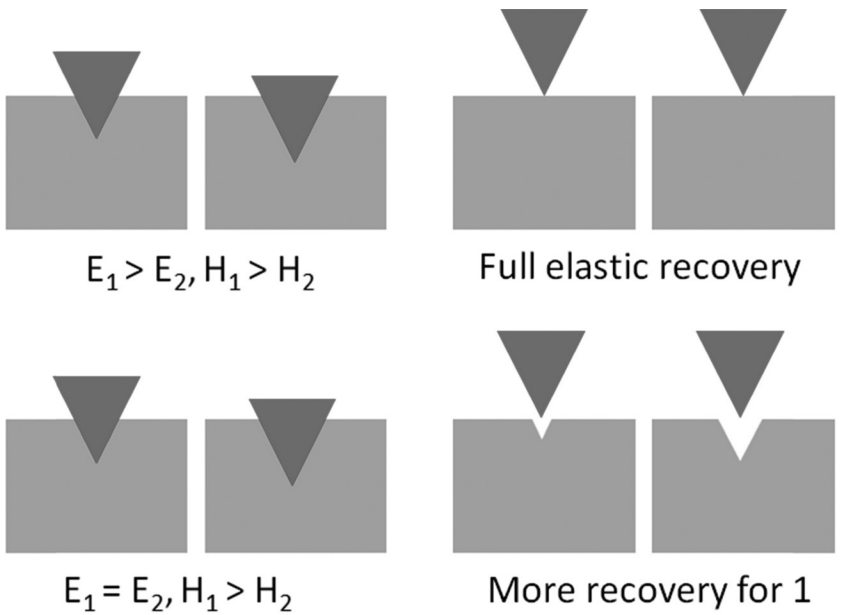


FIGURE 8.14. *For purely elastic materials, hardness (H) follows modulus (E). For plastic materials, for the same modulus, the harder material shows less indentation depth and a fuller recovery.*

tomatically) is to know both modulus and hardness. In order to optimize a formulation for scratch resistance, look out for how the two numbers are playing relative to each other. Brute force increase of modulus is a thankless task and generates the risk of loss of flexibility. The significant breakthroughs come from those variants that increase the hardness beyond the modulus.

The section in Chapter 4 on balancing chemical bonding, entanglement and cross-linking points towards formulation changes that are likely to increase hardness. Any micro-crack at a nanoparticle/matrix interface is going to allow plastic deformation, so make sure that the fracture energy at these interfaces is maximized. The ability of the coating to deform plastically is then drastically reduced.

In summary, if the nanoparticles aren't strongly locked into the hardcoat then the improvements to scratch resistance will be disappointing. By understanding the physics of locking them in (entanglement and the right balance of chemical bonds) and by using nanoindenter tests to measure modulus and hardness, the probability of producing a scratch-resistant coating increases strongly.

Sadly, once the good science has been done, customers (and Marketing) will still ask for pencil hardness data (or other equally unscientific tests). Happily, the chances of obtaining good values are significantly increased by getting the science right in the first place.

It might seem odd not to be discussing the fact that nanoindenters can be made to produce controlled scratches. An x-y stage on the machine can move the sample while the pressure on the stylus is steadily increased. Various analyses can then be performed on "scratch resistance". As far as the authors are aware this technique isn't of much direct help in coming up with improved formulations. The "scratch test" becomes an exercise in defining what the scratch test should be. Changing speeds, pressures, and indenter tip can all give different scratch characteristics and it is also not obvious which characteristic is important.

The "simple" load/unload technique provides objective measures of modulus, elastic recovery and hardness, and from these the formulator can send the formulation into the regime with the right balance (the highest hardness while preserving a low-enough modulus to avoid cracking) for the best practical resistance to scratching. Arguably, this is achieved by optimizing the level of nanoparticles and the robustness of their adhesion to the matrix, using the ideas on adhesion from Chapter 4.

The greatest current barrier to the adoption of nanoindentation testing as a norm within the surface coatings industry (along with lack of understanding) is probably cost. This should not discourage the formulator from investigating the technique; the tests can be outsourced and, as the demand for such tests becomes universal, manufacturers will be able to deliver routine test equipment at affordable prices.

8.10. ADHESION

The most important thing needed to make progress on a key product parameter is a meaningful test which can be performed rapidly and easily. With adhesion it is possible to have rapid tests and meaningful tests, though it is difficult to obtain both.

The first complication is that in order to rip things apart it is necessary to gain purchase on the coating and the substrate. This means that adhesives have to be found for both surfaces that are just slightly weaker than the adhesion required for the test sample. If they are stronger, then the test will always fail. If they are substantially weaker then they will always fail first, giving very little information about the true adhesion other than in cases where the adhesion is very much smaller than required.

Even given that this problem has been solved, another key problem has already been alluded to in the section on adhesion in Chapter 4. If the samples are pulled apart vertically, the chances are that even surface energy adhesion will be strong enough to resist the tug (see the picture of Abbott and the rubber plates). Therefore, a test needs to be found that seriously challenges the interface by forming a potential crack at the interface that will propagate if the work of adhesion is too low.

In rare cases (or sometimes by deliberate introduction of an adhesion inhibitor), an edge of the coating can be isolated, some adhesive or clamp can be attached to it and a tensile tester can tug at 90° and measure the work of adhesion from the stress/time curve or—it comes to the same thing—the average peel force.

More usually, tape adhesion is quick, easy and probably the only practical choice. It is also frequently meaningless. For example, those who create hardcoats that feature a slip agent on the surface in order to reduce scratching find that their coatings seldom fail a tape adhesion test because the tape fails to stick firmly to the surface and thus is easily removed without damaging the coating. Similarly, a tape that is applied to the middle of a sample has little chance of starting a crack, so there is no test of the resistance to crack propagation.

The "least worst" tape test can be performed at reasonable speed while providing reasonably consistent and relevant data. The recipe is:

- Apply super-strong double-sided adhesive to a strong surface
- Stick the substrate onto the adhesive using a repeatable, strong force
- Cross-hatch the coating in order to create a known and repeatable number of sites where cracks can start to propagate
- Wipe away any residue from creating the cross-hatch, so it doesn't interfere with adhesion
- Apply a known, strong tape with a known, vigorous force (weight, roller, rubbing tool)
- Gently pull one end of the tape until it is vertical and provides a good grip
- Pull sharply while moving the hand forward to preserve a near-vertical pull—this is because the peel force goes as $\cos(\theta)$ so when θ is near 180° (which is the more natural way to pull), the peel force is halved.
- Count the number of failed squares

Whatever the test for adhesion, it is important not to jump to conclusions if adhesion fails. In a multilayer system, *adhesive* failure can be at any one of the interfaces and *cohesive* failure can be in any one of the layers. Cohesive failure can also occur randomly within the layer or at a specific point. A typical example of the latter comes from excessive corona treatment to a surface. Such treatment can provide excellent adhesion to the top layer, while reducing that top layer to rubble, which has little adhesion to the rest of the substrate.

Given that the coatings are often thin, finding unambiguously whether failure is adhesive or cohesive is a real problem. In an ideal world, a simple measurement of water contact angle should be enough to distinguish between most substrates and most coatings. If you are unlucky, then a more rigorous test via Owens-Wendt or Wu analysis of the surface energy components (Dipolar and Polar) should reveal differences as discussed above. There are multiple problems with this approach:

- The definition of adhesive and cohesive failure can depend on a monolayer in either direction, because surface measurements reveal information only about the first few Ångstroms. It is entirely possible that the measured surface energy of the failed surface is equivalent to neither the substrate nor the coating.
- What is the surface energy of the coating? This may seem an odd

question. At an extreme, a coating that contains a surface-active component (e.g., to induce slip) will have a measured surface energy (at the top surface) that bears no relationship to the "true" surface energy at the interface with the substrate. Even without such an extreme, there are many ways in which a coated surface may not be representative of the "true" value at the interface.

Finally, it is worth noting that the strength of PET film lies mostly in the x-y plane (via orientation during production), so strength in the z direction is not so large. This means that failure can occur within the PET film, especially if scratches from the cross-hatch go deep into the PET, allowing a crack to form.

8.10.1. Between Adhesion and Brittleness

A different type of adhesion test looks for failure at an interface after prolonged flexing and/or flexing around a tight radius. This is the mandrel test and it has been used for many years within the surface coatings industry.

Failure can occur, at the extreme, either through adhesive failure due to the strong local tensile stresses or through fracture of a coating that is too brittle. The boundary between these two effects is more philosophical than scientific. Importantly, the thickness of the coating is a key element in this test, even though it has (or should have) no direct impact on the adhesion itself.

A coating of thickness d bent into a circle of radius R has a difference in circumference of $2\pi(R + d) - 2\pi R$, and therefore strain of $2\pi d/2\pi R = d/R$, Figure 8.15.

The stress along the interface is therefore Modulus*d/R. Whether the

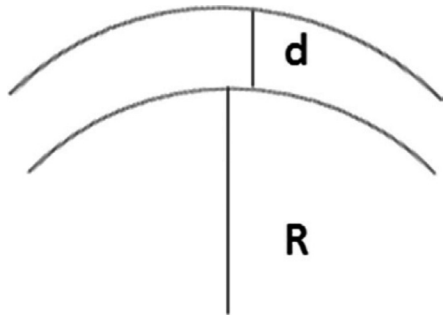


FIGURE 8.15. *The d/R strain relationship for a curved film.*

failure is of adhesion or from cracking of the film depends on whether the yield strength of the film is greater or smaller than the adhesion.

Note: Unlike many of the tests of this chapter, the recipe for success is clear. For a given desired thickness d, find an R_{min}, however small, that causes failure in a simple automated test. If this R_{min} is seriously smaller than any likely real-world scenario and if likely changes in d will make little difference to the result then everything is OK. If things are near a borderline then it is relatively easy to make controlled formulation changes to influence adhesion, brittleness or both and obtain an objective measure (R_{min}) of changes. If there are grounds to believe that one sort of change affects adhesion (or brittleness) only, then extra insight is gained into the failure mode depending on whether R_{min} is altered by those changes.

Freedom to modify d is important for other aspects of the product; thicker coatings are, for example, generally easier to make than thinner ones. If the fracture tests show that d must be smaller than the coating team might like, then at least the whole team can reach a compromise decision and work with it rather than waste time blaming the other side that (a) the coaters cannot coat properly and (b) the formulators cannot formulate properly.

Purists will note that the flex failure formula is somewhat over-simplified. As mentioned in the context of the pencil hardness test, it is easy to find more complicated formulae involving elasticity mismatches and so forth. Given that most of us have little idea about elasticity mismatches, the extra complication is not matched by extra practical insight, so the team can focus on the profound insights of the simple formula to reach an agreement on the combined optimum d and formulation.

Another way in principle to induce known stresses at the interface is via a nanoindenter. The plastic deformation caused by the stylus meets a blockage from untouched material around it, so there is a net upward "popping" force which can cause rupture at the interface. Although there are methods claiming to be able to calculate interfacial adhesion energies from the nanoindenter tests, they do not seem to be readily applicable to the sorts of coatings relevant to this book—or at least the authors cannot find a convincing way to transcribe the results from academic papers into recipes that they can put into practice.

8.11. UV RESISTANCE

Major efforts have been expended in the surface coatings industry

to produce products that do not degrade in an outdoor environment. A wide range of conventional UV stabilizers and anti-oxidants are available, in commercial quantities, to meet this challenge, often with demonstrably good results; so why consider a nanoparticle option? There are several possible answers to this question; these include:

- Conventional materials can leach from the coating, producing a cosmetically unacceptable appearance and failure of the material. This is a phenomenon that is familiar from the failure of polyethylene telephone cables due to the loss of the added anti-oxidant.
- Many of these conventional materials will themselves degrade over a period of time, thereby reducing the resistance of the coating.
- Often, combinations of materials are added to provide the maximum possible protection. It is not unknown for such combinations to be themselves unstable, with consequent breakdown of a supposedly protected article. An example of this is the reaction that can occur when hindered amine light stabilizers (HALS) are used in combination with thioester stabilizers. Acidic species produced from the latter can protonate the HALS, thus reducing their efficacy.

Such problems have been reviewed in some detail by J. Scheirs [6].

These are sound reasons for considering an alternative approach to improving outdoor stability of coated films. Addition of UV-absorbing particles overcomes many of these objections—provided, of course, that the particles meet the "invisibility" requirements (combination of small size and RI matching) discussed previously.

What nanoparticle materials are capable of providing outdoor stability? First, they are likely to be oxides so that they themselves are stable. Second, they need to have a UV-absorbing band close enough to the visible cut-off to provide protection and far enough away to avoid yellowness in the coating. Thirdly, the particles must not themselves generate destructive species when they absorb UV. In practice this comes down to titanium dioxide (though some forms generate potent radicals), zinc oxide and cerium oxide.

Each of these materials behaves as a semi-conductor; this means that there is a well-defined band gap between the valence band which is occupied by electrons and the conduction band which is unoccupied. Band gaps are generally expressed in terms of electron volts, eV. The table shows the values for the three oxides along with a translation into wavelength λ via:

$$\lambda = \frac{1204}{\text{eV}} \tag{8.10}$$

The refractive index is included to make the point that other things being equal, the ease of hiding the presence of the oxide goes as ZnO > CeO_2 > TiO_2 because light scattering increases with the RI mismatch with the matrix, which will typically be ~1.5, Table 8.2.

8.11.1. Nanoparticle UV Absorbers: All the Same or Completely Different?

These metal oxides all rely on the same basic mechanism of protection; harmful UV radiation is absorbed by the oxide before it can inflict significant damage to the surrounding organic phase. The absorbed radiation promotes an electron from the valence band of the oxide to its conduction band, creating an electron-hole pair. From this point the picture becomes rather more complicated and there are distinct differences in the behavior and the efficacy of the various metal oxides with different coating formulations. Once created, holes and electrons can recombine and be eliminated; the energy released in the recombination is given up as light (radiative recombination) or heat (nonradiative recombination). Alternatively, the electron/hole pair can trigger further (undesirable) reactions. The situation is complicated by the fact that most practical oxides are doped to tune their absorption properties and the fate of the electron/hole pair—usually by ensuring that there is no further reaction.

The use of the individual oxides is discussed below, along with the various advantages and disadvantages that they possess. As a Case Study, there will be full discussion of a step-by-step production of zinc oxide nanoparticle dispersion from bulk powder to formulated use, the problems encountered and their resolution discussed.

TABLE 8.2. Optical and RI Properties of Oxides.

Oxide	Band Gap eV	Wavelength nm	RI
TiO_2 (Anatase)	3.2	376 nm	2.5
TiO_2 (Rutile)	3.02	398 nm	2.5
ZnO	3.3–3.4	354–364 nm	2.0
CeO_2	3.0–3.2	376–401 nm	2.2

8.11.2. Titanium Dioxide

At the macro scale, titanium dioxide has a long history of use as a protective pigment (and whitening agent) in the paint industry. For micron-sized particles, there is little evidence for significant physiological risk. The chemistry and applications of nanoparticle titanium dioxide have been extensively reviewed by Chen and Mao [7].

Nanoparticle titanium dioxide has a relatively long industrial history and is the subject of a long-standing and often acrimonious debate between the green lobby and the nanoparticle developers concerning its use in sunscreens. The reason for its use is that it is one of the few materials that will provide protection against both incident UVA and UVB radiation. The central points of the debate are whether the small particle size provides mechanisms of damage that are unavailable to large sized particles, and whether the oxidative behavior of the material (which will be discussed later) poses a significant health hazard. Matters are further complicated by the fact that the grades of titanium dioxide used are themselves stabilized to prevent such oxidative behavior. Possibly the best advice is that if you are concerned about overexposure to the sun, don't overexpose yourself!

The chemistry of titanium dioxide is complicated by the existence of different crystalline forms. The Brookite form of the oxide is relatively unimportant, but both the Anatase and Rutile forms of the oxide show significant photolytic response, but the former shows the greater activity. The band gap of Anatase is 3.2 eV while that of Rutile is 3.02 eV. A number of explanations have been proposed to explain this difference in activity; these are summarized as:

- The rate of combination of holes and electrons is dependent on the crystal size. The unit cell volume of the Anatase crystal is greater than that of Rutile.
- Differences in the structure of the band gap (Fermi level) leads to a lower oxygen affinity for the Anatase compared to the Rutile. This potentially leads to a greater number of hydroxyl groups on the surface of the Anatase structure, with a corresponding increase in photocatalytic activity.
- Anatase possesses an indirect band gap, while the band gap of Rutile is direct. This means that the mismatch in position between the highest levels of the valence band and the lowest levels of the conduction band allows a promoted electron to maintain its position in the

conduction band for a little longer. The increased lifetime enables the electron-hole pair to demonstrate increased mobility.

As mentioned above, it should be noted that the surface activity of both the Anatase and Rutile forms of the oxide can be modified by surface treatment. To improve the dispersion characteristics and reduce the photo activity, Anatase has been coated with a variety of materials. Depending on application these include aluminosilicates, zirconates, manganese silicate and phosphates.

8.11.3. Titanium Dioxide: Photostabilizer and Photocatalyst

The photo activity of titanium dioxide does not only concern electron-hole generation and combination; the nature of the structure of the titanium dioxide (especially the Anatase form) results in relatively long lifetimes of the electron-hole pair. This allows the pair to migrate to the surface of the structure where reaction with oxygen, water and hydroxyl groups can take place, generating free radicals. These free radicals can take place in further reactions with organic species at the surface.

This effect, known as photo-catalysis, is being utilized in areas such as solar cell development, water purification and self-cleaning surfaces. As mentioned, the effect has given rise to worries about the use of nanoparticle titanium dioxide in sunscreens. This particular issue will be discussed more fully in Chapter 9 in regard to the safe use of nanoparticles.

The photo-catalytic effect must be considered when the use of nanoparticle titanium dioxide is proposed as part of a surface coating formulation. Potential formulations should be carefully tested to ensure that no photo degradation takes place.

The effects can be insidious. If the photo-catalysis destroys or yellows the coating, the tests will show the effect quite clearly. If, however, the photo-catalysis destroys whatever dispersant is around the particle, this could have more subtle effects. For example, if the coating requires the particles to be locked into the matrix to give hardness, then photo-destruction of the bond with the stabilizer will lead to a reduction in the particle-matrix bonding and, therefore, a reduction in hardness.

It is probably best to place the burden of proof onto the nanoparticle supplier. They have the ability to add dopants to inhibit the photo catalysis and in general are also supplying the dispersants. They will need considerable help from their customers to ensure that their tests are both valid and relevant.

8.11.4. Cerium Oxide

Cerium oxide is widely used as a protective agent against UV radiation and much of the early development of nanoparticle cerium oxide for industrial use was carried out by Rhodia. Nanoparticle cerium oxide dispersions are available commercially and are used widely in wood coatings. The other major areas of use include fuel additives, catalytic converters and in fuel cells.

The band gap of cerium oxide is similar to that found for titanium dioxide and the mechanism of protection is the same, but unlike the Anatase form it is a direct gap. This means that lifetime of the electron-hole pairs is much shorter and consequently there is no photo catalytic effect, although if this is required for other purposes it can be generated if the oxide is doped with the correct materials. This means that the issue of free radical damage to organic tissue does not arise; indeed, Ivanov and co-workers have proposed the use of cerium oxide as a therapeutic additive for sunburn treatment [8].

The lower RI of cerium oxide is, as discussed above, an advantage for achieving low degrees of scatter in the coating.

Illustrative of the behavior of nanoparticle cerium oxide is work originally carried out by D. Fauchadour and co-workers at Rhodia [9]. An addition of 1% (wt/wt) nanoparticle cerium oxide to typical wood stain formulations was claimed to provide excellent UV protection (1200 hours, Xenon test) when compared to the equivalent unprotected formulation and with the same formulation containing a conventional UV absorber. The results are independent of whether the formulations are water- or organic solvent-based. It was also reported that the addition of the cerium oxide improved the mechanical properties of the coated film.

8.11.5. Zinc Oxide

Commercially available dispersions of nanoparticle zinc oxide are readily available in both water and organic solvent. More recently, dispersions comprising radiation curable acrylates have also become commercially available.

Nanoparticle grade zinc oxide is commonly recommended for protection against UV radiation and is commonly found in two crystalline forms, Wutzite and Zincblende. There appears to be no significant difference in the photolytic behavior between them. The band-gap of zinc

oxide is a little greater than that of titanium dioxide and cerium oxide, and is direct in nature.

Zinc oxide, like cerium oxide but unlike titanium dioxide, shows no significant photo-catalytic activity and is therefore considered preferable for use in formulations where photo-catalytic degradation of the coating might become an issue. This is not to say that nanoparticle zinc oxide is inert. The earlier discussion on stain testing showed that a specific chemistry (such as a reaction with acetic acid) can cause problems in the coating.

The choice of UV stabilizing nanoparticle is complex, though the above summary provides some indication of key issues. The following case study is the longest in this book. Because the information is all in the public domain it provides a rare example of the typical ups and downs of a nanoparticle project. What seems a rather simple question—"Can we get UV protection by putting in some ZnO?"—turns out to be vastly more complicated.

8.11.5.1. Case Study: Looking More Closely at Zinc Oxide

An interesting study into the manufacturing, dispersion, formulation and behavior of zinc oxide nanoparticles in the dispersion form has been carried out as a joint project under the aegis of NanoCentral [10]. The original members of the consortium and their proposed roles within the scheme are described below. The structure of the proposed work illustrates the number of issues that need to be considered for a thorough

TABLE 8.3. Responsibilities within the ZnO Project.

Task	Who
Nano ZnO Preparation	Johnson-Matthey, Intrinsiq Materials
Commercial ZnO sample: disperse in acrylate	Lubrizol, Primary Dispersions, U. Liverpool
Formulate and apply acrylate coatings on plastic films	MacDermid Autotype
Characterize physical and chemical properties of powders and simple dispersions (TEM, Gas Ads, PSD, and XRD)	Johnson Matthey, Intrinsiq Materials, Intertek MSG
Characterize quality; compare R&D and commercial ZnO samples	Lubrizol, U. Liverpool, Intertek MSG
Evaluate performance of coated films (durability, UV screening, physical and mechanical properties)	MacDermid Autotype, Intertek MSG

study. It is also interesting to note that there is not a single member of the consortium who possesses expertise in all of the areas that require consideration. This illustrates the mutual benefits that can accrue from a good relationship between nanoparticle suppliers and users, Table 8.3.

8.11.5.2. Production of the Nanoparticles

Samples of nanoparticle zinc oxide were manufactured by Johnson-Matthey using the flame spray pyrolysis process. The particles were uncoated and were found to possess a typical surface area of 20 m^2/g. This was equated to a primary particle size diameter of 55 nm. The material was presented in powder form to Lubrizol for surface treatment and dispersion.

8.11.5.3. Dispersion of the Nanoparticles

The nanoparticle powder was milled and dispersed into hexanediol diacrylate (HDDA) by Lubrizol using their well-established experience and technology. The powder was dispersed using an agitator mill.

There are a number of different kinds of dispersion aids that can be employed to stabilize the nanoparticle dispersion. For this particular study the low-polarity polyester Solsperse 28000, the high-polarity polyester Solsperse 71000, and the acidic alkoxylate Solsperse AX5 were chosen from Lubrizol's extensive range of dispersants.

The particle size diameter (by volume) present in the dispersions was determined by laser diffraction. The results obtained by Lubrizol are shown. They are probably rather larger than ideal for coatings that have to be free from haze, Table 8.4.

8.11.5.4. Using the Dispersions

The Lubrizol dispersions were sent to MacDermid-Autotype for evaluation in radiation curable coatings.

TABLE 8.4. Particle Sizes for Three Samples of ZnO.

Sample	D50	D90
R44	72 nm	125 nm
R56	64 nm	115 nm
R58	80 nm	133 nm

TABLE 8.5. Four ZnO Formulations.

Item	Formulation 1	Formulation 2	Formulation 3	Formulation 4
Acrylate oligomer	51%	51%	51%	54%
HDDA	32%	32%	32%	43%
Lubrizol R44	14%			
Lubrizol R56		14%		
Lubrizol R 58			14%	
Photoinitiator	3%	3%	3%	3%
Zinc oxide Wt %	6%	6%	7%	0%
Zinc oxide Volume %	~1%	~1%	~1%	0%

The dispersions were added to model formulations comprising HDDA, a multifunctional polyurethane acrylate oligomer and an α-hydroxyl benzophenone photoinitiator. A formulation containing no zinc oxide, but with an equivalent weight of HDDA was prepared as a standard. The formulations are summarized in Table 8.5.

The zinc oxide was added to the oligomer/HDDA mixes with gentle stirring without incident. The mixes were then allowed to stand for twelve days at ambient conditions, after which time they were examined for signs of instability. Formulation1 showed significant soft settlement of the zinc oxide, while the other pigmented mixes showed no settlement. The zinc oxide appeared to disperse again quite readily when Formulation 1 was remixed.

The formulations were coated onto treated polyester film by a web coating process and the properties of the coated products were determined. Oxygen was excluded from the UV curing process.

Haze and transmission measurements were made on a Byk Gardner Haze Guard Plus photometer, while yellowness measurements were made on an X-Rite SP68 spectrophotometer. Thickness measurements were made with a microscope from sections of the coated films. A summary of these results is shown in Table 8.6.

TABLE 8.6. Properties of Coatings from the Four ZnO Formulations.

Sample	F 1	F 2	F 3	F 4	F 5
Thickness (μm)	5 μm	5 μm	5 μm	4 μm	5 μm
% Haze	8.1%	0.9%	1.0%	0.8%	0.6%
% Transmission	90.7%	91.7%	91.4%	91.5%	91.1%
Yellowness (E313)	3.53	4.08	4.12	3.18	3.73
Yellowness (D1925)	3.88	4.57	4.61	3.52	4.14

The results clearly indicate that Formulation 1 is seriously compromised in terms of clarity. From this and the settlement data it is apparent that the Lubrizol dispersion R44 is unstable when combined with the acrylate medium. This is an excellent demonstration of the difficulties in determining how to produce a nanoparticle dispersion that is useful to the formulator. There is little or nothing to indicate, prior to formulation, that the zinc oxide dispersed in Solsperse 71000 would show so little affinity for the acrylate medium. In retrospect, this is possibly another case where HSP analysis would have been useful.

A further series of coatings were prepared with greater quantities of zinc oxide being added. For this work, the Lubrizol R44 dispersion was discarded. No problems were encountered on production or storage of the new formulations which are described in Table 8.7.

It is also important to remember that UV absorbers such as zinc oxide (as may titanium dioxide and cerium oxide) can reduce the UV dose delivered to the photoinitiator and therefore compromise the hardness and abrasion resistance of the cured coating. This was investigated by carrying out a simple solvent rub test using a paper towel soaked in methyl ethyl ketone (MEK). The wet towel is rubbed across the surface of the coating with moderate pressure until the coating begins to dissolve, Table 8.8.

The increased volume of zinc oxide in the coating seriously compromises the chemical resistance of the coated film and illustrates an important point—that the addition of UV stabilizers to radiation curable formulations is a matter for careful consideration. Is there a way of remedying this seemingly impossible situation? Fortunately, matters can be improved by changing the photoinitiator from an α-hydroxyl benzophenone to a Bis-acylphosphine oxide (BAPO). The BAPO photoinitiators were developed specifically for use with pigmented coatings and their absorption spectra stretch into the visible regions of the spectrum.

TABLE 8.7. More ZnO Formulations.

Item	F 6	F 7	F 8
Acrylate oligomer	30%	30%	39%
Lubrizol R56	68%		
Lubrizol R58		68%	
HDDA			58%
Photoinitiator	2%	2%	3%
Zinc oxide Wt %	30%	31%	0%
Zinc oxide Volume %	5.5%	5.5%	0%

TABLE 8.8. More ZnO Formulation Test Data.

Sample	F 6	F 7	F 8
Thickness (μm)	4 μm	4 μm	4 μm
Haze (%)	2.2%	2.7%	0.7%
Transmission (%)	89.1%	89.6%	90.5%
Yellowness (E313)	Not tested	7.43	3.23
Yellowness (D1925)	Not tested	8.11	3.56
MEK resistance	5 rubs	5 rubs	50 rubs

On substituting such a photoinitiator into the formulation, the MEK resistance of the coatings containing the zinc oxide is increased to > 50 rubs. This further illustrates the necessity of comparing the absorption spectra of photoinitiator and nanoparticle additive when formulating a radiation curable coating. This will enable the formulator to ensure that the correct choice of ingredients is made. The UV-visible transmission spectra for the coatings containing the Lubrizol zinc oxide dispersions can be viewed later in this chapter.

The results obtained with these coatings highlight several important issues that the formulation scientist must confront to be successful:

- The stability of the dispersions towards additions of the other components of the formulation must be examined. The failure of dispersion R44 is a classic example of this. Considering the case of acrylates alone, it is perfectly possible for the dispersions R56 and R58 to be stable in the multifunctional urethane acrylate chosen for this work; it is not guaranteed that they will be stable in alternative acrylates. This was found to be the case when an alkoxylated pentaerythritoltetraacrylate (PPTTA) was substituted into the mix.
- It is quite possible that any instability of the mix will not be immediately apparent. Therefore, timed studies should be made to observe for any signs of developing instability. If the formulations become subject to "soft settlement" and can be re-dispersed readily on stirring, then care should still be taken. It is quite possible that the material will appear to be usable; but significant changes may have taken place to the particle size distribution. This is well illustrated by the result obtained with the redispersed R44.
- In some cases a glass-clear coating is required; this means that any slight haziness in the final coating cannot be tolerated. Generally speaking, a haze figure of about 1% is considered to be the upper

limit. It can be seen from the data that the 30% (wt/wt) additions of zinc oxide exceed this figure considerably; visually these coatings possess a slight but distinct haziness. As discussed above, the relatively large size of these particles makes them harder to disguise.

- It is important to remember that if a UV curable coating is being formulated, the presence of a UV absorber in the formulation can pose a problem in terms of curing of the coating. It is vital that the absorption spectra of the UV absorber and photoinitiator are examined to ensure that there is a sufficiently large window for curing to take place.

8.11.5.5. Analysis of the Nanoparticle Powder, Dispersion and Coatings

To obtain a fuller picture of the zinc oxide nanoparticles' behavior in the powder, dispersion and coating samples of each item were forwarded to Intertek MSG for analysis. The analytical techniques chosen for the study are shown below. This investigation presents the fullest picture of the behavior of the zinc oxide nanoparticles from their production in the powder form to their incorporation within a coated polymer film.

The dispersions were all found to be readily diluted with no instability problems apparent. The diameters of the particles were studied using photon cross correlation spectroscopy (PCCS), which measures the hydrodynamic diameter. This measurement includes the contribution from the dispersant surrounding the nanoparticle.

The particle size distributions were also measured by the disc centrifuge method; this measures the Stokes diameter which, in this case, is smaller than the hydrodynamic diameter measured by PCCS. The full range of measurements, including those originally obtained by Lubrizol, are summarized below.

The larger size of the Lubrizol R58 is attributed to the nature of the dispersant, which is a single chain anchor compared to the branched multi-chain anchors present on samples R44 and R56.

There is a reasonable degree of agreement between the Lubrizol results and those obtained by Intertek, Table 8.9.

The dried dispersions were examined by transmission electron microscopy (TEM). Three distinct particle types were observed: zinc oxide nanoparticles, larger zinc oxide platelets & organic material in the form of rods. An image of this arrangement is shown below. The large

TABLE 8.9. Particle Size Measurements Using Different Labs and Techniques. The Sizes are in nm Representing Mean Volume and Mean Number Values.

Sample	Laser diffraction D50 Lubrizol	Intertek MSG PCCS	Intertek MSG Disc Centrifuge
Mean Volume			
Lubrizol R44	72	75	59
Lubrizol R56	64	73	57
Lubrizol R58	80	91	56
Mean Number			
Lubrizol R44	33	46	39
Lubrizol R56	34	35	38
Lubrizol R58	41	35	35

particle visible in the central area of the bottom part of the photograph is approximately 200 nm in length. It was found that more clusters of particles were found in the Lubrizol R44 dispersion than in the other two dispersions. The discussion on sizing in Chapter 2 shows that the relatively large rods will have had a disproportionate effect on the value of D50. Perhaps the disk centrifuge results are somehow missing out on these larger particles, so D50 is significantly reduced, Figure 8.16.

The UV-visible transmission spectra for the coatings containing zinc oxide nanoparticles were determined by Intertek. The results covered coatings that contained 6%, 16% and 30% by weight of nanoparticle

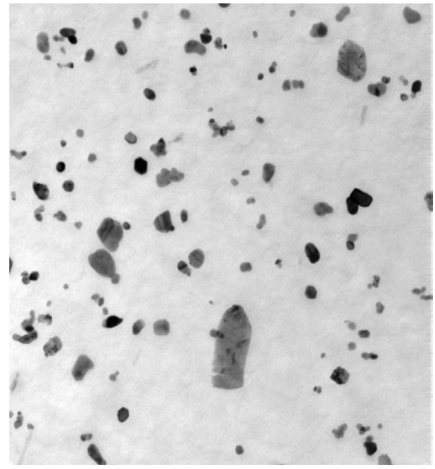


FIGURE 8.16. *The ZnO particles showing a wide variety of sizes/shapes.*

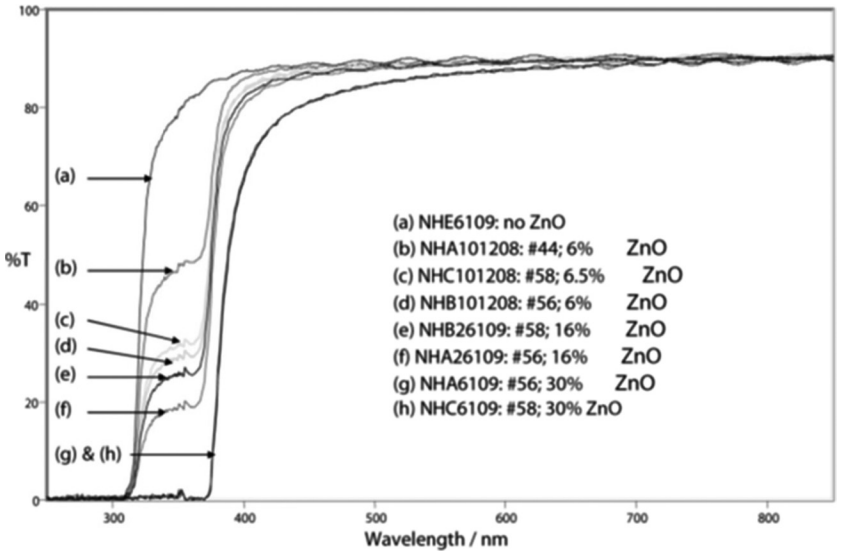


FIGURE 8.17. *UV-Visible Transmission Spectra for Different ZnO Coatings.*

zinc oxide derived from the Lubrizol dispersions R44, R56 & R58. The spectra are shown in Figure 8.17.

The increasing level of zinc oxide in the coatings is readily observed in the changes to the transmission spectra of the coatings. While there is no transmission of radiation below 380 nm for the coatings containing 30% zinc oxide, a possible curing window exists for the coatings containing less zinc oxide. The transmission curve shown as (b) is rather different in shape from the rest of the series and is explained by the unstable nature of the formulation used to produce the coating.

8.11.5.6. Abrasion Resistance

The solvent resistance of the coatings all appear satisfactory, but does the addition of the zinc oxide compromise their abrasion resistance? To determine whether this is the case, a series of Taber abrasion tests were carried out over 100 revolutions using two 500 g weights and CS10F wheels. The resulting change in haze values on abraded areas was measured and no serious deterioration in the abrasion resistance of the coatings was observed. This indicates that it is possible to formulate a UV curable coating containing a nanoparticle UV absorber without compromising the cross-link density (assumed to be necessary for such resistance) of the coating.

8.11.5.7. Getting Real with Zinc Oxide

The previous work carried out as part of the NanoCentral partnership is a valuable exercise, but how does this relate to commercially available zinc oxide? Fortunately, there are grades of material that are commercially available as dispersions, and these were tested in a similar manner to the experimental materials provided by Lubrizol.

The NanoCentral study was concluded by an evaluation of the accelerated environmental ageing of model systems prepared for the study. The study was extended to include other zinc oxide nanoparticle dispersions which were becoming commercially available around the time of this study. These included two dispersions in HDDA (20% and 35% solids content) which were provided by IbuTec with a particle diameter claimed to be 10 nm. A dispersion was provided by Byk comprising a 30% (wt/wt) dispersion of zinc oxide in tripropyleneglycol diacrylate (TPGDA).

All of the formulations were coated onto adhesion-treated polyester film using the MacDermid-Autotype UV pilot coating machine. A single consistent coating thickness of 5 μm for all of the formulations was sought and achieved. Because of differences in solids and viscosity the model formulations are not identical; however, they are sufficiently similar for general observations and comparisons to be made.

Two types of testing were carried out: combined humidity and UV testing using an Atlas UVCON test cabinet and accelerated sunlight testing using a Heraeus sunlight test cabinet. The Atlas unit cycles between 100 hours of applied humid conditions followed by 100 hours of exposure from the UV lamps. The Heraeus unit runs for a set time (100 hours in this case) and is then reset for further evaluation.

The samples were tested for changes in adhesion using the Sheen cross-hatch test method and their optical properties were measured as above.

To summarize a vast amount of work in a few sentences:

- Adhesion and abrasion performance remained good for formulations containing 30% ZnO—with adhesion falling off significantly with lower levels
- Yellowness remained unchanged even with modest levels of ZnO
- The biggest, and most unexpected, problem was the large increase in haze during the ageing. Presumably the particles are sufficiently mobile that they can clump together.
- Similar coatings containing no additives or HALS alone failed much earlier in the testing cycle.

- A surprising way to combine the best of both worlds was to have a 3% loading of ZnO along with 4% loading of HALS.

The reader may not particularly care about the specifics of ZnO protection. The point of the Case Study is to emphasize that life with nanocoatings is not easy—even when a large team of specialists focuses a lot of resources on getting things right.

Although these results are encouraging, it is important to remember that accelerated ageing tests do not fully replicate the conditions of real time outdoor ageing. Consequently, real time ageing results are required before drawing firmer conclusions—or, possibly, smarter ways to predict lifetimes are required. This theme will be explored after a major diversion into the issues of wood protection.

8.12. WOOD

This section appears a little out of character for this chapter. The reason for including it is that possibly the largest single current use for nanocoatings is in the wood industry, and it seems better to cover this important topic within one area rather than scattered through the book.

During 2008, a debate was begun in the letters page of Nature Nanotechnology into whether the commercial impact of nanotechnology mirrored the investments made. The number of negative responses prompted a response from Prof. Phillip Evans of the University of British Columbia, who made the point that the use of nanoparticles in the commercial treatment of timber appeared to have escaped the attention of most commentators [11]. The market for treated wood in North America was estimated to be of the order of $4.9 billion and about 20 million m^3 of wood were treated with aqueous copper preservatives. Evans continued by asserting that an increasing number of companies were producing commercial products whose active agents included nanoparticle copper carbonate.

This is a fascinating piece of information, which if correct, proves that nanotechnology can be harnessed to produce commercially viable products. This is also reflected in the commercial literature available from nanoparticle suppliers which specifically advertise nanoparticle additives for use in the wood industry.

Evans made the further interesting assertion that the number of scientists working in the field of wood protection is relatively small and that they potentially have little experience with nanotechnology. This

may underestimate numbers, for it is quite possible that a significant presence is invisible; commercial organizations tend not to advertise their technical expertise directly through publication.

So here is cutting-edge technology allied to a mature commercial market. It might seem almost boring compared with the excitement that is generated by the latest graphene paper; the technical problems mundane—it is only paint, after all. The truth could not be more different. Ignoring for the moment the complexities involved in the successful introduction of nanoparticles into paint, wood is an extremely difficult substrate with which to work. For simplicity, while acknowledging that wood coatings come with affixed commercial names like primer, stain, coats, etc. these will be combined together under the general heading of paint, apologizing to paint technologists who make subtle distinctions between these designations.

Wood is a natural material and provides a serious challenge for formulation scientists because there are many species of wood that are used commercially, both soft and hard. The number of different species of wood used commercially is greater than the number of commercially available plastic films. Besides considering the physical and chemical differences between different species of wood, there are many differences in how these surfaces have been treated and cared for prior to the coating process. It is also important to realize that the choice of the variety of wood might well differ with geographical location. Finally, the fact that the wood is porous provides both opportunities (the "coating" can be fully imbibed by the wood for greater protection) and problems (a fine gloss coating might be absorbed before it reaches its final gloss form).

A key reason for using nanoparticle coatings for wood is that active chemicals such as copper carbonate are in "insoluble" form and therefore do not leach out. The quotation marks around insoluble are a reminder of the issue discussed in terms of antimicrobials: insolubility is "good" because it keeps the active material in the coating; it is "bad" because the active material might have no chance to migrate to where it can kill a microbe.

The nanoparticles exist to improve durability of the wood. This can involve the performance of the coating as a barrier to moisture, fire retardation, the reduction of exposure of damaging UV radiation, the prevention of fungal growth and the improvement of abrasion resistance. Where possible, the different requirements for improved durability should be met with multi-functional formulations. The applications

TABLE 8.10. Nanoparticles used for Wood Treatment.

Property	Nanoparticle
Fire retardment	Clays, Silica, Titanium Dioxide
Hydrophobicity	Clays, Cerium Oxide, Titanium Dioxide
UV Protection	Titanium Dioxide, Zinc Oxide, Cerium Oxide, Clays
Biocidal activity	Silver, Copper salts
Abrasion resistance	Silica, Alumina
Photocatalytic activity	Titanium Dioxide

and possible approaches to these issues have been summarized by Fufa [12], Table 8.10.

It should be noted that not all of these ideas have reached commercial fruition. Fufa also discusses the efficacy of two paints, each claiming to contain nanoparticles. The treated wood was found to possess unimpressive moisture resistance. However no details as to the composition of the paints used is presented.

It is interesting to note that Evans' observations concerning the understanding of nanotechnology are reflected in the number of published papers within the field of wood coating technology. Most of the relevant papers have been published relatively recently—mainly within the last 5 years. This of course does not take into account "in-house" knowledge that is not publicly available.

A further general point is the reporting of particle size in many papers; it is not unusual for particle size ranges from tens of nanometers to tens of microns to be quoted. At the lower end of the particle size range, this rather conflicts with the study carried out by Schilde that is referred to in Chapter 3. This is encapsulated in a study carried out by Sow and co-workers into the efficacy of silica and alumina nanoparticle addition to UV-curable, waterborne, polyurethane acrylate-based wood coatings [13]. The purpose of this study was to investigate the possibility of improving the abrasion resistance of wood coatings—a common aim of nanoparticle research.

The study used commercially available raw materials to produce the waterborne dispersion, while the nanoalumina and silica used was Aeroxide Alu C and Aerosil 7200. The nominal particle size reported for these materials is 13 nm and 12 nm respectively, though these tend to be the "primary particle size" and the real particles tend to be agglomerates of much larger size. Humidity conditioned Sugar Maple (Acer Saccarum) was used as the substrate. The study found that significant

agglomeration of the nanoparticles occurred during the high speed dispersion process; agglomerates of >10 μm being present. The presence of these agglomerates significantly reduced the hardness of the coating. The difficulty in dispersing the nanoparticles within the coating medium was highlighted and this should come as no surprise to the reader. Improved performance was claimed when sonication was used as a dispersion method.

Improvements in performance were claimed when a surface treatment comprising a methacrolylsilane was grafted onto the surface of silica. The amounts used in these studies were rather low (1% to 5% by weight) compared with the relatively large additions considered necessary for abrasion enhancement on plastic films. There is, however, no reference to the thickness of the coatings tested and it is also a little unclear as to whether the tests were carried out on wood or glass. Further studies in this field would be of interest.

The preservation of wood against microbial attack is an active area of research and nanoparticle-based formulations have been examined in this role. A study by Bak and co-workers has shown that the resistance of various hard and soft woods to a brown rot fungus (*Poria placenta*) can be significantly improved by the addition of small amounts of nanoparticle zinc (20 nm–40 nm size distribution) [14].

It is known that copper salts can act very efficiently as wood preservatives. In a paper, Evans and co-workers set out to investigate the distribution of copper carbonate and iron oxide nanoparticles in treated wood [15]. The source of the nanoparticles was an un-named commercially available wood treatment paint. It is important to realize that previous studies have shown that the efficacy of wood preservatives is related to the ability of the active ingredients to penetrate cell walls and bind to the structural components (cellulose, etc.) of the wood.

The typical size of the diverse structures present within the wood studied (Southern Pine) was of the order of about 1 nm to > 1 μm, while the particle size distribution within the commercially available paint was claimed to range from 1nm to > 20 μm. It was found that the copper particles were not uniformly distributed within the wood, but accumulated in voids that act as flow paths for liquids within the wood. Particles were also found to deposit on cell walls, but not within the walls. However, elemental copper was detected within the walls, although at lower levels than are found with conventional copper paints. It is postulated that the copper carbonate nanoparticles are converted to ionic copper, which then penetrates the cell wall.

It is unfortunate, in scientific terms, that a more controlled particle size distribution was not available for the investigators' use. It would seem that, at present, the commercially available paints show a surprisingly wide particle size range. This might seem a little unnecessary as commercially available nanoparticles are available with controlled size distributions. However, it is an open question whether these materials would be suitable for formulation as paint. Nonetheless, nanoparticle cerium oxide of controlled size distribution is commercially available and is advertised as a UV protective agent for wood paints. The whole question of nanoparticle usage within the industry is probably best addressed by the paint manufacturers, rather than the wood scientists.

Returning to the previous theme of UV resistance, there are numerous papers on TiO_2, ZnO and CeO_2 nanoparticle coatings that provide good protection at relatively low levels—presumably because the formulations can disperse within the upper layers of the wood so that the total amount of nanoparticle is as large as in a few μm coating on a plastic surface. A typical example is that provided by the Riedl group, who show the importance of dispersion quality on performance [16]. Using predispersed ZnO nanoparticles (typically in the 20 nm range), it was possible to obtain high resistance characterized by color change, adhesion and gloss retention. Using nanoparticle powders, with their much larger (agglomerated) particles sizes the results were generally much inferior.

To conclude this brief survey of the field it is clear to the authors that wood coatings provide an example of the industrial application of nanotechnology. It might not garner headlines, but it is a field of work that is as challenging as more trendy proposed uses of nanoparticle technology, and more importantly it appears to be successful. It is therefore a little frustrating that more papers on the topic are not publicly available. This probably reflects Evans' comments at the start of this section and also the strong possibility that expertise is buried out of sight within the companies operating in this field. Perhaps even more progress would be made if this field of work were more widely recognized within the nanoparticle community.

For wood nanotechnology as with many others, one of the key barriers to progress is finding ways to prove that the technology delivers lifetime benefits. Although there are many standard tests used in wood research, they suffer from the familiar problems of providing too little information for too much work. The following section returns to the theme mentioned previously about the need to find a general purpose, more rational and high-throughput approach to lifetime testing.

8.13. LIFETIME TESTING

As customers we love to have guarantees from our suppliers that our raw materials will have an infinite storage lifetime. As suppliers, we hate the immense amount of hard work required to be able to assure our customers that our product will have an adequate lifetime within their application.

There are two interconnected reasons for hating the work of providing lifetime data. The first is that no one really knows what test criteria to apply. The second is that no one really knows how long to do the test within a practical timescale so that a real-life lifetime is guaranteed.

Tests emerge from within an industry and they tend to require 30-60 days at elevated temperatures and humidity. The test equipment (big, energy-consuming cabinets) is invariably full with other tests and/or set at the wrong temperature or %RH for the specific test required by the project we are working on, so we either have to wait for our own control of the cabinets or buy in more. Even worse, at the end of the 60 days we open the cabinet and find that our sample has failed—sending us back to the lab and at least another 60 days of testing.

The whole system would be much more rational if we could apply some simple Arrhenius rule such as "a 10°C increase in temperature equates to a doubling of the rate of deterioration". Generally we know that this doesn't apply to our particular system, and even if it applied for temperature it certainly doesn't apply to %RH, where going from, say, 60% to 80% can lead to a catastrophic rise in failure rates.

The result of this is that the coating industry is full of dissatisfied project teams feeding samples into expensive test equipment and getting relatively unhelpful Yes/No data after unacceptably long periods of testing.

What is needed is a predictive method that produces meaningful data in, say, 14 days which then allows extrapolation into an arbitrary future. For example, it would be nice if we could look at our shipping system and ask: we have 10 days at 25°C and 50%RH, 1 month at 35°C and 70%RH, 2 weeks at 40°C and 80%RH then 4 months at 30°C and 60%RH, (representing various aspects of shipping and storage before use), so will our product still be in spec?

The pharmaceutical industry faces the same difficulties as the surface coating industry. They need to put potentially unstable drugs into various forms of packaging, guaranteeing that they will be safe to use after, say, 2 years, then ship them across the world. They have, for de-

cades, used the same crude accelerated ageing techniques—with large numbers of temperature/humidity cabinets generating small amounts of data over large amounts of time.

A team at Pfizer, headed up by Dr. Ken Waterman, reinvigorated a technique that has been semi-known for some time and demonstrated its power and simplicity in providing lifetime predictions within, typically, 14 days [17]. Although they could have kept the technique as a form of competitive advantage, it was decided that bringing it into the public domain would encourage, for example, the FDA to accept results gained by this methodology.

It seems to the authors that the Pfizer technique, called ASAP (Accelerated Stability Assessment Program) is of profound importance for readers of this book. With warm acknowledgement to Pfizer and to Dr. Waterman's present company, FreeThink Technologies, we present our own take on ASAP [18]. Those who think the technique might apply to themselves can follow up with the literature and real experts in ASAP if they so wish.

A simple diagram describes the typical 30-day or 60-day approach, Figure 8.18.

The graph shows that after 30 days and 60 days some data are measured about the extent of bad stuff happening: loss of adhesion, increase in yellowness, degradation of a coating, etc. The experiments have been

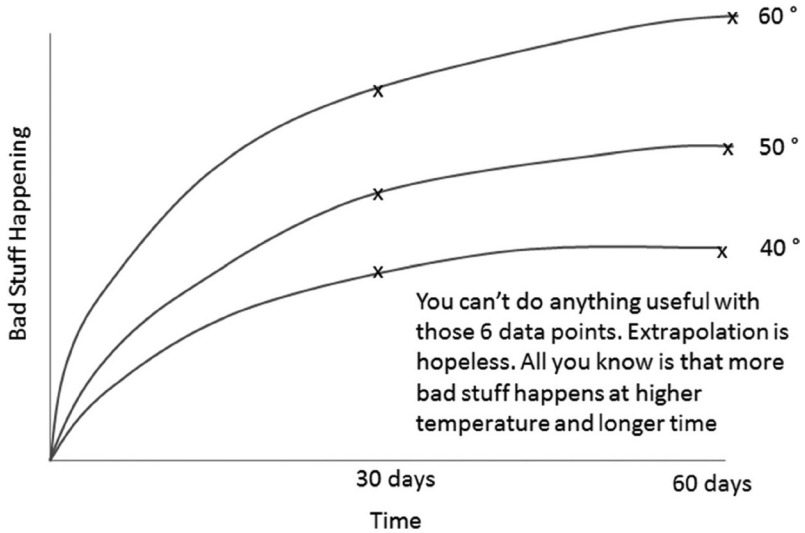


FIGURE 8.18. *A typical set of data from lifetime tests; the curves are "guides for the eyes", as in reality nothing much is known about what is going on.*

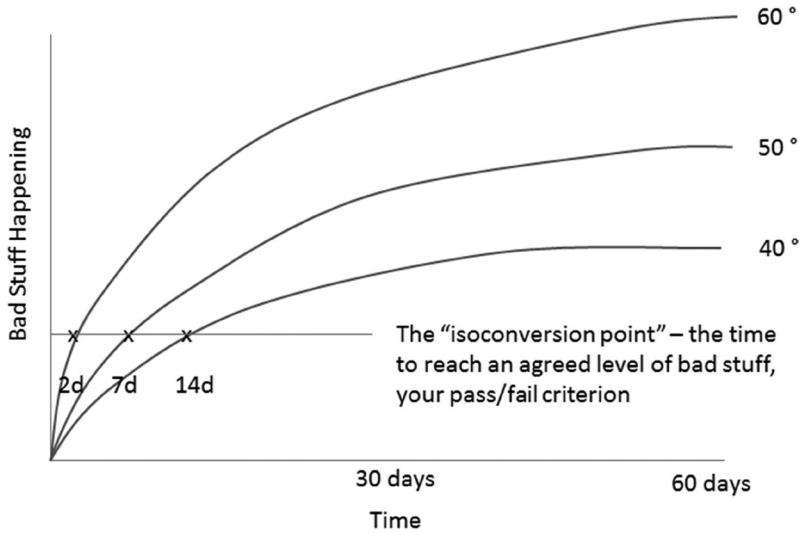


FIGURE 8.19. *Measuring the isoconversion point gives data that can be used for further analysis, even though the curves are merely "guides for the eye".*

done at 40°C, 50°C and 60°C and sure enough the bad stuff happens faster at higher temperatures. The curves are purely speculative—there are no intermediate data points. Importantly, there is no way that any meaningful extrapolation can be made from these data. Other than telling us whether we've passed or failed some criterion at *X* days and *Y*°C, we have learned essentially nothing after 60 days.

Now we do something slightly differently. Set a failure criterion: adhesion loss, color change, physical degradation—anything that can be measured objectively. Now see how long it takes to reach that criterion under different test conditions—deliberately designed to ensure that failure happens rather quickly. In the graph below all that is known are the times to failure—the curves are "guides for the eye", Figure 8.19.

From this experiment the "isoconversion point" at which the samples failed is respectively 2 days, 7 days and 14 days at the different temperatures. What the ASAP team recognized is that irrespective (within reason) of the precise shape of the deterioration curves, the data can now be meaningfully extrapolated within the context of Arrhenius. So activation energies and extrapolations to different times and temperatures are scientifically meaningful. Rational lifetime testing starts to become a possibility.

The problem is that variation of %RH is not included in the plot. In fact, in order to get failure early on, the chances are that the %RH

had to be at a high level. If the 40° test had been run at low %RH the isoconversion point would not have been reached within the reasonable 14 day timescale.

The beauty of ASAP is that for most systems that Pfizer have tested (and it seems reasonable that their assumption will apply to many other systems) the rate of deterioration, k, towards the isoconversion limit can be approximated by:

$$\ln(k) = \ln(A) - \frac{E_a}{RT} + B(\%RH) \tag{8.11}$$

Here there are three unknowns. A is a constant. E_a is the activation energy for the deterioration process (i.e., this is the Arrhenius part of the formula) and B is a constant controlling the effect of %RH. The equation looks to be linear in B, which is not our intuition. But k is *exponentially* dependent on B, so small increases in B can, as we know, lead to large increases in deterioration.

We have an equation in three unknowns. So in principle, just 3 experiments at appropriate T and %RH values can characterize the whole system. In practice, a few more experiments will help with statistical certainty and it needs some skill and experience to find a range of T and %RH values that give statistically meaningful results—and will do so within both the meaningful timescale (a maximum of, say, 14 days) and a meaningful range of temperatures (e.g., 40–90°C) and humidity (e.g., 10–90% RH). This in turn requires a meaningful definition of failure within the context of the product.

Using sample data from Waterman and following things within the ASAP spreadsheet, here is how the whole thing works in practice. A set of experiments at the following temperatures and humidities failed after the given number of days, Table 8.11.

TABLE 8.11. Time (Days) to Isoconversion Point for Five Sets of T And %RH Values.

T	%RH	Days
50	75	14
60	40	14
70	5	14
80	75	1
90	40	2

TABLE 8.12. Fitted Values from the Above Dataset.

ln*A*	E_a kJ/mole	B
34.49819	99.64665	0.031032

Using the Excel Solver and the above equation, it is possible to fit the data to the following values, Table 8.12.

Note: Like all numerical solvers, the results are meaningful only if the initial guesses are not too bad. The user enters values that give at least vaguely reasonable fits, then asks Excel to fit. If the guesses are too wild, the fitting algorithm will head off into bizarre territory. This is not a problem with ASAP; it is a general issue with non-linear optimization.

Using the built-in Quick Calculator it is possible to enter a constant T and %RH and get an estimate of how many days it will take before failure, Table 8.13.

In this example the material would fail after about 1 year at these conditions.

Just as importantly, a performance scenario can be entered of so many days at a given temperature and humidity. The scenario can be of arbitrary complexity. The only thing required is a definition of "failure"; in this case the value is set at "10" and all the previous fitting has been based on that value, Table 8.14.

In this case, failure took place ~130 days, with the largest single contribution to failure being the 30 days at 35/60 which contributed 2.18 towards the total of 10.

More sophisticated statistical techniques (not included in the spreadsheet) can make use of extrapolations or interpolation to find the isoconversion point, so data gathered on either side of the isoconversion

TABLE 8.13. A Prediction of Time to Failure at a Given Set of Conditions.

	Quick Calculator	
	T	**%RH**
	30	**50**
k	0.028079	per day
Days	356	to Failure

TABLE 8.14. Time to Failure (~130 Days) When Conditions Change During Shipping/Storing, etc.

Days	T	%RH	In Period	Sum	Total Days
10	25	50	0.14	0.14	10
2	45	75	0.79	0.94	12
30	35	60	2.19	3.12	42
50	30	45	1.20	4.32	92
2	40	80	0.51	4.83	94
3	45	80	1.39	6.21	97
30	35	60	2.185876	8.40	127
2	40	80	0.505239	8.90	129
3	45	80	1.385185	10.29	132
30	35	60	2.1185876	12.48	162
50	30	45	1.202186	13.68	212

point can be put to good statistical use. Techniques can also address an important question: what happens if, after 14 days, essentially no change in properties is noticed? The answer is that it is possible to produce reasonable combinations of the three parameters that *just* give a minimal change after 14 days, and from those estimate that the probability of failure after, say, 2 years is <1%—in other words, from negative results it is still possible to draw positive conclusions.

This brief introduction and a relatively simple (though still surprisingly powerful) spreadsheet give only a taste of what ASAP can provide if it can be shown to apply to the nanocoating industry. The intent of this section is to alert the industry to the potential of the technique so that it can be discussed internally, with customers and with standards committees. There is no way that anyone can adopt ASAP without a lot of industry-wide work. Given that a journey of 1000 miles starts with the first step, maybe it is time for that first step to be taken.

8.14. CONCLUSION

A journey into nanocoatings can still be an adventure into the unknown. Many issues can be resolved rationally, while there are many issues for which there is no clear methodology for ensuring that the nanocoating formulation is optimal. When the team is clear on what is and isn't possible to verify objectively, then it can better decide on the risks/rewards of pursuing a specific strategy. By identifying the gaps in understanding, new techniques for filling them will emerge.

8.15. REFERENCES

1. Jonathan Moghal, Johannes Kobler, *et al.* (2012) High-performance, single-layer antireflective optical coatings comprising mesoporous silica nanoparticles. *ACS Applied Materials and Interfaces* 4:854–859.
2. B. Gornickaa, M. Mazur, *et al.* (2010) Antistatic properties of nanofilled coatings. *Acta Physica Polonica A.* 117:869–872.
3. Siegfried Schmauder and Markus Meyer. (1992) Correlation between Dundurs' parameters and elastic constants. *Zeitschrift für Metallkunde* 83:524–527.
4. W.C. Oliver and G.M. Pharr. (1992) *Journal of Materials Research* 7:1564–1583.
5. S.V. Hainsworth, H.W. Chandler and T.F. Page. (1996) Analysis of nanoindentation load-displacement loading curves. *Journal of Materials Research* 11:1987–1995.
6. John Scheirs. (2000) *Compositional and Failure Analysis of Polymers: A Practical Approach.* New York: John Wiley & Sons.
7. Xiaobo Chen and Samuel S. Mao. (2007) Titanium dioxide nanomaterials:? Synthesis, properties, modifications, and applications. *Chemical Reviews* 107:2891–2959.
8. N.M. Zholobak, V.K. Ivanov, *et al.* (2011) UV-shielding property, photocatalytic activity and photocytotoxicity of ceria colloid solutions. *Journal of Photochemistry and Photobiology B: Biology* 102:32–38.
9. D. Fauchadour, T. Jeanson, *et al.* (2005) Nanoparticles of ceranium oxide—Applications to coatings technologies. *Paint & Coatings Industry* 21:88–93.
10. http://www.nanocentral.eu/ Accessed June 2012.
11. P. Evans, H. Matsunaga and M. Kiguchi. (2008) Large-scale application of nanotechnology for wood protection *Nature Nanotechnology* 3:577.
12. S. Fufa. (2010) Nanobased coatings and the influence on the hygroscopic properties of wood. *International Conference on Nanotechnology for the Forest Products Industry.* September 27–29, 2010, Espoo, Finland.
13. C. Sow, B. Riedl and P. Blanchet. (2011) UV-waterborne polyurethane-acrylate nanocomposite coatings containing alumina and silica nanoparticles for wood: Mechanical, optical and thermal properties assessment. *Journal of Coating Technology Research* 8:211–221.
14. M. Bak, B.M. Yimmou, K Csupor, R. Nemeth. L. Csóka. (2012) Enhancing the durability of wood against wood destroying fungi using Nano-Zink. *International Scientific Conference on Sustainable Development & Ecological Footprint.* March 26–27, 2012. Sopron, Hungary.
15. P. Evans, H. Matsunaga and M. Kiguchi. (2009) Microdistribution of copper carbonate and iron oxide nanoparticles in treated wood. *Journal of Nanoparticle Research* 11:1087–1098.
16. M. Vlad, B. Riedl and P. Blanchet. (2010) Enhancing the performance of exterior waterborne coatings for wood by inorganic nanosized UV absorbers. *Progress in Organic Coatings* 69:432–441.
17. Kenneth C. Waterman, Anthony J. Carella, *et al.* (2007) Improved protocol and data analysis for accelerated shelf-life estimation of solid dosage forms. *Pharmaceutical Research* 24:780–79.
18. http://freethinktech.net/. Accessed June 2012.

Dealing with Nanosafety

THERE is persuasive evidence that once the nanoparticles are bound into a nanocoating the nanosafety issues are very much reduced. This is helpful to know but the team still has to grapple with a large number of safety issues that the word "nano" raises. The aim of this chapter is to put nanosafety into a broad context so the whole team is aware of the issues involved.

9.1. IS NANOTECHNOLOGY SAFE OR UNSAFE?

> "There is nothing inherently sinister about nanoscience or nanotechnology, it just refers to the study of things on the scale of one-millionth of a millimetre." —Lord May, president of the Royal Society: 2005.

> "We are on the right path to dealing with the problem (of nanomaterial safety), but we're sauntering down it when a sense of urgency is required" —Sherwood Boelert, Chair of the US House Science Committee: 2006.

In some respects, this chapter of the book has proven to be the most difficult to write; health and safety assessment takes up an increasing amount of time within industry (and life in general), but the question "Is a new material safe to use?" is not quite as straightforward as it might seem. It is not the authors' intention to answer the question "are nanoparticles safe?". The context within which the particles are being used, their size, their shape, exposure times, mobility, solubility, and degradability are just some of the factors that must be considered to gain an idea of the scale of the question. It is a fact that materials that are considered innocuous on the macro scale may be toxic at the nano scale. An example of this is elemental gold, which is relatively inert in

its bulk form, but which is extremely active at the nano scale. The situation is nicely summed up in a Nature editorial entitled (after Paracelsus): "The dose makes the poison" [1].

Of course, this is not only a question that should be asked of nanoparticles; previously, commonly used materials such as CFC refrigerants, lead paints, tobacco and asbestos were only discovered to be threats to health and the environment after many years of general use. It is therefore a good thing that nanoparticle safety is taken seriously and that much current research in the field of nanotechnology is directly related to studying the impact on health and the environment. Obtaining an understanding of how these materials interact physiologically and ecologically is a worthy goal.

There are various strains of argument against nanoparticle technology and its use. One strain is the perfectly accurate statement that there is much that is unknown about these materials and that care should be exercised before application. Another strain is one of total rejection. This latter view is related to a belief that nanotechnology is inherently unnatural, possibly evil and that only natural (or "green") means should be used to solve the problems that presently confront us. Technology is seen as the problem, rather than the solution to the problem, the case against being presented as a polemic rather than a reasoned critique. This line of argument is exemplified by the reaction to safety concerns about the notorious Magic Nano bathroom cleaning product, which was finally shown to contain no nanoparticles at all [2].

Associated with the above example of willful mis-marketing (and its consequences), there are also scare stories with headlines that are designed to grab attention. A typical example of this is the following headline from 2011: "Biosecurity Expert Fears a Nanoparticle Attack. Microscopic bits of metal that float in the air in the aftermath of a terrorist attack could become a threat to national security" [3]. As this chapter shows, a terrorist hoping to create a nanoparticle attack would probably be much better served by purchasing an old diesel burning car to drive around the streets.

A well-educated friend of one of the authors firmly believed that nanoparticles were synthetic materials and was concerned that they presented a significant health risk. She had gained this "understanding" from general articles presented in the media and was shocked and surprised when informed that she was at that moment surrounded by nanoparticles that were not manufactured commercially. Answering her concerns about health risks requires the rest of the chapter.

Leaving aside the debate as to the general level of science literacy of the general public, this small incident illustrates the problems that need to be addressed if a sensible debate on nanoparticle health and safety is to be publicly held and where the "public" includes the team, their colleagues and their customers. This is not a platitude for, unfortunately, there is a far darker response to the debate concerning nanoparticle safety and indeed science in general. Attacks by terrorists acting under an anarchist umbrella have been reported against scientific facilities and also against individual scientists. This retreat from reason is discussed by Phillips in Nature [4].

This chapter should dispel some of the misconceptions concerning nanoparticle safety and, more broadly, nanotechnology in general, yet also highlight the fact that there is still much that is unknown. The latter part of the previous sentence should not be interpreted as a fear of the future, but a simple illustration of the fact that science operates at the border with the unknown.

9.2. UBIQUITY

Let us first remove some misconceptions; nanoparticles were not developed by evil industrial corporations bent on ruining the lives of blameless happy peasants. Neither were they created by mad scientists whose pride overwhelmed their humanity. They are natural. They have been around since the formation of the earth and they surround us in the form of dusts, mists and fogs.

The safety and environmental issues associated with such natural phenomena garner fewer headlines than what might be described as Grey-Goo stories. This is unfortunate, as it overlooks a serious question: do these materials, that are ubiquitous in nature, show significant biological and environmental effects?

The last question is not rhetorical; it confronts serious issues. At the individual level it is not only the inherent physical and chemical properties of the nanoparticle that must be considered, but also the possibility that reactive free radical species present in the atmosphere might become bound to the surface of the particle, creating further reactive species, which are then carried into the lungs or through the skin [5]. We must however also consider the bigger picture: how do these materials interact with the environment and is their degradation and accumulation pathway threatening in any way?

Attempts are being made to provide some answers to the questions

posed and to improve our understanding of these matters. To deal with this subject comprehensively would require a separate book, but short discussions on several interesting areas are provided to illustrate some pertinent points.

9.2.1. In the Home

An interesting paper from Wallace and Ott discusses personal exposure to ultra fine particles (UFP) while carrying out such every-day activities such as cooking, driving, ironing and dining in restaurants [6]. Needless to say, significantly large figures were obtained for activities related to smoking.

The diameter of the particles defined in this study range from 1μm down to 10nm. Cooking on gas and electric stoves gave rise to a peak personal exposure value of 10^5 particles/cc, with estimated emission rates of about 10^{12} particles/min. A variety of restaurants were monitored during the trial; UFP exposure values were found to be maintained between 5×10^4 and 2×10^5 particles/cc for the duration of the meal. Typical background counts without activity taking place were found to be of the order of 2 to 3×10^3 particles/cm.

The figures presented in the study show conclusively that those everyday activities, which one might find difficult to imagine being injurious to health (except of course smoking), are significant sources of nanoparticle production. This study does not define the chemical nature of the UFPs detected, nor does it define the shapes of the particles that are detected. Nevertheless, the ubiquity of nanoparticles arising from everyday activity is well demonstrated; you do not need a bottle of an exotic substance to expose yourself to nanoparticles; just switch on your oven, enjoy a barbeque, or light some candles.

9.2.2. On the Road

A further source of environmental contamination from nanoparticles that has been extensively studied is pollution associated with road traffic. A study by Vinzents and co-workers carried out in central Copenhagen attempts to relate personal UFP exposure to oxidative damage of the subject's DNA [7]. Healthy non-smoking subjects of similar age were monitored over a period of 6 days cycling through busy traffic in the city, after testing the same subjects on an equivalent exercise regime in a clean environment. At relatively modest levels of exposure, a con-

nection was demonstrated between UFP concentration and oxidative DNA base damage.

It should also be noted that earlier studies by Dockery and by Pope have shown that particles that are small enough to penetrate the tissues of the lungs can have serious effects on the health of the subject [8,9]. These results complement the later studies and point to a significant risk to health arising from this source of nanoparticles.

The composition of products from diesel exhausts has been studied widely and it has been recognized for many years that organic chemicals form a significant proportion of this material. The presence of nanoparticles, especially metal and metal oxides, has become a matter of study more recently. As late as 1985, no mention was made of nanoparticles within diesel exhaust by the US Department of Labor Occupational Health & Safety Administration, although this is definitely no longer the case.

It should be noted in passing that the physiological effects of nanoparticles are not invariably bad. A study carried out by Schubert and co-workers on cerium and yttrium oxide nanoparticles claims that these materials are neuroprotective [10].

When considering such results it is worth remembering that it is not only the chemical and physiological properties of the nanoparticles that require consideration. It is possible for the particles to act as carriers for materials in the atmosphere that can be absorbed onto their surface. This of course applies to both indoor and outdoor environments.

Thinking about these results raises the question: which is more injurious to health—the manufactured Titanium Dioxide or Zinc Oxide nanoparticles present in a sun cream, or the nanoparticle cocktail present as you walk by a busy road?

9.2.3. Thinking Globally

Over the last few years it has become apparent that atmospheric nanoparticle pollution might play a significant role on the global climate. The case of “black carbon”, which is a dark soot produced from nanoparticle agglomeration, is a case in point. This pollutant arises from a variety of sources both man-made and natural. Its effects have been claimed to be dramatic, ranging from being a significant contributor to arctic warming to affecting the moisture supply to the Indian monsoon [11]. Note that it should not be confused with the industrial pigment “carbon black”.

Recent work by Bisiaux and co-workers has traced the sources of black carbon nanoparticles around and in Lake Tahoe between 2007 and 2009 [12]. The study concentrated on particles of between 60 nm to 400 nm in diameter, which themselves are comprised of chains of smaller particles (< 40 nm). Significant natural sources of this contamination were forest fires within the Tahoe basin, seasonal snowmelt and storms. Man-made sources also contributed, with water run-off from the surrounding towns also playing a significant role.

A further interesting result of this study was the observation that sudden increases in levels of the black carbon particles, arising from forest fires, decreased over a fairly rapid time interval. This suggests that the particles are rapidly agglomerating or degrading within the environment of the lake. The authors point out that this result has implications for the transport of such materials in a global aquatic environment.

This book cannot hope to deal thoroughly with this fascinating and important issue, but the interested reader can gain some insights from the references provided. In 2009, Nature commented that even if the environmental benefits turn out to be less than hoped, cleaner air would save hundreds of thousands of lives and would probably be easier to institute than reductions in carbon dioxide emissions [13]. Given the experience of the London smog of December 1952—with its overall death toll of possibly 25,000 persons—this is probably not an exaggeration [14]. It is a sobering thought that much of the smog may well have arisen from black carbon nanoparticles of less than 100 nm in diameter [15].

To illustrate the fact that the London smogs were not geographically isolated incidents, a similar "killer smog" was experienced in St. Louis on November 28th, 1939 as result of meteorological temperature inversion and the use of poor quality coal, domestically and industrially [16]. As was the case in London, the result of this incident was the introduction of regulations and a concerted public education campaign that significantly improved the air quality of the city. A famous series of photographs of this pollution were taken by staff photographers of the St Louis Post-Dispatch; these can now be viewed on-line [17].

It is interesting and important that both regulation and education were applied together to resolve these problems. This engagement of the general public is a lesson that is pertinent today and is discussed in the next section.

The take-home lesson of this introduction is that the combined effect of the nanoparticles from traffic, fires (natural and cooking),

smoking and just "life in general" dwarfs any conceivable exposure to nanoparticles from nanocoatings. Thus in terms of quantity there is no *a priori* concern about introducing a few more particles to the general background. Or, to put it another way, in terms of coatings, synthetic nanoparticles *per se* are not an issue. The concern shifts to whether low levels of the specific particles resulting from their use in coatings are a risk.

9.3. NANO IS HERE: PRODUCTS AND DEGRADATION

In 2006, Maynard and colleagues published an article stressing the need for strategic and targeted risk research into nanomaterial technology, taking the view that research being carried out at the time was possibly taking a lower priority, relative to product development and technology [18]. In what amounts to a call to arms, the authors proposed 5 areas of research to be carried out over a time-scale of 15 years to provide a firm basis on which future nanoparticle safety research could be carried out. These challenges, as the authors described them are:

1. Develop and validate methods to evaluate the toxicity of engineered nanomaterials within the next 5–15 years.
2. Develop robust systems for predicting the potential impact of engineered nanomaterials on the environment and human health within the next 10 years.
3. Develop robust systems for evaluating the health and environmental impact of engineered nanomaterials over their entire life, within the next 5 years.
4. Develop instruments to assess exposure to engineered nanomaterials in air and water within the next 3 to 10 years.
5. Develop strategic programs that enable relevant risk-focused research within the next 12 months.

It is now 6 years since this paper was published and a huge number of papers have been published during this time. The author's fear that health and safety issues were relatively disregarded appears to have been unfounded. However, whether this increase in information has met the challenges posed by the authors is open to question. Much information concerning nanosafety can appear contradictory and the methodology employed is by no means uniform.

Perhaps we should not be surprised. Science is about discovery, and

the interactions between nanomaterials and biological and environmental systems have been found to be more complex than might have been imagined at the end of the 20th century.

Are manufactured nanoparticles present to any significant extent in consumer goods? To help answer this question, the Woodrow Wilson International Center for Scholars and the Pew Charitable Trust established the Project on Emerging Nanotechnologies in 2005. The project involves the maintenance of an inventory of consumer goods containing nanomaterials which have been identified by their manufactures, or from other sources [19].

In March 2006, the number of products claimed to contain nanoparticles was just over 200 items; in 2011, this had increased to 1300. This is in fact a little lower than an earlier prediction that had assumed that there would be 1500 such products by 2011. The full list can be viewed on the website referenced above. The major contributor to this list is nanoparticle silver, which according to the authors appears in 24% of the listed items. Nanoparticle silver is mainly used for antimicrobial purposes; typical examples include athletic socks impregnated with silver nanoparticles.

Are there safety issues concerning nanoparticle silver? Given the claimed antimicrobial efficacy of the material it has been the subject of wide study and the matter has been reviewed by Brasuel and Wise. Socks that were claimed to contain nanoparticle silver were found to contain between 0 and 1360 μg silver per gram of sock. The leaching rate of silver from the socks was found to vary considerably and to depend on the manufacturing process. Most of the colloidal silver released in the washing process was converted into ionic silver and modelling suggests that treatment plants will be able to cope with the increased level of these contaminants, despite their potential antimicrobial effect on activated sludge. A Swiss study (using a 2008 high emission scenario) predicts the following silver concentrations: 4.4 μgM^{-3} in air, 0.08 μgM^{-3} in water and 0.1 μgM^{-3} in soil. American studies have predicted sewage effluent levels of 21 μgL^{-1} with 6 μgKg^{-1} nano silver in sludge-treated soil and 1.6 μgKg^{-1} in sediments. All of these concentrations are below the present accepted toxicological threshold for nano silver [20,21].

There is however a complication: recent studies have shown that nanoparticle silver possesses a different toxicological profile compared to micron-sized silver particles or silver ions. The toxicity appears to be associated with the nanoparticle size, shape, surface functionalization

and the kinetics of oxidation [22]. This has proved to be of sufficient concern for many interested parties to keep a close watch on the situation. In December 2011, the European Commission asked its Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) to provide a scientific opinion on nanosilver. This cautious approach has not met with universal approval; Hansen and Baun consider that the case for regulating nanosilver is sufficiently strong and should be implemented without the need for further reports [23].

It may seem a little surprising that the efficacy of nanoparticle silver as an antibacterial agent has been questioned. There is however evidence to suggest that the actual antibacterial mechanism involves silver ions released from the nanoparticles. A recent paper by Xiu and colleagues at Rice University investigated the antibacterial efficacy of spherical nanoparticles of well-defined size under both aerobic and anaerobic conditions [24]. In the latter case the oxidative pathway from Ag(0) to Ag^+ is prevented and it was found that the antibacterial activity of the silver nanoparticles under these conditions was significantly reduced. This study was carried out using *E. coli* as a target organism; the authors speculate that it would be interesting to extend the study to higher organisms such as algae and zebra fish, which have been frequently used in nanotoxicity studies.

Given that the "cleanliness hypothesis" makes it plausible that we are currently experiencing an environment that is too free of microbes, it might also be that "safe" nano antimicrobials (along with bleaches, soaps, etc.) are causing damage totally unrelated to their nano size. It is also pertinent to consider whether less than optimum doses of nanoparticles might induce the relevant organisms towards an increased resistance to these biocides. The team at Rice University present some evidence for this in their paper.

Carbon nanotubes (CNT) occupy a significant place in the Project list, because they are being used as stiffening agents in a number of sporting goods and as additives in electronic goods. Of all manufactured nanoparticles, these have seen the greatest and most extreme press. Claims have ranged from them being a revolutionary breakthrough in electronics to being potentially the new asbestos.

Except in the case of medical applications, public exposure to carbon nanotubes from bulk objects is potentially less of an issue than is the case for nanoparticle silver, or indeed titanium dioxide. It is unlikely that anyone will try to eat a tennis racquet. How CNT might be released into the environment from the waste stream is also a little hard to en-

visage, as the propensity for the materials to agglomerate is very great. The question of exposure to CNT during the product manufacturing process is a different matter entirely and deserves closer examination. It is during the nanotube manufacturing process, and early processing stages, that inhalation pathways with resulting physiological effects can be envisaged.

There have been many studies on the physiological behavior of CNT using isolated cell lines, simple cellular organisms and small mammals. The results can be contradictory and it is difficult to extrapolate the behavior to human biology [25,26,27]. The toxicological potential of these materials appears to be very much dependent on the specific particle shape and size, surface chemistry and environmental interaction—as is probably the case for all nanomaterials. This matter is discussed in more detail later in this chapter with special reference to the fiber paradigm.

It is important to put CNT into context. Organic carbon particulates are environmentally ubiquitous; analysis of ice cores taken from Western Greenland reveal the presence of carbonaceous material arising from the late Holocene period. Weintjes and co-workers make a distinction between what they describe as "Organic Carbon" (low and medium molecular weight materials emitted directly to the atmosphere or generated as secondary organic aerosols) and "Elemental Carbon". The latter is equivalent to the materials that comprise Black Carbon, arising from incomplete combustion of biomass or anthropogenic sources. Both forms of material were found in the ice cores [28].

The modelling of the behavior and effects of such materials is not a simple task; attempts have been made to separate the effect of Black Carbon from that of naturally occurring dust and to attempt to quantify the results of Black Carbon deposition on the ice sheet. A recent paper describing the use of such models has been presented by Goldenson and colleagues [29].

Rivalling CNTs in the health and safety publicity stakes is nanoparticle titanium dioxide. The chemistry of this material is discussed elsewhere in this book and will not be repeated here. The UV absorption properties of TiO_2 have ensured that the material has been used as an active ingredient in sunscreens along with zinc oxide.

A wide range of products containing TiO_2 are presently to be found on the Nanotechnology register [19]. Typical claims for these materials include the phrases such as natural and chemical-free (to distinguish them from organic chemical UV absorbers). One favorite is a piece of

marketing blurb which advertises a product as "a chemical free sunscreen which contains titanium dioxide that has been micronized into nanosized particles".

The situation concerning the safety of nanoparticle TiO_2 is complicated by the fact that only one form of the material (Anatase) is photocatalytically active and that this form is often stabilized with other materials to reduce this activity. A review of the situation regarding safety by The Nanodermatology Society in 2011 came to the conclusion that the passage of nano TiO_2 (and indeed ZnO) through healthy skin does not take place to any significant extent. However penetration might be increased if the skin were damaged or diseased in any way (for example, eczema) [30].

A further complication is the fact that different cell lines are used in many of the published toxicological studies on nano TiO_2. In these circumstances it is unsurprising that results differ and further reinforces the need for careful selection of the exposure route and the cells most likely to be affected [20].

It is also known that nanoparticles are capable of entering via open hair follicles which might allow entry to the dermis layer of the skin. Mitigating this effect is the upward pressure exerted by the growing hair and sebum and the fact that the follicles are lined by a tough thick coating that might make it difficult for any nanoparticle to exit the area and reach the surrounding dermis [31].

In the midst of contradictory results, two Australian organizations have taken a robust and pragmatic approach to the problem. The Cancer Council of Australia and the National Dermatology Society have issued statements reviewing the potential risks of nanoparticles within sunscreens against the benefits of reducing the risk of skin cancers, concluding that the benefits outweigh the risks [32]. In this respect the council is echoing the advice of the Australian Therapeutic Goods Administration, who state that "the current weight of evidence suggests that nanoparticle titanium dioxide and zinc oxide do not reach viable skin cells" [33].

A recent report by Zhang and co-workers concludes that although toxic effects can be observed in the testing of cell and animal models, the findings cannot be assumed to apply directly to human exposure [34].

Of course, as described earlier there are other means of bodily entry; particularly by breathing. This probably constitutes the most potentially damaging scenario for workers (as free particles are a possibility) rath-

er than consumers (where free particles are less likely). Studies have shown that a harmful inflammatory response in the lungs can occur and that exposed workers may not be receiving adequate protection. If this last point is correct then there could indeed be serious repercussions for manufacturers [35].

In terms of the ease or difficulty of detection of nanoparticles in the atmosphere, an extensive state of the science review was published by the US Environmental Protection Agency in 2010 [36]. The study concentrated on five types of nanoparticles which are of interest to the EPA: cerium oxide, titanium dioxide, nanostructured carbon (tubes & graphene), zero-valent iron and metallic silver.

Most sampling techniques used for the collection of nanoparticles derive from methods developed for the collection of ultra-fine particles. It is reported that although these methods can often be adapted, complicating factors can also arise that may lead to inaccurate results. It is also important to mention the difficulties experienced in determining the levels of cerium and titanium dioxide present because of relatively high levels of these materials present from natural sources.

To tackle the problem of exposure, monitoring the development of a personal nanoparticle sampler for the workplace has recently been reported. The device comprises a respirable cyclone and micro-orifice impactor. Filters are placed so as to result in uniform particle deposition and collection [37].

9.4. WORKING WITH NANOPARTICLES: WHAT SHOULD I DO?

Let us first of all be quite clear: in view of the present state of the evidence, it is best to err on the side of caution. When working with nanoparticles, personal protective equipment (PPE) should be worn—as it should for working with any chemical composition. Of course, if exposure is prevented by use of the correct protective equipment, then the associated risk is much reduced. The discussion on what constitutes the correct protective equipment is presented later in this chapter.

The place where one most expects to find nanoparticle contamination is, of course, in the manufacturing laboratory or factory. The incidence of nanoparticle release over a wide range of materials and industrial processes has been reviewed by Kuhlbusch and co-workers [38]. This extensive study, with many references, provides an excellent up-to-date review, illustrating the problems and difficulties involved in the mea-

suring methodology, while presenting interesting results. It deserves to be better known.

The results of the review indicate that, irrespective of what material was being considered, the major release of nanoparticles during the manufacturing process was associated with problems associated with poorly maintained equipment and inappropriate working practices. The release of agglomerated particles of > 300 nm was observed regularly; this has implications for the users of fumed silica and carbon black. The release of nanoparticles of < 100 nm was only observed in incidents associated with process malfunction.

The previously-mentioned review of the health hazards potentially present in the environment highlights the two beliefs that dominate present thinking about how nanoparticles might enter and interact with the body: the fiber paradigm and the ultrafine hypothesis [15].

The fiber paradigm is based very much on the known behavior of asbestos and the rather less known but equally deadly Erionite, a naturally occurring material that is prevalent in areas of Turkey and gives rise to a disease very similar to asbestosis. Human toxicity is related to the following four factors:

- A fiber length of 15 μm or greater, below which the fiber can be removed by the macrophages in the lung.
- A fiber diameter of less than 3 μm, allowing the fiber to be inhaled to the gas exchanging part of the lung.
- Insolubility or resistance to dissolution within the lung.
- A sufficient dose to the target organ.

It is believed that these effects are independent of chemical composition, other than as a determinant of solubility. The mechanism of damage generally appears to be frustrated phagocytosis, as the macrophage is injured while attempting to engulf the long fibers. This releases the cytokines, mitogens and oxidants that initiate the process of fibrosis and carcinogenesis.

How materials such as carbon nanotubes (CNT) and graphene fit the fiber paradigm is a matter for debate. The proportion of airborne CNT that satisfies the paradigm is unknown. The CNT may be originally long but will have a tendency to agglomerate and also curl into tangled balls. These particles may be larger than the 15 μm of the paradigm, but also possess very low density. In turn, this would produce a particle of low aerodynamic diameter which would allow the particle to reach and deposit itself in the deepest portion of the lung.

Despite the unknowns, it is important that this debate is taking place now; the epidemiological research into the asbestos-related mesothelioma took almost half a century to establish the conclusion that it was virtually impossible to prevent the disease unless the use of asbestos was prohibited. No one wishes for a repeat performance with nanoparticles as the culprit.

An attempt to provide proof of the danger of CNT of specific lengths, with respect to the fiber paradigm, has been made by Donaldson and colleagues at the University of Edinburgh [39]. In an extensive review of the evidence the authors claim that long fibers are retained in the tissues comprising the parietal pleura as a consequence of their inability to navigate the normal stomatal clearance system. This in turn leads to frustrated phagocytosis and the damage that ensues. Short or tangled fibers appear to be cleared through the stomata.

This is a paper that requires careful study by any manufacturer or user of CNT and similar shaped materials, especially since it provides a comprehensive review of the literature of mesothelioma arising from fibrous particles and frustrated phagocytosis.

The ultrafine hypothesis states that the toxic component of particulate air pollution resides in the nanometer size components. This is despite the fact that present epidemiological evidence is unconvincing. Nevertheless, it is suggested by Atkinson (quoted in the review paper by Seaton *et al.*) that "particle count, which reflects the sub-100 nm component, is the metric that best relates to the risk of heart attack" [15]. It is assumed that toxicity arises from the increased particle surface area, which arises from smaller particle size and the consequential increased total surface area that will be generated by a large number of small particles. The mechanism associated with the physiological effects might also be associated with the production of reactive chemical species, such as superoxide radicals, on the surface of the particles. Of course, larger surface area potentially means a greater number of reactive species.

The assumptions made are not unreasonable, but it is not an easy task to provide absolute proof. Real-life studies must also take into account the possible interactions of larger particles which are present in the contaminated air. What is known is that there is a link between exposures to combustion-derived particles and atherothrombosis. Exhaust fumes from which particles have been filtered do not behave in this manner. Such an observation is not fully informative to the debate until it becomes clear how much the specific atherothrombosis effects are due

to fine particles and how much to the chemistry of the specific "dirty" particles in diesel exhaust.

In 2008 Nel and co-workers published a study of the effect of ultrafine particles (aerodynamic diameter < 180 nm) on cardiovascular activity [40]. The results of this study were compared with those obtained for particles with an aerodynamic diameter of < 2.5 μm. The study was carried out in a mobile inhalation toxicology laboratory in downtown Los Angeles on apolipoprotein E-deficient mice. The mice exposed to the ultrafine particles exhibited significantly larger early atherosclerotic lesions than those exposed to the fine particles.

The composition of the particulate matter was analyzed and found to comprise the components shown in the table below. It is interesting to note that the ultrafine particles contain a greater weight percentage of polycyclic hydrocarbons and it is these species that Nel and colleagues believe to be implicated in the more damaging results obtained with the ultrafine particles. Note that the different categories contain overlaps, so totals exceed 100%, Table 9.1.

Although the results might be interpreted as providing support for the ultra-fine hypothesis, it is also important to consider the significant differences in composition of the particle size sets. Are the differences observed composition-driven rather than being purely related to particle size? How general is this kind of composition for particle size distributions of this kind? Does composition and distribution vary significantly with sampling location?

The authors hope that the reader will now begin to understand some of the many factors that inform the debate on nanoparticle safety. The next part of this chapter will deal with what might be described as wear and tear—do products containing nanoparticles degrade over the product lifetime, releasing the nanoparticles back into the environment?

TABLE 9.1. Composition of Particulates in Downtown Los Angeles.

Fine Particles	Ultrafine Particles
Metals: 25%	Organic Carbon: 52%
Organic Carbon: 25%	Elemental Carbon: 10%
Nitrates: 23%	Metals: 17%
Sulfates: 19%	Nitrates: 9%
Elemental Carbon: 3%	Sulfates: 3%
Unknown: 5%	Unknown: 9%

9.5. PRODUCT WEAR AND TEAR: ASSOCIATED NANOPARTICLE RELEASE

For items that are coated with formulations containing nanoparticles, a key question is whether abrasion and wear will create nanoparticles in the users' environment. So far the results are encouraging, as testing by a variety of methods has so far not led to evidence for the production of free nanoparticle dust. Any of the initial nanomaterial that was released was found to be embedded within the matrix material. This is encouraging and not surprising, but backs up the other reasons (such as hardness) for ensuring that the nanoparticles are firmly integrated into the coating matrix.

A study with which the authors are familiar was carried out at the behest of NanoResins AG (now part of Evonik) into the potential release of silica nanoparticles from abraded coated articles [41]. The results from such testing are important, as this is the kind of experiment that is carried out on a routine basis in many industrial laboratories.

The study was carried out to determine whether dust generated from the drilling and grinding of composites comprising Nanopox (an epoxide-modified formulation containing silica nanoparticles) contained discrete silica nanoparticles. The results obtained from the collected dust revealed only agglomerated silica particles. Further work revealed that the silica nanoparticles within the dispersion were not contaminated with crystalline silica, which is agreed to be a serious health hazard.

A similar study was carried out by Vorbau and co-workers in 2009, who studied the release of particles from coatings containing nanoparticle zinc oxide [42]. The test samples underwent Taber abrasion testing, and particulates released during the test were collected and examined. The authors declare that further refinements to their test method need to be carried out, but that no particles of < 100 nm were detected in the dust released during the abrasion process.

A study carried out by Gohler and co-workers on sanded samples of polyurethane coatings revealed the presence of nanoparticles after sanding; however, the study could not detect a significant difference between those samples which originally contained a nanoadditive and those that did not [43].

The nanoclay-coated barrier from Chapter 6 has also been tested in a similar manner and again shows no sign of individual free nanoparticles.

Not unnaturally, it is considered that working with liquid dispersions

of nanoparticles should carry less risk than working with dry powders. However, to the authors' knowledge, limited studies have been carried out on the generation of aerosols containing nanoparticles during industrial coating processes. The associated risk might be even higher during spray coating processes. This last point is illustrated by a report published in 2006 by the German Federal Institute for Risk Assessment (BfR), who received several reports of health disorders (especially pulmonary edemas) arising from exposure to "nanosealing" sprays [44]. Subsequent analysis revealed that none of the products actually contained nanoparticles, although the effects were genuine. Nevertheless, it is possible that in liquid systems that actually contain nanoparticles, exposure might be increased if aerosols are generated during the handling process. This is an area of industrial health and safety that deserves further investigation.

9.6. PROTECTING YOURSELF

Nanoparticles may enter the body by inhalation, dermal exposure and ingestion. It therefore makes good sense to wear personal protective equipment (PPE) while working. Although this is plain common sense, the authors, in the past, have visited respectable research laboratories working with nanoparticles where the wearing of PPE appeared to be optional. In the author's opinion, PPE within a laboratory or manufacturing area is a necessity; it is even more so when the associated risks are relatively unknown.

To those who wisely accept the need for PPE, the question then arises as to what equipment will actually protect against nanoparticles.

9.6.1. Inhalation

For air velocities prevalent in the workplace, the inertia of suspended nanoparticles can be considered to be zero. They can therefore be considered to behave as a gas, remaining airborne and diffusing rapidly. The first line of defense against inhalation is filtration. As an aerosol interacts with a filter, the trajectories of the particles will deviate from the streamline regime, collide with the filter material and be collected. The science of nanoparticulate filtration is surprisingly complex [45]. It seems obvious that small particles would be filtered less well than larger ones, but the intuition is incorrect, as another effect enters within the nano domain. For particles of a diameter of < 100 nm, Brownian dif-

fusion becomes the dominant mechanism and as particle size decreases, the Brownian diffusion increases and filtration efficiency increases because the probability of a particle coming into contact with the filter element is increased. Once the particle contacts the filter surface then van der Waals forces will ensure that it remains stuck to the surface. It is also possible for the filter to act as an electret and this provides an alternative adhesion mechanism for the particles (electrostatic attraction). It should be noted that behavior and performance of electret filters can be rather different from filters that show no electrostatic activity. The balance of different effects throughout a mask means that there is a MPPS (Most Penetrating Particle Size) for each specific mask—i.e., both smaller and larger particles are blocked. This is why there is no correlation with a filter's nominal pore size. Different types of filters show different values for the MPPS, so it is important to check for which particle size(s) any given mask is optimized. It is also important to note that the lungs also have a MPPS value. It happens to be the case that the typical size of tobacco smoke particles is close to this MPPS for the "filters" in the lungs, so smoke particles are especially effective in getting deep into the lungs. If the MPPS of the filter is different from the MPPS of the lungs, then both mechanisms reduce the amount of particles reaching deep into the lungs.

A problem with this comforting scenario is the possibility that very small particles (< 2 nm) may not behave in this manner. These small particles might be able to undergo thermal rebound away from the surface of the filter [46]. Despite this caveat it is accepted that for diffusion-driven processes, the filter efficiency will be high.

A wide comparative study of different full face respirators (FFR) is unfortunately not a simple task. There are two major standards that are recognized internationally: the National Institute for Occupational Safety and Health (NIOSH) based in the USA and the European CE (Norms Conformité Européen). Matters are complicated further by the choice of different test protocols by the two organizations. Despite the difficulties, a study has recently been carried out by Rengasamy and co-workers using specifically designed test methods on differently rated FFR equipment with NIOSH and CE accreditation [47]. The equipment was sourced from two different manufacturers. The testing was carried out with monodisperse silver and sodium chloride nanoparticles.

The results indicated that all the filters met the criteria demanded by the accreditation agencies, despite the specified test regime not being identical to those specified by the accreditation agencies. Very small

particles (<10nm) showed very little penetration through the FFR filters that were tested. The filtration characteristics were similar for the filters bearing NIOSH accreditation and those with CE accreditation. The degree of particle penetration through the filters differed slightly between the products obtained from different manufacturers.

The MMPS values for the FFR equipment was found to lie in the particle size range 30–60 nm. Once again, slight differences in performance were observed between FFR equipment obtained from the two different manufacturers. To investigate whether the filters were behaving as electrets, the filter media were treated with isopropanol to remove residual charge. On retesting, it was found that MMPS range was significantly shifted to between 200–300 nm. The results strongly indicate that the equipment tested shows the characteristic behavior of electret filters.

It is an interesting question whether this result suggests that filter performance will degrade more rapidly in different environments. Will extended use in atmospheres that contain small quantities of polar species (alcohols for instance) alter the performance characteristics of the FFR? The authors do not know the answer to this question.

It should be noted that the above discussions assume a tight seal around the filtration apparatus. Given the airborne behavior of nanoparticles, it is quite possible that increased inhalation may take place around the edges of a poorly fitted filter mask or cartridge. This will markedly reduce the efficiency and the effectiveness of the PPE.

The above observations are drawn together in a recent NanoSafe dissemination report concerning PPE and represent a consensus across different jurisdictions such as the US and Europe [48]. In other words, wearing the right mask during (as mentioned above) the most hazardous steps of production of dry nanoparticles offers good levels of protection.

9.6.2. Dermal Exposure

It is an unfortunate fact that this kind of exposure scenario is likely to take place in the later stages of nanoparticle production, such as bagging and packaging. At this late stage of processing, it is relatively easy to overlook and under-invest in what might be considered to be a relatively mundane part of manufacturing. However, it is at this stage of the manufacturing process that contamination and exposure can potentially become a serious issue. This is certainly the conclusion that

one can draw from the studies of Kuhlbusch mentioned earlier [38]. It is not only airborne contamination that requires consideration but also the possibility of exposure to contamination that has settled and stuck to surfaces; this point is easy to overlook.

What kind of PPE should be used to combat this possibility of dermal contamination? Testing various clothing material against 40 nm and 80 nm graphite nanoparticles, NanoSafe reported that non-woven type materials such as Tyvek performed much better than standard cotton or polypropylene-based clothing. The relatively air-tight nature of Tyvek (a non-woven high density polyethylene textile) makes it the favored whole body protective garment.

The other important protective against dermal exposure is the glove. A series of trials carried out by NanoSafe using high concentrations of 20 nm to 100 nm graphite aerosols revealed no particle penetration. This suggests that against aerosols, gloves provide a good degree of protection. However, this is not the whole story; how well do gloves protect the wearer against dispersions of nanoparticles?

To understand the problem more fully requires a holistic view, rather than concentrating solely on the nanoparticle. Gloves can become degraded by chemical exposure and there are increasing numbers of solvent- or acrylate-based nanoparticle dispersions commercially available. Some formulations can be particularly aggressive in their behavior towards protective gloves.

A study of the effect of various radiation-curable acrylates that are commonly used to produce nanoparticle dispersions on nitrile rubber gloves of varying thickness revealed some disturbing results [49].Thin gloves (0.1 mm) were penetrated by HDDA (hexanediol diacrylate) after only 10 minutes' exposure. For medium thickness gloves (0.56 mm), penetration occurred after 2 hours. Although the report is not concerned with nanoparticle dispersions, these results indicate that the protection afforded by nitrile rubber gloves towards nanoparticle dispersions of (for example) titanium dioxide will be compromised much more readily if the dispersion medium is HDDA rather than water. Clearly, the chemical resistance of the glove must enter into consideration when choosing PPE and these results reinforce the need to take a holistic view, rather than considering the nanoparticles in isolation.

Concerning nanoparticles themselves, Dolez and colleagues studied the penetration of dispersions containing 15nm titanium dioxide through a variety of gloves [50]. The study revealed that commonly used materials such as nitrile rubber and latex can, under regular stretching/defor-

mation, create pores and cracks that are of a size that can allow ingress of the nanoparticle dispersion over long time periods. The findings reinforce the need for good laboratory practice: protective gloves should be discarded after use and new gloves used for the next round of work. It is not a good idea to use the same pair of gloves until they fall apart.

Clearly, there is scope for further work in this field, probably using the Hansen Solubility Parameter (HSP) approach discussed elsewhere in this book. Indeed, HSP are used extensively in the analysis of glove protection issues, with the ProtecPro glove safety web service in Canada being a recent example of the approach [51]. In the meantime, the advice is:

- Choose the glove material carefully with respect to the chemical/solvent containing the nanoparticle.
- If the chemical/solvent is known to be aggressive to the only-available glove material then wear two pairs (in case the first gets compromised unexpectedly) and change them if there is significant exposure.

But then such advice applies as much to the chemical as to the nanoparticle—the skin may well be more open to the chemical than the particle and thus the nanoparticle may be introduced into the sensitive areas of the skin more readily. Again, it is necessary to stress that the contaminating material must be viewed as a whole and not just as a collection of non-interacting individual components.

9.6.3. Ingestion

This is possibly the least studied means of nanoparticle entry into the body. The action of wiping a hand across the face can mechanically transfer contamination, and it is possible to imagine how nanoparticle silver might be ingested in this way from treated surfaces or clothing. Of course, the more covered the face; the less likely such transfer is to occur. A further simple means of reducing the risk of such transfers is to ensure that hands are regularly washed.

A study carried out at Cornell University reports that in *in vitro* tests, high doses of carboxylated polystyrene 50nm nanoparticles interfered with the uptake and transport of iron in the intestines of chickens [52]. The report simply suggests that the study might have uses for further toxicological studies; however, this has not prevented comments arising from the blogosphere suggesting that the results indicate a serious nanohazard [53].

9.7. WHAT DOES THE PUBLIC THINK?

This whole section once again illustrates the richness and diversity of the term nanotechnology, for the discussion now enters the realm of social science and ethics. Again, there is no intent to present a comprehensive account of the work that has been carried out in this field, but the following discussion will hopefully prove informative and useful.

How are nanoparticles and nanotechnology viewed by the general public? Is there the groundswell of opposition to this technology, as was and is the case with genetically modified foods? Efforts have been made to answer this question and this has brought into focus the issue of public engagement with the technology. This last point is important, for it was probably a lack of engagement that led to the problems that became associated with genetically modified food and crops.

A short, but very useful summary of the present situation concerning the "democratizing of nanotech" is provided by Toumey [54]. The correct approach to this problem must avoid the twin perils of politicizing the debate at the expense of the science and imposition of a science policy on an unwilling public, even in the case of supportive scientific data. Steering this narrow channel is not easy, even in an ideal world. A sudden firestorm can easily erupt, the embers of which are difficult to stamp out, even when the scientific basis of the incident is thoroughly and publicly discredited.

But who are the public? This useful question was posed by Wickson, Delgado & Kjølberg [55]. They claim that the public can be defined in three dominant ways: as laity, as consumers and as stakeholders. Within these definitions lies a different perception of what nanotechnology means and different arguments as to why engagement is important. For the acceptance of nanotechnology to be achieved, there must be an informed and democratically engaged citizenry. Further discussion along these lines is presented by Cormick, who reported that in an Australian survey as many as 35% of the public are not actually interested in science and technology, and that there is no best way to engage with the public other than to try and do so in as many ways as possible [56].

Do scientists suffer from the distrust that presently afflicts politicians and other professions? If this is the case, then the task of presenting the facts about nanotechnology becomes more difficult. The prevailing viewpoint that scientists should learn to communicate with non-scientists who would then listen and learn (the deficit model) suffered serious challenges in the UK over the outbreaks of mad cow disease, foot and

mouth disease and the response to sheep contaminated by Chernobyl fallout. The deficit model suggests that the public distrusts science out of ignorance, but according to Wynne and others, the public distrusts science because it has good reason to [57,58].

It is worrying that Wynne's suggestion may be uncomfortably close to the truth. There is undoubtedly a prevailing cynicism concerning the actions and motives of organizations and professions that had previously enjoyed at least a degree of trust. When this decay of trust is coupled to the world of instant communication, the difficulty in presenting a true and fair picture of any new technology becomes more difficult. The old joke holds true; using the Internet, it is possible to be simultaneously well informed and completely ignorant. How great a step is this mistrust to the terrorist activity reported earlier is a question that is now undergoing debate [4].

Contrary to this pessimistic view is a study by Satterfield and coworkers. This particular study found that judgements concerning nanotechnology were quite positive, with approximately a 3/1 ratio in favor of benefits outweighing risks [59]. However, a significant number of subjects interviewed were unsure of their standpoint, indicating the malleability of attitudes. An interesting finding of the study was that unfamiliarity with nanotechnology did not automatically equate to an aversion to the perceived risks associated with it.

The previous comments have taken a global view as to the anticipated response to nanotechnology; however, is it possible to take into account cultural and political differences that might exist between different sections of the public? An attempt to answer at least part of this question has been made by Pidgeon and co-workers, who have tried to make a distinction between the attitudes to risk in the UK and the USA [60]. Interestingly, the study did not discover any particular concerns about perceived medical risks, but rather privacy issues and control of personal information. The American and British participants' attitudes showed more similarities than differences, with a general optimism towards the technology being expressed. No mistrust of science, as discussed earlier in this chapter, was raised or discussed. Subtle differences were uncovered in the study; the American participants being more ready to believe that the benefits of nanotechnology would initially benefit the more prosperous members of society, but that "trickle down" effects would become apparent. The British members of the study were more at pains to stress the potential benefits in terms of community, or indeed internationally.

Summing up, it appears that despite the negative potential that is associated with risk inherent with all new technologies, the situation regarding nanotechnology is rather better than might be expected.

9.8. THE TEAM AS STAKEHOLDERS

The team is a very important group of stakeholders in this issue. Each member brings their own sets of preconceptions which can be influenced by actions at work and within their personal lives.

The authors' personal experience is relevant here. At one stage in a nanoparticle project the issue of nanoparticle safety came up at a high-level site safety review. The site was already comfortable with the handling of Aerosil nanosilicas because good practices for these materials have been in place worldwide for decades. But as a team, we had no good ability to give assurances about the new particles we were using. Around this time, the first influential UK scientific report on nanoparticles was published via the Royal Society and one clear recommendation stood out: particles in forms other than powders posed, on the face of it, no extra significant risk. That made it easy for the team to reach the decision that no nanoparticle powders (other than the current Aerosils) would be allowed on site. This decision was widely accepted at site level and the safety issues around the nanoparticle dispersions could then be resolved via the site's standard safety processes.

With more experience, we recognized that this approach could be further improved. As mentioned above, the general risks of well-handled nanoparticles are small—it is the little details (e.g., a failure to clean up properly) which might transform a comparatively safe nanoparticle dispersion to a dry unagglomerated powder on a bench, piece of equipment, or any adjacent surface.

As with so much of safety, the best approach is a methodical analysis typical of any Quality system. With the assistance of one provider of nanosafety advice (Assured Nano) it was possible to think through the risks in a methodical manner which could be externally audited.

Because there are no guarantees that today's safe nanoparticle won't turn out to be tomorrow's asbestos, the team needs to protect themselves and their co-workers both practically and legally. If all the right steps are in place, are being regularly audited and updated, then if the level of risk of a specific nanoparticle is raised, there is a very high chance that in hindsight the level of protection provided was adequate even for that higher level of risk.

In fact, the above advice is not at all specific to nanoparticles. Just about anything handled in the context of coatings has the potential to be reclassified as more hazardous than previously thought; therefore, the above process should be part of everyday life within the company.

9.9. NANOPARTICLES: A QUESTION OF DEFINITION

What is the definition of a nanoparticle? This is another question that is not as straightforward as it might seem; for instance, the scientist working with quantum dots (radius ~2.5 nm) might consider that colleagues in the surface coating industry are working with additives (radius ~25 nm) that are boulders. We must also consider the shape of a particle. It is not unusual for a material to have dimensions in the range of microns in the *x* direction, but dimensions of nanometers in the y direction. How are we to deal with this difference? Words are important, as without definition, emotion and inaccuracy can override sensible discussion. It is also the case that if a material can be defined, then it is possible to regulate it. The input into these discussions has not been small, with many "concerned parties" and "stakeholder groups" each contributing their own personal and often contradictory viewpoint.

In 2011, after lengthy discussion, the EU provided the following definition of a nanomaterial as:

> "a natural, incidental or manufactured material containing particles, in an unbound state or as an aggregate or as an agglomerate and where, for 50% or more of the particles in the number size distribution, one or more external dimensions is in the size range 1 nm–100 nm."

Notice that this definition takes into consideration the point that was raised in the previous paragraph; materials that are anisotropic (such as nanotubes and nanowires) are included within the definition. This lengthy document provides an overview of the health and safety issues that were thought to be worthy of discussion by the panel of scientists who were responsible for preparing the document [61].

It is probably unsurprising that this definition and the associated report have not been greeted with unanimous agreement. To take two examples from what might be considered to be opposed parties: the European Chemicals Industrial Council (CEFIC) criticized the document for using number particle distributions, rather than weight distributions, while Friends of the Earth Australia criticizes the use of the 50% limit-

ing definition [62,63]. There are many more instances of responses that find the EU definition unsatisfactory [64,65].

The definition provided by the EU was unlikely to meet with universal acceptance and approval. Given the diversity of the materials, science and technology involved, it is likely that no single definition can be described as appropriate. To the authors, the situation is analogous to attempting to regulate and define the term organic chemistry, without taking into account the subtleties and nuances present within the subject. So at present we seem to have the situation where everyone knows what a nanoparticle is, but cannot provide a concrete definition.

Nanomaterial regulation in the USA is less advanced than is the case for the EU. At present there is no single accepted definition for the term Nanomaterial. The situation is further complicated by the involvement of several different federal organizations; these include the Environmental Protection Agency (EPA), the Occupational Safety & Health Administration (OSHA) and the Food & Drug administration (FDA), to name a few.

A further level of complication is added by involvement of individual states: California, Massachusetts, Pennsylvania and South Carolina are attempting to introduce regulations for what are perceived to be nanomaterials. Historically, this is a regulatory route that has been travelled before in the USA. The use of phosphate additives in laundry detergents was regulated by individual states when no federal legislation was forthcoming. With enough states enacting legislation it became economically unviable for the detergent manufacturers to provide versions of their product containing the offending additives. It is possible that such situation might be repeated, claims John DiLoretto, CEO of NanoReg [66].

The situation in Asia appears even less well developed, although Japan's Ministry of Economy, Trade & Industry (METI) has established a Committee for the Safety Management of Nanomaterials. This follows from earlier meetings which culminated in a document entitled The Expert Meeting on Safety Measures for Nanomaterial Manufactures, etc. This committee was scheduled to hold its first meeting during December 2011, so presumably some definitive report will emerge [67].

9.10. IS IT NANO OR NOT?

From the industrial point of view, there are several materials that fall into the EU definition of nanoparticles and which are manufactured on

an extremely large scale. These materials, in their primary form, can be considered to comprise nanoparticles, but which in their delivered form comprise agglomerates. Classic examples are Carbon Black and Fumed Silica.

Carbon Black, used for industrial purposes (tires, ink, etc.), should not be confused with the environmental pollutant "black carbon" discussed earlier in this chapter. It is claimed that Carbon Black comprises a much greater proportion of elemental carbon than is found in the environmental pollutants, despite both materials being produced by an incomplete oxidation process [68].The primary particle size for Carbon Black depends on the manufacturing method employed, but can range from about 10 nm to 300 nm. These primary particles agglomerate into larger particles which comprise the bulk of the material used in manufacturing processes; the size of these larger particles is of the order of 100 nm to 1 μm. The primary particles appear to agglomerate into structures that have been described as resembling "bunches of grapes".

There have been significant safety studies into the industrial use of Carbon Black in its agglomerated form. Whether similar results would be obtained if identical studies were carried out on the nanoparticulate form of the material is unknown. It is, however, a fair comment that the occurrence of the non-agglomerated form of the material is unlikely outside the manufacturing reaction vessel. One rather unusual safety concern arises from the report by Høgsberg *et al.* in the B*ritish Journal of Dermatology,* who report that black inks used for tattooing comprise mainly particles with dimensions of < 100 nm [69]. Although most studies carried out on Carbon Black have not produced significant sources for concern, a series of experiments carried out by Koike and Kobayashi in 2006 did reveal that exposure to Carbon Black nanoparticles could produce deleterious effects on bone marrow cells [70].

The situation concerning fumed silica is rather similar to that of Carbon Black. There is the possibility of confusion with a more dangerous (and natural) physical form of the material, in this case crystalline silica. It has been known for many years that this material can create serious health issues [71]. Fumed silica is amorphous in nature and is not implicated in the cause of silicosis. Aerosil is the product name for fumed silica produced by Evonik and it is widely used in processed foods, cosmetics and paints. Decades of use in a wide range of industries has not produced any significant health and safety issues concerning this material [72].

Fumed silica is produced by a flame pyrolysis process; the primary

particle size ranges from about 5 nm to 50 nm. These primary particles rapidly agglomerate into larger secondary and tertiary structures which can have dimensions ranging from 1 μm to > 10 μm. As discussed in Chapter 3, reducing agglomerates of, say, 200 nm down below even 100 nm is a difficult challenge, so accidental production of < 100 nm particles seems unlikely.

As masks optimized for handling fumed silica have been around for decades, the correct PPE for handling the dry powder is not difficult to find.

9.11. CONCLUSION

Hopefully the team will now understand a little more fully that the question "Are nanoparticles safe?" is neither straightforward, nor easy to answer. There are, however, a few simple rules that should be followed by any organization that thinks there is a need to use these materials.

- Keep abreast of the literature on the subject: this may take up development time but it is worth the expenditure.
- Try to become involved in the debate on nanoparticle safety via trade organizations and learned societies.
- Do not claim to produce nanoproducts when you don't; it will only bring you pain.
- If you intend to use nanomaterials in your process, ensure that all the parts of your organization that are concerned with the process are kept informed, not just the lab team.
- It follows from the previous point that if you are using nanomaterials in your processes, ensure that they are labelled as such so that coworkers are clear about their status. It takes no time to produce such a safety label and it enables you to monitor what materials you have on site. This of course presupposes that your raw materials are being adequately managed from a health and safety viewpoint. If they are not, then rectify the situation immediately.
- It is useful to have a dedicated waste stream for mixes and formulations that contain nanoparticles. It reveals that you are taking a responsible attitude to these matters and it provides a degree of protection if there are problems with your waste management procedures.
- Attempt to become an accredited company. Organizations such as Assured Nano provide a valuable service in assessing practices and

procedures that are being used in the application of nanomaterials within industry. They also provide regular auditing of the procedures that have been put into place [73].

Going back to the opening remark, nano merely refers to the size, not to the properties. There is nothing intrinsically unsafe or unnatural about nanoparticles. Just as with most of what we do, there are plenty of unknowns. When dealing with unknowns, a proportionate process is required. With the information in this chapter, the team will be able to handle those unknowns as just another part of the day job.

9.12. REFERENCES

1. The dose makes the poison. (2011) *Nature Nanotechnology Editorial* 6:329.
2. *No magic in Magic Nano.* www.nanotechbuzz.com. Accessed July 2012.
3. David Ake. (2011) Biosecurity expert fears a nanoparticle attack. *National Defense Magazine,* August 2011. www.nationaldefensemagazine.org. Accessed July 2012.
4. Leigh Phillips. (2012) Armed resistance. *Nature* 488:576–579.
5. Golam Sarwar, Richard Corsi, *et al.*(2002) Hydroxyl radicals in indoor environments. *Atmospheric Environment* 36:3973–3988.
6. Lance Wallace and Wayne Ott. (2011) Personal exposure to ultrafine particles. *Journal of Exposure Science & Environmental Epidemiology* 21:20–30.
7. Peter S. Vinzents, Peter Møller, *et al.* (2005) Personal exposure to ultrafine particles and oxidative damage. *Environmental Health Perspectives* 113:1485–1490.
8. Douglas W. Dockery. (2001) Epidemiological evidence of cardiovascular effects of particulate air pollution. *Environmental Health Perspectives* 109:483–486.
9. C. Arden Pope, Richard T. Burnett, *et al.* (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* 287:1132–1141.
10. David Schubert, Richard Dargusch, *et al.* (2006) Cerium & yttrium nanoparticles are neuroprotective. *Biochemical & Biophysical Research Communications* 342:86–91.
11. Jeff Tolson. (2009) Atmospheric science: Climate's smoky spectre. *Nature* 460:29–32.
12. Marion M. Bisaux, Ross Edwards, *et al.* (2009) Stormwater and fire as sources of black carbon nanoparticles to Lake Tahoe. *Environmental Science & Technology* 45:2065–2071.
13. Time for early action. (2009) *Nature editorial* 460:12.
14. Michelle L. Bell, Devra L. Davis, *et al.* (2004) A retrospective analysis of mortality from the London Smog Episode of 1952: The role of influenza and pollution. *Environmental Health Perspectives* 112:6–8.
15. Anthony Seaton, Lang Tran, *et al.* (2010) Nanoparticles, human health hazard and regulation. *Journal of the Royal Society: Interface* 7:119–129.
16. Raymond T. Tucker. (1941) Smoke prevention in St Louis. *Industrial Engineering Chemistry* 33:836–839.
17. Tim O'Neil. *A Look Back: 1939 Black Tuesday Spurred Crackdown on Coal Pollution.* http://www.stltoday.com/news/local/article_00c3b6cd-ba69-5a19-b498-fbc29f9630c4.html?mode=image. Accessed August 2012.
18. Andrew D. Maynard, Robert J. Aitken. (2006) Safe handling of nanotechnology. *Nature* 444:267–269.
19. www.Nanotechproject.org/inventories/consumer. Accessed August 2012.

20. Murphy Brasuel, Kelsey Wise. (2011) The current state of engineered nanomaterials in consumer goods and waste streams: The need to develop nanoproperty-quantifiable sensors for monitoring engineered nanomaterials. *Nanotechnology, Science & Applications* 4:473–486.
21. T.M. Benn and P. Westerhoff. (2008) Nanoparticle silver released into water from commercially available sock fabrics. *Environmental Science & Technology* 42:4133–4139.
22. P.V. AshaRani, Grace Low Kah Mun, *et al.* (2009) Cytotoxicity and genotoxicity of silver nanoparticles in human cells. *ACS Nano* 3:279–290.
23. Steffan Foss Hansen and Anders Baun. (2012) When enough is enough. *Nature Nanotechnology* 7:409–411.
24. Zong-ming Xiu, Quin-bo Yong. (2012) Negligible particle-specific antibacterial activity of silver nanoparticles. *American Chemical Society Nano Letters* 12:4271–4275.
25. L.E. Murr and K.F. Soto. (2005) Combustion generated nanoparticulates. *International Journal of Environmental Research & Public Health* 2:31–42.
26. Elijah J. Petersen, Jarkko Akkanen, *et al.* (2009) Biological uptake and depuration of carbon nanotubes by Daphnia magna. *Environmental Science & Technology* 43:2969–2975.
27. Barbara J. Panessa-Warren, Matthew M. Maye, *et al.* (2009) Single walled carbon nanotube reactivity and cytotoxicity following extended aqueous exposure. *Environmental Pollution* 157:1140–1151.
28. I.G.M. Wientjes, R.S.W. Van De Wal, *et al.* (2012) Carbonaceous particles reveal that Late Holocene dust causes the dark region in the western ablation zone of the Greenland ice sheet. *Journal of Glaciology* 58:787–794.
29. N. Goldenson, S.J. Doherty, *et al.* (2012) Arctic climate response to forcing from light-absorbing particles in snow and sea ice in CESM. *Atmospheric Chemistry and Physics Discussions* 12:5341–5388.
30. Adnan Nasir. (2010) Nanotechnology and dermatology: Part 2—risks of nanotechnology. *Clinics in Dermatology* 28:581–588.
31. J. Lademann, H.-J. Weigmann. *et al.* (1999) *Skin Pharmacology & Applied Skin Physiology* 12:247–256.
32. http://www.cancer.org.au/cancersmartlifestyle/SunSmart/nanoparticles_sunscreen.htm. Accessed August 2012.
33. www.tga.health.gov.au/alerts/sunscreens.htm.—site undergoing maintenance.
34. Ruinan Zhang, Yuhong Bai, *et al.* (2012) The potential health risk of titania nanoparticles. *Journal of Hazardous Materials* 211–212:404–413.
35. Brian J. Majestic, Garnet B. Edarkos, *et al.* (2010) A review of selected engineered nanoparticles in the atmosphere: Sources, transformationsand techniques for sampling and analysis. *International Journal of Occupational and Environmental Health* 16:488–507.
36. *EPA Nanomaterial Case Studies: Nanoscale Titanium Dioxide.* Washington DC, 2009. Accessed August 2012.
37. Chuen-Jinn Tsai, Chun-Nan Liu, *et al.* (2010) Novel active nanoparticle sampler for the exposure assessment of nanoparticles in workplaces. *Environmental Science & Technology* 46:4546–4552.
38. Thomas A.J. Kuhlbusch, Christof Asbach, *et al.* (2011) Nanoparticle exposure at nanotechnology workplaces: A review. *Particle & Fiber Toxicology* 8:22.
39. Ken Donaldson, Fiona A. Murphy, *et al.* (2010) Asbestos, carbon nanotubes and the pleural mesothelium: A review of the hypothesis regarding the role of long fiber retention in the parietal pleura, inflammation and mesothelioma. *Particle & Fiber Toxicology* 7:5.
40. Andre E. Nel, Jesus A. Araujo, *et al.* (2008) Ambient particulate pollutants in the ultrafine range promote early atherosclerosis and systemic oxidative stress. *Circulation Research* 102:589–596.
41. *NanoResins. Toxicological aspects of nanocomposites.* January 2006. http://www.nanoresins.ag/images/download/nanotox_engl.pdf. Accessed August 2012.
42. Manuel Vorbau, Lars Hillemann, *et al.* (2009) Method for the characterization of the abrasion-induced nanoparticle release into air from surface coatings. *Aerosol Science* 40:209–217.

43. Daniel Göhler, Michael Stintz, *et al.* (2010) Characterization of Nanoparticle Release From Surface Coatings by the Simulation of a Sanding Process. *The Annals of Occupational Hygiene* 54:615–624.
44. Bundesinsitut für Risikobwertung. (2006) *Exercise caution when using "nanosealing sprays" containing a propellant.* Press release 08/2006, 31st March 2006. http://www.bfr.bund.de/en/press_information/2006/08/exercise_caution_when_using__nano_sealing_sprays__containing_a_propellant_-7699.html. Accessed August 2012.
45. R.C. Brown. (1993) *Air Filtration: An Integrated Approach to the Theory and Application of Fibrous Filters.* Oxford, UK: Pergamon.
46. K.W. Lee and R. Mukund. (2001) Filter collection. *In: Aerosol Measurement: Principles, Techniques, and Applications,* pp. 197–228. P.A. Baron & K. Willeke, *et al.* (eds.). New York: John Wiley & Sons.
47. Samy Rengasamy, Benjamin C. Eimer, *et al.* (2009) Comparison of nanoparticle filtration performance of NIOSH-approved and CE-marked particulate filtering facepiece respirators. *Annals of Occupational Hygiene* 53:117–128.
48. http://www.nanosafe.org/home/liblocal/docs/Dissemination%20report/DR1_s.pdf. Accessed August 2012.
49. Rob Zwanenberg. Adequate Protective Gloves for Working with UV/EB-Curing Acrylates. *RadTech Europe* www.radtech-europe.com/knowledge-centre/articles/consumer-electronics. Accessed August 2012.
50. Patricia Dolez, Ludwig Winches, *et al.* (2011) Development of a test method for protective gloves against nanoparticles in conditions simulating occupational use. *Journal of Physics: Conference Series* 304:012066.
51. http://protecpo.inrs.fr/ProtecPo/jsp/Accueil.jsp?institut=IRSST. Accessed September 2012.
52. Michael L. Shuler, Gretchen J. Mahler, *et al.* (2012) Oral exposure to polystyrene nanoparticles affects iron absorption. *Nature Nanotechnology* 7:264–271.
53. http://theextinctionprotocol.wordpress.com/2012/02/17/new-study-finds-ingested-nanoparticles-more-dangerous-than-previously-thought/. Accessed August 2012.
54. Chris Toumey. (2011) Democratizing nanotech, then and now. *Nature Nanotechnology* 6:605–606.
55. Fern Wickson, Ana Delgado, *et al.* (2010) Who or what is "the public"? *Nature Nanotechnology* 5:757–758.
56. Craig Cormick. (2012) The complexity of public engagement. *Nature Nanotechnology* 7: 77–78.
57. Brian Wynne. (1992) Public understanding of scientific research: new horizons or hall of mirrors? *Public Understanding of Science* 1: 37–43.
58. Chris Toumey. (2006) Science and democracy. *Nature nanotechnology* 1: 6–7.
59. Terre Satterfield, Milind Kandlikar, *et al.* (2009) Anticipating the perceived risks of nanotechnologies. *Nature Nanotechnology* 4:752–758.
60. Nick Pidgeon, Barbara Herr Harthorn, *et al.* (2009) Deliberating the risks of nanotechnologies for energy and health applications in the United States and United Kingdom. *Nature Nanotechnology* 4:95–98.
61. Scientific Committee on Emerging and Newly Identified Health Risks SCENHIR. *Scientific Basis for the Definition of the Term Nanomaterial* July 6, 2010. http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_030.pdf. Accessed September 2012.
62. CEFIC-The European Chemical Industry Council. *Nanomaterials Safety and Legislation.* Final Position Paper June 2012. http://www.cefic.org/Documents/PolicyCentre/PositionPapers/Position%20paper%20-%20Nanomaterials%20safety%20and%20legislation%20-%20June%202012.pdf. Accessed August 2012.
63. Cam Walker. (2012) New nano research reveals massive regulatory failure. *Friends of the Earth Australia* February 10th 2012. http://www.foe.org.au/articles/2012-02-10/new-nano-research-reveals-massive-regulatory-failure. Accessed August 2012.
64. The Centre for Environmental Law. (2012) *Nanomaterials "Just Out of Reach" of European*

Regulations. 6th February 2012. http://www.ciel.org/Chem/JustOutofREACH_Feb2012.html. Accessed August 2012.

65. Nanowerk. *More Criticism Of EU's Nanomaterial Definition From ANEC.* 20th October 2011. http://www.nanowerk.com/news/newsid=23117.php. Accessed august 2012.
66. John DiLoreto. (2010) We should have seen it coming: States: Regulating nanotechnology. *Nanotechnology Now,* July 2010. http://www.nanotech-now.com/columns/?article=484. Accessed July 2012.
67. Ministry of Economy Trade & Industry. (2011) *Establishing the Committee on Safety Management for Nanomaterials,* November 2011. http://www.meti.go.jp/english/press/2011/1130_02.html. Accessed August 2012.
68. Cabot. *Specialty Carbon Blacks* http://www.cabot-corp.com/Specialty-Carbon-Blacks. Accessed August 2012.
69. T. Høgsberg, K. Loeschner, *et al.* (2011) *British Journal of Dermatology* 165:1210–1218.
70. Eiko E Koike, Takahiro Kobayashi. (2006) Chemical and biological oxidative effects of carbon black nanoparticles. *Chemosphere* 65: 946–951.
71. George Rosen. (1943) *The History of Miners' Diseases: A Medical and Social Interpretation.* New York: Schumans.
72. Evonik. *Responsible Handling of Nanotechnology at Evonik.* http://nano.evonik.com/sites/dc/Downloadcenter/Evonik/Microsite/Nanotechnology/en/Nano%20Guideline_e.pdf. Accessed August 2012.
73. *Assured Nano*™. http://www.assurednano.com/. Accessed August 2012.

Index

About the Authors

DR STEVEN ABBOTT has worked in the coatings industry for many years and is now an independent consultant working on a wide range of coating and nanoformulation technologies whilst writing technical software and apps. He is a Visiting Professor at the School of Mechanical Engineering at the University of Leeds, and is based in Ipswich, UK.

DR NIGEL HOLMES has extensive experience of the paints and coatings industry and is Senior Project Scientist at MacDermid Autotype where his responsibilities include the development of high-performance nanocoatings. He works extensively with the NanoCentral network to help ensure that the development chain from particle to end-product is effective and mutually beneficial. He is based in Wantage, UK.