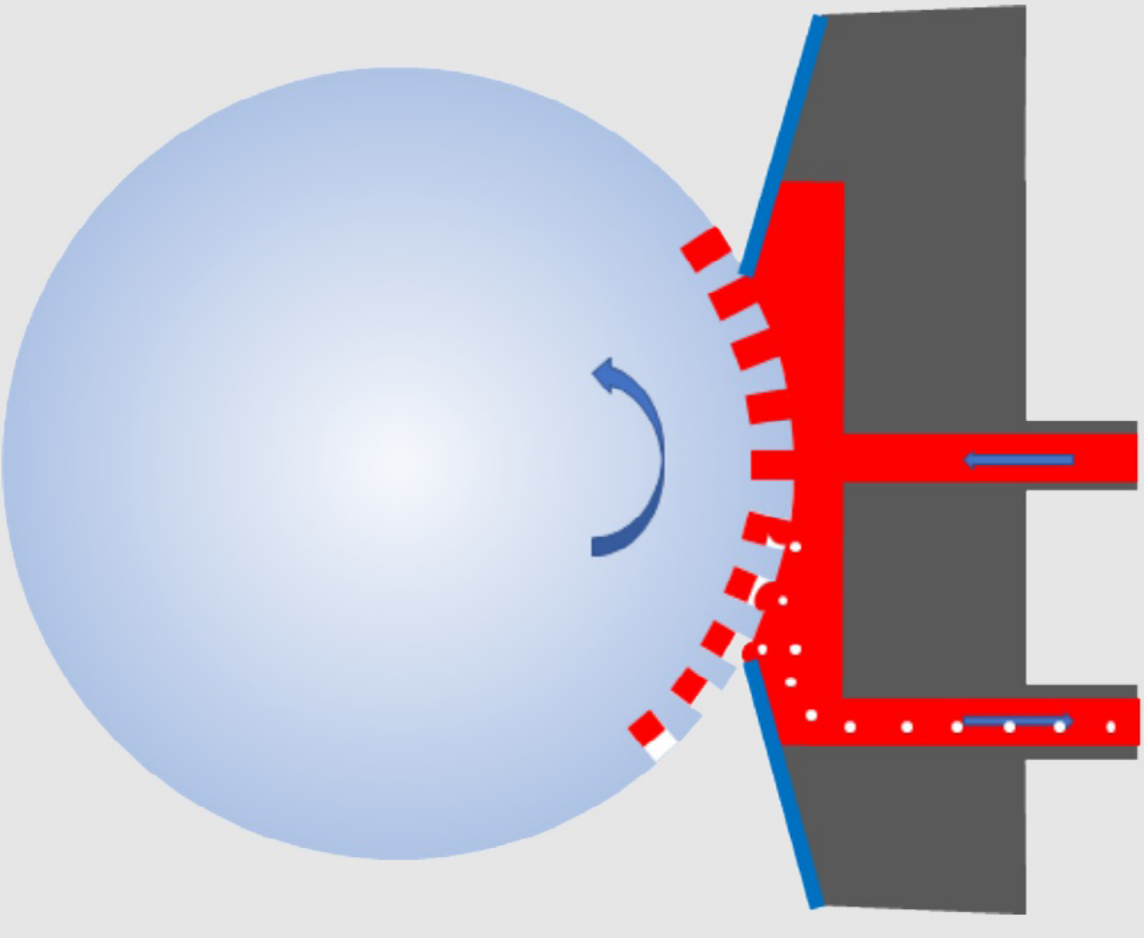
Printing Science

Principles & Practice



Prof **Steven Abbott**

Printing Science: Principles and Practice

Prof Steven Abbott

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

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

Version history:

First words written	21 Sep 2017
Version 1.0.0	26 May 2018
Version 1.0.1	23 Sep 2023

Contents

Preface

Modern printing produces vast quantities of printed materials on impressively automated machines at astonishing speeds. It is a routine and mature industry where the vast majority of things are printed with very little apparent trouble. So why write a book on printing science when most of those in the industry have no need of anything other than knowing which buttons to press?

There are two answers to that question. The first has general agreement among the industry experts with whom I've discussed it: don't do it; printing is so mature that a book is not necessary and will go unread.

The second answer to that question is that ever since I became involved in the printing industry I have been fighting the notion that printing is an art rather than a science. So even if no one reads the book, I will be happier for at least trying to put the record straight on the matter.

The resistance from printers to the scientific approach always struck me as strange because fighting the laws of physics (which they were) is always hard work to no benefit. Screen printing, for example, is remarkably straightforward and reproducible when done scientifically because basically there is nothing to go wrong. Done the way most screen printers do it is, conversely, a continual fight because they are trying to fix one problem they have caused via adjustments that are guaranteed to cause other types of problems - with no awareness that all the hard work and frustration is caused by themselves and not by screen printing itself.

I was rather slow to realise that there were positive reasons why Joe Printer wanted the technique to be an art rather than a science: job security. As long as everyone believed that only Joe had the magic to make the process work, Joe's job was safe; until, that is, the entire company went bankrupt because Joe's inefficiencies and waste made them unprofitable.

The good news for those who want to do things properly is that none of the scientific ideas is especially hard (at least at the level of broad understanding). And in most cases, if there is a formula that provides insights, I provide an app that brings the formula to life to make it clearer as to what it all means. I am not very good at understanding the meaning of formulae just by looking at them, so I have always had to write programs, spreadsheets and apps for my own understanding. With modern app infrastructure being so powerful, even on a phone or tablet, as well as being safe to run on corporate networks, I am happy to write the apps and share them with everyone; they are free, free of adverts and entirely open source. As this is an eBook, each app comes with a link that you click to take you straight to the appropriate app page.

There is a little irony that a book on printing science should be published only as an eBook. But the fact that the book can be downloaded freely, and can be updated regularly to fix problems or add new ideas, far outweighs the slight negative contribution to the world of ink-on-paper printing.

Although the various printing techniques have many differences, they share many common bits of science. So we start with the few common key principles shared across printing before looking at the specifics of each of the main techniques.

It is my pleasure to thank numerous people who helped me in many different ways. Tine Hennissen from Lawter taught me much about the modern realities of the water balance test. Prof Yanlin Song and Prof Junfei Tian each provided hard-to-access papers on their deeply insightful work on offset film splitting. John Bodnar at TAGA kindly assisted me with access to the TAGA Proceedings which are such a fine source of scientific data used extensively throughout the book. Ian Hargreaves provided a most helpful update on typical modern offset printers that allowed me to clarify some key ideas. Victor Ramon Soto Del Toro kindly provided an excellent critique of a draft version of the Offset chapter. Prof Edgar Dörsam and Dr Hans Martin Sauer at the IDD in TU Darmstadt were generous with their time and ideas and helped me avoid some egregious errors. I have an historical debt to the work on screen printing of Prof Tim Claypole's team at U Swansea and the Welsh Centre for Printing and Coating (WCPC). Papers from the team's work on other printing techniques are mentioned throughout the book, as are papers from the TAGA Journal, skilfully edited by Prof Claypole and the team. Prof Glen McHale and Dr Rodrigo Ledesma-Aguilar at Northumbria U provided important theory and support for drop wetting and dewetting phenomena.

In general, the images are my own, from the apps or from Excel. Where I have taken images from Wikipedia, these have been acknowledged.

Steven Abbott
Ipswich, 2018

Abbreviations and Definitions

AM/FM/XM	Amplitude, Frequency, miXed Modulation
FFT	Fast Fourier Transform
ICC	International Color Consortium
IPA	Isopropyl alcohol
LPC/LPI	Lines per Centimetre, Lines per Inch
PE	Polyethylene
PEI	Polyethyleneimine
PET	Polyethyleneterephthalate
PP	Polypropylene
PSA	Pressure Sensitive Adhesive(s)
PU	Polyurethane
Stress	Force divided by area, in N/m^2 or Pa or MPa
Strain	% increase in length caused by a stress
TAGA	Technical Association of the Graphic Arts
TTE/TTS	Time Temperature Equivalence or Superposition
VOC	Volatile Organic Compound
WCPC	Welsh Centre for Printing and Coating
WLF	Williams, Landell, Ferry time-temperature equivalence

1 The Core Ideas

A conventional book on printing will start by telling the reader how important printing is, will describe the different printing techniques and give some generalities about inks. My assumption is that readers don't need such a general background and have at least some knowledge of inkjet, flexo, offset, gravure and screen printing, plus basics like CMYK colour. I am concerned with the *science* of printing so I want to start with the core ideas that impact *all* printing processes. That means that when we examine each process, we already have the ideas in place and don't have to repeat them for each process. Electrophotography does not feature strongly in this chapter because many of the basic mechanisms are far removed from the other techniques.

Most of the core ideas are to do with the ink: how it flows (or not) within the printing process, how it splits from the initial bit of ink to the final bit on the substrate, and how it behaves on the substrate once printed - for example, will the dot spread, will a pinhole form? In the next chapter come ideas of how the ink dries/cures/absorbs. The chapter after that addresses the longer-term issue of how the ink sticks to the substrate.

By comparison there are relatively few core ideas connected to the rest of the process. One example is the fact, surprising to many, that rubber is effectively incompressible, so it cannot do many of the jobs people expect rubber to do. When it comes to the colour of inks we can certainly come up with a few core ideas; these are kept for a later chapter. Other specific properties of ink such as conductivity will be dealt with in an ad hoc manner if they fit in with more generally important themes.

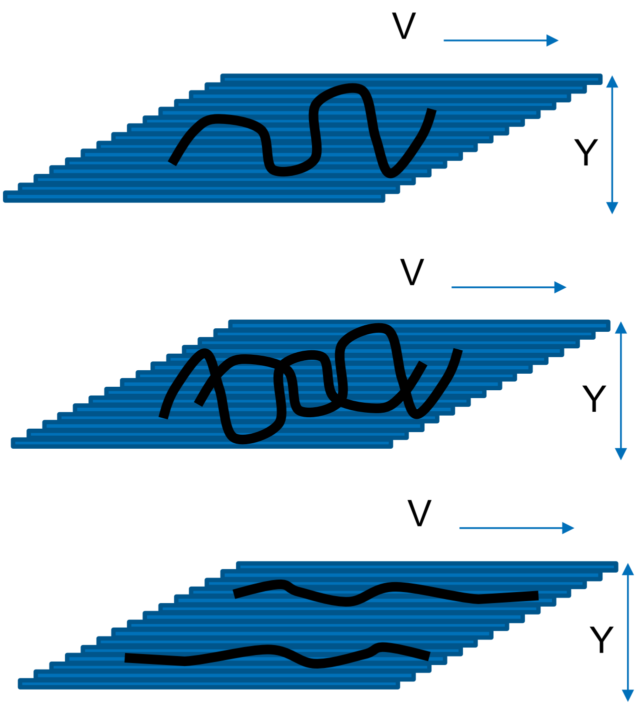
Let us start with some real basics.

1.1 Viscosity

We are all familiar with the basics of viscosity. Water, at 1cP, flows easily, olive oil at 100cP slows more slowly, honey at 5000cP definitely is slow to flow and ketchup at 50,000cP is a real problem, requiring us to bash the bottle to make it come out. Printing inks span that range: inkjet inks are in the 10cP range, flexo and gravure are in the 10-300 cP range, screen inks in the 10,000 cP range and offset inks are in the 10,000-100,000cP range.

I have used the rather strange unit, cP, simply because it is the most common and the most natural. The units of viscosity should be (for reasons we will see shortly) Pa.s, and because 1000cP=1Pa.s I could have said that screen inks are in the 10Pa.s range. I could also have said that water is 1mPa.s. Confusingly, (and I regularly get confused), viscosities can be reported in Poise, hence the use of cP which are centiPoise. As you can work out, 10P=1Pa.s, which is why 1cP=1mPa.s.

In reality, the numbers for ketchup, screen and offset inks are meaningless because the values depend strongly on how you measure them, i.e. on the shear rate used for the measurement. We therefore need to understand shear. To do so, I will start with where viscosity comes from.



In the diagram we have the fluid flowing in streamlines. The fluid near the bottom is stationary, and near the top it is flowing with velocity V, for example the flow between a stationary and moving roller. A polymer happens to be in the solution and one part of it is in the fast-moving stream, while other parts are rather slow moving. The portion of polymer in the fast-moving stream is being slowed down by the portion in the slow-moving stream so energy is dissipated by the fight between the bits of polymer in the different streamlines. This dissipation of energy means that we must provide more force to achieve the same relative flow, i.e. we have a higher viscosity. If we add more polymer, so that the chains get tangled, there is even more dissipation of energy across the streamlines and the viscosity is even higher.

If there is a large V or the difference in velocity is created over a smaller gap, Y, then the differential forces on the polymer chains are so large that the chains will end up being stretched out along the streamlines. So now the different parts of each polymer chain experience similar velocities and there is little dissipation, so the viscosity has decreased.

Shear is defined as V/Y, so higher shear means lower viscosity if you have things like polymers or particles that can span the streamlines. For fluids containing molecules that can only interact over short distances, then viscosity is independent of shear because nothing changes when shear is increased.

Newton showed that the force, F, (per unit area, i.e. the "stress") needed to shear a liquid with shear S is given by $F=\eta.S$ where η is a constant, the viscosity. So although, for simple fluids, the viscosity stays constant, the force needed to shear the liquid increases if the velocity increases and/or the gap decreases. Such fluids are Newtonian fluids and for inkjet, and many flexo and gravure inks, the concentration of pigments and polymers are so low that there is little chance for interesting alignment effects to affect viscosity therefore they are essentially Newtonian. Screen and offset inks are much more concentrated and viscous, and their viscosities decrease by a few orders of magnitude over the range of shear found in their printing processes.

The reduction in viscosity with shear is crucial to practical printing of high-viscosity inks. Pumping them or getting them to flow through nips would require much more energy if they retained their low-shear viscosity. Because we can estimate the shear in most parts of the process, we need to know what the viscosity will be in those parts. Let's have a go at some estimation.

When pumping some fluid at velocity V through a tube of diameter d, shear is V/d. At 10cm/s through a 1cm diameter pipe the shear is 10/1 = 10/s. The units are m/s divided by m, i.e. 1/s. Because the relevant force is force per unit area, measured in N/m^2, i.e. in Pa, his means that viscosity which is Force/Shear is Pa/(1/s) = Pa.s.

An impeller blade in a mixer going at 10m/s, acting across 1cm, the shear is 10/0.01 = 1000/s.

In a three-roll mill turning at 0.5m/s with a gap of 50μm, the shear is 0.5/5e-5, which is 10,000/s.

As these simple examples show, inks are subjected to a very large range of shears in quite basic processes.

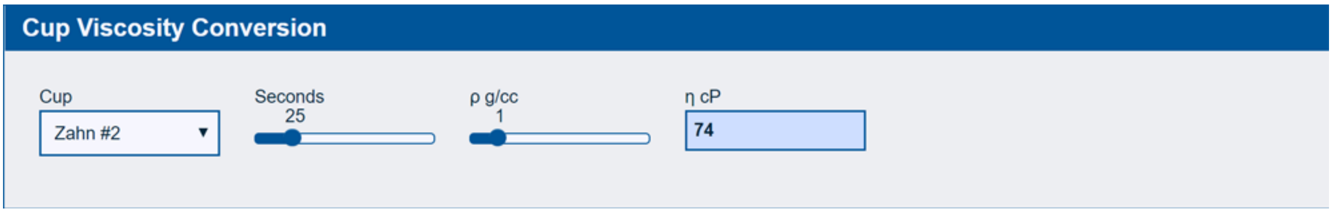
1.1.1 Measuring viscosity badly

Because viscosity is so important to all aspects of printing it is unfortunate that the industry relies on inadequate methods for measuring it.

Very popular is a Ford or Zahn cup[1], where a known volume of ink flows through the hole in the bottom of the cup and the time to empty the cup is measured. One ink might be 25s from a Zahn #2 cup while another might be 19.5s from a Zahn #4 cup. Although one can make a rough conversion between these values and real viscosities, it is really unacceptable that many people in an organisation have no idea what the real viscosities are and are unaware of how inaccurate and uninformative these numbers are. When I wrote the original draft of this paragraph I had looked up the viscosity values as being 50 and 250cP respectively. When I got round to writing a conversion app, they came out as 74 and 356 cP respectively; even for the (nominally) same cups, there is disagreement about the viscosity conversion factors.

---

[1] Image from the Wikipedia article


Cup Viscosity Conversion

Cup: Zahn #2

Seconds: 25

ρ g/cc: 1

η cP: 74

App 1-1 https://www.stevenabbott.co.uk/practical-coatings/Cup-Viscosity.php

Incidentally, the app received a severe rebuke from a highly-respected rheologist from the ink world - saying that it would encourage people to continue using these cups.

The cups are bad in many ways. If, for example, one ink is shear thinning and the other is Newtonian, they might give exactly the same measured time, but because the shear through the cup hole changes from large when the cup is full to small when the cup is near empty, there could be a very large effective difference between the two inks. For one gravure printing issue in which I became involved, the difference in viscosity between good and bad quality was less than the operator variation in making the measurements, though in this specific example, the shift that did regular measurements got better results than the shift that decided not to bother to do the measurements.

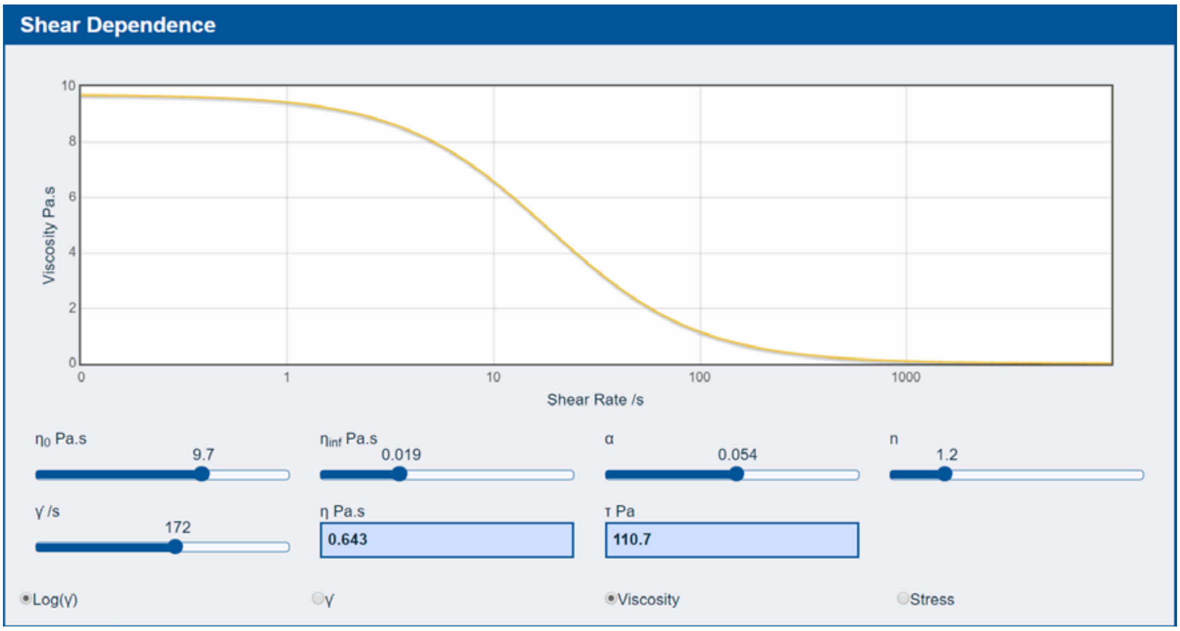
Slightly more sophisticated, but equally unacceptable if used simplistically are the spindle viscometers. The force needed to spin a disc or cylinder of known diameter at a known speed within the liquid is measured and the viscosity can be calculated directly in cP. For a given ink, the calculations are meaningful only within a modest range of rotation speeds and spindle diameters, otherwise the machine is in an out-of-control range where forces are too large or small and/or unhelpful flow patterns are created around the spindle. Set up properly, these viscometers can give a reliable single value for the viscosity within a meaningful control range. Perhaps the ink is specified to be runnable between 15 and 20cP. If the given spindle and speed range are adequate for 5 to 50cP then the measurements are going to be fine. But if the arrangement is good only for 1 to 20cP then a batch that is 22cP might be measured incorrectly.

The real problem with these spindle viscometers is that the speed/diameter restrictions discourage us from measuring each ink over a range of shear rates. So we might have two inks, both of which are "within spec" as measured at the single combination of speed and diameter point, yet perform very differently on press because one is strongly shear thinning, something which would have been apparent if the inks had been measured at a very different combination of speed and diameter.

1.1.2 Measuring viscosity properly

Printing would be far less full of surprises if we all, routinely, measured the viscosity of our inks properly.

The first necessity is to measure the shear dependence. In a rheometer this is easy. A rheometer (such as a cone and plate or simply a double plate device) can rotate or oscillate in just about any desired pattern to generate the required information. The simplest mode spins the top plate (for convenience we will use "plate" even if it is in fact a cone) at a given speed, measures the force required and, from knowledge of the precision gap, provides viscosity and shear. Then the speed is increased, another measurement taken, and so on from the lowest to the highest shear rates, say 1 to 1000/s. From such measurements we have a curve shown in the shear dependence app:



η_0 Pa.s 9.7

η_{inf} Pa.s 0.019

α 0.054

n 1.2

$\dot{\gamma}$ /s 172

η Pa.s 0.643

τ Pa 110.7

● Log($\dot{\gamma}$) ○ $\dot{\gamma}$ ● Viscosity ○ Stress

App 1-2 https://www.stevenabbott.co.uk/practical-rheology/Shear-Viscosity.php

Playing with the app gets you used to the approach I adopt for myself when trying to understand what is going on. I slide sliders and click options to see what happens. In this app we use the Cross model for viscosity which is simply a fitting function. The two viscosities, η_o and η_{inf} are the low- and high-shear limits, while α and n affect the shape of the curve between the two limits. It is conventional to show the plot as a function of the log of the shear and it looks as though stuff happens across a broad range. If you select the non-log plot (try it now!) it is quite a shock - everything happens very quickly in the low-shear range and after that nothing much happens. Most of us like to see the plot of viscosity versus shear. The machine measures stress and rheologists usually plot stress versus shear. Again, try the Stress option so you can get used to seeing how others choose to look at the data:

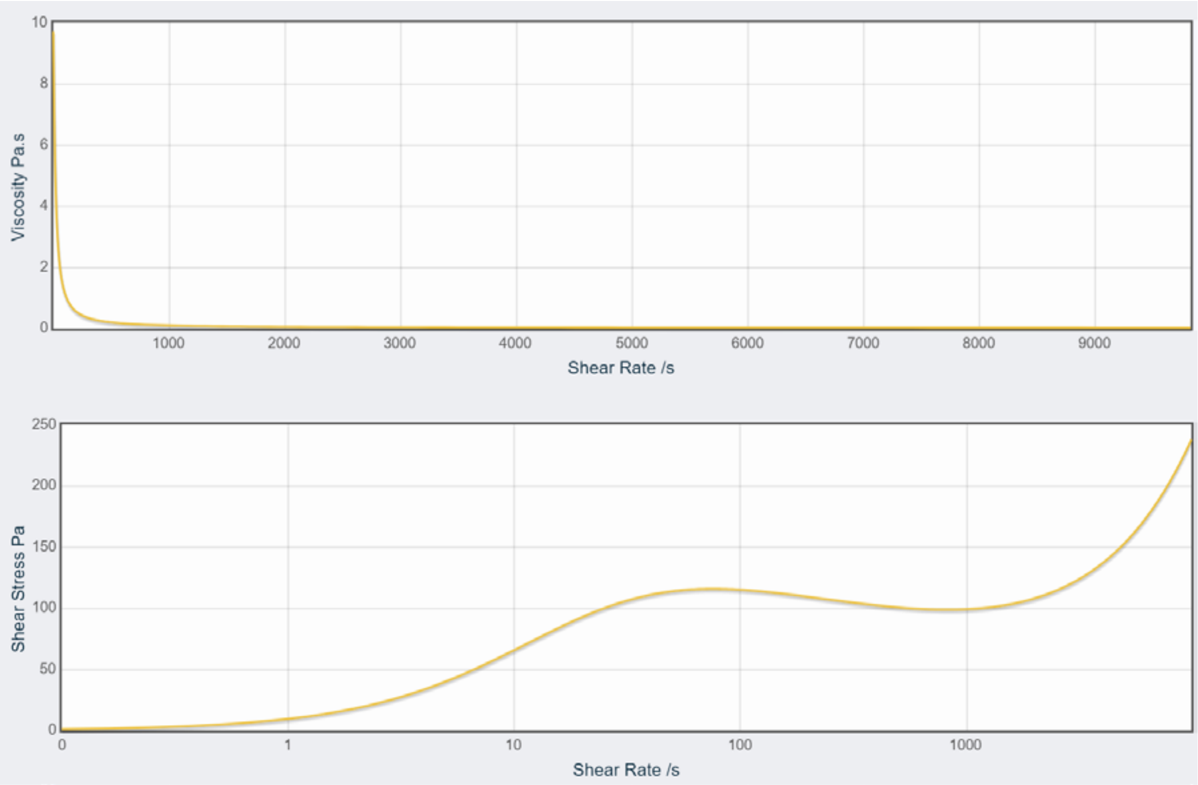


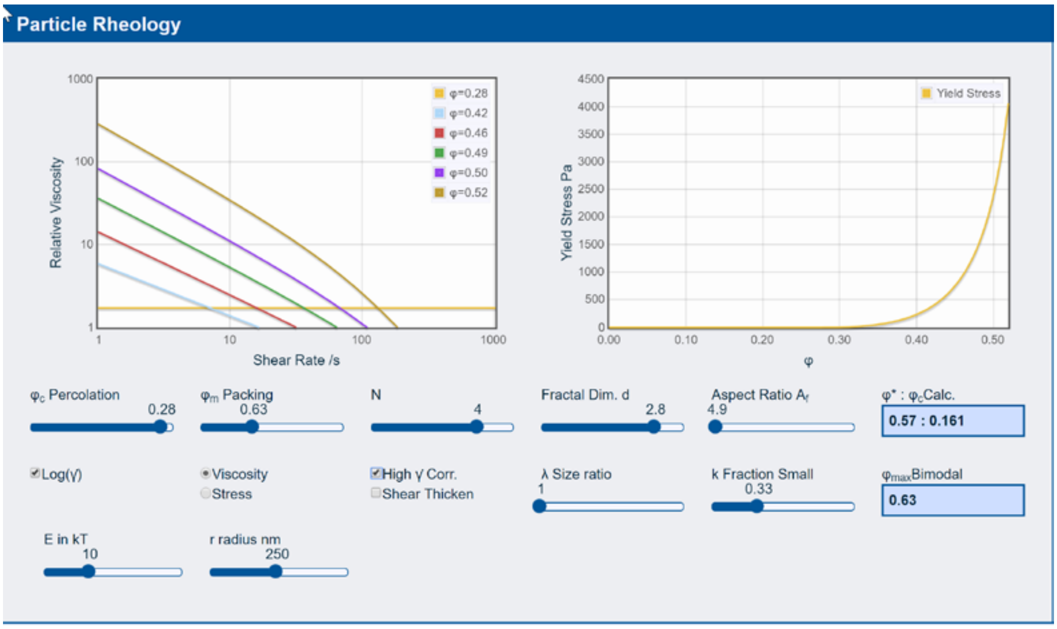
Figure 1-1 The same data plotted non-log (top) and in terms of stress rather than viscosity (bottom)

The rheometer, fortunately, can tell us much more about the ink. A perfect liquid always flows, no matter how gently you push it. Some ink formulations behave as elastic solids if you push them with a stress less than their "yield point". An ink with a high yield stress will not spontaneously flow, a serious inconvenience in some situations. For an offset ink (or, rather, paste) a significant yield stress is useful to avoid excess flow within the ink supply duct. A rheometer can measure the yield point by, for example, gently raising the stress on the ink until the plate starts to rotate, or by an oscillatory test (discussed below) which looks for the transition from elastic to viscous behaviour.

Another usually unhelpful aspect of an ink is thixotropy. The viscosity of many inks, whether shear-thinning or not, are a constant when measured at a given shear. A thixotropic ink has extra structure within the ink so that the viscosity at the start of the measurement is high and decreases over time. This dependence of viscosity on shear history means that different parts of the ink (e.g. old ink mixed with fresh ink) behave differently, so the system is not fully in control. Again, the rheometer can measure this behaviour by keeping the plate rotating at a constant speed and watching the decrease in viscosity over time, and/or doing a jump to high shear rate then looking at the hysteresis as the shear rate is returned to its previous value.

Given that most inks contain particles, it is especially important to understand how the particles' shape, concentration, size distribution and interactions affect the rheological behaviour. Generally speaking, the more particles we can put

into the ink, the thinner the printed dot can be and the more problems we avoid such as dot gain (discussed in terms of Stefan's law) and dot-on-dot physical effects (a specific example is discussed in offset). Even for offset inks where high viscosities are the norm, getting the most particle for the least particle-induced viscosity is a good idea, and for flexo and gravure it's a necessity.
A rather powerful app encapsulates the key issues involved with particulate systems at reasonably high loadings.



φ_c Percolation 0.28
φ_m Packing 0.63
N 4
Fractal Dim. d 2.8
Aspect Ratio A_f 4.9
φ^* : φ_cCalc. 0.57 : 0.161
☑ Log(γ̇)
● Viscosity
○ Stress
☑ High γ̇ Corr.
☐ Shear Thicken
λ Size ratio 1
k Fraction Small 0.33
φ_{max}Bimodal 0.63
E in kT 10
r radius nm 250

App 1-3 www.stevenabbott.co.uk/practical-rheology/High-Shear-Particles.php

This has every possible option included and you will have to read the content on the page to get the full picture. In essence, we have a non-spherical particle (aspect ratio A_f=5) which should start showing "interesting" effects at the percolation volume fraction φ_c=0.158. Viscosity should go to near-infinite at the packing limit not 0.63 (normal for spheres) but 0.54 - i.e. the packing effects kick in very quickly so at low shear the viscosity at φ=0.5 would be increased by a factor of >1000. The reason for the 0.54 is that we have some agglomeration of the particles (an average, N, of 4 per cluster) with a fractal dimension, d, of 2.6 partly because we have a non-spherical particle and partly because clusters automatically have a lower fractal dimension. Fortunately, we have a bimodal distribution which raises the spherical packing limit to 0.67 so at 0.5 we "only" have an increase of 450.

However, these particle systems shear thin rather dramatically because they move as a single block of particles, with no chance of wasteful dissipation of energy. So you see the viscosity decreasing (linearly in this log-log plot) according to a power law. The linearity breaks down at high shear and we lose even more viscosity. That's because we've added the High γ̇ Correction option. This accounts for two effects. First,at high shear, the limiting packing fraction increases significantly - for spheres from 0.63 to 0.7, because they can give close packing instead of random packing. Second, at high shear, the fractal clusters break apart.

Clearly no single app can get all those aspects right; the results are more "illustrative" than "precise". Yet the app captures a lot of complex features that intimidate most of us. I only gained the courage to write the app after reading an especially clear paper by Campbell et al[2]. This encouraged me to read other papers such as one from Bicerano et al[3] to build up the fullest possible picture within the constraints of a viable app.

In terms of take-home messages, there are no surprises: use spherical particles with a >10:1 bimodal distribution with 33% small particles, and have a great dispersant so there is no tendency to create a flocculated fractal structure.

The app has an option that shows what happens if you don't add the right level of the right dispersant. At high shear rates the formulation can shear thicken for complex reasons discussed in the app.

The app also gives an idea of the yield stress, as discussed above.

Finally, the rheometer can give us profound insight into that trickiest of properties: tack. While most low-viscosity inks do not show tack, many high-viscosity inks show an extra "stickiness" beyond the high viscosity. Picking up some ink on a spatula or stretching the ink between two (gloved) fingers allows the inexperienced to detect tack and for the experienced to have a good idea whether the level of tack is correct. Tack can be measured in a "tackmeter" though everyone agrees that the measurement leaves a lot to be desired as it is rather indirect. The rheometer allows us to disentangle pure viscous effects from the cohesive effects that constitute tack. To understand how, we have to learn about the infamous G' (pronounced G-prime) which is the elastic (cohesive) component, G'' (G-double-prime) (the ordinary viscous component) and tan δ ("tan delta", the ratio of the two, G''/G'). Fortunately, using an app to explain them means that they are not as scary as most people assume them to be.

1.1.3 G', G'' and tan δ

We still have our cone and plate, or two plates, but this time we oscillate rather than rotate. Imagine that the sample is a pure elastic solid. As one plate twists, the solid instantly stretches, and the force of that stretch is detected on the plate beneath the sample. The bigger the twist, the bigger the detected force. As the plate twists back, the measured force decreases, then as the plate starts to twist

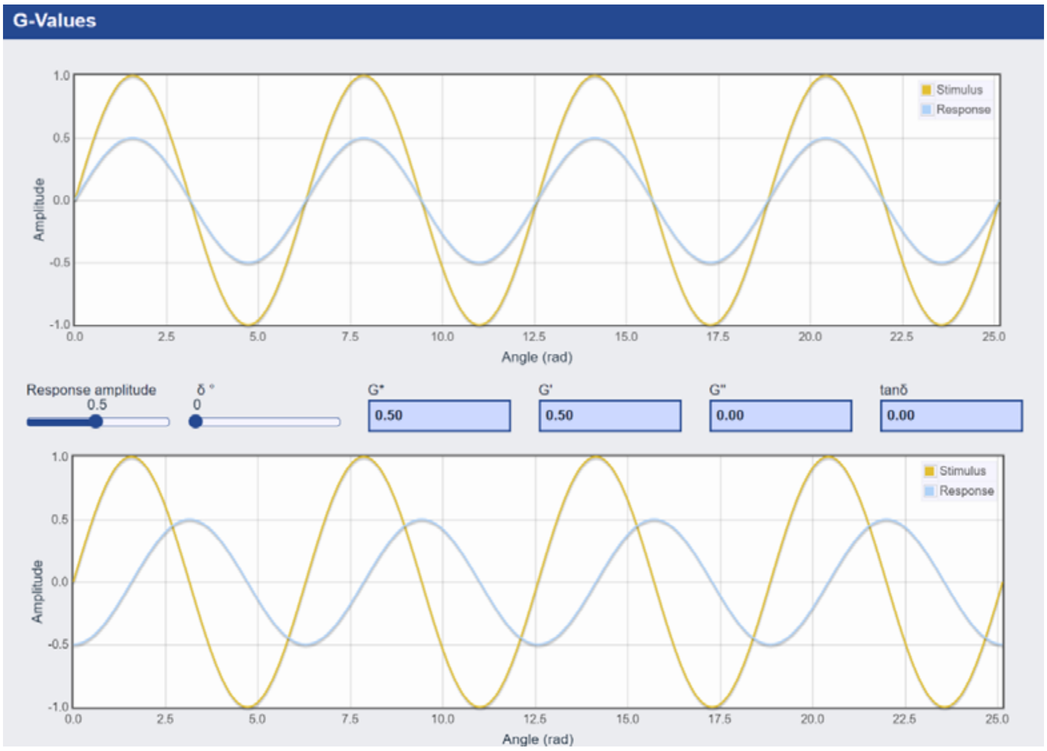
---

[2] Gregory A. Campbell, Michael E. Zak, Mark D. Wetzel, *Newtonian, power law, and infinite shear flow characteristics of concentrated slurries using percolation theory concepts*, Rheol Acta(2018) 57:197–216

[3] Jozef Bicerano, Jack F.Douglas, Douglas A. Brune, *Model for the Viscosity of Particle Dispersions*, J. Macromolecular Science, Part C, 39, 561-642, 1999

in the opposite direction, the measured force becomes negative. Everything is in perfect sync.

Now repeat the experiment with a perfect liquid. If the oscillation is at the maximum or minimum extent, the relative velocity is small, so the shear is small so the force is minimal. If the oscillation is passing from positive to negative, the velocity of the plate is maximum, the shear is maximum so the response is maximum. In other words, the stimulus and response are exactly out of phase[4].


G-Values

Response amplitude 0.5

δ ° 0

G* 0.50

G' 0.50

G" 0.00

tanδ 0.00

App 1-4 https://www.stevenabbott.co.uk/practical-adhesion/g-values.php

The top graph shows the stimulus (yellow) and response perfectly in phase, so the top sample is perfectly elastic, such as a thin layer of rubber. The lower graph shows the response 90° out of phase, so the sample is perfectly viscous, like water or glycerol.

In reality, fluids show some mix of the two behaviours. For example, if the measured response is 45° behind the stimulus then there is exactly equal amounts of elastic and viscous behaviour. The elastic response is designated G' (G is the symbol for modulus when measured in shear mode) and is often called the storage modulus because energy is stored (like stretching an elastic band) and released (letting go of the band) during the cycle. The viscous response is designated G" and is often called the loss modulus because the energy is lost as viscous heat. The tangent of the angle between the stimulus and response is called tan δ and is defined as G"/G'. The tangent of 45° is 1, which means

---

[4] I had always explained the effect differently. One day a physicist colleague, Dr Dilwyn Jones, happened to sit in on one of my classes where I "explained" the out of phase behaviour. During the coffee break, he kindly took me to one side, told me I was completely wrong and gave me the above correct explanation.

that G''=G' so G''/G'=1. A purely elastic material has an angle of 0°, and G''=0 so G''/G'=0. A purely viscous material has an angle of 90° and tan δ is infinite because G''/0=infinite.

The values of G' and G'' depend strongly on the frequency with which the device oscillates, so you often have curves of G' and G'' dependency on measurement frequency, with values decreasing at lower frequencies. The values also depend strongly on temperature, with both decreasing with increased temperature. The decreases with frequency and with temperature are actually the same thing because in rheology there is "time temperature equivalence" (TTE), meaning that a values measured at a certain frequency at a (say) high temperature can be found at a low temperature with a lower frequency. The relationship between them, TTE or TTS (Time Temperature Superposition) is described by the WLF (Williams, Landell, Ferry) equation discussed, for those who are especially keen, at https://www.stevenabbott.co.uk/practical-adhesion/wlf.php. WLF removes one classic excuse from print researchers. It is said that they cannot measure tack values at speeds/frequencies relevant to the press. With a bit of WLF theory, they can do the measurement at a suitably low temperature to get closer to the on-press situation. In the science of pressure sensitive adhesives (note that the WLF app is on the Practical Adhesion site) the use of low temperatures to investigate high-speed phenomena is routine. Sadly, this is not the case in printing.

Who cares? Printers (should) care. Tack arises when the polymers in the ink are so tangled that when the ink is stretched they act somewhat like an elastic band, so there is a large G' element in a tacky ink. By knowing G', G'' and tan δ values for different batches of ink we know much more about them than a feeling that "this one is tackier than that one". Importantly, the tackiness depends strongly on both the speed and the temperature at which the measurements are taken. As we have seen, thanks to TTE, at very low speeds or at high temperatures everything is viscous, so G'' and tan δ are large. At very high speeds or at low temperatures, everything is elastic so G'' and tan δ are small. From the rheometer we can directly see how the tackiness will change if the press speed and/or temperature change. This is a very powerful way to understand your ink.

There is one more trick about interpreting G' and G'' data. Because G'' is just viscosity, we should be able to read viscosity from a G'' plot. G'' is in units of Pa and viscosity is in units of Pa.s so we need to multiply by time or, rather divide by the frequency of the oscillation. If G'' at 100/s is 5000 Pa then the viscosity is 50 Pa.s. Purists will (rightly) point out that this is a misuse of the Cox-Merz "rule", but it points in approximately the right sort of direction and is valid when tan δ >> 1.

1.2 The unimportance of surface tension/energy

The majority of those in the printing industry would regard the title of this section as either an unfortunate misprint or the result of ignorance and delusion on my part.

Happily, I can prove to you that except for some very specific cases, surface tension and surface energy are largely unimportant for printing. It is astounding that so much time and energy has gone into studying these non-important factors, distracting everyone from the important factors.

Let us start our demolition with one of the fundamental values in printing and coating science, the capillary number, Ca which is given by:

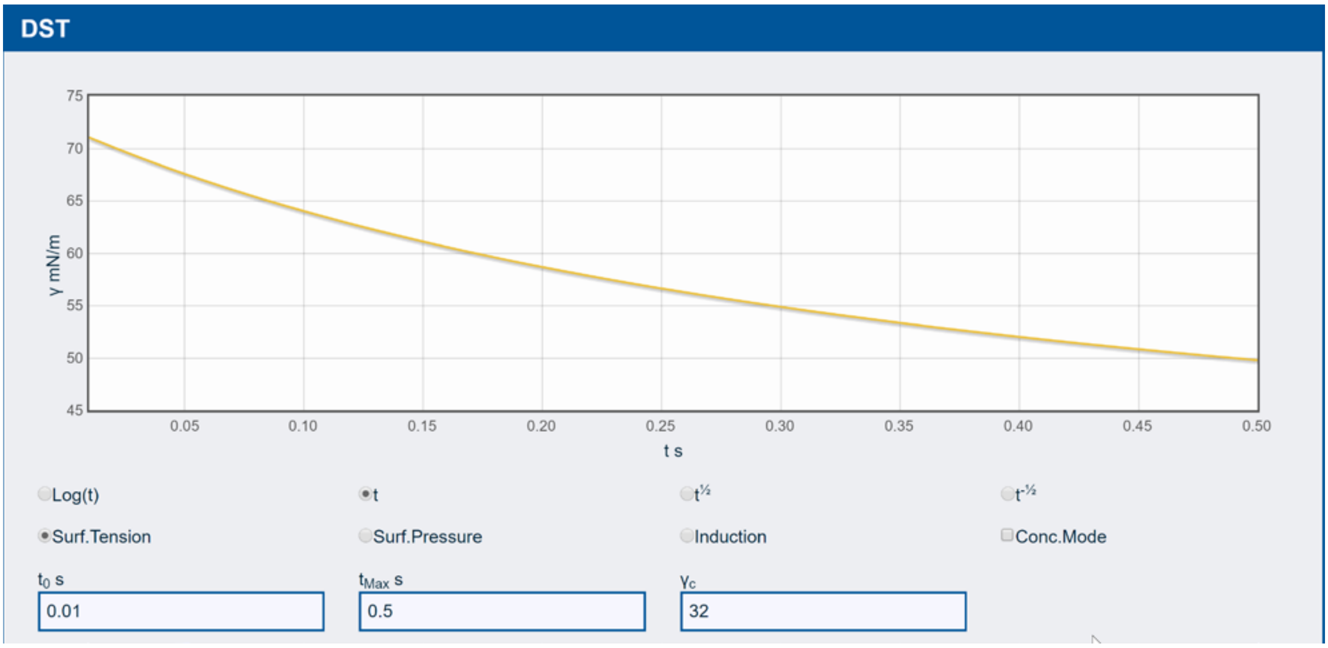
$$Ca = \frac{\eta V}{\sigma}$$

Equ. 1-1

The capillary number tells us the relative strengths of viscous drag forces, (viscosity, η, times velocity, V), versus surface tension forces, σ. From Ca we get a good idea of which factors are in control of any given step in the process. And it is no contest. Viscosities can range from 1 to 100,000 cP, velocities can range from 0.01 to 1000 m/s, while weak surface tension can change by a factor of ~3, from 25 to 72 mN/m or (more commonly) dyne/cm. If we are using just solvent-based printing then surface tensions can change only from 25 to 40, less than a factor of 2. You can find many papers where they try to see if a surface tension change from 32 to 36 dyne/cm is significant, a 12% difference, when the relevant process involves inks that can change viscosity by a factor of 100 simply by a change of speed or gap.

For water-based coatings, surface tensions are regularly measured to be in the 30 dyne/cm range using de Nouy rings or Wilhelmy plates. These values are then compared to printing performance. This has to be a complete waste of time, because during most printing processes we can guarantee that the surface tension will *not* be anything like those values. When we print we suddenly create lots of fresh surface and the surface active molecules that have been added to give the 30 dyne/cm values have not had a chance to get to the surface, especially if they are in any way a surfactant. It could well be that in the time-scale of a classic measurement, the surface tension of A will be 30 and that of B will be 33 dyne/cm, yet in the, say, 1ms timescale of the printing process, B might be 48 and A might be 54 dyne/cm. So the majority of papers that discuss the role of surface tension are using measured values that probably have little relevance to the effect in which they are interested.

The timescales are important to grasp. Let us print at a slow 60m/min, 1m/s. A dot will be in contact with a 1mm wide nip for 1ms. If the distance from printing head to drying oven is 0.5m then there are 500ms where surface tensions have time to evolve. Now let us look at a typical example of dynamic surface tension using an app:



○ Log(t) ● t ○ $t^{½}$ ○ $t^{-½}$

● Surf.Tension ○ Surf.Pressure ○ Induction ☐ Conc.Mode

t_0 s: 0.01

t_{Max} s: 0.5

γ_c: 32

App 1-5 https://www.stevenabbott.co.uk/practical-surfactants/dst.php

When you start the app, the default conditions are different from what we see here. The main thing is to change to a non-log plot and to set t_{Max} to our timescale of 0.5s. In this case, with a conventional surfactant, the surface tension has decreased from 72 to 50 mN/m, nowhere near the advertised value for this surfactant which is 32 mN/m which takes many seconds to reach. If we are printing at a more plausible 10 m/s then the mouse read-out shows that surface tension will have reduced only to 68 mN/m.

Of course, you can measure dynamic surface tensions using, say, bubble pressure measurements, and if there are some slight surface tension trends then these might correlate. But the number of studies using such dynamic surface tensions are only a tiny proportion, and, in any case, because the effects are going to be so small compared to inertia, it really isn't worth the effort except in very specific cases discussed below. Indeed, it is often difficult to get relevant dynamic surface tension measurements because the bubble method is not too good dealing with relatively viscous pigmented solutions.

Now let us turn our attention to surface energy. Again, for practical printing surfaces this changes by a factor of 2, from 25 to 50 across the whole range of good and bad printing, when ink adhesion can change by a factor of 1000. Indeed, adhesion from surface energy is typically 40 mN/m while even a sticky note is 4N/m and a strong adhesive tape (used to test ink adhesion) is 400N/m, 10,000x stronger than surface energy.

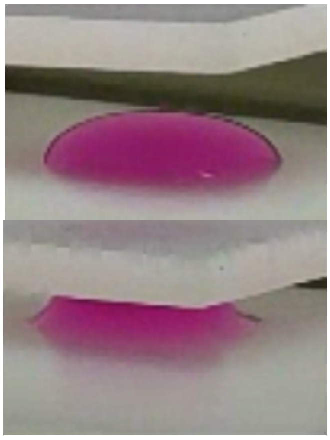
While many people focus, erroneously, on surface energy in the search for adhesion, even more focus on it, wrongly, because of naive ideas about wetting.

1.2.1 What about wetting?

In some very specific circumstances, surprisingly little known, the wetting characteristics as exhibited by contact angle can have a profound effect. These will be described below. The rest of the time, wetting and contact angles are irrelevant. How can this be?

Let us first start with an example where wetting *does* make a difference. If a liquid is going to flow into a fibrous material such as paper then the rate at which it flows depends on $\cos(\theta)$ the cosine of the contact angle θ. Across most likely fluids and fibres this angle will change from, say, 20 to 60, which means that the change in flow will be cos(20)/cos(60) = 0.94/0.5 ~1.9, basically a factor of 2. Now think of the other influences. Printing speed can easily double, halving the contact time. Yet the amount that flows depends on t^2 so this is a factor of 4. It also depends on the porosity of the substrate, which can change dramatically from substrate to substrate, and it depends on the applied pressure. Finally there is the viscosity which can change by a factor of 2 simply from speed (shear thinning) irrespective of the large range of real-world viscosities. In any case, when we examine the Walker-Fetsko equation we will find that the extra absorption into the substrate generally makes rather little difference to the amount of ink that is printed, and if you print under conditions where it makes a significant difference, you are likely to be out of control. So the surface energy factor of 2 that in theory affects travel into the fibre of the paper ends up being in the noise in terms of overall printing.

In defence of surface energy, everyone points to the fact that a drop of water will not spontaneously wet PE, a fact with which I do not disagree. However, other than for inkjet printing, this fact is irrelevant. We don't print by putting drops onto surfaces. Printing happens when one wetted surface is squashed onto another surface when the inked plate or roller comes into contact with substrate or blanket *and surface energy makes no difference to the wetting - a fact that is astonishingly little known while being rather obvious when you think about it*.



Suppose you have a drop of coloured water sitting unhappily on top of some PE. The images are from a video on my YouTube channel http://youtu.be/XXIGb6XFELY which also provides the credits for the experiment. Now bring another piece of PE closer and closer to the drop. What happens? Think of it from the drop's point of view. Does the drop like being in contact with PE? Obviously no. Does the drop like being in contact with air? Well, no. In fact, it dislikes air even more than PE because at least with PE there are some van der Waals (dispersive) attraction forces. So when the second PE gets close, the water positively jumps into contact and spontaneously wets out to form a continuous layer. The same thing happens between two pieces of Teflon. Liquids spontaneously wet out between *any* two surfaces, so surface tension and surface energies are entirely unimportant for one of the key aspects of any printing process (excluding inkjet). For those who don't believe me on this, there is a delightful paper called Teflon is Hydrophilic[5]. It points out four examples (one of which is similar to the PE drops example) where Teflon absolutely prefers to be with water than with air and so is "hydrophilic".

If we then use our PE and water system to simulate the splitting process that is intrinsic to all printing, and if we do things slowly we will certainly see a difference between how a PE-water-PE sandwich will split compared to a PE-water-(corona-treated PE) sandwich. There will probably be a bit more water ending up on the corona-treated side. So in terms of amount printed, in this artificial set-up, the wetting characteristic will affect the amount of ink printed. But here we are talking big blobs of liquid being pulled apart slowly. With any real ink in any real printing process, the splitting process is so rapid and violent that insignificant surface energy forces play no role, so that 50:50 splits are found overwhelmingly in the experimental data, even when there are deliberate attempts to change the surface energies.

1.2.2 What about offset litho?

Surface energy fans will point out two disproofs of my position. The standard offset plate has a low-ish surface energy hydrophobic ink-receptive region and a high-ish surface energy hydrophilic ink-repellent region. Surely offset printing is absolute proof that surface energy is important. The reader will have to wait till the offset chapter before discovering why surface energy is *not* the explanation; it is not even close. For those who like to be intrigued, one of the key experiments showing the irrelevance involved printing onto ice.

Even more so, waterless offset uses silicone and non-silicone regions to provide the non-printing and printing areas and, of course, silicone has a low surface energy. Again, the low surface energy is not the cause of the non-adhesion. A plate using Teflon with the same low surface energy shows no ability to produce waterless printing.

In both cases, the physics behind the non-attachment of the ink has been known for some time, is relevant both to silicone release liners and to ice-phobic surfaces, yet the blind belief in the magic of surface energy (just a factor of less than 2) has stopped many people from looking for, understanding and using the true explanation. This is all the more strange because conference papers

---

[5] Lichao Gao and Thomas J. McCarthy, *Teflon is Hydrophilic. Comments on Definitions of Hydrophobic, Shear versus Tensile Hydrophobicity, and Wettability Characterization*, Langmuir 2007, 23, 9125–9127

showing that surface energy was irrelevant to offset plates were relatively common in the 1980's.

1.2.3 Drop (non)spreading

We have all seen images of drops being applied to surfaces in contact angle measuring devices and a video often shows the drop spreading, with the amount of spread strongly dependent on surface tensions and surface energies, and, therefore, on contact angle θ. This is a powerful support for the idea that surface energy strongly influences drop spread in printing.

Think about it. Those drops in those videos are a few μl. Let us be generous and assume that the drop is a tiny 1μl and with a contact angle of 90° it is a hemispherical cap. Elementary geometry tells us that the drop would have a radius of 0.8mm. So the drop is 0.8mm high and 1.6mm across. What we see in those videos is, therefore, totally unlike anything relevant to real printing. The volume of a 1.6mm diameter offset dot, 2μm high is 4nl. And it starts its life not as a hemispherical cap 0.13mm high, but as a 2μm high cylinder so the effective contact angle is immeasurably small. Dot spread theory tells us that this printed dot is not going to go anywhere in any relevant timescale. Dewetting theory is discussed shortly because, at first, it indicates the necessity not just of a smaller equilibrium contact angle (the goal of many who worry about surface energies) but one that is very close to zero.

How *did* the offset dot get to its shape and size? By being a dot essentially the same shape but twice the thickness being smashed into contact with the substrate at a speed and viscosity that makes surface energy completely irrelevant, and then being ripped apart by a film splitting process featuring forces way beyond mere surface energy.

These facts about offset require no advanced science. Common sense tells us that surface energy is irrelevant (the same applies to gravure and flexo). So why has the printing industry wasted so much time worrying about unimportant parameters?

In fact, two printing techniques have been right to worry about it, because contact angle makes a huge difference, but not in the naive sense that is generally understood. Drop spread theory is profound, straightforward and surprisingly little-known.

Place a hemispherical cap drop onto a surface which can be completely wetted, i.e. the equilibrium contact angle is 0. This drop scenario is relevant only to inkjet, and the complete wettability is unlikely, but it illustrates the point. At the start it has a radius r and a contact angle θ which comes from the geometry, not surface energies. Remember - we have just put this drop into position and want to know what happens to it over time.

If you place the drop manually with a syringe, then put down the syringe and have a good look at the drop you will be disappointed. Not much will be happening. When you think about it, not much should be happening. Gravity is several orders of magnitude too small to be relevant and the only driving force is the surface tension of the drop which, if it's wetting, is likely to be rather low. If you stare long enough you will find that the drop is slowly creeping outwards and given a very long time will spread out fully.

However, if you look at the drop in a contact angle device with high-speed video, you will see a very different picture. The drop spreads rather rapidly at first, then slows down. The theory behind this was developed many years ago by Tanner and the basic equation could hardly be simpler:

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

Equ. 1-2

This tells us that the velocity, v, at which the drop edge moves depends on the cube of the contact angle θ and is driven by surface tension σ and held back by viscosity η. This simple equation has complex implications. Our initial high contact angle means that the initial velocity is high. But if the drop is spreading and the volume is constant, the height of the drop will decrease and the contact angle also decreases. Because the velocity depends on θ^3, the spreading slows down very rapidly. As Tanner showed, the radius of the drop goes as time, t^{-10} which is such a strong time dependence that if you look at a drop after a relatively short time, most of the action is already over. We can see this in the Drop Spread app:

Drop ●	Wedge ○	Cylinder ○			
Viscosity Pa.s	ST Dyne/cm	Radius µm	θ_{Start}°	$\theta_{Equilibrium}$°	
0.01	50	100	60	0	
T_{Max} s	R_{Max} µm	Height µm	Width µm	Volume	End Rad. µm
2	200	15	25	1.0nl	282.9
Start Height µm	End Height µm	Start µm	Start Angle°	End Angle°	End µm
57.7	8.0	200.0	n/a	n/a	n/a

App 1-6 https://www.stevenabbott.co.uk/practical-coatings/drop-spread.php

In this example, a 10cP inkjet ink with a surface tension of 50 dyne/cm forms a drop with an initial contact angle of 60° and with a radius of 100µm. The height of the drop is 57.7µm. This 1nl drop is large by printing standards. I use it for convenience. As the graph shows, after 0.2s the radius has increased by about

125μm, after 1s by 160μm and after 2s by 180μm. This is why I mentioned that casual observation shows a drop that hardly move. You would have to be very good to get a good view of the drop within 2s and by that time the t^{-10} effect has kicked in powerfully so the drop is expanding very slowly. The radius of the drop at 2s is 283μm and the height of the drop is 8μm.

What is fascinating about drop spread theory is that because σ can change only by a factor of 2, the driving force hardly changes. What about the resisting force? If you change the viscosity from 10 to 100cP (impossible for inkjet, but we are just exploring) instinctively you would think that the spread would be 10x smaller. Try this in the app and you find that the drop has expanded to a radius 223μm instead of 283μm. The 10x increase in viscosity has not even halved the drop spread! That is because the θ^3 term stays larger for longer.

If we want very little drop spread, what can we do? Because the driving force (unlike the powerful squeeze flows of other printing methods) is only surface tension then surface energy can be enough to stop the drop. If the equilibrium contact angle is 40°, then as soon as the drop spreads to that angle, spreading stops. In this example, the drop spreads by 17μm in 0.03s and stops.

So, at last, we have a use for all that work measuring contact angles. And we don't have to do anything complicated - we are not looking for dispersive and polar surface energies, none of those matter. We just want to know when our drop will stop. Will the ink then adhere? Because surface energy is 1000x too small, we don't care about it - we get adhesion by other means.

Thanks to the pioneering work of Prof Glen McHale (who kindly helped me back in 2004), we can do a similar analysis for printed lines in screen printing. For the large power densities of solar cells, a thick layer of silver ink is required. Because the ink blocks the sunlight, it needs to be as thin as possible. High and thin is the requirement, and those who have been following the story will immediately see the problem - the initial printed line will have a high contact angle and therefore a fast initial drop spread. The algorithm is rather complex, which is why we have apps, and the results are clear - that line spreading is a serious problem.

Drop ○ Wedge ● Cylinder ○

Viscosity Pa.s: 10
ST Dyne/cm: 50
Radius μm: 5
θ_{Start}°: 76.0
$\theta_{Equilibrium}$°: 0

T_{Max} s: 2
R_{Max} μm: 50
Height μm: 20
Width μm: 40
Volume: 0.9
End Rad. μm: 42.6

Start Height μm: 20
End Height μm: 10.9
Start μm: 50.0
Start Angle°: 76.0
End Angle°: 14.3
End μm: 125.2

Here we have a 10 Pa.s ink, 20μm high with a width of 40μm and a "wedge" of 5μm either side to simulate some sort of line shape. The exact details don't matter because the results are always similar. The starting contact angle is 76° so even with a 10 Pa.s ink the spread of the line is rapid. Within 0.4s the nominally 50µm line is 100µm wide.

Again, increasing the viscosity or decreasing the surface tension makes no significant difference; the only way to stop the spread is via a high equilibrium contact angle. If we set it to 40° the drop stops at 75µm wide.

As far as I can tell, there are only two other cases where surface energy is important for printing and, again, because there has been too much naive focus on surface energy itself, the real issue in the first one (pinholes) has been largely ignored. The real issue in the second case is very recent and raises an interesting question.

1.2.4 Pinholes

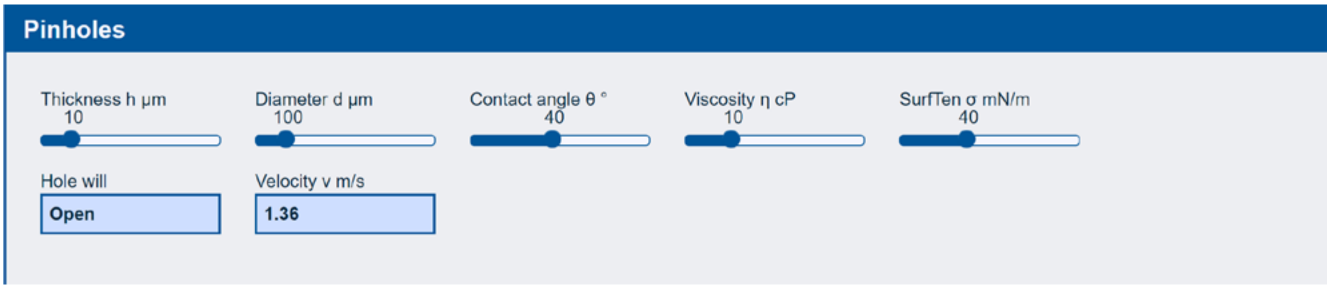
Let us return to the sort of printing that gives a thin layer and for which surface energy is irrelevant to drop spreading because the contact angle is so low (because the ink was squashed between plate and substrate) that there is no driving force. As long as the ink fully covers the substrate that is the end of the story. The surface energy is completely irrelevant to what happens in the short time before the printed ink is dried/cured.

Suppose, however, that because of dust, or some imperfection on the substrate surface or a fault on the plate, there is a small unprinted hole in the ink. As intuition suggests, the contact angle is important in influencing what now happens to that hole. Everyone knows that. What most printers do not know is that the hole diameter and the ink thickness are just as important. The focus on the popular surface energy factor, without knowledge of the other two factors has produced plenty of incoherent papers on pinhole and mottle formation. This is inexcusable because the theory is very simple. A defect of diameter d in an ink of thickness h will spontaneously grow if the contact angle θ is such that:

$$\frac{h}{d} < 2\left(1 - \cos\theta\right)$$

Equ. 1-3

The calculation is done in the app:



Thickness h µm 10

Diameter d µm 100

Contact angle θ ° 40

Viscosity η cP 10

SurfTen σ mN/m 40

Hole will: Open

Velocity v m/s: 1.36

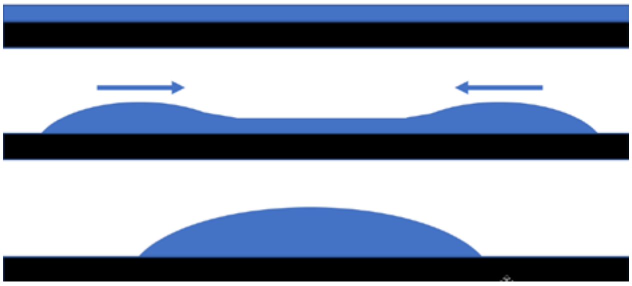
App 1-7 https://www.stevenabbott.co.uk/practical-coatings/pinholes.php

If the ink is perfectly wetting then cos(θ)=1 so h/d can never be less than zero so any small pinhole in the ink will spontaneously heal. If the contact angle is 60° then pinholes will spontaneously open when h/d<1, i.e. when the ink is thin and the starting defect is large. Of course, if the hole doesn't have time to grow then it won't be a significant problem. We already know how fast it grows, $\theta^3\sigma/\eta$, exactly the Tanner spreading velocity, with θ being the actual contact angle rather than the equilibrium angle used in the pinhole formula.

For relatively thick and high viscosity screen inks, pinhole spreading is not a massive problem (whereas dot spread is). For thin flexo and offset prints it is a borderline issue (and for gravure which prints in dots, it is irrelevant). For anything other than super-low equilibrium contact angles, the pinhole formula almost guarantees that any hole greater than a few µm will expand, making such printing nearly impossible. But because the actual contact angle is low, because of the thinness of the ink, the hole might not expand before drying. Offset inks are so thin and so viscous, that the balance is in favour of non-pinholes. For non-viscous flexo inks just after printing when they are relatively thick, the balance seems to be in favour of pinholes if the substrate is not treated (e.g. corona) to reduce the contact angle. Again, although people will seize on this admission that surface energy is of some relevance, there is a much bigger reason why offset solids are solid and flexo solids are full of holes. These reasons will be discussed when we get to the big issue of splitting of ink films. Once again, issues far more important than surface energy are at play, yet are largely unknown.

1.3 Dot unspreading

The pinhole problem is one aspect of how a printed area can fail. A problem that is related to both the drop spreading and pinhole issues is that of a flat, printed dot (effectively a very short cylinder, say 2µm heigh and 100µm radius) printed onto a substrate where its equilibrium contact angle is sufficiently high for the dot to want to bead up into a spherical cap with a much smaller radius. For a reason that are only obvious in retrospect, there was little experimental work on this subject; there was no self-evident way to arrange for a printed cylinder at t=0 and then follow its progress as it beaded up. A paper from a team that includes McHale who created the drop spreading theory used in that app, found an elegant solution[6]. By applying an electric field through a series of electrodes it is possible to pull a spherical cap of liquid down to a flat cylinder. As soon as the dielectrowetting field is turned off, a suitable video system can record from above and from the side how the dot returns to a spherical cap.



The title of the paper starts with "Not spreading in reverse". What they mean is that the dewetting process is not at all like the spreading process. Instead, a rim is formed at the edge of the dot and this rim moves in quickly, at a constant velocity, until it collides with itself, after which it relaxes exponentially into the final spherical cap. Despite this difference in mechanism, for the key process of the rims moving inwards we still have the same general velocity dependence as the other two phenomena, all we need to know is the constant k:

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

Equ. 1-4

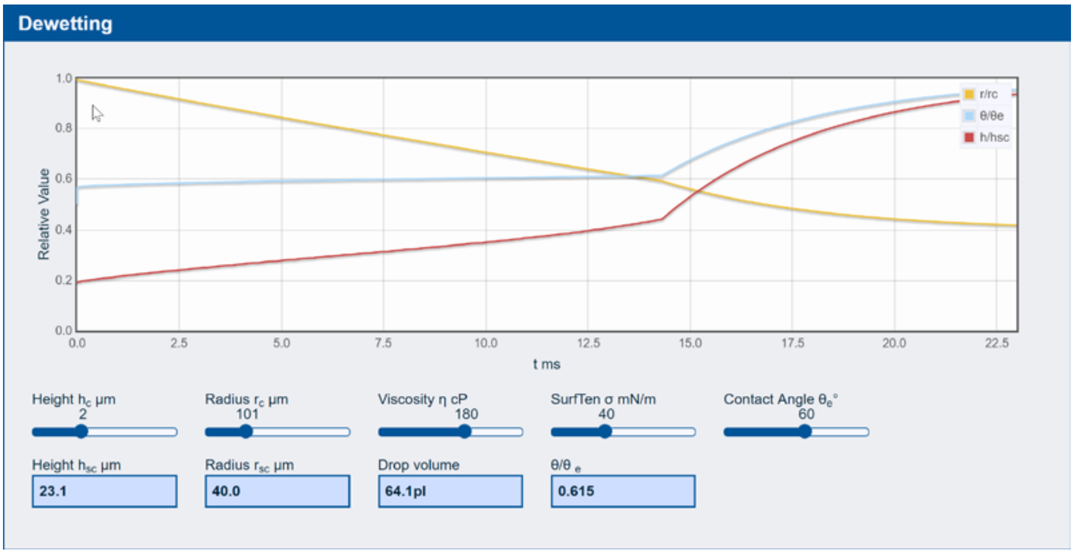
The θ³ term helps to explain why relatively modest changes in surface energy can transform unprintable dots into printable ones. Changing from a 60° contact angle to 30° reduces the velocity by a factor of 8 even though classical surface energy calculations based on cos(θ) give only a factor of 1.7. As always, surface tension can make only a small difference between 72mN/m of water and 36 of a typical formulation. Viscosity effects can be huge.

So, how do we calculate k, and what happens when the rims meet so that the near-spherical cap has to sort itself out? The second question is more easily answered - it is an exponential process with a timescale depending (approximately) on the inverse of the above equation. In the app the calculation is precise but the details don't concern us too much because we are not over-concerned about the precise shape and size of our dot at this point. Our real interest is the main dewetting, for which we need k.

Unfortunately, although v depends on k, k also depends on v so the calculations involve some complex loops. Those interested can read the paper. What is important is the results, and the app does all the hard work of producing them. I found them quite shocking.

---

[6] Andrew M. J. Edwards, Rodrigo Ledesma-Aguilar, Michael I. Newton, Carl V. Brown, Glen McHale, *Not spreading in reverse: The dewetting of a liquid film into a single drop*, Sci. Adv. 2016; 2 : e1600183



Height h_c µm 2

Radius r_c µm 101

Viscosity η cP 180

SurfTen σ mN/m 40

Contact Angle θ_e° 60

Height h_{sc} µm 23.1

Radius r_{sc} µm 40.0

Drop volume 64.1pl

θ/θ_e 0.615

App 1-8 https://www.stevenabbott.co.uk/practical-coatings/dewetting.php

In the example shown in the screen shot (you will need to set it up for yourself as the app's default values are those of the original paper) a 2µm high, 100µm radius printed dot takes only 22ms to shrink back to a 40µm radius spherical cap, 23µm high with a contact angle of 60°. If (not shown) the contact angle is 30° then it takes 120ms to reach a 53µm radius spherical cap 14µm high. Some might argue that a 60° contact angle isn't too good, but most of us would be very happy with a 30° contact angle, yet still the app shows that a printed dot would dewet before any solvent had a chance to evaporate - even for a 180cP ink which is higher viscosity than most flexo or gravure inks.

But we simply do not routinely see this amount of dewetting - we tend to have dot gain, not dot loss. So either the theory is wrong or there is some other effect of importance. As the theory seems to me to be sound, and in the absence of any data either way, I suggest that there is another effect influencing the dewetting. The original research was done with a pure liquid which will show more-or-less equal advancing and receding contact angles because it has no interaction with the pure substrate - basically Teflon. I suspect that real, complex inks, with substrates into which they have some "bite" (polymers, primers) or penetration (papers) have a much lower receding contact angle, especially when the dot has been smashed violently into place between the printing plate and substrate. Pigment particles and polymers might be relatively comfortable on the surface and, of course, interact well with the ink, so the effective contact angle will be 0, so there is no dewetting. My search for relevant literature has not provided any proof of this, but it is clear that it is relatively easy to get a 20° difference between advancing and receding contact angles when there is even very modest interactions with the substrate.

I might be wrong. If anyone does the experiments to prove/disprove the hypothesis I will be delighted to update the book with the results.

1.4 Speed, density and length

In addition to viscosity and surface tension, speed, density and length (or length scale) are the fundamental parameters behind printing science.

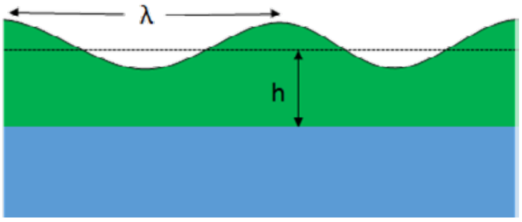
We have already seen that speed (or time) has a big impact on the values of the viscosity and a modest effect on the surface tension during the various stage of the printing process Also, "viscosity" is able to change to "elasticity" at high-enough speeds (temperature-time equivalent). Speed also has a large effect on the capillary number, so inertial effects rapidly overwhelm surface effects.

Density is not of massive importance, simply because it does not change all that much, from 0.8 for a hydrocarbon solvent up to 1.2 or so for a pigmented ink. It is included here for intellectual completeness as discussed in the next paragraph.

Length scales (plus density) feature in a number of key fluid dynamics properties such as the Reynolds number and Ohnesorge number which, in turn, affect whether we can produce a good inkjet drop. The first example of the importance of length scale is its 4th-power effect on levelling.

1.4.1 Levelling

Suppose that the feature you have just printed has some imperfections in terms of peaks and valleys. For a high-quality print it is important that the imperfections quickly level out. Our key quantities of viscosity, surface tension and length are important in determining whether there is time for the levelling to occur. The theory which describes the effects comes from Orchard.



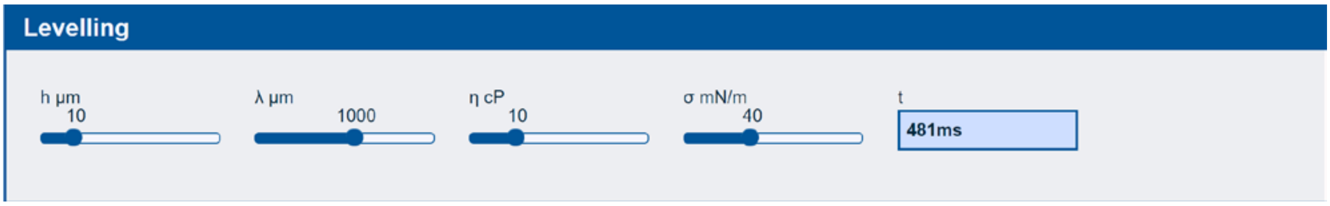
We have an idealised defect of length λ in a print of thickness h (this is *not* the height of the defect). The ink has a viscosity of η and surface tension of σ. How long does it take for the height of the defect to reduce by 1/e, i.e. by a factor of ~3? The levelling is driven by surface tension, so the time will depend on 1/σ and is held back by the viscosity so the time will depend on η. Note, incidentally, that contrary to popular mythology, a high surface tension coating will level out faster. Adding silicones to lower the surface tension will make this problem worse; giving us a clue that their popularity in terms of producing a smooth coating is related to a different problem discussed next.

Orchard tells us that the levelling time, t, is given by:

$$t = \frac{3\eta\left(\lambda / 2\pi\right)^4}{\sigma h^3}$$

Equ. 1-5

The calculations are done in the app:



Levelling

h µm 10

λ µm 1000

η cP 10

σ mN/m 40

t 481ms

App 1-9 https://www.stevenabbott.co.uk/practical-coatings/levelling.php

The first things to note are that the viscosity effect can be significant because viscosities can change by many orders of magnitude, while the surface tension effect is modest because at most it can change by a factor of 2. The other two items are hardly mentioned in the printing literature yet are far more important. First, an apparently modest reduction in thickness from 4µm to 2µm will give an 8x increase in time needed to level, so a well-intentioned reduction in ink thickness can have devastating consequences for print quality. Second, if the defect spacing happens to double, then the levelling time increases by a factor of 16. So a small change to a gravure cylinder pattern or to the creation of defects during the splitting process can have a large effect on print quality.

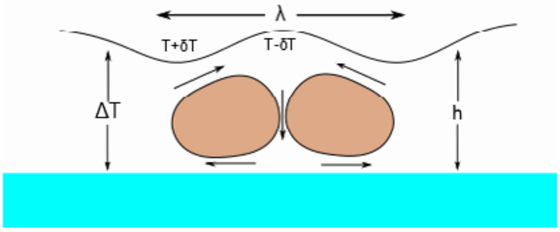
1.4.2 Marangoni (and orange peel)

One day, as a relatively inexperienced manager in charge of a new coating on a new coating machine, I was startled to see a perfectly good substrate go through a perfectly good coating nip and pass into a perfectly good oven - to emerge with the whole coating showing a delightful pattern of hexagons. Everyone looked at me expecting me to know what to do, but I had no ideas. It was time to halt the trial while I went to seek help.

I started to talk to our resident expert and no sooner had I said "hexagons" than he said "Marangoni". With his instant diagnosis, a quick fix to the problem was suggested and we soon had a high-quality coating.

Many years later I was coating a customer's formulation on my machine and there were an unfortunate number of streaks in the end product. The customer was not impressed with our incompetence. Something about the streaks made me take my customer to our microscope and as soon as the image appeared I could say "Ah, hexagons at the start of each streak - this is a Marangoni problem and I know how to fix it". A slight tweak to their coating solution and the coating was perfect.

How does the presence of hexagons confirm that the root cause is Marangoni, and what is the Marangoni effect?



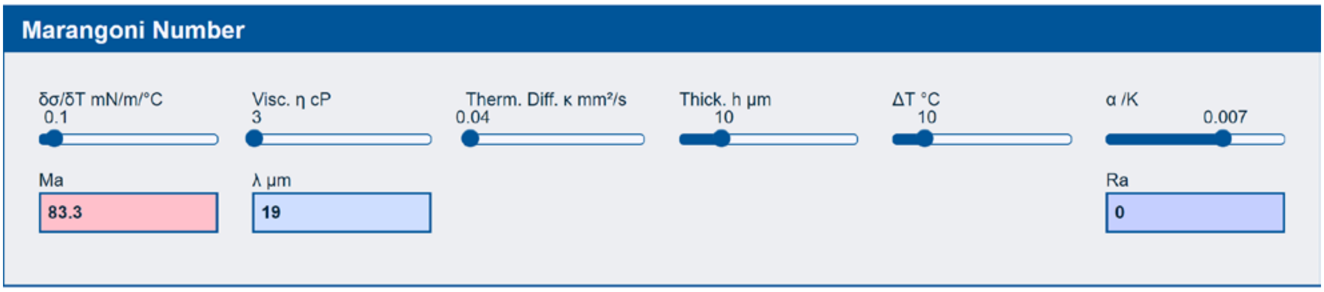
As solvents evaporate from a coating, a temperature difference, ΔT can appear between the top and bottom of the coating. Because surface tension is temperature sensitive, this sets up a surface tension gradient which drives a flow. These Marangoni flows appear in random places, caused by the slightest difference in the printed surface. Some of these flows will interfere with each other - a flow going down to the substrate meeting a flow coming up from the substrate. Others will start to support each other. With remarkable speed, these flows self-coordinate in the only way possible - as a series of hexagons. You can sometimes see similar patterns in a pan of water being heated - these are called Bénard cells.

Such patterns can only form when the conditions are right - and these are captured in the Marangoni number Ma which depends on the coating thickness h, the change of surface tension with temperature, δσ/δT, the viscosity η and thermal diffusivity α. They form a characteristic wavelength λ:

$$Ma = \frac{\delta\sigma}{\delta T} h \frac{\Delta T}{\eta\alpha} \qquad \lambda = \frac{2\pi h}{\sqrt{Ma/8}}$$

Equ. 1-6

Why do we see Marangoni effects comparatively rarely? The common rule of thumb is that if Ma<80 the effect is unlikely to appear. Calculating this number is tricky if you happen not to know how surface tension changes with temperature or the value of the thermal diffusivity. Fortunately the app not only does the calculations but gives you some ideas of the values to use.


Marangoni Number

δσ/δT mN/m/°C 0.1
Visc. η cP 3
Therm. Diff. κ mm²/s 0.04
Thick. h μm 10
ΔT °C 10
α /K 0.007
Ma 83.3
λ μm 19
Ra 0

App 1-10 https://www.stevenabbott.co.uk/practical-coatings/marangoni.php

In the app example, the thickness is relatively large and viscosity low so Ma is greater than 80 and the red warning appears in the box. Thinner coatings and higher viscosities obviously reduce Ma. A high thermal diffusivity means that the driving force is swamped by rapid thermal equilibration across the layer.

Although I have described the effect in terms of thermal surface tension gradients, the same effect ("Gibbs cells") can be created via surface tension gradients created by local concentration differences if there is faster evaporation in local areas.

Those of you who have never seen hexagons may be wondering why your time is being wasted with Marangoni. Even if you have not seen hexagons, many of you will have suffered with prints or coatings showing "orange peel" effects, a hard-to-pin-down pattern that has a hint of regularity about it. Although there are other "orange peels" out there, my view is that the most common is caused by Marangoni effects, without the full hexagon formation that would make the fault easier to identify and cure.

There are two different approaches to curing Marangoni-like effects. The first recognises that local surface tension differences can also appear via composition fluctuations in a mixed-solvent formulation. In both of my Marangoni outbreaks this was the root cause and was fixed by reducing the difference in evaporation rates of the solvent blend. If, for example, the MEK in the blend was evaporating much faster than another component, then swapping to MIBK might reduce the mismatch.

The second approach is to drown out the surface tension gradient. With water-based coatings, formulators throw in surfactants to do this, often with disappointing results. First, the surfactants might be too slow (remember the discussion on dynamic surface tensions), second because surfactants can build up their own gradients, making things worse. That is why throwing in some silicones so often fixes Marangoni and orange peel problems. The silicone floods the surface, giving a uniform surface tension because silicone surface tensions change very slowly with temperature. If you casually read articles about orange peel you often see that the surfactant "helps the coating to flow better". This is a meaningless statement, and the "flooding" of the surface tension gradient seems a much better explanation. Note, too, that in terms of levelling, a lower surface tension increases the time required. So a surfactant might fix Marangoni flows to cure orange peel, yet reduce the tendency to level out other coating non-uniformities. Understanding the different causes and cures helps find the right package of solutions.

1.5 Ink splitting

Those reading explanations of how printing works might hardly be aware that the single biggest puzzle in printing is the ink splitting step. Indeed, many printers imagine that ink is just "transferred" from plate to substrate via a process that is self-evident. Many printers seem to think that applying a bit more pressure will transfer more ink, somehow "pushing" it onto the substrate.

The two exceptions to this general feeling are inkjet - where most people know that ejecting lots of perfect drops of inks is not entirely easy, and screen printing where most people think they know that the ink is pushed out by the squeegee.

These folk notions of printing are correct only for inkjet; they are comprehensibly wrong for the other techniques. Being wrong about the core step in printing means that attempts to fix problems will invariably be misguided. So it is rather important to show what is really happening with ink splitting. Before looking at splitting in a printing nip, we see why it is an important process within inkjet.

1.5.1 Inkjet drops

An inkjet nozzle doesn't eject drops, it ejects a stream of fluid at a carefully controlled set of velocities (the actuators follow complex driver waveforms), with some sort of abrupt halt when enough liquid has been ejected. The ejected volume has to separate from the liquid that should remain in the head. We want this ink splitting to be as perfect as possible, where the initial drop plus tail should end up as a pure drop. Done incorrectly, the drop plus tail ends up as drop plus satellite which itself can fly off at a strange angle.

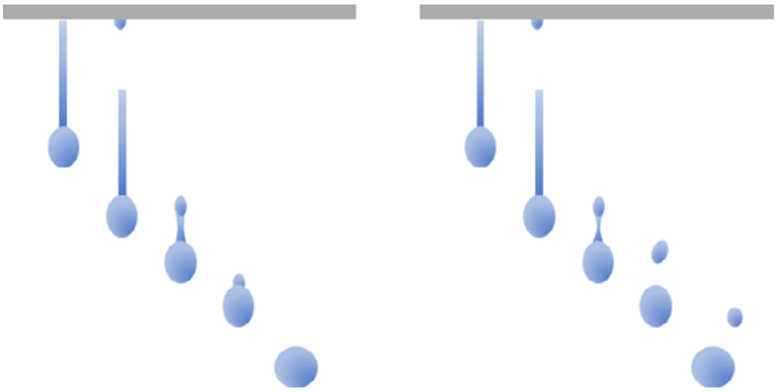


Figure 1-2 Desired (left) and undesired (right) inkjet drop splitting

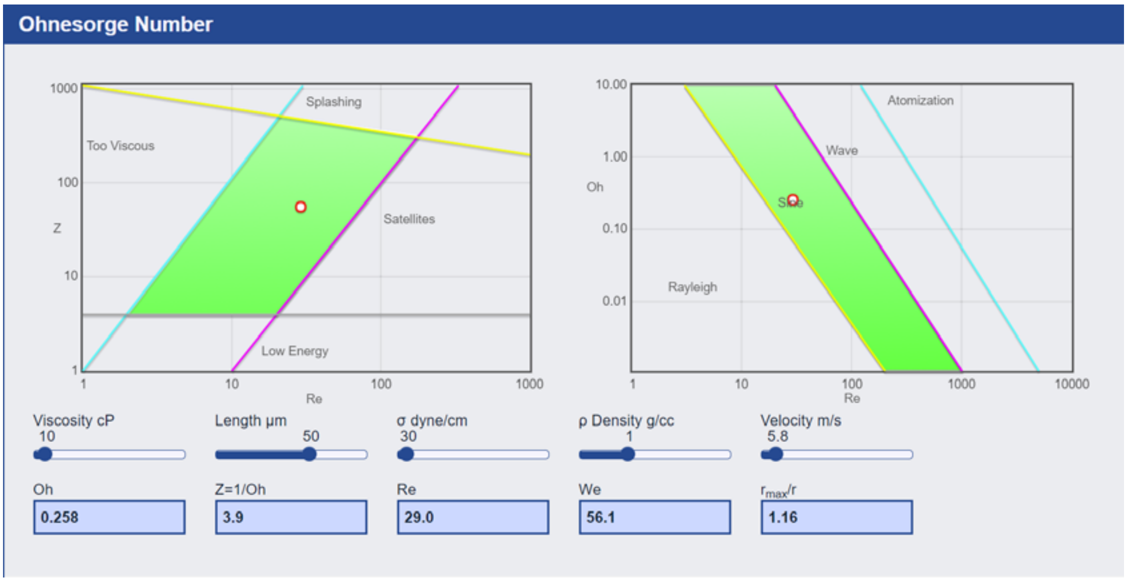
The natural instability of any stream of fluid can readily be observed when slowly turning off a flow of water with a tap (faucet). The steady stream becomes an unsteady stream then, thanks to the Rayleigh instability, breaks up into individual drops. Although an entire chapter could be written on the relevant phenomena, we can summarise what happens in two dimensionless numbers: the Ohnesorge number, Oh, and Reynolds number Re plus the related Weber number, We:

$$Oh = \frac{\eta}{\sqrt{l\rho.\sigma}} \quad \mathrm{Re} = \frac{V\rho l}{\eta} \quad We = \frac{V^2 \rho l}{\sigma}$$

Equ. 1-7

The Ohnesorge number gives us a ratio of viscous, η, to inertial and surface tension forces via the length, l, (in this case the diameter of the inkjet nozzle), the density (which is always going to be close to 1) and surface tension, σ and tells us (mostly) what will happen when the drop hits the substrate. The

Reynolds number is the ratio of viscous to inertial forces and tells us whether it will or will not form a single drop. At low Re the ink simply won't form a drop and at high Re it will tend to form satellite drops which fly off in all directions, degrading print quality.


Ohnesorge Number

Viscosity cP 10

Length µm 50

σ dyne/cm 30

ρ Density g/cc 1

Velocity m/s 5.8

Oh 0.258

Z=1/Oh 3.9

Re 29.0

We 56.1

r_{max}/r 1.16

App 1-11 https://www.stevenabbott.co.uk/practical-coatings/ohnesorge.php

The app shows an inkjet drop comfortably in the safe zone in Oh/Re space, not surprising given its modest viscosity, surface tension and density along with a typical nozzle diameter of 50µm. By playing with the app you will find just how restricted inkjet printing is by the conflicting demands of speed, viscosity, surface tension and nozzle size.

While we are here, it is interesting to look at the final output box, r_{max}/r. Many of us imagine the inkjet drop smashing into the surface with spectacular effects. Reality is disappointing, the velocities are so low that the drop hardly deforms on impact. The calculation gives the ratio with respect to the original radius of the maximum radius reached during the impact process, and for most inkjet-relevant conditions the value is very close to unity. The formula uses both We, Re and the contact angle θ:

$$\frac{r_{\max}}{r}=\sqrt{\frac{We^2+12}{3\left(1-\cos\theta\right)+4We^2/\mathrm{Re}^{0.5}}}$$

Equ. 1-8

1.5.2 Normal splitting

The general extent of discussions on splitting shows this highly idealised picture of two rollers with ink between them and the cavitation, filament formation and misting (if the filaments snap) that can occur when an ink film splits.

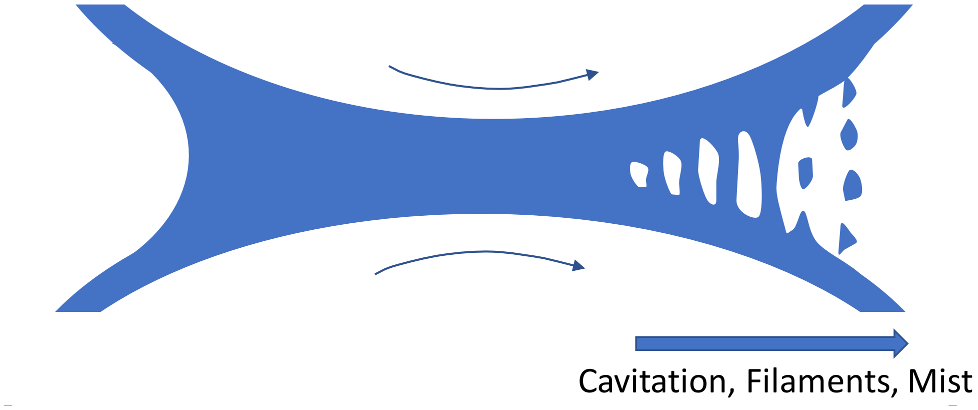


Figure 1-3 The standard view of ink splitting. Note that it is spectacularly out of scale so is a highly deceptive image.

This is fine for those who have multiple roller trains, such as in the ink supply to an offset plate. It teaches that 50% is transferred onto each downstream roller, so to reduce the ink thickness by 256 times you must have 8 rollers. It is also a warning to ink designers to be aware of the effects of, say, viscoelasticity on the splitting behaviour: the split might just be a nice meniscus, it might be just cavitation (holes in the liquid), it might be filaments (creating cobwebs over the machine) and it might be that the filaments spontaneously break into fine mist particles. Because everyone shows a diagram like this and most people know that you can hear the ink splitting on high-speed rollers (there are academic papers on measuring these acoustic emissions), you would imagine that the rules for ensuring that the ink split in the least undesirable manner would be well-known. My search through the literature has found a few patents that link sensibly to viscoelasticity, a number of papers that try some correlations of some parameters with other parameters to rather little effect - and that's about it. It seems bizarre to me that customers are not demanding that their ink suppliers provide inks with a desired low (or high) viscoelasticity when measured at certain speeds that match the splitting process. Instead there is some demand that the ink should have the right tack, with no clear idea of how tack does or does not relate to whatever splitting issue is bothering the press operator.

We might, however, argue that a bit of misting in the ink train is no big deal and can be managed as a problem, so it is possible to forgive the ink community for not being too clear about the issues. What *is* unforgivable is the near-total silence about the issues of ink splitting when the printing plate (or screen stencil) starts to separate from the substrate. This *is* the printing process and is generally ignored in discussions of what is going on. There are acres of papers on paper surfaces, acres of papers on surface energies, and almost no papers on the violent, high energy, short time-scale processes that control printing.

Even for roller splitting there have been arguments that the surface energy of the roller (e.g. ceramic versus steel) might make a difference. This is wrong for many reasons, one of which is that the surface energy of ceramic or steel that has encountered even a few molecules of ink is very different from the surface energy carefully measured on a perfect, fresh roller. Anyone who measures surface energies knows the importance of cleaning the surface to avoid false values. Yet this well-known fact is ignored when people try (and fail) to control ink splitting via changing surface energies of rollers. [For what it's worth, my opinion is that if there are differences, they are likely to be due to microstructural surface effects which alter the cavitation behaviour discussed below.]

When we come to offset printing we will even find that the obvious explanation of how the hydrophobic ink avoids the aqueous plate or (equally) the silicone layer of a waterless plate has *nothing* to do with surface energy. Shortly we will discuss how some ideas from pressure sensitive adhesives can illuminate some key issues in classic ink splitting. It turns out that some key experiments in another aspect of these adhesives provide the explanation of how the ink reliably refuses to remain attached to the wrong part of the offset plate - with surface energy playing no part at all in the explanation. This will all be dealt with in the Offset chapter.

Because the splitting process differs between flexo/offset, gravure and screen, the three different processes will be examined separately. The common elements uniting them will be discussed afterwards, though one such element needs to be faced from the start.

1.5.3 Splitting in flexo/offset

One of the many reasons that the standard diagram of splitting is misleading is, as the caption mentioned, that it is hopelessly out of scale. If you imagine a 100µm offset dot being printed, with an ink layer of 4µm (leaving a printed dot of 2µm because the split is always close to 50:50) then the splitting process, to scale, looks more like the following. In the Offset chapter a different (complementary) way of looking at the split in terms of the different length scales is discussed:

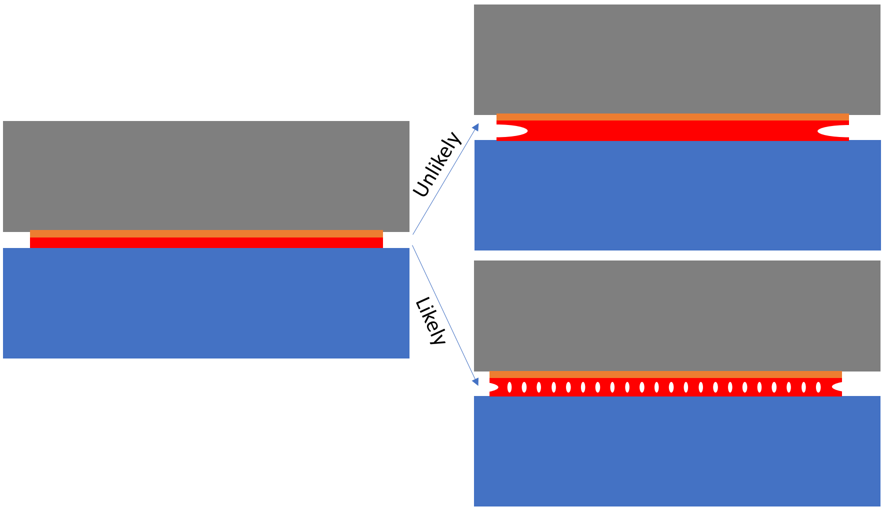
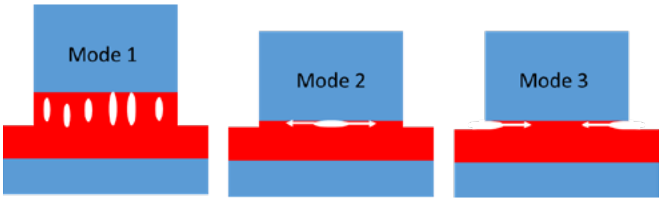


Figure 1-4 Splitting in a flexo or offset 100µm dot: gentle meniscus versus violent cavitation

If the splitting were the process of meniscus travelling from either end of the dot then the printed dot would be near-perfect. Far more likely is that there is sudden cavitation ripping the dot in half, leaving a deeply unhappy printed dot in the millisecond after printing. The fact that 100% solids in flexo are usually filled with holes follows quite straightforwardly from this cavitational form of splitting. Yet this elementary truth seems largely unknown to the flexo world who carry on insisting that any holes must be due to the wrong surface energy. As we saw with the discussion of pinholes, the wrong surface energy can make a small defect enlarge, so a higher surface energy can, in some cases, make flexo solids somewhat less bad. But this is not attending to the root cause.

Those readers who happen to know about pressure sensitive adhesives, PSA, may know about the probe tack test[7]. The adhesive is squeezed between a glass plate and a steel probe, then the probe is pulled vertically, monitoring the forces versus distance, while taking videos through the glass and from the side. From such experiments one can find that there are three modes of splitting.



Mode 1 is the classic, stringy failure required of most PSA. It is formally identical to the filament creation shown in the classic splitting diagram shown earlier. Mode 3 is the perfect cracking from the edges that can be found with rigid adhesives. Mode 2 involves cavitation in the centre with cracks propagating out rapidly.

It is important to note that the Mode depends not only on the rheology of the material being split but on the ratio of thickness to height. So a material that might split easily in Mode 1 as a thick sample readily imaged in a microscope on a µl-scale scale will fail in Mode 3 when it is a super-thin printed dot on the pico-litre scale.

My belief (though I have little data to support this) is that for flexo we are somewhere between Modes 1 and 2 and for offset between Modes 2 and 3. There is evidence in gravure printing for a slightly different Mode 2, called (Saffman Taylor) viscous fingering[8], but these cavitation/fingering modes are basically similar.

---

[7] Although I'm biased, the best description I know of for these effects is in my book *Adhesion Science: Principles and Practice*, DEStech Publications, 2015. The 3-mode analysis comes from Creton, one of the big names in PSA science.

[8] It is discussed, and evidence provided, in terms of sophisticated gravure printed electronics, in the thesis by Nils Bornemann, *Characterization and Investigation of Large-Area, Ultra-Thin Gravure Printed Layers*, TU Darmstadt, 2014

Why am I offering such vague ideas? Because the printing science community, with few exceptions, have not even begun to think about what is happening during splitting, and why offset gives rather clean solids and flexo gives such poor solids. In particular, the community has not taken rheology seriously, because rheology teaches us that mechanical properties change drastically with the measurement speed. So an ink that might have a high G'' and low G' (i.e. be nearly purely viscous) at time-scales that are convenient to measure (a few Hz) might have a much larger G' component at the time-scales involved with ink splitting. If this shifted a flexo ink from Mode 3 towards Mode 2 then the quality of the solids would improve dramatically.

Or if the plate were designed to initiate lots of small cracks during the splitting process, the failure mode would automatically be closer to Mode 2 and the print quality would thereby improve. By this, I take the analogy from the extensive research on PSA. By carefully changing the surface, the generation of cavitation voids can be enhanced or suppressed, allowing some control over the precise mode of splitting. Such thinking is behind my comments that splitting might be different between chromed and ceramic rollers depending on their microstructure and, therefore, their ability to initiate cavitation.

As mentioned above, there are hints in the patent literature that some of the better ink manufacturers have some grasp of how, say, relative G' and G'' can be altered to improve quality while having to cope with, say, misting in the ink roller train. One can imagine that there are conflicts between a high G' being desirable for clean ink splitting and a low G' optimal to reduce filaments and misting. If the printing community were more aware of these effects and their trade-offs (with different parts of the process operating at very different timescales), a more rational debate could be held about how the ink, press, plate and substrate could be optimised for the highest quality with least effort. Or maybe the better ink manufacturers already know all this and are just keeping their know-how to themselves.

1.5.4 Splitting in gravure and anilox

Some may regard this section as unnecessary. Surely gravure cells just "empty". OK, so they might only partially empty, but why do we have to consider splitting? The answer is that you cannot fight the laws of physics. At the moment of contact between the cell, the ink and the substrate the ink is as much part of the substrate as it is part of the cell, so the only way for it to be sub-divided is via splitting. A cartoon version shows the basic steps. The first time Prof Nik Kapur at U Leeds and I drew the cartoon we had no evidence for it other than the fact that it agreed with the known laws of physics - unlike the alternatives. Subsequently, detailed simulation work by groups such as those of Prof Carvalho's in Rio have shown that this is how things happen in reality and diagrams like this have become routine in the work (discussed in the Gravure chapter) of those concerned with high-performance ultra-small gravure printing for electronics.

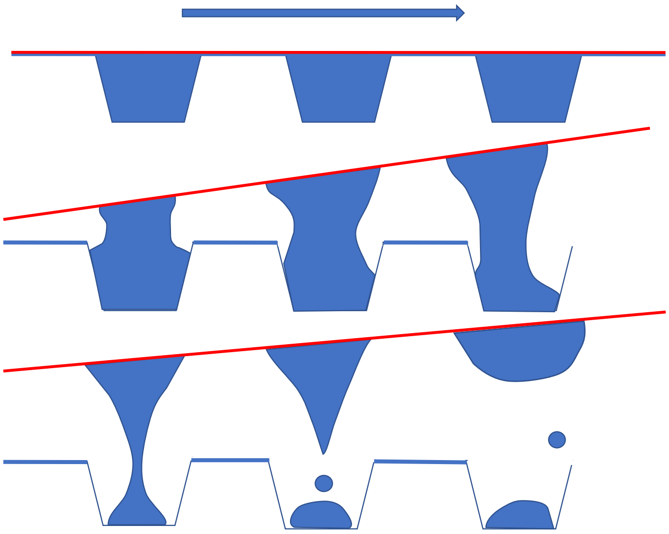


Figure 1-5 Splitting in gravure. The positions of the contact lines within the cells can vary greatly depending on conditions.

As the substrate separates from the cells the ink is dragged up and a bridge is formed. At the same time the edge of the ink moves down the cell, at various contact angles depending on speed and viscosity. Finally the liquid bridge breaks with, potentially, a mist drop (as shown) or a filament that appears as a spider's web on the print.

The reason we have, in general, no idea how much ink remains in the cell after transfer is because so many things are happening at the same time and complex issues such as dynamic contact angle and speed/time dependent rheology come into play. A cell of a given shape and depth might work well for one ink at one speed and badly for another. The ability to model this complex transfer process is a rather recent triumph of computational fluid dynamics. We have to hope that some usable rules will start to emerge from such studies. Until then we have to continue with the rule of thumb that ~40% gets transferred unless proven otherwise, and think about how to optimise the ink to avoid filaments and mists.

1.5.5 The Walker-Fetsko Equation

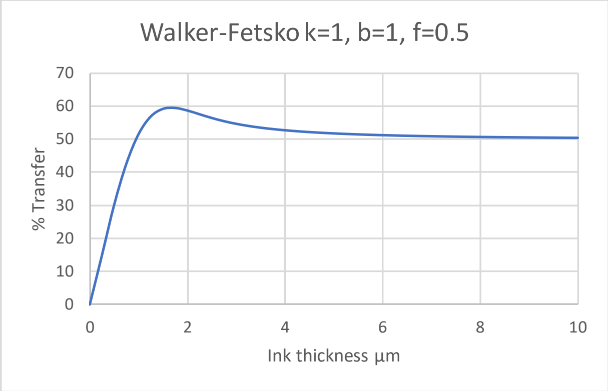
So far the implications have been that splitting is 50:50 except for gravure where it might be 40:60. There is one obvious exception to these values - when printing onto an absorbent substrate it is possible that during the time in contact some of the ink goes into the substrate before splitting, so the overall split is higher.

In fact, given that absorbent substrates such as paper tend to be rough, thin layers of ink on the printing surface may make imperfect contact with the substrate, so the split will be much *reduced*; you cannot transfer ink without contact. These two contradictory effects were noted by Walker and Fetsko who produced a formula to describe them. There have been many subsequent attempts to improve on Walker-Fetsko but they all suffer from the fault of the original - that the parameters needed for the equation come from fitting the experimental data and do not derive from primary physics. In other words, if you have enough data to use Walker-Fetsko, you don't need Walker-Fetsko because you already have the answer to whatever question you needed to ask.

It is still nice to have an idea what the equation entails, so here it is:

$$y=\left(1-e^{-(kx)^2}\right)\left\{b\left((1-e^{-x/b}\right)+f\left(x-b\left(1-e^{-x/b}\right)\right)\right\}$$

Equ. 1-9



The equation (with its app) is discussed in more detail in the Offset chapter, but, briefly printed ink thickness, y, depends on the original thickness x and the constants k (which depends on the paper roughness), b (how easily the paper can accept the ink) and f (the pure splitting ratio in the absence of absorption). The term in the first bracket is how much of the ink is in contact with the paper so controls the curve at low values of x. The $(1-e^{-x/b})$ term describes how much ink gets into the paper. It is usual, as here, to plot the % Transfer, y/x.

In practice, if you assume that it is unwise to print with values of x where the first term is significant, then the formula in terms of fraction transferred becomes:

$$\frac{y}{x}=f+\frac{b}{x}(1-f)$$

Equ. 1-10

which shows a higher fraction for small values of x, levelling out to 50% for higher thicknesses if f=0.5.

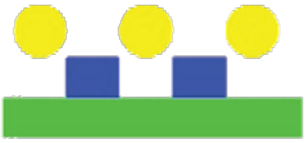
The problem with high values of b is that a, say, cyan dot printed onto paper will have a stronger colour than the same dot printed onto a previous, say, magenta dot which has already used up the ink absorbency of the paper. The question of what controls b is discussed in the Absorption part of the Drying section.

Of the many papers I've read on Walker-Fetsko, the curves from any real-life printing are rather boring; there is a small region where the printing pressures are absurdly, irrelevantly low, so you get less ink transfer, then there may or may not be a slight boost from paper absorption, and the equilibrium splits differ slightly from 50:50 perhaps because the paper substrates cause an earlier cavitation than on the plate, leading to a slight asymmetry. But the effects are tiny compared to a tweak of, say, the ink supply in the offset system, so they are of some modest intellectual interest. Some experiments show an anti-porous effect - the ink goes into the pores but because they are so large the ink comes straight out again, so a more porous paper might show less printed ink. Whether you call this Walker-Fetsko or common sense is up to you.

1.5.6 Splitting in screen printing

One of the standard questions in screen printing was: "How does the ink come out of the mesh?". To this question the usual answer was that the energy from the squeegee forced the ink out of the mesh. This answer *has* to be wrong for a very simple reason: when the squeegee is near the point of printing, the mesh is in contact with the substrate so no ink transfer can take place. Only when the squeegee is mm or cm away (and its influence is long past) can the ink transfer happen. That this obvious fact was ignored by the screen print community for decades is another sign of how printers refused to engage with real, useful science.

For those keen to know how the ink comes out of the mesh, the answer is that it doesn't. Instead the mesh comes out of the ink. How else could printing take place?


How they think it works


How it really works

Although I am co-author of an academic paper on the screen print process and can describe the science in detail, the whole process can be simply visualised. The first task is to distinguish between how most printers think the ink behaves and how it really behaves. The classic statement is "Ink isn't stupid, it knows that it can't flow through the mesh". Hence the diagram on the left where the ink has been pushed through the mesh by the squeegee and by some magical process will even-out beautifully. Note how clean the mesh is in most printers' imaginations. What really happens is the diagram on the right, showing that the mesh comes out of the ink and, just like a spoon coming out of honey, brings some ink with it - ink always splits one way or another, and this is how it splits with screen. Well, almost, the image shows that we still have a liquid bridge which is just about to split.

When we discuss screen in general, the whole process of flooding, squeegee filling/scraping, printing will be described. Here I simply show the mesh coming out of the ink, forming the bridge, leaving a little drop (or, if the ink is badly designed, forming a cobweb filament), and ending up with about 40% of the original ink wrapped around the mesh with the 60% left on the substrate.

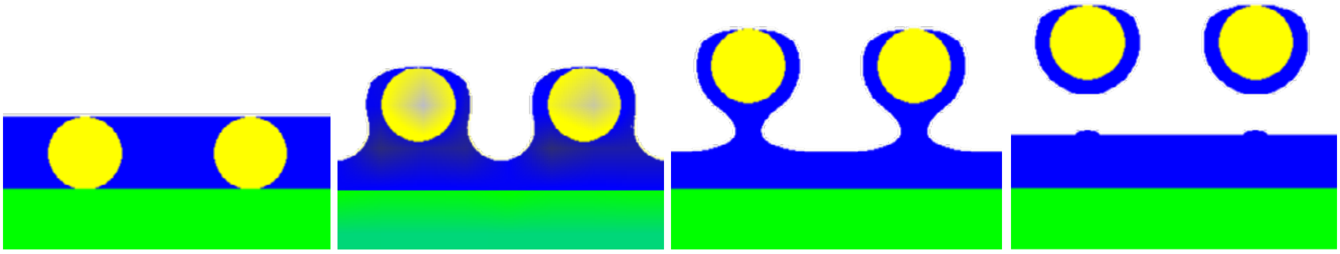


Figure 1-6 Screen printing: the mesh comes out of the ink

The idea that there is more ink underneath the mesh than between it is totally in contradiction to the folk mythology version, yet can be easily verified by anyone who cares to look. The pattern observed is of a bigger deposit under every second mesh knuckle, meaning that over+under behaviour is different from under+over. I never had access to the right 3D modelling to simulate a real mesh to properly understand what is going on. The simple 2D animation proved good enough for the screen print troubleshooting I had to carry out.

When my colleagues and I first proposed this model it was obvious to us (based on some coating theory we happened to know) that the faster the mesh rose from the ink, the larger the amount remaining on the mesh and therefore the lower the ink deposit. This prediction was quickly proven to be false by our colleagues in Swansea: the ink deposit depends only on the mesh, not on the ink, the press or the printing speed. A hurried search of the literature showed that our "faster = more" theory was correct up to a capillary number ($\eta V/\sigma$) around 1 and was constant after that. Because surface tension is so small and screen ink viscosities are so large, the capillary number in screen printing is always much larger than 1, so the amount around any given mesh is always a constant. This is one of the many reasons why screen printing, done scientifically, is so straightforward and immune to press settings, i.e. is (or should be) an in-control process. It is only the efforts of screen printers that convert an in-control, scientific process into an out-of-control "intuitive" process.

Splitting in real screen printing with a stencil is discussed in the Screen Print chapter, though the reader with a pencil and paper could quickly sketch out what must happen, based on the mesh-only diagram.

1.6 Squashing

The instincts of many printers is that higher pressure means more ink printed. This is true only in the early part of Walker-Fetsko if you are trying to print onto a rough paper surface which gives little ink contact under low pressure. If you have perfect contact plus porous paper then higher pressures might push a little more ink in, but the effects are modest at best and there are far better ways of getting a few % more ink on the substrate. On smooth substrates for packaging, there is no way that pressure can cause more ink to transfer.

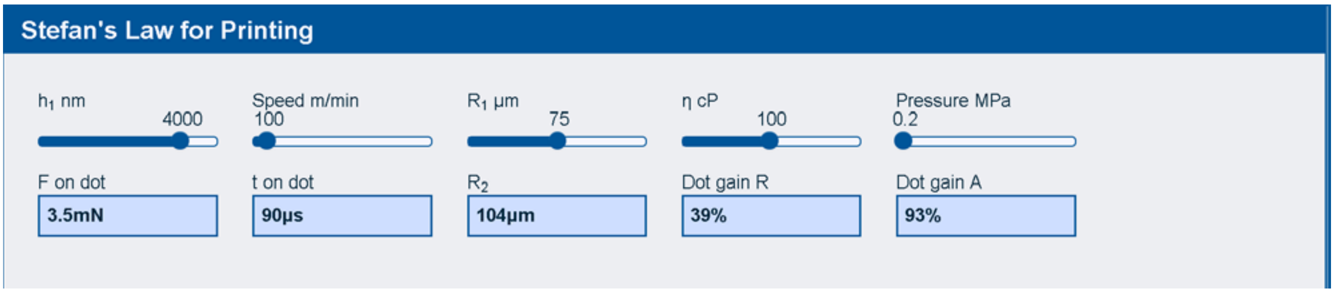
So the only thing that higher pressure regularly does is provide higher dot gain by squashing the ink. Those higher pressures not only strain your machine unnecessarily, they reduce the subtlety of tonal range by crowding 0-100% into, say, the 0-80% range of dots on your plate.

The classic equation to describe what happens when a dot of ink is squeezed comes from Stefan and is traditionally quoted as the time, t, for a sample of viscosity η and radius R_1 to be squashed from height h_1 to h_2 when a force F is applied. By a slight rearrangement, we can calculate h_2 after time t and by assuming that the dot is a cylinder, from the new h_2 we can calculate the new radius R_2 via $R_2=\sqrt{(h_1R_1/h_2)}$. So we just need h_2 which is given by:

$$\frac{1}{h_2^{\ 2}}=\frac{1}{h_1^{\ 2}}+\frac{16Ft}{3\pi R_1^{\ 4}\eta}$$

Equ. 1-11

What are F and t? If we know the nip pressure, P, typically 0.1-1MPa (readily measured with nip force sensors) then we know that this pressure is applied over the area of our dot, $\pi R_1^{\ 2}$, and F=P*Area. We also know that for printing speed V, $t=2R_1/V$. So it is not too hard, from a few typical inputs, to get an idea of the sort of dot gain expected via the squeeze. Of course there is a chain of suppositions behind this and the calculation cannot be accurate, but it gives a good idea of the relative changes on dot gain if you change the key printing parameters. The app allows us to explore these effects



Stefan's Law for Printing

h_1 nm 4000

Speed m/min 100

R_1 µm 75

η cP 100

Pressure MPa 0.2

F on dot 3.5mN

t on dot 90µs

R_2 104µm

Dot gain R 39%

Dot gain A 93%

App 1-12 https://www.stevenabbott.co.uk/practical-coatings/stefan.php

The straightforward trends tell us that halving pressure, doubling viscosity and doubling speed each cause a 1.4 decrease in dot gain. Offset's super-high viscosities are clearly an advantage compared to flexo. Because the dots are thin and being squeezed quickly, the shear rate will be high, so offset's viscosity advantage will not be quite as large as suggested by the ratio of their low-shear values. The absolute pressure for a small flexo dot, when the dot itself is being compressed, will be significantly smaller than the typical nip pressure, so, again, flexo isn't quite as disadvantaged as the headline figures suggest. In any event, the consistent message about printing with the lightest-possible pressure is once again confirmed.

In the formula, larger dots have a larger F (R^2) and t (R) giving an R^3 increase, outweighed by an R^4 decrease, giving a 1/R dependence. The net result is that halving the radius increases gain by 1.4. This confirms what we already know, that smaller dots give larger gain relative to their original size. Note that in normal dot gain curves you see rather small % dot gain for small dots. This is because the definition of % is different. If a 20% dot on the plate prints as a 30% dot that is recorded as a 10% dot gain. In the Stefan app, that is recorded as a 50% gain in radius.

The biggest effect is from halving the ink thickness which halves the gain. So going to a more pigmented ink requiring a smaller thickness gives significant dot gain benefits.

I need to stress that Stefan's theory was intended to apply to large, thick layers of liquid and not the very thin dots within compressible systems that are typical of printing. From what little literature I can find that methodically looks at dot gain in relevant systems, the trends are correct but the magnitudes are too large. One example that shows the effect of speed gives a trend of higher dot gain at lower speeds and higher pressures, but for a 50% dot, the gains changed by just a few % around a 25% average dot gain.

1.7 Two more core sets of principles

We now have a core set of principles about fluid flow that can be used in whichever printing processes are of interest to us. Before getting to those processes we must explore two other core areas. However perfect our printed dots, if they don't dry/cure correctly they will be worthless. And if they are beautifully printed and dried but do not adhere, they are also worthless.

As is the case in this chapter, each of those core areas suffers from misinformation and mythology. So it is worth the effort to find out what is really going on, especially given that the principles are not at all hard.

So keep going through the next two chapters before rushing to read about your favourite printing technique.

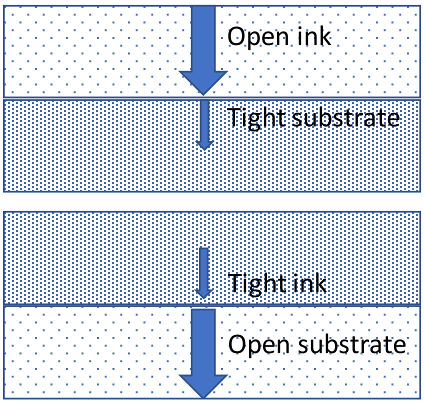
2 Drying and Curing

The only printing technique that requires no consideration of drying or curing is wax inkjet. The drops hit the substrate, cool and are solid. It is rather a shame that the drawbacks of the technique (e.g. the print feels "waxy" and adhesion is often questionable) have not made it a front-runner.

So, what is the science of going from liquid/viscous to solid? There are, broadly, three ways which we will discuss in turn: absorption; evaporation; curing

2.1 Absorption

Two techniques use absorption on a grand scale as the way to go from liquid to solid. Offset can do it because, being a process that relies on highly viscous pastes, it has a low % liquid that can readily be absorbed in the so-called cold-set method and is used on vast scale for paper printing where the absence of any drying step is a huge economic advantage. Inkjet is not naturally well-suited to absorption because the % solids is so low because, in turn, the inks have to be super-low viscosity. However, because water-based inks have to use lots of essentially non-volatile glycols, the technique has no real choice, so most practical inkjet media have either to be paper or to have a special coating that gives controlled absorption of the large amount of solvent, with the hope that most of the water will evaporate fast enough not to swell or cockle the substrate.



We can imagine two extreme cases where it would be possible to calculate how quickly the solvent moves into the substrate. The standard assumption seems to be that absorption is limited by the substrate - that flow out of the ink is easy and flow into the substrate is hard. The second case contains an idea that is familiar to those who dry inks conventionally, that it can be hard for all the solvent to escape from the ink. So if we wish to model the overall process we need two formulae, one for each of the sub-processes.

Even two formulae are not sufficient. We are supposed to be talking about drying, of the solvent leaving the ink. Yet if the substrate is sufficiently open, the pigment particles from the ink can also flow into the substrate. When we later discuss Kubelka-Munk theory about the influence of light-scattering from the substrate, we will see that ink absorption is a bad thing in terms of colour because the scattered white light turns blacks into greys, with similar effects on the CMY colours.

For inkjet this extra problem has led to the need for nanoporous substrates that quickly absorb the solvent, but leave most of the pigment particles at the surface. This also helps with the fact that absorption is (usually) isotropic so movement is equally fast in the horizontal and vertical directions, i.e. fast absorption would have meant large dot-gain if the substrate had not filtered out the pigment. The move towards nanopigments (which can travel to some extent through the nanomedia) for inkjet makes it harder to get this balance right.

Leaving aside the extra complexities of the ink particles, let us look at the two relevant formulae for the basic processes.

Focussing first on movement of fluid into the substrate, we can invoke a Darcy Law flow that depends on a pressure difference between the surface and the pores of the paper. The pressure can be both that of the pressure from the nip (for offset) and the "capillary pressure", the force that drives liquids into small capillaries. As we are concerned here with drying, the effect of the nip pressure can be ignored. Although there are a lot of formulae I could provide, in practice they are of little use to us because we seldom know the parameters to sufficient accuracy. So I will use one of the simpler Darcy formulae which tells us that at time, t, the volume of ink absorbed V(t), from a drop of Radius R, of surface tension σ and viscosity η into a substrate with pores of radius r, with a contact angle of θ, and a porous fraction f (i.e. the amount of open space) is given by:

$$V(t) = \pi R^2 f \sqrt{\frac{r\sigma \cos\theta t}{2\eta}}$$

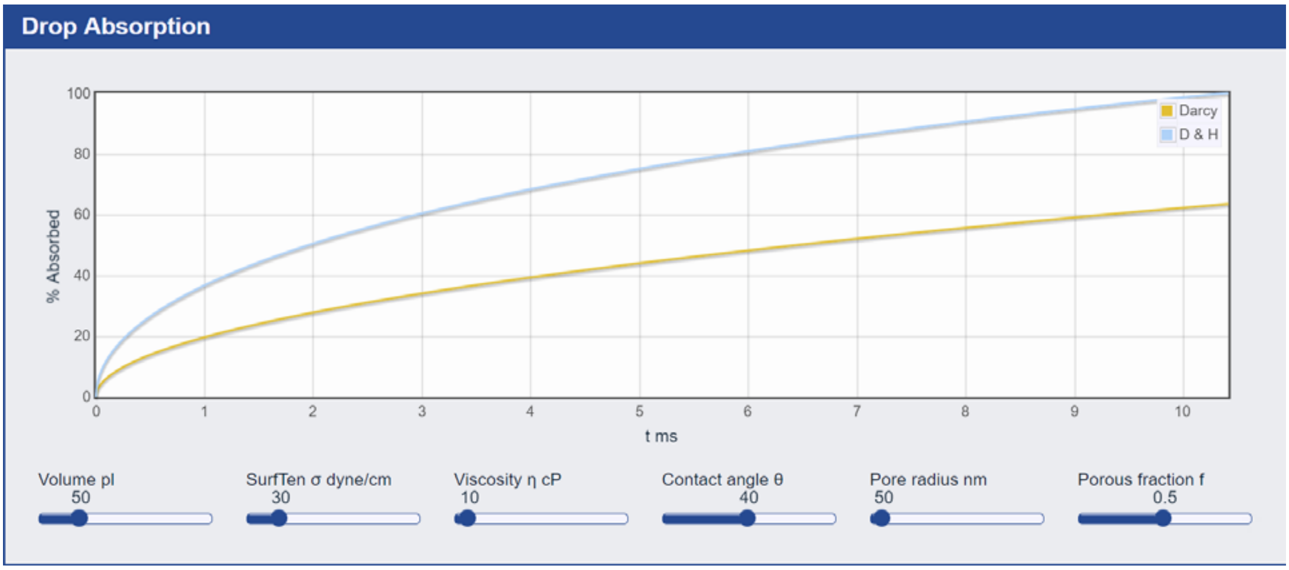
Equ. 2-1

The experimental data suggest that although popular, the Darcy equation is not suitable, at least for dye-based inkjet inks. A different model, Davis-Hocking, is found to be better:

$$V(t) = \frac{\pi R^2}{2}\left(2d(t) - \frac{d(t)^2}{D}\right) \quad d(t) = \sqrt{\frac{r\sigma \cos\theta t}{2\eta}}$$

Equ. 2-2

Here d(t) is the depth of the liquid after time t and D is the depth of the paraboloid which would accommodate the ink which depends on the porosity of the substrate.



Volume pl 50

SurfTen σ dyne/cm 30

Viscosity η cP 10

Contact angle θ 40

Pore radius nm 50

Porous fraction f 0.5

App 2-1 https://www.stevenabbott.co.uk/practical-coatings/drop-absorption.php

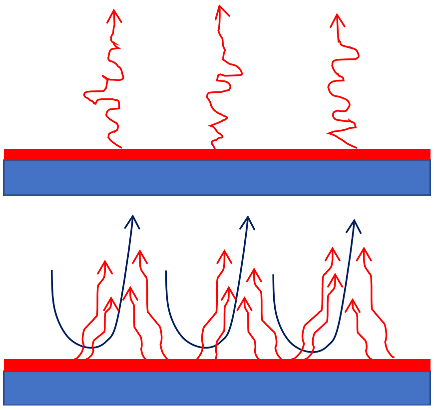
The app does both calculations. For dye-based inkjet inks the Davis-Hocking calculations seem to be realistic, and it is worth noting that the absorption into a typical nanoporous inkjet receiver is very fast. For pigmented systems the calculations are worthless because the pigments form a "filter cake" blockage and we get into the "tight ink" regime.

For this tight ink situation we can use a classic diffusion equation which tells us that the diffusion coefficient, D is calculated by fitting two parameters, A and B to the equation $D=Ae^{-s/B}$ where s is the % solids. As the % solids increases (as the solvent leaves the ink) the diffusion coefficient decreases. Although I could provide an app for this, we would be little better off because A and B are impossible to calculate from basic physical properties and if we could measure D independently we would already know how our ink behaved.

Although we have not ended up with a nice formula to solve all our absorptive drying issues, at least we have a clear set of principles which can be used for an informed debate about which factors to tweak to improve absorption. In terms of ink-related parameters, we can change σ by a factor of 2, viscosity might change by a factor of 2 and going from a contact angle of 0° to 60° likewise is a factor of 2 (cos(60)=0.5). As these factors appear inside the square root, these only change things by ~1.4. Changing the pore diameter and open volume of the substrate is overwhelmingly more important, but in general we prefer glossy substrates with small pores, so even here we cannot do much.

The reality is that absorptive drying is reluctantly accepted for conventional inkjet because there is no viable alternative other than very hot ovens to evaporate all the glycols. And for offset we either accept that cold-set is adequate for markets such as newsprint or we change the chemistry to provide some degree of cure over time (quick-set) or we use conventional drying (heat-set). This implies that the liquid solvents are in fact reactive chemicals. Although this concept obviously works well for UV curable "solvents" (as discussed later), it is only with highly viscous offset inks that we can have realistically curable "solvents" such as alkyd resins.

2.2 Evaporation



We need to get one fact straight right from the start: heat cannot dry, it can only make things hot. To dry a solvent means moving it from the vapour just above the ink to somewhere far away. It happens to be the case that the speed with which solvent vapour diffuses passively away from the surface (shown at the top) is many orders of magnitude too slow to be of any use to us. So we have to make sure that fresh air is sweeping away the vapour. Again, it happens to be the case that it is very difficult to get a large (laminar) flow of air close to the surface and the only effective way to displace the solvent vapour is via a turbulent flow that can reach very close to the surface (shown at the bottom). To get a turbulent flow requires plenty of air velocity.

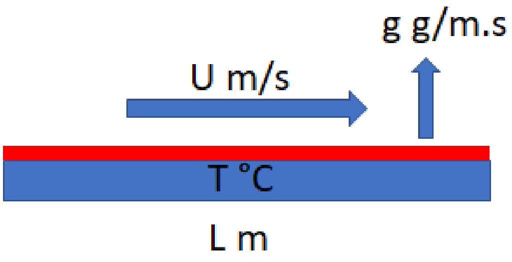
Unfortunately, because it takes energy to evaporate the solvent, the flow of air leads to a very cold ink, which means very few molecules in the vapour above the ink. So the drying requires plenty of heat to overcome the cooling effect.

That is the scientific way to put it. The more compelling way is to ask you to choose between drying your hair in front of a hot radiant heater or in a warm summer breeze. The former technique creates a very hot head and very slow drying. The summer breeze dries the hair quickly with a refreshing touch of coolness. Of course the optimal way is with a hair drier that delivers plenty of warm air or, if you are in a great hurry, plenty of hot air. Even if you don't touch your hair directly to know if it is dry, you can tell that the water has mostly gone when the evaporative cooling has reduced, so your head starts to feel hot.

My insistence on the idea that airflow is vital and that drying is about removing solvent molecules from the air above the ink comes because I have had to spend a lot of unnecessary time arguing against two myths about so-called IR driers.

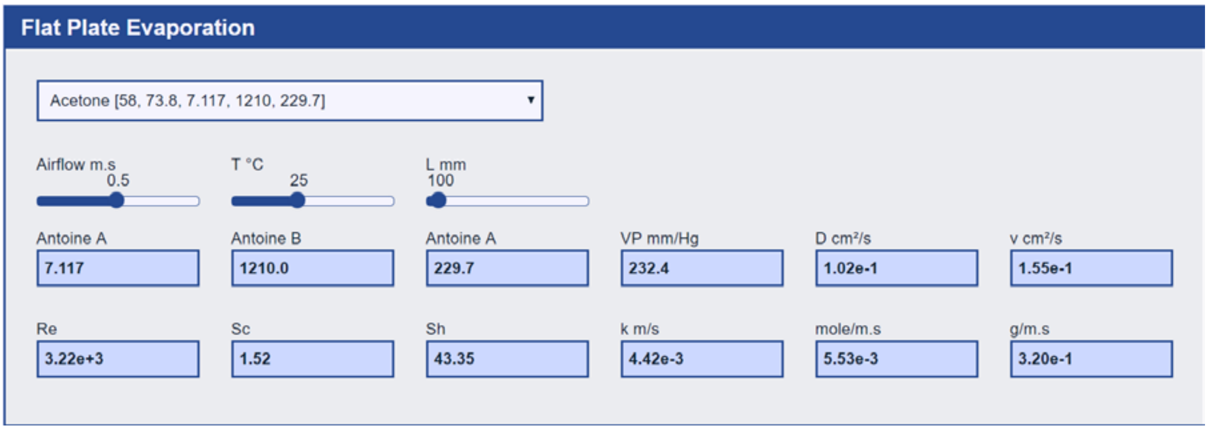
The first myth is that there is such a thing as an IR drier. There are plenty of excellent driers with lots of airflow where the primary heat comes from an IR source. I would have no problem with these if they were called IR-heated driers, to distinguish them from externally heated air driers. Boxes with lots of IR and little airflow are not driers, they are heaters.

The second myth says that IR driers are efficient because they "dry from the inside". This is nonsense because drying can only happen from the surface, irrespective of how heat is supplied. There is a companion myth about air-based driers that have the ability to supply air to the rear surface. These, too, are said to "dry from the inside" because heat comes through the ink from the rear surface. This is equally nonsense.



Knowing the basics and knowing the myths, we can now take a proper look at drying. A good starting point is an apparently simple model system: evaporation of a uniform layer of liquid on a flat plate with a constant stream of air passing over at a fixed temperature. The liquid is at T °C, air is flowing at U m/s and the plate is of length L. What we want to know is how many grams of solvent are evaporated over unit length and time. We come back to the types of dimensionless numbers discussed earlier, though this time we have the Reynolds, Schmidt and Sherwood numbers, which involve U, L, viscosity, diffusion and heat transfer. Then we need to know how much solvent vapour, at rest, would be sitting above the liquid. This is calculated from the Antoine Constants of the solvent and the temperature.

The text of the app discusses the constants and the calculation, which are not going to be repeated here. The various dimensionless numbers are provided for those who are interested. The important point is that even this simple model is complex enough, and real-world drying is significantly more complex.



Flat Plate Evaporation

Acetone [58, 73.8, 7.117, 1210, 229.7]

Airflow m.s 0.5

T °C 25

L mm 100

Antoine A	Antoine B	Antoine A	VP mm/Hg	D cm^2/s	v cm^2/s
7.117	1210.0	229.7	232.4	1.02e-1	1.55e-1

Re	Sc	Sh	k m/s	mole/m.s	g/m.s
3.22e+3	1.52	43.35	4.42e-3	5.53e-3	3.20e-1

App 2-2 https://www.stevenabbott.co.uk/practical-coatings/evaporate.php

The same app page includes a simpler model which actually tells you he drying time of your in. It uses (of course) the same theory but hides some of the complexities.

The point of the app is to provide a feel for the trade-offs inherent in all cases of evaporative drying. It is obvious that a volatile solvent such as acetone is much easier to dry than low volatility isophorone, but we often need to include some high boiling solvents to avoid drying so quickly that the ink has not had time to level, to produce a gloss finish, or to relax. A stressed ink is much more likely to crack or fail an adhesion test.

Similarly, although we know that we need both air and temperature, it is a difficult judgement about where the balance lies. While higher temperatures are more obviously wasteful in terms of energy, an efficient near-IR heater (to be discussed shortly) might be more efficient than an inefficient fan system. The app at least gives an idea of these trade-offs.

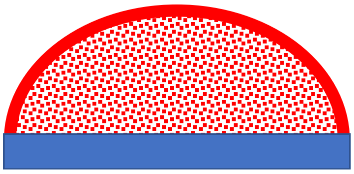
2.2.1 IR heaters

We need heat for two purposes. The first, as discussed, is to provide the energy to evaporate the solvent. The second is to raise the temperature of the substrate plus ink so that the first type of heat can do its job. In an all-air oven the hot air serves both functions, but the air doing the heating of the substrate is not doing much evaporation while the substrate is cool. So a nice compromise is a very short IR heater zone to bring the substrate up to temperature so that the (expensive) air in the main zone is doing what it does best - removing solvent.

A general IR heaters heats whatever comes beneath it, so it can work with any substrate and any solvent. The main practical problem is that if there is an emergency stop, the temperature of whatever is below the heater can rise catastrophically unless there is some extra safety system to shield the substrate and coating from the large, hot thermal mass.

A smart-sounding alternative is a heater in the near-IR region tuned to heat only a specific solvent - in general that means water. In theory, putting the heat only into the aqueous ink is more efficient than having to heat the substrate as well. Unfortunately, the thermal conductivity of inks and substrates are so large that in many cases the substrate gets heated too so there is little advantage in terms of efficiency. You can test this idea with the Flash Cure app discussed shortly - you will be very surprised how quickly heat migrates through these relatively thin layers. And because large amounts of hot air are required to remove the water vapour, the advantages of the near-IR system are further diluted. These near-IR systems *can* be very effective in some circumstances. But beware of any salesperson who tries to tell you that they are near-IR *driers,* that they heat *only* the ink or, worse still, that they avoid skinning because they *dry from the inside*. If they make these claims then tell them that they are contradicting the laws of physics.

2.2.2 Skinning and blistering



The diagram shows one of the key problems when we dry. The layer of ink on the outside of the drop has dried so fast that it has formed a skin. The liquid ink inside no longer has direct access to the air flow and solvent molecules have to diffuse through the skin - a much slower process than simply evaporating from the liquid surface. Skinning, therefore, slows down drying. Worse than skinning is what happens next. Because the drying is slow, the operators do the obvious thing which is to increase the oven temperature (with air or IR, it makes no difference). To their dismay they then find that the ink comes out full of blisters. When the oven temperature is higher than the boiling point of the solvent, and without the cooling effect of rapid evaporation, the solvent boils, explodes through the skin and creates a blister.

An alternative response to slow drying when skins are formed is to increase the amount of air through the drier in the hope that that will do something. Of course, this is a waste of energy because the speed limit is not due to removal of molecules from the ink, it is due to the diffusion barrier. Any extra air when you are in diffusion limited mode is a waste of energy. In fact, the only thing you *can* do to remove more solvent is to raise the temperature to has high as you can without causing a blister.

The standard IR mythology that IR "dries the ink from the inside" is, as I said, nonsense because ink can only dry from the outside. Perhaps an IR drier with low air flow will not form a skin, but that's because it is drying very inefficiently, without enough airflow to remove the solvent molecules from their slowly diffusing zone. No manufacturer of IR zones is going to say "Our ovens don't skin because they are inefficient" and salespeople prefer to say that their ovens "dry from the inside" because it sounds sciencey. In case you think I am being over-hard on IR oven suppliers, I can attest that I continue to hear such claims from some of them. I was recently guilty of arguing in public with an unfortunate scientist from one such company who had bravely taken over a more sales-like pitch from a marketing colleague who had been taken ill. The scientist was torn between defending the sales pitch and admitting that it was scientifically inaccurate.

There is some positive science to be taken from all this. First, anything which can keep the surface "open", such as a few % of a high-boiling solvent, can keep the ink in evaporative cooling mode for as long as possible, greatly helping the overall speed and, incidentally, avoiding "cooking" the ink or the substrate, whilst also producing an ink with higher gloss and less stress. Second, you can use two types of oven zone for the two different modes of drying. If your ink goes solid with plenty of trapped solvent then there is no point in keeping it in an oven with a vast air flow. A hot zone with a minimum air flow is all that is needed to remove the residual solvent.

For those who have catalytic systems (such as urethanes) it is very important to get most of the solvent out via evaporative drying before the catalyst kicks in and starts to slow the escape of the solvent. By tuning the system intelligently, the majority of the solvent will have evaporated before the ink temperature rises to kick off the final cure. By doing it this way, there is no need for a long, hot zone to remove residual solvent.

2.2.3 The multilayer ink problem

It has happened to many of us. We print a number of inks on top of each other. Because we are good scientists we have carefully checked that printing ink A on top of ink B does not cause any adhesion problems. Despite our checks, we find that adhesion is failing after printing, say, the 5th ink. The logical conclusion is that something about this 5th ink is causing the problem. So we make this 5th ink the 2nd ink and print the "good" 2nd ink last. To our astonishment, the adhesion fails again. It isn't the 5th ink itself, it is the fact that we have printed 5 inks. So what is going on?

When I first met this problem I had no idea what to do. Fortunately I knew a wiser, older scientist and he explained what was going on. His advice has proved useful in many other situations.

I won't bore you with diffusion science (you can go to my apps at https://www.stevenabbott.co.uk/practical-solubility/diffusionintro.php) and instead will explain what is happening. A good solvent must have some "bite" into the substrate to allow the polymers in the ink and the substrate to get tangled together for good adhesion. I had always assumed that when the solvent diffused quickly into the surface of the substrate it would also diffuse out quickly during drying. But that's not how diffusion works (there are "concentration dependent" effects discussed in the apps). Each layer of ink leaves a little of the solvent behind and it can diffuse to the interface with the substrate. After 3 or 4 or 5 prints there can be just enough residual solvent to weaken the interface and adhesion fails.

The fix for the problem is to be calmly aware of this "residual solvent" problem and to carefully tune your solvent blend and/or your drying process (heat and time to let the solvent diffuse out) so that any residue is below the adhesion failure tipping point. Although I admit that this is not an easy process, it is far superior to the more usual approach to the problem which I and many others have followed which is to rush around testing for the wrong problem.

To those who don't believe me, I suggest the following simple experiment. Print a few layers of ink as normal and, assuming you have a solvent with a distinctive odour, smell your dried samples. They will have no odour because you are good at drying inks. Now carefully wrap the samples in some thick polythene sheeting which happens to be a good barrier for most solvents other than things like toluene (in which case, wrap in PET). Leave the package for a week. Then quickly open the wrapping and cautiously (please think through the safety implications) sniff the contents. You will be surprised at the strong odour from the solvent that eventually escaped from deep within the print.

2.2.4 Don't forget to cool

Having emphasised the need for a hot, low air-flow zone to remove residual solvent, it is important to emphasise the need for cooling before the next process - especially it the next process is the final winding of the web into a large roll. I happen to have a computer model for what happens inside a large, hot wound roll in terms of temperature and in terms of the quality of the final wound roll (https://www.stevenabbott.co.uk/abbottapps/RDC/index.html). The model makes it clear that it is far, far better to have added the technology to bring the web to ambient temperature before winding. This can be done via cool air in an oven zone, though this is relatively inefficient in terms of space and energy. Far better is a chill roller with a large wrap angle. I mentioned in the context of near-IR heaters that thermal conductivities are rather high in terms of typical ink and web thicknesses. That is a problem for near-IR and an advantage for chill rollers. They can be super-efficient, provided there is not a pesky layer of air between the roller and the web. At high speeds such a layer is spontaneously dragged in and it is rather a challenge to get the good contact for high-speed removal of the heat. You can find an app addressing this issue on my Practical Webhandling site, https://www.stevenabbott.co.uk/abbottapps/ALC/index.html.

The need for massive banks of chill rolls is clearest in heat-set offset where the high boiling solvent (remember, the solvent cannot be allowed to evaporate significantly in the inking roller train) is removed in a high temperature oven at super high speeds.

2.2.5 Aqueous inks

It seems obvious that in this greener age we should all be using aqueous inks. Yet solvent-based inks maintain a large market share. There are two reasons for this. The first is that most printed articles have a chance of being exposed to water so the ink should be water-insoluble. The standard way to achieve that is via an emulsion-based ink. The polymer in emulsion form is perfectly happy in the aqueous environment, then coalesces (film forms) once the water has been removed. It is then not redissolvable. This is a problem on the press. With solvent-based inks the ink itself (or a quick wipe with solvent) can redissolve a dried-in ink. This is not possible for most water-based inks. The second reason is that it takes around twice as much energy to evaporate a gram of water than a gram of a typical solvent. This is thanks to the high "latent heat of evaporation" of water. So, as has often been noted, the environmental footprint of an aqueous ink may not be as wonderful as we might assume.

The other big problem of aqueous inks are those involving inkjet. A tiny amount of evaporation of water from the ink on the inkjet nozzle can lead to a tiny amount of insoluble ink stuck on the edge of the nozzle, which in turn leads to drops flying off in the wrong direction and, eventually, a blocked nozzle. So such inks must contain plenty of glycols. Glycols can be so water-loving that they positively attracts water to themselves from the air, allowing inkjet nozzles to stay open for longer. The problem with these glycols is that, of necessity, they evaporate even slower than water. So when these inks are "dried" by absorption, although the absorbed water might disappear over time, the glycols tend to remain causing, among other problems, dot gain and strange "contact" effects when the print is framed behind glass. When such inks are printed onto non-absorbent substrates (e.g. for printed electronics) then it needs a lot of temperature and modest airflow to persuade the glycols to evaporate.

2.3 Curing

2.3.1 Chemical curing

Common house paints are a classic example of an ink-like composition that dries (partially) by curing. The paints contain alkyd resins which polymerise when catalysed, conveniently, by oxygen in the air. Readers who are chemists will recognise that this is a simple chain reaction process of the (usually) unsaturated fatty acid side-chains on a polyester resin. So no diagram is required. For readers who are not chemists, such a diagram would be of little use.

It is no surprise, therefore, that many offset inks, the so-called quick-set inks, contain plenty of viscous alkyds (remember, offset inks are high viscosity) which, unlike cold-set, harden with time to avoid the annoying ink rub-off onto one's fingers during reading The trick with these inks is that they contain plenty of lower viscosity oils that absorb into the substrate, leaving the alkyds in a concentrated state where they can react relatively quickly (2-30min) with the oxygen to "set", i.e. to not transfer too badly to the reverse side of the print when they come into contact. On the press, the alkyd is sufficiently dilute that little polymerisation takes place. Like real paints, the proper "drying" to a robust finish takes from hours to days. So quick-set inks are a handy compromise for systems that do not require a press with attached ovens. Of course you can add IR heating zones to encourage the alkyd reaction, making it closer to a Heat Set machine. Such a compromise system might be called ultra-quick-set, with the aim of getting the whole process complete on press so that the end product is ready to be shipped fully cured, not requiring the hours or days of quick-set. The difference is one of degree rather than kind, with different alkyds and, perhaps, added heat-activated catalysts (peroxides) to help.

Another popular curing system creates urethane inks which can be formulated for just about any desirable property, though mostly they are thought of as tough and flexible. The ink contains a mix of diols and isocyanates which react to form a polyurethane. The challenge for the printer is to get the rate of reaction right. You can buy one-pot inks that are stable on press but, unsurprisingly, take a long time to cure using, perhaps surprisingly, small amounts of water from the air or the substrate to cure. Alternatively you can use a two-pot ink which is mixed just before printing and choose a level of, and type of, catalyst to give you the desired speed of cure at high temperature with the necessary long open time on press. This necessary compromise means that many urethane inks are rather like quick-set inks - not ready for end-use until hours or days have passed. This makes it challenging to test if the ink is OK. If you find after a day that a large print run has not cured because someone forgot to add the right amount of catalyst, it is too late to do anything other than throw the batch in the skip.

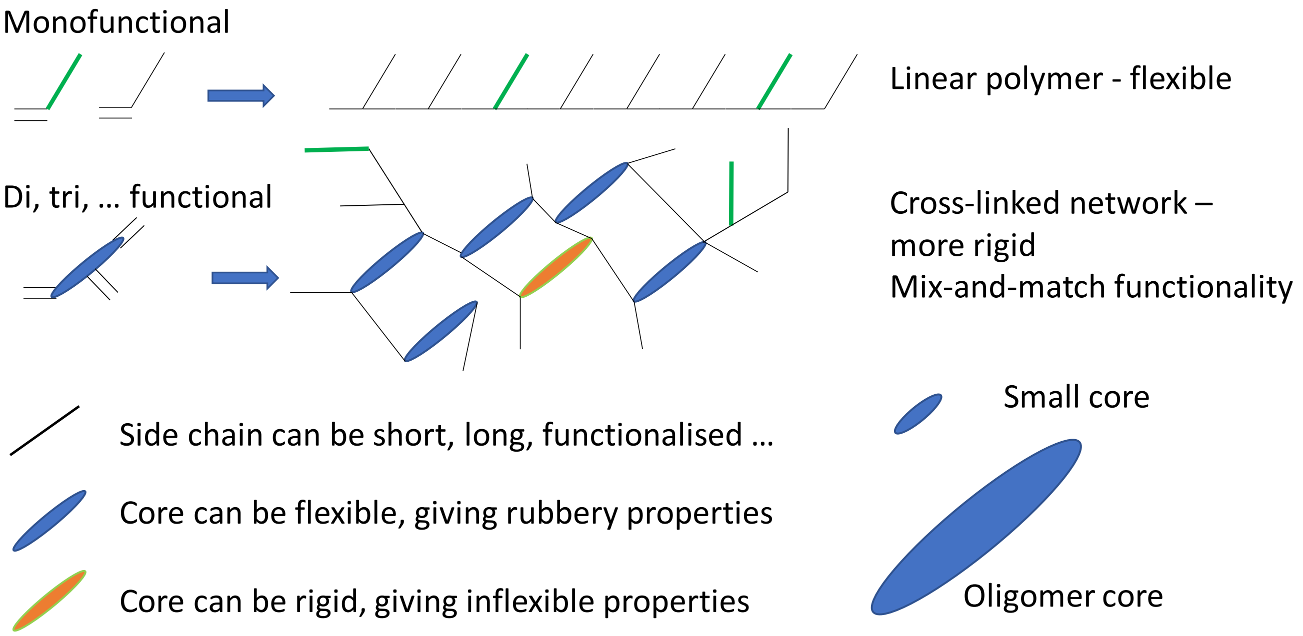
Although chemical curing systems have many desirable features, despite decades of work it does not seem to be possible to get a system that is sufficiently stable on the press whilst being fully curable with a short pass through a typical press oven.

Which is why, for all their many problems, UV (and e-beam) inks continue to grow in popularity.

2.3.2 UV and e-beam

For convenience I will use "UV cure" to describe both UV and e-beam as the basic principles are identical. There are just two differences between the techniques. The first is that e-beam does not require a photoinitiator so the resulting print contains fewer potentially undesirable residual chemicals. The second is that e-beam is massively more complex (the machines produce vast amounts of x-rays so need to be properly shielded) and expensive and can only be justified by those who need the monstrous speeds of which they are capable.

A large chunk of UV chemistry can be accommodated in a single diagram:



The sub-units can be monomers that polymerise to long chains with variations in the side chain length or functionality depending on the proportions of different monomers added to the mix. More normally, the monomers are mixed with dimers, trimers etc. to give a cross-linked network that is tougher and, generally, more rigid. The cores of the multi-functional acrylates can be varied to be flexible or rigid and can be small-molecules or larger oligomers or, even, polymers.

The reaction is started via photoinitiators which generate free radicals which react with an acrylate group which in turn carries on reacting with further acrylates till they get stopped by some undesired reaction or the system runs out of accessible acrylates - points discussed shortly.

The freedom to mix and match such a wide variety of properties is both an advantage and a problem. The advantage is that in principle you can create just about any desired overall package of properties. The problem is that it is very easy to get lost in acrylate space, not knowing quite how to get that extra adhesion without being too brittle and without a strong residual odour and without using an acrylate with an unacceptable health and safety profile. Similarly, there are a variety of photoinitiators with different light absorbing and reactivity characteristics, which also have different residual chemicals and odours.

The last points have resulted in the Nestlé list of acceptable acrylates and initiators. This came about from a perceived health risk of some "contaminants" in some printed food packaging, with Nestlé taking the initiative to construct a list of what might be acceptable and with tests for levels of "extractables" - i.e. residual acrylates and initiator residues. It is assumed that if you use ingredients on the Nestlé list and if you pass the standard extractables test then your product is safe to use.

Because we tend to want the fastest possible cure with the fewest number of expensive UV photons, the logical thing to do is take note of a surprising law of polymer networks: you don't have significant cure till everything has at least one connection to everything else - which is the gel point. If you measure, say, viscosity versus UV dose you find that nothing much happens for a while (even though the acrylates are reacting happily) then suddenly the viscosity starts to increase before reaching a plateau. The sharpest response curve is achieved with oligomeric (or polymeric) multifunctional acrylates because they need the fewest links before network connection and the network becomes complete very quickly. When there are n functional groups per molecule this gives, by Flory's definition, a functionality f=2n and the gel point occurs when conversion reaches 1/(f-1). For a monomer, f=2, i.e. you don't get a gel till 1/(2-1) = 1 = 100% conversion. For a dimer, f=4 and you need 33% conversion before anything significant happens. For a 4-mer, f=8 and the gel point is at 14%.

There is, therefore, a strong temptation to use lots of multifunctional acrylates. The downside is that the resulting network can be very brittle and, as we shall see later, this is not good news for strong adhesion.

2.3.3 Oxygen, temperature and photoinitiators

The biggest single enemy of acrylate curing systems is oxygen. The oxygen molecules react with the polymerising radicals to form relatively stable peroxides which can only slowly, if at all, continue polymerisation. The most satisfactory fix for the problem is to add an inert atmosphere of CO2 or N2, or for the curing to take place against a roller or cover material. Because a little oxygen provides a lot of inhibition, the inerting system needs to be quite sophisticated to get down to a low level of O2 whilst not consuming too much expensive gas. For some demanding applications (e.g. food or baby-product) using some printing techniques (offset seems to be an exception), the fight against oxygen is a losing one and including an inerting system by design, right from the start, is ultimately the best strategy. It is so much easier to formulate an excellent acrylate system when O2 is no longer an interfering factor.

I have listened to talks for at least two decades describing UV curing systems with low sensitivity to oxygen. All I will say is that if they were good, we would all be using them, and we're not. The thiolenes have been promoted as being exceptionally good, but there is some reluctance to use anything containing a (potentially odorous) thiol, and there simply aren't the wide range of cost-effective raw materials taken for granted with conventional acrylates.

One alternative is to use raw UV power. If you blast the system so it floods with radicals then the cure can take place before fresh O2 can diffuse in to interrupt it. In the early days of UV LED systems, although it was easy to provide the same *total* amount of power as a conventional Hg system, the cures would be much worse because the power was delivered over a longer time, allowing more access to O2. Modern LED systems are now more intense so can provide the necessary raw power if other aspects of the system are suitably optimised. Raw power is not cheap - another reason for considering an inerting system right from the start.

The other approach is smart ink design. With some compromises, it is possible to get full cure, especially with very thin offset inks which contain relatively little UV material because of the high pigment loadings possible for these high-viscosity inks. Those who print with lower viscosity inks often have to use lots of mono-functional acrylates to keep the viscosity low, and these are much harder to make non-extractable. An acrylate with 3+ functional groups will not fully cure (discussed next) but if at least one group has reacted then it is not extractable. And polymeric acrylates are intrinsically less extractable, though their generally higher viscosities exclude them from most inkjet or flexo formulations.

The other effect is not so obvious. With infra-red (IR) spectrometers it is easy to measure the amount of unreacted acrylate after cure (FTIR, (confocal) Raman and Near-IR can all be used). Suppose you have a good formulation cured at room temperature with plenty of power in an inert environment. It can be a shock to find that at least 25% of the acrylates are unreacted. You increase the initiator level, the UV power, the exposure time - yet nothing happens.

Now repeat the experiment at, say 50°C. The amount of uncured acrylate might be down to 10%. Why? It is nothing to do with *reactivity*. Instead it is a result of increased *mobility*. At lower temperature, as the cure proceeds, the matrix itself gets more rigid, so it becomes harder for a polymerising chain to move to an unreacted acrylate group. At higher temperature, the matrix remains more mobile so the polymerising chains can carry on finding unreacted groups. This will be discussed more precisely below.

Because we can generally get all the crosslinking we need, using a higher temperature to reduce the amount of uncured acrylate may seem unnecessary. But remember that acrylates are fairly unpleasant molecules so having any of them sitting around unreacted is a potential cause for concern in your end product. Problems include odour, health and safety, and the unintended ageing consequences of the oxygenated molecules produced when excess acrylates slowly react with air and moisture. I, and many others, have had excellent formulations that fell apart after a few months equivalent of ageing tests with cycles of heat and humidity.

As mentioned above, a distinct advantage of e-beam is that it does not require photoinitiators - the electrons do the job themselves. For UV curing, the temptation is to throw in a lot of initiator "just in case". Experience shows that reducing initiator levels, while focussing on inerting and temperature, reduces costs, improves cure and reduces the levels of sometimes unpleasant initiator residues.

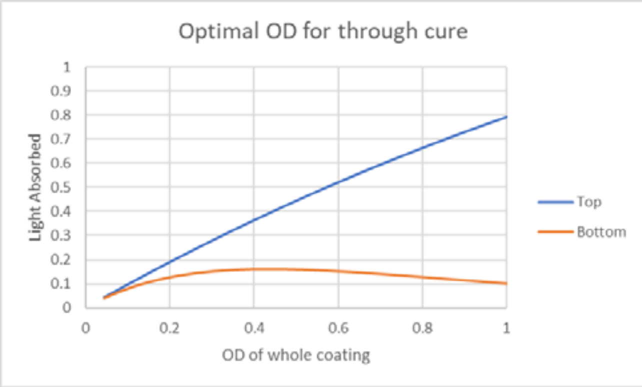
One of the great mysteries of photoinitiators, observed by many of us, is that when we objectively know that very little light is reaching them because of absorption by other factors (UV stabilizers, ink pigments, curing through absorbing base film), we can still get good cure - but only with a system designed correctly and with inerting and thermal effects under control. This is another way of saying that a very little initiator goes a long way if the rest of the process is under control.

The science of the speed of cure versus photoinitiator level shows that there is decreasing value in terms of speed as the level goes up. The app discussed next gives a general idea of the dependencies of speed and degree of cure on photoinitiator level and on the idealised functionality of the molecules.

My advice, therefore, is to build inerting and higher-temperature cure into your system from the start unless you have special reasons (such as high-viscosity offset inks) to be certain of a low level of extractables. This up-front investment saves you from many problems during the life of your production setup.

Although it is said that it is hard to devise a system that can efficiently inert at high speed, it is equally clear that smart designs based on good fluid dynamics can deliver a good balance of simplicity, speed and low N2 consumption.

2.3.4 Modelling UV curing



The first issue to model is how much photoinitiator is optimal. Too little and UV light is wasted, but you get good through-cure. Too much and, without the self-bleaching trick described below, the UV is absorbed in the top layers with none getting down to the bottom layer. It is often stated that the optimal level is an optical density (OD) of 0.43. It was fascinating to trace this statement back to its source. Repeating the calculation which is simply how much light gets absorbed at the surface compared to the bottom of the layer, I got the graph shown in the image. The claim does have merit, but the optimality is very broad around OD=0.4, which means, as has been observed by many formulators, that the amount of photoinitiator is a surprisingly non-critical element of a formulation. So given that it is expensive and leaves potentially unpleasant residues, less is better than more, unless you need to compensate for O2 inhibition at the surface which consumes initiator without crosslinking. If the initiator is self-bleaching (i.e. it doesn't absorb once it has reacted) then the question doesn't arise.

Some photoinitiators (such as the phosphine oxides) "self bleach", i.e. they lose their absorption once they react. This means that you can use a higher concentration at the start if you need it for other reasons.

The UV curing process is complicated and random, which makes it hard to model - unless you have an algorithm that is naturally at home with complicated, random processes. One such algorithm is the Gillespie SSA, Stochastic Simulation Algorithm which is a Monte Carlo method - a fancy way of saying that it explicitly uses random numbers as part of the algorithm. Thanks to a paper that gave a detailed explanation of applying the SSA to UV curing[9], I was able to create an app that allows the exploration of the key parameters involved. The screen shot gives an idea of what it can do and readers who are interested can learn more details by visiting the app.

---

[9] Gökçen A. Altun-Çiftçioğlu, Ayşegül Ersoy-Meriçboyu and Clifford L. Henderson, *Stochastic modeling and simulation of photopolymerization process*, Polymer Engineering & Science, 51, 2011, 1710-1719

% Initiator 1

UV Power (Rel) 4

Acrylate Functionality 3

N Molecules 10000

t_{Max} s 5

k_P 0.3

k_T 1.8

cyFact 1

Data

```
% Conv=38, tGel=0.39s  % Prop:Term:Cyc=32.3: 1.4: 66.3  MaxL=1015  tCalc=378ms
```

App 2-3 https://www.stevenabbott.co.uk/practical-coatings/uv-cure.php

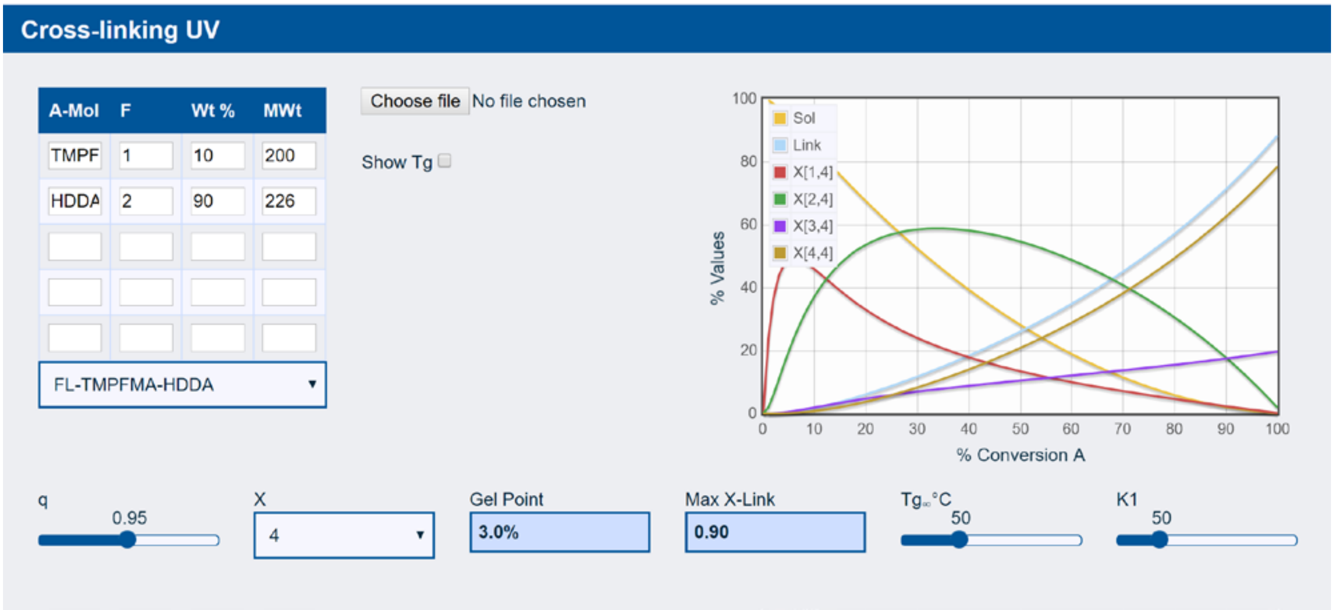
Briefly, there are tradeoffs between the % initiator (which you want to be low for cost and low extractables), UV power, the propagation rate, k_p and the termination rate k_t, along with a factor (cyFact) that controls how much of the reaction turns back on itself rather than extending the polymer network.

The screen shot shows the typical relatively low % conversion over long exposures, even with plenty of spare initiator. It simply becomes impossible for a new radical to move to find an unreacted monomer. With a monomer (try this in the app with Acrylate Functionality set to 1), conversion can be complete, but who wants an uncrosslinked formulation? With an unrealistic Functionality of 6, conversion plateaus at 8%, but the resulting formulation may well be too brittle to be useful.

The text box shows the final % conversion, the time when the system gels (i.e. there is on average one polymer-polymer link), % propagation versus termination of radicals versus cyclic (intramolecular) reactions and the maximum length of a polymer chain.

While the app can't readily be applied to any given formulation problem, my experience is that playing with it helps build an understanding that I, for one, never managed by intuition or by reading the literature.

Another view of the curing process can be found by looking at the fates of the different functionalities within the formulation. The methodology behind this was developed in 1976 by Miller and Macosko[10] but is astonishingly little-known. I only discovered it by accident in 2018. Even then I could not properly understand it and for UV, before I could implement it I had to find the only paper[11] that seems to have been published about using it. I am grateful to Prof Meichsner of HS-Esslingen for sending me the paper that could not be found in the journal's archive.



Cross-linking UV

A-Mol	F	Wt %	MWt
TMPF	1	10	200
HDDA	2	90	226

FL-TMPFMA-HDDA

Choose file No file chosen

Show Tg ☐

q 0.95

X 4

Gel Point 3.0%

Max X-Link 0.90

Tg∞°C 50

K1 50

App 2-4 https://www.stevenabbott.co.uk/practical-coatings/x-link-uv.php

As you can see from the screen shot, the app is rather complex. But it is describing a complex process. Most of the lines in the graph are showing the X[m,f] values. A di-functional molecule such as HDDA can be linked to 4 chains (each double bond has a chain coming in to it and one coming out.) X[1,4] shows how many HDDA molecules are linked to just 1 chain. This rises quickly at the start then tails off as X[2,4] takes over then X[3,4] and finally X[4,4]. The q value of 0.95 means that 95% of radicals manage to continue to grow, meaning 5% terminate somehow. Which is why X[4,4] never reaches 1.

This all looks very nice till we realise that the x-axis is degree of conversion and few of our formulations ever reach 100%. So if this example had stopped at 50% conversion (because the temperature was too low), the mouse readout tells us that there are only 20% of X[4,4], with 60% X[2,4]. Once you get used to reading the graphs, this sort of information tells you a lot about what's inside your formulation.

The phrase "because the temperature was too low" links to the idea, mentioned above, that once the network is too restrictive, curing slows down dramatically. This can be rendered more precise theoretically and experimentally. The rate of

---

[10] Douglas R. Miller and Christopher W. Macosko, *A New Derivation of Post Gel Properties of Network Polymers*, Macromolecules, 9, 1976, 206-211

[11] Georg Meichsner et al, *Netzwerksgrößen UV-gehärteter Lackfilme*, Farbe & Lack 103, 45-50, 1997

cure decreases once the network has reached its glass transition temperature, Tg. The Tg of a lightly crosslinked network might be 30°C, so curing at that temperature will stop at that light network. A more tightly crosslinked network might have a Tg of 50°C, so if you cure at that temperature you will attain that degree of crosslinking. Because Tg values can readily be measured using (say) differential scanning calorimetry or dynamic mechanical analysis, it has been possible to confirm that this is the case.

So an extra feature of the app is a Show Tg option which makes a plausible estimate of the Tg as a function of % conversion and, therefore, tells you what sort of temperature you would need to have in order to reach that degree of cure. This is only a guide. But I can assure you that I wish that my colleagues and I had had access to such an app when we were trying to optimize difficult UV formulations. The app is far from perfect but is much better than the hand-waving we (and the industry in general) had to use instead. The theory behind the Tg calculations is described in a further app: https://www.stevenabbott.co.uk/practical-coatings/x-link-tg.php

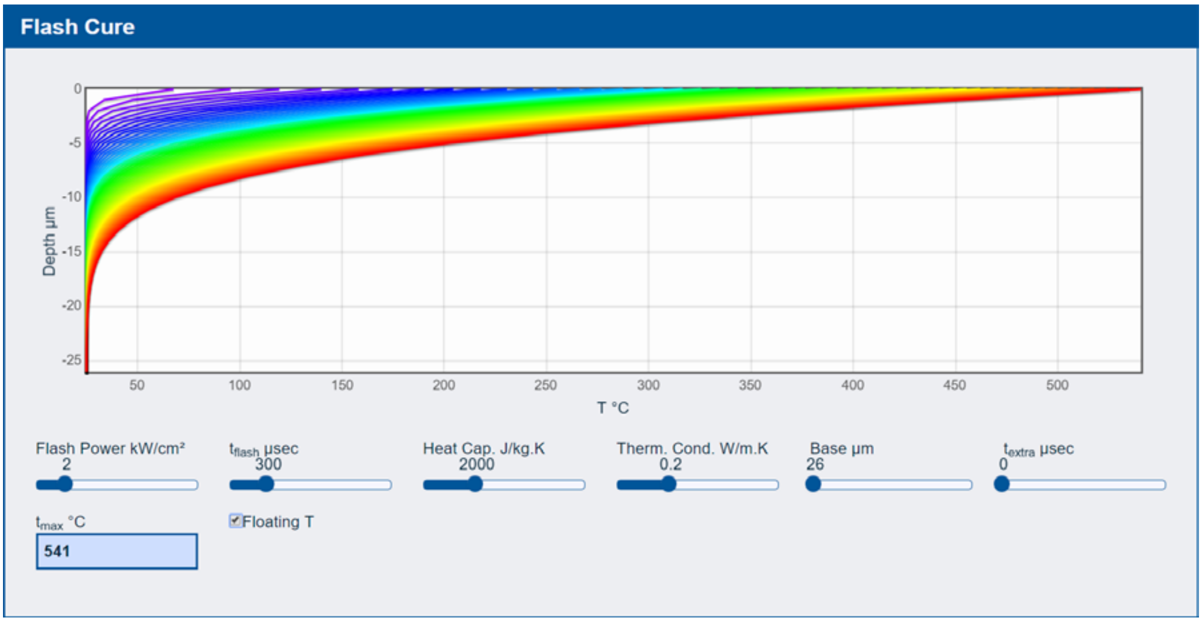
2.3.5 Flash curing

If you can get a lot of thermal energy very rapidly into the ink itself then it might be possible to thermally "cure" the system via some sort of reaction or via sintering/melting. Putting in a lot of energy slowly into the ink itself is a disaster because thermal conductivities are, in practice, so large that even in the millisecond time-scale a lot of heat will flow into the substrate, causing the problems (melting, distortion ...) that the flash process is designed to cure.

The best way uses a xenon flash (or maybe an excimer laser) with a system where the ink absorbs all the energy. The so-called "photonic curing" of silver inks for printed electronics is a good example of this, though because all curing with light is photonic, the name is rather unhelpful. A xenon flash can render a low-conductivity print of individual silver flakes/particles into an instantly high-conductivity ink with the silver particles forming a continuous path.

An alternative, far cheaper way, is to use a near-IR system. This isn't quite the same because instead of relying on the ink to absorb all the energy, it relies on the fact that near-IR wavelengths are not absorbed by most transparent substrates, so there can be a large differential heating of inks such as silver.

The app allows you to explore what power densities are required from your flash source, over what sort of timescale.

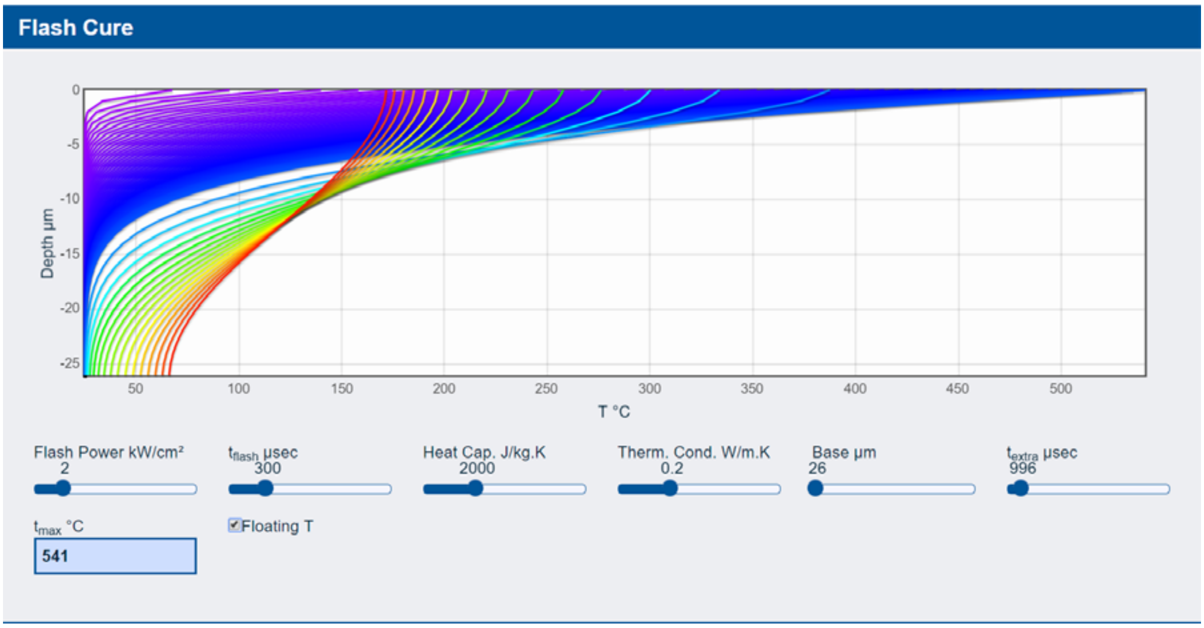


Flash Power kW/cm² 2
t_{flash} μsec 300
Heat Cap. J/kg.K 2000
Therm. Cond. W/m.K 0.2
Base μm 26
t_{extra} μsec 0
t_{max} °C 541
☑ Floating T

App 2-5 https://www.stevenabbott.co.uk/practical-coatings/flash-cure.php

In this example a 2kW/cm² pulse for 300μs raises the top few microns to 500°C, and below 10μm from the surface the temperature does not exceed 70°C. It takes a moment to work out what the graph means. On the Y-axis we have the depth into the substrate (the coating is always assumed to be 1μm thick) and temperature is on the X-axis. The rainbow colour coding shows how temperature changes over the time of the pulse, with blue being near the start and red at the end.

The question arises as to what happens in the time after the pulse.

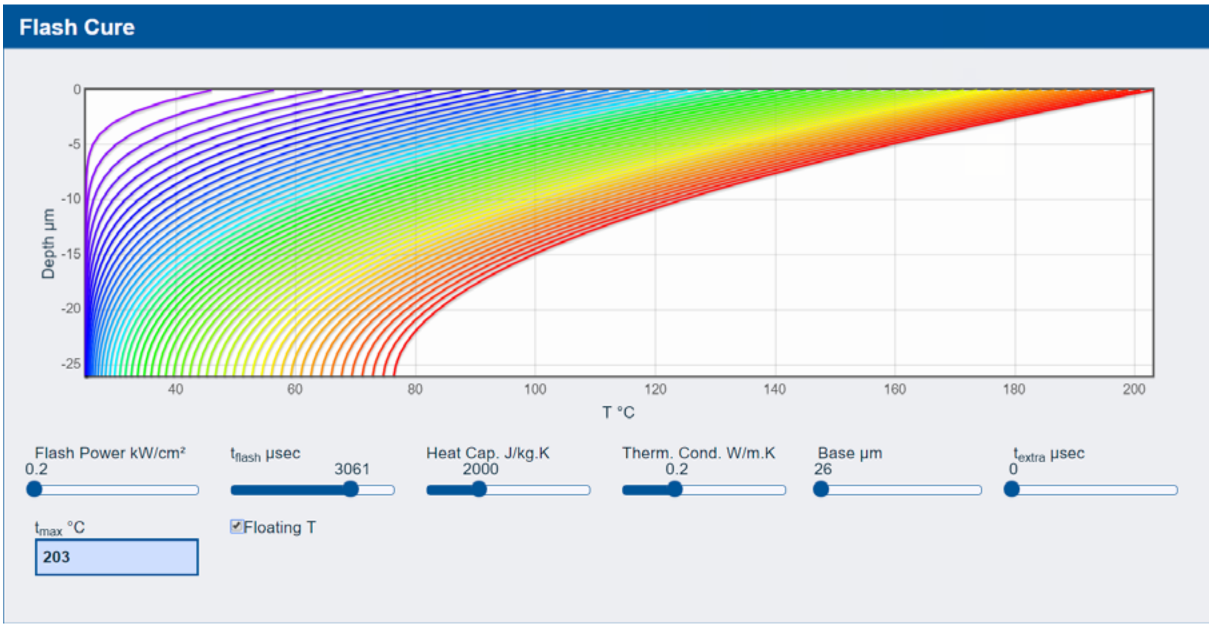


Flash Power kW/cm² 2
t_{flash} μsec 300
Heat Cap. J/kg.K 2000
Therm. Cond. W/m.K 0.2
Base μm 26
t_{extra} μsec 996
t_{max} °C 541
☑ Floating T

Figure 2-1 After an extra ms the temperature ranges from 75-150°C

The figure answers that question. In this example the substrate is "floating" so the rear surface temperature is not pinned at 25°C. The rainbow colouring shows the rapid decrease of temperature at the surface plus the more gentle rise in temperature on the other side of the substrate. If the "floating" option is turned off (try this yourself), the surface temperature is hardly affected, so it makes no practical difference after this timescale. Clearly in longer timescales the removal of the delivered energy needs to be thought through.

The power density is typical of a xenon flash. For a near-IR system, my understanding is that the density is an order of magnitude lower.



Flash Power kW/cm² 0.2

t_{flash} μsec 3061

Heat Cap. J/kg.K 2000

Therm. Cond. W/m.K 0.2

Base μm 26

t_{extra} μsec 0

t_{max} °C 203

☑ Floating T

Figure 2-2 With near-IR power densities, it is much harder to get a flash cure at the top without roasting the substrate.

The root cause of all flash cure problems is the relatively high thermal conductivity of most polymer substrates. If (not shown) the same near-IR flash is used on paper with a thermal conductivity of 0.05 W/m.K, then the top temperature reaches 350°C with far less heat within the substrate.

Remember, even the "failed" near-IR system is 2MW/m², a rather awesome power density. So getting really good flash cure over a large area requires a lot of very expensive xenon flash bulbs at 20MW/m².

2.4 Summary

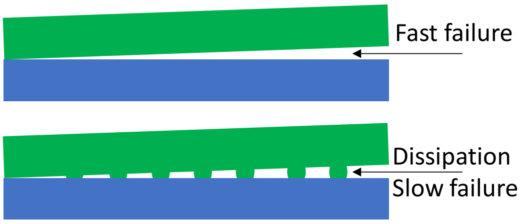
All three methods for curing inks are used extensively in printing and they each have their advantages and disadvantages. The key to whichever method you choose is to avoid the myths that are all too common and to focus on the few principles that can allow you to optimise your method.

3 Adhesion and Cohesion

We all need our inks to stick to their substrates, so we should all want to know the real science behind adhesion. Unfortunately, the science of adhesion got hijacked by irrelevant considerations of surface energy, so a large part of the industry has devoted considerable effort to irrelevant issues, while being largely unaware of what really matters.

What really matters is polymer physics, which is unfortunate because for most of us, polymer physics is hard to grasp. Thanks to a tip from a physicist friend, I found a way for me, a chemist, to begin to understand what is going on, and the good news is that it is all rather easy. For those who want a full guide to adhesion science, I have written a book[12] (not, unfortunately, a free eBook) and a set of apps https://www.stevenabbott.co.uk/practical-adhesion/ (which are free) which takes the reader through everything that a practical formulator will need. The short summary in this chapter should give those in the printing world a good grounding in what is going on, how to improve adhesion and, perhaps most importantly, how to avoid trying too hard. It is not often that you find good reasons for being more relaxed in your formulations, and this bit of science gives you freedom to optimise other aspects rather than putting all your formulation energy into trying to get a bit more adhesion.

The key to adhesion is dissipation of crack energy. What does this mean? For adhesion to fail, a crack must start (this might be deliberate or from an imperfection in the system) and then must run along the failure line. It takes energy to do this, and anything which absorbs (dissipates) this energy will make it less likely that the crack will propagate and, therefore, increase adhesion.



Take a thin sheet of glass which is held together by strong silicate bonds. Now put a tiny crack into the glass. It is trivial for the crack to travel along (top image) because the rigid glass has no way to absorb the crack energy. Now take a thin sheet of chewing gum which is *much* weaker than glass. Put a tiny crack in the chewing gum and try to make the crack propagate. Of course it cannot propagate, because as you pull on the chewing gum it stretches, (bottom image) dissipating the energy you are trying to put into the crack. Although the analogy with chewing gum is deliberately extreme, this dissipation is the reason why the adhesion from household adhesive tapes is so strong, even though they are based on weak polymers. Such tapes are called PSA (Pressure Sensitive Adhesives) and for good adhesion in printing, it often helps to be more flexible (closer to dissipative PSA) then rigid (closer to brittle glass).

---

[12] Steven Abbott, *Adhesion Science: Principles and Practice*, DEStech Publications, 2015

I am not suggesting that all inks should have the consistency of chewing gum. Instead I am saying that in terms of ink adhesion, formulating for dissipation can often be more effective than formulating for perceived strength.

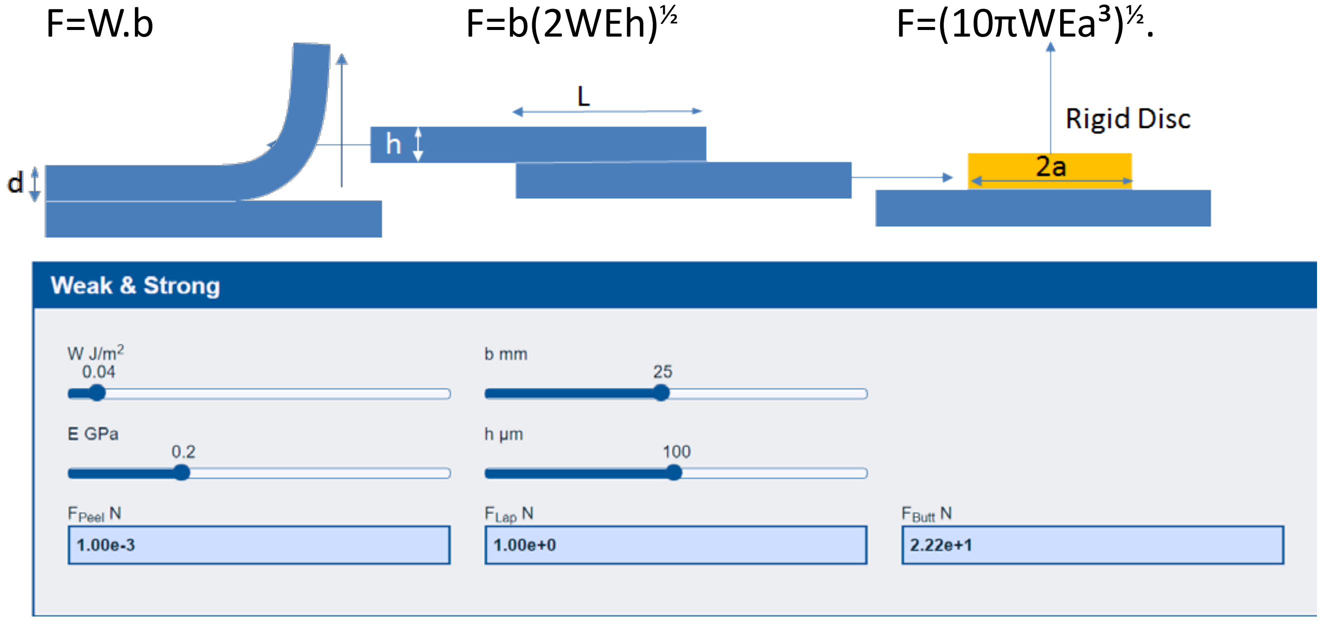
Which brings us to a problem. Not only do we want our ink to stick to the substrate, we also want our ink to stick to itself so that it is not easily damaged and so that components are not easily extracted (e.g. during food contact) or rubbed off (e.g. pigments from within the ink). For this we need strong cohesion. Balancing the needs of cohesion versus adhesion is one of the trickiest tasks for an ink formulator. Fortunately, we can use the laws of polymer physics to find a rational balance between adhesion and cohesion. Because adhesion is the basic priority (we don't care about cohesion if the ink just falls off), we will focus first on adhesion.

3.1 Adhesion

Before we worry about how to *get* adhesion, we have to worry about how to *measure* it. This is not at all obvious because, to invoke a mantra, "Adhesion is a property of the system". A simple example, involving only surface energy, shows how profound this idea is.

3.1.1 Adhesion is a property of the system

Take some nice sheets of super-smooth rubber (made by casting the rubber onto a smooth glass surface), cut them into pieces, create an adhesive joint by putting the two smooth surfaces together and test them in 3 ways that ensure that their contact areas are about the same. If adhesion were a property of the adhesive then, because we only have pure surface energy sticking the pieces together, it would not matter how we measure the adhesion - it would always be something like 40mJ/m² (equivalent to the surface energy of 40dyne/cm or 40mN/m) which we would calculate from the work required to pull them apart divided by the contact area. In practice we simply measure the force required, but we would expect that force to be similar for the three cases.



W J/m² 0.04

b mm 25

E GPa 0.2

h µm 100

F_{Peel} N 1.00e-3

F_{Lap} N 1.00e+0

F_{Butt} N 2.22e+1

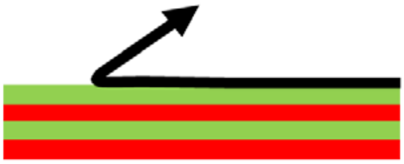
App 3-1 https://www.stevenabbott.co.uk/practical-adhesion/weak-strong.php

The first measured force, 1mN is exactly what we would expect from a 40mJ/m² surface energy (or Work of Adhesion), W, and a 25mm strip of rubber. Now look at the lap joint. It takes a force exactly 1000x stronger to pull it apart. Finally, the butt joint, with the vertical pull upwards requires a force 22,000x greater. Clearly, our intuitions about the same "adhesive" (surface energy) giving the same "adhesion" are very wrong. The key to the differences is that the physics of bond breaking involve different components. For the lap joint the most obvious component, the length of the overlap L plays *no* part in the calculation. It is shockingly the case that above a certain small value of L (below which the theory behind the calculations breaks down), the overlap is irrelevant. Instead, the modulus *and* thickness of the rubber play an extra role beyond W. For the butt joint, only the modulus plays an extra role.

Finally, although the middle joint is called the "lap shear joint" it fails in peel for reasons you can explore via Goland-Reissner theory in https://www.stevenabbott.co.uk/practical-adhesion/g-rlap.php should you wish to know more.

Once you realise that adhesion is a property of the system, you can use that knowledge for good or bad purposes. Suppose, for example, that your customer demands that your ink passes "the tape test". If you know that your ink adhesion is rather poor, there are multiple ways to gain a "pass" on this test if you choose to be deceitful. Obviously, using a very light tape helps you to pass and if your ink has a thin layer of silicone on the surface, even a strong tape will have little grip.

More subtle is to do the test in the middle of a solid layer of ink.



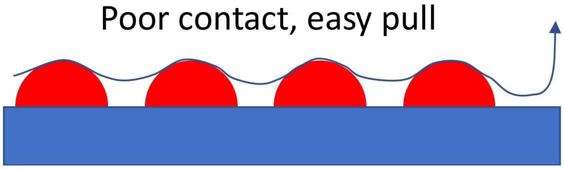
No tape peel

Easy tape peel

Imagine that your ink has only the adhesion offered by a roll of weak tape. If you apply a very strong test tape onto the outside of a roll of tape (the green layer), you will find perfect strong adhesion (between red adhesive and lower green layer) *as long as you keep the test tape away from the free end on the roll of tape.* If you apply the strong tape over the free end then it is easily pulled off. Doing a tape test on the middle of a solid patch of ink is very similar; there is no crack defect to start the process, so the test tape causes no problems. That is why a cross-hatch tape test is a common requirement for those who want a more demanding adhesion test for internal testing or for testing offerings from their suppliers. The cross-hatching creates potential cracks for the test tape to open up.

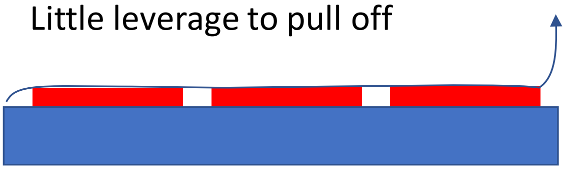
Going back to the example of the three tests with the same material, the test on the free end of the tape is equivalent to the peel test. The test on the middle of the tape (or on the centre of a large area of ink) is more like a butt joint. Given that a butt joint can require a force 10000x larger than a peel test, it is no surprise that many inks that pass a test within a solid area can fail when tested at the edges.

Poor contact, easy pull



We can also imagine a different trick. Apply the tape to a rather rough layer of printed ink. Now the tape is able to touch only small areas of the ink and there is an automatic crack provided - between the ink and the tape rather than the ink and the substrate. The tape will be happy to fail by pulling away from the ink.

Little leverage to pull off



Finally we can do something which is not cheating and is one of those "often observed but seldom explained" phenomena: that thin layers of ink stick better than thick layers. The basic reason is simple: there is very little leverage to pull up the ink!

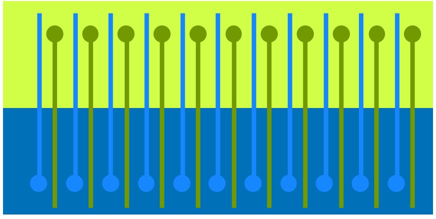
This quick exploration of the fact that "adhesion is a property of the system" leads us to an important question. If the customer demands "good adhesion" they can be told that this is a meaningless demand. They should be saying that they want good adhesion when the sample is subjected to whatever stresses are likely in use, and then should be able to specify tests that are relevant to those tests. There is no point in them asking for (effectively) a butt joint test if the failure mode they fear is peel.

Sometimes the request is made to measure the "real" adhesion of an ink. After all, if we can measure the "real" adhesion then we can work out how to optimize it. Sadly, the answer is that there is no such measurement. Even the apparently simple peel test contains many complexities, such as the fact that the measured adhesion depends on the speed of the test. Those who have tried to remove a sticking plaster slowly and painfully then found that the nurse rips it off quickly and painlessly will have some practical experience of this fundamental truth.

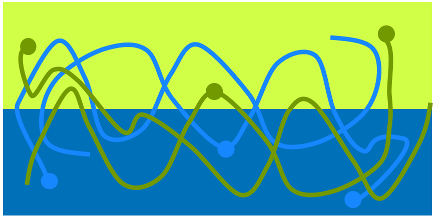
Because adhesion is a property of the system, and because adhesive systems contain all sorts of complexities, the chances that you can measure the real adhesion of any commercially relevant system are zero. Therefore, think as clearly as you can about the system for which you are testing and find the least bad test for it. Often that turns out to be some variation of a tape test. To be able to pass the test without using any of the tricks above, the best advice is to ensure you have plenty of intermingling or, even better, entanglement, without making your system too brittle. If the test is a rub test then whether the failure is adhesive or cohesive, intermingling and entanglement are equally important. Given that most of us have no idea of what is meant by intermingling or entanglement, a crash course in this aspect of polymer physics is necessary.

3.1.2 Intermingling and entanglement

Imagine your ink as a perfect layer sitting on top of a perfect substrate. With a sharp boundary between the two, we can imagine that the crack can easily propagate along the boundary. This is the case for all hard inks so unless we have some "PSA-like" adhesion from dissipation in the ink itself, adhesion will be very poor.



The alternative is for polymer chains from the substrate to be reaching into the ink, and polymer chains from the ink reaching into the substrate. If the substrate is something like glass or aluminium then it must have some extra molecules locked onto the surface which can mix with the ink's polymers. With polymers crossing the interface, any crack has more of a problem because it has to drag polymers out from either side in order for the crack to continue. This soaks up (dissipates) crack energy and if the effect is strong enough, the crack stops.



If the chains crossing the interface are relatively short then they are merely "intermingled". If they can reach a significant length then they become "entangled", exactly like a ball of string. Now any attempt by a crack to propagate is resisted as strongly as if you are trying to pull out a tangle from a ball of string. The adhesion increases strongly when you go from intermingling to entanglement.

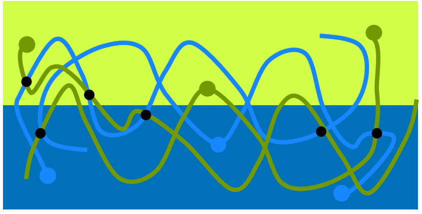
When you carefully select the solvent system for your ink, one key factor is that the solvent is able to provide some "bite" into the substrate. This means that the polymers in the ink and substrate have a chance to intermingle or entangle. With too much bite, the substrate surface becomes damaged. Too little bite and there is insufficient adhesion.

Some polymer substrates are not touched by any (reasonable) solvent because they are highly crystalline. The only way to get good intermingling and entanglement is via a process that destroys the crystallinity without destroying the polymer surface. The right level of corona, flame or plasma treatment can achieve this for PE, PE and PET. For PET a xenon flash can also be used to quickly melt the surface which freezes rapidly into the amorphous state.

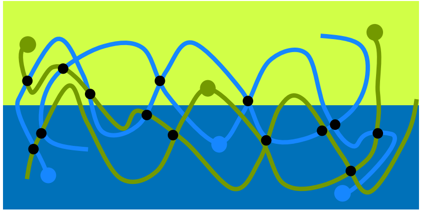
3.1.3 Entanglement via crosslinks

In some cases, it is impossible to get intermingling or entanglement with the substrate. For example it is hard to functionalise an Al surface with long polymer chains. Or, as we shall see with a primer such as PEI, it requires a weak intermediate polymer that itself cannot entangle with either substrate.

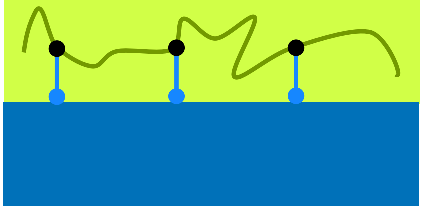
The trick here is to gain entanglement via crosslinks. Here we come to a major misunderstanding that has caused many of us major problems. It is obvious that creating chemical bonds across an interface will create strong adhesion. Unfortunately, although it is obvious, it is wrong. If you have a whole interface of strong chemical bonds, the adhesion is very small (1J/m²) because all the crack energy is focussed on that row of chemical bonds and a simple calculation (visit https://www.stevenabbott.co.uk/practical-adhesion/chemical.php if you are curious to know more) shows that breaking these chemical bonds is no problem for a typical crack. This is why glass is so weak - the crack energy is focussed along the strong chemical bonds which are no match for such concentrated forces. So when we get strong adhesion using chemical bonds, the mechanism has to be something different. Fortunately, we already know the one thing that can give really strong adhesion, and that is entanglement. We get effective adhesion via chemical bonding if these bonds create entanglement.



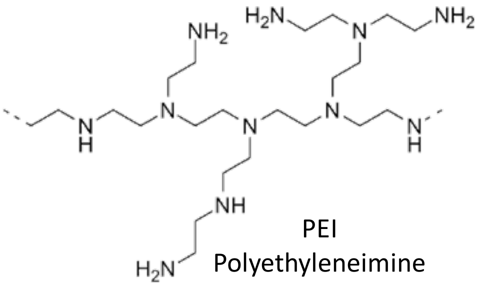
In this version of the entanglement diagram I have added some black dots at some of the crossing points. These represent chemical bonds at the overlap point, rather than mere physical crossing points. The important point is that if you try to separate the two surfaces, you still have exactly the same entanglement as before. Changing from physical to chemical crossing points does not alter the fundamental principle that the entangled polymers have to stretch and distort if the crack tries to pull them apart. And this brings us to a super-important point.



If a few chemical crosslinks give us strong adhesion, surely more crosslinks will give us even stronger adhesion. No! By adding lots more crosslinks we have stopped the polymer strands from stretching past each other - we have made a brittle interface. Trying too hard to get strong adhesion gives us poor adhesion. Many of us have found this to be true, and to me and many others I have known it was a puzzle as to why more crosslinking could give worse adhesion. Once we understand that strong adhesion comes from dissipation, which requires entangled chains to stretch and absorb energy, we can learn to find the minimum amount of crosslinking to give us the maximum stretching with the minimum brittle failure. There happens to be a precise formula (Lake & Thomas) to identify this sweet spot; you can find it in my book or on the Practical Adhesion website.



Entanglement explains how we can use short-chain molecules (such as APTES, aminopropyl triethoxysilane) to get good adhesion to substrates such as Al, something vital for those who print onto aluminized packaging. A low level of the primer, stuck well to the substrate, can interact with a few chemical groups in the ink. This forms an entangled network. If you add too much of the primer, creating a solid array of molecules along the substrate, then adhesion fails because the interface is too brittle. It has often been observed that a low level of APTES is superior to a higher level; the need for entanglement and dissipation explains this observation.


PEI
Polyethyleneimine

And we can explain how PEI can stick PE to, say, PU, something vital for many printers in the packaging and laminating industries. First the PE is corona treated to add a low level (few %) of -C=O and -CO2H functionality (any -OH functionality is a waste). Then a very thin (50nm) layer of PEI is applied. The amine groups in the PEI react with the -C=O and -CO2H groups to form a loose network. Finally when the PU ink is applied, remaining amine groups in the PEI can react into the PU network. We now have an entangled network across the PE-PEI-PU interfaces and adhesion is super strong. If the PEI layer is too thick then we merely have a weak PEI polymer layer which is easily broken. If the functionality is too high (or the PEI is supplied with too high a level of internal crosslinks) then the interface is brittle and adhesion fails. Because the amine groups in PEI can also react with acrylates, we have a similar large adhesion to UV inks.

3.2 Cohesion

If you have to solve an adhesion problem it is always a good idea to have some idea what *type* of problem it is. I was once asked to help with a real crisis where adhesion failure was causing major problems with some consumers of the printed product. Given a description of the problem via email, I was able to make many sensible suggestions about how to improve adhesion for this specific product, yet all of them were a waste of time. Because when I was shown a sample and saw the test for adhesion failure, it was immediately clear that the adhesion was excellent. The failure was a *cohesive* failure. For this product, the relevant "system" test was a rub test with a range of test liquids. The rub tests showed the coloured ink coming on to the test pads, and everyone said "adhesion failure". Yet even a casual glance showed that the inks remained locked hard onto the substrate and all that was happening was that pigment particles were being dragged out of the ink matrix, a classic cohesion failure.

We have done the hard work to be able to understand how to improve cohesion - and the dangers of trying too hard.

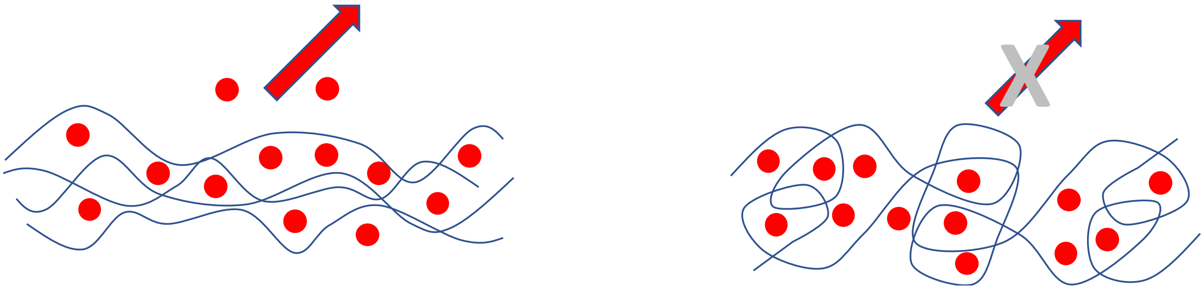


Figure 3-1 Pigments in a polymer matrix compared to pigments in an *entangled* polymer matrix.

On the left of the figure, our pigment particles are nicely sitting within a polymer matrix. Unfortunately the polymer is not tangled with itself, so it is very easy for a bit of rubbing and suitable solvent to drag out some pigment particles. On the right, the polymer is entangled, so it is very hard for the rubbing and the solvent to reach in and drag out a pigment particle.

Just as we found that for adhesion, chemical crosslinks are an exact equivalent of physical polymer tangles, so we can create a cohesive network using a crosslinked polymer. We can even go one stage further and lock the particles to the polymer matrix with extra bonds.

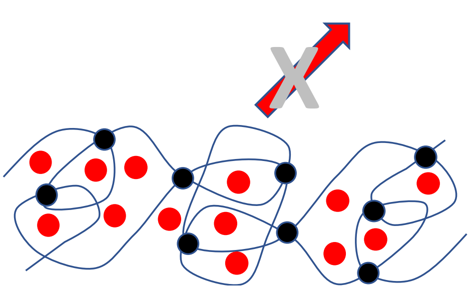


Figure 3-2 A crosslinked matrix achieves the same locking in, and additional bonds to the pigment itself are possible.

The problem with these crosslinked systems arises if you try too hard. Focussing on locking in the pigment might give a brittle ink which fails to absorb crack energy at the interface with the substrate. It might, therefore, be better to cut back on chemical crosslinks between the polymer chains and increase the links between pigment and polymer. If the pigment has multiple links then these are just as effective as chemical crosslinks within the polymer itself.

The trick is to avoid trying too hard.

If we are going to ensure that the pigment is, at the very least, compatible with the matrix then we need some dispersants. From the above discussions, it is clear that there are multiple options, shown in the diagram:

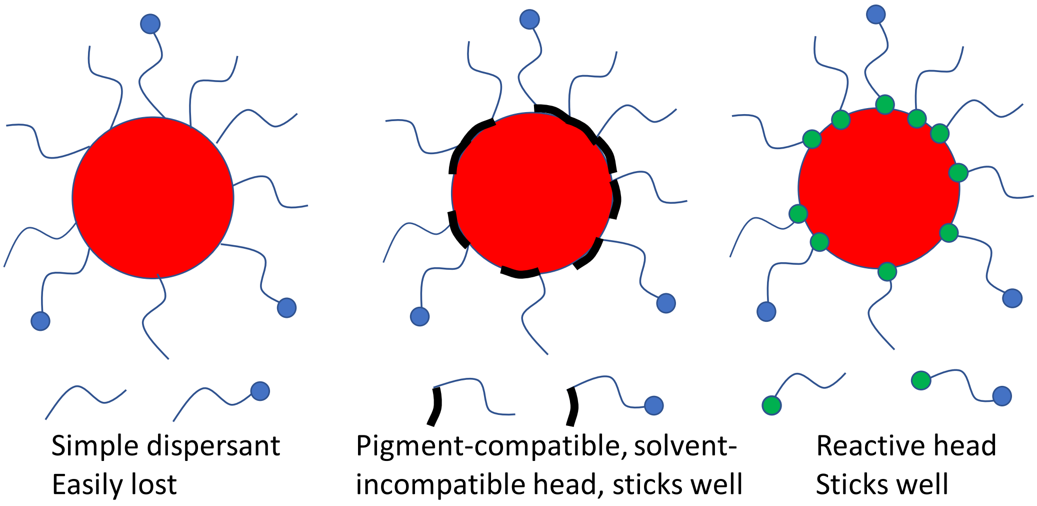


Figure 3-3 A range of options for pigment dispersants.

The simplest option is a long-chain molecule that you hope will stick sufficiently to the pigment and yet be nicely compatible with, and entangled with, the ink matrix. It might feature a reactive group on the end, capable of locking in to the matrix giving cross-linked entanglement. The obvious risk is that the dispersant is so happy in the matrix that it deserts the pigment, allowing it to become unstable during processing but, in the context of cohesion, making it easy for a crack to propagate by zipping past the pigment particle.

The next option is to make a two-part dispersant (again with an optional reactive group on the end). One part dislikes the matrix and loves the pigment with the other part having the opposite tendencies.

The final option is to react the dispersant with the pigment, still keeping the option of a reactive group at the other end.

For those who simply buy in a pre-formulated ink, this might all seem rather remote. But you have a choice of pre-formulated inks, and by pushing your supplier to say what types of dispersants are being used, you are more likely to discover either that your supplier has no idea (which would be worrying), that your package has all the features you require, or that the package is over-engineered for what you need.

In terms of the balance between adhesion and cohesion, the ideal is to have the best of both - entanglement within the ink for good cohesion, then entanglement with the substrate either as a polymer or as a functionalised surface such as aluminium. If you cannot do both, there is a choice of strategies, depending on the end-user requirements. The problem isn't one of choosing the optimal strategy - that can be rather straightforward when you pose the problem clearly. Instead, the problem is being unaware that there is a choice.

Finally, whatever choice you make, it is important to remember that trying too hard makes things positively worse by making the coating or the interface too brittle.

3.3 Using the core principles

These first three chapters are in many ways more important than the chapters on the individual printing techniques. Modern presses are so sophisticated that you can think of them as magic boxes into which you feed a substrate and out of which come a set of perfect prints.

And if you are printing "standard" inks on "standard" presses for "standard" purposes then there is really no need for you to be reading this book. Everything will work just fine if you just assume that it happens by magic.

If you *are* reading this, the chances are that life isn't quite so easy and that you have experienced failures and problems. My own experience (with plenty of failures and problems!) is that a core set of principles is a valuable asset. Very often people (including me) jump into problem solving with a great idea, only to find that the idea doesn't solve the problem because the problem isn't what we all assumed it to be. Having a set of principles allows a number of ideas to be considered, rejected or prioritised, and when the first ideas don't work out, then there are more ideas available, along with extra evidence and data from the tests that didn't succeed.

There is often a tension between "experience" and "science". The best of all worlds is to be able to learn from the experience and translate it into science by extracting the correct bits and eliminating the more common myths and legends. More normal is for the senior people with "experience" that happens to be wrong, to block the junior people who have the right basic science but lack the experience to implement it. There is no easy way around such people, but in the end, if your organisation is serious about success, the science will win.

4 Offset Litho

There are 4 key processes within offset and we will examine each in turn. Because we have done all the difficult groundwork, we can focus on the interesting issues behind offset. The processes (excluding the drying/absorbing/curing of the ink after printing) are:

1. Going from a thick layer of highly viscous ink in the ink duct to an even coating of a few µm of ink on the relevant part of the plate.
2. Going from a supply of low viscosity aqueous fount solution to an even distribution of a few µm of fount on the relevant parts of the plate.
3. Ensuring that the ink sticks only to the relevant part of the plate.
4. Ensuring that the ink is nicely transferred onto the intermediate blanket, without picking up anything that is already on the blanket and then that the ink transfers nicely from the blanket to the substrate without any of the substrate staying on the blanket or going further back up the chain.

Readers may object that there is a further process: producing the plate. The science of producing the plate either via old-fashioned analogue photoresist processing or via direct laser writing is fascinating and important. But my concerns in this book are with the science of the key printing steps. As we shall see, after decades of research into the offset process there are still many confusions remaining. So let us accept that the plate appears by some sort of magic and focus on what happens once we have it.

Just one comment is needed about the plate itself. The assumption in this chapter is that the plate is "planar", i.e. there is no need to take into account any height differential between the image and non-image areas. To justify this assumption, I have seen no hints in the literature that the small height difference on real plates plays any role in what happens on the press.

4.1 The ink train

As we discussed in the first chapter, printing is all about splitting inks. Given that we are all happy with ~1µm of printed ink[13], and given that the ink split from the plate and from the blanket are both ~50%, that means that we need to deliver an even layer of 4µm onto the plate (which delivers 2µm to the blanket). It happens that it is near impossible to deliver such a thin coating at such high speed (>500m/min) by any direct method, so the only viable route is via a train of ink rollers. If we have 8+2 splits in the chain then to deliver 1µm to the image we need 2^{10} µm of ink (approximately) delivered at the top of the train; that is ~1000µm, i.e. 1mm thickness of ink that can be delivered with the simplest of ink ducts.

---

[13] Throughout the book I will use the idea of 1µm of printed ink going back via a simple splitting ratio. In practice, inks vary in % solids from, say, 10 to 100 so the absolute thicknesses need to be adjusted accordingly. For a10% solids ink we would need 40µm on the plate to deliver 10µm wet for a 1µm final deposit.

This logic suddenly imposes a major constraint on offset inks, one that has stopped it from being just about the only print process anyone would ever use for high-volume work. Each contact between the rubber rollers at each stage of splitting (plus the energy dissipation within the ink itself) generates a lot of heat. And a train of rollers has a large open area. This means that an ink with a volatile solvent cannot be used for offset. Which means that there must be very little solvent in the first place. Which means that the ink is super viscous - a paste rather than an ink.

The reason this chain of logic has stopped offset from being all-conquering is that for the large flexible packaging market, solvent has to be used both to gain some "bite" into the substrate for good adhesion and to be evaporated quickly because it cannot be absorbed into the substrate. Normal offset cannot work with such high-volatility systems

That one downside comes with the many upsides which explain why offset has been so dominant. The high viscosity of the ink means (as everyone knows) that the dot gain is very low, allowing super-high LPI plates with exquisite print quality. What is less well-known is that these high viscosity inks must split via sudden catastrophic failure between plate and blanket and between blanket and substrate. Anyone who has compared a 100% solid from offset and from flexo will see that the flexo "solid" is (or, till recently, was) full of holes, which must be caused by cavitation. With the near-solid offset ink, there isn't even a chance for large bubbles to appear via cavitation and a sharp fracture, giving a smooth solid, is produced. My faith in this explanation was shaken by excellent arguments in the literature that the fount emulsion drops within the ink would allow rapid splitting as they provide instant positions for cavitation and crack propagation. The argument against this is that, as far as I know, the print quality of waterless litho is, in terms of splitting defects, indistinguishable.

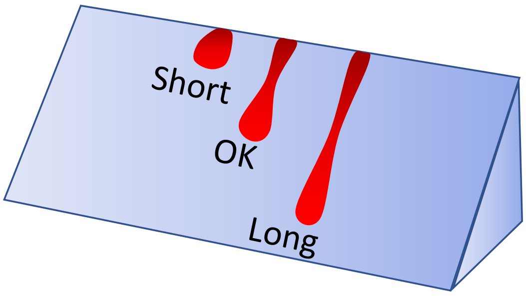
If the rheology of the ink had to cope only with the key steps (discussed in more detail below) of the printing splits, the challenge would be hard yet manageable. The problem is that the ink has to split beautifully within the ink train without producing one of the grand challenges of (high speed) offset, misting.

I have found very little useful work on the issues of rheology and misting. Maybe people have tried but found the problem intractable. One reason for the difficulty links to the reason for the excellent splitting of the printed ink. The timescale for splitting a 2μm film is very much faster than splitting a, say, 200μm film near the start of the ink train. And this makes a difference to the rheology, as a thought experiment can reveal. In rheology, temperature = time. A short timescale process is equivalent to a low temperature process because any polymers in the ink are effectively frozen at high speed. So the 2μm film might split as a brittle ink, while the 200μm film has the time to show its viscoelastic properties. And as we discussed in the splitting section, viscoelasticity can produce large liquid bridges that can create large mist particles.

Having said that, because of the need for multiple splitting, a complex ink train is the only way to deliver an even film of ink to the plate, there is an extra complication. At high speeds it has proven difficult to ensure that a single roller contact with the plate can reliably deliver the even amount of ink, so the train splits into sub-trains which contact the plate at different points.

Having explained why offset has such an advantage, it is worth pointing out that standard offset cannot use water-based inks. This *should* have been a major problem as water-based inks sound so nice in theory. In practice, water-based inks used in other techniques have had decades of promise with only modest increases in market share, so offset's disadvantage in this respect is not (yet) a crushing one.

4.1.1 Yield stress



There is one more complexity for the ink design. The ink must flow readily within the ink duct, but not *too* readily. The ink therefore has to show some degree of yield stress - it will not flow without a modest stress applied to it. Hence the presumed need for a test for whether the ink is too short (too self-entangled and with a high yield stress, or too long (no tangles or structure to provide a yield stress) by judging its flow down a slope of known angle. Such ad hoc tests cover up many issues and can miss key differences between different inks, so their value lies only in quick tests of batches of the same ink: i.e. if the result one day is different from all the other days, that flags up a warning.

4.2 Fount delivery

The many variants of fount delivering ("damping" or "dampening") systems use some common elements and two broad choices. The choices are, first, between delivering the fount straight to the plate or delivering it to one of the inking rollers in contact with the plate. The second choice is whether there is a continuous link from the fount solution trough and the plate or whether a gap crossed by brushes or sprays is used.

By isolating the system with a gap there is no chance of contaminating the fount with ink, but control of the system is less responsive. And conversely, by going straight to an ink roller, the control is most responsive, with a higher chance of the ink coming into the fount tray or fount going up the ink train. There are plenty of articles extolling the virtues of either approach, and neither is wrong or right. You just need to be aware of which system you have and, therefore, its strengths and weaknesses.

Whereas the ink control is about getting the right printed density, fount control is about getting the right relative amount of fount to ink. Here our interest is in how the relative amount can be controlled without a long, complex train which would evaporate the water and alcohols. The answer is that by varying the relative speeds of rollers in close contact, the split ratio can be controlled far from the standard 50%. For those who are interested, the formula for the relative flux, F, transferred at a speed ratio of the two rollers of S, depends on an intermediate value R=S(S+3)/(1+3S) according to:

Equ. 4-1

$$F = \frac{SR}{1+RS}$$

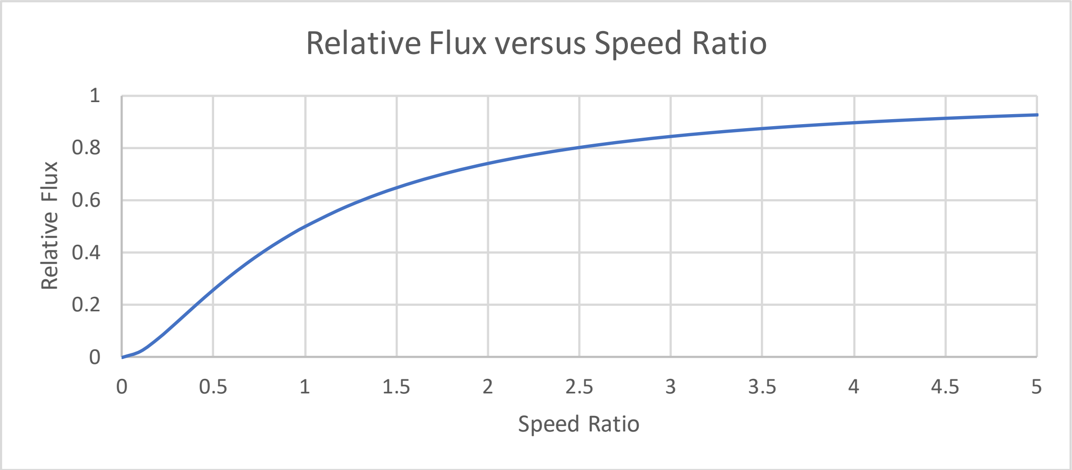


Figure 4-1 How the flow of the fount can be controlled by the roller speed ratio.

This trick only works with low viscosities - it cannot be used effectively for the highly viscous inks.

In addition to the speed ratio, if a hard roller is pressed against a rubber roller then the amount squeezing through, H can be controlled by the speed, V, viscosity η, rubber modulus E and the applied load W according to:

Equ. 4-2

$$H = \frac{3\left(\eta VR\right)^{0.6}}{W^{0.2}E^{0.4}}$$

The reason for showing the formula is that it is surprisingly resistant to all the controlling parameters which are each to a power less than 1. If you double the speed then the thickness getting through increases by 1.5, which is not too bad and could easily be corrected by adjusting the roller speed ratio. However, if you wanted, instead, to increase the applied load to bring it back to the original thickness, W would have to increase by a factor of 8.

The formula also explains why you can't reduce the ink train significantly by using pressure nips. The ink viscosity is so high that there is no viable set of parameters that can deliver the necessary small values of H.

4.3 Differential ink adhesion on the plate

Having set the scene with the fount and the ink arriving at the plate, I want, first, to throw away the fount and talk about waterless litho. This is necessary because I need to remove the core myth behind offset. As soon as the myth is removed, it becomes possible to have a sensible discussion about the litho process. The myth is so deeply part of the offset story that it gets in the way of a rational discussion of what happens with conventional, fount-based printing. By discussing the myth in the context of waterless litho we will be better able to understand conventional litho.

4.3.1 The surface energy myth

Here is the myth expressed in conventional fount-based language: "Ink is hydrophobic so sticks to the hydrophobic portion of the plate. The water is attracted to the hydrophilic part of the plate to keep it nicely hydrophilic so that the ink won't stick." For those who love surface energy arguments, the story goes: "Ink doesn't wet the high surface energy non-image area so it only sticks to the lower surface energy hydrophobic part."

Now let us try to explain how waterless litho works. Instead of low and high surface energy surfaces we have low (image area) and super-low (silicone area) surface energies. The ink does not stick to the silicone area "because silicones have low surface energy".

This myth can be dismissed in an instant, and by destroying the waterless myth it does terminal damage to the water-based myth. It is not at all hard to create a plate with normal printing areas and fluorinated ("Teflon-like") areas with a surface energy as low or lower than that of the silicone equivalent. Yet you cannot get a single offset print from such a plate. The ink happily sticks to everything. Yes, it sticks rather less well to the fluorinated parts, but the differential adhesion is nowhere near large enough to give a printing plate.

If surface energy is irrelevant to differential adhesion, what is the cause of the differential adhesion and high-quality printing of a silicone-based plate? Although I have no data to support me in the realm of waterless offset, I can make a confident prediction from another area of differential adhesion. I predict that if you substitute the normal silicone for increasing amounts of so-called MQ silicone, you will get approximately the same surface energy along with a surface that is as useless as a fluorinated version. MQ silicones are rigid silicones, used extensively by those who want very good silicone adhesives. The standard silicones are very flexible and are used extensively in release liners for, say, adhesive tapes.

It has been shown (the experiments are described in my Adhesion Science book) that silicone release liners on PSAs work not because of their low surface energy (fluorinated liners with the same low energy do not release) but because the silicone is highly flexible.

And that is the secret of waterless litho, it is acting like a silicone release liner because it is flexible. The explanation is somewhat subtle. There is a fundamental law of fluid dynamics called the "no slip condition". This means that a liquid (such as an ink or a viscous PSA) in contact with the surface cannot move at the surface, no matter how large is the speed of the fluid flow away from the surface. If you try to pull a liquid ink (or PSA) away from any "no slip" surface you have to set up a differential flow between the ink starting to pull away and the "no slip" ink. By the laws of viscosity, this means that energy gets dissipated. And as we know, energy dissipation is the key to adhesion. The magic of silicones is that they are so fluid at the surface that the no slip condition no longer applies to anything else at that surface. There is no longer any reason for the ink (or PSA) at the surface to resist motion and it can come off easily.

I do not know if the inventors of waterless litho were aware of how it really works. I have only found one paper[14] which comes close to the right explanation (and showing that fluorosurfaces are useless), though the authors were unaware of the analogy with PSA. The fact that this paper is by Shen's group at Monash U is not a coincidence. We can now return to conventional litho using the truly elegant experiments[15] from the Shen group that reveal how conventional litho works.

4.3.2 Fluidic release

To understand the conventional mechanism, let us go back a few hundred years when litho printing really was done with stones. In those days, the non-printed areas were covered with gum arabic. Well, in those low-tech days they had to use something, and why not some arbitrary exudate of the acacia tree; they had no access to high-tech polymers. So what is the single most usual polymeric additive found in 21st century founts? Yes, it is gum arabic. It turns out that the people who invented litho had chosen gum arabic because it had specific properties that gave good differential adhesion. And we continue to use it for those same reasons. Yet until the paper from Shen's group I don't think that anyone understood what it was doing that made it so good.

---

[14] Junfei Tian, Yu Mao andWei Shen, *Ink Transfer and Refusal Mechanisms in Waterless Offset Printing*, J. Adhesion Sci. Technol. 23 (2009) 281–296

[15] Wei Shen, Brenda Hutton And Fuping Liu, *A new understanding on the mechanism of fountain solution in the prevention of ink transfer to the non-image area in conventional offset lithography*, J. Adhesion Sci. Technol., 18, (2004), 1861–1887

The Shen team did some wonderfully ingenious experiments. They tried printing with ice. Because, in surface energy terms, the surface of ice is rather similar to the surface of water, if ink doesn't stick because of the high surface energy of water, it won't stick to ice. Yet the ink sticks nicely to ice. By slowly raising the temperature of the ice they found that adhesion was lost when a small layer of liquid water started to form on the surface of the ice.

So it is the *fluidity* of a surface that makes it non-stick, exactly as found for silicones. The magic of gum arabic is that it traps a thin layer of water at the surface ("everyone" knows that), and that the fluidity of the water means that adhesion is minimal (the bit that is missed out by those who use surface energy explanations).

This fluid boundary explanation helps with another key aspect of the aqueous fount system. Sometimes a thin layer of ink does stick to the non-image area and a function of the fount is to help that layer to clear spontaneously. An analogy with another classic experiment in adhesion science is useful. Take some clean glass and cure a layer of epoxy onto the glass. Test the adhesion. The epoxy sticks well. Now put the glass plus epoxy into some water. The epoxy floats away. The reason is that glass allows water to penetrate along the silicate surface, getting between the epoxy and the solid glass. Yet again, by having a fluid interface, adhesion becomes minimal. The gum arabic performs a similar function. If it happens to be too dry at any given instant (e.g. during startup) and adhesion is too large (no fluidity), water can easily penetrate through the polymer underneath the ink, producing the required fluidity, and the ink is easily removed when it next contacts a roller.

4.3.3 All that stuff in the fount solution

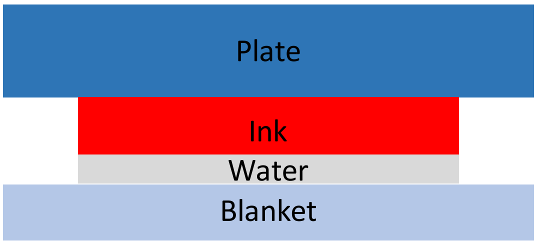
Maintaining a fluid release system at 1000m/min is a tough challenge. So fount solutions contain many additives to address different issues. The first imperative is to make sure that the gum arabic (or its equivalent) is stuck firmly to the aluminium plate so that it can provide its fluidic effect. Although we all know that the surface of aluminium is "alumina" via oxidation, it is hard to know what the actual surface is because it can have a continuum of oxide and hydroxide variants. It seems that "somewhat acid" or "somewhat alkaline" provide some sweet spots, so some founts are mildly acidic and some are mildly alkaline. Going too far in either direction can cause destruction of the plate or press, so using phosphates or silicates as buffers seems to help.

Then there is need for the fresh delivery of water to wet out quickly, so that it can get to where it is needed as quickly as possible. The high surface tension of water slows down this process, so a low surface tension is required. As we discussed, surfactants are generally useless for this because they respond (get to the surface) too slowly; the DST (dynamic surface tension) is too high at short time-scales. So IPA (isopropanol) has, for a long time, been the standard way to reduce surface tension given that it is cheap, effective and relatively safe. Its volatility is an advantage in that it rapidly disappears when it is no longer required, though that advantage has become a disadvantage with the need to reduce VOC emissions. So IPA has been replaced by less-volatile alcohols (such as the Dowanols) which can do a similar job.

All this would be complex enough to make the life of a fount formulator hard enough. Then we hit the problem of emulsification. Reading papers from the decade of the big transition from IPA it is interesting to see what a tough task this was. If it was just a question of dynamic surface tension then the transition should have been easy. Presumably IPA also helped with the correct emulsification and random substitutes did not provide the required functionality. The question is, therefore, what is required from the emulsification of the fount within the ink.

4.3.4 Emulsification

We know from waterless offset that there is no need for an ink to contain 20% or more water in order to print. We can read about the positive benefits of water being in the ink, e.g. to provide some welcome cooling in the heat-generating ink splitting process between plate and blanket. This seems to me to be a rationalisation about an unwelcome side-effect of the offset process. Rather than use water to indirectly cool the plate, temperature is far better controlled via a cooling system in the roller, which is what they do on a waterless press.



My reading of the literature suggests that emulsification is necessary to fix two problems inherent to a water/ink system. The first is rather obvious. Because the ink and water repel each other, and because the whole plate must be covered by the fount solution (the delivery system has no knowledge of where water is most or least wanted), the natural outcome would be a thin film of water sitting on top of the ink. During the splitting process between plate, ink and blanket, the fluid layer of water would break first, giving 0% transfer to the blanket, so no printing.

The second reason I found in a discussion site on the internet. I have not been able to find any academic analysis, but it seems to me to make sense. When the plate transitions from a fount-flooded, non-printing region to a printing region, it is easy for the ink to find the not-very-wetted part of the plate. At the trailing edge, in the transition back to the wetted region, the argument goes that there is a shortage of fount, so the fount emulsified into the ink is a key additional resource. Without it, trailing edges, it is claimed, are not at all clean.

The absence of clear reasons for the emulsion does not mean an absence of interest in the emulsification. However, as you will find in the next few pages, everyone is working hard with variants of a test (Surland) which has no obvious direct link to the press, and those doing the tests are often varying individual components in isolation while trying to optimize for the only thing which matters which is how the whole system actually prints.

Since Surland published a set of classic emulsification curves or "water balance" curves, everyone cares passionately about them. It is said that the ideal balance curve translates to ideal on-press behaviour which is defined by a large distance (measured by the different levels of fount) between two points The first point is "catch up" (where the fount is reliably removing excess ink on the non-printing areas) at a comfortably modest feed rate of fount. The second point is "wash out", the point where the fount is obviously interfering with the print quality by diluting the ink too much. If the distance between the two points is large, then the press has a comfortable, stable printable zone.

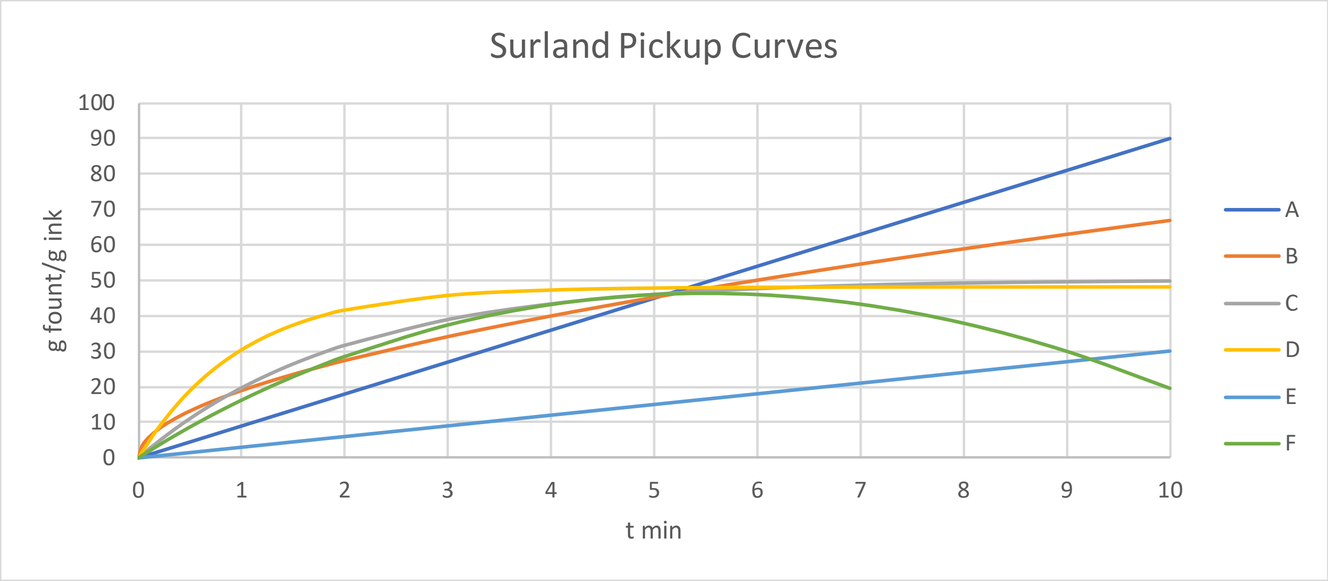


Figure 4-2 The classic Surland fount pickup curves

The curves show the amount of fount emulsified into the ink following the shown number of minutes of emulsification with some sort of blender blade, with some minutes of free standing to allow non-emulsified water to separate.

Everyone agrees that a Type C curve is optimal in general because it reaches saturation not too quickly and remain immune to all the extra water added after, say, 5min. Curves such as A show no ability to reject excess fount and curves like F show odd behaviour at high found content. Yet I have found very few explanations of how these curves directly relate to on-press performance and why an emulsification capacity of, say, 50% is better or worse than, say 40%. Why does *any* test based on measuring water pickup over a 10 min timescale have *any* relevance to what happens on a high-speed press? If the average bit of ink+water survives for five revolutions of 1m circumference before being completely replaced and the press is at 500m/min, the ink is emulsified for 60/100s, i.e. ~1s, so how is the difference of C (good) and D (bad) of any relevance just because D picks up fount after 3min instead of 5? Even the mixing method in the test has no relationship. On older Duke testers, the mixing is with a "stirrer". On newer devices the mixing is with a high-speed blade. On a press, the mixing is via squeeze flows, the dynamics of which are very different.

[Feel free to skip this small digression about squeeze flows. With conventional mixing, a blob of (say) oil can, under some conditions, just spin faster at higher mixing energies rather than breaking into smaller blobs. For those who find this an interesting notion, I have a web page about it: https://www.stevenabbott.co.uk/practical-surfactants/critcap.php. Under squeeze flows there are no such limits. Some emulsions of higher-viscosity oils into lower viscosity fluids are impossible to make via conventional stirring but present no problem when produced in a high pressure "squeeze/stretch" homogeniser.]

A further puzzle is that the older literature relies on the measure of how much fount is emulsified over time, while on modern machines the measure is how the torque changes with time, with a steady increase being indicative of a gradual emulsion formation (see https://www.stevenabbott.co.uk/practical-surfactants/emuvisc.php for how fast viscosity increases with vol% emulsion). The torque measure can also identify when the emulsion breaks down at the "emulsification capacity". Again, how does this process over 10 min relate to on press issues?

An alternative or complementary test measures the torque of a shearing disk in the ink, with water gradually being added once the torque has settled down and been tested for consistency over a few minutes' running.

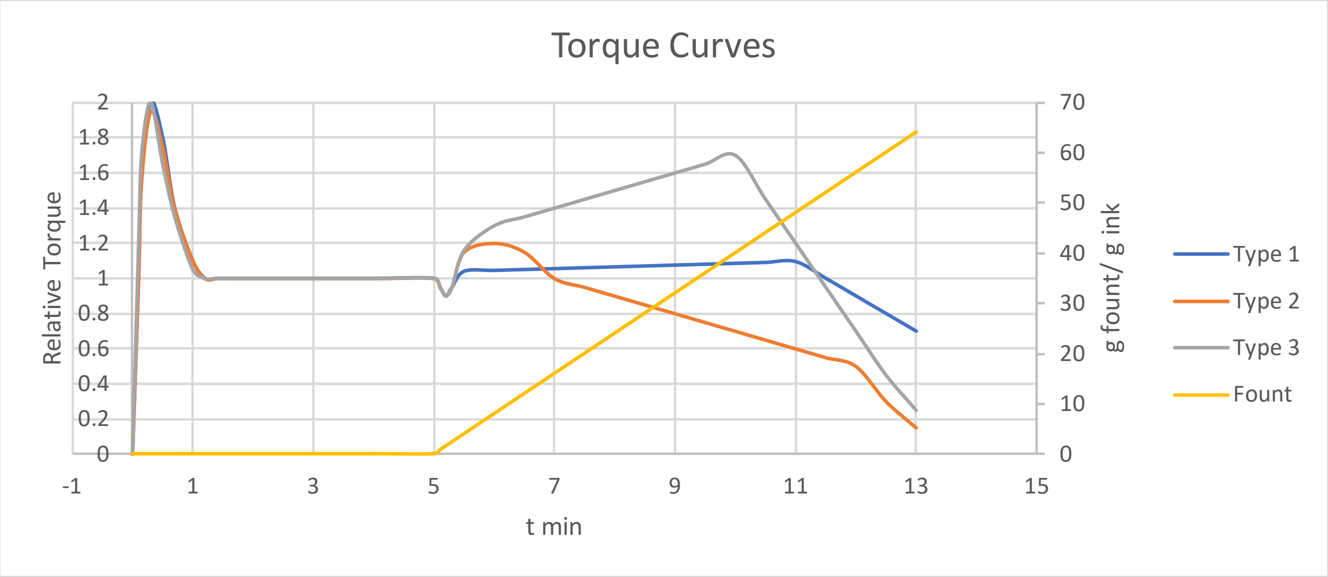


Figure 4-3 A modern torque curve to replace the Surland curve

There are, broadly, three types of torque curve. Type 1 is the optimum and shows, as theory suggests, a modest increase in torque as the percent fount increases. It is worth noting that although the viscosity increases, the shear modulus, G', decreases, making the ink at the same time more viscous but weaker. This will be discussed further when we come to tack.

The point at which the torque decreases indicates the "emulsification capacity" which, as with the Surland curves should be neither too small nor too large. Type 2 indicates an unstable system that changes rather rapidly with water content. Type 3 shows a large ΔT, change in torque, as the water is added and this is generally seen as a bad thing because such large changes indicate unnecessary complications.

There seems to be a good correlation between Surland curves and torque curves, with the latter being much easier/faster than the Surland technique.

There are publications (especially in TAGA in the 1990s) pointing out that proper rheology on inks+emulsions should be providing deeper insights into what is going on during emulsification. But once more, it is not obvious why there should be a link between a rheology scan done over several minutes and an on-press emulsification step taking μ-seconds or milliseconds.

Finally, some tests include a look at the quality of the emulsion being formed. It is generally agreed that an even distribution of small (2-3μm) drops is a better sign than a wide distribution with plenty of large drops, with, again, no obvious link between cause and effect.

The situation is even more bizarre because although it might be rational for fount suppliers to tweak their fount to give an "optimum curve" with a given ink (or, more likely, a given set of inks as some colours have very different performance), raw materials suppliers are urged to obtain good curves by tweaking, say, a polymer additive with a given fount (or just water!) and with just some components (say just the alkyd components) of the ink, without other components such as the pigments. My most charitable view of this is that it is a way to check if there are any components that are excessively surface active in the manner discussed when I summarize what I think is going on. For example, a relatively pure vegetable oil might have an emulsification curve very different from the "same" vegetable oil that contains lots of (polar) monoglycerides or glycerol.

There are plenty of academic publications that explore the Surland water balance curves (or equivalent torque curves) and the relationship to print quality, yet the majority just make observations such as "there was more tinting with this oil, less with this pigment, the same with this set of additives and some strange behaviour with this fount" with no attempt to provide a theoretical framework which would help someone with a different oil, pigment, set of additives or fount. 35 years after Surland's 1983 paper, we seem to be none the wiser as to how to control the ink/water balance via rational methods.

So let us have a go at seeing what factors are relevant to what really goes on during the printing process. The next few sections are entirely negative; they show that many of the classical ideas about emulsions are irrelevant to offset. They are included because once we know which issues are irrelevant, it becomes easier to find which aspects of the system are really important.

4.3.5 Emulsification science

Readers who are interested in the complexities of emulsification can read all about them in my (free) Practical Surfactants eBook and/or look at the relevant apps on https://www.stevenabbott.co.uk/practical-surfactants. I can save you the bother of studying these complexities because they are all irrelevant to offset emulsification. Strangely, they *are* relevant to most of the devices used for exploring the water balance curve, but because these devices do not mimic the key features of offset emulsification we can ignore the emulsion science.

The reason that emulsification science is irrelevant is that for offset we have the very peculiar case of super-high shear in super-confined spaces and the few academic papers that come close to investigating these conditions simply say "Small emulsion drops will form whatever happens".

Remember, by the time we come to the interface between ink/fountain rollers and the plate we are talking about 4µm ink films (let us say 5µm once we have 25% fount). If the press is at 600m/min, i.e. 10m/s, shear rates are $10/4e^{-6}$, i.e. 2.5 million! As long as we have some approximation to an emulsion (rather than a layer of water on top of the ink) the water has no choice but to be smashed and stretched into 2µm droplets. As those who study "high" shear emulsification in "narrow" (100µm!) gaps would say, it is trivial to transform a low-quality, large drop size/variation emulsion into a high quality monodisperse emulsions under these sorts of conditions.

Clearly the Surland-style emulsification test is orders of magnitude irrelevant to the question of whether an emulsion will form, and whether the droplet size will be the ideal few µm. At the same time, the idea that these high shear systems will inevitably form a perfect emulsion goes against our experience, especially with Type E inks (the lower line in the Surland graphs) that refuse to accept any water in the Surland-style test and which, all investigators agree, instantly scum beyond printability.

4.3.6 Emulsion capacity

There is agreement that if an ink cannot accept, say, 10% of fount then it will not print and that if it accepts, say, 60% fount then it is probably too absorbent. There seems little agreement about what constitutes the ideal amount that can be accepted: some inks seem happy at 25%, others at 40%.

What everyone agrees on is that this "emulsion capacity" as measured by these sorts of tests is an important parameter. Because the phrase "emulsion capacity" did not occur in any of the 100s of references I used for my surfactant science book (so it is not a favourite topic of surfactant scientists) I did a literature search that revealed a woolly and subjective term.

In the context of standard surfactant theory, there is some sense that the ability to create a good emulsion depends on the ratio of surfactant to the amount of dispersed phase and also on key properties of the surfactant such as the head area - with larger heads able to surround more drops for a given number of surfactant molecules. As you can explore in the Emulsion Surface Area app, https://www.stevenabbott.co.uk/practical-surfactants/emusa.php, 0.3g of a typical surfactant can fully surround 100g of dispersed phase with an average radius of 1µm, i.e. the typical 2µm drop formed within the few µm of ink on the plate. If we allowed the drops to be 10µm then we would need only 0.03g of surfactant to achieve full coverage.

So, according to surfactant science there is no obvious "emulsion capacity" until you specify the size of the drops which will be so large that they will rapidly (depending on your desired timescale and viscosity of the continuous phase) coalesce, giving you effective separation of the two phases. Given that the continuous phase has a very high viscosity, even a very bad emulsion, irrespective of the type of emulsifier, will appear to be stable, so the "capacity" is highly subjective.

The use of surfactant science in this section was merely to try to put the debate onto some rational grounds. We know that in the relevant timescales, surfactants are of no use (dynamic surface tension response too slow) in the context of fount solutions. However, by accident or design we have many surface active elements, varying from large quantities of IPA, through smaller quantities of low-solubility IPA replacements, through to the various resins which may or may not contain "impurities" which act at the interface and even to the oils where "vegetable oils" can readily contain significant amounts of monoglycerides (one fatty acid attached to a glycerol molecule) which are the low-grade versions of the recently popular polyglycerol surfactants.

Interestingly, one definition I encountered defined emulsion capacity as the amount of dispersed phase that can be accepted before emulsion inversion takes place. Which nicely brings us to the next issue.

4.3.7 Emulsion Inversion

The natural emulsification process when there is relatively little water and a lot of oil is to form the desired water-in-oil emulsion. As it happens, getting a low viscosity drop phase into a high viscosity continuous phase is also the "natural" way for emulsification. At large water to oil ratios (typically >75%) there is not enough oil to create a continuous phase, so the drops of water join up to create a new continuous phase, giving an oil-in-water emulsion. This is emulsion inversion and is what can happen on a bad press day.

Unfortunately, the science of emulsion inversion shows that it is a complex process and the inversion point can move in either direction from the theoretical 75% value. A fundamental law (see https://www.stevenabbott.co.uk/practical-surfactants/critcap.php) tells us that for conventional shear-induced emulsification, it is impossible to create a good emulsion when the ratio of viscosities of the drop phase to the continuous phase is larger than 4, so for the extreme (1000s:1) ratio of offset inks, it should be impossible to create an inverted emulsion. However, the reason for the fundamental law disappears when the shear takes place in a confined space, so early inversion becomes possible.

If we define emulsion capacity as the amount of dispersed phase that can be added before inversion takes place then we solve a major problem noted in the previous section. However bad the fount-in-ink emulsion might be, phase separation will be very slow because the ink is so viscous. As soon as inversion takes place, we have an oil-in-water emulsion where the continuous phase is low viscosity so separation can take place easily

If the "inversion" definition of emulsion capacity is the intended meaning, it would be very helpful if those who wrote about it would make it explicit. We shall see that because inversion implies scumming in the non-image areas, a large emulsion capacity is necessary to avoid this problem. However, Type A and B (the top curves in the Surland graph) seem to have a very large emulsion capacity (presumably with no phase inversion) so the term remains ambiguous and unhelpful.

4.3.8 Emulsion stability

We need the emulsion to be stable; if the water separates out to the surface then printing will be poor. Those who are familiar with emulsion stability issues will be aware of Ostwald ripening as one process (https://www.stevenabbott.co.uk/practical-surfactants/ostwald.php), and coalescence (https://www.stevenabbott.co.uk/practical-surfactants/coalescence.php) as another process that affects emulsion stability - in both cases small drops naturally grow into bigger drops. As with emulsification science, we can say with complete certainty that these mechanisms are entirely irrelevant. Under the most extreme conditions in my Ostwald ripening app, a 2µm drop will grow to a 3.4µm drop in 48hrs. Coalescence is typically slower.

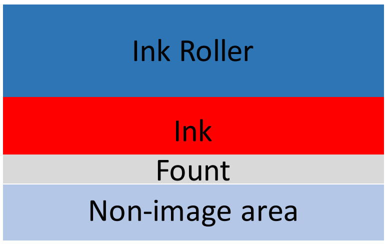
Our emulsion has to be stable for the 0.5m travel time between plate and blanket which, at 10m/s is 50ms.

Clearly the Surland-style emulsification test, waiting minutes for emulsification and separation is orders of magnitude irrelevant, even if we are talking about emulsion inversion where the separation is relatively fast.

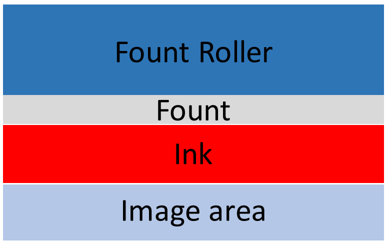
4.3.9 A coherent story

For me, the first coherent story of what is really going on comes from Thomas Fadner[16], published in a conference proceedings from 1982 titled: *Surface Chemistry Control in Lithography.* It seems regretful that Fadner's insights were somehow ignored or forgotten. Had his ideas been followed up, offset printing would be in a much better intellectual space. What follows is my own version of Fadner's views, taking advantage of a few decades more academic literature on emulsification and adhesion science.

The first thing to note is that Fadner's chapter is not about Surface Energy. He quickly dismisses (as I do) all those who think that surface energy, which is orders of magnitude too weak, is of the slightest relevance to offset.



Instead, what we have to imagine is two separate aspects of what happens on the plate. The first is when the inking roller meets the non-image area covered with fount. At its simplest we have the fact that the ink and fount do not mix much so as the roller separates from the plate, the splitting is between water and plate, so there is no ink contamination of the non-image area. We know from earlier discussions that it is more subtle than this, that the gum arabic or equivalent provides a mobile slippage location. And we know from the small section on emulsification that the ink is going to be forced, somewhat, into the fount solution creating an oil-in-water emulsion and, therefore, the risk of ink scumming of the fount solution.



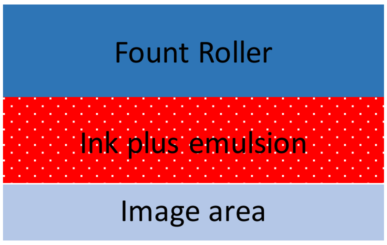
At the other extreme we have the case where the fount roller contacts the inked part of the plate and creates a fount-on-ink layer which, if it split leaving fount on the surface of the ink would produce a non-printing area as described earlier with the picture of plate, ink, fount in contact with the blanket. Again we know that violent

---

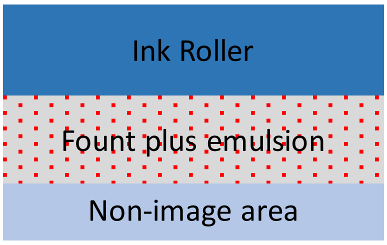
[16] Thomas A Fadner, *Surface Chemistry Control in Lithography*, Ch 17 in *Colloids and Surfaces in Reprographic Technology*, ACS Symposium Series Vol 200, 1982

emulsification theory tells us that this should be a perfect water-in-oil emulsion, giving us a perfect print.

The Fadner paper is about surface *chemistry*. He was working at a time where memories of printing without IPA were strong and he was actively involved in finding non-volatile alternatives to IPA. Because he knew that surface energy arguments were irrelevant, he focussed on what the surface chemistry implications were for the process. In particular he noted that IPA replacements such as 2-ethyl-1,3-hexanediol were optimal near their maximum solubility in water, suggesting that it was their presence at the water/ink interface that made them effective.



Translating his language into a more modern surfactant language, his insight translates into: "If the interfacial energy (at very short timescales) between fount and ink is very low then violent emulsification can take place rapidly, allowing the fount to end up in the ink as a fine emulsion". This would mean that for easy printing all we have to do is have lots of fast-acting surface-active agents which allow this emulsification.



The problem, of course, is that if the process is *too* easy then the reverse effect becomes possible. The ink can now emulsify into the fount solution and we no longer have the ability of the boundary slip layer to force a break at the ink/fount boundary. The break could now be in the middle of the ink-in-fount emulsion, leaving plenty of ink on the non-imaging area to be transferred to the blanket and thereby produce a scummed image. As this is an inverse emulsion then if the "emulsion capacity" means the onset of inversion we can see that a low capacity would assist this breakthrough into the fount.

We can see how the water balance curves might have some general link to the issue of finding a broad, stable regime where fount going into the ink is much easier than ink going into the fount. Type E fails because the fount simply does not go into the ink - the interfacial tension must simply be too high. Types A and B fail because it is too easy for ink to go into the fount. Type D is trickier to manage because it is a little too easy for ink to go into the fount. I do not know how to fit Type F into this picture and I don't really care, because the link between the water balance curves and on-press performance is so indirect.

We know, for example, that the rheology of the ink makes a difference. Even given a low interfacial tension, if the relevant shear modulus is too high then there will not be time for the fount to flow into the ink. Similarly, if the shear modulus is very low then it becomes easier for the ink to flow into the fount on the non-imaged areas.

Although the effect of viscosity ratio on classic emulsification is well-known, it has been shown that the effect can be overruled by emulsification in a confined space. Nevertheless, there are some hints that the viscosity ratio does make a difference. It is frequently pointed out that the modest rise (1.5-2x) in viscosity of the fount through the addition of IPA seems to make things easier, though I have not been able to track down what "easier" is supposed to mean. Similarly, there is a feeling that if the polymers added to ensure splitting from the non-image area also increase the viscosity by a factor of 2, this seems to help. Maybe a factor of 2 reduction in the high ratio of ink to fount viscosities is enough to make a difference.

An alternative, though ill-defined view is that the extra viscosity in the fount is irrelevant to the emulsification and simply helps the found to "flow better through the fount deliver system" whatever that means.

4.3.10 Mass balance

An important point was made to me by expert Erik Nikkanen. Everyone, including me, obsesses about the complexity of the ink delivery system and worries both about how the ink gets through the ink train to the plate then, as in the next section, from plate to print. But there is an underlying simplicity which we forget: if you feed X grams of ink per second then you end up with X grams of ink per second on your prints. This is the law of mass balance: ink going in = ink going out. Given that you know the text/image/coverage required from each ink, you know in advance precisely how many g/s you should be feeding. By controlling the *quantity* of ink being provided, any adjustments to the press can only be affecting the *quality* of the print resulting from that quantity.

If you are tempted to adjust *quality* by adjusting *quantity* you are caught in a doom cycle - in the short term, changing the ink feed might fix a quality problem, but within seconds or minutes, the system will catch up on itself and your print will suffer from too little or too much ink on average, so you'll have to twiddle some other setting which, in turn, might give a short-term fix but is doomed by the laws of mass balance.

Although this law of mass balance is trivial and obvious, many try to fight the laws of physics by twiddling with the ink feed. And anyone who fights the laws of physics is guaranteed to lose.

4.4 From plate to blanket and blanket to substrate

At last we can discuss the two steps that get the ink from its correct place on the imaged part of the plate to a near-perfect replica on the substrate.

Before discussing all the things that can go wrong in this process, we first focus on what is going on when things go right.

4.4.1 Splitting and tack

The first point to make is that although we talk of transfer from plate to blanket and blanket to substrate, there is *never* transfer from plate to blanket after the first revolution of the press. Working backwards from the known thickness of ink printed on the substrate (as mentioned above, for simplicity we shall call this 1μm) and assuming that we are in a suitable Walker-Fetsko zone (discussed later), we know that the ink split to the blanket is 50:50 so the returning blanket has a 1μm ink deposit, which means that to be refreshed from the plate back to 2μm (ready for the next 50:50 split) there must be 3μm of ink on the plate so that the overall 4μm ink film splits to 2+2μm.

As a consequence, we can say with certainty that neither the plate nor the blanket nor (excluding, for the moment, the complexities of fibrous paper) the substrate has any effect on the quantity (split ratio) or quality of the split, providing that any topology in the blanket or the substrate is significantly below the splitting thickness.

And because the split ratio is always ~50%, our only concern at this stage is with the quality of the split. If (as happens in flexo) the split involves lots of cavitation and/or stringing then the print quality will be poor. If it is a short, sharp split then the only remaining issue is the 25% level of defects we have carefully provided - the drops of fount.

Let us first address, and dismiss, the fount drop defects. Suppose that, magically, all the fount disappears leaving 0.5μm deep holes with a 3μm spacing in what should be a 1μm layer. Although the Levelling app, https://www.stevenabbott.co.uk/practical-coatings/levelling.php, described earlier does not allow you to enter these extreme settings, the formula tells us that the levelling time is ~0.1s, i.e. the defects will never be visible. Removing the need for magic, the water, given a reasonable airflow from the movement of the substrate, will have evaporated in less than 1s so the levelling and evaporation will take place with no problem.

This leaves us with the one issue that is of relevance: the forces involved in splitting the ink. At this point, the convention is to start talking about tack. I do not want to follow convention because it is evident that most of what is written about tack, in the context of plate-to-blanket and blanket-to-substrate splitting is wrong, for the simple reason that the timescales are so small.

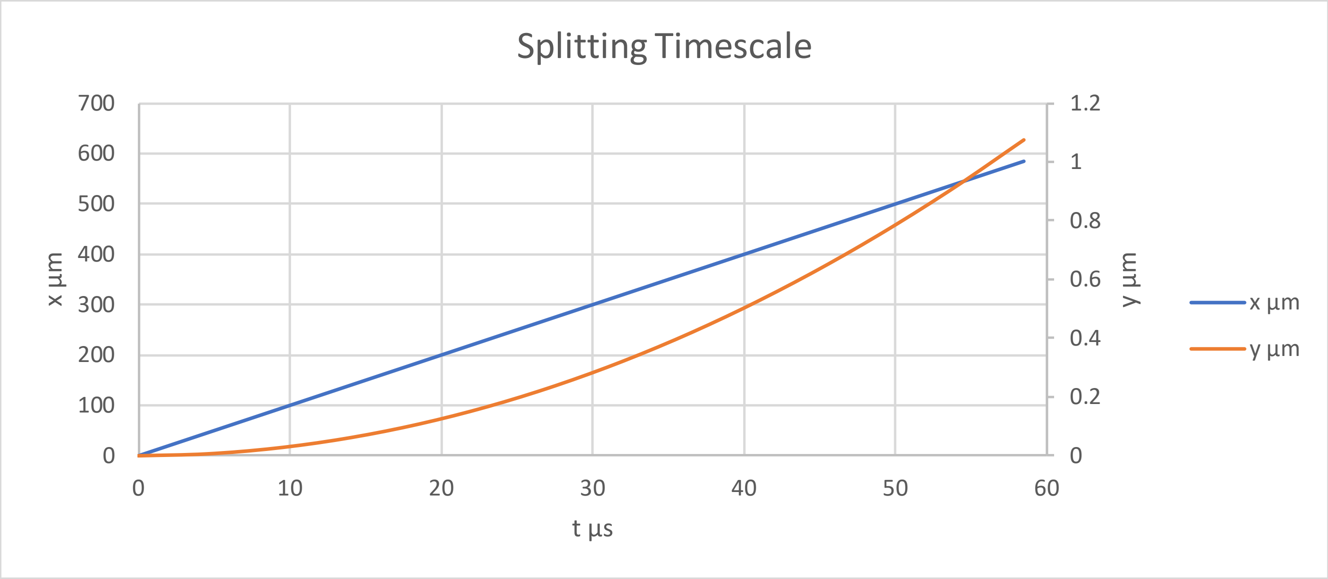
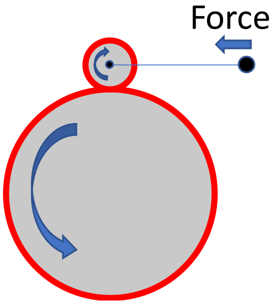


Figure 4-4 Movement in the x and y directions during 60µs. Notice the very large differences in x and y scales.

Taking our 1m circumference setup at 10m/s then it takes 55µs for the film to separate by 1µm (50% extension of the film or 100% extension of the split film), more than enough to consider that the 2µm film is "split". Because the first 15µs produce no effective separation, the real splitting time is closer to 40µs. It is interesting to note that a feature of ~400µm in the x-direction will be undergoing this split. This difference between the two dimensions is at the heart of the idea that splitting cannot be the nice meniscus opening phenomenon shown in most splitting diagrams (including mine in Chapter 1). It is much closer to the catastrophic, cavitational splitting discussed in more detail in the Flexo chapter and has close parallels to similar splitting phenomena within pressure sensitive adhesives. Those who argue that a PSA is much more viscous, and that splitting times are slower than in offset, are forgetting the temperature = time, so high speed offset splitting is equivalent to a normal splitting done at a much lower temperature, where the viscosity/modulus will be much higher. This point brings us nicely onto the rheology.

In the offset literature I have seen the rheology (G' and G") measured up to "high" speeds such as 100/s. And in terms of viscosity measurements, values at, for example, 2500/s are used for the shortness ratio (the ratio of yield stress at 2.5/s and viscosity at 2500/s). If the splitting is taking place in 40µs then the equivalent timescale is $10^6/40$ = 25000, a factor 250x larger than typical rheometry timescales and a10x factor for viscometry.

Although there have been brave attempts to measure splitting forces directly using parallel plates, the acceleration required for comparable timescales is daunting - 500m/s² (I have been told that modern "extensional rheometers" can manage 200m/s²). Some papers on drops splitting at relatively high speeds are fascinating but irrelevant because the drops are in the µl range that makes them readily visible in a camera rather than the pl range typical of printed 2μm ink layers. As pointed out in the discussion of splitting Modes, the geometry can be as important as the rheology. So we have little idea of what is really going on with any real ink at any relevant timescale.



The standard answer to this problem is to use a tackmeter, where the force exerted by the ink on the floating measurement roller has some relationship to the force needed to split the film in the nip. The problem with tackmeters is summed up nicely in a key paper by Lyne and Aspler[17] that will be cited a few other times: "Unfortunately, laboratory [tackmeters] do not measure tack force, but rather a complex combination of tack force, ink viscosity, and the rheological properties of the elastomeric covering on the measuring roller."

It has been especially noted that tackmeters are prone to "slippage" at the interface, especially when attempts are made to include water to simulate the emulsified ink.

Lyne and Aspler are convinced that the ink at the moment of splitting is basically elastic, and rheology shows, even at a 100/s measurements that there is a large G' element, though G" is still larger. Torque measurements of ink emulsification show (as mentioned earlier) the slight *increase* in viscosity created whenever an emulsion is formed. Rheology measurements of emulsified inks show a *decrease* in G', following the intuition that fount drops should reduce the elastic modulus. One paper quotes a 20% decrease in G' with a 20% water content. There is much confusion. It can be resolved, to some extent, using the approach of Palierne[18] (tweaked with the modification of Pal[19]) to take a known rheology curve with its G' and G" values and compute the effect of adding a φ volume fraction of an emulsion which shows pure viscous behaviour. The output is a new pair of G' and G" curves which show (usually) that the effects on G" are modest (so the emulsion effect is small when checked in a viscometer), while the effects on G' are significant, providing a large increase in G' at low frequencies and a modest decrease in G' at high frequencies.

The Palierne app is not for the faint-hearted. The calculations use complex factors (complicated and involving complex numbers) which are hidden from the user. The real problem is providing the input G' and G" curves. After much

---

[17] MB Lyne and JS Aspler, Ink-Paper Interactions in Printing: A Review, Ch 20 in *Colloids and Surfaces in Reprographic Technology*, ACS Symposium Series Vol 200, 1982

[18] JE Palierne, *Linear rheology of viscoelastic emulsions with interfacial tension*, Rheol Acta 29:204-214 (1990)

[19] Rajinder Pal, *A New Linear Viscoelastic Model for Emulsions and Suspensions*, Polym. Eng. Sci., 48:1250–1253, 2008

experimentation I found that the least bad method was to specify the low and high frequency range, then to set three (quasi-Prony) factors for G' and for G'' that influence the low, middle and upper parts of the curves.

Once you have the representative curves, enjoy playing with the three factors that really matter:

1. The volume fraction, φ. This is related to the emulsification curves, with some ink/fount combinations allowing a higher φ than others.
2. The drop radius. The choice is really rather limited. The emulsion drops *must* be in the 1μm range, otherwise they could not exist properly in the 4μm film that gets printed.
3. The interfacial tension between drop and matrix. This is the surprising factor which speaks to the issues of why adding IPA or its substitutes might have surprising effects on the overall process. My reading of the literature is that this effect (which is significant) is little known.

Palierne Rheology

G'_m0 -500

G'_m1 -600

G'_m2 33500

ω_{from} 0.5

G''_m0 1220

G''_m1 -6000

G''_m2 100000

ω_{to} 2.4

φ_d 0.2

R μm 1

σ mN/m 10

App 4-1 https://www.stevenabbott.co.uk/practical-coatings/Palierne.php

In the app, G'_m is that of the ink and G' is that of the 20% emulsion. Note the ~10x increase in G' at low frequencies and the 25% increase (not obvious from the log-log plot so you need to use the mouse readout) at the highest frequency.

Whether the Palierne theory is helpful depends on your attitude to tack. If you regard it as something that can only be measured with a tackmeter, then rheology is irrelevant. If you think that rheology is (or should be) relevant, then it is useful to have some fundamental insights into the effects of the emulsion.

This brings us back to the shortness ratio, yield stress at 2.5/s to viscosity at 2500/s. There is a steady stream of evidence that it says something about the printability of the ink. I have never seen a reason why this specific ratio should be of *any* value. Nothing in offset (other, perhaps, than ink duct flow) has any relation to yield stress at low shear rates and clearly tack is *not* directly related to pure viscosity (G''), because it is a G' phenomenon. But a paper by Bassemir and Shubert[20] tells us that in many cases it is the dependence of the shortness ratio on % emulsion that can be a better guide than the Surland curves themselves. Bassemir shows over a number of papers that emulsification slightly decreases the 2500/s viscosity and significantly increases the yield stress, so the shortness ratio of emulsified inks increases significantly. If this increase is out of control, then, despite an acceptable emulsification curve, the print shows scumming and a tight water balance. What has this to do with splitting and tack? There seems to be agreement that shorter inks (i.e. ones with more structure/tangles), up to a point, show cleaner splitting that gives nicer-looking prints, though there are always exceptions because like is not always being compared to like. For those who remember the PSA splitting modes from the Core chapter, a shorter ink takes us closer to Mode 1, with a short, sharp split, rather than a messy, stringy split. Too much of a good thing is always a bad thing - make it too short and it might just split away from the blanket or substrate or previous dot of ink on the substrate.

I am the first to admit that these recent sections are highly unsatisfactory. That is partly because the problems are genuinely difficult. It is also partly due to my own lack of understanding. My belief, however, is that it is mostly due to decades of wasted opportunities, of funding going into phenomenological studies with no clear scientific hypotheses to be confirmed or refuted. The scientific questions are relatively clear and a few more papers focussed directly on these questions rather than on vaguely related issues such as emulsification curves, would help us all to better understand, and therefore control, what is going on.

4.4.2 Dot gain

Within offset, the use of Stefan's law in terms of the forces involved in splitting is rightly, discredited. As discussed in the Core Concepts chapter, this simple law works very well to describe how relatively thick layers of fluid are squashed in compression or (the original intent behinds Stefan's paper) resist being pulled apart over relatively long timescales. It is not useful for the splitting conditions of offset. However, it is a useful shortcut for pointing out the obvious thing about dot gain - that most of it takes place from plate to blanket. Stefan's law is a subset of the Poiseuille flow law which tells us that for a given applied pressure, the flow of a fluid depends on h^3/η. Doubling viscosity, η, not surprisingly, halves

---

[20] Robert W. Bassemir and Fred Shubert, The Rheology of Lithographic Inks and Their Press Performance, TAGA Proceedings, 1985, 298-308

the flow. Halving the thickness, h, reduces the flow by a factor of 8. So when our 4µm thick ink dot on the plate encounters the blanket it can flow (give dot gain) of amount X, and when the 2µm dot on the blanket meets the substrate, its flow is X/8 so the dot gain is very small.

The h^3 effect is so powerful that even modest changes to the system can make a significant effect on dot gain. If the ink absorbs a lot of fount then for a given amount of pigment to be printed the total (ink+fount) has to be thicker so dot gain is larger (any viscosity effect will be small by comparison to h^3). If a more intense ink can be used, and if on-press control of ink delivery is good (the h^3 effect impacts every process) then the thinner ink deposit generates less gain.

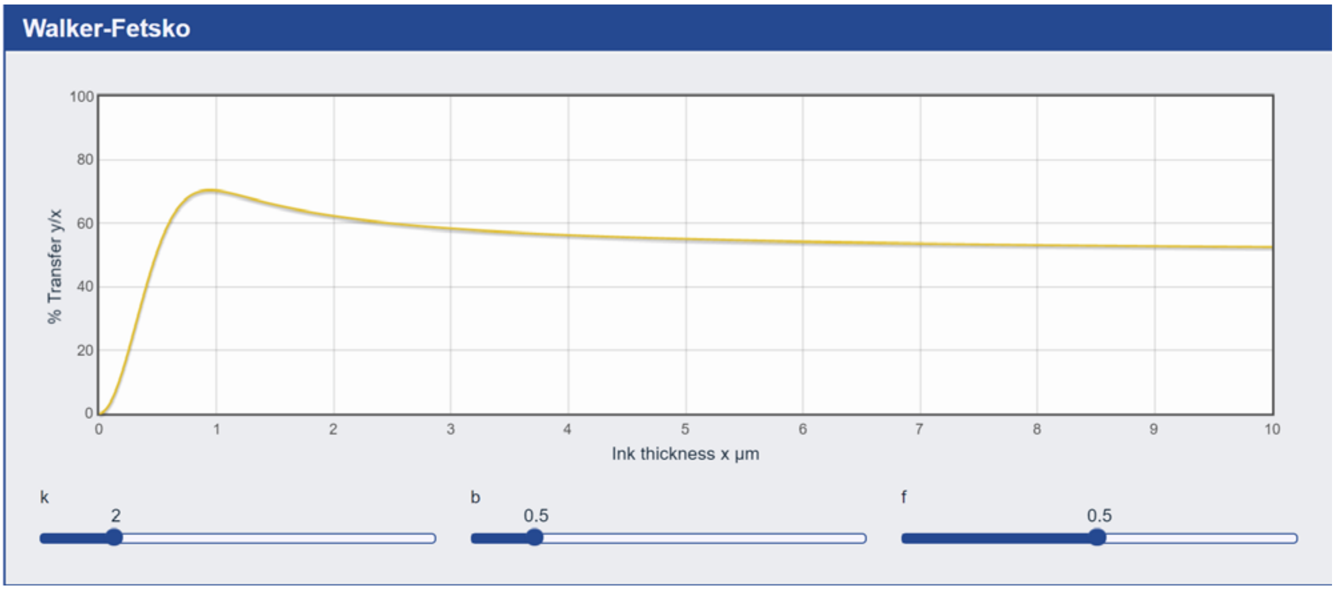
The fact that the ink needs time to flow also tells us that higher-speed printing gives lower dot gain, a welcome bonus.

4.4.3 % Split and Walker-Fetsko

The drive to thinner ink deposits for lower dot gain has a direct impact on the issue of split from blanket to substrate. The whole issue is dominated by the split onto paper substrates because with smooth substrates the result is straightforward: as long as the ink and substrate are in good contact, there is no reason for the split to be anything other than 50:50.

As soon as paper is involved there are three splitting zones. First, at very low ink deposits, there is a good chance that some parts of the ink will make no contact with the paper, and without contact there is no transfer. Next, on porous papers (though not too porous) the ink can start flowing into the paper, so when it comes to the real split which remains 50:50, that is 50:50 of the ink in the gap to which should be added the ink already in the paper. The caution about "not too porous" is that if it is too easy for the ink to go in, it can also come out as the blanket and substrate separate. The third zone is when there is so much ink on the blanket that the small amount absorbed into the paper is almost irrelevant and the split asymptotes to 50:50.

There have been many papers on the issue, but because they all show the same general effect and because there seems little actionable benefit from any of them, we can show the one most frequently used, Walker-Fetsko.



k 2

b 0.5

f 0.5

App 4-2 https://www.stevenabbott.co.uk/practical-coatings/Walker-Fetsko.php

The app shows the % ink transfer, y/x, for ink thickness x on the blanket. It requires three inputs. The split ratio, f, should be 0.5 for a smooth blanket and substrate as there is no reason for any other ratio. For a rough surface that induces cavitation it might be that the split is nearer the substrate so f might be 0.4. The parameter k controls what happens at low values of x. If the substrate is perfectly smooth and blanket contact is perfect then k is large (in practice values larger than 10 make little difference) so the ink transfer is close to ideal even for very small original thicknesses on the blanket. For a substrate such as paper, contact is poor for small ink thicknesses, so k is small and the transfer is low. The parameter b takes into account absorption into the substrate. For a smooth polymer or a highly-coated paper, b is 0. For a porous paper, b might be 1. The effect of a higher b is to give a higher effective split ratio: you still have the factor f controlling the actual split, but must add the extra ink inside the surface. Walker and Fetsko used this logic to construct their equation:

$$y = \left(1 - e^{-(kx)^2}\right)\left\{b\left(1 - e^{-x/b}\right) + f\left(x - b\left(1 - e^{-x/b}\right)\right)\right\}$$

Equ. 4-3

The ideal ink transfer formula would have just a few parameters that relate directly to the printing setup. Walker-Fetsko is not ideal. For example, k and b are inter-related (you cannot have a high k and high b), f can also depend on whatever influences b, and press pressure and speed both influence k and b. Then ink viscosity and whatever influences tack/cavitation affect all three parameters on rough surfaces. So Walker-Fetsko is not much practical use. It is, however, a familiar tool across the offset industry and none of the many alternative formulae have made much impact. Use the general ideas to identify where you might be in the curve, without worrying too much about the specific values.

4.4.4 What can go wrong?

If offset were about ink transferring from plate to fresh blanket and blanket to fresh substrate, there would be few issues to worry about. Instead we have the two problems of ink transfer from plate to blanket with ~50% of the original ink, and transfer from blanket to substrate that (apart from the first colour down) is partly covered by fount solution and previous printed dots.

Our first generic problem arises if any of the three relevant surfaces develops a thin layer of fount before encountering fresh ink. A thin layer between fresh ink and the 50% remaining ink on the blanket could, in principle, split at the point where the layer of fount appeared - meaning that it would split at the original 50% or the new 33% point. A thick layer of fount on the substrate or a previously-printed colour would mean essentially no new ink being printed.

Our second generic problem appears if the splitting forces during the blanket-substrate separation are sufficient to pull out fibres from the substrate. These fibres can pass up to the plate itself. If the fibres have pulled out some parenchyma cells ("ray" cells, shown in the little diagram) these large cells cause a more significant problem as they are relatively hydrophilic and can give a local fount-rich spot on the plate, causing a missing-ink defect in the print.

There is something of a competition between these two effects. An ink that can readily accommodate an extra bit of moisture will avoid one defect while being better able to flow into and around fibres in the substrate making it more likely that they can be pulled out. Although the conventional discussions dispense the true, but unhelpful, advice that too much or too little tack is a bad thing, this does not address the real issues. Interfacial energy's influence on the interaction of the ink with the fount on the substrate could be one feature to be optimised. And a lower-tack ink able to flow into the (perhaps pre-wetted) fibres may give more fibre pull-out than a higher-tack ink which makes poor contact with the fibres. Because these effects are complex, I have no complaints that they have not been fully resolved. My complaint is that in general the various complexities are not even mentioned so whatever data is obtained is analysed against whatever limited hypotheses (if any) have occurred to the investigators.

The properties of the previously-printed dots also influence the outcome. Suppose we have a previously printed 1µm dot with properties totally unchanged from the moment it was printed. Now let it contact 2µm of ink on the blanket. If we assume that the new ink is identical except for a different pigment, the ink does not know that the 3µ of total ink is not 3µm of the new colour, so it will give a 50:50 split, returning 1.5µm and leaving just 0.5µm of new ink (plus the 1µm of old ink). If, however, the previous dot has become instantly solid, then the split is within the fresh 2µm of ink on a solid substrate, leaving the desired 1µm. So a slight difference in drying or absorption could give a 50% reduction in ink deposit

If there is previously absorbed ink (ignoring the dots on top) then there is no possibility of a Walker-Fetsko peak compared to an unprinted area where the standard absorption can take place.

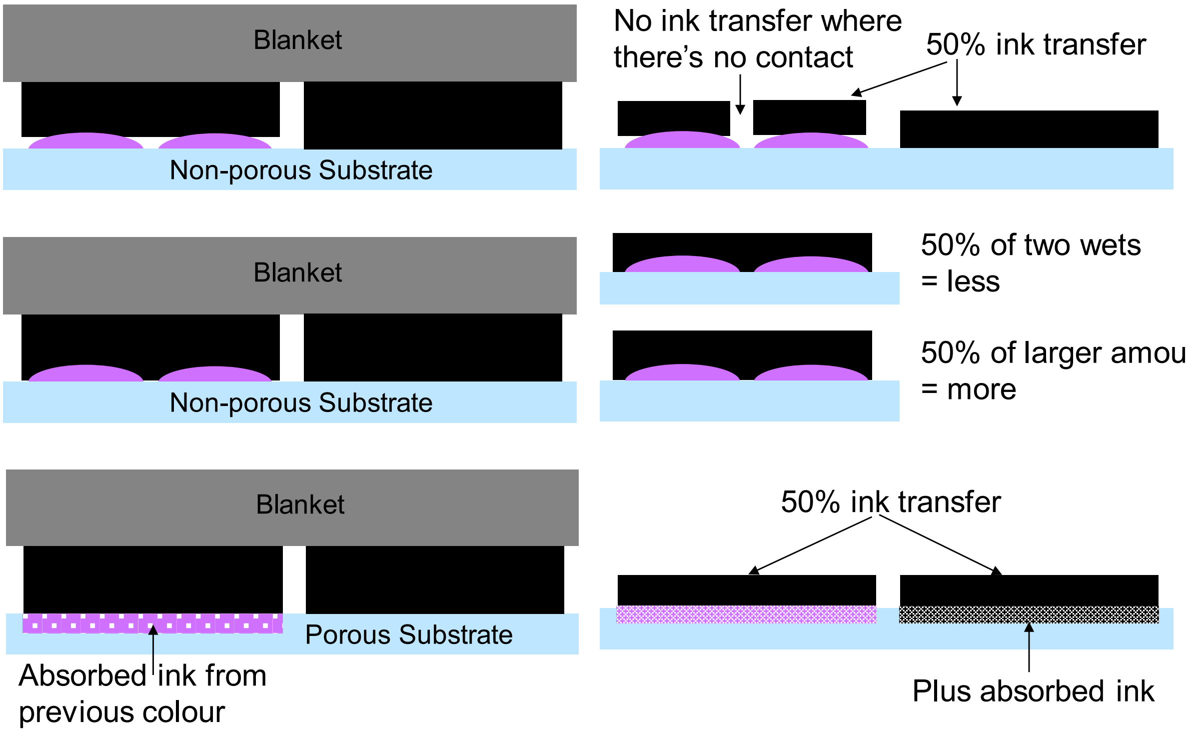


Figure 4-5 Some possible trapping outcomes.

The figure shows some of the possible outcomes. This change of printed amount on a previous dot compared to the amount on plain substrate is called trapping and depending on the specific case gives a higher or lower deposit. Correcting for it in pre-press is possible if the printing system is in a stable regime. If everything is at the borderline of printability (e.g. the different splitting regime for wet-on-wet and wet-on-dry) then trapping can shift significantly, making the pre-press adjustments invalid. Because so many published papers are phenomenological, attention is seldom paid to whether the situation being studied is in a stable domain (e.g. plenty of time for the first colour to dry) or an unstable one (borderline drying). So conclusions based on broad terms such as tack measured on a tackmeter are usually unhelpful or wrong.

4.4.5 Blanket (in)compressibility

We have been assuming that the blanket was always in perfect contact with the plate and with the substrate. In real life, the blanket has to accommodate variations in thickness so it has to be compressible. The obvious, but incorrect, intuition is to have a nice thick, soft rubber blanket that can absorb all the bumps in the process. Apart from the fact that the surface needs to be relatively hard to survive the process, this squishy blanket idea suffers from the fact that it goes against the laws of physics. The only way to get it to work is to crank up the nip pressure, with all the downsides that that causes.

Here's the problem: an ideal rubber is completely incompressible and a real rubber is hardly compressible.

A simple equation explains this fact. Rubber is, of course, easily stretched, with a tensile modulus E with which we are all familiar. So what is the modulus, K, you would measure if instead of stretching you tried to compress the rubber? Here is the answer:

$$K = \frac{E}{3(1-2v)}$$

Equ. 4-4

The v on the bottom of the equation is the Poisson ratio, the amount the material moves in one direction if you pull it in the other. For a perfect rubber, v=0.5 so K is infinite (E/0) and for a real rubber we can give it a value of 0.499. This means that K=E/(3*0.002), i.e. 166E, so the rubber is hardly compressible. To put it into words, as you try to squash the rubber locally to accommodate some imperfection in the system, it has to go sideways, but the rest of the blanket is in the way so it cannot easily deform.

Given that the press must accommodate the imperfections in the system, what strategies can be adopted? The printer's instinct is "higher pressure is better", especially given the false intuition that higher pressure will force more ink to be transferred. The blanket does have *some* compliance so this strategy will work at the price of unnecessary stresses on the machine and larger dot gain. The physics tells us that a foam *is* highly compliant, so a thin foam layer combined with a thin blanket is the ideal combination. That is why most modern blankets are foam-backed.

Those who want a fuller discussion on foam-backed rubbers can read the relevant section in the Flexo chapter where the myth of "a thick rubber will absorb non-uniformities" is demolished with some real experimental evidence.

4.5 Summary

Maybe none of this matters. Offset is wildly successful despite the myths believed by many and the numerous complications unearthed when the process is thought through scientifically. Why should anyone bother to research difficult issues for, probably, little reward in terms of improved inks or founts? In gravure and flexo there is a modest research community using good science because they are forced to by their need to find improved methods of printing electronic devices. The case for printed electronics via offset seems weaker so there are fewer papers providing new insights.

But if the thoughts here inspire just one harried formulator to come up with a better solution to a tricky problem, then I am content.

5 Flexo

Flexo has an image problem and a linguistic problem.

The image problem is that it is a crude process, suited for cardboard boxes and little else. The very real faults in printed flexo images assist this image problem.

Fortunately, the image problem is receding because the many strengths of flexo mean that it is continuing to grow in use. The growth in capabilities is largely due to the refinements in the plates themselves. The intrinsic shortcoming of the printing step are a more formidable challenge and are the focus of this chapter.

The linguistic problem is a hangover from its past when the only viable inks for the process were based on aniline. So we ended up with the name "aniline printing" and the process included a trade-name, Anilox, for the roller that delivered the ink to the plate. When aniline inks became classed as being toxic, it was not good business sense to keep calling the process aniline printing. It is interesting to read in Wikipedia that although the term Gummidruck ("rubber printing") has been in long-term use in Germany, various marketing attempts at names such as Lustro and Transglo failed to catch on, and it needed an industry-wide process in 1951 to come up with a choice between three names: Permatone; Rotopake; Flexographic, with an overwhelming majority for the last of these.

There are three steps in flexo printing which we will discuss in turn:

1. Creating a controlled amount of ink within the cells of the anilox roller
2. Transferring some proportion of that ink onto the raised portions of the rubber printing plate
3. Transferring ~50% of that ink onto the substrate in the printing stage

As with offset, the making of the plate will not be discussed as a separate topic. Unlike offset, the physical features on the plate are of considerable importance, so those aspects of the plate-making process which affect the physical features, and therefore the print, will be included.

5.1 Metering the ink

As with offset, the aim is to get ~2μm of ink onto the surface of the plate so that with a 50:50 split onto the substrate the final ink is ~1μm. Instead of the large, complex train of ink rollers, in flexo the metering is done by filling the cells of the anilox roller and scraping off any excess. This process (and the subsequent transfer to the plate) cannot happen at relevant speeds with viscous inks, so the inks contain lots of solvent (which may be water) which is subsequently evaporated.

Although the ink can be taken out of a pan by a roller and then transferred to the anilox roller, it is better to describe the more rational chambered doctor blade method.

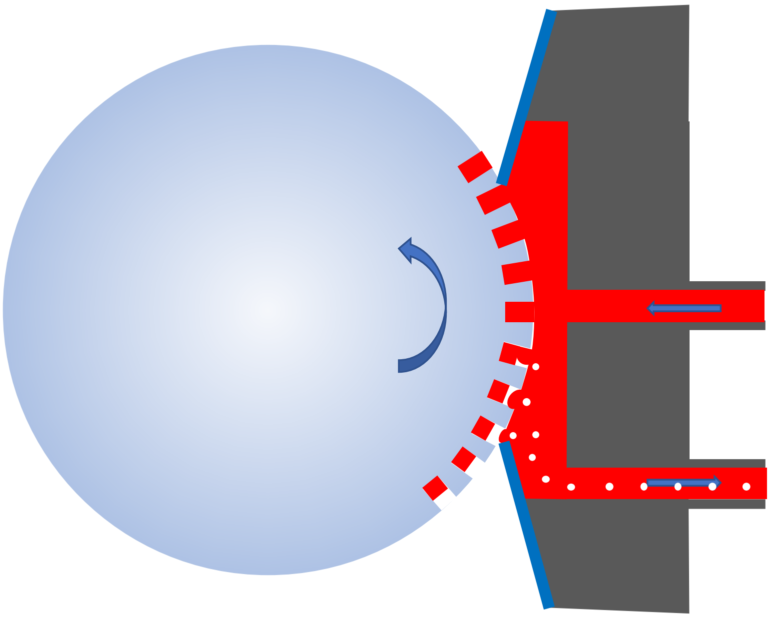


Figure 5-1 The chambered doctor blade and the messy act of filling the cells which necessarily adds bubbles to the chamber.

Simple diagrams show the roller in contact with the ink in the chamber with the "walls" of the chamber defined by two flexible doctor blades. The lower blade simply defines the bottom of the chamber, the top blade also scrapes off any excess ink from tops of the cells and from the roller surface between the cells. The ugly (and not to scale!) diagram here shows the messy reality that is seldom talked about. Everyone likes to say that the chamber fills the cells with ink. What is not usually mentioned is that the air in the returning cells (shown as being 30-50%-filled, as discussed later) has to escape when the cell is filled. The bubbles of air are potentially catastrophic if they invade the upper parts of the chamber. So the flow of ink should be as shown with fresh ink having a chance to push out the bubbles into the return loop which, of course, must have some way to remove most of the bubbles before being pumped back in. The diagram is a great oversimplification. Real-world systems require considerable sophistication, involving 3D flow with fluid coming in through a single pipe in the centre and air-filled fluid flowing out via two pipes at the edges.

The whole process of filling, removing bubbles and scraping off excess is far easier with lower viscosity inks.

Why are the blades called "doctor blades" instead of, say, "scraper blades". Wiki says that the name derives from "ductor blades", because at one time they were used to wipe the ductor rollers that "lead" inks to the distributor roller in letterpress. Others say that to "doctor" is to remove defects, and excess ink can be thought of as a defect.

I know of only three bland rules for doctor blades:

1. They should wear out (they are cheap) slightly faster than the roller (which is expensive)
2. They should be sharp, free of defects and applied with uniform pressure
3. That low blade pressures are bad because ink can leak through and that high pressures are bad because they cause the blade to twist, allowing the ink to leak through.

I always find it surprising that the direction of the blade makes little difference, though for high speeds the resistance to hydroplaning is greater when the blade is pointing (as with the top blade in the diagram) against the direction of rotation.

The general effects of the size of the chamber, of speed, viscosity and blade angle on the pressures and (sometimes) circulation zones (bad) within the chamber have been modelled[21] by the team at WCPC, showing, fortunately, that our intuitions are correct and that the key pressure field before the exit doctor blade is highest (which presumably helps ensure filling of the cells) near the blade and is higher for higher speeds, higher viscosities, higher blade angle and smaller chamber. The paper acknowledges that it would be good to be able to understand how/where the air escapes but that such a study would be immensely complex.

Now we need to decide what the roller surface should be made from (the basic shell is steel), and what shape and size of cells to use.

The only criteria for the surface are that it should be easy to create accurate cells, whilst preserving a smooth surface between the cells, and that the surface should be resistant to the constant scraping of the doctor blades.

In practice that means chrome or ceramic. There are many solemn discussions about the respective surface energies and wetting characteristics. These are, of course, totally irrelevant because at flexo speeds, surface energies are many orders of magnitude too small compared to viscous forces. The choice is one of cost and convenience in having the shape/size/volume of cells that you require. The undoubted differences between the "same" cell pattern on chrome and ceramic is due to the fact that the patterns are *not* the same - there will be differences in many of the subtle (and not so subtle) properties depending on the precise nature of the process that creates the cells within the roller.

A reader kindly pointed out one issue that I'd missed. The distance from the doctor blade to the printing nip can be too long for a volatile ink and a hot surface from the action of the doctor blade. If the ink starts to dry in this short

---

[21] E.H. Jewell, T.C. Claypole, M.F.J. Bohan & M. Von Grol, Modelling the Flow in a Flexographic Ink Chamber, TAGA Proceedings 2000, 497-502

space then either the heating from the doctor blade has to be decreased or the volatility of the solvent has to be decreased.

5.2 From anilox to plate

Here we hit the big problem There is no rational theory describing how the different shapes and sizes of the cells combine with the rheology of the ink to deliver a known amount onto the plate surface. This is not through lack of trying. There are plenty of papers reporting on "correlations" with different parameters. I have read many of them and the results are all the same: "it's complicated so you should assume 40% unless your ink is way too viscous or your cell is way too deep". A welcome exception to this trend is the very insightful paper from the Swansea team[22] which does careful measurements and shows that the 40% is OK for 100% solids but that in some circumstances a 30% dot can pick up 70% by dipping into cells either side of the target. The discussion that follows does not take such extra effects into account. The paper also points out that speed made no obvious difference to % transfer.

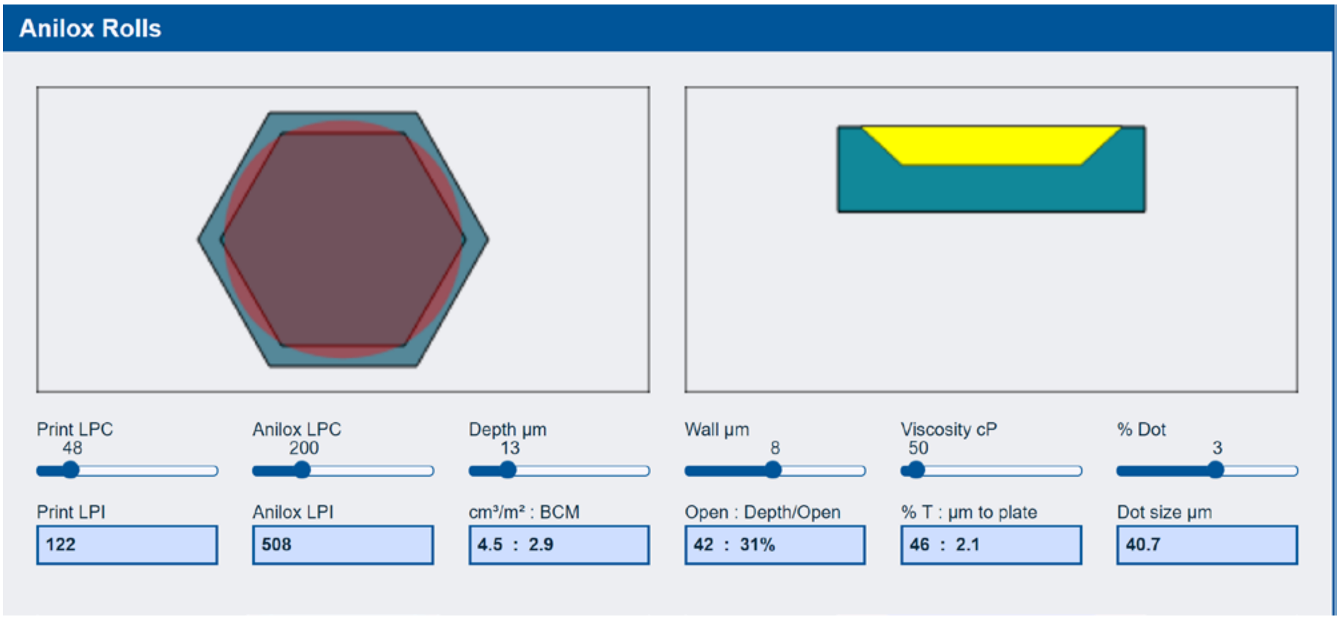
If we go with the 50% then our question is "50% of what". One of the world's spectacularly bad units that the US like to invent is BCM, billions of cubic microns per square inch. I don't know why the absurdity of this unit has not caused it to instantly fall out of use. The BCM part is metric and the unit, as stated, does not include the fact that it is per square inch, an archaic unit of area; so it should be BCMpSI. And there is no obvious link from this to how much might appear on your plate. The more sensible unit is cm^3/m^2 (or, equivalently, ml/m^2). That happens to translate directly to µm, so 50% of $4cm^3/m^2$ means that 2µm get transferred to your plate and a further 50:50 split means 1µm on your print, which is a typical target. BCM values are 0.65 smaller (i.e. a BCM is larger), so you need a 2.6BCM for a 1µm print.

A standard 60° hexagonal cell is a good general-purpose default. Because UV flexo inks tend to be higher viscosity (the best acrylates for good cured properties are usually not the ones with the lowest viscosity), some alternatives might be required, with trade-offs necessary against two other print properties: moiré and minimum dot. The moiré issue is discussed in a later chapter. So what about the minimum dot?

If a dot on the plate is smaller than the size of the anilox cell then it can "dip" into the cell and will print a messy dot. By going to a high line-count cell, you avoid the dipping problem, but to achieve the required cm^3/m^2 you need very deep cells which might give less than 50% release. If you are happy that the smallest printed dot is, say, 4% then you can use a lower line-count compared to the choice of 1% dot as the minimum. All these trade-offs become apparent with the app.

---

[22] John Cherry, Tim C. Claypole and David T. Gethin, *Measurement of the ink release from the Anilox Roll,* 2006 TAGA Proceedings, 395-408



Print LPC 48

Anilox LPC 200

Depth µm 13

Wall µm 8

Viscosity cP 50

% Dot 3

Print LPI 122

Anilox LPI 508

cm³/m² : BCM 4.5 : 2.9

Open : Depth/Open 42 : 31%

% T : µm to plate 46 : 2.1

Dot size µm 40.7

App 5-1 https://www.stevenabbott.co.uk/practical-coatings/anilox.php

At a low Print LPC it is relatively easy to find an Anilox LPC setting that with a modest 13µm depth gives a theoretical 4.5cm³/m², i.e. 4.5µm which, with a 46% transfer provides the ~2µm on the plate required for a 1µm final print. But only if the smallest dot is 3%. At 2% the dot is 33µm which is smaller than the 42µm cell opening. So what choice do you make? A higher LCP anilox requires a deeper cell for the same nominal volume, but the % transfer will start to decrease. And this is for a typical 50cP ink. A higher viscosity, e.g. for UV flexo, will have a lower % transfer making the problem even harder. Cell shapes that extend the hexagon in one direction allow higher % transfer so they are popular for UV.

The Depth/Opening ratio is critical. When it is too large (>33% as a general guideline) then it is harder to make the cell, the % transfer decreases, and it is more likely that ink will accumulate at the bottom of the cell and be difficult to clean. When it is too small (<23%) then it becomes difficult to make the cell to the required accuracy as a small absolute error in cell depth becomes a large % error.

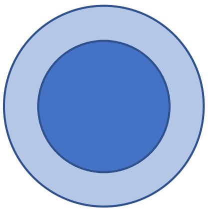
The app contains many simplifications but the numbers map well onto the figures provided on many of the anilox suppliers' websites, so they are a good starting point for discussion and for the whole team to understand that a decision about one part of the system (e.g. the Print LPC) has many implications for other parts (e.g. the minimum % dot).

And for UV, the alternative to a different anilox design is to run everything at a higher temperature which significantly reduces the viscosity.

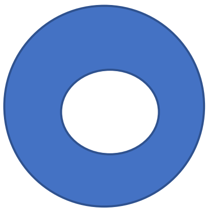
5.3 From plate to substrate

Up to this point, flexo was doing OK apart from the compromises around the anilox cells and the finest printed dot. The real flaws in flexo appear in the transfer process. The familiar diagrams of this step are rather poor at describing all the problems. So I will use some screen shots from a simple animation I created many years ago when I worked for MacDermid (one of the major flexo plate manufacturers)[23]. The effects are exaggerated for visual clarity but they do a good job at describing the key flaws in the process.

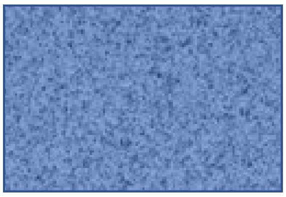
Our problem is to describe four classic flexo defects. The first, large dot gain, is easy. The second is donut dots, the phenomenon where, say, 30% dots have a hole that is usually a bit skewed. The third is mottle which can be so severe for 100% printed areas that printers choose a 98% solid instead which seems to give a more solid solid than 100%. The fourth is the edge line or "halo" often found around printed lines or text.



Dot Gain

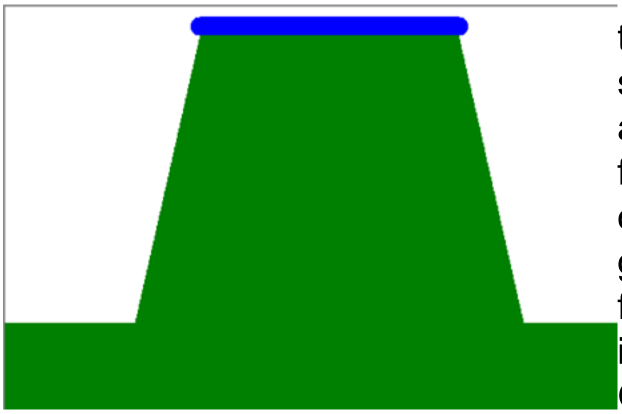


Donut Dot

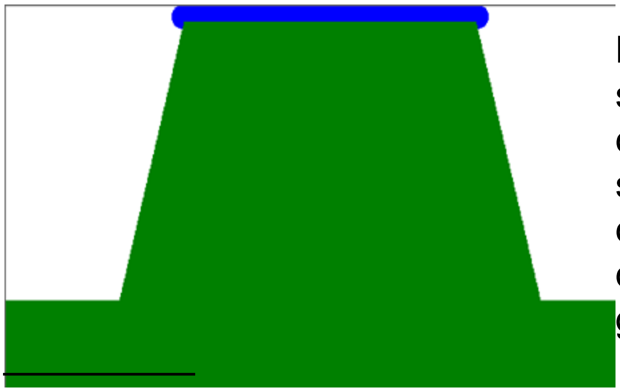


Mottle

Halo



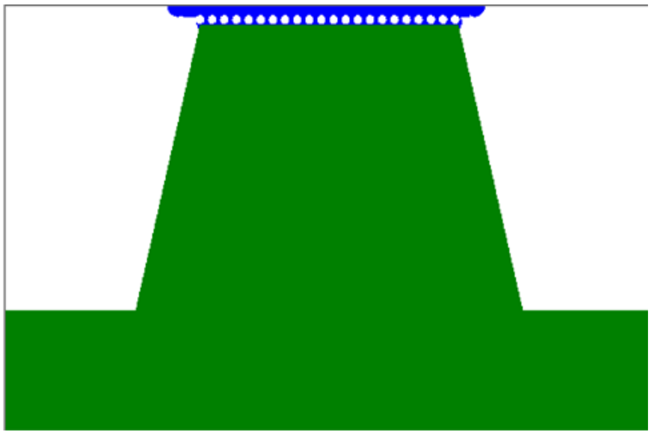
We start with the blue ink on the surface of the rubber dot just about to touch the substrate which we can imagine as being along the top of the image. Because ink from the anilox can flow slightly, it is shown overhanging the edge of the dot, ready to give a little bit of dot gain. However it is not flowing onto the side of the dot - it is inhibited by a phenomenon known as Gibbs pinning.



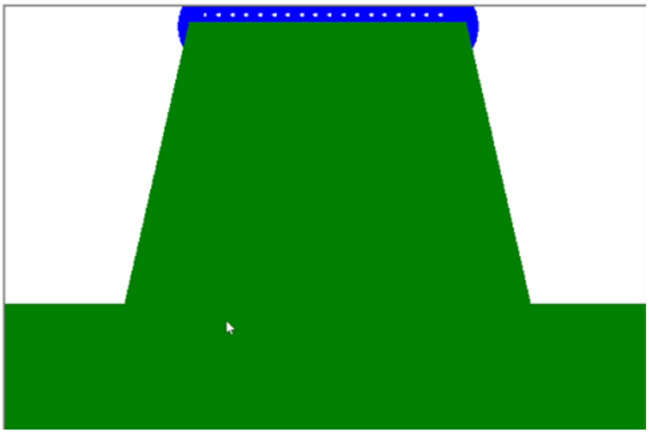
Now the dot comes into contact with the substrate. Two things happen. The rubber dot itself spreads a little and the ink is also squashed out. This gives us two forms of dot gain and, at the same time, pushes ink down the side of the dot, a problem that will grow with repeated prints.

---

[23] Readers can download it (with kind permission from MacDermid), along with the Screen modeller from www.stevenabbott.co.uk/_downloads/SPAandFlexo.zip.



As the plate is pulled away from the substrate we see that the rubber dot returns to its original shape and the dot gain in the printed ink is rather clear. More seriously we see that the ink has had to split in the only way it can - via cavitation. So the printed dot starts off very rough, ready to produce mottle. Although a lot has been written about flexo mottle (and we will discuss it further shortly), only rarely has the issue of cavitation been identified as a key driving force for the process[24]. When the dot is rather small, the cavitation forces are concentrated in the centre so there might be one big cavitation bubble which can cause the classic donut dots seen so often in flexo. There is some directional asymmetry from the process, so the hole is slightly off centre and skewed; it can even become a half moon. An alternative explanation for the donuts became apparent when I was researching the gravure chapter. It states (and will be discussed in more detail) that the filling process leaves a dip in the liquid in the middle of the cell. This means that that part of the liquid does not contact the substrate, leading to the missing bit of ink. The cause of the dip naturally produces it off-centre and non-circular.



Here is another shot which happens to show the cavitation process, caught a little earlier on in the splitting. The point of interest in the image is the ink going down the edge of the dot. This image was taken after a few cycles of printing so that the small amount squeezed out each time has accumulated to produce a rather large deposit of ink. With one more bit of physics we can see how this might cause the edge line, or halo, around letters.

One ideal for the rubber dot would be that it was so rigid that it contributed very little to dot gain. However, this is impractical for every-day printing surfaces. A soft dot gives more dot gain. Worse than that, if the angle of the dot and the pressure combine in the wrong way, the dot can "barrel", forcing the edge upwards and depositing the ink along the edge of the print. Barrelling is more likely for linear features than round ones - it's easy to buckle a linear fold of

---

[24] Whether it is "cavitation" or "viscous fingering" is not such an important debate. This paper from TU Darmstadt shows that the Saffman Taylor viscous fingering approach helps explain mottle. Hans Martin Sauer, Dominik Daume, Edgar Dörsam, *Lubrication theory of ink hydrodynamics in the flexographic printing nip*, J. Print Media Technol. Res. 4(2015)3, 163–172

paper but not a cylinder of paper. The barrelling phenomenon has been nicely described by the Swansea team[25], though in the context of dot gain rather than the halo effect. With no authoritative analysis of the edge phenomenon[26], the illustration is meant merely to be suggestive of what is going on.

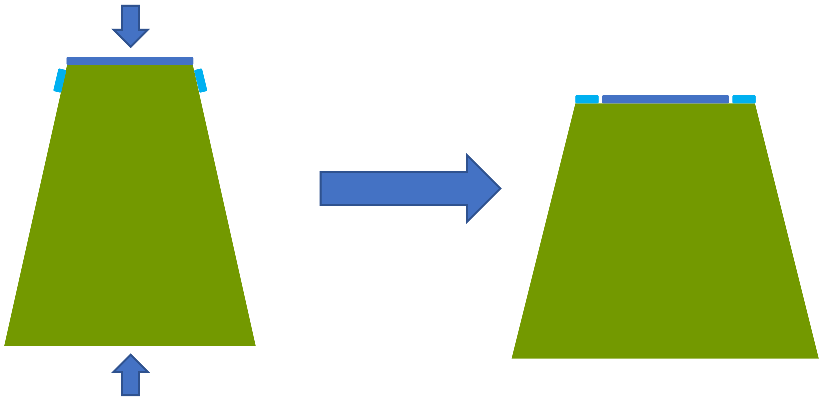


Figure 5-2 How barrelling might bring edge ink in to form an edge line or halo. The ink on the side is shown as a different colour to show where it goes when the dot barrels into its new shape. There is a natural reason for a line between the two bits of ink.

Returning to dot gain, the question arises as to how much is due to the rubber spreading sideways and how much is due to the ink spreading under pressure. Fortunately the same paper answers that question: it is mostly ink spreading. As we saw, flexo requires low viscosity ink to enable a sensible set of compromises for the anilox transfer, so dot gain is an inevitable consequence. The paper also shows that for a given dot diameter, a short, stubby dot, will give more barrelling than a long thin dot.

The Darmstadt paper gives a different explanation for the halo effect. It states that the ink in the middle of the dot can deform the rubber, but at the edge it is easy for the ink to be squeezed out, so the rubber is in contact with the substrate, accounting for the gap between the halo and the print. The Darmstadt thesis makes clear that there have been numerous potential explanations (including the "barrelling" one), whilst favouring the squashed edge idea. The proposed explanation is visually compelling but the sad truth is that currently there is no definitive evidence to support or reject any specific hypothesis and there may, indeed, be multiple mechanisms. It is hard to know how to identify the root cause(s); whether the cause is barrelling or squeezing, a harder dot would tend to show less halo.

---

[25] D C Bould, T C Claypole and M F J Bohan, *An investigation into plate deformation in flexographic printing*, Proc. Instn Mech. Engrs Vol. 218 Part B: J. Engineering Manufacture, 2004, 1499-1511

[26] An excellent thesis by Constanze Ranfeld at TU Darmstadt actually uses the halo effect for some clever Ag line printing. She describes a number of halo theories from papers I cannot get hold of. The thesis is freely available on-line: Constanze Ranfeld, *Wet etching of printed silver layers using an etch resist structured by flexography*; TU Darmstadt, 2015

5.3.1 Flexo mottle

The single word "mottle" is used in many confusing ways and sometimes what one calls mottle another calls, for example, UnCovered Areas (UCA), or, separately washboarding (on corrugated board) which is a very different issue, discussed in the next section on compressibility.

Even measuring mottle is contentious. There is the ISO 13660:2001 standard which is the standard deviation of density or reflectance measurements across a large sample, with "graininess" being defined as the standard deviation of higher-frequency variations within a measurement patch and "mottle" being the standard deviation of average measurements across many separate patches.

As this is the 21st century it is arguably better to just take a 2D-FFT of densities and from the spectral density curve, integrate the "power" in the frequency range (scaled to the human visual system) of interest, perhaps via a weighting system calibrated to give the best match between visual judgement and calculated values. This is, indeed, the preferred method (and joint best scoring method) from a paper[27] that performs an analysis of 11 different mottle measurements.

Most of the papers describing mottle on paper/board are describing the rather obvious facts

1. that if the ink cannot touch fibres in a hole in the paper, there is no ink transfer - that is the key idea in a paper on UCA[28]
2. that if the fibres/fillers form denser and less dense regions, then the ink will flow differently into them and give different densities/scattering.

In terms of the first of these, there is not much to be done, except perhaps squeeze a little harder at the price of extra dot gain in the half tones. For the second, although there are plenty of interesting things one can say about flows into fibres (some of these were discussed in the chapter on drying), the general conclusion is obvious, that a smoother coated paper will give a nicer print.

The key interest here is that flexo shows plenty of mottle when printed onto smooth polymer surfaces in the large flexible packaging market. This is an intrinsic, fundamental form of mottle that can only be addressed by understanding the physics behind it.

My reading of the literature trying to understand mottle phenomenologically (by changing plausible parameters with no specific theory) is that it is largely

---

[27] Carl-Magnus Fahlcrantz and Per-Åke Johansson, A Comparison of Different Print Mottle Evaluation Models, TAGA Journal, 2,148-160, 2006

[28] Gustavo Gil Barros, Carl-Magnus Fahlcrantz and Per-Åke Johansson, *Topographic Distribution of UnCovered Areas (UCA) in Full Tone Flexographic Prints*, TAGA Journal, 2, 43-57, 2006

unhelpful because it does not correctly identify the root cause - the defects generated by cavitation or viscous fingering as shown in the animation, nor does it discuss two relevant extra pieces of physics.

The first of these is about levelling: if the ink can level before drying, mottle will not be present. As we saw in Chapter 1, the Orchard theory tells us that the levelling time, t is given by:

Equ. 5-1

$$t = \frac{3\eta\left(\lambda / 2\pi\right)^4}{\sigma h^3}$$

The dependency on viscosity and surface tension are linear and typical values do not change much between formulations, so (omitting the highly viscous UV inks) we can largely ignore them. If the ink cavitates very easily, or the plate or substrate encourage cavitation, λ will be small and its 4th power gives us rapid levelling. And doubling the ink thickness reduces levelling time by a factor of 8, so thicker deposits should be less mottled - in addition to the fact that a 0.1μm variation in a thinner ink will be more visible than in a thicker one.

The second piece of physics is about hole opening. As we saw, a defect of diameter d will grow larger if:

Equ. 5-2

$$\frac{h}{d} < 2\left(1 - \cos\theta\right)$$

Again, a thicker ink and fine cavitation which gives smaller d combine to give more resistance to the creation of large-scale mottle. Because a lower contact angle θ helps to reduce hole opening, raising the surface energy of the substrate via, say, corona treatment might help.

With this chain of reasoning from cavitation through the physics of levelling and pinholing, it should be relatively straightforward to analyse mottle experiments. My problem with the literature is that experiments have not been based on a coherent set of ideas so have tended to be phenomenological rather than useful. Usually there is no way to transfer any learning from one set of experiments to one's own setup, because there are no meaningful numbers and formulae that can at least provide some guidance. This is, unfortunately, all too common in the world of printing, so large amounts of precious experimental data are unusable.

Apart from the one exception mentioned earlier there have been no physics-driven experiments getting to the heart of mottle. Such experiments would help us to understand the influences of solvent, pigments, plate surfaces, speed, thickness, print surface topography etc. If anyone rises to this challenge, I predict that the impact on flexo will be significant.

One fruitful scientific possibility is to try to induce lots of cavitation nuclei via a micro-roughened plate or substrate. The more cavitation points, the smaller the wavelength of density variation and, via Orchard, the faster the smoothing out to an unmottled print.

5.4 Dot on dot

If we have 2µm of ink on the plate and print onto a smooth substrate then we will get a 50:50 split and 1 µm will be printed. Now let us print directly onto a previous dot, say a C dot directly onto a Y dot. How green will that dot be? If the C dot had previously dried, then we would have the standard situation of 2µm of ink to be split, giving 1 µm and a perfect green. If, however, the Y ink was perfectly wet then we have 3 µm of wet ink which will split 50:50 to give a 1.5 µm dot. Given that we started with 1 µm of Y, this means that we have printed only 0.5 µm of C, so we will have a very yellowy-green.

Although there are plenty of presses that print individual colours with drying stages between them, it is far more efficient to have a single station (CI, Common Impression) that prints all inks (at least 4, usually 6 or 7) in one go before drying. This variable splitting issue is, therefore, a significant problem.

Hence the attraction of ingenious inks that allow wet-on-wet printing. An example are inks that gel near instantly after printing so that the next ink experiences something closer to wet-on-dry printing and, therefore, better colour control.

My reading of the literature is that there is little more useful to be said about flexo, except for the final key topic for flexo which is that of compressibility (or lack of it).

5.5 Compressibility

I remember a time when there was a large transatlantic divide about the desirable overall thickness (as opposed to the dot height) of a flexo plate. In the USA the preference was for a thick plate because this, obviously, would more readily compress and accommodate variations in substrate thickness (especially for printing onto corrugated board). This preference came at a considerable extra cost in raw materials, processing time etc. In Europe, the preference was for the thinnest possible plate because it was self-evidently more cost effective in time and material. The difference seemed to be that the Europeans better understood the laws of physics which say that rubber is effectively incompressible. This means that all the extra rubber on the US plates was not absorbing anything for the simple reason that it was not at all a nice compliant material.

A simple equation explains this fact. [Because the physics is the same, this section repeats the text in the blankets section in Offset]. Rubber is, of course, easily stretched, with a tensile modulus E with which we are all familiar. So what is the modulus, K, you would measure if instead of stretching you tried to compress the rubber? Here is the answer:

$$K = \frac{E}{3(1-2v)}$$

Equ. 5-3

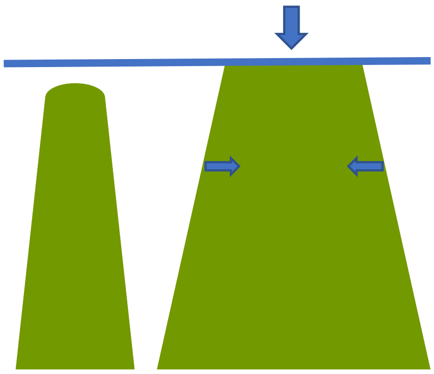
The v on the bottom of the equation is the Poisson ratio, the amount the material moves in one direction if you pull it in the other. For a perfect rubber, v=0.5 so K is infinite (E/0), and for a real rubber we can give it a value of 0.499. This means that K=E/(3*0.002), i.e. 166E, so the rubber is hardly compressible. To put it into words, as you try to squash the rubber locally to accommodate some imperfection in the system, it has to go sideways, but the rest of the plate is in the way so it cannot easily deform.

Given that the printer must accommodate the imperfections in the system, what strategies can be adopted? The printer's instinct is "higher pressure is better", especially given the false intuition that higher pressure will force more ink to be transferred. The plate does have *some* compliance so this strategy will work at the price of unnecessary stresses on the machine and terrible dot gain. The physics tells us that a foam *is* highly compliant, so a thin foam layer combined with a thin plate is the ideal combination. A thesis shows[29] that for a specific flexo setup, a plate without foam backing gave a 1% compression under 1MPa nip pressure and the foam-backed equivalent gave a 9% compression. For printing on corrugated board it is not possible to completely remove the influence of the corrugations on the print, but two pieces of academic work, discussed next, show that a foam-backed plate works well to reduce the effect, whereas a thick plate does nothing and high pressures squash the corrugations, but still show the repeating defect.

The dots themselves *are* compressible because their sides can expand - with smaller dots being more compressible, presumably slightly reducing the dot gain compared to bigger dots. The big problem with the larger dots is that their lack of compressibility often comes with the fact that the smaller dots are smaller in height - the peaks get washed away during the plate development. This was especially true for direct-written dots where the plates are exposed to UV in air. Oxygen inhibits the acrylate cure so the tops of the dots are very rounded and eroded.

---

[29] Johanna Johnson, *Aspects of Flexographic Print Quality and Relationship to some Printing Parameters*, Dissertation, Karlstadt University, 2008



These lower, smaller dots are especially difficult to print if they are near bigger, higher dots which are less compressible so stop the smaller dot from making good contact. The foam backing is not sufficiently compressible over small areas to accommodate this sort of non-uniformity so the only solution is to print with higher pressure, causing unnecessary dot gain on the larger dots.

The newer flat top dots are produced using exposure under nitrogen or a thin oxygen-excluding sheet. Because the dot heights are much more even, printing at high quality is far easier at the desirable lower pressure.

This leaves us with the question of how tall the dots should be and at what shoulder angle. For those who print correctly, at the lowest possible pressure, using a thin, foam-backed plate, using flat-top digital dots the answer is that the dots should be of small height and steep angle. This is the fastest exposure with the least wasted material, giving dots that are relatively stable to lateral movement and where the compressibility differences between small and large dots are minimized. For those who insist on printing at high pressures, well, good luck fighting all the compromises that this requires.

5.6 Printing on corrugated board

The problem with printing onto corrugated board is that a solid tone can show distinct bands of density variation, with lower density coinciding with the deformable part of the board. A thesis[30] that investigates how to minimise the banding effect fully supports the ideas in previous sections. The following applies to the printing of 100% solids

1. Glossy boards show greatly reduced banding for the simple reason that ink transfer does not depend on pressure once above a certain minimum contact pressure.
2. Rougher (uncoated) boards show stronger banding because the ink transfer depends on the degree of contact with the board fibres (Walker-Fesko), which is reduced in the areas of lower contact pressure where the board can deform.
3. Using a "softer" plate is of no benefit. The overall pressure in the nip goes down somewhat, so a larger nip impression is needed and the differences between the zones remain.
4. A foam-backed plate significantly reduces the banding because the genuine large-scale compressibility allows the whole system to respond to the variations in the board and, to a significant extent, correct for them.

---

[30] Martin Holmvall, *Nip Mechanics, Hydrodynamics and Print Quality in Flexo Post-Printing*, PhD Thesis, Mid Sweden University, 2010. The phrase "post-printing" means printing onto the board after it has been fabricated rather than printing onto one of the layers of the board which is subsequently made into board.

For half-tone printing, the visible print defect is more to do with dot gain which depends on pressure, rather than fibre contact. Although no relevant experimental data were shown, this would imply that the halftone variations would be less dependent on whether the board was coated or uncoated.

Another paper[31] does a more detailed comparison with thick and thin foam and thick, thin and absent "mounting foils" between the foam and the plate. The mounting foils either make no difference or make things slightly worse, and the thicker foam is always better in terms of high period "mottle" (i.e. washboarding corresponding to the spacing of the corrugations) and dot gain, with removal of the mounting foil giving the lowest dot gain, with little negative impact on solid or halftone density.

Although the desirable effects of foam compliance are especially obvious in corrugated board, the principle of welcoming compliance applies to the whole of flexo. A light printing pressure is desirable in every condition except when it is necessary to smash the ink into a poor-quality fibrous surface. The danger with a light pressure is that thickness differences in the substrate or runout in a roller can lead to lack of printing in some areas. With a compliant backing, you can have a light pressure *and* accommodation to changes in the nip gap.

A little bit of physics goes a long way to help produce better prints with less effort.

5.7 Summary

The compromises that are intrinsic to flexo means that it will never be a trivial process to get right. But the past decade has seen a big boost in quality thanks to steps that, knowingly or otherwise, accord with what science tells us:

1. Use a sophisticated chambered doctor blade with special attention to removing air at high speeds
2. Use tools like the anilox app to reach an intelligent compromise on viscosity, speed, minimum dot, ink strength to get an optimal cell shape and frequency in a well-etched roller.
3. Use the thinnest possible plate, with relatively hard rubber, with a foam backing for compliance and with short, sharp, flat dots for greater precision and the ability to print small and large dots in close proximity
4. For Common Impression printing, find inks that instantly gel to allow easier wet-on-wet printing

Finally, given that mottle and halos are still relatively poorly understood, find a research group who have a good grasp of the essential physics and some imaginative research kit and see if they can find the solutions.

---

[31] Erik Hallberg, Peter Rättö, Magnus Lestelius, Fredrik Thuvander, Astrid Odeberg Glasenapp, *Flexo Print of Corrugated Board: Mechanical Aspects of the Plate and Plate Mounting Materials*, TAGA Journal, 2, 16-28, 2006

6 Gravure

My colleagues and I once had to introduce a gravure printing process on our production line. We were unfamiliar with gravure so we hired Old Joe who was an experienced gravure printer. He was one of those people with magic fingers and trained eye who knew how to get the job done. After months of frustration at failing to get our shiny new, high quality machine to print reliably we found the root cause: Old Joe. This discovery came out of frustration. A QC technician had been watching our failures and suggested one day that he should be allowed to set up the printer as he had a theory of what was going wrong. We had nothing to lose so put him in charge. The machine started, print quality was perfect and as we stood there for an hour, the quality remained perfect. I went off for a cup of tea and came back to find that the quality had slightly deteriorated. Why? Because Old Joe could not resist the temptation to tweak the system a little. I regret to say that I got very angry with Old Joe. But that incident was a formative moment. The secret of the technician's success will be revealed shortly.

Since then I have been alert to the fact that the Old Joes who are in most organisations are not, as they think, the people who keep the machines in good shape. Their intuitions are built on misunderstandings of their processes, and their meddling often makes things positively worse. Once when I joined a coating company I quickly got the impression that what the machine operators were doing was largely making things worse - and I was no better. But I happened not to know the physics so this was only an impression. Once I got to know the physics I found that my impression was correct - we were doing just about everything wrong.

Although the day-to-day problems are caused by Old Joe, the root cause is that more senior people have not taken the trouble to find out how their printing and coating processes should work. Or, if they have done so, they have not (and I speak from my own failings) worked out the best way to show Old Joe how to do things properly.

All this is leading up to the point that a chapter on gravure really has nothing much to say other than some details about the cell filling and emptying that were not discussed in the Flexo chapter, as their effects in flexo are diluted by the intermediate step onto the plate.

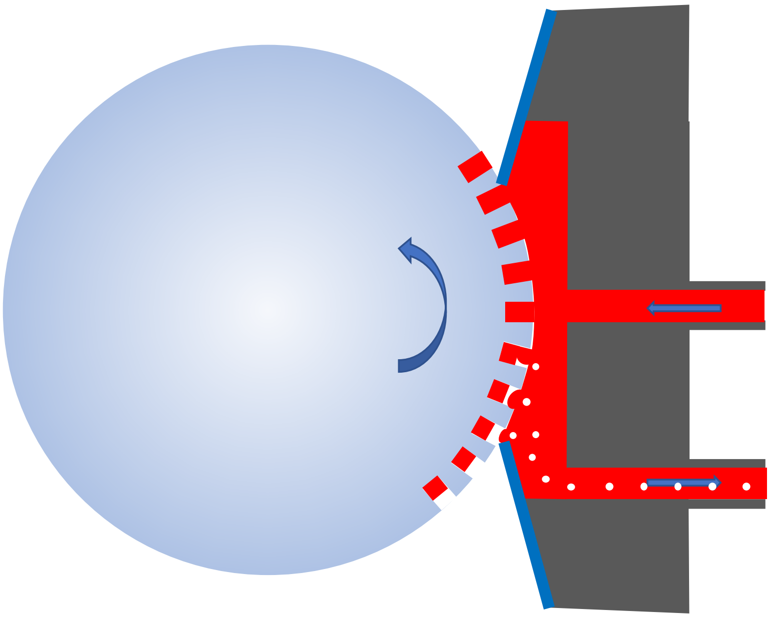


Figure 6-1 The chambered doctor blade and the messy act of filling the cells which necessarily adds bubbles to the chamber.

If, as in the discussion of the anilox for flexo, you have a good chambered delivery system with the doctor blade in good shape and with the problems of air bubbles solved, there is little more to be said that is relevant to basic print quality. A bad doctoring will leave streaks of ink on your print, and air bubbles will produce obvious defects.

A reader kindly pointed out one issue that I'd missed. The distance from the doctor blade to the printing nip can be too long for a volatile ink and a hot surface from the action of the doctor blade. If the ink starts to dry in this short space then either the heating from the doctor blade has to be decreased or the volatility of the solvent has to be decreased.

My experience from Old Joe and around the world is that the proposed answer to most doctor blade issues is to increase the pressure. This is, of course, fighting the laws of physics. A light pressure applied with the minimum of distortion to the blade is more than enough to clean the cylinder (and generate the least amount of heat). The insight of the QC technician was that Old Joe spent more time talking about his skills than he spent carefully and gently aligning the doctor blade on start-up. Without the attention to detail, the blade started out with minor stresses and strains which could only be temporarily fixed by twisting one of the adjustment knobs - which in turn would cause a strain elsewhere. The technician reversed the priorities; he said very little and spent the time getting the clean blade fitted evenly with no stresses.

In the screen-printing chapter I will further discuss this need to fiddle to fix mistakes that are inherent to the set-up. The laws of physics tell us that a machine with the fewest adjustment controls is the most likely to be long-term successful.

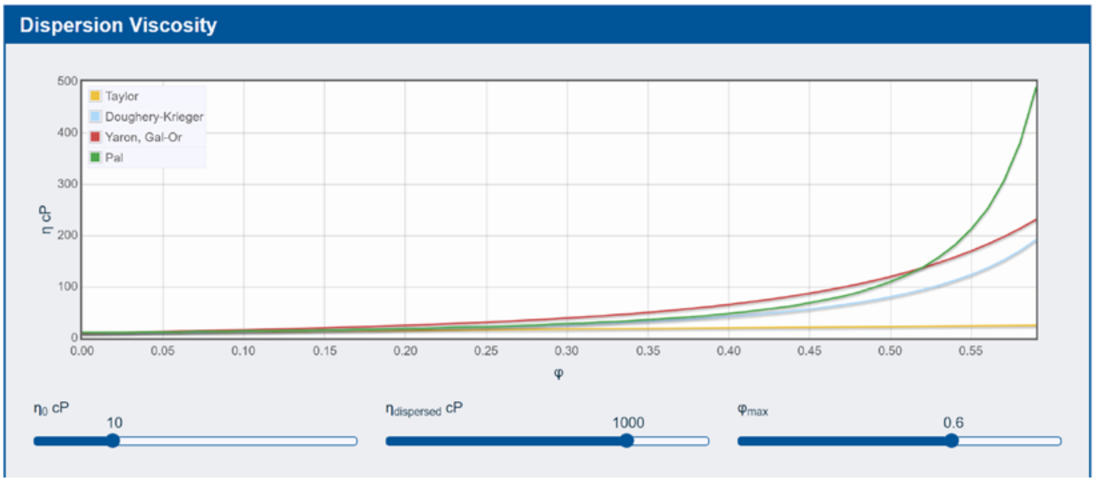
A confirmation of this view on the doctor blade adjustment is echoed by the very wise Dr Werner Kunz in a 1975 review[32] of gravure science. Apart from the issue of coping with the gradual wearing away of the blade, the whole of the doctor blade issue was summed up in this sentence, in slightly Germanic English: "These deviations in the print are well known as wiping deviations, and they are extensively avoidable, if the doctor blade is led very precisely and without getting loose."

In terms of bubbles induced in the chamber at high speed, you can try to get some de-aeration in the circulation system but my impression is that this only fixes very large bubbles which are less of a threat than cell-sized bubbles. Apart from that the only choices are to reduce viscosity (the bubbles get dragged more by higher viscosity) or purchase a superior design of chamber from a better supplier. I once got involved in a "we can't run fast enough because of bubbles" issue and had a chance to inspect the chamber carefully. It seemed an entirely mad design and I could not work out what was going on. But it started to make sense and when I found the manufacturer's patent with a precision drawing it became clear that a lot of good thought had gone into the design. The trick was to trap the air bubbles in a vortex that was constantly swept out to the edges of the chamber by incoming fresh ink, with the printing side of the chamber almost blocked off from the air bubble side. Only at high speeds or excess viscosity would the bubbles get dragged over to the printing side.

Viscosity effects are clearly dominated by pigment (or emulsion particle) interactions and some smart chemistry can help minimize this aspect of the viscosity. There is another aspect that is less well-known and turned out to be crucial in the "can't go fast enough" problem. We all know that viscosity increases with percent pigment or emulsion. We generally want the maximum amount of pigment with the minimum viscosity and there is always a compromise between adding a little extra pigment to get, say, a denser print and adding a little extra solvent to get the viscosity down to a printable level. As the app below shows, at some point the viscosity takes off quickly and it is often the case that our formulations are near that point. [A rather complex app (https://www.stevenabbott.co.uk/practical-coatings/Particle-Rheology.php) discusses the influence of particle shape and tendency to self-associate on viscosities. At high shear, ensuring that particles stay un-associated is critical.] If neither part of this compromise works, is there anything you can do about it? The answer, surprisingly, is "Yes".

---

[32] Werner Kunz, Ink Transfer in Gravure Process, TAGA Proceedings 1975, 151-176



η_0 cP 10

$\eta_{dispersed}$ cP 1000

φ_{max} 0.6

App 6-1 https://www.stevenabbott.co.uk/practical-coatings/viscosity.php

The app shows how viscosity depends on the volume fraction of the pigment (if you set η_{drop} to a high value) or emulsion (a lower value for η_{drop}). There are a few different equations and for our purposes here it doesn't matter which one is used. Via the mouse I can find that at 0.4 volume fraction for the Pal curve, the viscosity is 59cP when the viscosity of the bulk solution (without the particles) is 10cP. Now use the app to see what happens if we make a 5cP increase to the bulk solution - not a very large absolute change and something that could easily happen through some special additive. Using the mouse, we find that the viscosity at the 0.4 point is now 84cP. Our "trivial" increase of 5cP has resulted in a 50% increase in viscosity! The reason is simple yet profound. Whatever the functional dependence on volume fraction, $f(\varphi)$ that is used, the viscosity is $\eta_0(1+f(\varphi))$. So going from 10cP to 15cP is a 50% increase in the overall viscosity. If you start with 1cP and make a trivial change to 2cP then the viscosity of the formulation will double!

The message from this is that apparently small changes in one part of the system can have surprisingly large changes in another part. In the specific example I was involved in, a 15% reduction in the base viscosity (rather easily achieved) allowed a 15% increase in production speed without getting the bubbles coming through the sophisticated doctor chamber and without having to alter the complex functionality of the particles.

6.1 The gravure cells

A paper each in the 1975 and 1976 TAGA conferences seem to me to be the high point of gravure science research. After that, papers tended to be much more about phenomenology rather than the scientific principles. This is a distressing aspect of the printing industry. With DoE techniques and modern data analysis techniques it makes a lot of sense to vary a few key parameters in some methodical manner and conclude that A has a big effect and B has a small effect on some printed value such as dot shape. The problem is that we end up with general hand-waving interpretations rather than some proof or disproof of some scientific mechanism. The two TAGA papers showed no lack of good science! The first is the Kunz paper mentioned earlier. The second is by Yuri Bery[33] and is full of mathematical analyses of cell emptying. I am not convinced by the *specifics* of Bery's analysis because modern finite element analyses of cell emptying confirm that it is a hugely complex process. However, I *am* convinced that his overall analysis of key aspects is correct. What follows is my melding of the Kunz and Bery ideas

The first issue is that the doctor blade does not produce a perfect flat liquid surface. It tends to drag the liquid to the downstream side, pulling some of the ink out. In the relatively short time before encountering the substrate, the ink cannot fully recover and, in any case, there is a dip in the ink which may not make contact with the substrate, causing (as mentioned in the Flexo chapter) an asymmetric donut dot. The problem is made worse thanks to "flash evaporation" - solvent from the surface evaporating under the extreme high temperature created by the doctor blade scraping on the cylinder. Keeping inks and the cylinder cool is crucial to minimizing this evaporation problem.

The implications for that dent in the liquid becomes apparent when we visualise the splitting process. The diagram, which describes three stages from perfect contact to separation of substrate from cylinder, shows the complex flows involved. The diagram is my own interpretation of what basic physics suggests, slightly modified by some recent computational papers that show the same general features. The diagrams are intended to tell a general story. The details of where the wetting lines actually are and how much ink (nm or µm?) remains on the walls of the cells are issues that are interesting but not the focus of the essential ideas that are shown here.

As the ink comes out it must slide down the cell wall, adjust itself on the substrate and start to form a liquid bridge. When the bridge gets long enough it snaps, with the invariable outcome that the splitting is less than 50% on the substrate.

---

[33] Yuri Bery, *Gravure Printing on Non-Adsorbing Materials*, TAGA Proceedings 1976, 207-220. The emphasis on non-adsorbing materials was deliberate - to understand the key processes without the complexities of transfer onto paper. The vindication of this idea comes from the Sprycha 2007 paper described later, which fully acknowledges the Bery paper.

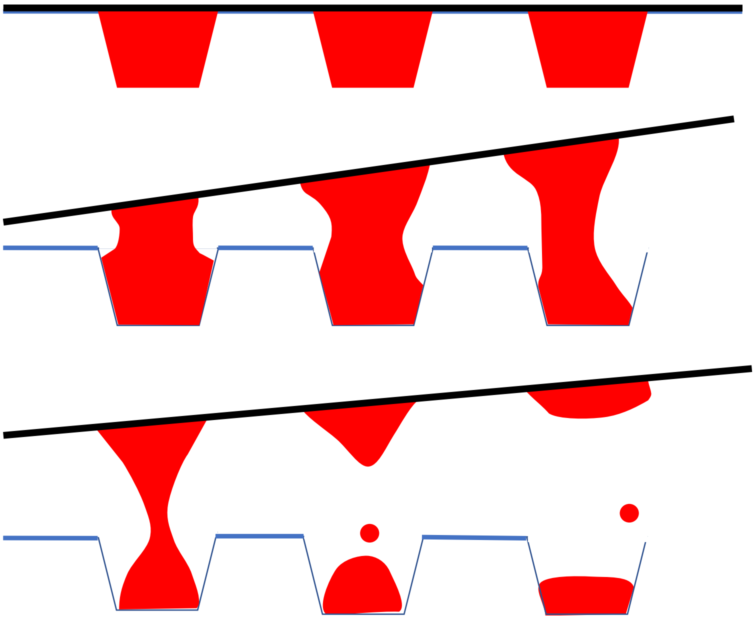


Figure 6-2 The gravure separation process when there is no trapped air.

The little satellite drops are typical of what happens when a liquid bridge finally splits. Whether they are there or not depends on the whole system of speed, shape, viscosity and viscoelasticty.

When there is a dip in the surface from doctor blade forces, the air gets trapped between the ink and the substrate and because splitting only takes place when there is contact with the substrate, the result is a donut:

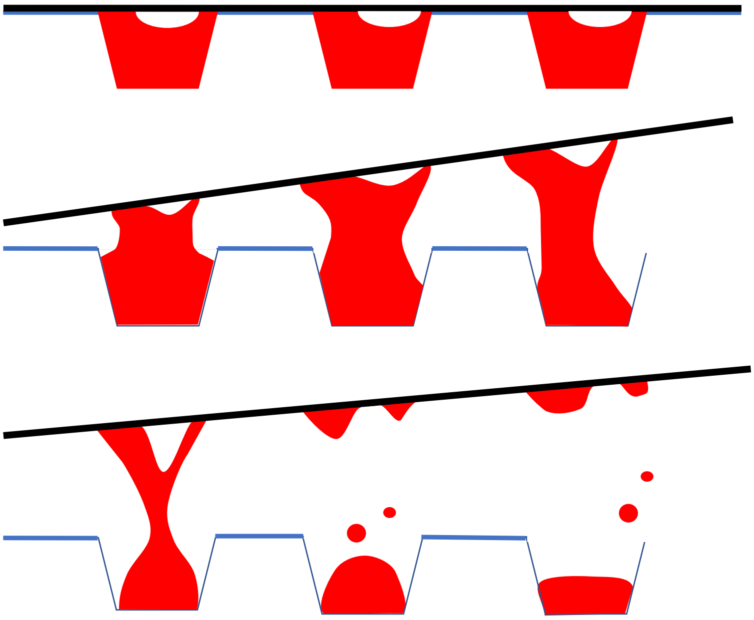


Figure 6-3 The gravure separation process when air gets trapped, resulting in a donut dot.

Although relatively recent academic models with powerful finite element techniques can answer questions about, say, relative contact angles, these are irrelevant because the surface of the cell is instantly "contaminated" by the ink so the contact angle will be small whatever the cell is made from. And the angle on the substrate will not change much, especially for high-speed printing. So the split depends on how the different processes unfold, depending strongly on cell depth and profile as well as on the viscosity of the ink. For a given ink, the tendency to form a liquid bridge will not change much if the depth of the cell gets larger, so breakage will take place at the same absolute amount of ink on the substrate which means a lower percent of the total ink if the cell is deeper. For anilox cylinders where all the cells are the same, we can make pronouncements about Depth-over-Opening ratios. For half-tone gravure the cell depths and/or openings change so the relative % transfer can change dramatically, adding to the difficulties of creating a proper half-tone curve.

Unfortunately, the refilling process of the cells is not especially violent, which means that the liquid at the bottom from the previous print has no special reason to be replaced. So if there is any solvent evaporation before the next filling, the viscosity will increase, making it even less likely that the old ink will be replaced. Over time, as is commonly observed, the cell can become half-filled with dried ink. Because the fluidity of the ink at the bottom of the cell must have at least *some* positive influence on the printing transfer, as the ink dries in, the % transfer (in terms of apparent cell volume) decreases.

Looking at the transfer process, I would have said that speed would make a difference to what fraction of the ink gets transferred. The data seem to suggest that the amount printed is rather independent of speed. Further data, discussed below, suggest that this is partly chance, that the cells fill less well, but the overall transfer is higher. Other data suggest that a larger impression roller, leading to a change in the overall transfer geometry, increases the % transfer, but I have not found enough evidence to know if this is either true or significant.

In the end, it seems to be the case that except for cells that are obviously too shallow or too deep (partly depending on how high the viscosity is), 40% transfer is a likely outcome.

6.1.1 Bad substrates

Although it is generally assumed that papers will extract more ink via Walker-Felsco-style absorption, a paper from Sun Chemical[34] shows how bad paper can be for gravure. At the extreme, some dots have nothing at all in the centre, with a halo of ink where the cylinder wall was. Their explanation, backed up by extensive white-light interferometry is a mixture of capillary flow in the paper and a sort of reversed donut dot phenomenon.

---

[34] Ryszard Sprycha, Richard Durand Jr. and Greg Pace, *Mechanisms of Abnormal Dot Deformation in Gravure Printing*, TAGA Proceedings 2007. 482-505

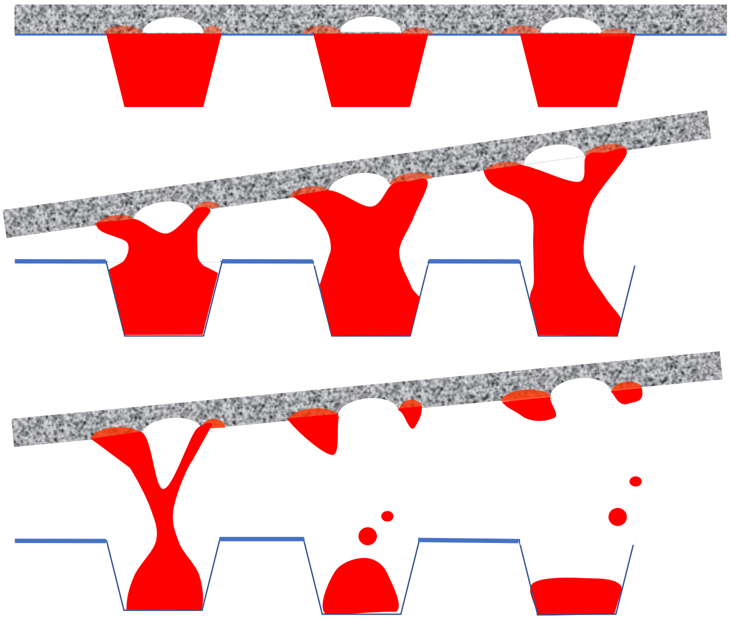


Figure 6-4 Things can go very wrong when the roughness of a substrate means that only the edges of the dot touch the paper.

My diagram attempts to capture the essence of what they observed. The cell happens to touch a part of the paper with a dip that does not come into contact. The ink starts to spread into the paper fibres at the edges. So ink continues to flow into the rim of the dot, giving a hollow centre with ink where the rim of the cell had been.

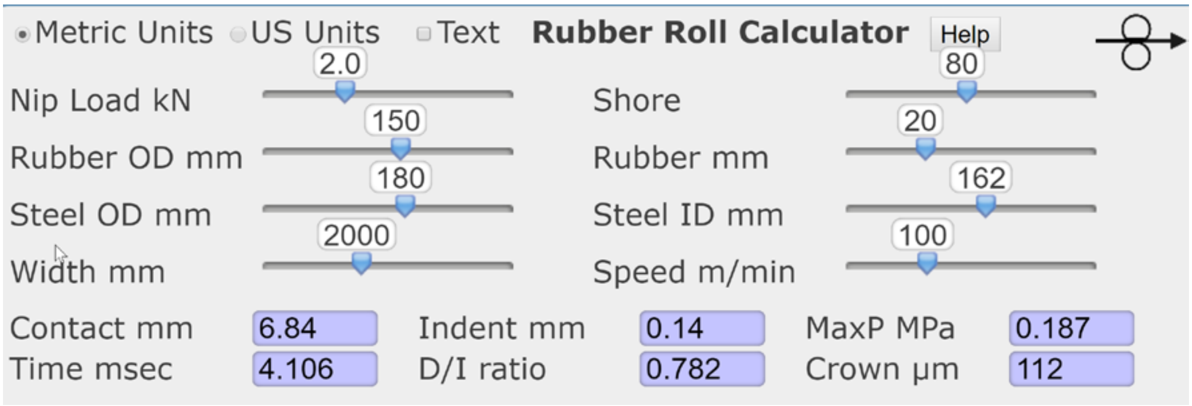
Papers, obviously, can go from being near-perfect prints when they are super-smooth like polymer films to super-poor quality when you get "dip in the paper" effects. The combination of ideas of capillary flow and "no printing without contact" can explain the whole gamut of observed effects.

It is sometimes stated that paper can pick up more ink because it can flex into the cell. As Kunz showed, this effect, while possible when paper is relaxed during sheet-fed printing, is less likely when the paper is under tension in roll-fed printing.

6.2 The impression cylinder

The only advice I have read about the rubber for the impression cylinder is that it should be neither too hard nor too soft. I can find no set of rational arguments for what constitutes the ideal Shore hardness. My guess is that the actual printing step is perfectly happy with just-in-contact pressure with just about any reasonable rubber and that the choice of Shore is more connected to the longevity of the rubber at high speeds, where a high Shore hardness (>70) resilient rubber (polyurethane) at modest pressure (sufficient to give an even nip impression) will do the job.

For those who would like to think through what is happening in their own setup, an analysis of pressures, widths and times within the nip can be useful. Rubber nip mechanics are fairly complicated. Fortunately, an app in my Practical Webhandling series for the converting industry (an older, cruder app format forced on me by the need to convert between metric and US units) does the calculations for you.



Nip Load kN 2.0
Rubber OD mm 150
Steel OD mm 180
Width mm 2000
Shore 80
Rubber mm 20
Steel ID mm 162
Speed m/min 100

Contact mm	6.84	Indent mm	0.14	MaxP MPa	0.187
Time msec	4.106	D/I ratio	0.782	Crown µm	112

App 6-2 https://www.stevenabbott.co.uk/abbottapps/RRC/index.html

This specific setup is chosen to illustrate a super-poor design choice. A 9mm steel shell (180 OD-162 ID, divided by 2) is not strong enough to resist its own weight and the deflection pressure from the 2kN load, so it deflects by 112µm which means (approximately) that the impression roller has to be crowned by this amount, or that another support roller has to be provided to stop the deflection of the gravure cylinder. It also tells you that there is a 7mm contact width, that the rubber is indented by 140µm and that the maximum pressure is ~0.2MPa. At the relatively low speed of 100m/min, the cell has 4ms in which to empty.

Using the app as an indicator of pressure/width/time/deflection in setups that work well for you it should be possible to gain insights into which of these parameters go wrong in the poor setups. If we ignore the crowning issue, then rather large changes in settings have relatively small effects on the nip. Doubling the nip pressure in the setup above, increases the contact width (and time) by 25% and both the maximum nip pressure and indent increase by ~60%. Keeping the original pressure but changing from 80 to 70 Shore gives a similar 25% increase in contact width, a 50% increase in indent and a 20% decrease in maximum pressure.

6.3 (Not) Joining the dots

So far, the cells have been shown as totally isolated from each other, thanks to a perfect doctor blade sweeping over a perfectly smooth cylinder between cells. If the doctoring process is faulty then an obvious thin layer of ink remains on the cylinder between the cells and appears as an unwelcome background tint.

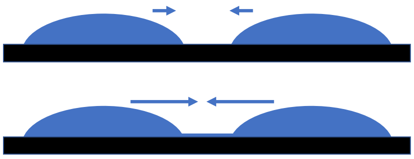
In some contradiction to the ideal is the fact that most practical gravure cylinders are deliberately roughened to allow some ink to survive within the roughness.

One explanation is that this lubricates the surface, otherwise the doctor blade gets so hot that it melts. A complementary explanation is that it reduces the contact area with the (oscillating) blade to reduce heating and wear. So what is the optimum roughness? Here I failed to find any coherent data, with vague phrases like "a roughness of 0.5μm" with, unfortunately, this roughness defined either as an Ra or an Rz, even though the latter is, statistically, ~5x larger than Ra for given surface (feel free to play with the Surface Profiler, https://www.stevenabbott.co.uk/practical-coatings/surface-profile-explorer.php, to explore the relationship between the two). Let us suppose that it is an Rz of 0.5μm. We can make two different guesses about the fate of the ink within the roughness.

The first is that the heat from the blade and/or the rapid airflows will be enough to evaporate the solvent, with the result that very quickly the roughness is filled with dried ink whose only purpose is to reduce friction and heating of the blade.

The second guess is that the ink remains liquid and some of it gets printed onto the substrate, giving, say, a 0.1μm layer of ink. For printed electronics this would be a disaster as 100nm is a typical thickness of many electronic structures. This is why those in printed electronics require smoother rollers and print at slower speeds.

So let us stay with graphic printing. If that 100nm of wet ink is not visible as a tint, might it have any other effects on the print? The diagrams of cell emptying suggest that a gravure dot is not like the squashed cylinder of an offset or flexo dot, and this can be a potential disadvantage. So it would be good for the dot to flow somewhat in order to level out. The Tanner law spreading discussed in the first chapter is controlled by contact angle and the "slip length" at the edge of the dot. This slip length was a mathematical way to get round the embarrassing fluid dynamic fact that the velocity of a fluid at a solid interface is always 0, the "no slip" boundary condition. It says that a thin (molecular) precursor film forms ahead of the spreading drop. Modern techniques have shown that this mathematical convenience is a physical reality, but it remains a tricky issue. If, however, the thin slip layer already exists then fluid can flow quickly and the contact angle of the drop is now with respect to this fluid layer and is zero - so the drop can spread quickly and without the limit imposed by equilibrium contact angles.



What this means is that if the ink from the roughened cylinder forms any sort of nm-scale coating on the substrate, then unlike the dots without a coating which spread slowly, gravure dots on the nm-scale coating can flow together freely till they either join or slow down because of evaporation. This means undesirable dot gain for some or desirable flowing together of "dotted" approximations to solids or lines.

This seems such a fundamental aspect of gravure (for good or ill), and the deliberate roughening of rollers is so intrinsic to the industry, that I had expected to find analyses of the issues in the literature. Maybe I haven't used the right search terms, but so far I have found nothing of relevance.

However, the team at TU-Darmstadt kindly pointed out to me a major issue commonly known in the industry and, unfortunately, not known to me. Depending on the ink and gravure pattern there is often a critical % dot where there is a transition between "single cell splitting" (what we've just been discussing) and "film splitting" where all the dots have joined together, losing the ability to print the remaining tonal range. The terms are taken from the Bornemann thesis from TU-Darmstadt[35] discussed (for other reasons) in the first chapter, and Bornemann refers with approval to Kunz's pioneering work in imaging this phenomenon. As far as I know, there is no current explanation or predictive tool, though the intuitions that larger cells and lower viscosities encourage the "film splitting" mode seem to be correct. Although the gravure *printing* industry has long known (presumably via custom and practice) how to avoid this transition in practical printing, for gravure *coating* and for gravure-based printed electronics, "film splitting" mode is highly desirable as it is hard to reliably join up individual dots.

6.4 A scientific graph

Work at Berkeley by Grau, Kitsomboonloha, Subramanian and others has seriously advanced our understanding of how to make gravure fit for printed electronics. An overview paper[36] makes excellent reading and a striking graph provides a rather neat summary of how ink deposit depends on the key parameters, viscosity and speed via the capillary number (so surface tension plays its usual small role too), $Ca=\eta V/\sigma$.

The key idea is that there are a series of fractional values, φ, which describe different parts of the gravure process and come together to form an overall % transfer onto the print, φ_p. Let us look, first, at my version of their graph, then an explanation of each term and how they come together to create φ_p.

---

[35] Nils Bornemann, *Characterization and Investigation of Large-Area, Ultra-Thin Gravure Printed Layers*, TU Darmstadt, 2014

[36] Gerd Grau, Jialiang Cen, Hongki Kang, Rungrot Kitsomboonloha, William J Scheideler and Vivek Subramanian, *Gravure-printed electronics: recent progress in tooling development, understanding of printing physics, and realization of printed devices*, Flex. Print. Electron 1 (2016) 023002

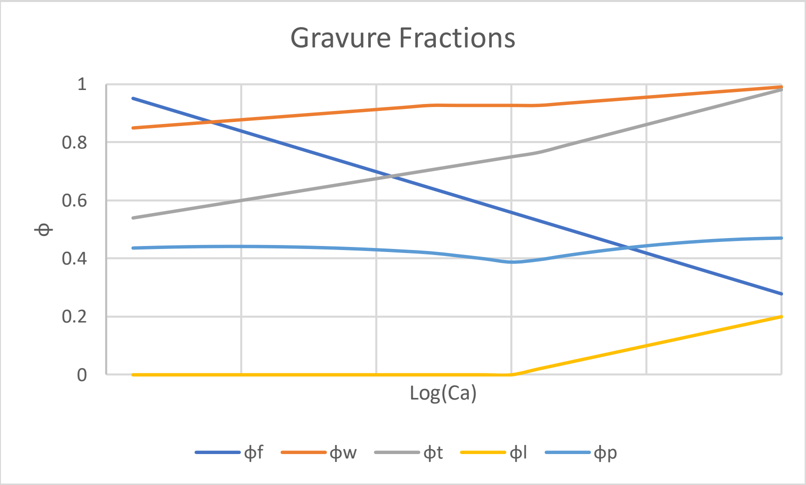


Figure 6-5 Four individual transfer fractions combine to give an overall fraction of the cell volume transferred to the print.

The first fraction is $φ_f$ which is how perfectly the cell is filled. At low Ca (low speed and/or viscosity) it is 100% filled but that diminishes steadily with increasing Ca. The next is $φ_w$ which looks at how much of a filled cell gets wiped away by the doctor blade. This is quite a lot at low Ca, where the liquid in the cell has time to flow up and around the blade (and re-deposit as a drag-out line behind the cell) but rather little at high Ca. The pause in the growth may, according to the authors, be some sort of viscoelastic effect. Then there is $φ_t$ the amount that is transferred from the (imperfectly) filled cell onto the substrate. This increases with Ca. These three factors are multiplied together to give an overall value - to which is added a fourth term, $φ_l$ which is the amount that leaks under the doctor blade at higher speeds and generally increases the printed volume but, of course, appears as scumming in the print.

One good rule of science is never to trust a graph with no units along an axis. So what is the diagram trying to hide? The usable range, according to other papers by the authors, is for Ca to be between 0.1 and 1. This is entirely acceptable to their area of interest which is super-high-quality printed electronics with features in the few µm size. To achieve this, any form of conventional gravure cylinder is far too rough, so they use special silicon originals transformed into hard nickel rollers via a standard electroforming process. Their inks are, say 40 cP with a surface tension of, say, 40 mN/m so this means that they can print between 0.1 and 1 m/s. This is very impressive for such exquisite electronics printing but not at all relevant to "normal" gravure which is happy at 10 m/s. Presumably the different scale of cells and of gravure blades (and acceptable pressures) allows the general scheme to stretch out to higher Ca values.

Their detailed analysis of blade tips and (by putting some PTFE tape under the doctor blade!) its surface energy shows that the low Ca drag out problems can be greatly reduced, though they admit that a super-sharp tip and a PTFE coating both have their practical issues for long print runs. Equally, this work shows that for higher speed conventional gravure, our current blades are fine and attempting to change their surface energies (or, indeed the surface energies of most other parts of the process) is scientifically, and practically, of little value.

Whatever the Ca scale, the work behind the graph is truly impressive. It really is Ca that matters - they get *identical* results if they test a 200 cP ink at 0.5 m/s or a 100 cP ink at 1 m/s. Their elegant experiments confirm what others have shown, that Ca rules apply to cell filling and to cell emptying. This means that those who care about pushing the boundaries of speed must pay special attention to the shear thinning aspects of their inks. The increase of Ca from higher speed will be somewhat compensated by the decrease in viscosity from higher shear.

And this is where it gets frustrating. For the glamorous part of the process, it is rather easy to formulate an ink which drastically reduces its viscosity at the high shear of the doctor blade and the high extensional flows of the transfer process (even if few of us have the means to measure such viscosities). But we still have to get the air out of our cells in the chambered doctor blade and the shears in the chamber will be relatively low, so the viscosity will be high enough to drag the air bubbles to the doctoring area and spoil everything. Which brings us back to yet another paper from the Swansea team.

6.5 No more theory available

It may be depressing that a century of gravure printing is summed up with the insights of a just few papers and that the reader is left with nothing other than the advice to get someone to run the press who understands the need for careful setup and then avoiding touching any of the controls. On the other hand, the fact that once you have your gravure cylinder nothing much can alter the print is a key strength that has made gravure a solid performer over the decades. A paper from the Swansea team confirms this[37]. Doctor blade angle, impression and a few other parameters had only small effects on print densities. The only thing that really mattered was that in a typical print run, viscosities varied significantly and these variations showed up in the print densities. So for good control it just required good viscosity control.

Gravure is gradually losing market share as new plate technology increases the quality of flexo, which is more flexible in terms of shorter runs in the age of mass customisation, and offset is also increasing its ability to compete on long runs. So it is hard to see anything that will be able to make it relatively more attractive. After 100+ years of good performance that is not too sad a fate.

---

[37] M.F.J. Bohan, T.C. Claypole and D.T. Gethin, An Investigation into Ink Transfer in Rotogravure Printing, TAGA Proceedings, 1998, 484-494

7 Inkjet and other digital techniques

This chapter is mostly about inkjet and, for reasons that will become clear, is surprisingly short. Although the other digital techniques are of great practical importance, they are effectively black boxes to us so there is not much we can do even if we understand the science.

7.1 Inkjet

Inkjet is such a delicate and fussy process that its success is surprising. The inks can only work over a tiny range of viscosity (8-20cP) and surface tension (25-50 dyne/cm) and are jetted through tiny holes that will clog up or gain some jet-displacing edge contamination at the slightest excuse. Tiny satellite drops reduce print quality and there is a contradictory need for the drops to spread somewhat (to avoid being too "dotty") without spreading too much (to avoid blurring) on substrates that are neither too absorbent (dilution of colour via Kubelka-Munk effects) nor too non-absorbent (especially for water-based inks which contain essentially non-volatile glycols).

I mention the problems right at the start because inkjet has been the subject of vast hype by those who hoped to use it for applications beyond those of producing nice-looking images, especially printed electronics. That most of these have come to nothing was entirely predictable (and I often made such predictions at conferences), and the relatively few specialised successes (in some parts of printing displays) are from processes requiring 100s of millions of dollars, not the disposable inkjet printers we see all around us. Even then, the drops are jetted into regions previously delineated by lithography, so the naive ideas of all-inkjet printed electronics hasn't worked out.

Following the precedents of the previous chapters, the inkjet head itself is largely ignored. There are crucial differences between the different types of piezo heads, and between piezo and thermal heads, but they are not the sorts of issues relevant to this book. All I care about is that a set of pressure waves of precise timing and intensity are able to eject a drop through a narrow orifice. My interest starts as the drop is coming out of the orifice.

7.1.1 A delicate balance

The fundamental reason we have cheap inkjet printers is that the whole process is delicate. So rather ordinary electronics can deliver the rapid pulses of energy needed to push some small drops out of a small hole. We can also have lots of those nozzles close together to get relatively high-resolution and high-speed (full width nozzle) printing. The downside is that the process is a delicate balance between surface tension, viscosity and inertia.

We have already discussed that balance in the first chapter, and the following is a simple copy/paste.

An inkjet nozzle doesn't eject drops, it ejects a stream of fluid at a carefully controlled set of velocities (the actuators follow complex driver waveforms), with some sort of abrupt halt when enough liquid has been ejected. The ejected volume has to separate from the liquid that should remain in the head. The initial drop plus tail should end up as a pure drop. Done incorrectly then the drop plus tail ends up as drop plus satellite which itself can be flying off at a strange angle

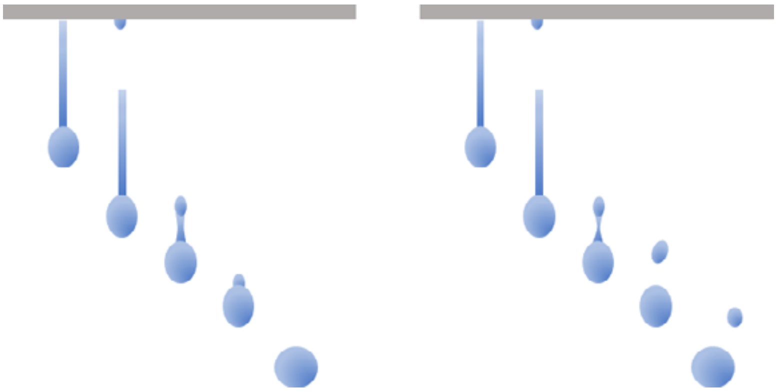


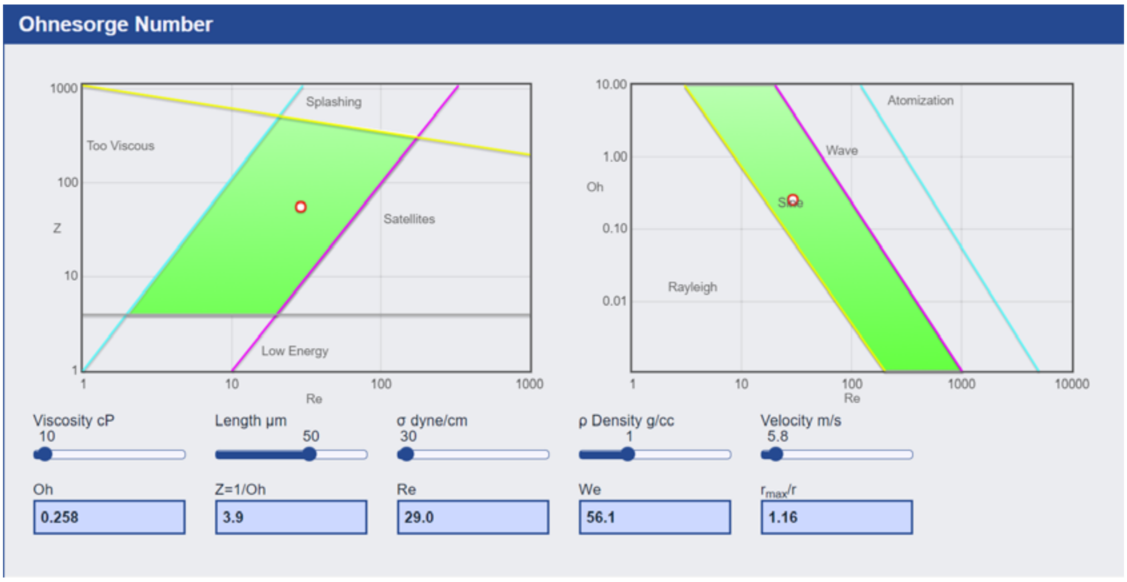
Figure 7-1 Desired (left) and undesired (right) inkjet drop splitting

The natural instability of any stream of fluid can readily be observed when slowly turning off a flow of water with a tap (faucet). The steady stream becomes an unsteady stream then, thanks to the Rayleigh instability, breaks up into individual drops. Although an entire chapter could be written on the relevant phenomena, we can summarise what happens in two dimensionless numbers: the Ohnesorge number, Oh, and Reynolds number Re plus the related Weber number, We:

$$Oh = \frac{\eta}{\sqrt{l\rho.\sigma}} \quad \mathrm{Re} = \frac{V\rho l}{\eta} \quad We = \frac{V^2\rho l}{\sigma}$$

Equ. 7-1

The Ohnesorge number gives us a ratio of viscous, η, to inertial and surface tension forces via the length, l, (in this case the diameter of the inkjet nozzle), the density (which is always going to be close to 1) and surface tension, σ and tells us more about what will happen when the drop hits the substrate. The Reynolds number is the ratio of viscous to inertial forces and tells us whether it will or will not form a single drop. At low Re the ink simply won't form a drop and at high Re it will tend to form satellite drops which fly off in all directions, degrading print quality.



Viscosity cP 10

Length µm 50

σ dyne/cm 30

ρ Density g/cc 1

Velocity m/s 5.8

Oh 0.258

Z=1/Oh 3.9

Re 29.0

We 56.1

r_{max}/r 1.16

App 7-1 https://www.stevenabbott.co.uk/practical-coatings/ohnesorge.php

The app shows an inkjet drop comfortably in the safe zone in Oh/Re space, not surprising given its modest viscosity, surface tension and density along with a typical nozzle diameter of 50µm. By playing with the app you will find just how restricted inkjet printing is by the conflicting demands of speed, viscosity, surface tension and nozzle size.

While we are here, it is interesting to look at the final output box, r_{max}/r. Many of us imagine the inkjet drop smashing into the surface with spectacular effects. Reality is disappointing, the velocities are so low that the drop hardly deforms on impact. The calculation gives a size of the maximum radius reached during the impact process, and for most inkjet-relevant conditions the value is very close to unity. The formula uses both We, Re and the contact angle θ:

$$\frac{r_{\max}}{r} = \sqrt{\frac{We^2 + 12}{3\left(1 - \cos\theta\right) + 4We^2 / \mathrm{Re}^{0.5}}}$$

Equ. 7-2

Having repeated the text from the first chapter, we can start looking at some detailed issues.

UV inkjet should be impossible given that the acrylates that give desirable printed properties tend to have higher viscosities than inkjet can manage. Fortunately, inkjet heads do not mind being heated to 60 or 70°C and that is enough to bring the viscosities down to a printable level.

More worrying to heads than mere temperature are specific chemicals or pH values that can attack one of the many parts of a head. My colleagues and I once worked on a novel inkjet imaging system that, for a week, showed amazing promise. Then, suddenly, the head stopped working. We had strictly followed the manufacturers list of rules about our ink and had prepared it super-carefully.

When we discussed the problem with the manufacturer they said "Oh, we forgot to say that the head cannot accept a pH lower than X". Unfortunately our process depended entirely on the pH being slightly lower than X and so our ink had eaten the head. We were neither the first nor the last inkjet formulators to destroy a head for unexpected reasons.

If the ink has been well-formulated so that it jets nicely with few satellite drops, there are three issues that are nightmares for inkjet systems:

7.1.2 The problem of bubbles

Take 1ml of an ink at 20°C and raise its temperature to 30°C. We know that air is less soluble at higher temperatures so it is possible that some of the air that was soluble at 20° will produce bubbles at 30°. A simple app at https://www.stevenabbott.co.uk/practical-coatings/bubbles.php allows us do the calculation if we assume that each bubble has a radius of 100μm. Most of us might guess that we would get a couple of extra bubbles. In fact, if all that air came out of solution and formed 100μm bubbles we would have 300 of them. In practice this does not happen, but it shows that even modest temperature rises are enough to create a serious risk of bubbles which could easily form a blockage or misfiring inside the inkjet head. The only way to avoid this problem is to de-gas the ink in advance, either by equilibrating it at a temperature higher than it will encounter in real life or by vacuum removal of excess air. In both cases the ink needs to be well-sealed from air so that it does not have a chance to re-equilibrate at, say, 20° before experiencing a time at 30°. Along with this de-gassing step before packaging, it is necessary to perform a very fine filtration step because every bit of contamination is a potential blocked head and an unhappy customer.

7.1.3 The problem of drying on/in the nozzle

We all hate the fact that inkjet printers are forever spitting our expensive ink into some cleaning unit. But we hate even more if any nozzle in the head becomes partially blocked/contaminated causing the jet to fire in the wrong direction, or becomes totally blocked so that it doesn't fire.

The problem, again, is that the forces involved in inkjet are very low, so there is not enough fire-power in a head to force the ink out of a blocked nozzle. Hence all the capping of heads, wetting, wiping, spitting - just about anything to ensure that each nozzle remains in perfect shape.

The key dilemma for the ink designer is that entirely non-volatile solvents are perfect for avoiding the risk of drying out in the nozzle, while volatile solvents are best for avoiding problems once the ink reaches the substrate. For water-based inks the compromise is to add whatever is currently considered to be a safe glycol and hope that the substrate can cope with absorbing it and, potentially, retaining it for ever. Because a glycol such as glycerol positively attracts water from the atmosphere, the nozzle has the greatest chance of remaining moist - though the ink designer has to be aware that the glycols can cause problems such as stripping dispersing agents from the pigments or reducing the solubility of components that are only happy in water.

7.1.4 The problem of drops on the substrate

So far we have imagined that inkjet is about firing a drop, letting it land and for it to dry on or absorb into the substrate. Reality is very different. Even for simple inkjet systems, multiple drops are required in close proximity and for modern systems that can deliver multiple sub-drops, the problems of drop-to-drop interactions are more severe. Increase of speed via multiple parallel heads comes at the cost of more complex firing patterns to avoid too many dots joining together to form a puddle of ink.

Up to this point in the book we could be relaxed about drop spread (because the printed drops from offset, flexo and gravure are so small), absorption into the substrate (other than thinking about Walker-Fletsko) and surface energy (usually a few orders of magnitude too small to be relevant). For inkjet all aspects are crucial.

Let us see what happens with each in turn. As we learned from the Drop Spread Modeller, the velocity of the drop spread depends on the contact angle θ^3. We can assume that the initial drop is a "hemispherical cap" with a high angle, so the drop spreads very fast, especially given the low viscosity.

The Drop Absorption app tells us what happens to a drop of a given volume on a medium with pores of a given radius. From this we can get an estimate of the time available for drop spreading.

The surface energy, along with the surface tension of the ink gives us a contact angle. This also imposes a limit on the drop spreading because the drop will stop when it reaches this angle. The "puddle of ink" phenomenon takes place if a previous dot has already wetted out the surface; this is because the contact angle will be, effectively, zero so the new drop will spread fully.

We can combine the first two apps to see that with a reasonably absorbent medium, drop spread is almost irrelevant which could actually be a problem:

Volume pl 50
SurfTen σ dyne/cm 30
Viscosity η cP 10
Contact angle θ 40
Pore radius nm 49
Porous fraction f 0.5
Diameter μm 45
Volume pl 47.7

Drop ● Wedge ○ Cylinder ○

Viscosity Pa.s 10
ST Dyne/cm 30
Radius μm 29
θ_{Start}° 90
$\theta_{Equilibrium}$° 40
T_{Max} s 0.1
R_{Max} μm 8
Height μm 15
Width μm 25
Volume 51.1pl
End Rad. μm 35.3

App 7-2 https://www.stevenabbott.co.uk/practical-coatings/drop-absorption.php and https://www.stevenabbott.co.uk/practical-coatings/drop-spread.php

The drop absorption app shows that a 50pl drop (~45μm diameter) is fully absorbed in 10ms. In that time the radius of a similar drop will have spread by just 1μm. This could be a problem for a digital image which might look rather too "dotty". It would also be a problem in terms of Kubelka-Munk effects as the medium will scatter the light.

In practice, pigmented inks tend to form a barrier to absorption so the timescale for dot spreading is likely to be longer. The image shows that after 100ms, the 29μm radius drop has spread to 35μm, illustrating just how rapid drop spread can be.

For those who do printed electronics onto non-absorbent media, then the drop can easily spread for 1s, giving a radius of 43μm if the equilibrium contact angle is 40°. Restricting drop growth via the contact angle is just about the only reliable strategy for printed electronics unless one chooses to defeat the whole purpose of the technique by pre-imaging non-wetting shapes around the drop to confine them that way.

What happens when two drops touch? Is there any way we can control this process so that, for example, drops don't grow together too fast? The short answer is "No". The literature might debate whether drop coalescence takes place in 0.1ms or 1ms or even 2ms, but in practice that means that coalescence is instantaneous, once the drops touch. The time taken for two drops, printed a certain distance apart, to coalesce is controlled by the drop spreading rules. Using the example above, if drops are spaced at 70μm apart then the 29μm radius drops have a gap of 12µm between them and this is covered in 100ms. In about 101ms the drops are essentially a single drop, though it might take a few 10's of ms more before the new drop has sorted itself out.

For printed electronics, a desirable basic structure is a nice straight line. Clearly if the drops are spaced too far apart, the line will never join up; and if placed too close together it will just produce a wet, fat (and, as it happens, bulging) line. What, then, is the correct spacing for a nice line? With modern multi-nozzle heads working in complex patterns, there is no simple answer. The analysis from Stringer and Derby[38] assumes that single dots have been printed sequentially and gives us a clear number, p_{crit}, which is the critical spacing between dots above which the line will look like a series of not-well-joined dots. To calculate p_{crit} we need the actual spacing, p, the contact angle, θ, the drop diameter, D, from which we need to calculate a factor β (the ratio of a fully expanded diameter over D) and can calculate the line width, w.

Equ. 7-3

$$\beta = \sqrt[3]{\frac{8}{\tan(\theta/2)\left(3+\tan^2(\theta/2)\right)}}$$

then:

Equ. 7-4

$$w = \sqrt{\frac{2\pi D^3}{3p\left(\frac{\theta}{\sin^2\theta} - \frac{\cos\theta}{\sin\theta}\right)}}$$

and finally:

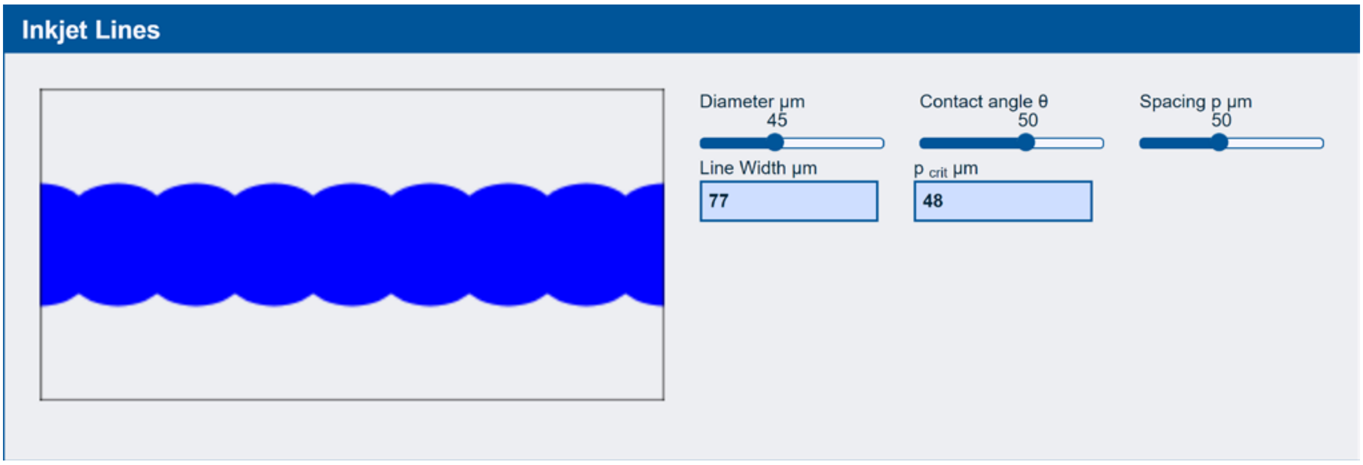
Equ. 7-5

$$p_{crit} = \frac{2\pi D}{3\beta^2\left(\frac{\theta}{\sin^2\theta} - \frac{\cos\theta}{\sin\theta}\right)}$$

The complicated geometrical terms come from the spherical cap shape and are rather hard to interpret. The app does the interpretation for you.

---

[38] Jonathan Stringer and Brian Derby , *Formation and Stability of Lines Produced by Inkjet Printing*, Langmuir 2010, 26, 10365–10372



Diameter µm 45

Contact angle θ 50

Spacing p µm 50

Line Width µm 77

p_{crit} µm 48

App 7-3 https://www.stevenabbott.co.uk/practical-coatings/InkJetLines.php

In the image, the spacing is 50µm which is just above the critical 48µm. If you slide the spacing to a lower value then a perfect line is formed.

If you drag the spacing even lower then (not shown) some bulges appear in the line. The paper describes how to calculate the onset of bulging, but it is rather too complex for the app so the bulges appear when the spacing is p_{crit}/1.5 to illustrate the point.

7.1.5 More!

The reader might expect at this point to be reading about more detailed scientific methods for the key aspects of real-world inkjet printing, especially for printed electronics. Having read many excellent books and papers on the subject, I can only conclude that it is impossible to say much more of value. There are many wonderful studies of specific types of drops doing specific things, imaged at high speeds and with careful automatic measurement of many interesting aspects of the development of the drop structures. Yet I have not found any plausible methodology for generalising any of these examples. Everything seems to specific to each ink. My advice from expert colleagues is that even the apparently simple task (given the constraints of Ohensorge and Reynolds numbers) of creating a new, jettable ink for a specific need is a matter of much trial and error. Once you add the complexities of multi-head, multi-nozzle, grey-scale droplets systems, plus the behaviours of different substrates with or without previous drops which, in turn, are wet or dry, then the whole thing becomes too difficult for helpful generalities.

I wish I could write more about inkjet. But I cannot.

7.2 Other Digital

The major techniques are all black boxes using techniques developed by well-funded and, presumably, scientifically minded teams who have brought the technique to an advanced state. My interest here is to identify the few scientific principles that we, as potential end-users of these techniques, might want to apply during the process of deciding whether to use them and, if so, to work to their strengths and within their limitations. I will start with offset inkjet.

7.2.1 Offset inkjet

Many of us were sceptical when the Landa Nanographic Printing® technique was announced at Drupa in 2012. Although it was clearly a splendid idea, we could see that it would have to solve multiple issues before it could be general-purpose and versatile. We were right in both senses. Its theoretical capabilities are impressive and bringing the machines to market has taken many years, with roll-out having started in late 2017.

The idea of offset inkjet is to manage two of the tricky parts of inkjet within a highly-controlled environment. Everything to do with jetting the ink is controlled, and only their special inks can be jetted. And by controlling the substrate (an endless belt with a carefully-designed surface) onto which the drops fall, and the temperature of the environment, everything about dot gain and ink drying can also be controlled. Although the inks are water-based, they still need extra additives such a glycols and removal of these can take place in the controlled drying zone.

The user's substrate of choice therefore has no compromises in terms of dealing with wet inkjet inks. In principle, therefore, a perfect, high-resolution print is simply transferred from the offset surface onto the substrate, irrespective of whether it is relatively rough paper or smooth polymer..

The "Nanographic" part is simply that the pigment particles are (by the definition of "nano") sub 100nm. At a size less than the wavelength of light, there is less scattering so the colour gamut (think Kubelka-Munk) is larger and the colours are more intense. The printed features can, therefore be closer to 0.5µm thickness than 1µm which, in turn, means that dot-on-dot issues are less of a problem. Because the inks are not absorbed into paper substrates there are fewer Kubelka-Munk compromises compared to conventional inkjet.

Our concern is not with the difficult juggling act of getting the ink dots to stick sufficiently to the offset substrate to withstand the printing and drying, then allowing the ink to transfer perfectly to a user's substrate. That, presumably, took up a good chunk of the years between announcement and real-world installations.

Our first concern, as with the Indigo process below, is adhesion of the ink onto the substrate. The advantages of custom inks onto custom substrates is that the full range of adhesion tricks can be used. For these generic inks the only possibility is good contact via a heat-induced flow onto the substrate. For very thin inks this is not a bad strategy as it is difficult to pluck them off in normal use.

However, the ink has a balancing act: if the polymeric carrier has a high MPt it will be relatively hard at room temperature and resistant to casual abrasion, but it will be hard to get it to flow well during the heat transfer process. With a low-MPt ink, the flow and adhesion will be much easier, but in general the printed ink will be more susceptible to damage.

The second issue arises because the technology is explicit about its use of nanoparticle inks. Although "nano" only means "small"[39] and there are plenty of safe unnatural nanoparticles and unsafe natural nanoparticles, there is legislation world-wide requiring those who use nanoparticles to specifically address the potential safety concerns that arise simply because the particles are sub 100nm. What is interesting is that to the best of my knowledge, a lot of conventional inkjet ink are (for obvious reasons) based on nanoparticles without (as far as I can tell) any special safety registration work or push-back from users.

7.2.2 Offset electrophotography

Direct electrophotography with dry toner particles comes with well-known strengths and limitations. The limitations mostly arise because the toner particles need to be relatively large, otherwise they become an uncontrollable fine dust when printed at high speeds. The large particles tend to give lower resolution and lower gloss, and the larger scattering reduces the colour gamut. Fusing these large particles onto the substrate requires considerable thermal energy from a hot roller or a radiant source. Adhesion is generally adequate but the relatively thick layers are more vulnerable to being picked off. Clearly it is a wildly successful technique and is not discussed further as I can find nothing interesting to add in terms of the science.

The offset technique (HP Indigo) uses µm-sized particles in a liquid toner. There are, therefore, no dust-related speed restrictions and the prints are glossier. The toner particles are transferred to a heated blanket (with the usual advantages of this being a controlled environment with fewer compromises) where the bulk of the solvent (a hydrocarbon) is evaporated and recycled, and the particles (partially plasticised by the residual solvent) fuse into a glossy print. Further details of all those processes in the black box are not my concern here. The issue is the same as with the offset inkjet - why does the ink then adhere to the substrate when it is placed into contact?

The story is similar - the molten ink flows into good contact with the substrate and cools instantly to give adequate adhesion. However, it is well-known that Indigo adhesion is generally not sufficiently robust without a specific priming system.

---

[39] A practical approach to nanoparticle safety can be found in Steven Abbott and Nigel Holmes, *Nanocoatings: Principles and Practice*, DesTECH Publications, 2013

The trick is that the ink itself contains an acidic polymer. If the substrate is primed with a basic polymer then there is a substantial acid-base interaction that, at suitably low levels (remember, in adhesion too much of a good thing is a bad thing) gives strong adhesion. It is well-known that polyethylene imines are often used as primers for Indigo, though other basic polymers are available.

I recall that in the early days of Indigo, the lack of reliable adhesion was a real problem on some substrates and it required a large effort to produce good primer systems. The usual mythologies about surface energies confused things. Adding a corona treatment would, at best, increase adhesion by a factor of 42/32 (the ratio of surface energies after/before treatment) which is insignificant in terms of getting reliable strong adhesion.

My understanding of the in-line priming system available on Indigo presses is that they are a conventional corona/coating/drying system and although the system is proprietary, one can create a plausible logic for what is going on. The corona system creates ketone and carboxyl functionality on the surface of the substrate. The primer coating is something a like a polyethylene imine which reacts permanently with those oxygenated functionalities, leaving plenty of amine groups ready to react with the acids in the Indigo ink. Because polyethylene imine is a rather hopeless polymer, the primer needs to be pre-crosslinked to give it sufficient strength, but obviously the level of cross-linking needs to be low enough to allow the polymer to be soluble or dispersed in the coating solution. Again, the adhesion relies on the fact that lower levels of functionality (above a certain minimum) enhance adhesion - allowing very thin coatings to react with low levels of corona-induced functionality and to provide relatively low levels of acid/amine interactions with the Indigo ink. The system as a whole provides plenty of crack energy dissipation which increases adhesion.

This is all rather straightforward, but for historical reasons the whole chain of reasoning has been steeped in mystery.

Because there is so little real-world experience (at the time of writing) with the offset inkjet process it is too early to tell whether it will require special priming tricks (assuming the ink has the requisite functionality) or, alternatively, some sort of all-over coating to ensure that adhesion does not become a barrier to adoption.

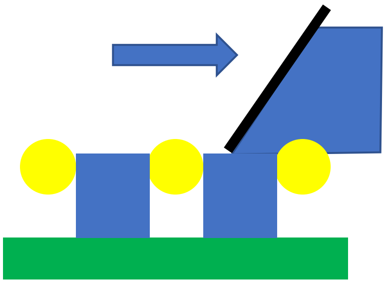
7.3 3D Printing

I had assumed that 3D printing would require a chapter of its own, not a subsection of "digital". To my surprise, I have found rather little science of the sort that is relevant to a printing science book. By this I do *not* mean that 3D printing is unscientific. There is plenty of science in sintering powders or extruding polymers through nozzles; it is just not the sort of science that is relevant to this book. There is also a large amount of mechanical engineering required in terms of analysis of the strengths (and, often, weaknesses) of 3D printed structures and lots of Finite Element work needed to understand the complex shrinkage effects inevitable in all the relevant techniques. This is all fascinating and important, but not something for this book.

So, with some reluctance, I finish what must be the world's shortest section on the science of 3D printing.

8 Screen Printing

The science of screen printing is simple and clear. It tells you that a well-set-up system needs almost no adjustments and that the print quality is predictable and reproducible. The reality of screen printing is that the science has been scrupulously ignored so that the majority of printers do the wrong things for the wrong reasons and make life unreasonably hard for themselves. Old Joe in gravure can do plenty of damage, but gravure is relatively simple so even Old Joe cannot be totally destructive. The Old Joes in screen printing have actively contributed to the decline and fall of the industry. In a period where screen printing saw excellent competition from inkjet, the last thing it needed was to stay stuck with bad habits. Yet this is what it did and so the market declined far faster than it should have done.



To most people in the industry the way screen printing works is simple: the squeegee (shown moving in the direction of the arrow) pushes the ink into piles between the mesh fibres which then flow together to give a slightly uneven print. This is wrong in every respect and is such nonsense that I had some difficulty in drawing the image. Dr Messerschmidt[40] in Germany had announced the correct mechanism in 1982 but his pioneering work had been entirely ignored. At a time when I was also unaware of his work, I came across a semi-official explanation of screen printing from a large organisation. It seemed to me to be entirely incorrect. I showed the document to colleagues (Professors Phillip Gaskell and Nikil Kapur) at U Leeds and they not only agreed that it was entirely wrong but in a few minutes worked out what the true mechanism should be. We were quickly able to confirm the essence of the mechanism, rediscovered the work of Dr Messerschmidt, and made a confident prediction based on the new theory. Colleagues at U Swansea quickly *dis*confirmed the prediction and we were able to dig deeper to find our error. A PhD thesis[41] confirmed the refined theory and a rather belated paper explained it in detail[42]. My then company, MacDermid Autotype fully backed the new theory and we tried our hardest to convince the screen printing world that printing was easier and more reliable if the theory was used. We even produced a book called *How to be a Great Screen Printer*[43] and gave it away to anyone who wished to read it. But in nearly every case, Old Joe won the battle and the theory has been mostly ignored.

---

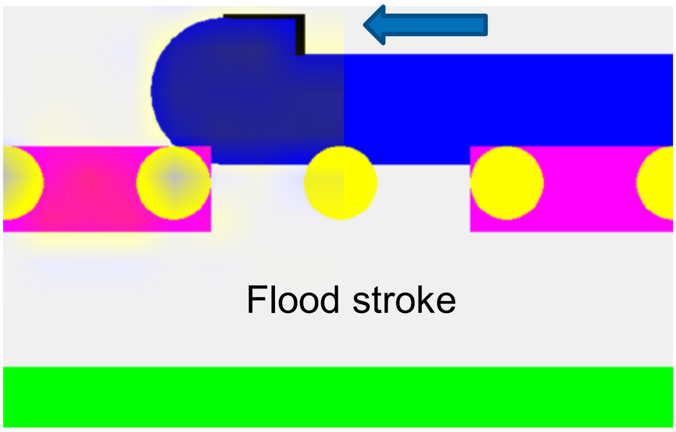
[40] E. Messerschmitt, *Rheological considerations for screen printing inks*, Screen Print., 72, 62–65, 1982

[41] E. D. Dolden, *Fundamental investigations into screen printing*, Ph.D. thesis, Univ. Leeds, Leeds, U.K., 2001.

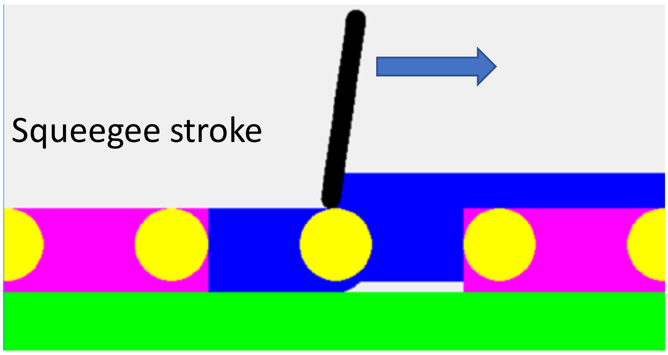
[42] Nikil Kapur, Steven J. Abbott, Elisabeth D. Dolden, and Philip H. Gaskell, *Predicting the Behavior of Screen Printing*, IEEE Transactions on Components, Packaging and Manufacturing Technology, 508 - 515, 2013

[43] MacDermid Autotype kindly gave me permission to host it on my website: Steven Abbott, Tricia Church, David Parker, Anna Harris, https://www.stevenabbott.co.uk/_downloads/How%20To%20Be%20A%20Great%20Screen%20Printer.pdf

If the theory was difficult or if it made the printer's life difficult, I might have some sympathy. Instead, the theory is easy and obvious, and once you follow it, printing becomes much easier and you can lock off many of the knobs that cause so many of the day-to-day quality problems in production. The images below are all from the Screen Print Animator (SPA) that I wrote when at MacDermid Autotype. Readers can download it (with kind permission from Autotype), along with the Flexo modeller from www.stevenabbott.co.uk/_downloads/SPAandFlexo.zip.

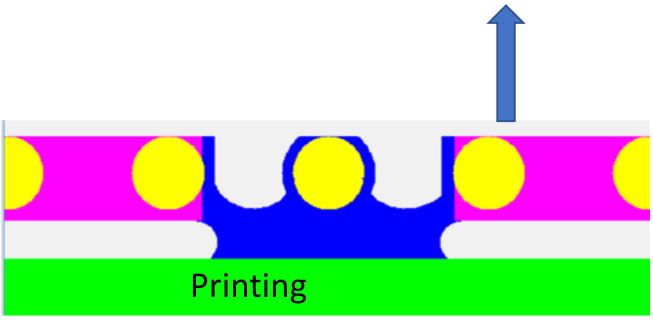


The process starts with the flood stroke. A relatively coarse blade pushes a fresh batch of ink over the top of the mesh and stencil. Many printers think that the flood stroke should fill the mesh, but this is clearly a mistake because the slightest over-fill will mean excess ink wrapping around the lower part of the stencil, causing dot gain. A more subtle reason for a light flood stroke is discussed later.



Next the squeegee does three things in one go. Note that although it is very busy doing those three things it does *not* do any printing. First, it pushes the stencil and mesh into contact with the substrate. This is quite hard work and imposes an unfortunate strain on the squeegee. Next, it pushes the ink into the gaps in the mesh and stencil. Finally, it scrapes off excess ink from the surface.

I once wrote an article that said that the squeegee had nothing to do with screen printing and there was a strong reaction against it, especially by squeegee makers. But it is obvious that the squeegee has not done any printing by the time it disappears off the edge of the image, while the mesh and stencil are still in perfect contact with the substrate. The only way printing takes place is when the mesh starts to rise from the substrate, by which time the squeegee is doing its three things on a distant part of the mesh. In some forms of electronic printing the mesh and stencil start off in contact with the substrate and after the flood, the squeegee has just two jobs - to fill the mesh and scrape off the excess. The printing only takes place when the substrate is peeled away from beneath. Although there are good practical reasons for the two different modes of printing, the results are identical. The squeegee really has nothing to do with the printing part of screen printing. We will come back to squeegee design shortly.



A famous question in screen printing that has caused many unhelpful discussions is "How does the ink come out of the mesh?" The true answer is: "It doesn't; the mesh comes out of the ink." If you pull a spoon out of a jar of honey you don't ask "How does the honey come off the spoon?" It is the same for screen printing - the mesh comes out of the ink. But exactly like the spoon coming out from the honey, it drags some of the ink with it. The ink flows to accommodate the motion and it ends up, as with the other splitting events discussed with other print processes, as a liquid bridge which eventually breaks.

If we look at the printing process in more detail we see something fascinating.

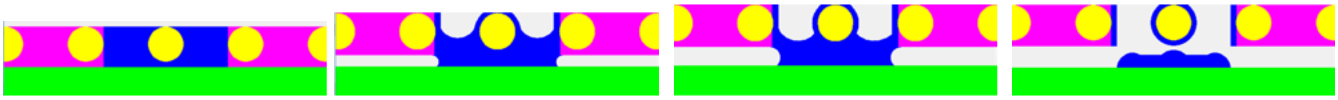


Figure 8-1 A small opening prints with liquid bridges between stencil and fibre

It is even clearer if we show what goes on with just the mesh

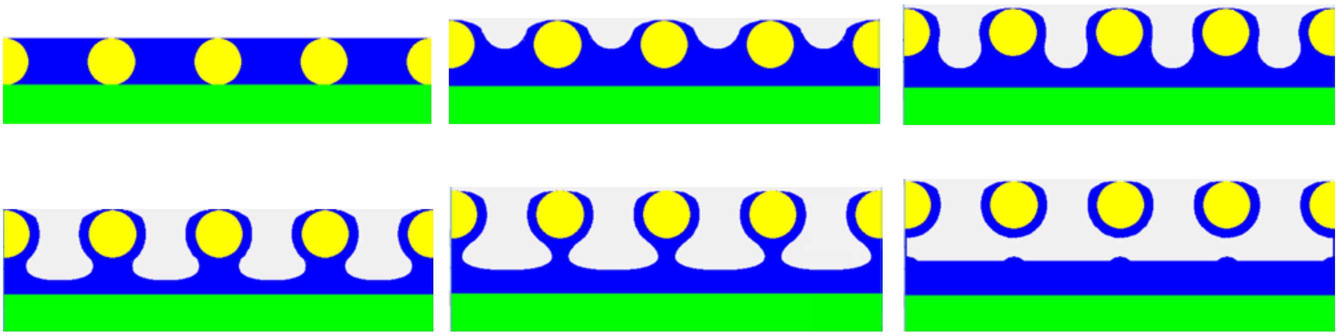


Figure 8-2 An open mesh prints, showing the steady development of a liquid bridge which snaps, leaving an extra dot of ink below the fibre.

The fascinating thing is that the largest ink deposit is exactly underneath the mesh fibre, in complete contradiction to the folk-physics idea of the ink being printed as columns between the fibres. It takes only a little bit of work with a microscope to confirm that the thickest ink is exactly under the fibre (actually, under every other fibre because of the under/over weave) rather than in between. The same process provides (unlike the folk-physics ideas) an explanation for ink splatter sometimes seen on prints, due to the unstable drops created when a liquid bridge snaps. It also explains why some inks give "cobwebs", i.e. when the liquid bridge carries on extending rather than snapping.

Note that at the end of the print, the mesh has plenty of ink wrapped round it. I mentioned that our original theory contained an error. If you think of pulling a spoon out of honey, you can imagine that pulling it very slowly allows the honey to flow back, while pulling it out quickly will leave a large amount on the spoon. We had thought that the amount of ink on the mesh would increase with print speed, and, therefore, that the amount printed would reduce at high speeds. The work at U Swansea showed that the printed ink deposit was independent of print speed (within a practical speed range). That caused us to re-examine the theory of how much ink gets dragged out. It turns out that we were right up to a certain combination of speed and viscosity. With reasonable speed and high viscosity, the amount dragged out is independent of either. Because of the high viscosity of screen inks and the relatively high speed with which the mesh is pulled out of the ink, screen printing is always far above the critical point. Which means that the print is entirely independent of any meaningful process variable - providing, only, that the squeegee is able to fill the mesh and scrape off the excess.

The true physics explains what had previously seemed rather mysterious: that the slowest print on a hand bench and the fastest print on a high-tech press were identical. This fact should be a cause for celebration. Any process that is basically independent of machine settings is intrinsically stable and should give reliable, high yield. It turns out that the Old Joes do everything possible to fight against this good news; otherwise (as they see it) they would be out of a job.

We can now introduce the key screen print equation. It is simple and obvious, yet printers like to resist it up to the point that their business closes down:

$$\text{Ink}_{\text{Printed}} = \text{Ink}_{\text{In Mesh at the Start}} - \text{Ink}_{\text{On Mesh at the End}}$$

Equ. 8-1

All the equation is saying is that you fill the mesh area (defined, if necessary, by the stencil) with a certain amount of ink; then the ink that remains on the print depends on how much remains wrapped around the mesh fibres (or on the edge of the stencil). Because of the physics which, for screen, makes the amount pulled out independent of speed and viscosity, the $\text{Ink}_{\text{On Mesh at the End}}$ is ~40%, i.e. screen is not so far from the 50:50 split common to all the other printing techniques.

As most printers had imagined that the mesh ended up with just a tiny % of the ink, this 40% value is quite a surprise. If anyone is interested, it is trivial to confirm it. Take a polymer A4 sheet, weigh it, then print a large rectangle of UV ink (it is easier than using a solvent-based ink), and re-weigh the sheet. You can now calculated $\text{Ink}_{\text{Printed}}$. Weigh a convenient wiping cloth, soak it in a solvent such as acetone and wipe the inky mesh clean. Allow the solvent to dry and re-weigh the cloth, from which you can calculate $\text{Ink}_{\text{On Mesh at the End}}$, and therefore the original total amount of ink. You will find a roughly ~60:40 ratio between the two weights. The thesis by Elizabeth Dolden contains many such measurements on a variety of inks and confirms this general finding.

Earlier I mentioned that a light flood stroke is desirable for a subtle reason. We now know enough to work it out. With a solvent-based ink (the solvent could be water), there is a danger that the ink wrapped around the mesh between prints will start to dry out. The only way to avoid the same ink sitting there print after print is to swish it away with fresh ink. If you try to fill the mesh with the flood stroke, you have to be rather gentle, so the filling process is unlikely to swish away much of the old ink. With the squeegee filling you can be (indeed, it is desirable to be) as fast as possible in the process, pushing the ink into and under the mesh (it's hard to draw it but it's obvious in 3D). This gives the best chance of displacing old ink with fresh ink. This approach to avoiding drying in of difficult inks had been known (but not to me) to some in the screen world and once we had a good theory of the whole process, it had a chance to become a more standard practice - though not with Old Joe who "knew" that the flood was to fill the mesh and that the squeegee did the printing.

8.1 The golden rules of screen printing

They are very simple:

1. Let the mesh do the metering
2. Let the stencil do the shaping
3. Let the ink do whatever is the primary function of the ink.

The amount of ink to be delivered should be controlled entirely by the mesh and the ink equation. Using the stencil to help control the amount of ink is a bad idea because the stencil can only increase the deposit at the edges of a print - with the danger that an insufficient flood or inadequate squeegee pressure failing to fill at the edges, giving "negative sawtoothing".

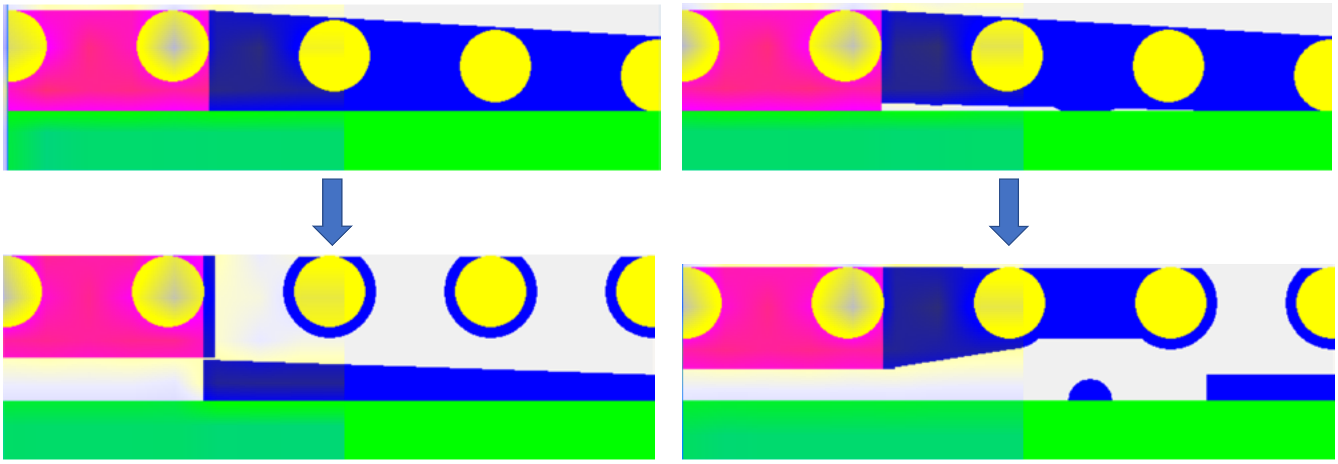


Figure 8-3 A thick stencil either gives a thick edge and thinner middle (left) or "negative sawtoothing" where the squeegee failed to fill to the extra thickness. Both are undesirable

So it is always the case that the thinnest-possible stencil is the ideal - provided it is nice and smooth. Roughness is usually measured via a stylus profilometer and the Rz values is the usual metric (see https://www.stevenabbott.co.uk/practical-coatings/surface-profile-explorer.php for an explanation). If it is rough (has a high Rz) then the squeegee can force the ink into the roughness and you get ragged dot gain ("positive sawtoothing") at the edge of your print.

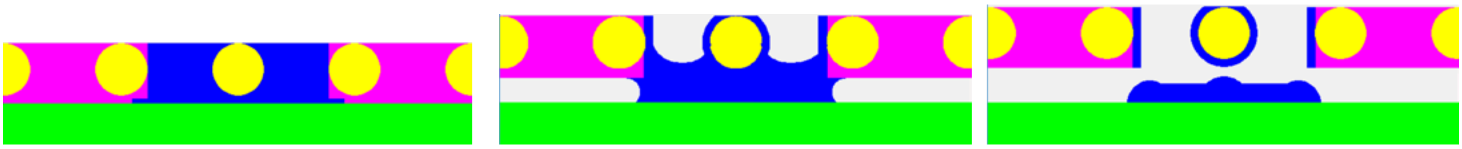


Figure 8-4 A high roughness (Rz) allows ink to get under the edge of the stencil, giving ragged dot gain ("positive sawtoothing").

It happens to be easy to make a thin, rough, stencil and easy to make a thick, smooth stencil. It is hard to make a thin, smooth stencil. However, by getting into the habit of making such stencils, the stencil is simply doing the shaping and not trying to influence (for good or bad) the metering.

If you insist on having a rough stencil then to reduce the risk of ragged dot gain you have to engineer the ink to be so thick that it hardly flows into the roughness of the stencil. But you want the ink to be relatively thin so that the squeegee has no problem filling and scraping. By having a more viscous ink, to help the squeegee you have to have a harder flood fill, but this gives the risk of over-filling and of ink drying onto the mesh. The extra stuff added to the ink to make it more viscous might fight against other desirable properties such as colour or conductivity. So a single bad decision ("I'll get the ink to help the metering and shaping") leads to multiple problems. By letting the mesh do the metering and the stencil do the shaping, the ink can be tuned to give the best colour, conductivity or whatever property it is intended to have.

Finally, those who choose the wrong mesh which delivers the wrong amount of ink have to play rather difficult games to get the correct deposit. If the mesh delivers too much then the squeegee has to try to dig out some excess from the otherwise smooth surface. This is hard to do and puts a great strain on the squeegee. If the mesh delivers too little then you have to get the squeegee to slide slightly above the mesh, not quite scraping off all the ink. Good luck trying to do this reproducibly. The physics is clear - you are dancing on a knife edge.

Your squeegee supplier will love you if you adopt either of these strategies as you will forever be trying new ones, throwing away old ones and generally spending a lot of money fighting the laws of physics.

With a thin, smooth stencil, the mesh with the right thickness to deliver the desired amount of ink, the right ink to give you the required properties, there is nothing much for the press or the operator to do. It is all very simple and reproducible.

Under these circumstances, just about any reasonably sharp squeegee that is resistant to your specific ink (you don't want the rubber to swell) under moderate pressure and a moderate angle will work all day, every day, with no need to fiddle, adjust, clean, sharpen etc. Your squeegee supplier will be unhappy, and Old Joe will be unhappy. But your customers will be happy and your business will be more profitable. It is as simple as that.

8.2 Dot gain in screen printing

Dot gain in offset, flexo and gravure is relatively small because the printed dot is just a few µm thick so dot growth by pressure in the nip or from the classic drop spreading is small. We saw that inkjet shows potentially large growth because the original dot starts with a high contact angle, θ, and the spread of the dot goes as θ^3, giving us a significant problem.

As the images from the SPA indicate, screen lines or dots also start off with a high θ, so spreading is very rapid. Every time I examined a dot or line immediately after printing, I confirmed my intuition that the viscous ink gave very little spreading. It was, therefore, a shock to discover that (especially) my printed conductive tracks had doubled in width in the short time between printing and examining them - in much less that 1s.

You can re-confirm this with the dot spread modeller. The majority of the spreading takes place in a very short time.

8.3 Printing on previous prints

With other printing techniques there are always problems of printing on top of previous parts of the image - CMYK dots or printed electronics. These problems are relatively simple because the previous printed portions are just a few µm thick. Screen printing cannot naturally print thin features. A typical mesh delivers 15-20µm of wet ink so if it is 50% solids (screen inks are relatively viscous so are rarely in the 10% solids range) that is still a final ~10µm thickness. An image from the SPA shows what happens when you try to print onto another dot:

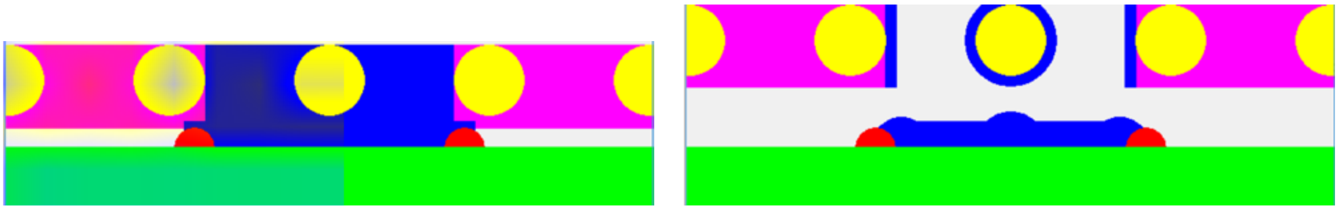


Figure 8-5 Printing the blue ink onto the previous red dot gives lots of dot gain.

The problem is that the rather high previous dot gives lots of room for the ink to be forced sideways and also to give a thicker deposit, i.e. giving double dot gain. If you try to fix the problem by providing less squeegee pressure then you fail to fill the mesh fully so that on separation you just have a little dot remaining, plus some extra ink on the side of the previous dot. This is the phenomenon of "skipping"

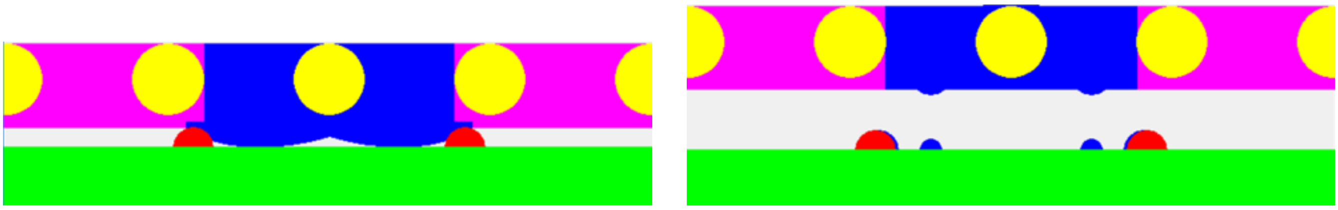


Figure 8-6 Trying to fix the problem with a lower squeegee pressure gives you "skipping" instead.

There is no cure for these problems other than using the finest-possible mesh that gives the thinnest-possible dot ("let the mesh do the metering"), with the thinnest-possible stencil to avoid printing extra ink. This requires the ink to be more highly pigmented to obtain the same printed colour strength, but this isn't a problem for screen as high viscosities are readily printed *if the ink is highly shear-thinning during the squeegee filling process.* This brings us back to the theme of making things as simple as possible for the ink. Adding things to try to fix other problems (such as too much dot gain from a rough stencil) make it difficult to do what a highly-pigmented system does very naturally which is to shear thin and then to recover viscosity quickly to avoid dot gain.

At one time, this point about highly-pigmented inks was theoretically obvious but I had never had a chance to prove it. One day I had the chance to do an exact side-by-side comparison of two silver conductive inks, using identical optimum stencils that had a low roughness and low thickness. The aim was to print very fine lines. The first ink was wonderful. With no effort we got beautiful fine lines. The second ink was very disappointing. No matter what we did, the lines showed a lot of dot gain. The first was a "cermet" ink that was mostly metal + solvent. It was designed to become conductive and adherent when fired at high temperatures. Such an ink shows dramatic shear thinning and very rapid recovery to a high viscosity, limiting the chance of dot gain before the solvent had evaporated. The second was a typical polymer-based ink for use on polymer substrates. Its shear-thinning curve was much less dramatic and it took a long time to recover its viscosity, so it gave high dot gain.

That was many years ago. Many modern polymer-based inks now use "associative thickeners" which show sharper reductions in viscosity and faster recovery. With a thin, smooth stencil these give low dot gain. But if you have a thick and/or rough stencil these advantages turn into a disadvantage because the low viscosity causes lots of spreading under the rough stencil.

It is very strange that printers are prepared to spend a large amount of time and money trying to get an improved print with a lower-cost stencil, rather than spending a little more time and money creating a thin, smooth stencil (via repeated wet-on-dry emulsion coating or a sophisticated capillary film) which solves many problems in one go.

I once gave a lecture on screen printing to a major high-tech company. Their experienced printers told me that I was wrong and took me to their printing shop to show me some of their poor-quality prints which, they said, refuted my claims. It was immediately obvious to me what was wrong - they had massive positive dot gain from a rough stencil. "What is the Rz of your stencil?", I asked. "We don't know, we get our stencils made by an outside company who told us that they were optimum". So here were some high-tech people, trying to print demanding electronic components, completely ignorant of a key parameter in their process - all because they had outsourced the "unimportant" task of creating the stencil.

8.4 Mesh myths

The mesh plays two positive roles:

1. Holds the stencil in place
2. Meters the amount of ink that is printed

It also plays four negative roles:

1. Getting in the way of fine lines
2. Swelling/staining from previous prints, which can mark subsequent prints
3. Interfering with the exposure of the stencil
4. Stretching (creeping) over time and spoiling registration

Clever surface treatment can help provide adhesion to the stencil without making the surface too easy to stain. Such improvements have been one of the significant steps from the polyester mesh manufacturers.

For polyester meshes, a UV absorbing colour can reduce the scatter and reflection that interferes with the exposure. For stainless meshes, a special blackening treatment gives a dramatic improvement in the quality of the exposed stencil.

Stainless meshes are far more resistant to creep than polyester, though advances in fibre technology have helped somewhat to reduce polymer creep. "Liquid-crystal" polymer meshes have much lower creep but unfortunately the strength in the lengthwise direction creates a weakness in the cross-thread direction so the fibres become very squashed, giving a relatively thin, fat thread which reduces the capacity for high-resolution printing.

Precision weaving and subsequent treatment (e.g. controlled stretching and calendaring) can give a uniform metering of the ink. From the $Ink_{Printed}$ formula it is obvious that we need to know the true volume of ink in the mesh at the start and the exact % of ink that remains on the mesh at the end, which will depend on the subtleties of the fibre shape and weave. You would think that the mesh manufacturers would provide these values. But they don't. The provide a bogus value that relies on a clever image but bad science. The reality is described below.

Making the fibres as thin as possible helps reduce the interference with the printing of fine lines. Clearly there is a trade-off because thinner can also mean more stretchy. But if you need to print very fine straight lines, for example, you have no choice but to use a thin mesh, as the following diagram explains.

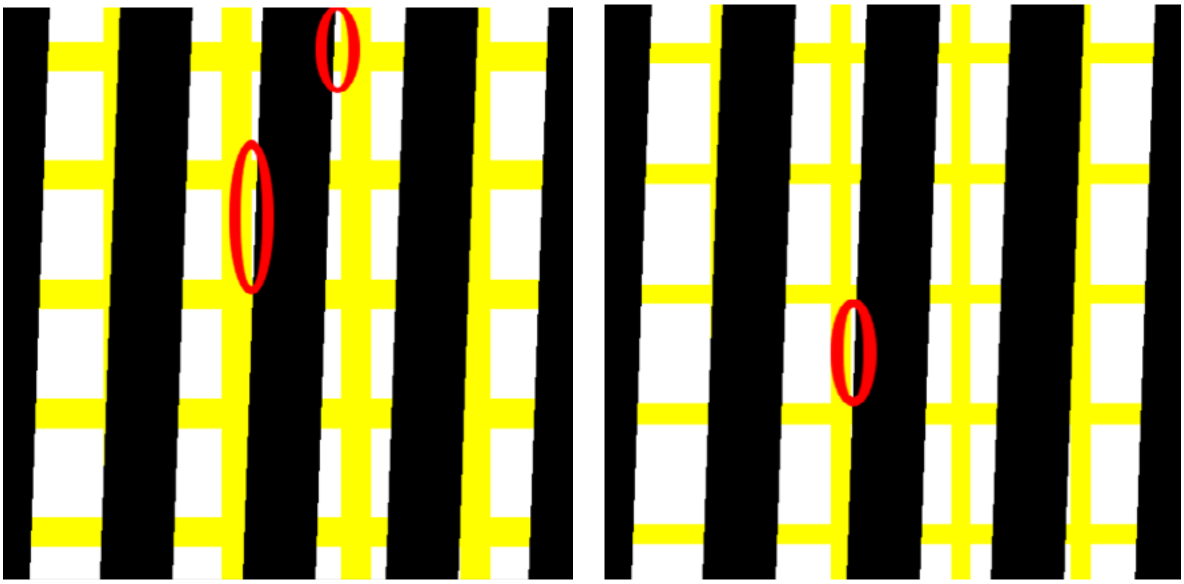


Figure 8-7 The thicker yellow mesh on the left has more points where the ink cannot squeeze between the mesh and the black stencil - so the lines will be more ragged then those with the thinner mesh on the right.

For serious fine-line printed electronics, the logic has been clear for a long time that the only possible mesh would be a blackened (for high resolution), sub-20µm stainless mesh. It was a considerable surprise to me that Western printers stayed stubbornly with thicker stainless while the finest Asian printers were heading down to 13µm meshes. The excuse used was that fine meshes were difficult and expensive. Well, if you insist on printing badly with the wrong stencils and the wrong physics (and, probably, excessive squeegee pressure), these meshes are too hard to use. But with a thin, smooth stencil (also said to be too difficult and expensive), and with "in contact" printing so that the squeegee does not have to force the mesh into contact with the substrate, printing fine lines on to such a mesh is not so hard, especially with a well-designed rheology for the ink (fast shear thinning and rapid recovery).

It is all very clear and simple. It is surprising, therefore, that mesh manufacturers have tried hard to focus on other, irrelevant or erroneous factors, rather than on giving their customers the meshes and data that can give them the best possible prints for the price. Two mesh myths have been especially bad for the industry.

The first myth took a number of years to rise and fall. Some manufacturers decided to promote a mesh that had a surface coating that "printed more ink". They demonstrated this at trade shows by contrasting normal mesh with treated mesh, showing that drops of water sat on the normal mesh but went straight through the treated one. Approximately 1 second's thought would have shown that this was nonsense. Even if more ink came through on the first print (which it wouldn't), after that first print, the surface is "contaminated" by the ink so subsequent prints only "saw" ink and would have no knowledge of the surface treatment. But Marketing is more powerful than science so the nonsense was promoted for a number of years.

Some companies could prove to me that their treated meshes gave a different amount of ink compared to the same specification (thread count and thickness) of untreated mesh, though the differences were rather subtle. It could readily be shown that these differences arose because the surface-treated fibres slipped and bedded slightly differently during the finishing processes so the volume of ink in the mesh at the start was slightly different and the amount retained on the mesh was also slightly different. Imagine what would have happened if all the useless money spent on Marketing of a non-existent effect had been spent in optimising the surface treatment to give whichever printed ink volume the customers actually required.

The second myth is still around. If you look at the mesh from above, the hole is approximately a square of an easily calculated open area, x^2 in the diagram below. Knowing the thickness of the mesh, 2D, an "ink volume" can be calculated by multiplying area by height, i.e. $2Dx^2$. A few seconds' of thought shows that this volume is a nonsense. You cannot calculate the volume of a complex 3D structure by an arbitrary square and height.

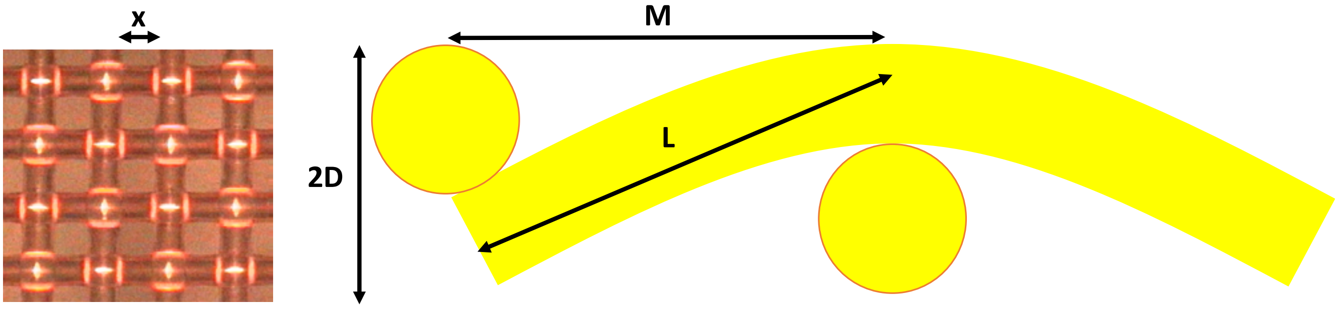


Figure 8-8 The common but erroneous ink volume calculation based on $2Dx^2$ on the left, and the geometry for the true ink volume on the right.

The real volume (TIV, the Theoretical Ink Volume) in the complicated 3D volume of the mesh can be calculated using elementary geometry based on the length of the segment L, the diameter of the fibre, D, and the mesh spacing, M. Conveniently we can get rid of L and M by using TPM, the Threads Per Metre value that is specified for all meshes

$$TIV = 2D - \pi \frac{D^2 TPM \sqrt{1 + TPM^2 D^2}}{2}$$

Equ. 8-2

The exact ellipticity of the mesh (a key factor in the manufacture), with an effective diameter in the vertical direction of $D_{\text{Compressed}}$ and a rather stretched TPM has a significant effect on the calculated volume.

$$TIV = 2D_{Compressed} - \pi \frac{D^2 TPM_{Stretched} \sqrt{1 + TPM_{Stretched}{}^2 D_{Compressed}{}^2}}{2}$$

Equ. 8-3

A final correction takes into account the amount that the squeegee will naturally remove from the top surface. This is somewhere between D/4 and D/2, with the latter giving the final formula:

$$TIV = \frac{3D_{Compressed}}{2} - \pi \frac{D^2 TPM_{Stretched} \sqrt{1 + TPM_{Stretched}{}^2 D_{Compressed}{}^2}}{2}$$

Equ. 8-4

Two meshes that have the same apparent pseudo volume might have significantly different real volumes. As printers print real volumes of ink, you might think that manufacturers would quote these volumes (and their own best estimate of the % ink remaining after printing) to help their customers. But they do not. I have no idea why this is so. I am certain that it has positively harmed the industry and stopped the better manufacturers from creating exciting new products through control of the real parameters.

8.5 What use is screen printing?

Screen could never compete with offset, flexo and gravure. For some time, the limitations of inkjet meant that screen was the choice for lots of prints on "difficult" substrates such as the sides of trucks or large signs in supermarkets. Gradually print runs decreased (customisation) and inkjet's capabilities increased, so a large chunk of the large-area market is lost forever.

Screen, therefore, is only defensible in areas that play to its two key strengths:

1. Ability to print at almost any viscosity above a certain minimum onto just about any substrate, therefore allowing just about anything (cake icing, diabetes sensors, thick-film ceramics, silver tracks on solar cells, flags ...) to be printed

2. Ability to print a relatively thick layer of ink where that is desirable (cake icing, diabetes sensors, thick-film ceramics, silver tracks on solar cells, flags ...) onto just about any substrate.

Because its two strengths are impressive, it is deeply unfortunate that the lack of good science in screen printing gave it a reputation as a technique just good for T-shirts. A great deal of money was wasted on inkjet as a no-hoper technique for much of printed electronics. The sales pitch of functional drop-on-demand was never matched by the need for many functional inks to be the sorts of thick pastes that inkjet can never print. If a fraction of the money wasted on inkjet had gone into improving screen presses and inks, printed electronics would be in a happier place.

It is still not too late. It took about 20 years for the hype of inkjet to become clear to almost everyone. Fortunately, the hype of printed electronics taking off "real soon now" just about matched it, so inkjet's failures were relatively unimportant (except to investors). Now that (beyond diabetes sensors and solar cell tracks) printed electronics is starting to be more realistic, taking place behind the scenes in modest, specific projects rather than in grandiose follies, an assessment of what high-spec screen presses can achieve should convince (and is convincing) more printers to use the intrinsic merits of screen.

An exciting recent development is "offset screen" for printed electronics. The printing is done onto a perfect substrate where everything (especially dot gain) is controlled and the ink is dried/cured to a satisfactory level. It is then, as with offset electrophotography and offset inkjet, transferred to the final substrate. The technique is expensive. But as part of a $billion line for producing display screens, the specific advantages of screen compared to the (non)alternative techniques mean that the costs are almost irrelevant.

It is, of course, not a question of screen versus inkjet versus gravure/flexo. It is a question of which technique, when analysed scientifically, gives the best price/performance capability for which part(s) of a device. If you need a 100nm layer of a special polymer, inkjet might well be ideal. If you need a significant conductivity in fine silver-based tracks then inkjet is less likely to be useful than screen.

And that is the beauty of modern printing systems. To a computer-controlled multi-layer printing system, it doesn't matter if one step is flexo, another is inkjet and another is screen. All the system needs is a set of rational instructions to control each step. If screen was a process requiring Old Joe to be twiddling knobs, it could not fit in. But a proper screen press has no adjustable knobs because there is nothing that can be adjusted if the mesh is doing the metering, the screen the shaping and the ink is delivering its desired functionality. Screen is then just another rational form of printing.

9 Colour science

The first time that a colleague and I realised that we needed to understand colour theory, we thought it would take us a few hours to learn. How ignorant we were! Having struggled to understand it ever since, this chapter is my attempt to make it as easy as possible for those interested in printing (including myself) to understand the main issues and perform whatever calculations are required.

If our colour vision worked like a camera, life would be much easier for colour theorists but harder for us. Over a surprising range of different lighting and shadow conditions, objects to our eyes retain a constant colour, even though an objective measure of the photons arriving at our eyes would show that colours are constantly changing. So our vision is based on an active model by the brain which (for good reasons) fools us into thinking that things don't change colour for no good reason. The reason we can be easily fooled by colour illusions (the same colour looking different depending on the colours around it) is that our internal model never had to cope with entirely artificial colour combinations that lack the context to tell us what the colour should "really" be.

Fortunately, for most printing purposes, the objective colour is our main concern so the complexities of how our brains perceive colour can largely be ignored. One amusing consequence of our brains knowing what the colour should be is that most of us would have no problem identifying the red of a well-known cola drink or the yellow of a hamburger brand, even if specific examples of that red or yellow covered a rather large range. The print buyers for such well-known brands spend a lot of time worrying about the precise colour match, even though the exact colour is of zero concern to the consumers who already know what it looks like.

As with previous chapters, I am going to assume that you have a certain basic knowledge, in this case you know the light is a spectrum of colours from the blue end above 400nm up to the red end towards 800nm and that you can objectively measure the amount of each colour with some sort of spectrophotometer which splits the light into individual bands of colour (with greater or lesser resolution) and measures their relative intensities.

9.1 What our eyes respond to

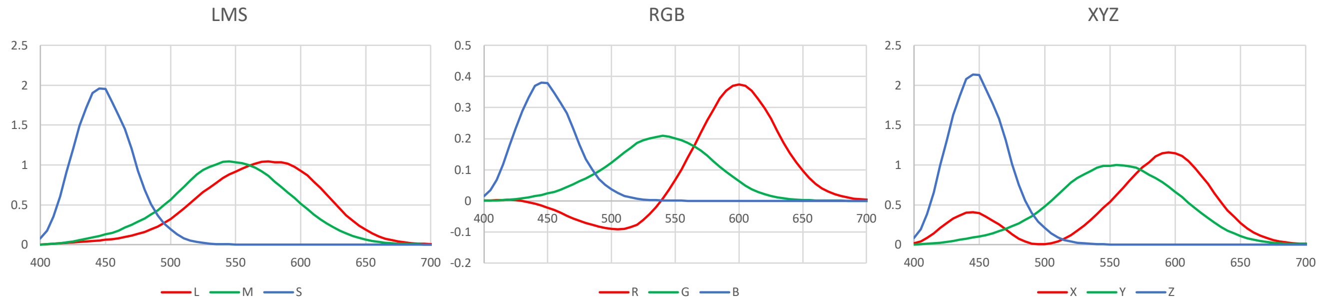


Figure 9-1 Our cones respond like LMS (Long, Medium Short wavelengths), we can simulate with RGB (Red, Green, Blue) light and we can calculate using XYZ (deliberately meaningless letters).

Here is the problem. If we expose our eyes to pure spectral light then we can work out the sensitivities of the three cones that are sensitive to different bands of colours. As shown in the LMS graph, the Short wavelength cones are sensitive to the blue area around 450nm, the Medium to the yellows and greens around 540nm and the Long wavelength cones are, well, not very different from the Medium ones, with a peak at 580nm. This is not a limitation of biology: bees, for example, have a "scientifically" optimal spacing, just as our RGB cameras do. Explanations for the "unscientific" spacing of the cone wavelengths cluster around ideas that we have superior red/green detection that is useful for detecting things such as fruit ripening. Whatever the reason, it is this odd spacing which causes so much trouble for colour theorists.

The next chart shows what mix of pure R, G and B light you need to shine to match the colour of a pure spectral line. The fact that the L and M cones overlap give us the uncomfortable truth that it needs a negative amount of R to reproduce some bright greens. What this means experimentally is that you have to contaminate a bright green with some extra red in order to get a visual match.

The RGB "colour matching functions" are all we need to be able to discuss any colour (possible and impossible), but they are not very convenient because of those negative curves and because they don't link in numerically to another fact that our overall sensitivity to light intensity follows the M curve. So the XYZ space was created, deliberately using letters that bore no resemblance to RGB. It is merely a mathematical transform of the RGB curves that (a) eliminates the negatives and (b) gives us a Y curve that defines "luminosity" - how bright we think a colour is.

Note that the problem of the negative values didn't disappear - it just shifted so that the X curve while being mostly in the red has an inconvenient peak in the blue region.

Note, too, that what I am calling RGB and XYZ are often labelled r̄, ḡ, b̄ and x̄, ȳ, z̄. There seems to be lots of inconsistency in how they are used. For our purposes the simpler versions are good enough.

9.1.1 What is *that* colour?

The XYZ curves are the objective way of talking about colours and can perfectly describe any pure, single-wavelength colour in a single [X,Y,Z] triplet. That is of no use to us in the printing world as we seldom have to worry about these pure colours. A real, printed colour, is reflecting a broad range of wavelengths at different intensities, yet comes across to us as being a single colour. How do we get an [X,Y,Z] triplet from this broad range?

The trick is to find how much total X-ness, Y-ness and Z-ness are in that light.

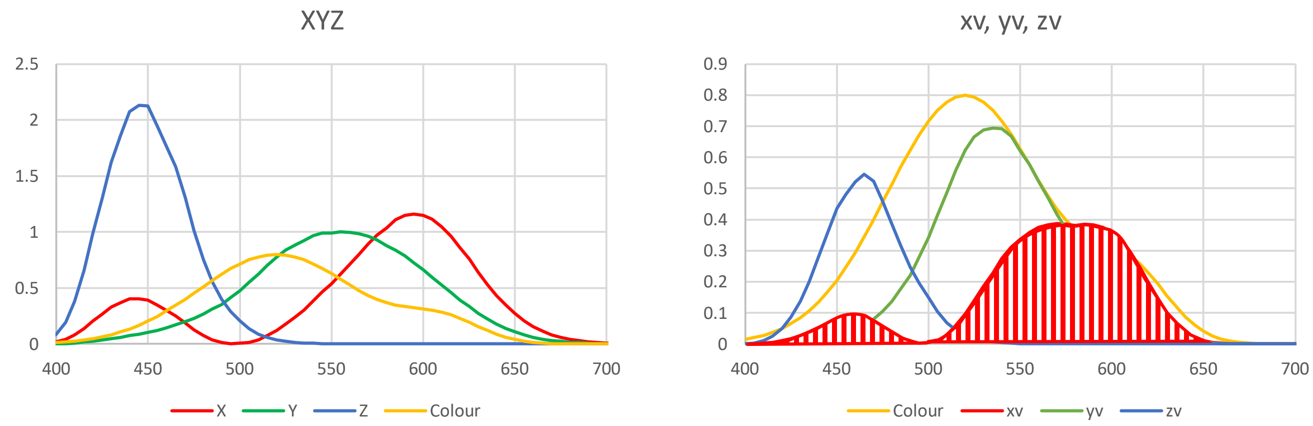


Figure 9-2 Our colour is compared to the XYZ curves and the individual xv, yv, zv values can be calculated and summed to give the XYZ value of that colour.

In the image, the colour shows a peak at ~520nm with an extra peak around 600nm so there is no obvious way to say what its colour might be. Instead we multiply each of the XYZ curves by the colour value to give xv, yv, zv curves, then sum the amount of X-ness in the xv curve (as shown via the hatched lines) and the Y-ness in the yv curve and Z-ness in the zv curve. The XYZ triplet is then those xv_{sum}, yv_{sum}, zv_{sum} values normalised against the sums of the X, Y and Z curves. In this specific example we find that the colour is [0.32, 0.51, 0.24].

So to get an objective measure of any colour you "simply" find the amount of light coming back at any wavelength and do this xv, yv, zv trick (technically called integration) and we have our XYZ triplet.

There are, inevitably, two problems. The first is that the amount of light being detected by our spectrometer or colorimeter does not depend only on what has been printed. It depends on the light falling on the sample. If you looked in pure R, G or B light your spectrometer would only see a single component. If you used a yellow-ish tungsten lamp you would automatically have more yellow than if you used the sort of bluish lights popular in stores selling white goods.

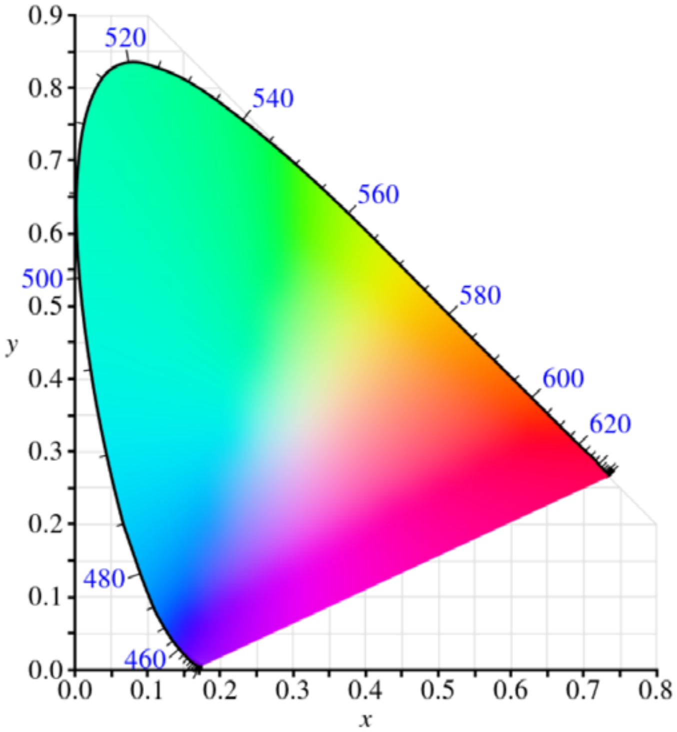
Your "objective" XYZ triplet is only objective if you specify the light source used for the measurement. Typically this is D65 light, a "daylight" specified at a colour temperature of 6500K (actually 6504K). If you want to measure it in D50 light (more yellow because it is at a lower colour temperature of 5000K), that's not a problem, as long as you specify that that is what you are using. If you have a D50 measurement and someone wants the values in D65, the conversion is straightforward.

The real problem is not so much getting the XYZ values from our specified light source. Instead it is that XYZ numbers mean nothing to most of us and we have no convenient way to plot them, given that colours in 3D space are very hard to grasp.

This means that we have to find some way to humanize XYZ by bringing it into a 2D space. And this gives us a whole new set of problems.

9.1.2 Colour in 2D

Because XYZ defines Y as the luminance, and because instinctively we have some idea of what that is, we can take Y for granted and define x and y via $x=X/(X+Y+Z)$ and $y=Y/(X+Y+Z)$. Although $z=Z/(X+Y+Z)$ it is trivially defined as $z=1-x-y$.



Using Wikipedia's version of the xyY colour space[44] (i.e. we plot x and y and assume that the luminance is controlled independently by Y) we have the familiar CIE xyY diagram:

The pure colours from single XYZ values define the edges, everything in between is a mixture that would be calculated via the integration procedure described earlier.

There is nothing especially wrong about the xyY plot; it just isn't very useful when we want to think about colour variations or colour differences. A small change in the low x,y region means large changes in blues, while

---

[44] Attributed to BenRG

such a change in the 0.1, 0.8 region just makes a slight difference to a set of rather similar greens. This means that we cannot have any uniform measure of colour difference (e.g. based on $(x_1-x_2)^2+(y_1-y_2)^2$. There is also no way to think about opposite colours. Moreover, although we think this is a familiar representation of the truth is is only of colours at large Y values. Other real-world colours are found at lower Y values, where the plots aren't as familiar.

So we have to do another transformation to XYZ for something rather more useful because (a) it defines two sets of opposite colours: magenta-green; blue-yellow and (b) it creates a (reasonably) uniform colour space where similar numerical differences imply similar visual differences; this in turn allows (c) a reliable measure of colour differences. There are plenty of such spaces, but the most common in printing is Lab. Or is that CIELAB (yes) or L*a*b* (yes)? As far as I can tell, we are supposed to use L*a*b* to distinguish it from an earlier Lab devised by Hunter. They both do the same thing, but L*a*b* does it slightly better and is the de facto standard. So I will talk about Lab, noting that when people talk about the Hunter version they call it Hunter Lab to distinguish it from Lab, meaning L*a*b*. Colour nomenclature is very confusing.

The most interesting thing about the Lab conversion formulae is that they rely on X_n, Y_n, Z_n which are the XYZ values for your desired white light, such as D65 or D50. This means that Lab values are meaningless without specifying which white is being used.

Equ. 9-1

$$L = 116 f\left(\frac{Y}{Y_n}\right) - 16$$

Equ. 9-2

$$a = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right)$$

Equ. 9-3

$$b = 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right)$$

The function f(t) distinguishes a normal zone where $f(t)=t^{1/3}$ if $t> 0.008856$ and $7.787t+0.138$ otherwise. If D65 is used then the X_n, Y_n, Z_n triplet is [95.05, 100, 108.88]

The ugliness of the xyY space is not fixed by conversion to Lab. That is why we generally hear only that -a values are green and +a values are red, while -b values are blue and +b values are yellow. Lab's real use is as a way to get linearised colour measurements where we have some idea that a shift in a *b* value of 6 is going to have an effect similar to a shift in an *a* value of 6. More importantly we can calculate a colour difference, ΔE (technically ΔE^*_{ab}) as a geometric distance between L_1, a_1, b_1 and L_2, a_2, b_2:

$$\Delta E = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2}$$

Equ. 9-4

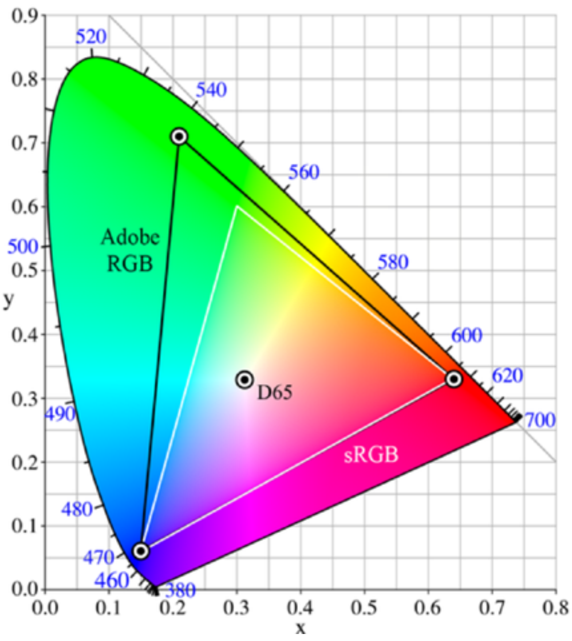
It is generally agreed that a ΔE<2.6 is not noticeable, which, in addition to measurement errors, stops us worrying too much about anything smaller than 3. This makes sense because L values go from 0 to 100 and a and b values go from (approximately) -100 to 100, so a difference of 3 in just one of the variables is in the 1.5 to 3% range.

Note that ΔE^*_{ab} is increasingly being replaced by ΔE^*_{94} and ΔE^*_{00} (for 2000) which use L*C*h* and L'C'h' respectively. There is also CMC (or ΔE(CMC)) which is also based on L*C*H*. Those colour spaces are rational alternatives to Lab, using different conversion formulae. The claim that a ΔE^*_{ab} <2.6 is not significant would be OK if Lab was a truly even space. But there are areas where the same ΔE^*_{ab} has a different impact to expert eyes. The alternatives are more even, and CMC has two variants [CMC(1:1), CMC(2:1)] which appeal to different industries or experts. The principles are the same. Use whichever ΔE measure is conveniently present on your own device and which is used by your supplier(s) and customer(s). The evidence is that ΔE^*_{00} and CMC more accurately show up differences visible to human eyes, and the criteria might be more exacting differences of 1 or 0.5, but the advantage is irrelevant if those you are dealing with only have ΔE^*_{ab}.

9.1.3 Specifying colours

At some point, someone makes a decision that *this* is the image that needs to be printed. In an ideal world, each pixel of that image would be specified as [XYZ] and the printed version could be checked by comparing the Lab values directly converted from the [XYZ] to the values measured with a spectrophotometer.

The real world does not operate like that. Even if the original image existed in the physical realm, it has to be brought into the digital domain as a set of triplet values that can be converted into printer-dependent instructions to create a plate or transfer some drops of ink.



In an ideal world our digital colour processing would all be in XYZ or Lab. In the real world, everyone uses RGB - though confusingly, not the RGB from which XYZ was created (which has negative R values), but an approximation which assumes that all colours can be created by mixes of R, G and B going from 0 to 1. The assumption is completely false; any computer RGB colour space is highly restricted compared to XYZ. And the one most regularly used across many industries, sRGB (standard RGB) is very highly restricted. This isn't quite as bad as it sounds. If you look at Wikipedia's article on the superior Adobe RGB space (common in the printing industry), there is this convenient plot of the gamuts of sRGB, Adobe RGB and the full xyY space. It makes sRGB look rather pathetic and even Adobe RGB is not great. However, the article points out that the benefits of Adobe RGB are not as large as it seems because of the distortion inherent to the xyY graph. If you download the famous Hoffmann CIELAB document[45], you will find an Lab comparison of the two spaces, viewed at different levels of L. The advantages of Adobe RGB are not so overwhelming, though those who do real graphics work will require an Adobe RGB (or, rather similarly, the NTSC RGB). Only super-specialists require a monitor to a higher gamut and a reproduction chain that can preserve that gamut.

The other point is that increasingly our idea of what things look like are dictated by what we see on screens which default to sRGB and these restricted, artificial colours seem to us to be more natural than real colours. If you do experiments where you show images to non-experts (or if you set up TVs in a showroom), our preferences are overwhelmingly for unnatural, saturated, sRGB colours over "proper" reproductions of what our eyes see.

If you are a professional print buyer, your job is to obsess about subtle colour issues and to push your printer to higher fidelity and higher resolution at, of course, lower cost. In terms of impact on the end users, this is largely a waste of time as most of us could not tell the difference between a high-spec job and an adequate job and in any case we glance at the image for a fraction of a second (if we notice it at all) rather than examine it with a spectrophotometer and magnifying glass.

There *are* cases where precise matching is important. For example, inside a car there are many different types of surfaces and in an expensive car those surfaces will blend together, causing any slight mismatch to be (potentially)

---

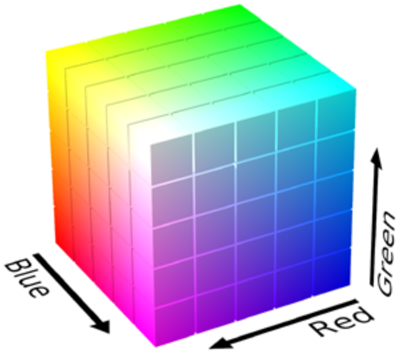
[45] Gernot Hoffmann. *CIELab Color Space*. http://docs-hoffmann.de/cielab03022003.pdf, 2003

visible. In a cheaper car, the colour matching spec can be lower because the different materials are deliberately separated by a different colour, so slight inconsistencies are not visible.

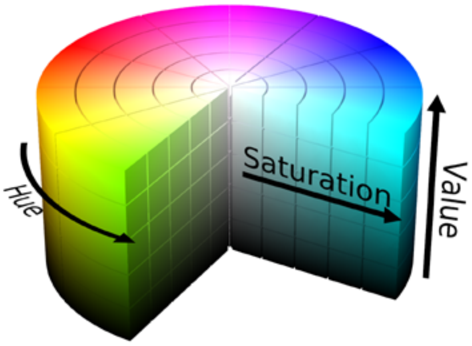
As the specialist differences between various xRGB formats are not important for our purposes (though it is important that you are clear on the specific one used in your chain of suppliers/customers), I will focus on sRGB, then drop the "s" for further discussions.

If you wanted to go from XYZ you would apply a matrix transformation which takes into account the D65 white point standard. And here is the important point; you would then throw away any values less than 0 or greater than 1 as these are not meaningful when you translate sRGB from its 0-1 range into 0-255 integer bytes. All those colours outside the sRGB triangle in the xyY chart are those values outside the 0-1 range. Having our 0-1 values we then artificially boost the lower values by applying a gamma function.

In reality, our starting point is the sRGB (or whatever equivalent you choose to use in Photoshop if you have a higher-quality monitor) and we fiddle on the screen till the image looks right. The assumption is that the screen has been calibrated to our standard so what we see is what we really want, or, rather, what we think we want, given that our screens cannot reproduce all the world's real colours.



The practical trouble with RGB is that it is not at all obvious how to specify a colour that you can visualise in your mind. All the encodable colours fit inside a cube[46] and it is hard to know how to navigate through the cubic space - either intellectually or with a tool in your graphics program. And so we create new colour spaces that seem to better capture how colours should seem.



Two examples are HSL and HSV, Hue, Saturation and Lightness or Value. By definition, these can only contain whatever colours are in your RGB space and the reason that most of us remain confused about the difference between Saturation and Lightness or Value is that these extra dimensions are designed to avoid the embarrassing fact that a pure HS chooser (which sounds much more intuitive) would create non-available colours. There is no colour-scientific basis to either Lightness or Value, they are simply convenient values to allow the user to navigate through a reconfigured RGB

---

[46] The RGB and HSV images are from Wikipedia and attributed to SharkD

colour space. On balance the world has chosen HSV as the least bad way to input colours. It quickly gets you into the right basic colour (Hue) with, we hope, the right sort of intensity (Saturation) and then its general brightness/darkness can be controlled with Value.

So, after a lot of work we have gone from a pure way of describing colours as seen by our eyes, found a pure way to measure those colours and differences between them, then rapidly lost most links to pure colours by choosing to work in one of the RGB variants with, we hope, a colour-calibrated screen, finally choosing colours via one of the least bad options such as HSV.

At this point our colour problems have hardly started. Now we have to print them.

9.1.4 Printing the colours

Even if our monitors are "only" sRGB quality (many are "70% sRGB") we are gazing at a highly controlled surface which emits discrete dots of relatively pure light, with carefully controlled back lighting and an excellent difference between the black and the D65 white. They will have been produced on a production line costing $billions, where the aim is to make the same product as often as possible.

Now we have to get a wide range of images from our highly-controlled environment onto some random piece of substrate, at the highest-possible speed and lowest-possible cost, using dots of finely-ground chemicals that have to not only produce the colour but be safe, low-cost, stable, well-adhered etc. And we do it via methods that allow blobs of ink to be pulled apart, squashed into unintended shapes and jammed on top of each other. And the colour has to come from whatever light is not absorbed by the colourants and manages to get reflected back to our eyes rather than scattered all around us.

With so much stacked against us, it is surprising that we get anything at all reasonable. In fact, we end up with a gamut that (on a good day) is about the same size as sRGB, but with a different shape that, happily, sits inside Adobe RGB, allowing those with good monitors (unlike those with sRGB monitors) to simulate what the print will look like.

Ignoring, for the moment, the problem of printing blacks, we could, in principle, get a perfect Adobe RGB gamut if we had a perfect C, M and Y, which, in their respective pairs, give us what we require.

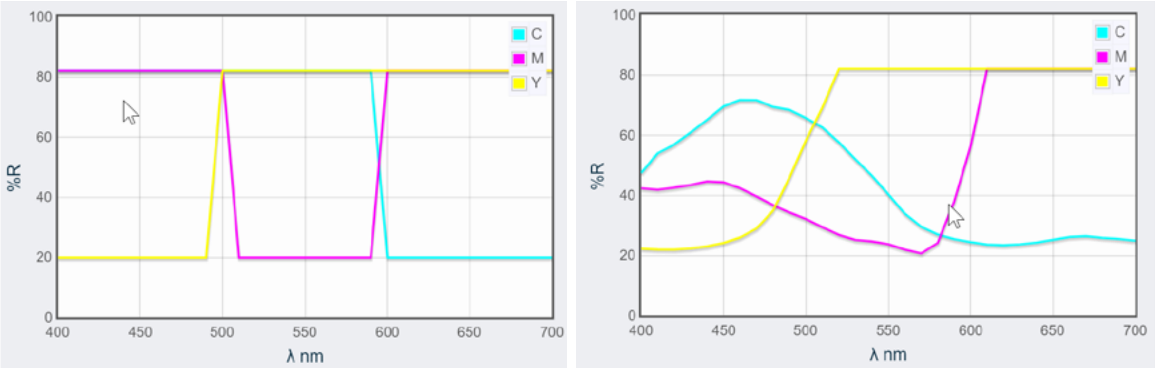


Figure 9-3 The reflectance spectra of ideal CMY inks are on the left; typical spectra of real inks are on the right. Y is OK, but there is too little blue reflectance for the M and C is poor in both the blue and green areas.

However, although a typical Y pigment does a great job of reflecting R & G wavelengths, as the figure shows (images from the Kubelka-Munk app) M is OK in R but poor in B and C is OK in B but not great into G. So we are off to a bad start, even without the other issues of real printed dots.

The K is added because inks based on carbon black are cheaper than other pigmented inks so it makes financial sense to replace as much C+M+Y=Gray-to-Black with the equivalent K. And because dot-on-dot printing raises multiple complications (dot gain, trapping, colour edge effects), it makes better colour sense to print as few dots as possible. Finally, there is far less solvent/water to dry or UV ink to cure. Although K offers such advantages, it adds another layer of complication to an already messy story.

We also have a set of trade-offs in our inks. For some purposes, an ink with great hiding power, or opacity is required. This can be achieved by providing lots of scatter. This, as we know from Kubelka-Munk, reduces our gamut because scattering essentially adds "whiteness" to our print. Scattering is optimal for pigments with sizes around the wavelength of light. The theory is rather complex because it depends on relative refractive indices and the optimum radius is ~450nm, i.e. a rather large 900nm diameter; you can have a look at a Mie scattering app for more details: https://www.stevenabbott.co.uk/practical-coatings/mie.php. The reason we use TiO_2 as a white pigment is that its refractive index is high, giving maximum opacity for minimum pigment.

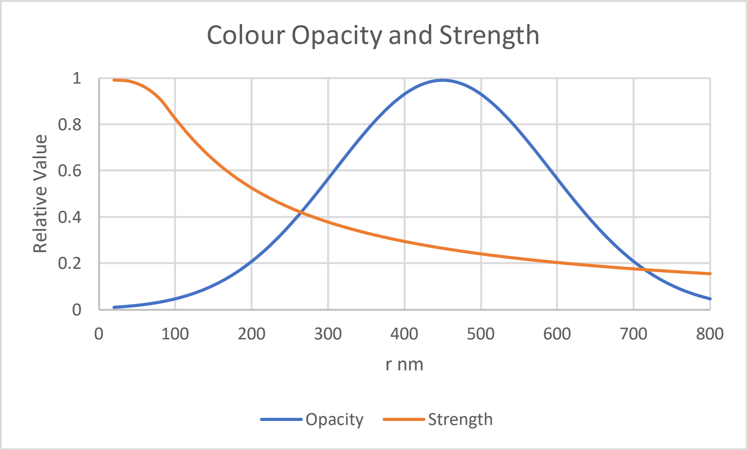


Figure 9-4 How colour opacity and strength might vary with particle size. The actual curves depend on the complex refractive index (refractive index and absorption) of the pigment and the refractive index of the binder.

For other purposes we want strong colours. For a given % pigment, strength is increased by going to smaller particle sizes which don't scatter and where a larger fraction of the molecules in the pigment get to interact with the light. By this I mean that if light is absorbed totally by 20nm of pigment then a 200nm diameter sphere would have inside it a 160nm diameter sphere of molecules that never saw a photon, meaning that $(160/200)^3$ is useless - that's ~50% wasted pigment.

An interesting paper[47] brings this trade-off alive in an unusual manner. The authors showed that during a printing run, the particle size of the pigments decreased, giving less opacity and more colour strength. The ΔE values between the start and end of the run showed significant differences that matched what would be expected from the significant decreases in particle sizes. Yet another complication for the printer!

Three things come to our rescue from all this complexity.

1. We have been using CMYK for over 100 years so we are used to its colours being "normal".
2. We can extend the gamut of CMYK with extra colours such as OG or OGV or use special pigments or especially dense pigments to produce various expanded range CMYK gamuts.
3. Although there is no simple Adobe RGB→CMYK conversion because of the imperfections in C and M, by going from RGB into XYZ then integrating against individual absorption curves of the C,M and Y, it is possible to do rational conversions for a given ink set on a given printing system.

---

[47] Andrea Frimova, Alexandra Pekarovicova, Paul D. Fleming and Jan Pekarovic, Ink Stability During Printing, TAGA Journal, 2, 122-131, 2006

In fact it is even simpler than that. Either you don't have colour correction software and what you get will be rather chancy (but good enough for many purposes, though you will not be able to compete for high-quality jobs) or you have a proper chain of colour correction (for example, using ICC profiles along the whole chain of design, proofing and printing) which handles everything; not just the colour correction but dealing with Kubelka-Munk effects discussed previously and dot gain and the other issues discussed in the next section.

ICC (International Color Consortium) profiles are a way of implementing all the transforms we have been discussing, and doing it in a standardised manner. If your computer monitor is fed a set of pure colours which are carefully measured, then it is possible to work out a profile that takes any given colour and manipulates it so that what appears on the screen is a close as possible to the desired value. Similarly, you print a set of pure colours, measure them and create a profile that pre-corrects the colours before printing them. As long as everyone in the chain has the right ICC profile, then the colours specified by the designer have a good chance of being observed in the print, once they've been passed through the various ICC profiles along the chain. Any wrong assumption or miscommunication down the chain will probably make matters worse. But the point of ICC is that everyone uses it and after a few costly errors, works out how to use it properly.

I *could* have described the whole process in mathematical detail and I *could* have written some apps. But what is the point? The correction process depends on your specific press, specific inks, specific RGB definition and whether you are using normal CMYK or have an extended gamut. The principles behind the correction process have been discussed so you get a general idea of what is going on and why it is so complex. The details scarcely matter; you have paid for your spectrophotometer, used it to calibrate your raw prints (a set of uncorrected CMYK colours that cover a large part of colour space), and providing you receive no warnings about parts of the original image being "out of gamut", you will print with the closest fidelity your system can provide.

This is an important point. Colour correction is powerful and impressive, but it is not magic. If, for example, your dot gain is so large that there is no difference between an 85% and a 100% colour then you have 15% fewer colour step possibilities so many subtle colours cannot be printed and cannot be corrected for.

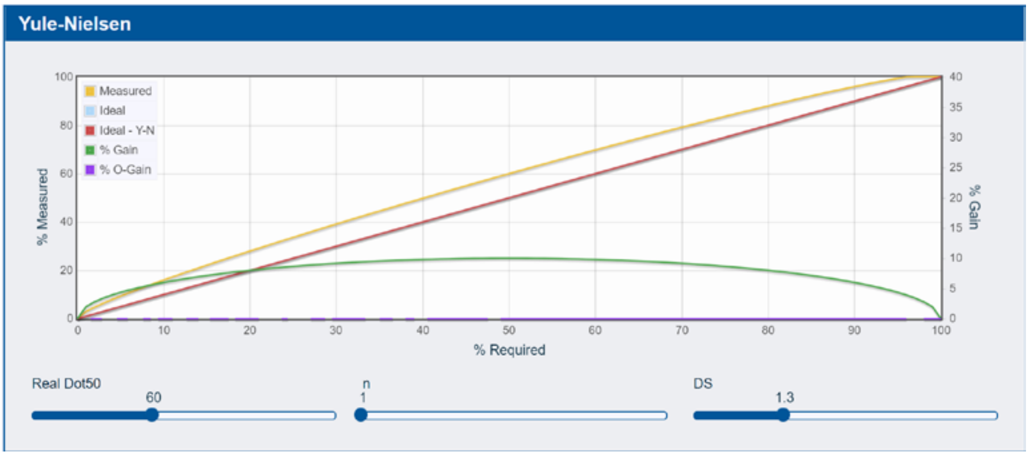
There is one big problem with the "leave it to colour correction" approach. The most reliable way is to print and measure a large number of colour patches which the system can then digest and convert into a high-quality ICC profile. This is fine if you print one type of ink and one type of job. If you have a larger variety, the alternative approach is to print far fewer samples and use physics to create the required interpolations for a high-quality correction process.

So rather than spend time discussing something over which you have almost no influence (because you will have bought in a colour correction system), let us spend time discussing those things that you can influence and which, when improved, will be fed back through your colour correction process (with fewer test patches if necessary) to give you a wider gamut and a better chance of capturing more demanding jobs from picky print buyers.

9.1.5 Dots

Note: as far as possible the discussions here are neutral about whether the dots are classic halftones (amplitude modulation, AM), stochastic (frequency modulation, FM) or mixed (XM). The term "screen ruling" (i.e. the LPC or LPI of the print) is used neutrally even though it is historically associated with AM. The term "densitometer" generally means "spectrodensitometer"; an old-fashioned densitometer which uses a few colour filters is not appropriate for proper colour work.

Our most obvious problem is dot gain. Yes, in the highlights we might experience dot loss but the causes/cures are not so relevant to this chapter. Dot gain is far more significant. We measure the densities of our supposedly X% dots where our densitometer tells us the density D of our dots compared to DS for our solids (normalised for our substrate white) and we calculate via the Murray-Davis equation (discussed more below) $\%Dot=(1-10^{-D})/(1-10^{-DS})$. We find that it is (in general) rather larger than X%. The gain is (%Dot-%X) rather than %(Dot-X)/X which is what I always thought was meant by "percent dot gain". A typical curve would look like this image from the Yule-Nielsen app described below. The settings mean that you see the ideal straight line, the actual measured % dot and the dot gain (on the right-hand axis):



Real Dot50 60

n 1

DS 1.3

Figure 9-5 A classic dot gain curve generated by the GRL equation. No optical gain effects are present because the Yule-Nielsen n=1

The curve uses Viggiano's GRL equation (after the Graphics Research Lab in Rochester Institute of Technology) based on a 10% dot gain at 50% solids. The GRL is one of the many attempts to describe dot gain with the minimum number of parameters - in this case the whole dot gain curve can be derived from a single reading of the X=50% print. The (limited) value of such equations is the main point of this section, but it is really useful for creating an insightful app. If the required dot area is A and the dot gain at A=50 is Δ_{50} then the gain at any value of A, Δ_A is given by the following, along with A_{max} where the gain gives a 100% print:

$$\Delta_A = 2\Delta_{50}\sqrt{A(1-A)} \text{ and } A_{\max} = \frac{1}{1+4\Delta_{50}^{\ 2}}$$

Equ. 9-5

Real Dot50 70

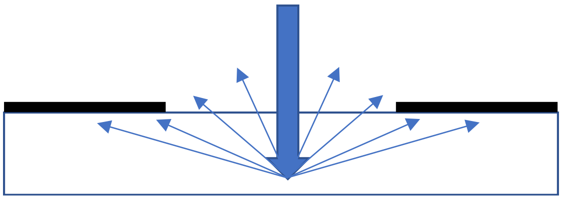
n 1

DS 1.3

%Tone Reduction 13%

We need to make a small diversion into the implications of A_{max}. With a 20% dot gain at 50% then A_{max} is 86%. That means that after correction, our 0-100% printed range is achieved with a 0-86% dot range. So what? If we can print any sized dot with perfect accuracy, this makes little difference. Real imagesetters have limits. Suppose, for simplicity, that we can meaningfully print dots at 5% intervals. If we could print 0-100% dots then we can distinguish 20 tonal levels. Now we are limited to 86%, we have 17 levels, a significant reduction in our ability to reproduce subtle tones or fine vignettes. The actual % reduction is shown in the app for your convenience.

There is a significant complication with dot gain. Those who print with refined offset plates which naturally give rather small dot gain (mostly in the plate-to-blanket step) will be dismayed to find that they have significant measured dot gain and might be confused to find that if they measure the printed dots, they are not as large as the dot gain suggests. This extra gain is less obvious for flexo and gravure where the real dot gains are larger because of the much lower viscosity ink, and for flexo the generally lower LPC do not show this effect so strongly.



The extra gain arises on white substrates because light entering between the dots does not get fully reflected; some of it is scattered and trapped beneath the dots. This "optical gain" effect does not exist when printing onto clear polymer films and is worse on uncoated paper (which scatters more) than on glossier coated paper. The variability in optical gain is a real problem for the printer because even with superb control over dot gain on the press, the differences in optical dot gain on different substrates will require a re-calibration of the whole colour correction chain.

The theory behind optical gain is called Yule-Nielsen (Y-N) and they were the first to stress that the theory isn't very good because some of their measurements required "impossible" values, as we shall see. There have been many attempts to improve on Y-N but they all hit the same problem, which applies to Y-N too: unless the technique is good enough to give an accurate prediction of the whole tonal curve from one or two measurements, you might as well measure the whole curve and correct it from the measurements. Back in Yule's times, anything which reduced the measurement and computational burden was welcome. These days it is hardly worth the bother.

It is still worth having an app to explore Y-N as it is a key part of our printing science heritage, it is something you come across regularly, it links to Murray-Davis and even allows us to discuss YNSN, Yule-Nielsen Spectral Neugebrauer theory.

The Yule-Nielsen equation is a more general form of Murray-Davis, linking the densitometer reading, D (where DS is that of the pure substrate), to what the eye thinks is the printed area of the dot, A. We can either get A from D or D from A:

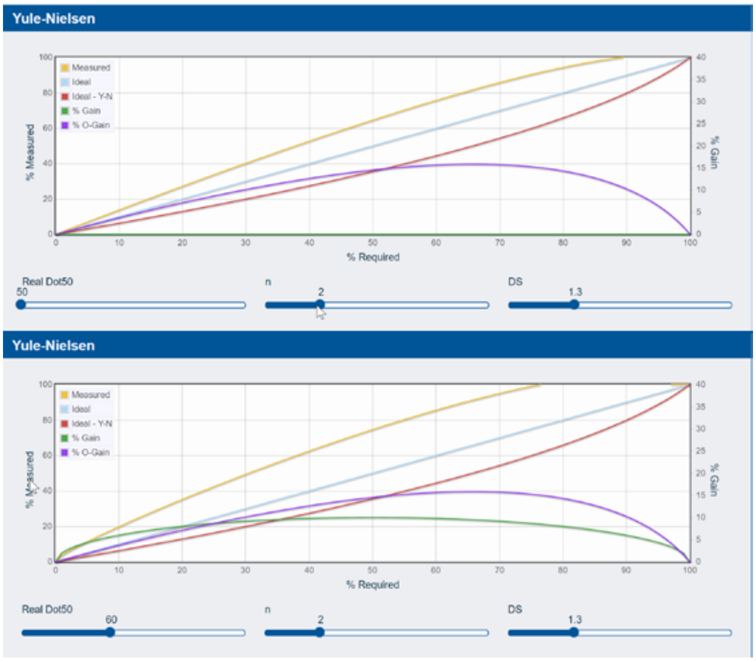
Equ. 9-6

$$A = \frac{1-10^{-\frac{D}{n}}}{1-10^{-\frac{DS}{n}}} \text{ or } D = -n\log\left\{1-A\left[1-10^{-\frac{DS}{n}}\right]\right\}$$

Murray-Davis (or Y-N on a non-scattering surface) has n=1, and Yule & Nielsen showed that on a highly scattering substrate, n should be 2. So all prints should have n to be somewhere between 1 and 2. The problem is that many prints require "impossible" values higher than 2. The most compelling explanation I can find for higher values is that the scattering of the ink (a sort of Kubelka-Munk effect) and scattering from a ragged dot edge can both send n towards high values. This is an important point for discussion after exploring the app.

A natural question to ask is how severe the Y-N effects are in real prints. The answer would come from accurate comparisons of densitometer areas with areas measured via a digital microscope. The problem is that although it is relatively easy to get an accurate digital measure of dot sizes of well-behaved inks on smooth substrates, these are the cases where n~1. As soon as you have rougher surfaces and more challenging inks, it is rather hard (I've tried this myself) to find an unambiguous dot versus no-dot image threshold. As a result, the number of theoretical Y-N papers far exceeds the number of papers with definitive values for the effect. The original paper (Yule was from Kodak) used photographic dots/stripes laid onto substrates so they accurately knew the size of the optical effect (after a few necessary adjustments).

The approximate answer is that it goes from 1 in low screen rulings to 2 in high rulings if the dots are "nice" and to 3 or more with bad inks badly printed onto bad substrates. That still doesn't properly answer the question - for that I needed the app.



App 9-1 https://www.stevenabbott.co.uk/practical-coatings/Yule-Nielsen.php

We have already seen the conventional dot gain curve where the 50% dot was actually 60%. The top curve in this image shows what would happen if the dots were perfect (so 50=50). It is fairly similar to the case of a real dot gain of 10% at 50%, though it is worse in the darker areas. The optical dot gain is shown. There is an extra curve labelled Ideal - Y-N. This is what the real printed dots would have to be if they were going to produce an ideal curve. So to produce a 50% dot you would have to print a 35% dot because the optical gain is 15%. Is 15% large or small? It depends on n. The value here is 2, which is a "perfectly scattering" substrate combined with perfect dots that add no extra Y-N effects. Maybe on a coated paper n=1.5 and you need a 40% printed dot to get a measured 50% value. But then again, that same coated paper might give a higher n for a higher-screen print.

Both the standard print with n=1 and the ideal print with n=2 are unrealistic. A more realistic curve is the second one shown in the app where the real dots have a 10% gain at 60% and the combined effects mean that an 85% dot measures as 100% - i.e. you have lost plenty of printing range because of these combined effects.

9.1.6 Dot on dot: YNSN

The dot gain curves for a single colour printed in a standard test gradation are a good starting point for understanding what is going on. It is useful background information to understanding the next level of analysis involving dot on dot; not because we are likely to use it, but because it gives us an idea of why things go wrong.

The starting point is Neugebauer's idea that CMYK prints can be analysed via a set of pure colours: CMYK plus W and RGB. The W is the pure substrate and the RGB arise from MY, CY, CM dot-on-dot combinations. Because we know the XYZ values of all these colours, we can know the colour of any halftone by summing the XYZ values of the individual colours scaled by the halftone area.

The flaws in this simplified version are obvious and there are two extensions which make it more powerful. The first is to use the individual spectra (rather than single XYZ) values for each colour. With 16 combinations of CMY&K the reflected light at any wavelength is the sum of those individual combinations The second is to correct the ideal version via Y-N. This double combination is YNSN, Yule-Nielsen Spectral Neugebauer.

A set of tests[48] of pure N, YNN then YNSN show that the mean errors in prediction (ΔE_{00}) after measuring the 16 standard combinations are ~7 for N (i.e. useless), 3 for YNN (i.e. not bad on average but some large individual errors) and 2-3 for YNSN, (with fewer large errors).

For those who are really keen, the KM-YNSN method adds Kubelka-Munk terms, but the benefits are modest compared to using the straight YNSN method.

This means that those brave enough to trust to theory can do all their colour prediction from measuring 16 patches rather than the 300-1000 (or even 1485) patches in many of the standard targets used within the ICC system. This really depends whether you have a manual spectrophotometer or an automated scanning one.

Such a decision would be nice and rational if you believed that your printing process was entirely under control. The point of this entire section is something very different.

9.1.7 Your dot quality

In the previous section I quoted typical variations in predicted versus actual ΔE values. What is of even more interest to me than the conclusion that YNSN is OK, is the variation in ΔE values between different printing techniques. For offset the mean is 1.75 (i.e. below visibly discernible) with the maximum error being 5.85. For thermal sublimation (or, rather, die diffusion thermal transfer), another high-quality technique, the values are similar. For flexo the mean is 3.2 but the maximum is 19, and for inkjet the figures are 4.4 and 14.3. In both these cases I would be certain that the dot quality is poorer and so any scientific method will be limited by those uncertainties.

---

[48] Kiran Deshpande, *N-colour separation methods for accurate reproduction of spot colours*, PhD Thesis, U of Arts, London, 2015.

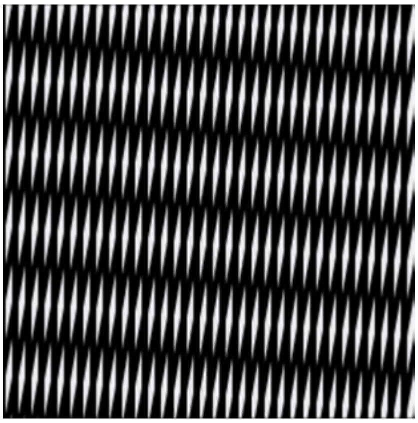
Although it sounds trivial to say that dot quality affects the ability to correct for dot gain, the insights gained from looking at the basic theories show that relatively small changes in control of dot quality, or in substrate or in screen rulings (I have also found papers where changes of the dot angles and dot placements can influence the effective dot gain) get magnified because the n value in Yule-Nielsen depends on all those values and the values in the spectral Neugebauer will also change significantly if the inks dry/flow differently to give different scattering and absorption profiles.

There is almost nothing you can do in terms of improving the colour science aspects of your printing. If you have a reasonable densitometer and make a reasonable choice in terms of calibrating patches and linking through via ICC profiles there is nothing more to be done.

However, what you do in terms of choice of inks and substrates, press conditions, locking off "tweaking" options on the press will have a large effect. If your choices give controlled dots with good edges and the minimum possible dot gain, and if you can persuade print buyers against higher screen rulings because they come at a scientific price in colour gamut and reliability, then you will have a deep impact on your colour printing capabilities.

It is always nice to know things like XYZ and Lab; but it is far more important to know how to keep your dots looking nice, day after day after day.

10 Moiré



There is a common myth about the funny moiré patterns we see when regular patterns are superimposed. An example of moiré is shown here, created by superimposing two grating structures angled at 10°. Like all the examples in this chapter, the image comes from the moiré app discussed shortly. The myth states that they are named after a Professor Moiré who worked out the theory. The reality is more interesting. The word is derived from the Arabic for the cloth woven from the fine wool of an Angora goat.

Here's what I quoted on the subject some years ago: “MOHAIR, the hair of a variety of goat originally inhabiting the regions of Asiatic Turkey of which Angora is the centre, whence the animal is known as the Angora goat. The Arabic mukhayyar or muhayyar مُخَيَّر, from which the word came into English probably through the Ital. moccacaro or Fr. inocayart, meant literally, ‘choice’ or ‘select’, and was applied to cloth made of goats’ hair. In the 17th century the word, which before appears in such forms as mocayare or mokaire, became corrupted to the English “mohair” from which the French adapted moiré, a watered silk fabric.”

Most of the time, moiré is an interesting optical illusion. And most of the time, a printer doing routine work should never see moiré, either because they are using stochastic (FM) screens or because their CMYK plates are at the standard angles that give the pleasing rosette pattern (itself a moiré, but an inoffensive one for reasons discussed later). However, flexo printers can get nasty moiré effects from the anilox roller, screen printers can get bad moirés via interactions with the mesh, and any systematic error in a plate (such as a stretching) can cause a pleasing rosette to turn into a moiré horror. Note that FM screens can give ugly patterns; one thing we can be certain of is that these are not moiré which is scientifically defined as the superposition of regular features.

The mathematics of moiré were made definitive by the wonderful book by Isaac Amidror[49]. This makes it possible to provide an app that helps to explain the basics of moiré (of some interest but mostly irrelevant to modern printing) then allows you to explore the more subtle effects that can mess up a printer's life. There is, for example, a magic number of 4 which helps decide whether a moiré will be worrisome or not.

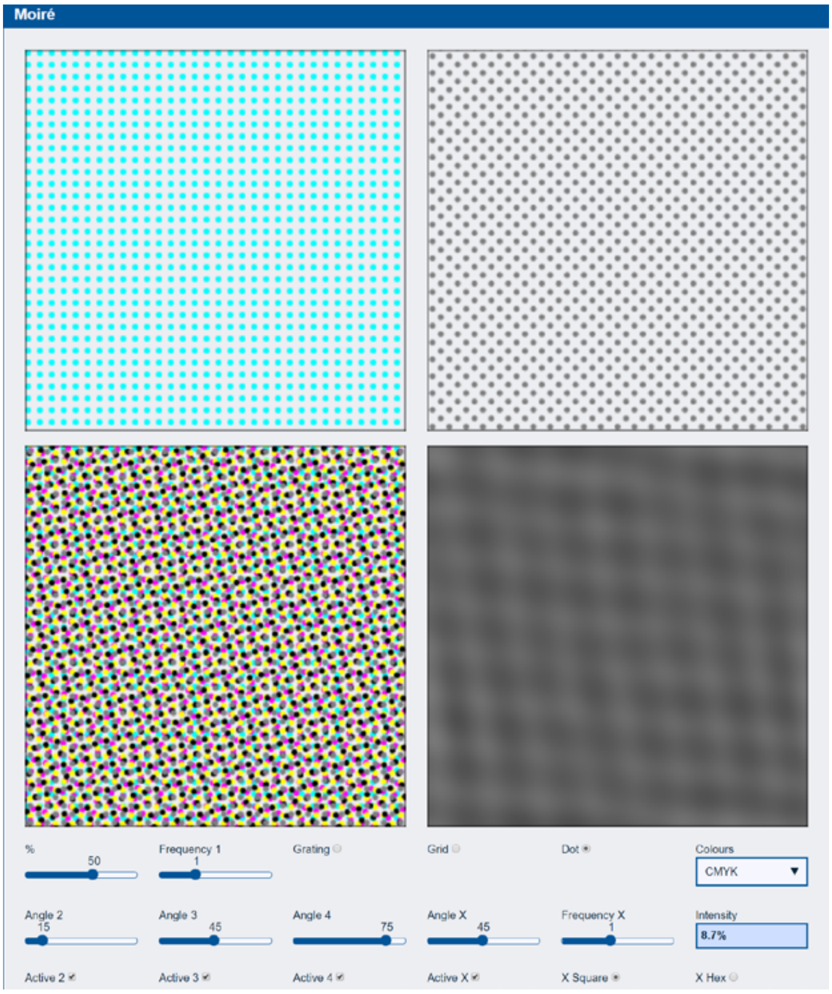
Just as important is to be able to say with confidence that a "moiré effect" on a print is *not* caused by moiré. In the world of screen printing I often had to use moiré science to prove that some ugly pattern on the print was caused by

---

[49] Isaac Amidror, The Theory of the Moiré Phenomenon, Kluwer Academic, 2000

something different from bad screen angles. By eliminating one possible root cause, it was possible to identify the true cause.

Because the app is central to this chapter, let us look at it in its full complexity before taking it one step at a time:



% 50
Frequency 1 1
Grating ○
Grid ○
Dot ●
Colours CMYK
Angle 2 15
Angle 3 45
Angle 4 75
Angle X 45
Frequency X 1
Intensity 8.7%
Active 2 ☑
Active 3 ☑
Active 4 ☑
Active X ☑
X Square ●
X Hex ○

App 10-1 https://www.stevenabbott.co.uk/practical-coatings/moire.php

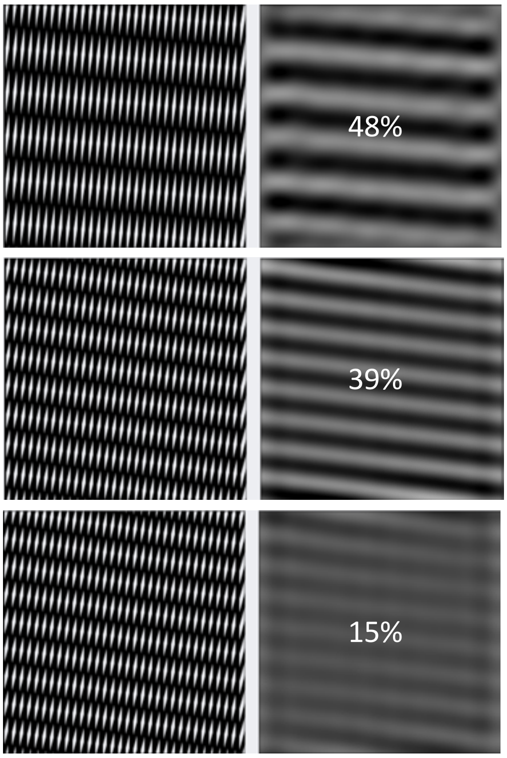
Note 1. The screen shots will (confusingly) contain their own moiré via the series of pixellated steps that bring them to your eyes. Please run the app to see the true image.

What we can see in the bottom-left is that we have CMYK 50% dots at angles of 0 (the first image is always at 0°, top-left), 15, 45, 75, with the frequency of the first pattern (1=C) set to 1, which means that it is the same as all the others. The bottom-right image shows that we have a moiré from the classical rosette, but the intensity is <10% which means it's not really objectionable. In the top-right we have an X pattern which (to give us an unobjectionable image) is set to be identical to the Y. This X pattern allows us to simulate what might happen if we have defects from a screen mesh (a Square pattern) or a flexo anilox roller (a Hexagonal pattern), at a Frequency X relative to the standard frequency.

Because everything is adjustable with sliders and we can turn angles on and off and can choose any ordering of CMYK or just have KKKK, we have a powerful, general-purpose moiré simulator.

Note 2: The calculations and graphics require quite a bit of processing power. If you have an older laptop or are using a mobile device, play with Gratings and a few angles (using up much less processing power) before attempting many angles and (especially) dots.

10.1 The Rule of 4



Reverting to the very basics with two gratings, KKK selected and only Angle 2 selected, we can quickly confirm the rather powerful Rule of 4. Each of the gratings has 32 lines. The three moirés are produced at angles of 10, 15 and 16° and you can see roughly (and this is known exactly) that the moiré lines are at angles of 5, 7.5 and 8°. You can also count the number of moiré lines which are ~5, 8 and 9 respectively. Divide the two frequencies and you get 6.4, 4 and 3.6. The Rules of 4 states that a moiré is visible when the ratio is >4, with 4 being borderline. Given that we have to use a 15° angle in printing you can see that the Rule of 4 is rather important. If we used less that 15° the moiré would be much more visible. Unfortunately we cannot use 16° because we cannot fit in 4 colours. That is why we sometimes have to change the 15° colour; if this is Y, which is not so visible, and if there is not too much Y in the image, then we are safe. If Y is key to features of the image and, say, M, is less used, then M can be set to 15°.

Along the way we have learned that moiré is mathematical; we know that if we see a strong moiré at angle X then it must be formed by features separated by angle 2X. Or, if we have features separated by angle Y we will see moiré at Y/2 and, if we have a grid or dots, 90+Y/2. The ratio behind the Rule of 4 is calculated from the angle, θ, between the lines:

$$Ratio = \frac{1}{2\sin\theta/2}$$

Equ. 10-1

These rules aren't just science for science sake. If you see a moiré then you can easily measure its angle and estimate the Ratio. This immediately tells you which elements of the same frequency at which angles are interfering. Just as importantly, if you do not have any elements at these angles then this is either a moiré with an element at a *different* frequency (such as a screen mesh or an anilox roller) or *is not a moiré at all.* This second point is so important that it requires a section to itself.

10.2 The importance of not being a moiré

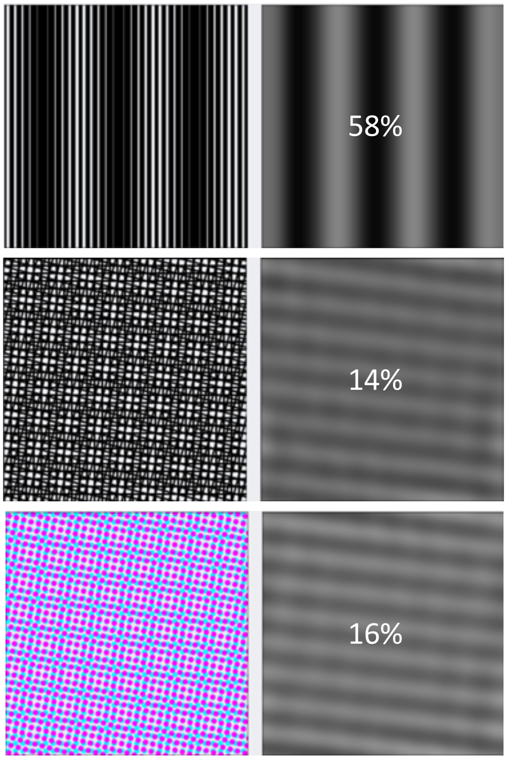
I first learned of the importance of not being a moiré when working on moiré in screen printing. At that time, the words " moiré" and "magic" were almost interchangeable. As few printers understood moiré, any time they saw some fairly regular pattern in their prints they put it down to this mysterious moiré force that loved to ruin their prints.

Rather quickly it became clear that a simple calculation using basic moiré maths made it possible to rapidly rule out moiré as being the cause of many of their imperfections. Very often, for example, it was due to a bad dot-on-dot printing setup that gave lots of "skipping" (incomplete dots) in the print. It is rather hard to tell a printer that they are doing something wrong; they far prefer to blame things on mysterious forces such as moiré. Being confident that the laws of physics ruled out moiré made it much easier to persuade the printer that the root cause lay elsewhere. Looking through a magnifying glass at a print with moiré may not be very different from looking at another type of defect. But looking at that defect when you are certain that it is *not* moiré makes it much easier to identify the true cause.

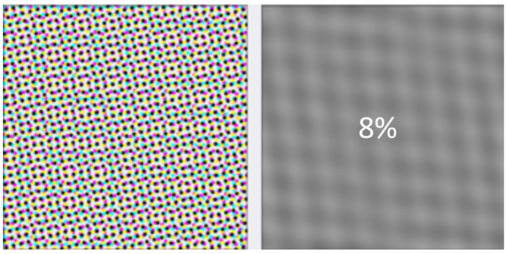
Negative proofs are very important in problem solving. My all-time-favourite (nothing to do with moiré) is when my company was accused of shipping bad product with a distinctive mark all over it. When our technical specialist visited the customer he was able to say "While it is perfectly possible that we could ship product with marks in it, it is impossible that our marks should all coincide with the top right corner of your printed sheet." By showing a negative, the customer could then focus on the fact that the most likely cause of the defect appearing in the top right of every sheet was something on their press - which was duly found.

Note 3. The easy disproof arises only when the frequencies are the same and the angles are known. We will return to the trickier case of moire caused by different frequencies once we have explored more of the app.

10.3 More complex moirés



Now let us look in quick succession at some more variants of 2-layer moiré. In turn they are the same angle but different frequency (setting Frequency 1 to 1.1), giving a very strong visibility, then a Grid which, because of the extra high frequency components, dilutes the intensity compared to the Grating equivalent, then to the classic 15% dot pattern (having selected the CMYK colour option) which seems so strong when viewed close up but is surprisingly acceptable when viewed from a normal distance. In fact the moiré image is calculated via a Fast Fourier Transform (FFT) filtering of high frequency components based on a typical viewing distance, and that filter kicks in at a factor of 4.



When we get to the full CMYK rosette, with equal portions of each colour, the calculations are clear - the moiré as a whole is not visible. Given that 15° moiré is borderline, whether you have to fiddle with a different print order depends on whether there are large patches of 15° areas, something the app cannot comment on.

10.4 The X factor

To understand why there are some unsettling moirés that come and go in screen printing, we have to turn on the X option, with a square grid at a Frequency X and angle different from that of the print.

Trying to include a screen-shot from the app wasn't too successful. So set up Frequency X at 2.1 and arrange a slight misalignment (something that might easily happen on a press) of 1.5° between the mesh and the image. There is not a strong moiré, but there is clearly something unsettled about the print. And this is what is found in screen printing. The mesh interferes with dots at (in my judgement) half the mesh frequency (i.e. "under" mesh crossings are different from "over" mesh crossings), so there is a slight bias (shown by "bites" out of the dots) which gives an unsettled look. The effect is strongly angle and frequency dependent. A separate program (freely available if you email me for a copy) does precise calculations of these effects using deep moiré theory.

Similarly, if I select a 60° hexagonal grid at a frequency of 1.9 I can get a quite ugly moiré that shows the possibilities inherent in a misjudged combination of anilox roller and image angle. At the time of writing, I have no feedback either way about real anilox moiré conditions, so the app should be used to illustrate the problem rather than be definitive.

10.5 Summary

This chapter is short because writing about moiré is far less important than exploring it live via the app. If you never see moiré in your prints then you can amuse yourself with some simple examples of this fascinating effect. If you have effects which you think are moiré, then you can use the app to help work out if the effect really is moiré, in which case it can be fixed by adjusting angles and frequencies, or is not moiré, in which case the solution lies elsewhere.

As with all my apps, if you would like improvements or changes, feel free to let me know.