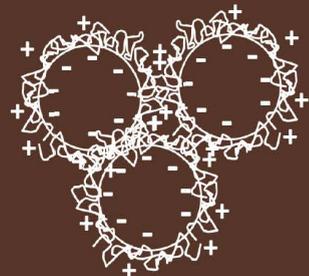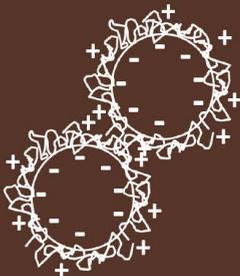# Introduction to Colloid Science

Applications to sediment characterization

Claire Chassagne

Introduction to Colloid Science

Applications to sediment characterization

Claire Chassagne

Contents

wie 't kleine
niet éert is 't
grote niet
wéerd

---

*"Who do not
honour the small things
is not worth the big ones"*

[Dutch saying]

Foreword

This book is meant as an introduction to the field of colloid science, i.e. the study of the behaviour of micrometric particles in a fluid (or a gas). The book was written with a special emphasis on sediment particles. Sediment particles are complex colloidal particles due to their composition, shape and interaction with their environment. Characterization of the colloidal fraction of sediment is done by recording, among others, the particles' size, shape and electric surface charge and evaluating their interactions. These properties are important for civil engineering applications, even though the size range of these particles and their interactions is microscopic. Large-scale sediment transport models for example require as input the settling velocity of individual particles. In concentrated areas, this velocity becomes a function of the particles' concentration and particle-particle interactions lead to the creation of larger particles, called flocs. These flocs can settle and, when reaching the bed, consolidate in time. All these aspects, and related models, are treated in the present book.

The book was written for students having no special beforehand knowledge in colloid science, and a limited knowledge in physics and chemistry, the two main disciplines relevant to this branch of science. The mathematics are also kept minimal. They are given when the derivations present no particular problem. The more lengthy derivations, often very elegant (but sometimes rather tedious) can be found in the references given as footnotes in the different chapters.

Key words are highlighted in the text, which will enable the reader to find further information by searching for them in a web browser.

I take the opportunity to thank the colleagues that reviewed this book, and helped with their comments to improve it. To the present reader, I would like to say: do not hesitate to report any mistake or unclarity to me.

Claire Chassagne
Delft, spring 2019

Chapter 1

Introduction, general definitions and some properties of clayey material

Introduction

Mud and clayey systems have from origin been studied in the field of soil science[1], i.e. the branch of science concerned with the formation, classification and fertility properties of soils. In soil science, the classification of soils is of prime importance, as the properties of a soil are used by farmers to decide which types of crops, livestock and soil management are best for their piece of land. The information obtained by a soil survey can also be used by an architect or a builder to determine whether a given soil is suitable for a specific type of construction.

In agriculture, soils are primary viewed for their chemical properties, i.e. their ability to bring nutrients to plants. At the beginning of the 19th century chemical studies were extensively conducted which confirmed for instance the role of phosphate, nitrogen and potassium in plant growth, and commercial fertilizers were developed. The composition of a soil is also determinant for the type of plants that are growing on it. The first report that clearly related the soil properties to the different climates, vegetation and parent rocks from which the soils were formed was written by Dokuchaiev in 1883[2]. Many of the chemical changes occurring in a soil are related to the

1350 BC, Egypt : thanks to the fertile soil of the river Nile, Egypt became a powerful and long-lasting civilisation

action of bacteria. At the end of 19th century it was for instance discovered that the bacteria which live in nodules on the roots of legume plants absorb nitrogen from the air and convert it to a form which plants can utilize. The differences in soil characteristics are connected not only to the (bio)chemistry but also to the mineralogical and physical properties of soils. These properties are important for construction, but also to the farmer as the porosity of a soil governs gas exchange with the atmosphere, water penetration and root growth. The porosity of a soil is strongly connected to the particle size distribution of the grains forming the soil and the soil's state of consolidation. The mineralogical composition of a clay, which can be determined from X-ray crystallography and electron microscopy tests, will tell if a clay soil is prone to weathering, swelling and retain or exchange cations[3]. This last property is pH-dependent. This shows that both chemical and physical properties

---

[1] We will focus on natural soils. Another important research concerns the use of clays in ceramics, cosmetics and drugs which will not be discussed here.

[2] Dokuchaiev, V.V. (1883) Russian chernozem, Monograph, Sankt-Peterburg. See also Greenland, Dennis James, and Michael Hilary Bermingham Hayes. *The chemistry of soil constituents*. Wiley and sons (1997).

[3] a cation is a positively charged ion; an anion is a negatively charged ion.

are of importance to understand the diffusion and retentions of ions within the soil fabric. Similarly, the consolidation, permeability, strength and aging of soils are related to chemo-physical processes: by oxidation (chemical process) the permeability of a soil (physical property) can for instance been altered.

In the context of sediment dynamics, i.e. the study of sediment transport, deposition and erosion, the first studies have been initiated starting from the work done on sand. We emphasize that we are here dealing with water-saturated bodies and not Aeolian transport. The studies we refer to are therefore relevant for fresh and salt water systems such as rivers, lakes, seas and estuaries. Aquatic sand transport, deposition and erosion has been monitored and modelled successfully in the past decades. The models have been tried on clays and clayey systems, but it was soon recognized that these systems were much more complex and the models and theories should be adapted[4].

In this first chapter (**Chapter 1**), clayey soils and suspensions are introduced from their classical definitions, derived primarily from soil science, in terms of particle size, clay mineralogy and physical properties (fluid, solid). We will start by defining *colloidal* sized particles. Colloids (= colloidal suspensions) will be the main topic of interest in the following chapters.

There are several differences between clay and sand particles. One of them is their size, which affects their diffusion behaviour in the water body, as detailed in **Chapter 2**. A major difference between sand and (colloidal) clay particles is that clay particles have the ability to *aggregate*. the reason behind this property is discussed in **Chapter 3**. The clay particles can aggregate between themselves, adsorb (poly)ions, stick to or interact with other particles (some of them are reviewed in **Chapter 4**). In **Chapter 5**, the different types of aggregation (also called flocculation) and the different types of aggregates (flocs) produced are reviewed. The rate of flocculation is modelled in **Chapter 6**.

From the concepts introduced in **Chapters 2 – 6**, it will become clear that the main reason sand models are not adapted to clayey systems is that because of their size, clay particles are extremely dependent on surface forces (interaction with other particles in suspension), whereas the dynamics of sand particles are mainly controlled by a volume force (gravity force). Sand particles can be assumed to have a constant Stokes' settling velocity[5], which can be calculated from their density and volume whereas clayey particles have a time-dependent density and volume. Sediment transport models have as input the concentration of particles of a certain size and their Stokes' settling velocity, it is therefore clear that some adjustment is

---

[4] Leussen, W. van, *Estuarine macroflocs and their role in fine-grained sediment transport*, PhD thesis, RIKZ (1994)
[5] Stokes' settling velocity is defined in Chapter 2

required to the models as aggregation implies that particles change both their size and their settling velocity.

The density and composition of the clay suspensions influence their rheology (**Chapter 7**) and the settling of their constitutive particles (**Chapter 8**). Once settled, mud suspensions will form a soil with a time-dependent permeability (**Chapter 9**). Given a specific relation between permeability and density, the early stage of consolidation can be studied (**Chapter 10**). This so-called early stage of consolidation, when mud is freshly deposited at the bottom of a water body, is important to study in connection to sediment dynamics: the strength of a freshly consolidated mud layer is a key parameter for sediment transport models as it will determine the amount of mud eroded from the bottom at a given water velocity. The link between **Chapters 8 and 10** is discussed. In the last part of **Chapter 10**, we will briefly review the later stage of consolidation, when forces at the contact points between particles become predominant. This stage is extensively studied in soil science and the reader is referred to books in that field of science for further reading.

Solutions, colloids and suspensions

For the remainder of the book, we will be concerned with what is technically defined as a two-phase system: water and particles. Already we have to be more precise: water itself does contain particles. First of all water molecules of course, but also dissolved ions, which (see **Chapter 3**), can play a key role in the aggregation of clay particles. Even fresh water contains ions, for example HCO_3^- and Ca^{2+} ions which are the predominant ions found in river waters, and originate from the weathering of limestone or feldspar (carbonate rocks). These type of rocks are typically found in Europe, North America and Asia. The weathering of silicate rocks will produce different types of ions. In chemistry, one does therefore speak of a *solution* (water + ions) and not of *water* (as is done in civil engineering). In a large branch of colloid science, solutions are seen as a continuum. This means that the solution is a fluid with given properties and that no distinction is made for the individual ionic particles or water molecules in the water. This fluid has bulk properties such as density and ionic concentration. In this book, water will be seen as a continuum.

The particles we refer to are predominantly clay particles, but, as we will see in **Chapters 4 and 5** mud is composed of numerous other particles. These particles can either be seen as individual particles which interact with each other and the gravity field or as a continuum with a given (time-dependent) density. This continuum approach will in particular be used in **Chapters 9 and 10** but it will become clear that all models presented in this book depend in a way or another on the properties of the microscopic constitutive particles.

Mainly depending on the size of the particles in suspension, three different types of fluids can be distinguished: solutions, colloids and suspensions. They are represented underneath:

	a solution	a colloid	a suspension
	size < nm	size < 10 μm	
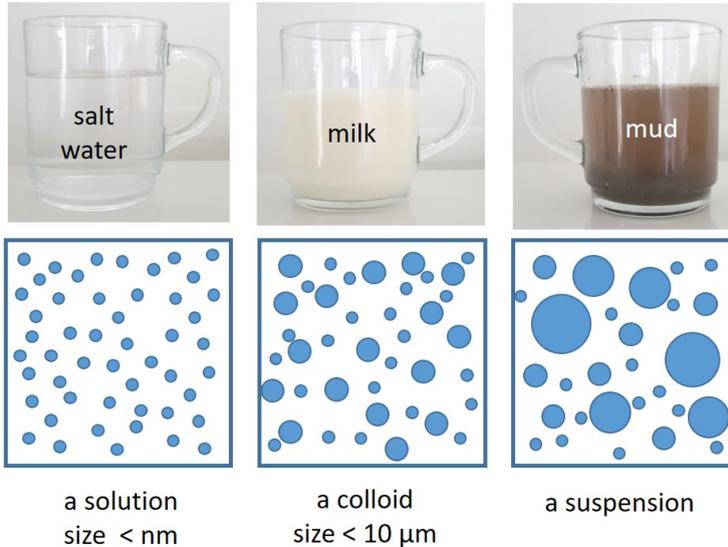
*A **solution** is a homogeneous mixture that appears clear, such as salt water in a glass. In a **colloid**, such as milk, the particles are much larger but remain dispersed and do not settle. A **suspension**, such as mud, is a heterogeneous mixture of suspended particles that can either remain in suspension (for the colloidal fraction) or can settle (for the coarser fraction).*

A **solution** is a homogeneous mixture resulting from the dissolution of a solute (ex: salt) in a solvent (ex: water). In the case of salt in water, one speaks of **electrolyte solutions (or electrolytes)**, as dissolved salts are charged and electrically conducting. It can happen that a solution is **saturated** in one of its solutes, in which case it cannot dissociate this solute anymore. This happens for example when too much salt is added in a glass of water: some salt crystals will remain at the bottom of the glass. Note that due to gravity, when the height of the water column is important, like in the sea, some stratification can be observed in the solution, and its density may vary as function of height.

A **colloid (or colloidal suspension)** is composed of particles or droplets (in which case one speaks of an **emulsion** – milk is for instance an emulsion) that are much larger than ions or molecules, but that are small enough not to settle out. These colloidal particles are dispersed in a solvent. We will only consider water or an electrolyte solution as solvent in the remaining of the book, as these are the most common solvents for sediment particles. If the colloidal particles or droplets are dispersed in a gas one refers to an **aerosol** (smoke and mist are aerosols).

To be classified as colloidal, a material must have one or more of its dimensions (length, width, or thickness) in the approximate range of 1-1000 nm. The particles in a colloid are large enough to *scatter* light, a phenomenon called the **Tyndall effect**. This can make colloidal suspensions appear cloudy or opaque. The Tyndall effect is an easy way of determining whether a fluid contains a significant amount of colloidal particles. When light is shined through a true solution, the light passes cleanly through the solution (ex: an electrolyte solution), however when light is passed through a colloidal solution, the substance in the dispersed phases scatters the light in all directions, making it readily seen (ex: milk).

John Tyndall (1820 – 1893) was a prominent 19th-century physicist. His initial scientific fame arose in the 1850s from his study of diamagnetism. Later he made discoveries in the realms of infrared radiation and the physical properties of air. Tyndall also published more than a dozen science books which brought state-of-the-art 19th century experimental physics to a wide audience. From 1853 to 1887 he was professor of physics at the Royal Institution of Great Britain in London.
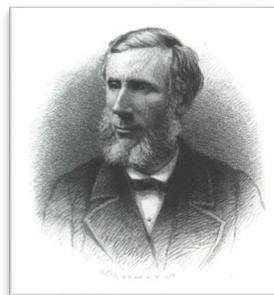
extract from the book The Glaciers of the Alps, by John Tyndall[6]

Expedition of 1857: the Lake of Geneva. Blueness of the water

"On Thursday, the 9th of July, 1857, I found myself upon the Lake of Geneva, proceeding towards Vevey. I had long wished to see the waters of this renowned inland sea, the colour of which is perhaps more interesting to the man of science than to the poets who have sung about it. Long ago its depth of blue excited attention, but no systematic examination of the subject has, so far as I know, been attempted. It may be that the lake simply exhibits the colour of pure water. Ice is blue, and it is reasonable to suppose that the liquid obtained from the fusion of ice is of the same colour; but still the question presses—"Is the blue of the Lake of Geneva to be entirely accounted for in this way?" The attempts which have been made to explain it otherwise show that at least a doubt exists as to the sufficiency of the above explanation.

It is only in its deeper portions that the colour of the lake is properly seen. Where the bottom comes into view the pure effect of the water is disturbed; but where the water is deep the colour is deep: between Rolle and Nyon for example, the blue is superb. Where the blue was deepest, however, it gave me the impression of turbidity rather than of deep transparency. At the upper portion of the lake the

---

[6] http://www.gutenberg.org/files/34192/34192-h/34192-h.htm#Page_33

water through which the steamer passed was of a blue green. Wishing to see the place where the Rhone enters the lake, I walked on the morning of the 10th from Villeneuve to Novelle, and thence through the woods to the river side. Proceeding along an embankment, raised to defend the adjacent land from the incursions of the river, an hour brought me to the place where it empties itself into the lake. The contrast between the two waters was very great: the river was almost white with the finely divided matter which it held in suspension; while the lake at some distance was of a deep ultramarine.

The lake in fact forms a reservoir where the particles held in suspension by the river have time to subside, and its waters to become pure. The subsidence of course takes place most copiously at the head of the lake; and here the deposit continues to form new land, adding year by year to the thousands of acres which it has already left behind it, and invading more and more the space occupied by the water. Innumerable plates of mica spangled the fine sand which the river brought down, and these, mixing with the water, and flashing like minute mirrors as the sun's rays fell upon them, gave the otherwise muddy stream a silvery appearance."

The Tyndall effect can be linked to theories developed by **Rayleigh** and **Mie**, in the frame of optics: when the particle is smaller than the wavelength, Rayleigh theory is used, which predicts that for this type of particles, short-wavelength light is scattered more than longer wavelengths. The blue colour (short wavelength) is therefore more scattered than the red (long wavelength): this is why the sky appears blue to us, as the gas molecules of the atmosphere are much more effective in scattering the blue wavelength. Some suspensions, containing nanoparticles, similarly have a blueish tinge. Clouds, on the other hand, are formed by water droplets that are much larger than gas molecules and scatter in the same way all the parts of the light spectrum: this is why clouds appear white (like milk does).

A solution is clearly scattering much less light than a colloidal suspension, and this is why the way a fluid is scattering light can be used to determine if this fluid contains a substantial amount of colloidal-sized particles.

Note that the fact that a true solution might be coloured is not due to the scattering of light, but on the contrary to the adsorption of light by the atoms and molecules that compose this solution. What we see is not the colour absorbed, but its complementary colour, originating from the removal of the absorbed wavelengths. For example, beta-carotene, an organic pigment abundant in plants and fruits, has a maximum absorption at 454 nm (blue light) and consequently what we see appears orange (the complementary colour for that type of blue). (Beta-carotene takes its name from the carrot (daucus carota) from which it can be extracted.)

A **suspension** (not necessarily colloidal) is a fluid composed of a solvent and particles in suspension. At low or at no shear rate, these fluids are usually containing colloidal particles to be classified as "suspensions". If the particles constituting the suspension

would be all much larger than a few microns the time the particles would remain in suspension would be very limited at low shear: sand particles for example, settle out very rapidly and large gas bubbles usually move rapidly towards the water/air interface. In natural systems, however, it is possible to keep fine particles in suspension by turbulent mixing. This mixing insures that sand can be transported over some considerable distance during storms for example.

Typical orders of magnitude for ions, colloidal and sand particles:

	Ions	colloidal particles	sand
typical particle size	10^{-10} m	10^{-6} m	10^{-4} m
number of particles in a spoon (15 mL)	10^{25}	10^{13}	10^{7}
equivalent in number to	number of sand particles in the Sahara	10000 times the population of Earth	number of inhabitants of Paris

The unit to measure the number of ions in a litre is mol/L. A mole (symbol: mol) corresponds to 6.02×10^{23} particles, which is the number of ^{12}C particles contained in 12 g. One speaks of Avogadro's number (symbol: N_A) to express the fact that there are 6.02×10^{23} particles in one mole ($N_A = 6.02 \times 10^{23}$ mol^{-1}). A solution of 1 mol/L therefore contains 6.02×10^{23} particles per litre. The molar mass of a product enables to determine how many grams of product is contained in one mol. The molar mass of NaCl salt (kitchen salt) is 58 g/mol. There are therefore 6.02×10^{23} salt particles in 58 g of NaCl, which will give, when dissociated in water, 6.02×10^{23} Na^+ particles and 6.02×10^{23} Cl^- particles. Comparatively, 58 g of clay particles which have a density of 2.6 kg/L (2600 kg/m^3) and a size of 1 micron will give approximatively 10^{13} clay particles, i.e. 10^{10} times less particles, or, in other words, if 58 g of clay is mixed with 58 g of salt in one litre of water, for each clay particle, there will be 10^{10} ionic particles. The typical concentration of salt ions in natural systems ranges from less than mM (millimol/L or 10^{-3} mol/L) for fresh water system to hundreds of mM for sea water. Clay concentrations are usual of the order of a few mg/L at sea to hundreds of mg/L and beyond in estuaries.

The difference between clay, silt and sand particle sizes

Sand is a naturally occurring granular material composed of finely divided rock and mineral particles. It is defined by size, being finer than gravel and coarser than silt. Here we differentiate sand, silt and clay based on *size*. Depending on the context, silt can refer to a type of mud or particles with a specific grain size. A similar confusing double definition exists for clays, see the table below. Particles that have a "clay" size may not necessarily be **clay minerals** (organic matter for instance can have a clay size but is not composed of clay minerals).

Sand feels gritty when rubbed between the fingers. Silt, by comparison, feels like flour. Even though the exact definitions, in terms of size, of sand, silt and clay are depending on countries, one assume in general that they can be defined as follows:

	pebbles 63 mm < size < 2 mm	**sand** 2 mm < size < 63 µm
boulder 63 mm < size	**silt** 63 µm < size < 2 µm	(invisible) **clay** size < 2 µm

Traditionally, sieves could not be made for particle sizes smaller than 63 µm. This is why the lower limit of 63 µm was adopted for sand. The size of silt particles is measured traditionally by **hydrometer** test. Nowadays, **Static Light Scattering** (laser diffraction) devices enable to measure the whole range of particle sizes (from 1 nm to 1 mm) with a single device, see **Chapter 2**.

The most common constituent of sand, in inland continental settings and non-tropical coastal settings, is silica (silicon dioxide, or SiO_2), usually in the form of quartz, which, because of its chemical inertness and considerable hardness, is the most common mineral resistant to weathering. In contact with water or electrolyte, the presence of hydroxyl (silanol) groups have however been observed on the surface of silicas and silicates, leading to the creation of a surface charge.

SAND   LOAM   CLAY

0 – 10% clay   10 – 30% clay   50 – 100% clay
0 – 10% silt   30 – 50% silt   0 – 45% silt
80 – 100% sand   25 – 50% sand   0 – 45% sand
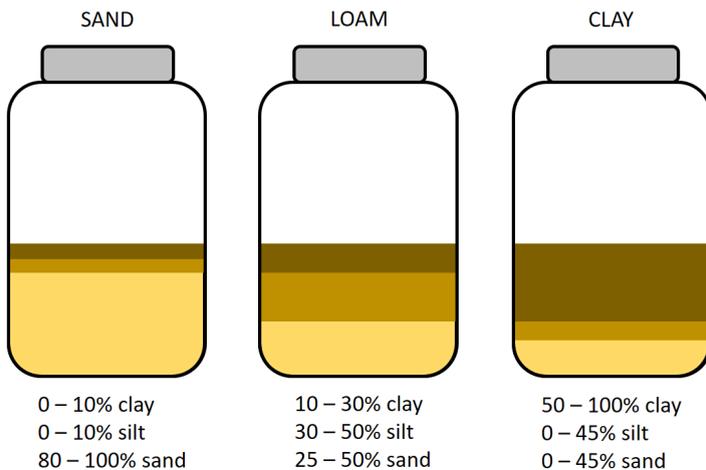
An easy way to find out roughly the composition of a sediment is to use a jar test: Fill the jar halfway with sediment. Add water to fill up the jar, close the lid, shake,

and set on the table. The sand will settle out quickly, in a matter of hours (like "Sand" in the example underneath). The middle jar ("Loam") represents the mixture of clay and silt, which takes up to 24 hours to settle out. The third "Clay" jar represents how long it takes clay content to settle out, in this example up to three days.

Loam is soil composed mostly of sand (particle size > 63 μm), silt (particle size > 2 μm), and a smaller amount of clay (particle size < 2 μm). Loam soils generally contain more nutrients, moisture, and humus than sandy soils, have better drainage and infiltration of water and air than silty soils, and are easier to till than clay soils. Loam, (combined or not with straw), has been used in construction since ancient times.

The definition of mud, silt and clay

Mud is composed, in part, of **clay mineral** and water. For the other parts, most types of mud are include ions and organic material. Particles of different sizes, such as **clay, silt** and **sand** particles can also be incorporated in mud. The most common definitions of mud, silt and clay in English, French and Dutch are given in the table below[7].

Clay particles are colloidal particles when defined by their size (< 2μm). The lower size range of silt particles can also be seen as colloidal, but in most cases silt particles (when composed of mineral clay) are non-cohesive, which implies that their dynamics in the water column are dominated by the gravity force and hydrodynamics similarly to sand particles.

Clay particles can be found in all water bodies (sea, lake, river), but their relative ratio compared to silt and sand varies of course from site to site. A coastal wetland where clays and silts can be found are **mudflats**. Mudflats are formed when mud is deposited by tides (tidal flat) or rivers. Mudflats (French: vasière; Dutch: wad) can be seen as exposed layers of (wet) mud. A tidal flat is submerged approximately twice daily enabling estuarine silts, clays and organic matter to be deposited. On the Baltic Sea coast of Germany in places, mudflats are exposed not by tidal action, but by wind-action driving water away from the shallows into the sea. These wind-affected mudflats are called *windwatts* in German. Mudflat hiking (Dutch: Wadlopen; German: Wattwandern; Danish: Vadehavsvandring) is a recreation enjoyed by Dutch, Germans and Danes in the Netherlands, northwest Germany and in Denmark. Mudflat hikers are people who, with the aid of a tide table, use a period of low water to walk and wade on the watershed of the mudflats. The study of the

---

[7] *See in particular for the French definitions: " La vase" by J. Bourcart and C. Francis-Boeuf (1942)*

morphology of mudflats is an on-going topic of research (see middle figure below where students fix instruments to a rig).

English	French	Dutch
mud liquid or semi-liquid mixture of principally water and clay minerals	**boue** from old Celtic word "baw" (cf gall. baw, "dirt"): creamy substance that is found, after rainfalls, on roads .	**modder** *Mud* and *modder* are etymologically related to Middle Low German and Dutch *mudde* and *modde* denoting something wet or dirty
	vase from old Saxon word "wase": wet mud that is to be found under water (in lakes, rivers or seas) or deposited on mudflats	**slib** probably derived from old Dutch *slik/slijk*: deposited wet mud (on mudflats)
silt a - mud carried by running water and deposited as a sediment, especially in a channel or harbour. b - particles with a size > 2 μm and < 63 μm	**limon** a - from latin "limo - limus" ("mud"): mud that accumulates due to water flows on river sides. b - particles with a size > 2 μm and < 63 μm	**silt** a - from Middle Dutch *silte, sulte* "salt marsh brine". Contrary to slib, silt is found in open air and can therefore oxidize. b - particles with a size > 2 μm and < 63 μm
clay[8] a - mud that can be moulded when wet, used in bricks, pottery, and ceramics b - particles with a size < 2 μm c - clay (minerals) are hydrous aluminium phyllosilicates	**argile** a - from latin "argila": mud having plastic properties, used in bricks, pottery, and ceramics b - particles with a size < 2μm c – *argiles (minéraux argileux)* are hydrous aluminium phyllosilicates	**klei** a - *potklei*: clay used in pottery b – *klei* is composed in majority of *lutum* (from latin "lut(um)": clay) i.e. clay particles with a size < 2μm c – *klei(mineralen)* are hydrous aluminium phyllosilicates

---

[8] Geologists and soil scientists call "clays" particles less than 2 μm. The Dutch *klei* (clay) is composed in majority of *lutum* particles, which are by definition particles smaller than 2 μm. Sedimentologists often use the limit 4–5 μm, and colloid chemists use 1 μm.
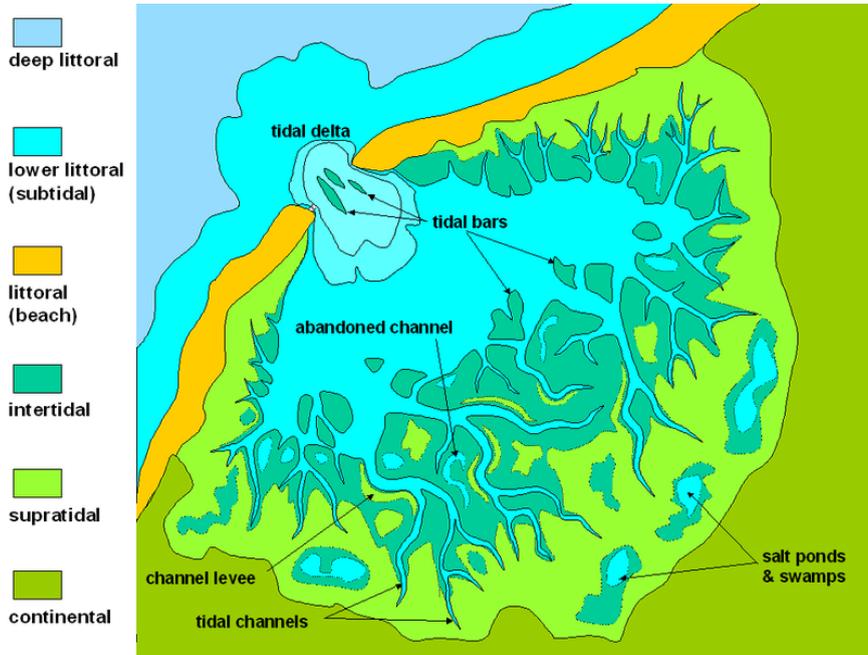
Tidal flat


Wadden Sea in the Netherlands[9]

Clay minerals

Clay minerals are common weathering products and low-temperature hydrothermal alteration products. Clay minerals are very common in soils, in fine-grained sedimentary rocks such as shale, mudstone, and siltstone and in fine-grained metamorphic slate and phyllite. Depending on the soil's content in which it is found, clay can appear in various colours from white to dull grey or brown to deep orange-red.

Clay minerals are hydrous aluminium phyllosilicates. Phyllosilicates are minerals formed by parallel sheets of silicate tetrahedra, sometimes with variable amounts of iron, magnesium, alkali metals, alkaline earths, and other cations found on or near some planetary surfaces. Depending on the academic source, there are three or four main groups of clays: **kaolinite**, **montmorillonite-smectite**, **illite**, and **chlorite**. Chlorites are not always considered a clay, sometimes being classified as a separate group within the phyllosilicates. There are approximately 30 different types of "pure" clays in these categories, but most "natural" clay deposits are mixtures of these different types, along with other weathered minerals.
Single mineral clay particles are in the clay (colloidal) size range. The identification of the type and structure of clay minerals requires special analytical techniques as x-ray and electron diffraction methods (XRD, XRF and SEM in short).

Clay mineralogy

The crystalline structure of clay minerals is built up from different types of sheets or layers. The fundamental building blocks of these sheets are the **tetrahedron** and the **octahedron** units. The tetrahedron is composed of either a central silicon or aluminium surrounded by four oxygen ions in a tetrahedral coordination. The octahedron is composed usually of a central polyvalent cation surrounded by six oxygen (O) or hydroxyl (OH) ions in an octahedral coordination. Whether a cation forms tetrahedral or octahedral coordination with oxygen depends on the relative size of the cations and anions involved.
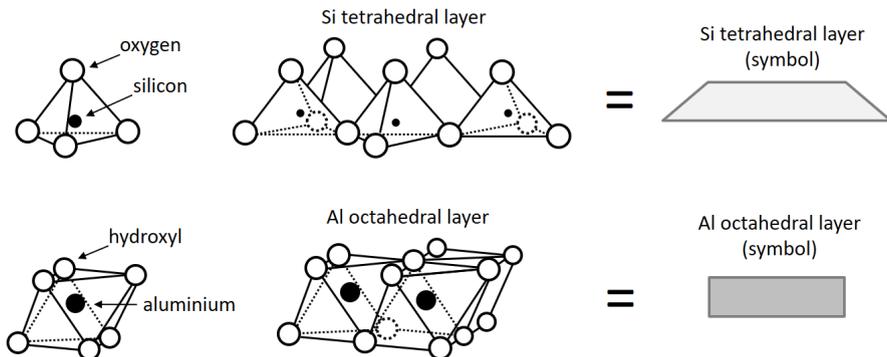
Tetrahedral layer: All the tetrahedrons in a tetrahedral layer are in the same plane and therefore their tips point in the same direction. The unit cell formula of the Si tetrahedron is $(Si_4O_{10})^{4-}$. Electrical neutrality is obtained by replacement of four oxygens by hydroxyls (-OH) or by union with a sheet of different composition that is positively charged.

Octahedral layer: This sheet structure is composed of magnesium or aluminium in octahedral coordination with oxygens or hydroxyls. If the cation is trivalent, then normally only two-thirds of the possible cationic spaces are filled and the structure is termed dioctahedral. In the case of aluminium the composition is $Al(OH)_3$. This composition and structure form the mineral gibbsite. If the octaedrally coordinated cation is divalent then normally all possible cation sites are occupied and the structure is trioctahedral. In the case of magnesium the composition is $Mg(OH)_2$, which forms the mineral brucite.

Combining the tetrahedral and octahedral layers, the following minerals are formed:

kaolinite	chlorite	illite/mica	montmorillonite
(non swelling)	(non swelling)	(non swelling)	(swelling)
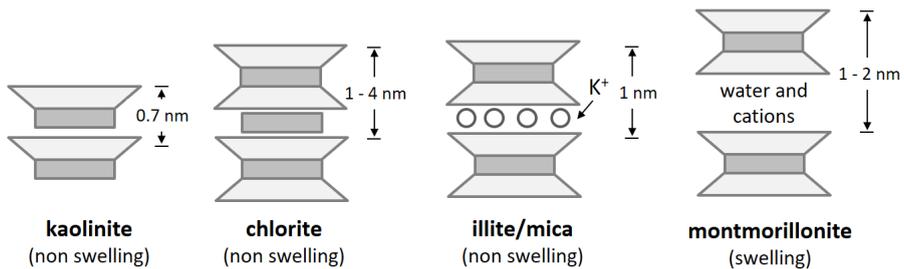
Each platelet has a very large length to width ratio. For commodity, the symbols used above for the clay minerals have a much smaller ratio.

The chemical structure of minerals is usually far from ideal, as the actual composition of minerals is frequently altered by isomorphous substitution, i.e. the substitution of ions within the structure. Weathering allows Si^{4+}, Al^{3+} and Mg^{2+} to be substituted with cations with comparable ionic radii in their respective tetrahedral and octahedral sheets.

In each of these minerals, the layers combine differently. For kaolinite, we get:
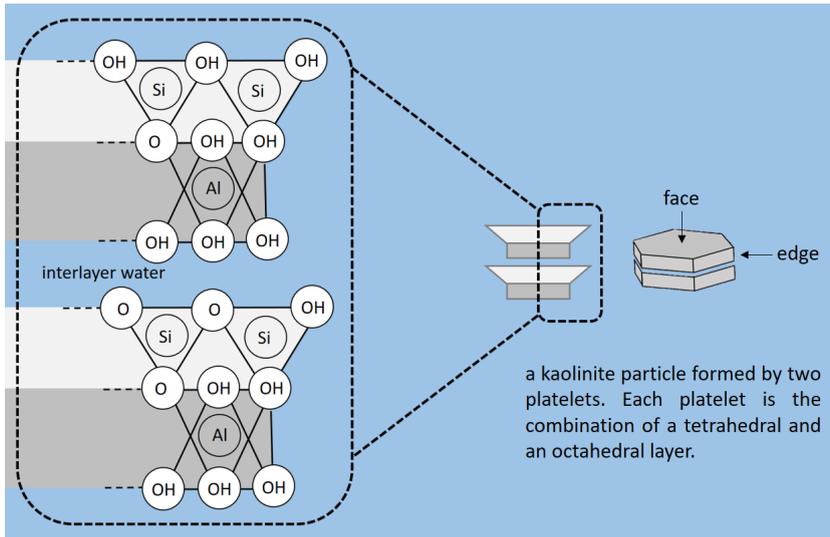
A kaolinite particle usually consist of a pile (stack) of platelets (tetrahedral + octahedral layers forming one platelet). We have here only represented a pile of two, but the number of platelets forming a single particle can vary significantly:
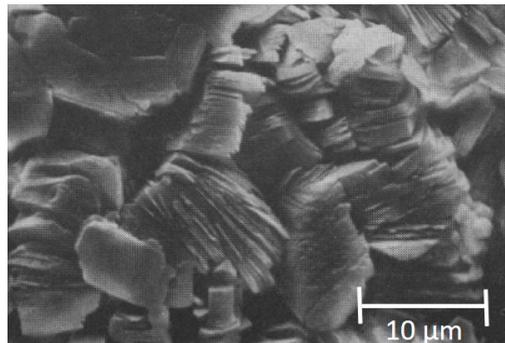
Scanning Electron Microscopy (SEM) picture of kaolinite platelets

The different classes of Phyllosilicates

Serpentine group

Antigorite – $Mg_3Si_2O_5(OH)_4$
Chrysotile – $Mg_3Si_2O_5(OH)_4$
Lizardite – $Mg_3Si_2O_5(OH)_4$

Clay mineral group

Halloysite – $Al_2Si_2O_5(OH)_4$
Kaolinite – $Al_2Si_2O_5(OH)_4$
Illite – $(K,H_3O)(Al,Mg,Fe)_2(Si,Al)_4O_{10}[(OH)_2,(H_2O)]$
Montmorillonite – $(Na,Ca)_{0.33}(Al,Mg)_2Si_4O_{10}(OH)_2 \cdot nH_2O$
Vermiculite – $(MgFe,Al)_3(Al,Si)_4O_{10}(OH)_2 \cdot 4H_2O$
Talc – $Mg_3Si_4O_{10}(OH)_2$
Sepiolite – $Mg_4Si_6O_{15}(OH)_2 \cdot 6H_2O$
Palygorskite (or attapulgite) – $(Mg,Al)_2Si_4O_{10}(OH) \cdot 4(H_2O)$
Pyrophyllite – $Al_2Si_4O_{10}(OH)_2$

Mica group

Biotite – $K(Mg,Fe)_3(AlSi_3)O_{10}(OH)_2$
Muscovite – $KAl_2(AlSi_3)O_{10}(OH)_2$
Phlogopite – $KMg_3(AlSi_3)O_{10}(OH)_2$
Lepidolite – $K(Li,Al)_{2-3}(AlSi_3)O_{10}(OH)_2$
Margarite – $CaAl_2(Al_2Si_2)O_{10}(OH)_2$
Glauconite – $(K,Na)(Al,Mg,Fe)_2(Si,Al)_4O_{10}(OH)_2$

Chlorite group

Chlorite – $(Mg,Fe)_3(Si,Al)_4O_{10}(OH)_2 \cdot (Mg,Fe)_3(OH)_6$

Lizardite

Margarite

The surface charge of clay minerals

As we will see in the forthcoming **Chapters 3 and 5**, the surface charge of mineral clays play a very important role in understanding their aggregation behaviour. The surface of a clay mineral particle is defined as the parts of the clay exposed to the bulk water (in contrast to the interlayer water, see figure above).

The surface charge of clays is very dependent on pH. In particular the hydroxyls (OH) are exposed on the surfaces and edges of the particles and they dissociate in water under the influence of pH:

$$SiOH + H_2O \leftrightharpoons SiO^- + H_3O^+$$
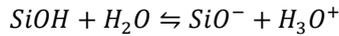
The higher the pH, the greater is the tendency to dissociate the hydroxyl OH into H_3O^+ (often noted H^+ for simplification) and SiO^-.

For kaolinite, the following pH dependence is observed:

| pH < 5.5 | 5.5 < pH < 7.5 | pH > 7.5 |
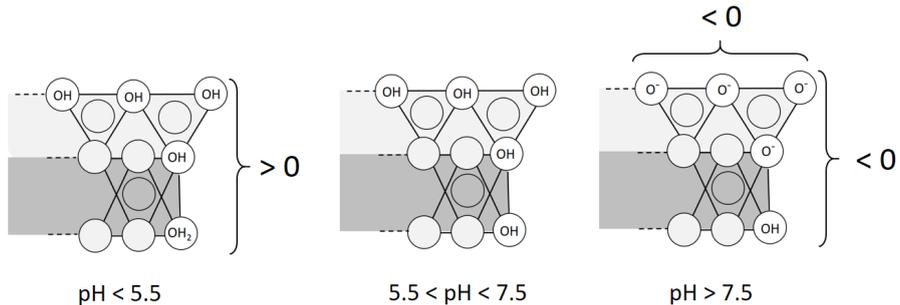
At low pH, the edges of a platelet are positively charged, whereas the faces are not (or very weakly negatively) charged. At medium pH, the edges are uncharged, and the faces are not (or weakly negatively) charged. At high pH, both edges and faces are negatively charged.

The way clay particles are interacting with each other will be studied in **Chapter 5**.

The different states of clay

If the water content of a sample made predominantly of water, clay and/or silt is varied, different states can be observed, for example:

| liquid | plastic | solid (after drying) |

In soil science, criterions have been established to quantify these states. These are traditionally based on the measurements of the **Atterberg limits** of the sample. Four states are defined: solid, semi-solid, plastic and liquid. The boundary between each state is defined on a change of the soil's behaviour. From the extended studies done over the years, it is now possible to distinguish between different types of silts and clays, based solely on the Atterberg limits of the sample. This is in particular possible

for "standard" soils, i.e. as commonly found in nature. Nonetheless, caution is required when the soil composition is complex (including a lot of organic material for example), as these non-mineral components, even in low amounts, can greatly influence the properties of the soil.

Atterberg and the Atterberg limits

Albert Mauritz Atterberg (1846–1916) was a Swedish chemist and agricultural scientist who created the Atterberg limits. He received his Ph.D. in chemistry from Uppsala University in 1872. At the age of 54, while continuing his work on chemistry, he began to focus his efforts on the classification and plasticity of soils. Atterberg was apparently the first to suggest the limit <0.002 mm as a classification for clay particles.

Atterberg found *plasticity* to be a particular characteristic of clay and as a result of his investigations arrived at the consistency limits which bear his name today. He also conducted studies aiming to identify the specific minerals that give a clayey soil its plastic nature.

The importance of Atterberg's work has never been fully realized in his own field of agricultural science. Its introduction to the field of geotechnical engineering was due to **Karl Terzaghi**, who came to realise its importance at a relatively early stage of his research. Terzaghi's assistant, **Arthur Casagrande**, standardized the tests in his paper in 1932 and the procedures have been followed worldwide ever since.

The Atterberg limits are defined as follows:

The liquid state and liquid limit

In the **liquid state**, the clay sample cannot withstand any type of loading, it is said to have no strength. When poured, It will flow like a fluid. The viscosity of the fluid can however change with clay concentration, this will be detailed in **Chapter 7**. The **Liquid Limit** (LL) is defined as the moisture content which soil begins to behave as a liquid material and begins to flow.

Casagrande cup with sample in which a groove has been made.

The liquid limit (LL) is obtained by putting a paste-like sample into a brass cup (**Casagrande cup**). A groove is cut in the paste using a standard tool. The cup is then bumped a standard number of times using a crank-operated mechanism. The LL is the moisture content at which the shear strength of

the sample is so small that the soil " flows" to close the groove. The LL can also been obtained using a **cone penetrometer**, where a pointy cone is made to free-fall on a sample from a certain height. If the cone penetrate a given distance (20 mm) inside the sample the LL is reached.

The plastic state and plastic limit

In the **plastic state**, as in the example of the potter's clay, the sample is in a state where it is possible to make shapes out of it. If the sample is left to dry out for a short time (so that it is not actually completely dry) it will lose its plasticity. If we then try to shape it, many cracks will appear, indicating that the clay sample is in its **semi-solid state**.

The **Plastic Limit** (PL) is defined as the moisture content at which the soil begins to behave as a plastic material.

The plastic limit (PL) is obtained by rolling down a dough-like sample into threads of 3 mm in diameter on a glass plate. Being able to roll such a thread is an indication that the PL is reached. If the water content is too high, the thread will be too soft and impossible to roll, if it is too low, some crumbling will take place, indicating that the semi-solid state is reached.

The shrinkage limit and the solid state

The **Shrinkage Limit** (SL) is defined as the moisture content at which no further volume change occurs with further reduction in moisture content.

The shrinkage limit (SL) is obtained when a semi-solid sample does not shrink anymore by losing moisture content.

In the **solid state**, the sample cannot be shaped, and becomes brittle.

Moisture content

In the representation of the Atterberg limits hereunder, w represents the **moisture content** (also called **water content**) usually defined as:

$$w\ (\%) = 100 \times \frac{m_{\text{water}}}{m_{\text{dry sample}}} = 100 \times \frac{m_{\text{wet sample}} - m_{\text{dry sample}}}{m_{\text{dry sample}}}$$

where m_{water} represents the mass of water in the sample, $m_{\text{wet sample}}$ the mass of the wet sample and $m_{\text{dry sample}}$ the mass of the sample after drying it at a temperature not exceeding 115°C.
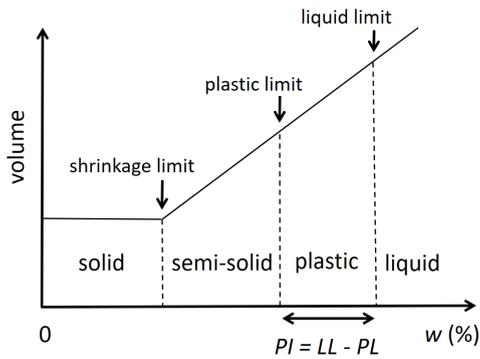
Atterberg limits: depending on water content, several states of a sample made of predominantly water, clay and/or silt can be observed. The transition between one state and another is given as a specific water content (in %). SL: Shrinkage Limit, PL: Plastic Limit and LL: Liquid Limit. The soil volume increases with moisture content once passed the solid state.

As a classification of fine-grained soils, **the plasticity index** (PI) is introduced:

$$PI = LL - PL$$

The water content of a soil can be related to the soil's **stress-strain** response[10]:
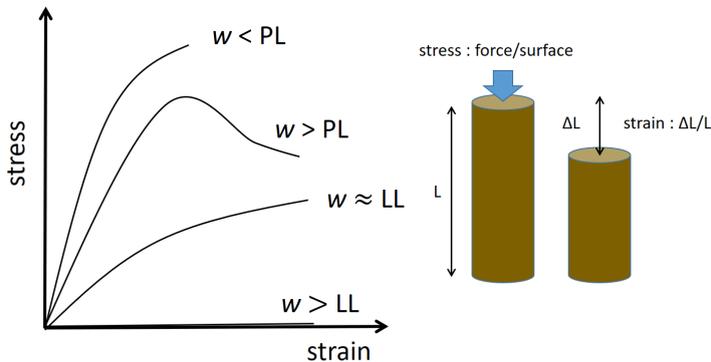
By increasing the water content (w) of a soil, the amount of stress to exert on it to get a given strain is reduced.

The stress-strain is obtained by measuring the amount of deformation (strain) at distinct intervals of tensile or compressive loading (stress). More about stress and

---

[10] See for instance *An introduction to Geotechnical Engineering*, by Holtz & Kovacs, Prentice-Hall inc.

strain is to be found in **Chapters 7 and 10**, where the rheology and consolidation of clays is discussed.

Mud found in water systems is mainly liquid, but sampled at the bottom of a water body its consistency can be plastic. Semi-solid and solid mud is found on land, as it requires the dissipation of a certain amount of water, which can be obtained by evaporation. To study and characterize the physical properties of liquid and plastic mud rheological tests are usually performed, see **Chapter 7**. These tests are to be preferred to Atterberg limits tests as they can give quantitative estimations for the stress/strain (or shear rate) responses. For more consolidated materials, other tests such as the **oedometer** tests are performed (see **Chapter 10**). Attenberg limit tests have therefore only an engineering value and are usually used in the preliminary stages of designing any structure to ensure that the soil will have the correct amount of shear strength. We nonetheless like to present them for their historical value.

Cassagrande's plasticity chart

Cassagrande (1932) studied the relationship of the plasticity index to the liquid limit of a variety of natural soils and proposed a plasticity chart, given here:
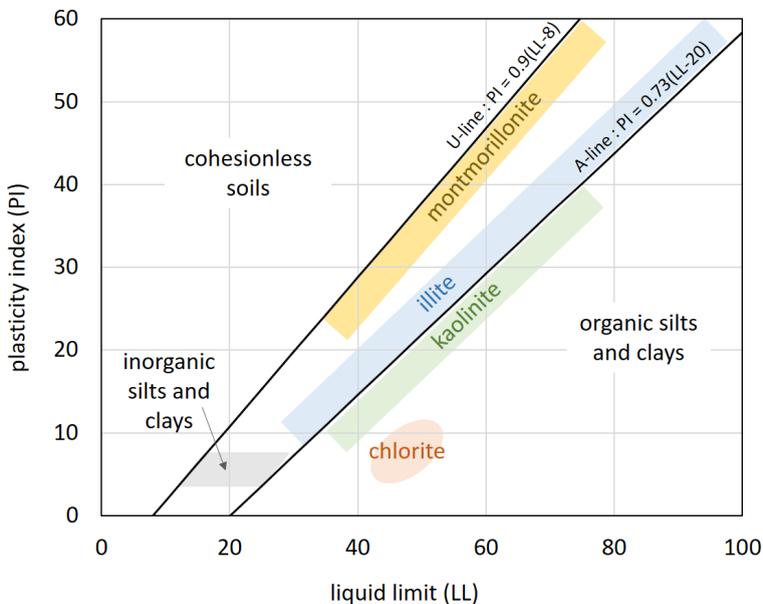
Casagrande's plasticity chart, showing several representative soil types

Two lines are defined in the chart:

The **A-line** which separates the inorganic clays (above the A-line) from the inorganic silts (below the A-line).

The **U-line** which is approximately the upper limit of the relationship of the plasticity index to the liquid limit for any currently known soil. Above the U-line the soils are therefore assumed to be cohesionless.

"Activity" of clays

It has been demonstrated that the plasticity index to the clay fraction content (< 2µm) is in good approximation constant[11]:
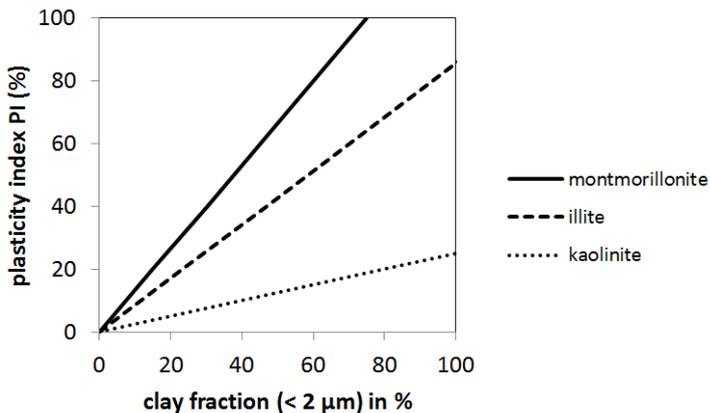
This led to the definition of "**colloidal activity**":

$$\text{activity} = \frac{\text{plasticity index}}{\text{clay fraction}}$$

In simple words, the clay (colloidal) content of a soil is directly proportional to its cohesiveness. We will see in **Chapter 5** that montmorillonite clay has special properties that explains the fact it is more "active" than other clays. The presence or not of salt within the sample also greatly influences the values of the plasticity index[12]. This can be understood from colloidal interactions, which strongly depend on salinity, as will be seen in **Chapter 3**.

---

[11] Skempton, A. W. "The colloidal activity of clays." Selected papers on soil mechanics (1953): 106-118.
[12] Bjerrum, Laurits. "Geotechnical properties of Norwegian marine clays." Geotechnique 4.2 (1954): 49-69.

Illustrations

the tomb –chapel of Nebamun : Nebamun hunting in the marshes, adapted from a wall painting currently to be seen at the British Museum; ca. 1350 BC (public domain)
https://commons.wikimedia.org/wiki/File:TombofNebamun-2.jpg

John Tyndall (public domain)
https://en.wikipedia.org/wiki/John_Tyndall

Mudflat (creative commons license)
https://en.wikipedia.org/wiki/Mudflat

lizardite photograph (creative commons license)
https://upload.wikimedia.org/wikipedia/commons/7/71/Lizardite%2C_Chrysotile-288581.jpg

margarite photograph (creative commons license)
https://en.wikipedia.org/wiki/Margarite

Albert Atterberg (public domain)
http://geotecnia-sor.blogspot.nl/2010/11/consistencia-del-suelo-limites-de_17.html

Casagrande cup (creative commons license)
http://labmodules.soilweb.ca/soil-compaction-atterberg-limits/

Chapter 2

Settling, diffusion and stabilisation

Fine (colloidal) particles can remain in suspension for an extremely long (sometimes infinite) time. In this chapter, we are going to explain why and when these colloidal particles remain in suspension and how they diffuse and settle. We will primarily address **dilute** suspensions i.e. suspensions containing a limited amount of colloidal particles. The settling behaviour of **concentred** suspensions will be discussed in **Chapter 8**.

Stokes' settling velocity

In **Chapter 1**, we have already stated that in colloid science the typical length scale for a particle is 1 µm. By this, we mean *of the order of* 1 µm which means a particle somewhere in the range between 0.01 µm and 10 µm.

Let us now consider a particle of any size > 1 nm in water. Three forces are exerted on this particle: the force of gravity, the force of Archimedes and the force of friction. We furthermore assume that

1) the water far from the particle is at rest
2) we have reached the regime where the particle's velocity is constant (the initial acceleration is not considered)
3) the velocity is small enough for the fluid to be in the **laminar** regime.

The fact that a flow is laminar or not can be evaluated by the estimation of the **Reynolds number**[13] *Re* which represents the ratio of inertial forces (which create turbulence) to viscous forces (which create friction):

$$Re = \frac{v\rho L}{\eta}$$

where v is the fluid's velocity, L is a characteristic length (in our case the size of studied particle, as the particle is setting the fluid in motion) and the kinematic viscosity is given by η/ρ where η is the viscosity and ρ the density of the fluid. The laminar regime is defined by a low Reynolds number ($Re < 10$ for a sphere), implying that the friction force is dominating, in accordance with our initial assumption. The balance of forces gives:

$$m\boldsymbol{g} = 6\pi\eta a\boldsymbol{v}$$

where m is the mass of the particle, compensated for Archimedes, g the gravitation constant, η the viscosity of the water, a the radius of the particle and v its velocity. The fact that the force of friction might be expressed as $6\pi\eta a\boldsymbol{v}$ for a spherical particle in a laminar flow in due to the work of **Georges Stokes** (1819-1903). We have

---

[13] The number has been invented by Stokes but is named after Osborne Reynolds (1842-1912) who popularised its use.

expressed the vectors in bold letters, implying that the vector $\boldsymbol{v}$ is directed along $\boldsymbol{g}$. The mass of the particle, compensated for Archimedes can be expressed as:

$$m = \frac{4}{3}\pi a^3 \left(\rho_p - \rho_w\right)$$

where ρ_p is the density of the particle and ρ_w the density of the water. From the balance of forces we get:

$$\boldsymbol{v} = \frac{2}{9}a^2 \frac{\rho_p - \rho_w}{\eta}\boldsymbol{g}$$

From this expression which is **Stokes' settling velocity**, we can verify that:

(a) the velocity scales as the size squared: a particle twice as big will settle 4 times as fast.
(b) if the density of the water and the particle are the same the velocity is zero, the particle remains in suspension. If $\rho_p < \rho_w$ the particle will float (this happens for certain types of algae for instance)

Several instruments make use of Stokes' settling velocity to assess the size of (sub-) colloidal particles, like the **sedigraph**, the **hydrometer** and the **sediment balance**. These techniques are explained below.
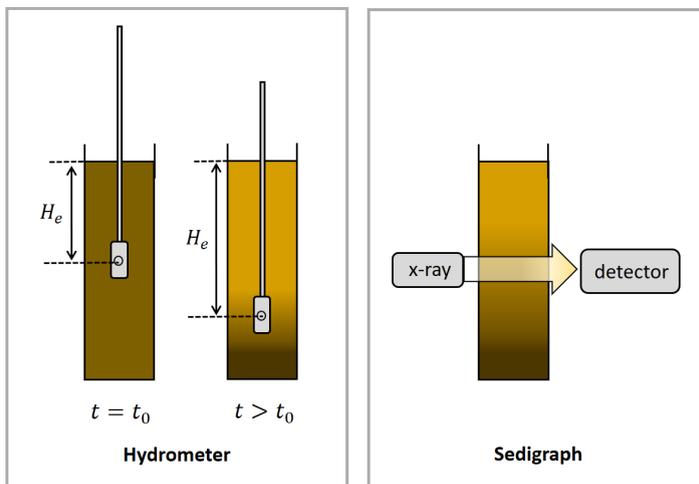
Hydrometer and sedigraph methods are based on recording the density evolution in a settling column as function of time.

Determining particle sizes experimentally

The hydrometer test

For particles that pass the smallest sieve (63 µm in general), their particle size can be assessed by the use of a hydrometer. The hydrometer consists of a cylindrical stem with a precise scale on it and a bulb weighted with lead to make it float upright. The suspension to test is poured into a column, and the hydrometer is gently lowered into the liquid until it floats freely. The hydrometer indicates the ratio of the density of the tested fluid to the density of water (this ratio is called the **specific gravity**). After calibration of the hydrometer, the height H_e (see figure above) can be related to the density of the suspension. The height H_e is increasing in time as the fluid becomes less and less dense. From Stokes' law it is possible to estimate the size of particles that, for a given t, will be above H_e . These are the particles with a radius smaller than:
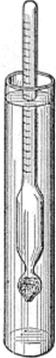
$$a = \sqrt{\frac{9\eta H_e}{2(\rho_p - \rho_w)gt}}$$

The density evaluated from the hydrometer at time t is given by: $\rho(t) = \phi(t)\rho_p + (1 - \phi(t))\rho_w$ where $\phi(t) = \sum N_i(t) V_i/V$ is the total volume fraction of suspended particles; $N_i(t)$ is the number of particles with volume V_i in suspension at time t and V is the total volume of fluid in the column. Assuming that all particles have the same density ($\rho_p = m_i/V_i$ where m_i is the mass of particle i) we see that $\rho_p\phi(t)$ represents the total mass in suspension at time t per unit volume. The percentage in mass of particles remaining in suspension (so with a radius smaller than a) can therefore be estimated from:

$$\% \text{ smaller than } a = 100 \frac{\rho_p(\rho(t) - \rho_w)}{\rho_0(\rho_p - \rho_w)}$$

where $\rho_0 = m/V$; m is the mass of dry soil in the column.

The sedigraph

The sedigraph uses a paralleled X-ray beam to detect changes in suspended sediment concentration: the X-ray attenuation is proportional to mass concentration, following the so-called **Beer-Lambert law**. The beam can be positioned at different vertical positions in the analysis cell, and records the changes at different times. In essence, the principle is very similar to the hydrometer (see above) except that the concentration is recorded through X-ray attenuation instead of the estimation of the specific gravity.

The sediment balance

The sediment balance records the mass M of particles falling on a balance as a function of time: a long rod with a disk attached to it is plunged in a settling column containing a suspension. The rod is attached to a laboratory scale. The assumptions made when performing a sediment balance measurement are that 1) the suspension is homogeneous at t = 0 (start). The suspension has been mixed just prior the measurements and 2) the particles settle according to Stokes' law. This implies that particles with a diameter $d = 2a$ will all have arrived on the balance at the time: $t = 18h\eta/[(\rho - \rho_W)gd^2]$ where h is the height of the liquid in the column. From t = 0 (start) until *t* these particles have been falling continuously on the balance, in a linear way, since the suspension is considered to be properly mixed. It is therefore possible to estimate the percentage of particles that at a time *t* have a settling time smaller than (or equal to) *t*, i.e. are larger than or equal to *d*. The relation is given by:
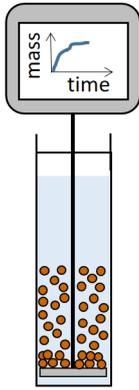
$$\% \text{ particles larger than d } = 100 \times \frac{M - t \cdot dM/dt}{M_{end}}$$

where M is the mass on the balance at time *t*, M_{end} is the final mass on the balance (after one day, typically) and dM/dt represents the slope of the curve *M(t)* at time *t*. The formula is called the **Oden formula**, as it has originally been derived by Oden in 1916[14]. Note that $t \cdot dM/dt$ represents the amount of particles <u>at time *t*</u> with a settling velocity smaller than or equal to *h/t* whereas M represents the amount of particles with a settling velocity smaller than or equal to *h/t* between t = 0 and t. If there would be only particle of one size, then $M - t \cdot dM/dt = 0$ until all the particles have fallen on the balance. When this happens, $dM/dt = 0$ and we find that :

$$\% \text{ particles larger than d } = 100$$

For each of the techniques (hydrometer, sedigraph and sediment balance), it is crucial to know the proper density *ρ* of the particles. In the case of aggregated clay particles, the density can vary significantly from the mineral clay density, in which case the measurements are not reliable. In most protocols it is therefore important to deflocculate the particles prior measurements. This is done by chemical agents or ultrasonication.

A colloidal mineral clay particle, exposed to gravity, will always settle according to the expression for Stokes' settling since clay minerals have a density of the order of

---

[14] Odén, Sven. "Eine neue methode zur bestimmung der körnerverteilung in Suspensionen." Colloid & Polymer Science18.2 (1916): 33-48.

2.6 kg/L (the density of water is of the order of 1 kg/L). As colloidal clay particles can also remain in suspension without settling, it implies that we missed important forces while setting-up the force balance. One is related to so-called Brownian motion. Another is related to the repulsion that can exist between colloidal particles. These forces and their principles are described in the next sections.
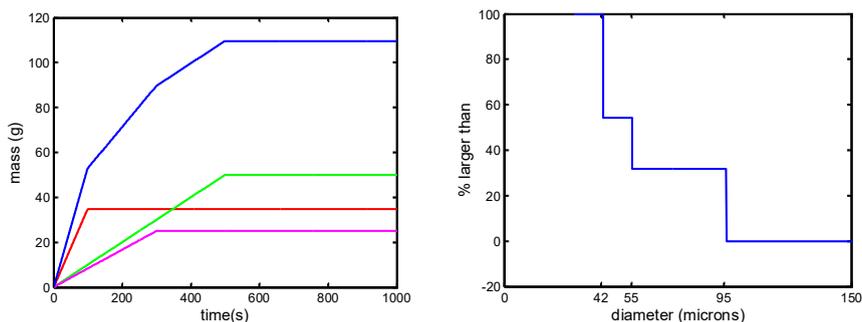
Sediment balance principle. Left: The coloured curves represent the mass on the balance as function of time for suspensions made of: (red): 35 g of particles with a diameter of 95 microns, (magenta): 25 g of particles with a diameter of 55 microns, (green): 50 g of particles with a diameter of 42 microns and (blue) a suspension made by the mixture of the three types of particles. After using Oden's formula on the left blue curve, one obtains the blue curve on the right figure, which gives the amount of particles larger than a certain size in %.

Brownian motion

This transport phenomenon is named after the botanist Robert Brown (1773-1858). In 1827, while looking through a microscope at particles trapped in the water of the interior of pollen grains, he noted that the particles moved through the water but he was not able to determine the mechanisms that caused this motion. It is Albert Einstein in a paper in 1905 that explained in precise detail how the motion that Brown had observed was a result of the particles being moved by individual water molecules. Atoms and molecules had long been theorized as the constituents of matter and therefore Einstein's explanation of Brownian motion served as convincing evidence that atoms and molecules exist. It was verified experimentally by Jean Perrin in 1908. Perrin was awarded the Nobel Prize in Physics in 1926 "for his work on the discontinuous structure of matter" (Einstein had received the award five years earlier "for his services to theoretical physics" with specific citations of different researches).
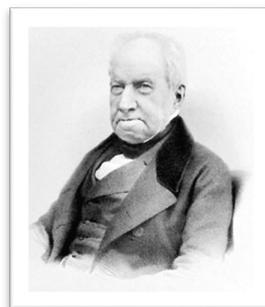
The fact that atoms or molecules (like water molecules) are moving randomly is associated to **temperature**: temperature is an indirect measure of the microscopic

movement of these particles. A whole branch of science (i.e. thermodynamics) is based on the phenomena associated to temperature and temperature fluctuations. From thermodynamic principles, one has defined a proper unit to measure temperature: the **Kelvin**, which is linked to Celsius degrees by:

$$K = °C + 273.15$$

The temperature corresponding to zero Kelvin is called the "absolute zero" , a temperature never to be reached as it would correspond to a state where molecules would have a zero velocity, which is in contradiction with Heisenberg's uncertainty principle in quantum mechanics[15]. In 1908, the Dutch scientist Kamerlingh Onnes[16] and his research group in Leiden were the first ones to be able to measure a temperature of 0.9 K, which made Leiden "the coldest place on Earth". This led to a Nobel prize in Physics in 1913 for K. Onnes.

trajectory of the particle
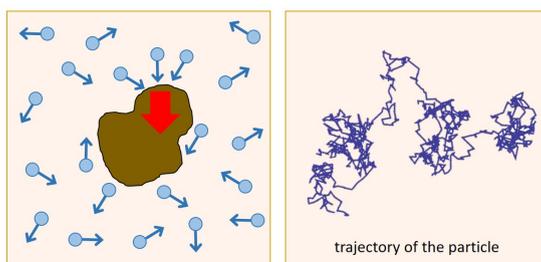
Brownian movement: due to the thermal agitation of the water molecules, colloidal particles experience a random motion (the trajectory of such a particle is given on the right panel)

The direction of the total force (red arrow in the figure above) on the particles due to the bombardment by water molecules is random, and therefore constantly changing, which makes it impossible to derive such a force. At different times the particle is hit more on one side than another, leading to the **stochastic** (random) nature of the motion of the particle. This type of motion is called **Brownian motion** or **random walk**. The trajectory of a particle experiencing a random walk is given in the figure on the right.

Statistically, if one would add all the velocities of all the particles in the water at a given time, the resulting velocity would be zero: there is no global movement (contrary to particles in a flow, where clearly when adding all the velocities, the total velocity will be in the direction of the flow). If one would follow a given particle

---

[15] The uncertainty principle states that it is impossible to know exactly at the same time both the momentum and the position of quantum (i.e. very very small) particles. If the velocity (and momentum) of particles would be zero, then it is obvious that their position should exactly be known.
[16] https://www.lorentz.leidenuniv.nl/history/cold/cold.html

submitted to Brownian motion as function of time, one would observe that this particle gets further and further away from its starting point, but that no direction is privileged. It is a talented mathematician, Louis Bachelier[17], who, in a thesis presented in 1900, demonstrated that what characterized the movement of such a particle was not the arithmetic mean of its displacement $\langle X \rangle$ but in fact its root mean square $\langle X^2 \rangle$:

$$\langle X^2(t) \rangle = \frac{1}{t} \int_0^t x^2(\tau) d\tau$$

where x is the position of particle and t the time. From this relation it can be shown that the root mean square displacement is proportional to time:

$$\langle X^2(t) \rangle = 2dDt$$

where d is the dimension of the movement[18] (linear d=1, planar d=2 or spatial d=3), D is the diffusion coefficient and t the time. Note that if the particle would undergo a regular translation of the type $x(t) = v_0 t$ where v_0 is a constant, one finds:

$$\langle X(t) \rangle = \frac{1}{t} \int_0^t v_0 \tau d\tau = \frac{v_0 t}{2}$$

and in that case, it would be the arithmetic mean of the displacement $\langle X(t) \rangle$ which would be proportional to time (at a given time t, the distance between the particle and its origin at $t = 0$ is *proportional to time*). For a particle under a 3D Brownian motion, one has $d = 3$, hence:

$$\sqrt{\langle X^2(t) \rangle} = \Delta r = \sqrt{6Dt}$$

where $\Delta r = \sqrt{\langle X^2(t) \rangle}$ symbolises the averaged position relative to the original position (at $t = 0$) of the particle (see illustration underneath). In this case, the (averaged) distance between the particle and its origin at $t = 0$ is proportional to the *square root of time*.

One unknown in the previous equation remains the diffusion coefficient D. It is Albert Einstein in 1905 and independently **Marjan Smoluchowski** in 1906, who derived an expression for this diffusion coefficient, based on kinetic theory. In fact the so-called **Einstein-Smoluchowski relation** is an early example of **the fluctuation-dissipation relation**, which is a powerful theorem in statistical physics:
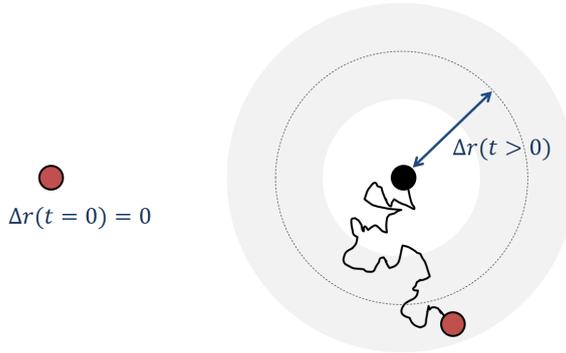
---

[17] https://en.wikipedia.org/wiki/Louis_Bachelier

[18] In Chapter 10 the dimension d will be represented by D. In the present chapter it could be confused with the diffusion coefficient D, so we adapted the notation.

$$D = \mu k_B T$$

where k_B is Boltzmann's constant ($k_B = 1.38 \times 10^{-23} \ J/K$) and T the temperature in Kelvin.

Illustration of a 2D Brownian displacement of a spherical particle. At $t = 0$ the particle is at origin, and $\Delta r = 0$. For a given time $t > 0$ one can infer that the particle will be, on average, at a distance $\Delta r = \sqrt{4Dt}$ from its origin. The direction of the particle relative to its origin is unknown.

The product $k_B T$ is called the **thermal energy**. The **mobility** μ of the particle is defined as the ratio between the particle's terminal velocity v to the applied force F:

$$\mu = \frac{v}{F}$$

Thanks to Newton's equation of motion (that we used earlier to establish Stokes' settling velocity), and using Stokes' s friction force, we have:

$$\boldsymbol{F} = 6\pi\eta a \boldsymbol{v}$$

We have already said that Stokes' frictional force (also called **drag force**) exerted on a spherical particle is only valid at very small Reynolds numbers (at small Reynolds numbers, the fluid flow round the particle is laminar, i.e. non turbulent. This happens for small flow velocities). The force $\boldsymbol{F}$ is due to the bombardment of water molecules on the particles. Without having to know this force (which is of the order of $k_B T/a$), we easily get $\mu = 1/(6\pi\eta a)$ and we obtain the **Stokes-Einstein** relation:

$$D = \frac{k_B T}{6\pi\eta a}$$

This is the theoretical expression for the diffusion coefficient of a sphere subjected to Brownian motion. It is from this relation that one has experimentally determined

the radius of ions and small colloidal particles. The radius thus obtained is called **hydrodynamic radius** (or Stokes radius) as it is known that moving particles always carry some molecules of water with them. The hydrodynamic radius of particles can be assessed by **Dynamic Light Scattering[19]**.

Static and Dynamic Light Scattering

Static and Dynamic Light Scattering (SLS and DLS) are two different techniques that both enable to assess the size of particles. Dynamic Light Scattering can be used for particles in the range [1nm – 1 µm] whereas Static Light Scattering can be used for the range of particle size, besides gravel: [1 nm – 1 mm]. DLS works on the following principle: when light hits particles, it scatters in all direction provided that the particles are smaller than the wavelength (about 600 nm for a red laser). The total scattering intensity of the light that is collected in a photodetector fluctuates over time, owing to the Brownian motion of the particles that produce interferences. This fluctuation can be linked to a characteristic time which is a function of the diffusion coefficient D. From D, using the Stokes-Einstein relation, the size of the particles can be obtained.
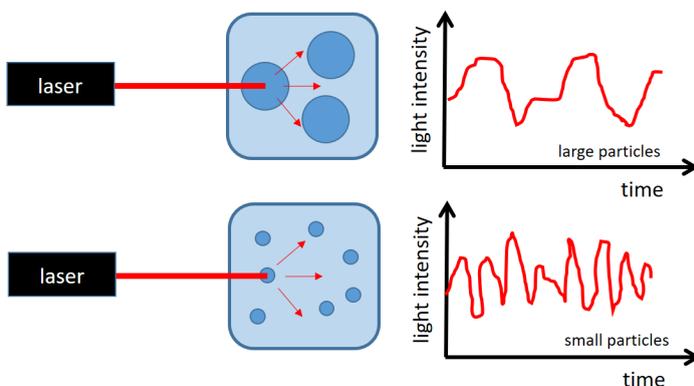
Principle of Dynamic Light Scattering (DLS). The scattered light is measured at an certain angle relative to the incident laser beam direction

Like DLS, Static light scattering (SLS) also measures the scattered light. But, instead of measuring the time-dependent fluctuations in the scattering intensity, SLS makes use of the time-averaged intensity of scattered light. For particle size measurements, SLS makes use of multiple photodetectors, positioned at various angles relative to the incident beam: depending on the size of the particles, the scattering intensity will be different in the different detectors (Rayleigh scattering occurs for extremely

---

[19] for particles larger than a few nm. For ions (of size about 0.1 nm), the hydrodynamic radius is determined using conductivity measurements and can nowadays easily to be found in Handbooks. See also the end of **Chapter 3**.

small particles, whereas Mie scattering occurs for particles larger than the wavelength) :
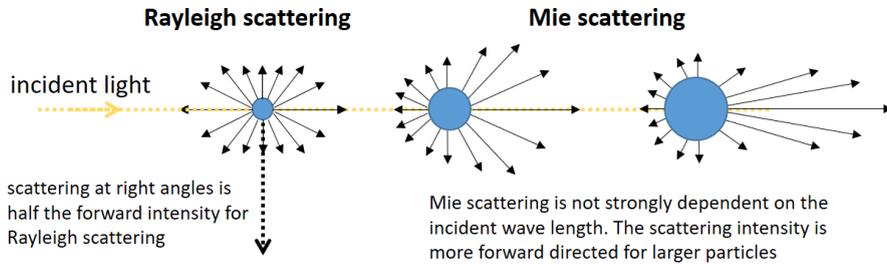
In the Rayleigh regime, the intensity of the scattered light is proportional to

$$I \sim \frac{1 + \cos^2(\theta)}{\lambda^4}$$

where θ is the angle between the incident beam and the observer and λ is the wavelength of the incident beam. It can be verified that

$$\frac{I(\theta = 0)}{I\left(\theta = \frac{\pi}{2}\right)} = 2$$

which is illustrated on the figure above: the scattering at right angles is half the forward intensity. The dependence of the intensity of the fourth power of the wave length leads to the fact that short wave lengths are more scattered than long wave lengths. As already discussed in **Chapter 1**, this is the reason why the blue colour (short wavelength) is more scattered than the red (long wavelength) for small particles.

Fick's laws

Fick's first law

The diffusion coefficient D expressed above is very important as it is used in fundamental relations, such as, for example Fick's law of diffusion (Fick's first law):

$$\boldsymbol{J} = -D\frac{\partial n}{\partial x}\boldsymbol{e_x}$$

where $\boldsymbol{J}$ is the diffusion flux, n is the concentration of particles (ex: colloidal clay particles) and $\boldsymbol{e_x}$ the unit vector in the x direction. The units of $\boldsymbol{J}$ are in number of particles per square meters per second. The law has here been written for a 1d

diffusion, it is easy to generalize to more dimensions (using the general relation $\boldsymbol{J} = -D\boldsymbol{\nabla}n$).

Fick's law expresses the fact that given a gradient of concentration, the colloidal particles will move towards the region of low particle concentration (hence the minus sign). For example, if a suspension of colloidal particles (or an electrolyte solution) is pipetted into a jar of water the particles will diffuse into the whole jar, and after some time no gradient in concentration will be observed anymore: the particle concentration in the jar will be everywhere the same:
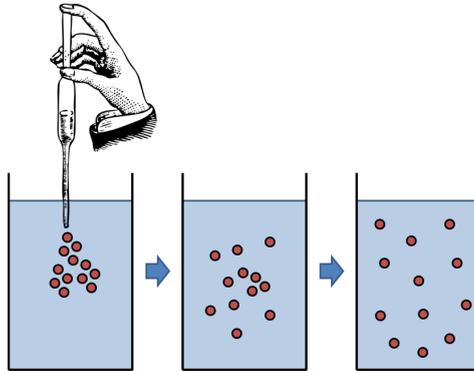
Colloidal particles obeying Fick's law and diffusing to regions of low particle concentration. When the particle concentration is everywhere the same, the particles will still move due to Brownian motion, but the particle concentration will not change anymore and remain uniform in the jar. (the action of gravity is neglected in this example)

Fick's second law

Conversation of matter dictates that for a given volume, during a small time dt, the difference between the fluxes of particles entering and going out of this volume must be equal to the variation of concentration of the same particles. In mathematical notations this implies that:

$$[J(x) - J(x + dx)] = \frac{\partial n}{\partial t}\,dx$$

and therefore:

$$-\frac{\partial J}{\partial x} = \frac{\partial n}{\partial t}$$

Using Fick's first law, one gets Fick's second law:

$$D\frac{\partial^2 n}{\partial x^2} = \frac{\partial n}{\partial t}$$

Fick's laws – and similar type of relations - are found in many branches of science, where diffusion plays a role (these equations can also be used to describe the transport of heat or momentum).

Fick's second law enables to give a good example of the connection between Brownian motion and diffusion processes. We consider the case where the diffusion of particles can be schematically visualized as follows:

The initial conditions of the experiment are that all the particles are located at $x = 0$ in an infinitely thin layer Δl. We have: $n(0,0) = n_0\Delta l$. To account for the fact that the layer is infinitely small, one expresses usually that: $n(x, 0) = n_0\Delta l\delta(x)$ where $\delta(x)$ is called the **Dirac distribution** which is a function such $\delta(x = 0) = 1$ and $\delta(x \neq 0) = 0$. Using this Dirac function insures that one has:

$$\int_0^\infty n(x,t)\, dx = n_0\Delta l$$

for any time t (conservation of mass principle). One can verify that the solution of Fick's second law, in the simple case considered here (i.e. the diffusion is only along x) is:

$$n(x,t) = \frac{n_0\Delta l}{\sqrt{\pi Dt}}\exp\left(\frac{-x^2}{4Dt}\right)$$

The characteristic timescale associated to the diffusion is:

$$\tau = \frac{x^2}{4D}$$

Note that the characteristic time τ can be linked to the measure of how far the substance has spread in a given time $t = \tau$:

$$\langle X^2(t) \rangle = 2dDt$$

where $d = 1$ as we have here a linear movement (dimension 1). This is the expression we have seen to be related to Brownian motion.

The functions plotted underneath were obtained numerically from the expression of $n(x, t)$ given above, and each function was normalized by dividing it by $n(x = 0, t)$:
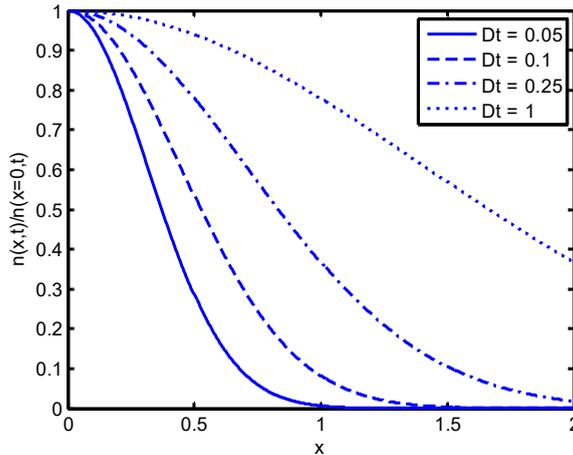
At large times ($t \gg \tau$), one will obtain: $n/n(x = 0, t) = 1$ as there will be no gradient in concentration anymore.

Important: It should be emphasized that the diffusion that has been discussed here is not the one defined in hydrodynamics: in that case, the diffusion term, still expressed as $D\frac{\partial c}{\partial x}$ as in Fick's first law, refers to a diffusion coefficient D that is orders of magnitude larger than the diffusion coefficient discussed above. In colloid science, the diffusion coefficient is associated to the thermal energy $k_B T$ whereas in hydrodynamics the diffusion is related to the water flow (usually in the context of turbulent mixing). The typical diffusion coefficient in colloidal science is of the order of 10^{-9} m^2/s whereas it is 10^{-1} m^2/s in hydrodynamics: in short, it is more efficient to stir a cup of coffee than to wait for the sugar molecules to diffuse according to Brownian motion!

This does not imply that Brownian motion should be neglected when dealing with turbulent mixing: it plays an important role for predicting the flocculation of colloids.

Application of Fick's law : cleaning colloidal suspensions

In order to study colloidal suspensions, it is often required to "clean" them, i.e. remove as much dissolved salt or small dissolved molecules as possible from the suspension. Doing so, it is then possible to perform accurate measurements as function of a chosen added salt for example and avoid that the background electrolyte concentration contaminates the measurements.

A well-established technique to clean suspensions is to perform a **dialysis**. The suspension is put in a flexible tube (Visking tube), made of a membrane that only allow the passage of small molecules. The principle of the membrane is based on diffusion: the tube containing the suspension is placed in a jar containing ultra-pure water (of extremely low conductivity). By diffusion the small molecules and ions will then tend to pass the membrane of the tube and invade the ultra-pure water. As the colloidal particles are too big, they cannot pass the membrane and remain in the tube. After an equilibrium state is reached, the tube is removed and the contaminated water is replaced by new ultra-pure water. This procedure is repeated several times, until the conductivity of the water at equilibrium is close to the one of pure water.

dialyzing tubing containing a colloid (brown spheres) and ions (blue spheres). The tubing allows only ions to exchange with the surrounding water

magnetic stirrer (the mixing facilitates the ionic migration through the tubing)
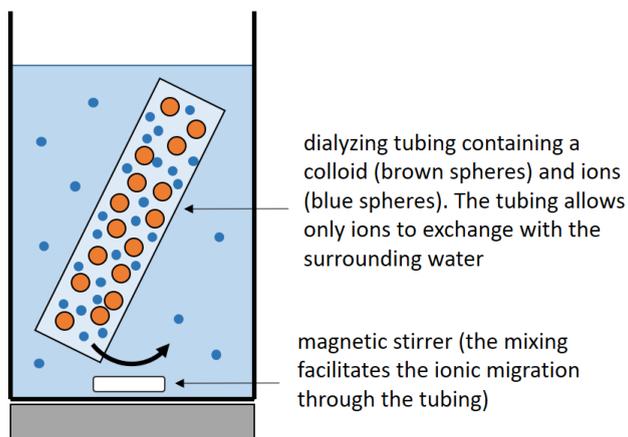
Dialysis of a colloidal suspension

The cleaning can also be accelerated by using **electrodialysis**, where the application of an electric field speeds up the migration of ions:
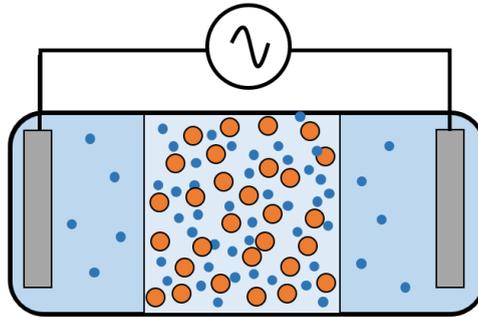
The ions move by **electrophoresis** to the oppositely charged electrodes, a phenomenon that will be discussed in Chapter 3.

Stable and unstable colloidal suspensions

An important condition to keep a colloidal particle suspended is that it cannot "glue" to another particle (and another, and another…) to form a larger entity that might eventually settle under its own weight. One says of such isotropic suspensions in which particles do not glue or settle that they are "stable". In the next chapter, we are going to review in detail the mechanisms leading to unstable suspensions. For the moment, we simply note that stable suspensions are composed of:

electrostatic attraction/repulsion
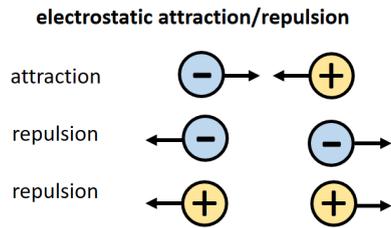
> ➢ particles with a surface charge of same sign, since these particles are repelling each other (whereas particles with opposite charges are attracting each other)

> ➢ particles in special mixtures (defined in the **Chapter 4**), or particles coated with a special type of molecules can also create stable suspensions, irrespective of the surface charge of the particles. (The mechanisms leading to stability are then not linked to surface charge).

In order to be stable, the particles should also be able to overcome gravity. This can be done when the particles are colloidal by Brownian motion, or, in special cases by electrostatic forces (discussed below). We will distinguish four different types of suspensions:

(A) Stable dilute suspension (B) Stable concentrated suspension (C) Unstable suspension (D) Unstable suspension
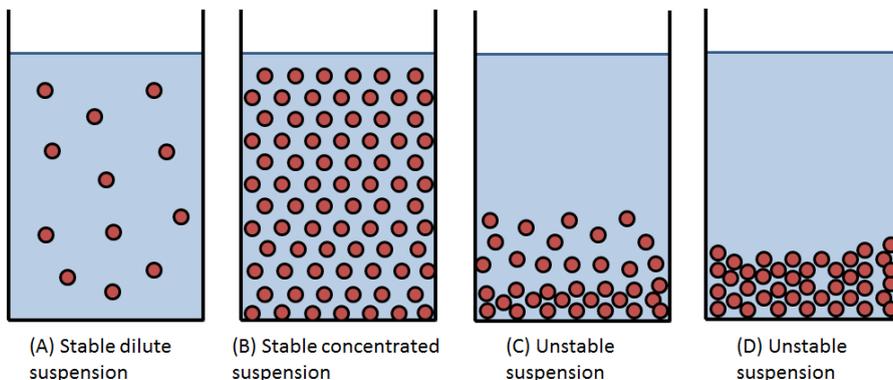
Stable suspensions: case (A) and case (B)

A stable isotropic dilute suspension of particles can be achieved when the particles are undergoing Brownian motion. Gravity then usually plays a negligible role (this is the case for a small jar or settling column – for some meters high suspended sediment, on the other hand, a gradient in concentration can appear). The particles should also repel each other, otherwise they would glue at some point , form a larger entity and settle. Ions in water, even though ions are usually smaller than the typical colloid size[20], satisfy all these conditions. Small colloidal particles, with a significant surface charge, also satisfy these conditions. A typical example is the colloidal gold suspension (see picture on the right) prepared by Faraday (1791-1867) that have remained stable over more than a century.

The difference between case (A) and case (B) lays in the fact that because of the high concentration the particles are " trapped" in their positions in case (B) and do not " walk around" easily (thanks to Brownian motion) as in case (A). An important consequence of this trapping is that the relative distance between the particles is rather constant, similar to what one has in crystals. This is why this type of suspension is called a **liquid crystal**. A lot of work in the 1960-70s has been done on colloidal crystals, as they were used as models for atomic systems that cannot easily been investigated. The research continues on this type of systems, especially on non-spherical colloids that can arrange themselves in very different ways, depending on parameters such as temperature, pressure, and concentration:

---

[20] Some ions can be of colloidal size: an example are ions called polyoxometalate (abbreviated POM), composed of metal and oxide; the size of these ions can reach 10 nm – the lower end of what is called a colloidal particle.
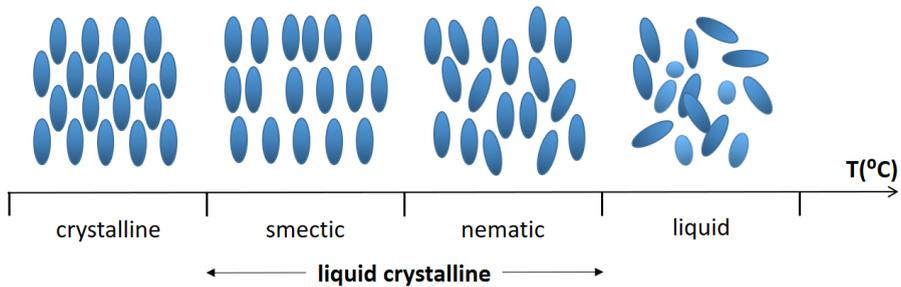
Liquid crystals made a great break-through in the 1970s as they became commercially available and were used in digital calculators as Liquid-Crystal Displays (LCDs). The principle of a LCD is similar to the one used in research to investigate the birefringence of complex suspensions. In birefringence studies, an electric field is applied to a monodisperse suspension, such that all particles will align with the electric field thanks to electric dipole-dipole interactions. When the electric field is switched off, the particles will relax because of their Brownian motion. From evaluating the time needed for the suspension to relax completely (i.e. become fully isotropic again), information about the particle's size can be deduced. The observation is done by analysing the amount of light going through a photodetector as function of time. The time for the particles to relax is extremely fast and the transition cannot be observed by the eye, which implies that for LCDs what one sees is a pixel switching directly from black to white (and vice-versa):
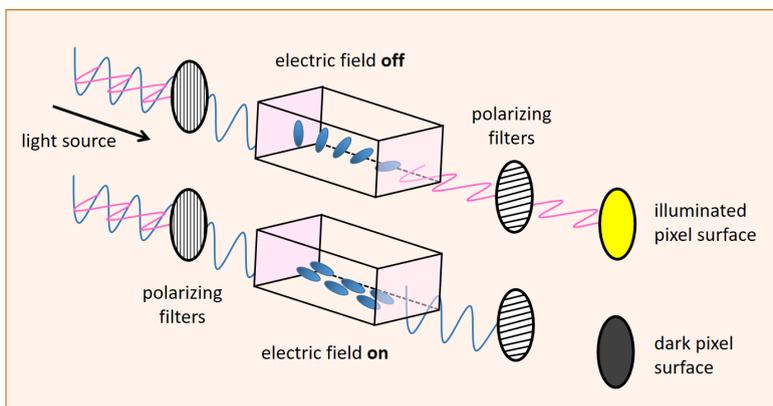
Liquid crystal principle: a cholesteric nematic liquid crystal is placed between two transparent electrodes (in pink). Cholesteric liquid crystals have the property that they organize themselves in layers and that each layer has a director axis, represented here by the direction of a blue rod. The distance between electrodes is tuned in such a way that the light exiting the cell can pass through the filter and illuminate a pixel. When an electric field is applied, the ordering of the crystal changes and no light can pass through the filter.

Liquid crystals made of clay minerals are an on-going topic of research[21]. The complexity of the clay mineral structures (see **Chapter 1**), explains why for years researchers have privileged simpler colloidal particles than clays. However, the understanding of the transition of one clay liquid crystal phase to another will enable a better predictability of the rheology of these liquids, and their consolidation behaviour. The applications in Civil Engineering are enormous: drilling fluids can be seen as examples of clay liquid crystals (they are a mixture of bentonite clays and polymers[22]), and so can cohesive sediments (a mixture of clay, silt and organic matter), found in many delta regions, rivers, river shores and lakes.

Unstable suspensions: case (C) and Case (D)

When the colloidal particles are noticeably sensitive to gravity in the jar or settling column, they will settle. We still only consider the case that the particles are not gluing to each other, meaning that they just settle because of their own weight.

Mud pool after a rainfall: the colloidal particles are in suspension and will slowly settle down in time.

Now an interesting question arises:

Let's take silica particles as an example. We consider two types of particles: small (colloidal) silica particles and large (sand) silica particles. The only difference between the particles is their size, which means they have the same surface charge (in C/m^2) and the same density. We suppose that the particles are highly charged. The question is: is it possible to define a size of particles such that the particles are so repellent that they remain in suspension like in case (B)? Intuitively, we know that this is not possible for sand particles: sand particles will always settle.

The answer lays in the comparison of the forces exerted on the particles. The dominant forces are gravity and the electric repulsion. We have already seen that the gravity force scales with a^3 (the radius of the particle to the power 3). The electric force is linked to the surface charge of the particles. This force is thus proportional to a^2. The ratio of these forces is therefore proportional to the radius of the particles a. Of course, there are more factors in this ratio besides the size but the general idea is that for a large particle (large a) *volume* forces will always be large than *surface*

---

[21] van der Beek, David, and Henk NW Lekkerkerker. "Liquid crystal phases of charged colloidal platelets." Langmuir 20.20 (2004): 8582-8586.

[22] Polymers will be defined in the next chapters. They are a special type of colloidal particles.

forces. A **criterion**[23] for determining if a particle can or cannot settle is therefore its size: the typical particle size for which the two forces balance is around 1 µm, i.e. the size of colloidal particles.

In Case (C), we have displayed the case where, even though the particles settle under their own weight, there is a small region above the bed where particles are not touching the bed, because of the strong repulsion between the particles and the gradient in concentration that tends to push them vertically up. This phenomenon has been studied in 2003 in an article called "Defying gravity with entropy and electrostatics: sedimentation of charged colloids"[24]. A particular case of this general theory is when only gravity and the gradient in concentration plays the dominant role (the repulsion between particles just prevent them from gluing). Supposing that the particles do not hinder each other, they all settle according to Stokes' settling velocity:

$$v = \frac{2}{9} a^2 \frac{\rho_p - \rho_w}{\eta} g$$

The corresponding flux of particle is given by $J_g = Cv$ where C is the concentration of particles (in mol/L or number/m^3 depending on one's choice). By settling, a concentration gradient will establish, and this will lead to a flux, according to Fick's first law (we take z as the coordinate along the vertical axis such that $\boldsymbol{g} = -g\boldsymbol{e}_z$):

$$J_C = -\frac{k_B T}{6\pi\eta a} \frac{dn}{dz}$$

At equilibrium (steady state), the two fluxes must be equal: $J_C = J_g$ from which we obtain:

$$\frac{dn}{n} = -\frac{4}{3}\pi a^3 \left(\frac{\rho_p - \rho_w}{k_B T}\right) g\, dz$$

Note that this equation is mathematically similar to the one obtained for a **Rouse** profile, where the balance is between the downwards settling (due to gravity, like here) and the upward turbulent diffusion, instead of the thermal diffusion. We have already stressed above that the two diffusion coefficients are extremely different in magnitude, implying that a Rouse profile can easily be observed in-situ in the case of clays, and the one derived here generally not.

The equation can be solved using $n(z = 0) = n_0$ and leads to:

---

[23] It is of course not the only criterion, surface charge and the ability to undergo Brownian motion are others.

[24] van Roij, R. (2003). Journal of Physics: Condensed Matter, 15(48), S3569.

$$n(z) = n_0 \exp\left(-\frac{4}{3}\pi a^3 \left(\frac{\rho_p - \rho_w}{k_B T}\right) gz\right)$$

A case where this equation leads to an easily observable profile is the case of gas. Assuming an ideal gas, the relation between particle concentration and pressure is:

$$P = nk_B T$$

where P is the pressure and n the number of gas particles per unit of volume. Furthermore, defining the volume of one gas molecule as $V_p = \frac{4}{3}\pi a^3$ and realizing that there is no Archimedes force in this case, one gets:

$$P(z) = P_0 \exp\left(-\frac{V_p \rho_p}{k_B T} gz\right)$$

The mass of one gas particle $m_p = V_p \rho_p$ can be substituted by using the molar mass M of the gas, and one obtains:

$$P(z) = P_0 \exp\left(-\frac{Mg}{RT} z\right)$$

where $R = N_A k_B = 8.314 \, \text{J} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$ is called the **gas constant**. This expression for the pressure with altitude is called the **barometric formula**. It is valid in the limit of ideal gases, for temperatures that do not vary with height. For air ($M = 0.02896$ kg/mol) at T = 288 K (15 °C), one gets that $P(z) \approx 0$ for $z = 8.4$ km − clearly an observable length.

Another case where this equation leads to an observable length is for particles with a density close to the one of water for which $(\rho_p - \rho_w)$ becomes very small. This case is illustrated on the right-hand-side where latex colloids are dispersed in ultra-pure water. The latex particles were deposited gently with a pipette on the bottom of the tube, see picture (t = 0) and left unstirred. The picture (t = 40h) was taken 40 hours after the sample was prepared. A fuzzy region of a few cm high can be observed above the bed.

(t = 0)   (t = 40h)
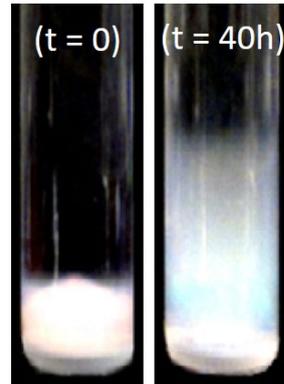
In Case (D), the particles are all settled on the bed. The structure of the bed, and in particular its porosity, will be extremely dependent on the way the particles have settled down: did they settle rapidly and just stick to the bed? Did they have time to rearrange their position relative to the particles already settled? Did the particles aggregate (flocculate) before they reached the bed? The evolution of a bed formed by settled particles will be discussed in **Chapters 9 and 10**.

Osmotic pressure

When a suspension is brought into contact with a semi-permeable membrane, another effect takes place to ensure that the suspension reaches a stable state. This effect is linked to a transport of water (not of colloidal particles) through the membrane, and is driven by what is called osmotic pressure.

We have derived above the barometric pressure for gases where we used the ideal gas equation of state:

$$P = nk_BT$$

where n is the number of gas atoms per unit of volume. An analogue of this relation, in the context of colloids, is the **osmotic pressure** of the suspension, defined for a dilute suspension as:

$$\Pi = nk_BT$$

where n is the number of colloidal particles per unit of volume. This equation is demonstrated in **Chapter 8**. Thus, a dilute suspension of colloidal particles behave thermodynamically like a collection of "giant atoms": they have the same equation of state and show the same sedimentation equilibrium as a classical ideal atomic gas (even though not on the same length scale as discussed above).

The osmotic pressure can be seen as the pressure that must be applied to a suspension in contact with a bath of pure solvent across a semi-permeable membrane (which the particles cannot cross) in order to stop the flow of solvent from the bath to the suspension. It can therefore easily be measured:

initial state   at equilibrium   at equilibrium with
                                                applied pressure
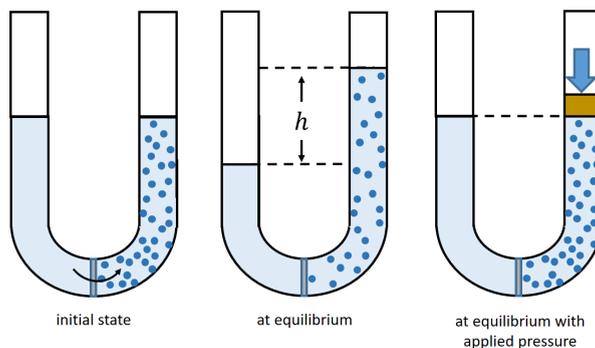
At the initial state, a suspension is set in contact with a semipermeable membrane that only allows the flow of water. In time, water will flow to the compartment with the suspension. At equilibrium, the osmotic pressure can be determined from the height difference between the water and the suspension or by applying a counter-pressure so as to have the same fluid levels.

The relation between height and osmotic pressure is given by:

$$\Pi = \rho g h$$

where ρ is the density of the solvent (water) and the measured height difference at equilibrium.

The concept of osmotic pressure plays an important role in the stability of suspensions. This will be discussed further in the **Chapters 3 and 4**.

Applications

Biology

An important domain of application for osmotic pressure is biology and medicine. As cells in biological bodies can be seen as fluids encapsulated in semi-permeable membranes, the properties of the solutions in which they are found determine their structure: when the composition of the cell's fluid and the suspending solution are such that water flows out of the cell, the cell will shrink, and it will swell when water will flow into the cell. Three situations are defined:

Hypertonicity: there is a greater concentration of solutes outside the cell than inside: water will get out the cell. Some organisms have evolved intricate methods of circumventing hypertonicity. For example, saltwater is hypertonic to the fish that live in it. They need a large surface area in their gills in contact with seawater for gas exchange, thus they lose water osmotically to the sea from gill cells. They respond to the loss by drinking large amounts of saltwater, and actively excreting the excess salt. This process is called osmoregulation.

Isotonicity: there is the same concentration of solutes outside and inside the cell. In this case the cell neither swells nor shrinks. Water molecules diffuse through the plasma membrane in both directions, and as the rate of water diffusion is the same in each direction that cell will neither gain nor lose water. Isotonic sport drinks contain similar concentrations of salt and sugar as in the human body. These drinks are marketed as soft drinks (they contain approximately 15 grams of sugar per 250 ml). They are supposed to help athletes replace "water,
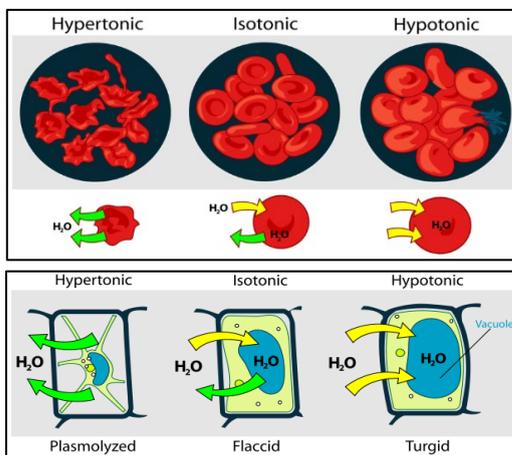
electrolytes, and energy" after training or competition but their efficacy is still not proven.

***Hypotonicity*:** there is a greater concentration of solutes inside the cell than outside: water will get in the cell. In the case of plants, the cell walls are rigid enough to contain the internal osmotic pressure and limit the cell expansion (for animal cells, like blood cells, the cells will eventually burst). It is hypotonicity that enables herbaceous plants to stand upright.

Water purification

Osmotic pressure is the basis of filtering ("reverse osmosis"), a process commonly used in water purification. The "reverse osmosis" compares to the case where a counter pressure is applied to equilibrate the liquid heights in the tube, as depicted above. The water to be purified is placed in a chamber and put under an amount of pressure greater than the osmotic pressure exerted by the water and the solutes dissolved in it. Part of the chamber opens to a differentially permeable membrane that lets water molecules through, but not the solute particles. The osmotic pressure of ocean water is about 27 atm (1 atm = 10^5 Pa). Reverse osmosis desalinates fresh water from ocean salt water.

Illustrations

Robert Brown (public domain)
https://en.wikipedia.org/wiki/Robert_Brown_(botanist,_born_1773)

Faraday gold
https://www.rigb.org/

Tonicity (creative commons license)
https://en.wikipedia.org/wiki/Tonicity
https://commons.wikimedia.org/wiki/File:Turgor_pressure_on_plant_cells_diagram.svg

Chapter 3

DLVO forces

Given certain conditions, colloidal particles can "glue[25]" (**aggregate**) to each other, hereby creating a bigger entity (an **aggregate** or **floc**) that may be able to settle down. Its settling velocity is then different from the settling velocity of the individual particles that created it. This, in turn, has important consequences in terms of sediment transport as Stokes' settling velocity is for example one of the parameters used in large numerical codes that predicts large scale sediment transport.
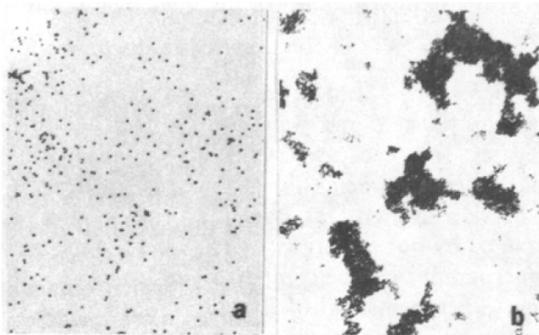
(a): stable suspension with non-aggregated (primary) particles; (b): unstable suspension with large flocs consisting of aggregated (flocculated) particles. In (a) the solvent properties are different than in (b), but the primary particles are the same in (a) as in (b). The flocs in (b) are made by aggregation of primary particles. The primary particles in (a) have a smaller settling velocity than the flocs in (b)[26]

In the present chapter, the conditions leading to the (non) aggregation of particles are reviewed. These conditions originates from electrical interactions between particles. Some of these interactions occur between molecules and atoms constituting each particle. The sum of all the interactions between these atoms and molecules leads to forces between particles. This is detailed below in the section "Van der Waals forces". The van der Waals forces are usually attractive[27]. Other forces ("Coulombic forces") originate from the interaction between the surface charges of the particles. When the particles have surface charges of same sign, the forces are repulsive. It is the relative strength between the Coulombic and the van der Waals forces that determines the stability of a suspension: if van der Waals dominates, the particles will aggregate and the suspension is said unstable as its composition changes over time. If the repulsive Coulombic forces dominate, the particles will stay away from each other and the suspension is said to be stable as its composition will not change over time. This is provided of course that gravity does

---

[25] The word "colloid" comes from the greek word *glue*.
[26] From Shih et al., Aggregation of Colloidal Particles with a Finite Interparticle Attraction Energy; Journal of Statistical Physics, VoL 62, Nos. 5/6, 1991
[27] The van der Waals force between two identical bodies in a medium is always attractive, while that different bodies in a medium can be attractive or repulsive.

not play a significant role, in which case the particles will settle and the suspension is again said unstable as its composition is changing over time.

Van der Waals forces

Microscopic observations of colloidal particles have led scientists in the 19th century to discover that colloidal particles have the tendency to form persistent aggregates through collisions induced by Brownian motion. This led them to conclude that there should be attractive forces between the particles. It is the Dutch theoretical scientist and thermodynamist **Johannes Diderik van der Waals** (1837-1923) who was the first to quantify these forces that nowadays carry his name.
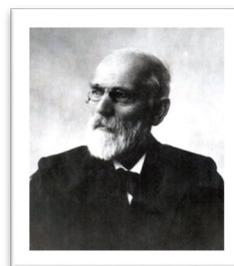
Van der Waals is famous for another discovery: the van der Waals equation of state that describes the behaviour of gases and their condensation to the liquid phase. This equation was published in his PhD thesis entitled *Over de continuiteit van den gas- en vloeistoftoestand* (On the continuity of the gaseous and liquid state). The original PhD thesis can still be found in the library of the University of Leiden where he studied. He got the Nobel prize in Physics in 1910.

The van der Waals forces originate from weak electrical interactions between atoms and molecules. These forces have 3 origins, and are all related to **electric dipole moments**.

Electric dipole moments are created by the fact that two (or more) charges of opposite sign are in the close vicinity of each other. Water molecules, for instance, have dipole moments as they have a slight positive charge on H-sides and a slight negative charge on the O-side (even though the molecule H_2O is not charged):
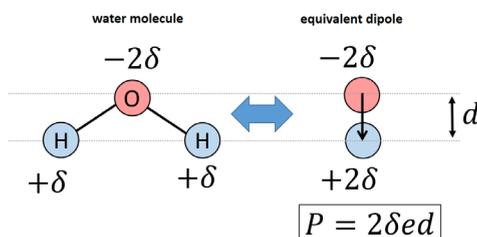

The water molecule, H_2O, is uncharged but possesses a dipolar moment P due to the asymmetry of the shared electrons' distribution in the covalent OH bonds. This results in the oxygen atom having a slight negative charge $-2\delta e$ (where e is the electron charge) whereas each hydrogen atom has a slight positive charge $+\delta e$.

The dipole moment P of two charges ($+q$ and $-q$) separated by a distance d is given by

$$\boldsymbol{P} = q\boldsymbol{d}$$

The bold notation indicates vectors, and $\boldsymbol{d}$ is oriented from $-q$ to $+q$. In the case of water, illustrated above, we first find the location of the barycentre of the positive charges which leads to $q = 2\delta e$ (illustrated above). The dipole moment P can then be evaluated. For water one finds it is equal to 1.85 D (1 D = 3.335 64 × 10^{-30} C m).

Dipole-dipole interactions can be either attractive or repulsive:

attraction   attraction   repulsion   repulsion
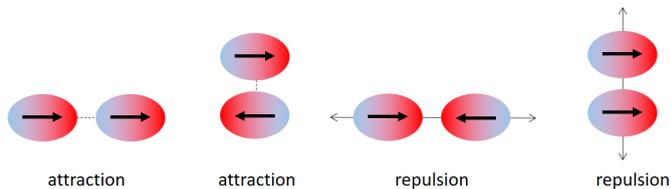
It is due to this polar nature of water that water molecules make **hydrogen bonds**, which are defined in **Chapter 4**.

The three forces that form the van der Waals forces are due to:

1 – the positive (or negative) interaction between permanent dipoles of the atoms/molecules. These forces are also called **Keesom** forces, in honour of the Dutch physicist who derived them mathematically.
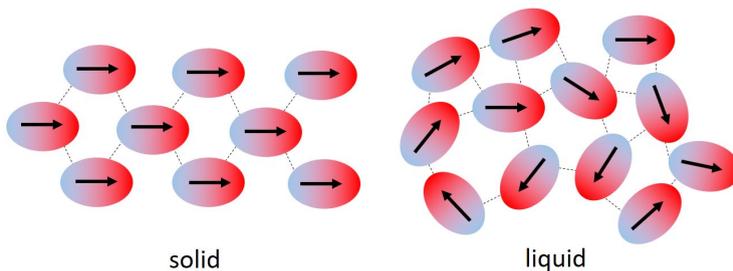
solid   liquid

Schematic representation of dipole-dipole interactions in a solid and a liquid

2 – the positive attraction between a permanent dipole and an induced dipole. An induced dipole is created when there is a temporary (extremely brief) deformation of the electronic atmosphere around an atom: as the nucleus of an atom is positively charged and the electrons are negatively charged, a temporary dipole is created. This deformation is induced by the proximity of another dipole or an ion. These

forces are called **Debye** forces, in honour of the Dutch physicist **Peter Debye**, who made significant contributions to colloid science, and hence his name is also associated with an equation and a length scale that will be introduced later.

3 – the positive attraction between two induced dipoles. These forces are also called **London** forces, in honour of the German physicist who derived them mathematically.

The van der Waals forces were primarily derived for modelling the interaction between two molecules. However, it was soon discovered that it was possible to extend them to larger objects. This was done in particular by the Dutch scientist **Hugo Christiaan Hamaker** (1905-1993), who gave his name to the **Hamaker constant** used to describe the van der Waals forces between macroscopic objects (macroscopic in the sense that they are larger than molecules). We will only give the examples of the van der Waals forces exerted in the cases of the two following geometries:
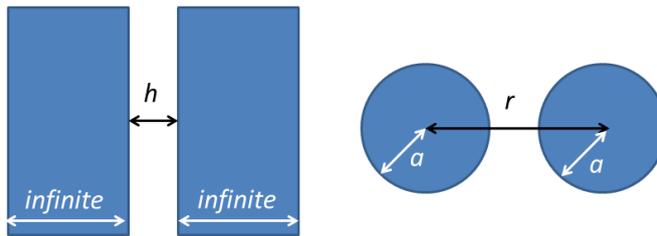
For the two semi-infinite plates separated by a distance *h*, the van der Waals potential per area of interface is given by:

$$\frac{\Phi}{area} = \frac{-A}{12\pi h^2}$$

where A is the Hamaker constant. For equal spheres of radius *a* at center-to-center separation *r* it is given by:

$$\Phi = \frac{-A}{6}\left(\frac{2a^2}{r^2 - 4a^2} + \frac{2a^2}{r^2} + ln\left(\frac{r^2 - 4a^2}{r^2}\right)\right)$$

The van der Waals force F_{vdw} can be estimated using the relation:

$$F_{vdw}(r) = -\frac{d\Phi}{dr}$$

The van der Waals interaction potentials can be calculated from the adding-up all the inter-atomic dipole contributions and the Hamaker constants are then obtained

by identification. These Hamaker constants are then compared with the ones found experimentally, by performing flocculation experiments (see **Chapter 5**). When the system is too complex, the Hamaker constant can only be determined experimentally. The Hamaker constants are often noted A_{123} symbolizing the fact that it is the Hamaker constant for media 1 and 3 interacting across medium 2. For instance A_{121} could represent a clay particle (medium 1) interacting with another clay particle (also medium 1) across water (medium 2). For typical clays, *across water* one finds[28]:

kaolinite	$A_{121} = 3.1 \cdot 10^{-20}$ J
illite	$A_{121} = 2.5 \cdot 10^{-20}$ J
montmorrillonite	$A_{121} = 2.2 \cdot 10^{-20}$ J

To give an idea, Hamaker constants are about 10^{-19} J for interactions *across vacuum*. Typical values are in between 10^{-19}-10^{-21} J. These values can decrease slightly with increasing salt concentration.

Peter Joseph William Debije (1884 - 1966) studied in the Aachen University of Technology under the supervision of the theoretical physicist Arnold Sommerfeld, who later claimed that his most important discovery was Peter Debye. Debije' s name is usually written **Debye** as the digraph of the letter i and j (ij is considered a letter in itself) is only known to the Dutch language.

Debye made major contributions to the field of physics and physical chemistry (including colloid science). He applied the concept of dipole moment to the charge distribution in asymmetric molecules in 1912 and developed equations relating dipole moments to temperature and dielectric constant. The units of dipole moments are termed Debye (D) in his honour (1 D = 3.335 64 × 10^{-30} C m). His name is also associated with the Debye frequency, the frequency above which an electric double layer does not polarize anymore and the Debye length that we are going to define in the next section. In 1923, together with his assistant **Erich Hückel**, he developed an improvement of Svante Arrhenius' theory of electrical conductivity in electrolyte solutions. This work resulted in the **Debye-Hückel equation**. In 1936 he got the Nobel prize in chemistry "for his contributions to our knowledge of molecular structure through his investigations on dipole moments and on the diffraction of X-rays and electrons in gases".

---

[28] Novich, B. E., & Ring, T. A. (1984). Colloid stability of clays using photon correlation spectroscopy. Clays Clay Miner., 32(5), 400. [Note that photon correlation spectroscopy = DLS, see previous chapter]

Coulombic forces

The first given reason for the stability of suspensions is associated to surface charge of the particles. These charges lead to **Coulombic** forces that are repulsive when two particles have charges of same sign (and are attractive otherwise).

In order to calculate these Coulomb forces, one needs first to know the electric potential around one particle, without any interaction with another.

Electric potential distribution around a charged sphere in an electrolyte solution
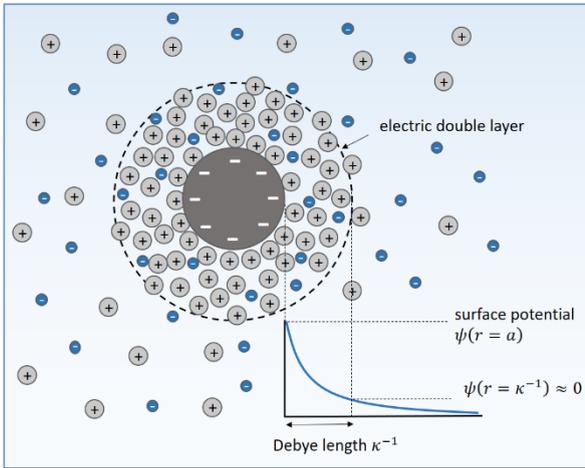
When a spherical particle of radius a, say negatively charged, is in an electrolyte solution, there will be a so-called **electric double layer** around the particle. This layer is composed of the (fixed) surface charges (the first layer) and its surrounding cloud of ions (the second layer). This cloud is mainly composed of counter-ions, i.e. ions with charge + (but not only: there are also co-ions, i.e. ions with charges -). The distribution of ions around the particle is given by the **Boltzmann distribution**:

$$n_i(r) = n_i(\infty)\exp\left(\frac{-q_i\psi(r)}{k_BT}\right)$$

where $n_i(r)$ is the concentration of ions i$(= +, -)$ in number / m^3 as function of the distance from the centre of the sphere. The other parameters are: q_i the electric charge of ion i and $\psi(r)$ is the electric potential around the sphere. The electric potential $\psi(a)$ is the potential at the surface of the particle. Note the presence of the thermal energy k_BT in the expression: it is linked to the Brownian motion of the ions in the cloud.

The expression for $\psi(r)$ is found using the **Poisson** equation, derived from electrostatics, given here in spherical coordinates:

$$\Delta\psi = \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi(r)}{dr}\right) = \frac{-1}{\varepsilon_0\varepsilon_r}(q_+n_+ + q_-n_-)$$

By combining the Poisson and Boltzmann equations one obtains the non-linear **Poisson-Boltzmann** equation for $\psi(r)$. Using some mathematics, it is then possible to derive the electric potential around one particle. Many expressions for this potential are available, depending on the approximations made for the mathematical derivations. The exact solution of the Poisson-Boltzmann can also be obtained numerically[29].

At low surface potentials: At low surface potential $\psi(a)$ it is possible to find a simple analytical solution by using the fact that $exp(x) = 1 + x$ when x is small. One then obtains:

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi(r)}{dr}\right) = \kappa^2\psi(r)$$

which can be solved into, using the convention $\psi(\infty) = 0$:

$$\psi(r) = \psi(a)\frac{a}{r}\exp\bigl(-\kappa(r-a)\bigr)$$

where the parameter κ^{-1} is called the **Debye length** (or **double layer thickness**) and can be evaluated from:

$$\kappa^2 = \frac{\sum n_i(\infty)q_i{}^2}{\varepsilon_0\varepsilon_r k_B T}$$

where $\varepsilon_0\varepsilon_r$ is the permittivity of the medium ($\varepsilon_0 = 8.85 \times 10^{-12}\,\text{F/m}$ is the permittivity of vacuum and $\varepsilon_r = 80$ the relative permittivity of water). The electric potential around a charged particle decays therefore to zero over a typical length equal to the Debye length. Beyond the double layer the ionic concentration is equal to the bulk concentration $n_i(\infty)$. The double layer thickness does not depend on the size of the particle, it is only depending on the salt concentration. Therefore, for a given salt concentration, κ^{-1} is the same for a 1 nm or a 1 mm particle.

At high ionic strength: When the ionic strength is large (a few tenth of mM to give an order of idea), one usually satisfy the condition $\kappa a \gg 1$, independently of the size of a. For a sphere with a thin double layer, i.e. for which $\kappa a \gg 1$, one can show that the electric potential around that sphere, for any surface potential $\psi(a)$, is given by:

---

[29] Chassagne, C., and D. Bedeaux. "The dielectric response of a colloidal spheroid." *Journal of colloid and interface science* 326.1 (2008): 240-253. See also Chassagne, C. "Dielectric response of a charged prolate spheroid in an electrolyte solution." International Journal of Thermophysics 34.7 (2013): 1239-1254.

$$\psi(r) = \frac{4k_BT}{q}\tanh\left(\frac{q\psi(a)}{4k_BT}\right)\exp\big(-\kappa(r-a)\big)$$

where the ion charge is usually given as $q = ze$ where z is the valence of the ion, and e the elementary particle charge ($e = 1.6 \times 10^{-19}$ C). We will here only consider a symmetric electrolytes for which $z = z_+ = -z_-$. For monovalent salts (ex: for NaCl, KCl) one has $z = 1$. For small x one has $\tanh(x) = x$ which implies that for small $\psi(a)$ the above expression reduces to:

$$\psi(r) = \psi(a)\exp\big(-\kappa(r-a)\big)$$

We have previously found that for a spherical particle with small $\psi(a)$ (and any κa) the potential takes the form:

$$\psi(r) = \psi(a)\frac{a}{r}\exp\big(-\kappa(r-a)\big)$$

For small κa one has:

$$\frac{a}{r} = \frac{\kappa a}{\kappa a + \kappa(r-a)} \approx \frac{\kappa a}{\kappa a + 1} \approx 1$$

since the characteristic distance over which the electric potential is non-zero is $(r - a) \approx \kappa^{-1}$. This implies that we recover the previous expression, valid for small $\psi(a)$ and large κa i.e.

$$\psi(r) = \psi(a)\exp\big(-\kappa(r-a)\big)$$

This expression is in fact also valid for two planar surfaces (for any $\psi(a)$ this time and large κa). For thin double layers ($\kappa a \gg 1$) the curvature of the interfaces does not influence the profile of $\psi(r)$:

For thin double layers (the thickness of which is given by the double arrow), the electric potential as function of the distance from the (here grey) surface is the same for planar or curved surfaces
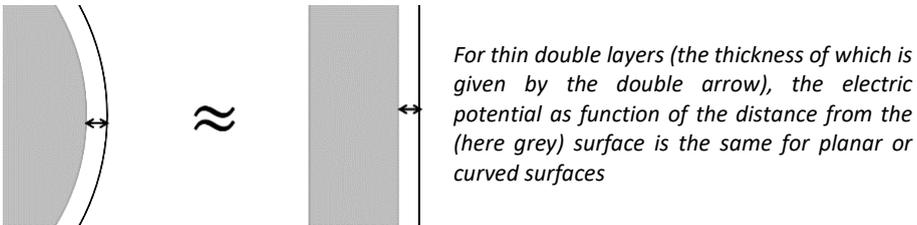
Particles approaching one another

Beyond the double layer, the ionic concentration is the same as if there would be no colloidal particles in the solution. This implies that the ion cloud screens the particle's charge: from a distance, the colloidal particle surrounded by its double layer appears uncharged (reflected by the fact that the electric potential is then zero ($\psi(\infty) = 0$). Two colloidal particles can therefore approach each other without "feeling" each other. They start to interact only when their double layers start to overlap, i.e. when their $\psi(r)$ overlap.

If the two particles would be in vacuum and therefore have no double layers, they would then simply undergo an electrostatic repulsion. It is now important to realize that in the present situation, the overlap of the double layers results in an increased ion concentration between the two particles. This implies that the local **osmotic pressure** is higher between the two particles, and it is this pressure that pushes the particles apart until their double layers are not overlapping anymore.

region of higher osmotic pressure
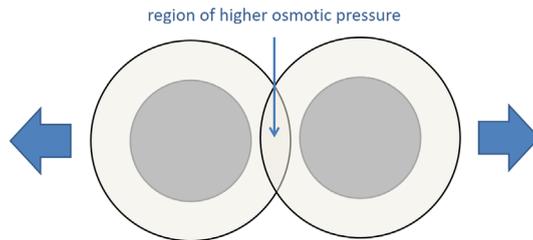
Slipping plane and Stern layer

When a particle is moving, it is possible to define a **slip(ping) plane** (also called **surface of shear**). This is the surface behind which everything moves with the same velocity as the particle. The electric potential at this plane is defined as the **zeta potential** ζ . In ideal situations $\zeta = \psi(a)$ but quite often some water and ions are tightly bound to the particle and the slipping plane is of the order of 0.1 nm away from the particle's surface: a small distance, but due to exponential behaviour of the electric potential one then finds that $\zeta \neq \psi(a)$. The region between the slip plane and the particle is called the **Stern layer** region.
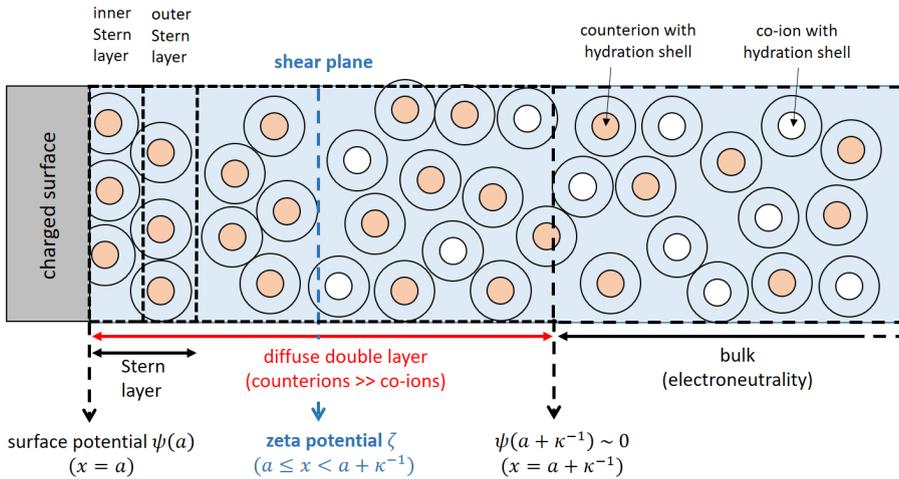
Schematic representation of the arrangement of ions close to a charged surface. Note that in some Stern layer models two regions are distinguished: the inner one where ions are strongly attached to the surface, with loss of part of their hydration shell and one where they are tightly bound but still possess their entire hydration shell. Between the charged surface and the shear plane the ions are not free to move. Beyond the shear plane it is assumed that Poisson-Boltzmann is valid.

Coulombic repulsion between spheres

The expression for the Coulomb repulsion potential between a couple of spheres with thin double layers, i.e. particles for which $\kappa a \gg 1$ is given by:

$$\Phi = 32\pi\varepsilon_0\varepsilon_r \left(\frac{k_B T}{q}\right)^2 a \tanh^2\left(\frac{q\psi(a)}{4k_B T}\right)\exp\bigl(-\kappa(r-2a)\bigr)$$

The Coulomb repulsion force F_{rep} can be estimated using the relation:

$$F_{rep}(r) = -\frac{d\Phi}{dr}$$

DLVO theory

This theory is named after the scientists Derjaguin, Landau, Verwey and Overbeek. Both Derjaguin and Landau, and independently Verwey and Overbeek proposed to explain the stability of colloidal suspensions by adding the van der Waals potential and the Coulomb repulsion potential, so as to give the interaction potential:

$$\Phi_{DLVO} = \Phi_{vdw} + \Phi_{rep}$$

One formulation for the DLVO potential between two spheres of equal size *a*, for any $\psi(a)$ and $\kappa a \gg 1$ is obtained from the expressions given above:

$$\Phi_{DLVO} = \frac{-A}{6}\left(\frac{2a^2}{r^2-4a^2} + \frac{2a^2}{r^2} + ln\left(\frac{r^2-4a^2}{r^2}\right)\right)$$
$$+ 32\pi\varepsilon_0\varepsilon_r\left(\frac{k_BT}{q}\right)^2 a \tanh^2\left(\frac{q\psi(a)}{4k_BT}\right)\exp\bigl(-\kappa(r-2a)\bigr)$$

By evaluating Φ_{DLVO} for a given suspension, it is possible to say if the suspension is stable (i.e. the particles will not aggregate) or unstable (i.e. they will aggregate): when the potential energy is below (approximately) $10\,k_BT$ then aggregation is possible, and the suspension is unstable.

Typical examples for Φ_{DLVO} are given in the following examples:
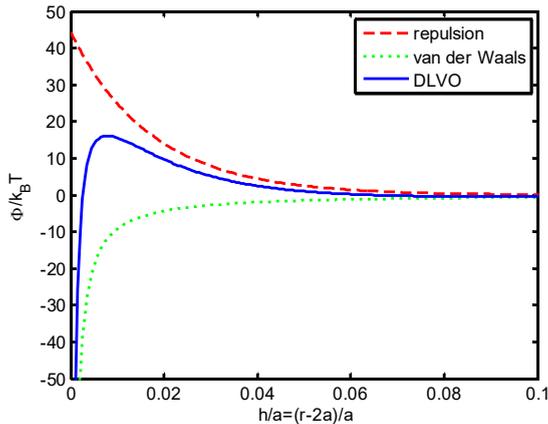
Interaction energy between two spheres of radius 250 nm, using A = 5.10^{-21} J and $\psi(a) =$ 12.5 mV. The spheres are immersed in an electrolyte solution made of KCl at a salt concentration of 5 mM.

Note that the distance at which the particles start to interact is very small: it is less than 1/50 of the radius of the particles. The Debye length in this case is less than 1/200 of the radius.
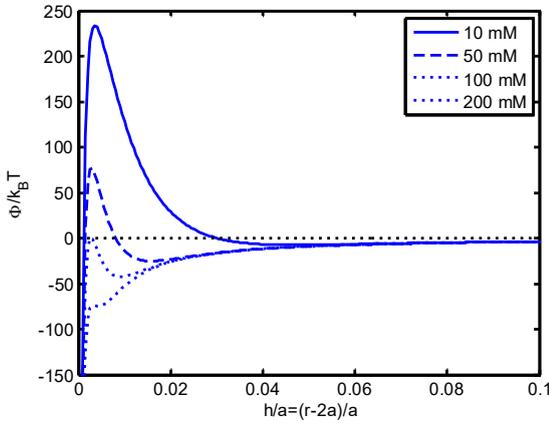
Interaction energy between two kaolinite particles of radius 370 nm, using A = 3.10^{-20} J, and $\psi(a) = 40$ mV. The spheres are immersed in an KCl electrolyte solution. The concentrations are indicated in the figure.

The values used for this example correspond to the ones given by Novich, B. E., & Ring[30]. These authors have indeed found the fastest aggregation for 200 mM of added KCl.

Between 50 and 100 mM of added KCl (where Φ_{DLVO} is of the order of some $10\ k_B T$) aggregation is still possible, but will take a longer time, as the particles have to overcome the energy barrier (the peak in the picture). This was also found experimentally by the authors.

By looking closely at the interaction potential curves, one can distinguish the following features:

At long-range separation, the particles experience a weak attraction, due to the long-range action of the van der Waals force: the van der Waals force acts on a longer range than the repulsion force, which becomes faster zero. This is best seen when plotting $|\Phi_{vdw}|/|\Phi_{rep}|$, see figure. It is clear that both for long-range (h/a > 0.06) and small-range (h/a < 0.001) the van der Waals potential is larger in magnitude than the repulsion potential ($|\Phi_{vdw}| > |\Phi_{rep}|$). At mid-distance, the repulsion potential dominates ($|\Phi_{vdw}| < |\Phi_{rep}|$).

---
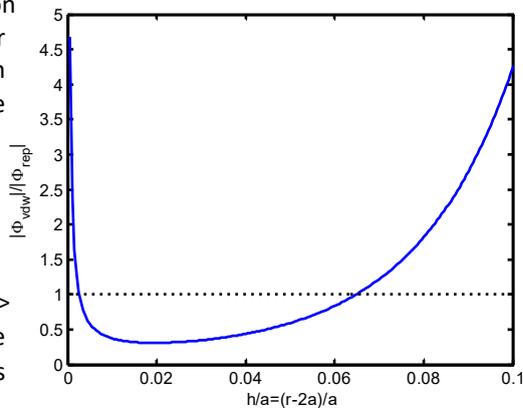
[30] See earlier footnote

At mid-range separation, a minimum is appearing in the curve, more or less visible, depending on the value of $|\Phi_{vdw}|/|\Phi_{rep}|$: if $|\Phi_{vdw}|/|\Phi_{rep}| \ll 1$ then the repulsion is so strong that the minimum is barely visible. In that case, the particles will simply be "pushed back" by the energy barrier. If, on the other hand, $|\Phi_{vdw}|/|\Phi_{rep}| \leq 1$, then the minimum is visible (see the curves for 50 mM and 100 mM on the picture of the previous page). If the minimum is sufficiently deep, the particles will be trapped inside. For the 50 mM case, the minimum is located at about 3.7 nm (so 1/100 of the particle's radius). The two particles will therefore, after having approached each other, stay located 3.7 nm apart. On the scale at which measurements are performed, this means that the detecting device will "see" the two particles as aggregated. This type of aggregation is reversible. By lowering the salt concentration (by diluting the sample with water for example), the energy barrier will increase again ($|\Phi_{vdw}|/|\Phi_{rep}| \ll 1$), the two particles will repel each other and no longer form an aggregate. By shearing or sonicating the sample, the particles will also be able to escape the minimum, but when the sample is then left to stay unstirred, the particles will slowly come back into the minimum, and aggregation will start again.

This minimum is called **secondary minimum** by opposition to the **primary minimum** that always exists at very small separations and that we are now going to discuss.

At small-range separation, a very deep minimum (the primary minimum) is always present, indicating the attractive action of the van der Waals forces at very low separations. When two particles are trapped in this minimum, they are nearly irreversibly aggregated, as the energy to get them out of the minimum has to be extremely large. The separation between the particles is then close to 0 (the particles are touching). The DLVO model does not account for what exactly occurs when particles are touching. Depending on the surface properties of the touching particles, the depth of the primary minimum and the height of the energy barrier, it is sometimes possible to de-aggregate (= **disperse**) the particles by sonication or using a **deflocculating agent** (also called a dispersing or destabilizing agent). This agent is a chemical product. For instance, in soil science, sodium hexametaphosphate[31] $(NaPO_3)_6$ is used in combination with sodium carbonate[32] Na_2CO_3 to disperse soil particles.

Even if the energy barrier is high (of some order of $10\ k_B T$ – but not too high i.e. not of the order of $100\ k_B T$), it is still possible for the particles to " jump" the barrier thanks to thermal fluctuation and get into the primary minimum. Particles in the secondary minimum are therefore called " kinetically" stable, implying that over

---

[31] It is also used as an active ingredient in toothpastes as an anti-staining and tartar prevention ingredient.
[32] also known as washing soda

time (days, weeks, years) (some) particles inside the suspension could get over the energy barrier and aggregate.

The scientists behind the DLVO theory

Boris Derjaguin   Lev Landau   Evert Verwey   Theo Overbeek

Boris Vladimirovich Derjaguin (or Deryagin; Russian: Бори́с Влади́мирович Деря́гин) (1902– 1994) was a Soviet chemist. He helped laying the foundation of the modern science of colloids and surfaces. He is at the origin of the **Derjaguin approximation** widely used in order to approximate the interaction between curved surfaces from the knowledge of interacting planar ones.

Lev Davidovitch Landau (Russian: Ландау, Лев Давидович) (1908 – 1968) was a Soviet physicist who made fundamental contributions to many areas of theoretical physics. He received the 1962 Nobel Prize in Physics. Together with Derjaguin, they published what would become known as the DLVO theory in an article entitled "Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes." by Derjaguin, B. V., & Landau, L. (1941). Acta physicochim. URSS, 14(6), 633-662.

Evert Johannes Willem Verwey (**Verweij**) (1905 – 1981) was a Dutch chemist, who also did research in physical chemistry. He obtained his PhD in 1934 under the guidance of **Hugo Rudolph Kruyt** (who is one of the pioneer in colloid science). In 1934 he moved to the Philips Laboratories (NatLab) in Eindhoven. He continued work on colloids, which was also the topic of his dissertation, and on oxides. The Verwey transition in magnetite is named after him.

Jan Theodoor Gerard Overbeek (1911 - 2007) was a Dutch professor of physical chemistry at the university of Utrecht. Like Verwey, he obtained his PhD under the guidance of Kruyt, and his PhD (obtained in 1941) was entitled *Theorie der Electrophorese, het Relaxatie-effect.* (Theory of electrophoresis, the relaxation effect). He then went to work for Philips, where Evert Verwey was his direct boss. The van der Waals forces had already been studied at Philips in 1936-37 by **Hamaker** and De Boer. On the basis of thermodynamic concepts, they were able to calculate

the free energies that would enable to derive the interaction potential. Their work, which led independently to the DLVO theory, was published in some key articles in 1946, and soon they wrote a book: "Theory of the stability of Lyophobic colloids" by Verwey and Overbeek, with the collaboration of K. van Nes (Elsevier, 1948).

It is clear that the DLVO theory has much to thank to the *Philips Natuurkundig Laboratorium* (**Philips Physics Laboratory**) or **NatLab** in short which was the Philips research department located in Eindhoven. Famous physicists like Hendrik Casimir and Balthasar van der Pol from Utrecht University also worked for the NatLab on experimental physics.

The Schulze-Hardy rule

The Schulze-Hardy rule is a classical well-known empirical relation that states that the critical **coagulation**[33] concentration (c.c.c.) of a suspension varies with the valence z of the counter-ions as:

$$c.c.c. \sim \frac{1}{z^n}$$

where n is an exponent which is generally observed to be between 2 and 6. The relation was found originally by Schulze in 1882[34]. In article of 1925, Weiser[35] states:

"From an investigation of the coagulation by electrolytes of negative arsenious sulfide and antimony trisulfide sols, Schulze concluded that the coagulating power of electrolytes is greater the higher the valence of the ion having a charge opposite to that on the colloidal particles. This conclusion has been supported by the later work of Prost, Linder and Picton, Hardy, Freundlich, and others and has come to be known as Schulze's Law."

One of the remarkable features of the DLVO theory is that it is in agreement with the Schulze-Hardy rule. The c.c.c. corresponds to the situation at which the maximum in the DLVO potential energy curve just touches the horizontal axis (which is the case in the example shown above for kaolinite particles at 100 mM of salt). The point on the curve where this happens is mathematically defined by:

$$\Phi_{DLVO} = \frac{d\Phi_{DLVO}}{dr} = 0$$

---

[33] when the aggregation of particles is due to ions, as discussed until here in the present chapter, it is called coagulation.

[34] H. Schulze, J. Prakt. Chem., 1882, 25, 431; 1883, 27, 320; 1885, 32, 390.

[35] Weiser, H. B. (1925). Adsorption and Schulze's Law. The Journal of Physical Chemistry, 29(8), 955-965.

The expression we have used so far for the DLVO potential is rather complicated to differentiate. To illustrate the fact that the Schulze-Hardy rule can be found from the DLVO theory, we will use another expression for Φ_{DLVO} which is valid for spheres with low $\psi(a)$ and at large separations:

$$\Phi_{DLVO} = \frac{-Aa}{12(r-2a)} + 2\pi\varepsilon_0\varepsilon_r\big(\psi(a)\big)^2 a \exp\big(-\kappa(r-2a)\big)$$

One then finds from $\frac{d\Phi_{DLVO}}{dr} = 0$ that:

$$2\pi\varepsilon_0\varepsilon_r\big(\psi(a)\big)^2 \kappa a \exp\big(-\kappa(r-2a)\big) = \frac{Aa}{12(r-2a)^2}$$

By substituting this expression in the expression for $\Phi_{DLVO} = 0$ one finds that

$$\frac{-Aa}{12(r-2a)} + \frac{Aa}{12\kappa(r-2a)^2} = 0$$

which is satisfied when $r = 2a + \kappa^{-1}$. Using this value in the first equation of the page and setting $\Phi_{DLVO} = 0$ one gets that:

$$\kappa^2 = \left[\frac{24\pi\varepsilon_0\varepsilon_r\big(\psi(a)\big)^2}{A}\right]^2 \exp(-2)$$

where the inverse of the **Debye length** squared for a symmetric salt of valence *z* is given by:

$$\kappa^2 = \frac{2n_\infty(ez)^2}{\varepsilon_0\varepsilon_r k_B T}$$

and where $n_\infty = n_+(\infty) = n_-(\infty)$. The number of ions per volume[36] and the concentration of salt C are related by:

$$n_\infty(\text{number/m}^3) = N_A(\text{mol}^{-1}) \times C(\text{mM})$$

One obtains:

$$C(\text{mM}) = \frac{\varepsilon_0\varepsilon_r k_B T}{2N_A(ez)^2} \left[\frac{24\pi\varepsilon_0\varepsilon_r\big(\psi(a)\big)^2}{A}\right]^2 \exp(-2)$$

---

[36] In physics the unit of volume is the cubic meter (m^3) whereas it is litre (L) in chemistry. In the formula for κ^2, n_∞ should be in number/m^3.

This relation clearly indicates that the salt concentration required to obtain the c.c.c. is proportional to $1/z^2$. From another expressions of the DLVO theory, based on other approximations[37], it is possible to show that the salt concentration required to obtain the c.c.c. is then proportional to $1/z^6$ or $1/z^n$ with n values between 2 and 6.

Note that Schulze-Hardy rule in the simple case we have illustrated above originates from the fact that $\kappa^2 \sim z^2$. This can be linked to an intuitive understanding:

For ions with a higher valence, the Debye length will be smaller for the same salt concentrations. As the Debye length indicate the distance over which the electric repulsive potential is going to zero, one can say that multivalent ions screen better the particle's surface charge than monovalent ions (for the same amount of salt):
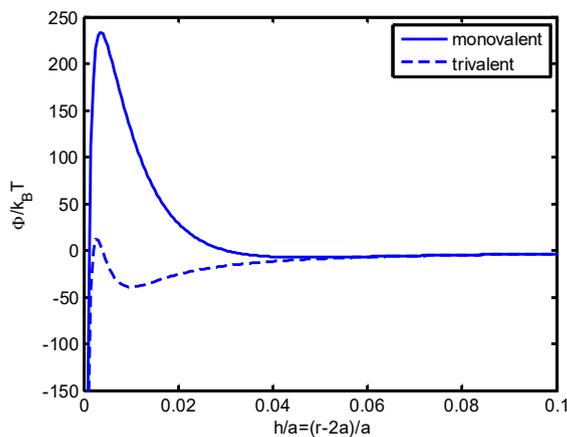
Interaction energy between two kaolinite particles of radius 370 nm, using A = 3.10^{-20} J, and $\psi(a) = 40$ mV. The spheres are immersed in an electrolyte solution made of 10 mM monovalent salt (z = 1 → $\kappa^{-1} = 3$ nm) and 10 mM trivalent salt (z = 3 → $\kappa^{-1} = 1$ nm).

Link between surface potential and surface charge

We have until now expressed the fact that the particle is charged by using its electric surface potential $\psi(a)$ in the equation. Usually, however, it is the surface charge of a particle that is known (the surface charge can be determined by **potentiometric titration**). The relation between $\psi(a)$ and the particle surface charge $\sigma_S (C/m^2)$ cannot be derived analytically, but numerical solutions are available. Empirical

---

[37] E. J. W. Verwey and J. Th. G. Overbeek, The Theory of the Stability of Lyophobic Colloids (Elsevier, Amsterdam, 1948) and I. M. Metcalfe and T. W. Healy, Charge-regulation Modelling of the Schulze-Hardy Rule and related Coagulation Effects, Faraday Discuss. Chem. SOC., 1990,90, 335-344

relations, based on the numerical results, can also be found. A very useful one[38], valid for all $\zeta = \psi(a)$ (implying the shear plane is at the surface of the particle) and $\kappa a \geq 0.5$ is:

$$\sigma_S = \frac{\varepsilon_0 \varepsilon_r k_B T}{ez} \kappa \left[2 \sinh\left(\frac{ez\zeta}{2k_B T}\right) + \frac{4}{\kappa a} \tanh\left(\frac{ez\zeta}{4k_B T}\right) \right]$$

At low surface potential, one finds, using $\sinh(x) = x$ and $\tanh(x) = x$ for low x :

$$\sigma_S = \varepsilon_0 \varepsilon_r \kappa \zeta \left[1 + \frac{1}{\kappa a} \right]$$

and we find a linear relationship between surface charge and surface electric potential. (This linear relation obviously breaks down for higher potentials). The linear equation can be related to the definition of a **capacitance**[39] C (as found in electrostatics), namely:

$$C = \frac{Q}{U} = \frac{\sigma_S S}{U}$$

where Q (in Coulomb, symbol C) is the charge on one condensator plate of surface S (it is $-Q$ on the other plate), U is the electric potential difference between the two condensator plates. In the present case, one can identify $C/S = \varepsilon_0 \varepsilon_r \kappa [1 + 1/(\kappa a)]$ and $U = \zeta - 0 = \zeta$ (the difference between the surface electric potential and the electric potential in the bulk (i.e. far from the particle).

As we have already discussed, often, there is a Stern layer in between the surface of the particle and the slip plane. This layer has an extremely small thickness *d*. Therefore, it is usually assumed that the relation between the surface electric potential $\psi(a)$ and the zeta potential ζ (i.e. the potential at the slip plane) is given by a similar linear relationship:

$$C_{Stern} = \frac{\sigma_S S}{\psi(a) - \zeta} = \frac{\varepsilon_0 \varepsilon_{Stern}}{d}$$

Where the permittivity ε_{Stern} of the Stern layer is not necessarily the same as water: it could be lower, because the water molecule dipoles in that region are strongly

---

[38] Loeb, A. L., Overbeek, J. T. G., Wiersema, P. H., & King, C. V. (1961). The electrical double layer around a spherical colloid particle. Journal of The Electrochemical Society, 108(12), 269C-269C.

[39] The capacitance, symbolised by C has as unit F (Farad, named after Michael Faraday). Not to be confused with C (Coulomb), the unit of charge, nor with C (in mol/L) the ionic concentration.

oriented and bound to the surface. Usually, C_{Stern} is used as an adjustable parameter in the models.

Measurement of the zeta potential

We will see in **Chapter 6** how the zeta potential can be related to the optimal flocculation rate. Measuring the zeta potential of a particle (as function of solvent property, like the pH, salinity or added polyelectrolyte concentration) therefore enables to predict the range (of pH, salinity, etc..) where flocculation is most likely to occur: this happens when the zeta potential is close to zero. A very well established method to assess the zeta potential of colloidal particles in suspension is the use of the **electrophoresis** technique. Electrophoresis occurs when a colloidal suspension is under influence of an electric field and each charged particle is moving to the electrode which has a charge of opposite sign. Usually the electric field is a fast alternating (AC) field, to prevent the electrodes from corroding. The particles then follow the reversal of electrode charges by oscillating at the same frequency as the electric field (for the low frequencies that are typically used in the measurements). The velocity of the colloidal particles in the electric field is recorded. This can be done with laser Doppler velocimetry where the phase shift in the scattered laser light can be related to the particle's velocity, or simply by video microscopy. In video microscopy, a movie of the moving particles is analysed to get their velocity. Laser Doppler velocimetry is suited for particles in the 1 nm - 1µm range, whereas video microscopy can be used for larger particles, in the range 0.5µm – 10 µm.
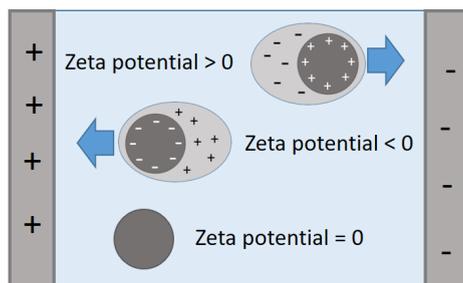
From the measure of the particle's velocity, the zeta potential can be deduced. Charged colloidal particles will move to the oppositely charged electrodes, whereas uncharged particles have no electrophoretic velocity and hence their zeta potential is zero. During electrophoresis, the electric double layer around a particle is deformed, and an electric dipole is created.

A simple relation between particle velocity and zeta potential has been derived by **Smoluchowski**. This relation is:

$$v = \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta} E$$

where v is the velocity of a colloidal particle, η is the viscosity of the solvent (water) and E the applied electric field. Usually the **electrophoretic mobility** μ_E of the particle is given instead of its velocity. The electrophoretic mobility μ_E is defined by:

$$\mu_E = \frac{v}{E}$$

The Smoluchowski formula is valid when κa is very large – in fact so large that it corresponds to salt concentrations so high that they are usually not measured in practise. However, it is often used in articles to convert the electrophoretic mobility (which has units of $m^2 V^{-1} s^{-1}$) into zeta potential units (V), which are more appealing: an electrophoretic mobility of 1.7×10^{-8} $m^2 V^{-1} s^{-1}$ corresponds to a zeta potential of 25 mV for example. One can therefore defined a measured "Smoluchowski" zeta potential by:

$$\zeta_{\mu_E} = \frac{\eta}{\varepsilon_0 \varepsilon_r} \mu_E$$

and plot ζ_{μ_E} instead of μ_E.

There exists nowadays advanced numerical and analytical models that enable to convert the electrophoretic mobility into zeta potentials, for any κa. A lot of work on this topic has been performed by Hiroyuki Ohshima professor at the University of Tokyo, who derived convenient, easy to implement, analytical formulas[40].

First, let us look at the simple case where we assume that the Smoluchowski relation would be correct. We will assume the following surface potential/surface charge relation:

$$\psi(a) = \zeta = 2 \frac{k_B T}{e} \text{arcsinh} \left(\frac{e \sigma_S}{2 \varepsilon_0 \varepsilon_r k_B T \kappa} \right)$$

This potential is in fact the zeta potential / surface charge relation for a flat surface (one can see, by comparing it with the relation given above for spherical particles, that many elements are similar). This relation is valid for spheres provided that κa is large.

---

[40] See the books: Ohshima, H. (2006). Theory of colloid and interfacial electric phenomena (Vol. 12). Academic Press, and: Ohshima, H. (2011). Biophysical chemistry of biointerfaces. John Wiley & Sons.

In the figure below, we have plotted ζ as a function of ionic strength (salinity) for a *constant surface charge* given by $\sigma_S = 3.9 \times 10^{-2}$ C/m².
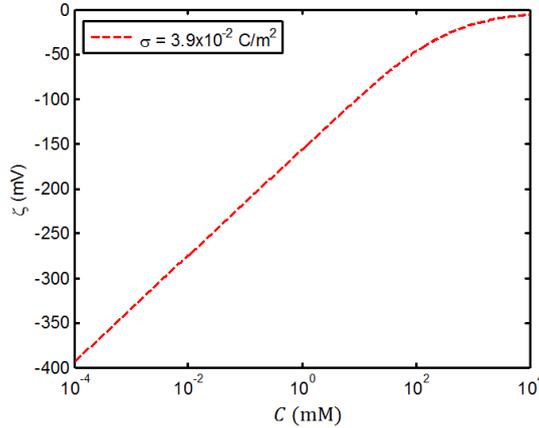
Variation of the surface potential assuming constant charge density as a function of ionic strength. We took $D_k = 2 \times 10^{-9}\ m^2/s$ which is representative for the diffusion coefficients of K^+ ions. The particle's radius was taken to be 300 nm.

One can clearly see that the surface potential (equal to the zeta potential as the shear plane is at the particle's surface) is varying as a function of salinity, whereas the surface charge is not. This is due to the fact that, by adding ions, the double layer gets compressed and that the potential close to the particle's surface is changing accordingly. The relation between the particle surface charge and the electric potential is given by Gauss's law:

$$\sigma_S = -\varepsilon_0 \varepsilon_r \left(\frac{d\psi}{dr}\right)_{r=a}$$

This is a fundamental relation, from which all the surface charge / potential relations given above are derived. This equation tells us in particular that if σ_S is constant, $d\psi/dr$ should be constant. For the sake of argument, we can make the assumption that:

$$\left(\frac{d\psi}{dr}\right)_{r=a} \sim \frac{\psi(a + \kappa^{-1}) - \psi(a)}{(a + \kappa^{-1}) - a} \sim \frac{-\psi(a)}{\kappa^{-1}}$$

as $\psi(a + \kappa^{-1}) \sim 0$. Therefore $\psi(a)$ *should* vary if κ^{-1} is changing.

If the Smoluchowski relation would be correct, this implies that we should have $\zeta_{\mu_E} = \zeta$ and the measured electrophoretic mobility would be equal to:

$$\mu_E = \frac{\varepsilon_0 \varepsilon_r}{\eta} \zeta = 2\frac{\varepsilon_0 \varepsilon_r k_B T}{e\eta} \text{arcsinh}\left(\frac{e\sigma}{2\varepsilon_0 \varepsilon_r k_B T\kappa}\right)$$

The "expected to be measured" ζ_{μ_E} should look like the dashed red curve given above. Note that for moderate to low salinities the expected zeta potential gets unrealistic values (in general $|\zeta| < 150$ mV).

A formula, derived by Ohshima, valid for $\kappa a > 10$ and monovalent electrolytes (z = 1), relates the particle's surface charge (assuming that the shear plane is at the particle's surface and no Stern layer exists) to the electrophoretic mobility or ζ_{μ_E}. We give it here (see the book given in footnote to find the full derivations), as it will enable us to understand an important feature of the mobility (μ_E or ζ_{μ_E}) versus salinity curve.
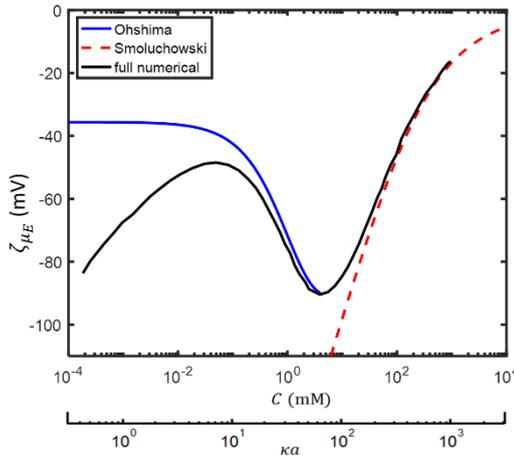
The formula found by Ohshima is given by:

$$\zeta_{\mu_E} = \zeta - \frac{2F}{1+F}\frac{k_B T}{e}\ln\left(\frac{1}{2}\left[1 + \exp\left(\frac{e\zeta}{2k_B T}\right)\right]\right)$$

with

$$F = \frac{2}{\kappa a}\left(1 + \frac{2\varepsilon_0 \varepsilon_r}{\eta D_k}\left(\frac{k_B T}{e}\right)^2\right)\left(\exp\left(\frac{e\zeta}{2k_B T}\right) - 1\right)$$

where D_k is the **ionic diffusion coefficient** of the counterion.

In the following figure, we have plotted the measured ζ_{μ_E} according to the formula of Ohshima, the formula of Smoluchowski and the full numerical solution, using the same values as given above.

ζ_{μ_E} according to the formula of Ohshima, Smoluchowski and the full numerical solution as function of ionic strength C (mM) and κa. Particles have a radius of 300 nm and a constant charge[41].

One can now see that the behaviour of the curve is complex. This is due to the fact that the double layer, when

---

[41] Chassagne, C., & Ibanez, M. (2012). Electrophoretic mobility of latex nanospheres in electrolytes: Experimental challenges. Pure and Applied Chemistry, 85(1), 41-51.

the electric field is applied, is deforming. This deformation is then changing the local electric field "felt" by the colloidal particle and hence the colloidal particle adapts its velocity. At high ionic strengths, the double layer is so compressed that it hardly deforms and the conditions used by Smoluchowski to derive his equation applies. This happens for salinities larger than 100 mM (larger than the salinity of sea water) which are seldom encountered in experiments.

The full numerical solution behaviour has been confirmed by experiments[42]:
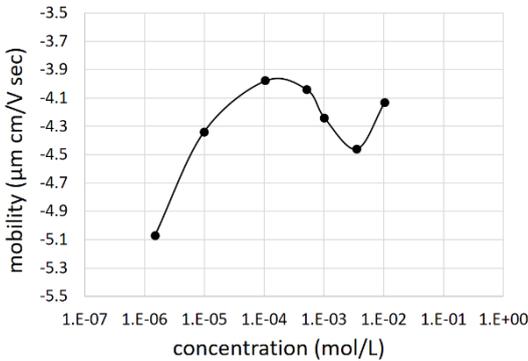
Electrophoretic mobility of 156 nm latex particles in NaCl solution: the counterions of the particles are nearly the only ions in solution. At very low ionic strength, the data follows the theoretical prediction and the mobility increases dramatically.

These experiments were done in very "clean" water.

In the following figure, NUM represents the full numerical solution and OHSHIMA is the curve according to the formula given above. No adjustable parameters have been used. The experiments were done using spherical, homogeneously charged, latex spheres. The symbols correspond to measurements. The values of the variables are the ones given in the previous theoretical figures. It would now appear that the formula of Ohshima works better than the full numerical

Measured ζ_{μ_E} as function of added salt concentration

solution at low ionic strength: this is an artefact. In the experiments performed here, below 0.1 mM of added salt, the ionic strength remained constant, due to the

---
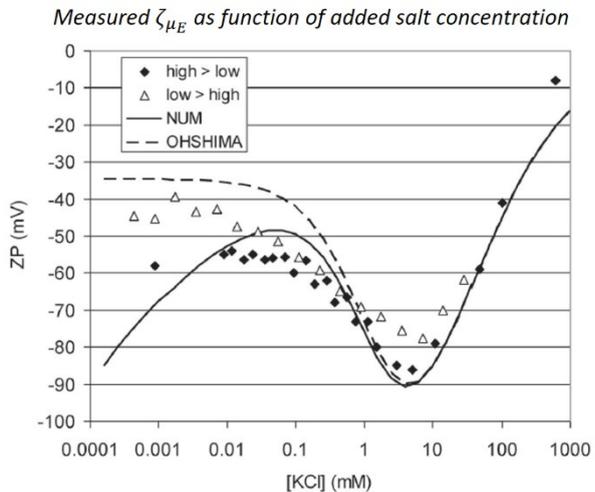
[42] M. R. Gittings, D. A. Saville , Electrophoretic Mobility and Dielectric Response Measurements on Electrokinetically Ideal Polystyrene Latex Particles, Langmuir, 1995, 11 (3), 798-800

presence of background ions. At very low concentrations of added salt, the added salt concentration is then not representative of the true electrolyte concentration. These background ions were released when the particles were added to the demi-water[43]. To remove these background ions, dialysis (see **Chapter 2**) can be performed. This was not done in the present experiments but was performed in the experiments performed by Gittings et al. given above. In most experiments on clays, dialysis is not often done, as it could trigger some unwanted effects, such as the **delamination** of the primary clay particles (see **Chapter 5**). This implies, in the present case, that the measured ζ_{μ_E} remains constant below 0.1 mM of added salt. The difference between white and black symbols are related to the experimental protocol. From this study, it was concluded that the high > low protocol was the most adapted, since it corresponds best to the expected theoretical values.

From the measured ζ_{μ_E} it is also possible to have an idea of the surface charge of the particles. One relation between surface charge and electric potential some pages above for $\kappa a \geq 0.5$. One can show that in most cases encountered in practice, this formula simplifies into:

$$\sigma_S = \frac{2\varepsilon_0 \varepsilon_r k_B T}{ez} \kappa \sinh\left(\frac{ez\zeta}{2k_B T}\right)$$

which is the transposed of the relation given at the beginning of this section, where it was given as zeta potential / surface charge relation for a flat surface. If we plot:

$$\sigma_{\mu_E} = \frac{2\varepsilon_0 \varepsilon_r k_B T}{ez} \kappa \sinh\left(\frac{ez\zeta_{\mu_E}}{2k_B T}\right)$$

We get:
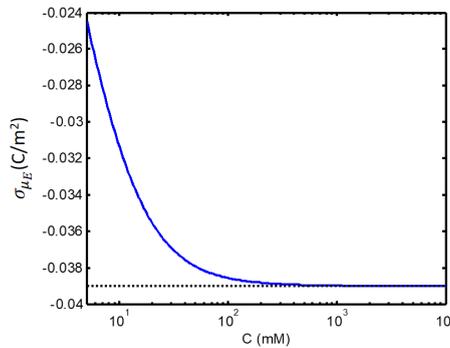
where the black dotted line represents the surface charge of the particle. For small C, σ_{μ_E} goes to zero (not shown) as κ becomes very small, but for C > 10 mM one has

---

[43] Demi-water is water with an extremely low conductivity.

in good approximation $\sigma_{\mu_E} \sim \sigma_S$. This is due to the fact that in that range of concentrations one has $\zeta_{\mu_E} \sim \zeta$ as discussed above.

Ionic conductivity and diffusion coefficient

When an electric field is applied to a solution, ions are moving like colloidal particles do via electrophoresis. The electrophoretic mobility of ions cannot be measured by the techniques described above, as they are too small to be seen, even by laser Doppler velocimetry (ions' sizes are below 1 nm). An alternative way to assess the electrophoretic mobility of ions is to measure **conductivity**. As ions are charged, their displacement can be related to an electric current which in turn can be related to the applied electric field *E* by:

$$J = K_e E$$

where K_e is the conductivity (in S/m) of the solution. The symbol S (1 S = 1 Ω^{-1}) stands for "Siemens" (in honour of Werner von Siemens, founder of the electrical and telecommunications company Siemens). Note that in the previous relation the electric current *J* is expressed in A/m². The electric intensity *I* is in A (Ampère[44]) in the standard relation $U = RI$, where *U* is the applied voltage difference and *R* the resistance.

In Chapter 2, we have seen that mobility of a particle of radius a (independently of the applied force) is given by $\mu = 1/(6\pi\eta a)$ if Stokes' s friction force applies. This expression gives the right order of magnitude for the electrophoretic mobility of a colloidal particle, however it is imperfect, as the application of an electric field induces corrections to the friction force on the colloidal particle. For ions, Stokes' s friction force is a good approximation. For example, for a typical ion radius of 0.1 nm, one finds that the **ionic diffusion coefficient** (see **Chapter 2**) defined by $D_k = k_B T/(6\pi\eta a)$ gives $2.1 \times 10^{-9} \, m^2/s$ which is very close to the diffusion coefficient of K^+ and Cl^- for example, which have radii of the order of 0.1 nm.

The conductivity of a solution is linked to the ionic diffusion coefficients of its ionic constituents by:

$$K_e = \sum_k \frac{e^2 z_k^2 n_\infty}{k_B T} D_k$$

where we recall that: $n_\infty (\text{number}/m^3) = N_A (\text{mol}^{-1}) \times C (\text{mM})$. The ionic diffusion coefficients can be found in handbooks, and are usually given as **limiting ionic conductivities** Λ_k^∞ (S m² mol⁻¹):

---

[44] In honour of André-Marie Ampère (1775-1836) a French physicist and mathematician who was one of the founders of classical electromagnetism.

$$\Lambda_k^\infty = |z_k| D_k N_A \frac{e^2}{k_B T}$$

A convenient relation between the **Debye length** and the **conductivity** is given by:

$$K_e = \varepsilon_0 \varepsilon_r \kappa^2 \frac{z_+ D_+ - z_- D_-}{z_+ - z_-}$$

Most electrophoretic measuring devices also measure the conductivity of the suspension. Knowing the type of salt present in the suspension, κ^2 can then be estimated. Note that electrophoretic measurement are always done for small concentrations of colloidal particles (to avoid particle-particle interactions), and that the contribution of colloidal particles to conductivity is negligible compared to the one of ions: in **Chapter 1**, we have seen that on average for 1 colloidal particle there is 10^{10} ions in solution.

Illustrations

van der Waals
https://en.wikipedia.org/wiki/Johannes_Diderik_van_der_Waals

Derjaguin
https://link.springer.com/article/10.1023%2FA%3A1020686631909

Landau
https://en.wikipedia.org/wiki/Lev_Landau

Verwey
https://chg.kncv.nl/geschiedenis/biografieen/v/verweij,-e.j.w.

Overbeek
http://www.ecis-web.eu/overbeek.htm

Chapter 4

Other colloids: polymers, surfactants, microorganisms …

So far, we have only discussed *electrostatic* stabilization, induced by charges (the charges on the particle's surface) and destabilization triggered by the ions in solution. There are however other mechanisms leading to (de)stabilization. These are related to the presence of special molecules in the colloidal suspension. These molecules are called **polymers** and **surfactants.** These molecules have also a colloidal size and combine with clay particles to create new colloidal particles (flocs, micelles…). Flocculation (the formation of flocs) will be the topic of **Chapters 5 and 6**. In the present chapter, we will review the properties of polymers and surfactants. These molecules can have an industrial origin, but can also originate from plants and animals. Polymers and other type of organic matter are produced or altered by microorganisms which are living in the mud. Some of these microorganisms, presented here, have also a colloidal size and can be part of flocs.

Polymers & polyelectrolytes

Polymers come in various size, shape and charges. They can be found in nature or made in the laboratory. On some of the next pages, we give an overview of some polymers. Polymers that are electrically charged are called **polyelectrolytes**. Polymers are colloidal particles, as one of their characteristic size (usually their apparent radius or their width) lays within the colloidal range.
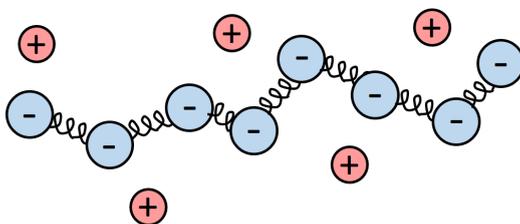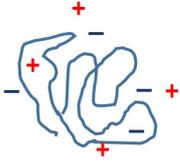

Illustration of a polyelectrolyte in water. The polyelectrolyte is seen as a long flexible chain made of repeating units i.e. the monomers (circles) connected by springs. The counterions (in this case cations) are in vicinity of the negatively charged units

Polyelectrolytes have been studied extensively over the years[45] as their behaviour can be quite complex. As the charged colloidal particles studied in **Chapter 3**, they are subjected to electrostatic interactions, but the fact that they are flexible chains adds a **conformation** behaviour. Depending on the affinity of the polyelectrolyte with the solvent and the ionic concentration, the polyelectrolyte conformation can be different.

---

[45] Dobrynin, Andrey V., and Michael Rubinstein. "Theory of polyelectrolytes in solutions and at surfaces." Progress in Polymer Science 30.11 (2005): 1049-1118 and references within (especially the books)

Very simple illustrations of polymers in solutions are given here:

random coil	*expanded conformation*	*coiled conformation*
the uncharged polymer configuration is close to a self-avoiding (=two monomers cannot occupy the same position at the same time) **random walk** (see **Chapter 2**)	the charges on the polymer are repelling each other through Coulomb forces leading to an extension of the chain in the water (at low salinity). The extension of the polymer depends on its degree of dissociation .	in the presence of salt, the charges on the polymer will be screened (see **Chapter 3**), and it is energetically favourable for the polymer to coil.

The conformation of polymers in suspension can be studied theoretically by numerical simulations. As in the picture above, one then represents usually the monomers by spheres that interact through **Coulomb** and **Lennard-Jones** potentials (see **Chapter 3** for Coulomb and **Chapter 7** for Lennard-Jones potentials). The numerical models are compared to different types of experimental data (light scattering, fluorescence or Raman spectroscopy, Nuclear Magnetic Resonance (NMR), etc...).

The behaviour of polymers and polyelectrolytes is also extensively studied from thermodynamic concepts, similar to the ones introduced at the end of **Chapter 8**.

Branched and crosslinked polymers

We have represented polyelectrolytes as long chains, but there exist also **branched** and **crosslinked** polyelectrolytes/polymers.

left: branched polymer; right: crosslinked polymer

As can be seen underneath, natural polymers are often branched or crosslinked.

**Some examples of polymers, polyelectrolytes and (de)flocculating agents**

By definition a **polymer** is a macromolecule, which consists of many repeated subunits (for example n repeated ethylene units $-(CH_2)_n-$). Natural polymeric material include fibres like cellulose (found in wood, and used in paper making), proteins, gums, DNA, etc… Synthetic polymeric material include plastics (latex – natural latex also exists), paints and glue.

Polyacrylamide ($-CH_2CHCONH_2-$) : It forms a soft gel when hydrated, and is used in applications such as gel electrophoresis. One of the largest uses for polyacrylamide is to flocculate solids in a liquid. Polyacrylamide can be supplied in **copolymer**[46] forms of acrylamide combined with other chemical species to form an acrylic acid or a polyelectrolyte. Common uses of polyacrylamide and its derivatives are in Enhanced Oil Recovery, in water treatment (for flocculation), and processes like paper making and screen printing.
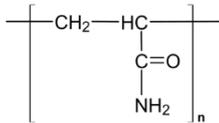
Xanthan gum ($-C_{35}H_{49}O_{29}-$) : This natural polymer is a polysaccharide secreted by the bacterium Xanthomonas campestris. It is composed of repeat units of glucose, mannose (sugars). In foods, xanthan gum is most often found in salad dressings and sauces. In the oil industry, xanthan gum is used in large quantities, usually to thicken drilling mud. In cosmetics, xanthan gum is used to prepare water gels, usually in conjunction with bentonite clays. It is also used in oil-in-water **emulsions**[47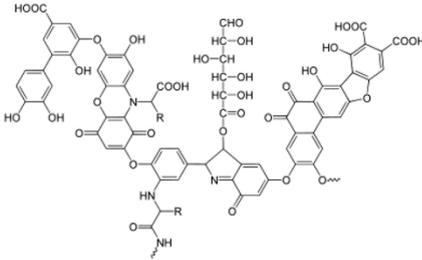] to help stabilize the oil droplets against coalescence. Natural polymers secreted by microorganisms are in general called **Extracellular Polymeric Substances** (EPS) – generally referred to as slime. They are composed of polysaccharides, and include other macro-molecules such as DNA, lipids and humic substances. In nature, they are important for the formation of **biofilms**. Biofilms (see picture) are communities of symbiotic micro-organisms (bacteria, fungus, algae, protozoa) which can stick to each other and are embedded within a self-produced matrix of EPS. Biofilms are an important food resource for invertebrates, and can be found at flood on mudflats. They can also help to stabilize a soil against erosion.

---
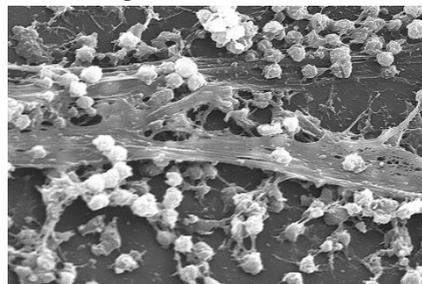
[46] A copolymer is a polymer having more than one type of monomer.
[47] See "surfactant" , below.

Humic acids : They are a major organic constituent of soil, and produced by the biodegradation of dead organic material. Each humic acid is a complex mixture of different acids containing carboxyl and phenolate groups. They can form complexes with ions commonly found in the environment creating humic colloids. Fulvic acids for instance (a special type of humic acid) are colloidal polyelectrolytes. Humic acids are able to interact with each other and create higher order complexes. They can also form complexes with metal ions, and hereby regulate their bioavailability. Modern investigations have found that humic acid is released from straw when mixed with mud and increases clay's plasticity. The bricks made with mixtures of straw and mud are stronger (they are less likely to break or lose their shape) than mud bricks[48]. The adsorption behaviors of humic acids and fulvic acids onto clay minerals in an on-going topic of research.

Biodegradation of polymers

The polymers used in industry are often synthetic and made from polyolefins. These are for example polyethylene and polystyrene. These polymers are produced from fossil fuels, and are considered to be undegradable. In recent years there has been a gain in interest in biodegradable polymers[49]. These polymers are defined as those that undergo microbially induced chain scission leading to their mineralization. The biodegradation is strongly dependent on environmental conditions, like pH, humidity, oxygenation and the presence of some metals. Biodegradable polymers are made from corn, wood cellulose or are synthetized by bacteria from small molecules like butyric acid or valeric acid that give polyhydroxybutyrate and polyhydroxyvalerate.

The degradation of polymers in soil is an on-going topic of research, and is strongly linked to the biochemical and physical processes occurring within the soil. The degradation of a polymer within a clayey fabric has for instance consequences for the consolidation, permeability and strength of this soil.

---

[48] Lucas, A.; Harris, J.R. (1998). Ancient Egyptian Materials and Industries. New York: Dover Publications. p. 49. ISBN 0-486-40446-3

[49] Ray, Suprakas Sinha, and Mosto Bousmina. "Biodegradable polymers and their layered silicate nanocomposites: in greening the 21st century materials world." *Progress in materials science* 50.8 (2005): 962-1079.

Polymers and clays

Polymers can interact in different ways with clay colloidal particles. A polymer can adsorb on the surface of a colloidal particle as a result of a Coulombic (charge-charge) interaction, dipole-dipole interaction, **hydrogen bonding**, van der Waals forces (see **Chapter 3**), or any combination of these mechanisms.

A **hydrogen bond** is the electrostatic attraction between polar groups that occurs when a hydrogen (H) atom bound to a highly electronegative atom such as nitrogen (N), oxygen (O) or fluorine (F) experiences attraction to some other nearby highly electronegative atom (see **Chapter 3**). These hydrogen-bond attractions can occur between molecules (intermolecular) or within different parts of a single molecule (intramolecular). Depending on geometry and environmental conditions, the hydrogen bond typically has between 5 and 30 kJ/mole in thermodynamic terms. This makes it stronger than a van der Waals interaction, but weaker than charge-charge (**covalent**) bonds. This type of bond can occur in inorganic molecules such as water and in organic molecules like DNA and proteins.
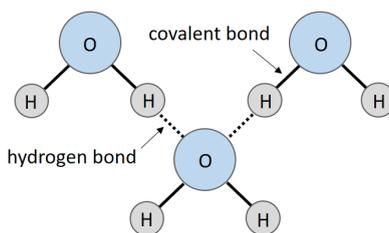

Water molecules have both hydrogen bonds and covalent bonds

As the polymer is a very long snake-like molecule, the way it sticks to a particle depends not only of its affinity with the particle's surface, but also its affinity with the solvent. The usual result is that the polymer sticks to certain points of the surface (as **trains**), separated from one another by **loops** and for much of its length it is able to trail out into the solvent (as **tails**).
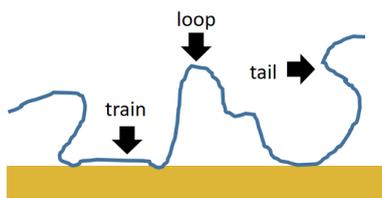
Aggregation by the addition of polyelectrolytes is a method widely used in industry in the treatment of mineral ores and in the purification of water.
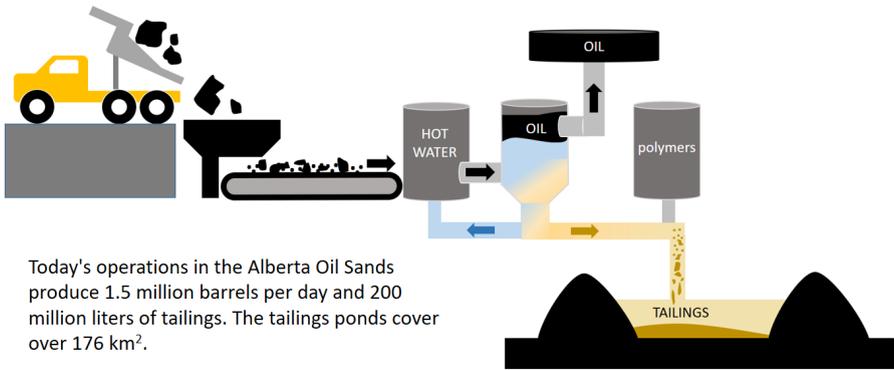
Today's operations in the Alberta Oil Sands produce 1.5 million barrels per day and 200 million liters of tailings. The tailings ponds cover over 176 km^2.

Tailings ponds in Alberta (Canada): these are by-products of the oil sand industry. After extracting the bitumen, the remaining fine particles (tailings) are flocculated by polymers, deposited in large ponds and left to settle.

When two particles, coated with polymers, are getting close to each other, this can lead to two different aggregation mechanisms[50]:

Bridging aggregation[51]: the loops and tails of one polymer of one particle will be able to attach to the other particle. This process is facilitated when the amount of polymers adsorbed on the particles is not too high. Usually, one finds that the optimum polymer concentration to achieve flocculation corresponds to half surface coverage for the polymer. Polymeric bridges are changing as function of shear (see **Chapter 5**). Bridging aggregation can even occur with polyelectrolytes having surface charges of same sign as the ones of the particles. In that case, aggregation is enabled by the presence of oppositely charged ions in the water, that will act as "binders" between particle and flocculant.
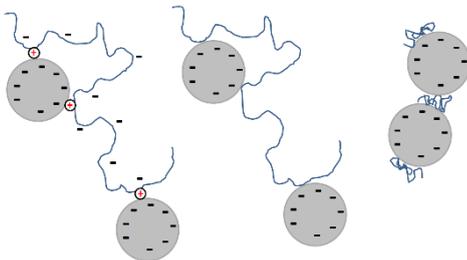
Left and middle: bridging flocculation. The anionic polyelectrolyte on the left needs a cation (in red) to bridge to the clay; right: patching flocculation

Patching aggregation: patching aggregation occurs usually when polyelectrolytes have a charge that is opposite in sign to the one of the particle. These polyelectrolyte then strongly bind to the particles, and their tails do not extend much into the

---

[50] Bergaya, Faïza, and Gerhard Lagaly. Handbook of clay science. Vol. 5. Newnes, 2013.

[51] When particles are aggregated through polymers, one usually speaks of flocculation (as oppose to coagulation, i.e. salt-induced aggregation).

solvent. Aggregation is then made possible between one polymer patch of one particle and the bare surface area of another particle.
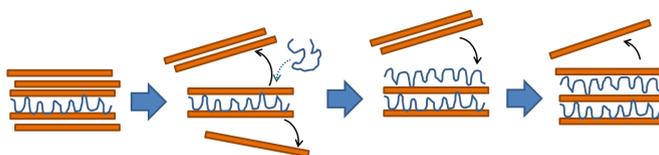
Interlayer interaction

Until now we have represented clay particles are spheres, but as we have seen in **Chapter 1**, clay particles are not spherical. Many types of clay particles have the form of stacks of platelets, and the interaction between clay and polyelectrolyte can also influence the stacking of platelets:
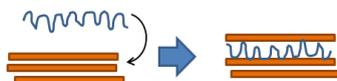
Clay interlayer with a cationic polyelectrolyte

Cationic polymers can strongly interact with clay minerals and penetrate between the layers if they are small enough (in particular montmorillonite due to its swelling properties, see **Chapter 5**).

Another mechanism is induced by the strong interaction between a polyelectrolyte coated face of a clay particle with the bare face of another: one silicate layer of one or each clay particle can then be peeled-of so that one (or two) new bare faces are exposed and able to adsorb polymers:

Interaction between polyelectrolyte and clay platelets. Two silicate layers are peeled-of so that one polymer can adsorb on one of the bare surfaces. Subsequently another stack of platelets can adsorb on the polyelectrolyte, and again a silicate layer is peeled-of. This leads to the intercalation of polyelectrolytes between the platelets of the clay.

Depletion effect

When small (or coiled) polymers are added to a suspension and do not stick to the suspended particles they can lead to their destabilization through a **depletion effect**. This effect occurs when two colloidal particles are close to each other, and that there is a region (indicated by the black arrow below) where the polymers cannot penetrate because they are too big to get inside the volume. The resulting gradient in polymer concentration gives rise to a lower osmotic pressure in that region, leading to aggregation.

region of lower osmotic pressure
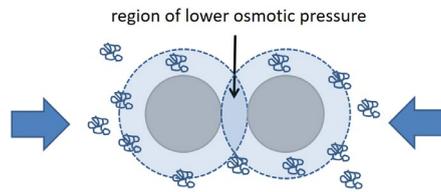
Even though they were not aware of its cause, early makers of inks and paints used

to add natural gums or other polymers to their pigments to promote the binding between the pigments via depletion effects[52].

Art work made of clay mineral suspensions, probably using saliva as a binder. Saliva is composed of 99.5% water (electrolyte) with several natural colloidal and polymeric agents. (Grotte de Lascaux, 17.000 BC)

Stabilization

In some cases, the presence of polymers in the suspension can enhance its stabilization. This occurs in particular when:

1 – the particles are fully coated by polyelectrolytes of opposite charge. The particles then get an effective charge of same sign as the polyelectrolyte they are coated with and start repelling each other, much like standard particles of same charge do:
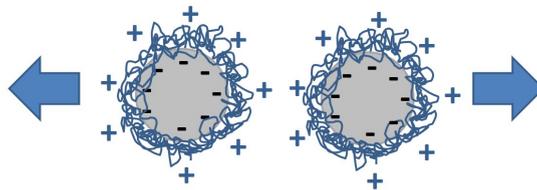
Charge repulsion between fully polyelectrolyte-coated particles

2 – the particles are coated with (uncharged) polymers. In that case, aggregation will be prevented by **steric repulsion.** Steric effects originate from the osmotic pressure in the region where the polymers overlap, due to the crowding of the polymer

---

[52] Lambourne, R., & Strivens, T. A. (Eds.). (1999). Paint and surface coatings: theory and practice. Elsevier.

chains, much like what we have seen for overlapping double layers (in that last case, it was due to the crowding of ions):

region of higher osmotic pressure
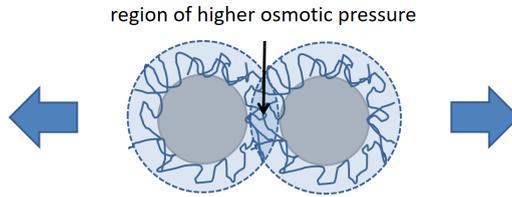
Like for DLVO theory, theories have been developed to account for steric effects, from which stability criteria are derived[53].

Surfactants (amphiphilic molecules)

Colloids that has not yet been reviewed and are of interest are **surfactants** (their name come from **surf**ace **act**ive ag**ents**, see **surface tension** below) .
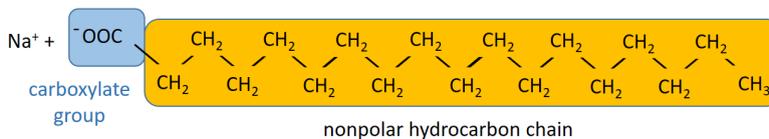
The surfactant called sodium stearate (sodium octadecanoate, $CH_3(CH_2)_{16}CO_2Na$) is found in most soap. Soaps represent about 50% of commercial surfactants. In water, sodium stearate dissociates to form an ionic (carboxylate) group and a counterion (Na^+)

Surfactants can be seen as a special class of polymers. Their main property is that they possess both a **hydrophilic** part (their "head") and a **hydrophobic** part (their "tail", which is an hydrocarbon chain).

---

[53] Stuart, MA Cohen, et al. "Adsorption of ions, polyelectrolytes and proteins." *Advances in colloid and interface science* 34 (1991): 477-535.
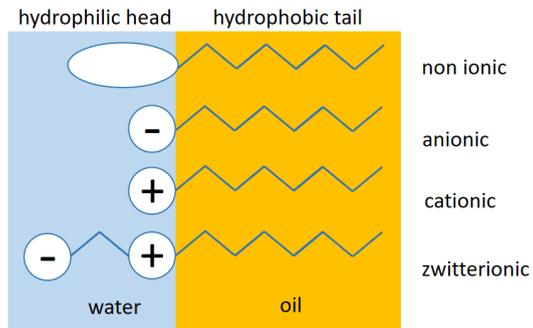
Surfactant classification according to the composition of their head

Surfactants can also be found in natural systems, usually in the group of **lipids**, molecules that include fats, waxes, glycerides, vitamins… Lipids can be either hydrophobic or **amphiphilic** (**amphiphile** is another word for **surfactant**). An abundant amphiphilic lipid is the phospholipid (note that this surfactant has two tails (in yellow) – there are surfactants which have more):
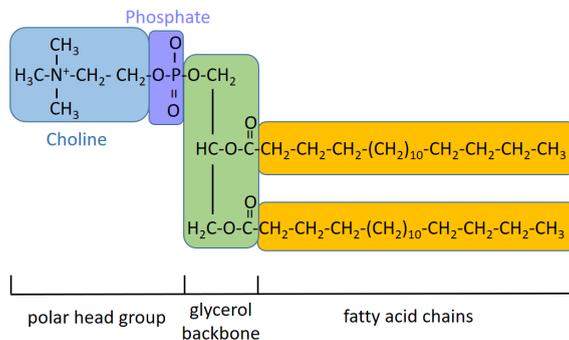
Thanks to their dual hydrophilic/hydrophobic properties, surfactants are able to form particular structures called **micelles**:
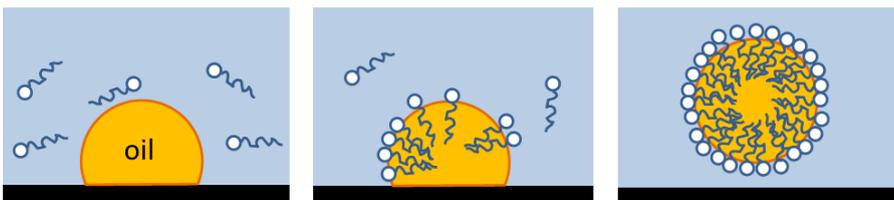
From left to right: adsorption of surfactant on an oil droplet and formation of a micelle (right figure)

We consider here a drop of oil adsorbed on a surface. Surfactant molecules are added into the water. Energetically, it is favourable for the molecules to go to the oil/water interface and position themselves such that their hydrophilic (from the greek "water-loving") head sits in the water, and their hydrophobic (from the greek "water-hating") tail in the oil. It is the need for the surfactants to sit at the interface that leads to the rapid removal of the oil droplet from the surface, so that more surfactant can adsorb on it. The entity such formed (the droplet coated with a layer of surfactant molecules) is called a **micelle**. This micelle will remain in suspension and not stick to the surface, as the hydrophilic heads prefer to remain in water and have no affinity for the surface. This is how soap is used to clean oily dishes or clothes. Note that shearing (using a brush or rubbing clothes) helps to detach the micelles from the surface.

Micelles do not have to be spherical. Their shape is dependent on the shape of the surfactant, its charge and its environment (pH, salinity, temperature…). For example, another extensively used surfactant in cleaning and hygiene products is **Sodium Dodecyl Sulfate** (SDS), which undergoes a transition from spherical micelle to rod-shape micelles upon increase of ionic strength:

spherical micelle   high ionic strength   rod-shaped micelle
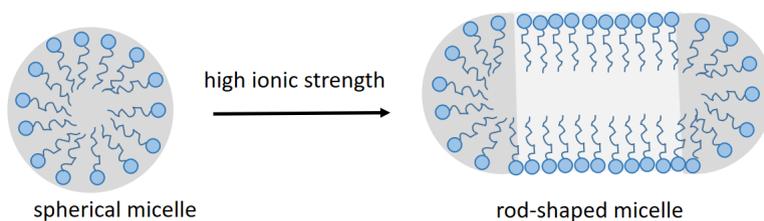
Structural (sphere-to-rod) transition in charged micelles (charges not shown) induced by high ionic strength. The structural transition takes place in charged micelles at concentrations well above the Critical Micelle Concentration (CMC, defined below). Note that the packing of the surfactant molecules is different in the spherical 'end caps' (darker shade) and the cylindrical central part (lighter shade). The spacing is reduced in the cylindrical part of the rod-shaped micelle due to attenuation of charge interactions at high ionic strength[54].

Surfactants in nature

In nature, **phospholipids** are extremely important amphiphilic molecules that form, in particular, the cell membrane. They can be part of the natural organic matter that is found in sediments. These molecules can arrange themselves in micelles, liposomes or bilayer sheets.

---

[54] from: A. Chaudhuri et al. / *Chemistry and Physics of Lipids 165 (2012) 497– 504*
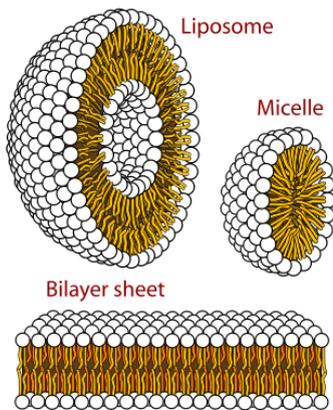
A **vesicle** is a large structure consisting of liquid enclosed by a lipid bilayer. A **liposome** is a spherical vesicle. It can be created by the disruption of biological membranes (which form bilayer sheets) by sonification. To maximize the contact of its hydrophilic head and water, a vesicle is then formed. Liposomes can be as big as 1000 nm (with a standard size between 20 and 1000 nm). Only some small mole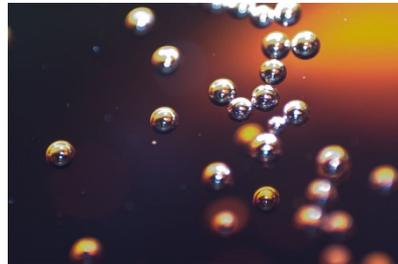cules can pass the bilayers of liposomes, and larger ones can be trapped inside. Liposomes are for example used to carry drugs which they subsequently deliver by fusing with the cell membranes.

Surface tension

Let us consider an **interface** between a liquid (water) and a gas (air).The cohesive forces between the water molecules are at the origin of the existing **surface tension**: in the bulk, each molecule is surrounded by water molecules but the molecules at the water/air interface do not have the same molecules on all sides. The molecules in the layer in contact with air experience other interactions than the molecules in the bulk, which results in a higher **surface energy** (J/m^2). This energy can also be expressed in a **surface tension** (N/m) by realising that 1 J = 1 N.m. A system will always try to minimize its surface energy, resulting in the fact that it will try to minimize its interface with the other **phase** (here one phase is water, the other air). This is why one finds many spherical, often colloidal, drops of liquid in gas (or gas in liquids, like the bubbles depicted here) in nature: a sphere is the volume that offers the smallest interface. This also explains why when two drops get into contact, they tend to form a larger drop (this process is called **coalescence**): again this is the smallest interface for a given volume:
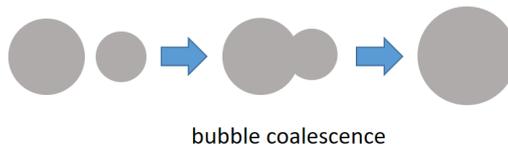
bubble coalescence

Measuring surface tension
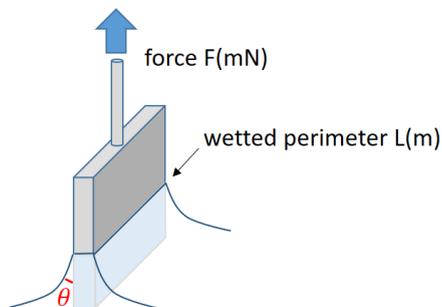
Principle of the Wilhelmy plate: the force needed to pull the plate out of the liquid is recorded; the perimeter of the plate is known : the surface tension can be estimated.

It is possible to measure a surface tension. One of the method is to use a **Wilhelmy plate**, which is a thin plate of a few square centimetres in area. The plate should be made in a material that is – obviously – able to be wetted by the liquid. The plate is pulled from the liquid, and the force is then measured (usually of the order of mN). The perimeter of the plate L is known: $L = 2d + 2b$ where d is the length of the plate and b its width. The surface tension γ is given by:

$$\gamma = \frac{F}{L\cos(\theta)}$$

In general the contact angle θ is unknown[55], but assumed to be close to zero, as there should be complete wetting. One then simply gets $\gamma = F/L$: the surface tension can be seen as the force, perpendicular to the interface, needed to deform this interface per unit of length.

Lowering the surface tension

When a surfactant is added in water in a clean jar (which therefore contains nothing else than water), the surfactant molecules tend to go to the water/air interface, where their hydrophobic tail can stick out of the water. This lowers the **surface tension** of the water, hence their name of surfactants. On the picture, an example of the evolution of the surface tension as function of concentration surfactant is given. Below a critical

---
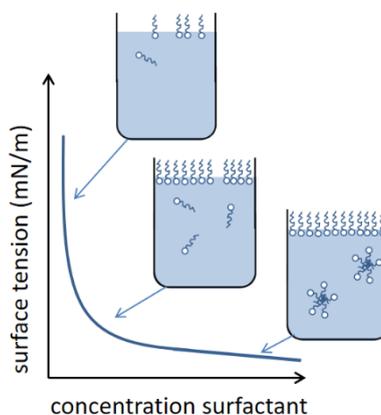
[55] Nowadays, if needed, cameras can record the angle with good accuracy.

concentration called **Critical Micelle Concentration** (**CMC**), the surfactant populate the interface. At the CMC, the interface is full, and micelles start to form: the surface tension does not vary much anymore, even though still some surfactant molecules will be able to pop in-between those who are already at the interface.

Both polymers and surfactants are important parameters to account for when clay aggregation is studied, but they can be used in other engineering applications as well. In petroleum engineering for example the oil production is increased when water containing surfactant is flushed in reservoir pores. By lowering the interfacial tension the oil mobility is increased thus allowing a better displacement of the oil by injected water. The addition of polymers on the other hand increases the viscosity of water injected into the oil reservoir enabling it to exert more pressure on the oil.

Biosurfactants

Besides phospholipids, there exist many more natural surfactants ("biosurfactants") produced by microorganisms from various substrates like sugars, oils and wastes[56]. These surfactants are synthesized as metabolic by-products , and their composition depends on pH, nutrient composition, substrate and temperature. These biosurfactants are usually either anionic or neutral. Only a few are cationic such as those containing amine groups. Most biosurfactant-producing organisms are aerobic, but a few anaerobic producers exist.

Biosurfactants are used for bioremediation of contaminated soils. For instance rhamnolipid surfactants were found to efficiently remove hydrocarbons from a sandy loam soil, and this was applied to the beaches in Alaska after the Exxon Valdez tanker spill. Standard surfactants, like sodium dodecyl sulfate (SDS), were found to be less effective than biosurfactants in removing hydrocarbons. Biosurfactants were also shown to better able than SDS to enhance the solubilisation of polycyclic aromatic carbons (PAH's). PAH's are uncharged, non-polar molecules found in coal and tar deposits. They are also produced by thermal decomposition of organic matter. The dominant source of PAH's in the environment comes from human activity by the combustion of biofuels and fossil fuels. Emissions from vehicles such as cars and trucks are a source of PAHs in particulate air pollution. Industrial activity such as aluminum, iron, and steel manufacturing, coal gasification and production of coke, tar distillation, shale oil extraction, road paving and asphalt manufacturing, rubber tire production can produce and distribute PAHs. Soil and river sediment near industrial sites can be highly contaminated with PAHs. PAHs have a strong affinity for organic carbon, and thus highly organic sediments in rivers, lakes, and the ocean can be a substantial sink for PAHs.

---

[56] Mulligan, C. N., R. N. Yong, and B. F. Gibbs. "Surfactant-enhanced remediation of contaminated soil: a review." Engineering geology 60.1-4 (2001): 371-380.

Plankton

Any water in a sea or lake that is neither close to the bottom nor near the shore can be said to be in the **pelagic zone**. The pelagic zone can be thought of in terms of an imaginary cylinder or water column that goes from the surface of the sea almost to the bottom. The pelagic zone can be contrasted with the **benthic** and **demersal zones** at the bottom of the sea. The benthic zone is the ecological region at the very bottom of the sea. It includes the sediment surface and some subsurface layers. Marine organisms living in this zone, such as clams and crabs, are called benthos. The demersal zone is just above the benthic zone. It can be significantly affected by the seabed and the life that lives there.

In the marine environment, polymers, polyelectrolytes, surfactants and humic acids are originating from microorganisms and plants. We here review some of the microorganisms present in the water body, which go under the generic name of "plankton". Some of these microorganisms have a colloidal size and can therefore experience colloidal interactions: algae for instance are known to aggregate between themselves.

Plankton (singular plankter) are a diverse group of organisms that live in the water column of large bodies of water and that cannot swim against a current. They provide a crucial source of food to many large aquatic organisms, such as fish and whales.

These organisms include drifting or floating bacteria, fungi, archaea, algae, protozoa and animals that inhabit, for example, the pelagic zone of oceans, seas, or bodies of fresh water. Though many planktonic species are microscopic (colloidal) in size, plankton includes organisms covering a wide range of sizes, including large organisms such as jellyfish.

Biologists have a peculiar way to define the size of plankton, as can be seen in the table underneath:

name	size	example
picoplankton	< 2μm	cyanobacteria
nanoplankton	2μm - 20μm	diatoms
microplankton	20μm - 200μm	dinoflagellates, diatoms
mesoplankton	200μm – 20mm	small molluscs
macroplankton	20mm – 20 cm	jellyfish, snails, krill
megaplankton	> 20 cm	jellyfish

Picoplankton and nanoplankton can be considered as colloidal particles, but it is important to remember that not only colloidal plankton but also larger planktons are able to secrete or decompose into smaller organic parts.

Phytoplankton

Phytoplankton are photosynthesizing microscopic organisms that inhabit the upper sunlit layer of almost all oceans and bodies of fresh water. They use solar energy to synthesize complex organic molecules from simpler inorganic compounds such as carbon dioxide (CO_2) and water (H_2O). A simple photosynthetic reaction is:

$$CO_2 + H_2O + light \rightarrow CH_2O + O_2$$

By this type of reaction carbon-based **polymers** $(CH_2O)_n$ are formed, typically molecules such as glucose or other sugars. These relatively simple molecules may be then used to further synthesise more complicated molecules, including proteins, complex carbohydrates, lipids, and nucleic acids (DNA, RNA). This happens when carbon-based molecules are eaten by living organisms (by small microorganisms, like bacteria, which in turn are eaten by larger organisms, or directly by larger organism, like fishes) that cannot themselves synthetize carbon[57].

The most important groups of phytoplankton include the diatoms, cyanobacteria and dinoflagellates.

Diatoms

Diatoms are algae with distinctive, transparent cell walls made of silicon dioxide hydrated with a small amount of water ($SiO_2 + H_2O$). Diatoms are abundant in nearly every habitat where water is found – oceans, lakes, streams, mosses, soils, even the bark of trees. These algae form part of the base of aquatic food webs in marine and freshwater habitats. Assemblages of diatom species are often specific to particular habitats and can be used to characterize those habitats. Nearly all diatoms are microscopic - cells range in size from about 2 microns to about 500 microns.

Planktonic diatoms in freshwater and marine environments typically exhibit a "boom and bust" (or "bloom and bust") lifestyle. When conditions in the upper mixed layer (nutrients and light) are favourable (as at the spring), their competitive edge and rapid growth rate enables them to dominate phytoplankton communities ("boom" or "bloom"). They can then also produce large quantities of polymeric slime (the freshwater diatom Didymosphenia geminate, for example, produces large quantities of a brown jelly-like material called "brown snot" or "rock snot").

---

[57] this type of organisms include humans, who rely on eating fruit, vegetables and meat to get their necessary carbon-based molecules.

Several species of fresh-water diatoms

When conditions turn unfavourable, usually upon depletion of nutrients, diatom cells typically increase in sinking rate and exit the upper mixed layer ("bust"). This sinking can for example be induced by a loss of buoyancy control or the synthesis of mucilage (= polymeric slime) that sticks diatoms cells together. Cells reaching deeper water or the shallow seafloor can then rest until conditions become more favourable again. In the open ocean, many sinking cells are lost to the deep, but refuge populations can persist near the thermocline[58].

In the open ocean, the diatom (spring) bloom is typically ended by a shortage of silicon. Unlike other minerals, the requirement for silicon is unique to diatoms and it is not regenerated in the plankton ecosystem as efficiently as, for instance, nitrogen or phosphorus nutrients. Because of this bloom-and-bust cycle, diatoms are believed to play a disproportionately important role in the export of carbon from oceanic surface waters. Significantly, they also play a key role in the regulation of the biogeochemical cycle of silicon in the modern ocean.

---

[58] The thermocline divides the warmer upper layer from the cooler (and calm) deep water below.

Cyanobacteria

Cyanobacteria are a group of photosynthetic, nitrogen fixing bacteria that live in a wide variety of habitats such as moist soils and in water. They may be free-living or form symbiotic relationships with plants or with lichen-forming fungi. They range from unicellular to filamentous and include colonial species. Colonies may form filaments, sheets, or even hollow balls.
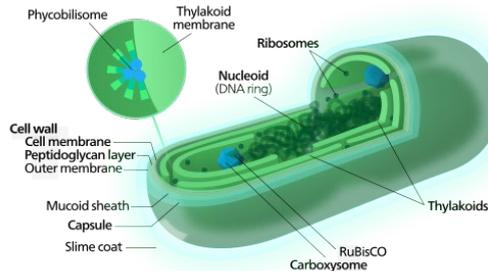
a typical cyanobacteria cell

Aquatic cyanobacteria are known for their extensive and highly visible blooms that can form in both freshwater and marine environments. The blooms can have the appearance of blue-green paint or scum. These blooms can be toxic, and frequently lead to the closure of recreational waters when spotted. Marine bacteriophages (which are viruses) are significant parasites of unicellular marine cyanobacteria.

Dinoflagellates

The dinoflagellates are marine plankton but can be found in freshwater habitats as well. Their populations are distributed depending on temperature, salinity or depth. Many dinoflagellates are known to be photosynthetic, but a large fraction combine photosynthesis with ingestion of prey. Dinoflagellates are unicellular and possess two dissimilar flagella arising from the ventral cell side. The flagellar movement produces forward propulsion and also a turning force.

A bloom of dinoflagellates can result in a visible coloration of the water colloquially known as red tide, which can cause shellfish poisoning if humans consume contaminated shellfish.

Dinophysis acuminata

20 µm
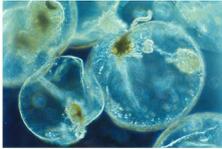
Noctiluca scintillans

Algal bloom by Noctiluca in Nagasaki

Left: photographs of dinoflagellates; Right: algal bloom

Zooplankton

Zooplankton are organisms drifting in water (oceans, seas or fresh water) that are usually microscopic, but some (such as jellyfish) are larger and visible with the naked eye. Many zooplankton have locomotion, used to avoid predators or to increase prey encounter rate. Zooplankton feed on bacterioplankton (the bacterial component of the plankton), phytoplankton, other zooplankton or dead organic material (detritus). Zooplankton are therefore primarily found in surface waters where food is abundant.

Some zooplankton have a symbiotic relationship with algae. *Paramecium bursaria* for instance has such a relation with the green algae called *Zoochorella*. The algae lives inside the cytoplasm of *Paramecium* and provides it with food while *Paramecium* provides the algae with movement and protection. *Paramecium* is larger than a colloidal particle, being 80-150 µm long. Other zooplankton can be much smaller.


Paramecium bursaria

The marine cycle

When considering colloids in natural environment, it is important to realize the realm of important (variable) parameters that has to be considered in order to correctly model their behaviour. For instance, predicting suspended sediment behaviour requires not only to be able to know physical parameters such as shear stresses, turbulent mixing conditions, erosion parameters, or chemical parameters (salinity, pH, presence of polymeric substances), but also to understand – at some extend - the role of biology and bio-chemistry: When and why does EPS become available? How does organic decomposition change the properties of the flocs? Is plankton able to bind to sediment particles? Can minerals, like silica, have a biological origin?
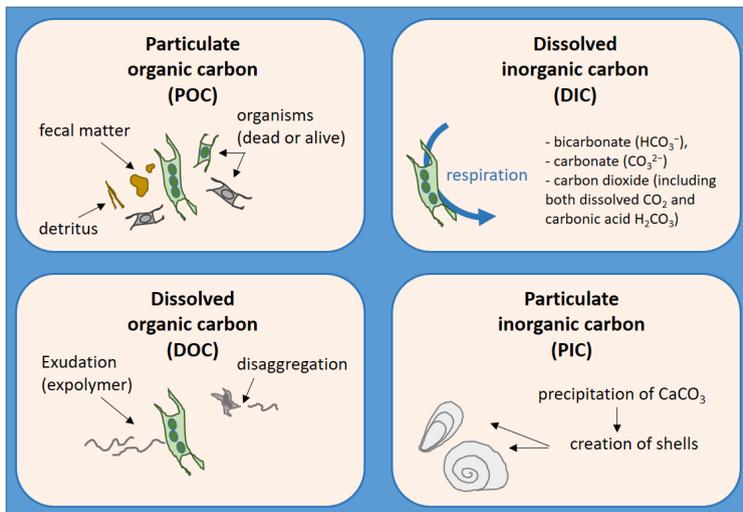
The oceanic carbon cycle (or marine carbon cycle) is composed of processes that exchange carbon between various pools within the ocean and between the atmosphere. There are four distinct carbon pools (POC, DIC, PIC and DOC). Phytoplankton are responsible for most of the transfer of carbon dioxide from the atmosphere to the ocean. Carbon dioxide is consumed during photosynthesis, and the carbon is incorporated in the phytoplankton. Most of the carbon is returned to near-surface waters when phytoplankton are eaten or decompose, but some falls into the ocean depths.

A part of the sediment (silica) found in the marine environment (as part or not of marine snow) comes from the decomposition of microorganisms:
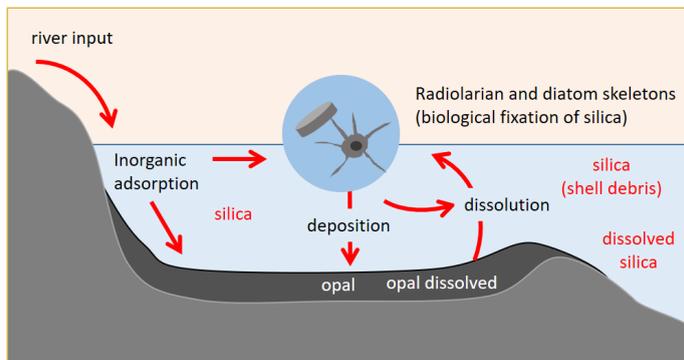
Cycling of silica in the marine environment: Silicon commonly occurs in nature as silicon dioxide (SiO₂), also called silica. It cycles through the marine environment, entering primarily through riverine runoff. Silica is removed from the ocean by organisms such as diatoms and radiolarians (a type of zooplankton) that use an amorphous form of silica in their cell walls. After they die, their skeletons settle through the water column and the silica redissolves. A small number reach the ocean floor, where they either remain, forming a silaceous ooze, or dissolve and are returned to the water column. Opal is a hydrated amorphous form of silica (SiO₂nH₂O)

Many suspended particles in the marine environment are aggregates of particles of different origin. **Marine snow** is a type of aggregate made up of a variety of organic matter, including dead and living phytoplankton, bacteria, fecal matter, sand, and other inorganic dust.

A majority of marine snow composition is actually made up of aggregates of smaller particles held together by a sugary mucus, transparent extracellular polysaccharides (TEP). These are natural polymers exuded as waste products mostly by phytoplankton and bacteria, but also by zooplankton. These aggregates grow over time and may reach several centimetres in diameter, traveling for weeks before reaching the ocean floor. Phytoplankton, microorganisms and bacteria live attached to aggregate surfaces and are involved in rapid nutrient recycling.

The composition of the flocs is also depending on the seasonal variations, linked to the natural lifecycles of the microorganisms involved. In general, most detritus linked to mineral particles are found in estuaries or higher up in the river stream. In winter, they form the most abundant type of aggregates in the whole system (from the sea to up in the river).
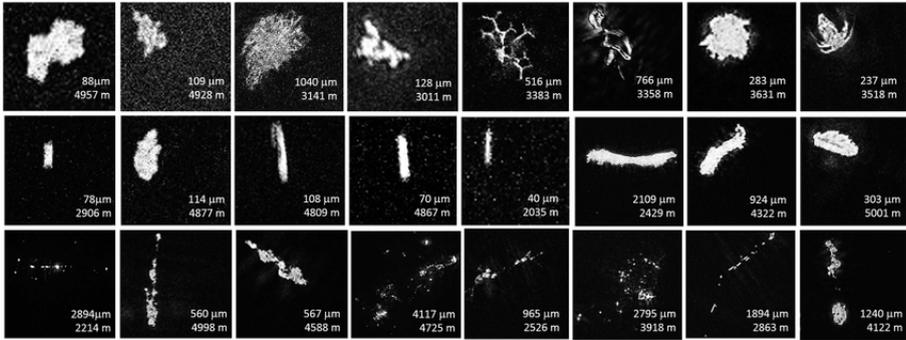
marine snow (bottom row), faecal pellets (centre row) and "others" (top row). The "others" category includes all recognizable planktonic organisms (alive and carcasses) and optically dense debris that does not classify as marine snow or faecal pellets. For each image, the size (μm) and depth sampled (m) are given[59].

Illustrations

Biofilm (public domain)
https://en.wikipedia.org/wiki/Biofilm#/media/File:Staphylococcus_aureus_biofilm_01.jpg

Lascaux (public domain)
https://fr.wikipedia.org/wiki/Grotte_de_Lascaux#/media/Fichier:Lascaux2.jpg

Phospholipides (public domain)
https://commons.wikimedia.org/wiki/File:Phospholipids_aqueous_solution_structures.svg

Micelles (public domain)
https://en.wikipedia.org/wiki/Micelle

Plankton (public domain)
https://en.wikipedia.org/wiki/Plankton

Diatoms (public domain)
https://en.wikipedia.org/wiki/Diatom
https://en.wikipedia.org/wiki/Diatom#/media/File:Diatomeas_w.jpg

Cyanobacteria (creative commons)
https://en.wikipedia.org/wiki/Cyanobacteria

---

[59] Bochdansky, Alexander B., Melissa A. Clouse, and Gerhard J. Herndl. "Dragon kings of the deep sea: marine particles deviate markedly from the common number-size spectrum." Scientific reports 6 (2016): 22633.

Dinoflagellates (creative commons)
https://en.wikipedia.org/wiki/Dinoflagellate

Paramecium_bursaria (creative commons)
https://en.wikipedia.org/wiki/Paramecium_bursaria

Chapter 5

Floc formation and break-up

Given certain conditions (see for instance DLVO theory, **Chapter 3**), colloidal particles can "glue" (aggregate) to each other, hereby creating a bigger entity called "**floc**" that is able to settle down. The settling velocity and size of flocs is recorded in-situ or in the lab, from which their density is evaluated using Stokes' settling velocity (defined in **Chapter 2**). In the present chapter we are going to see that the density of flocs is a complex function of their size, related to their structures and composition. We discuss why these structures are found. We also discuss the effect of shear on flocs as well as **delamination** and **swelling**, which are processes that break the structures of mineral clay aggregates.

Cluster aggregation

Before forming large clusters, particles first have to form small aggregates (mainly **doublets** at early stages), then they grow larger and larger:

| stable suspension | early stage of aggregation | later stage of aggregation |
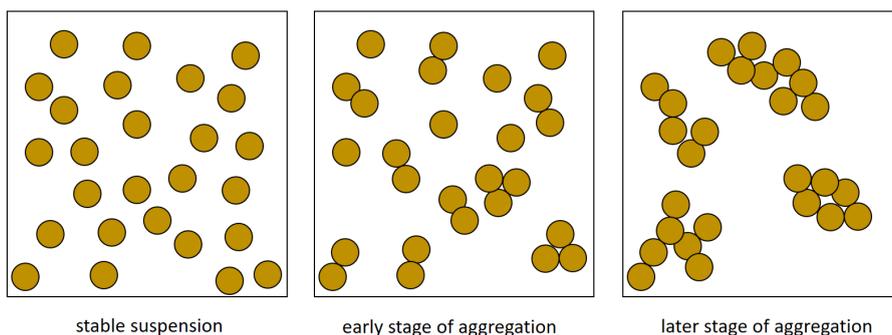
Two important aggregation mechanisms are linked to the sticking probability of two particles:

DLCA (**Diffusion Limited Cluster Aggregation**): the particles stick at first contact (ex: can they immediately get in the "primary minimum" defined by the DLVO theory[60], as there is a very low energy barrier)

RLCA (**Reaction Limited Cluster Aggregation**): the particles need to position themselves in a comfortable way to stay attached (ex: there is a substantial energy barrier to be overcome first)

The results of these two types of aggregation lead to different **floc** structure:

---

[60] See **Chapter 3**
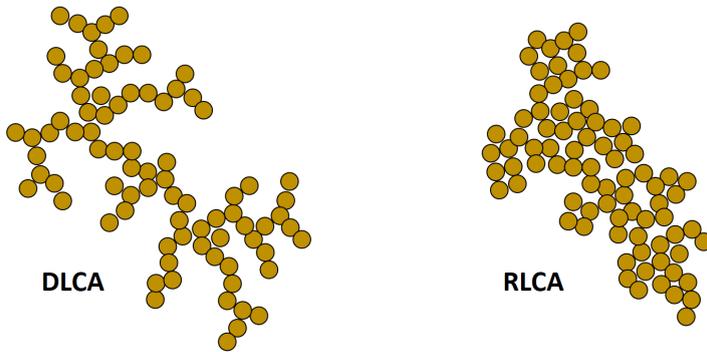
From the figures it is evident that the density of these two types of flocs is different: RLCA flocs are denser (i.e. they contain less water) than DLCA flocs. We will see later that this can be quantified in terms of a (pseudo) **fractal dimension**.

In the previous pictures, spherical particles are shown, as this geometry is the easiest to investigate from a theoretical point of view. In nature, most of colloidal particles are not spherical. Clay minerals in particular are highly anisotropic: they can have the shape of platelets (used in the examples underneath), but some are also cylindrical-shaped:
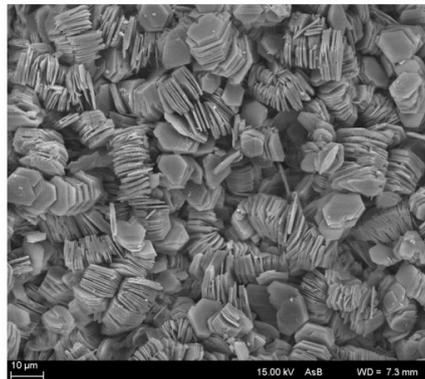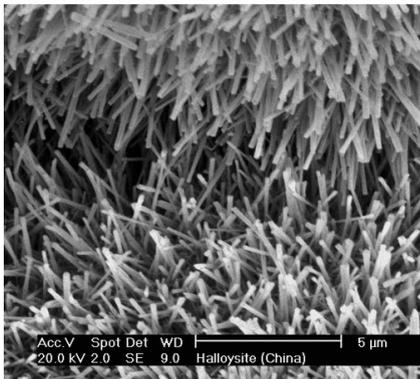
Shape of microscopic clay particles. Left: halloysite and right: kaolinite

Depending on the clay mineral type, clay particles will disperse in different way in water – and also flocculate in a different way.

Stable clay suspensions

Here we show two examples, one of a non-swelling and one of a swelling clay, kaolinite and montmorillonite respectively, as they are the most encountered in engineering, and natural environments. Kaolinite is used in ceramics (it is the main

component of porcelain), in toothpaste, in paint, as adsorbents in water and wastewater treatment. Montmorillonite is a component of drilling mud, used in the oil-drilling industry, as it makes the mud slurry viscous. Its swelling property makes montmorillonite-containing bentonite useful as a seal for water wells and as a protective liner for landfills. Montmorillonite has also been used in cosmetics and sodium montmorillonite is used as the base of some cat litter products, due to its adsorbing and clumping properties.

montmorillonite

Montmorillonite clays (part of the so-called swelling clays) can usually disperse easily in water as their layers are only weakly bound together. They then form a suspension of very small platelets (usually 1 nm thick and 100 nm long). In the illustration, it is said that the layers are held together by van der Waals forces, but this point is controversial. It is more generally assumed that electrostatic attraction is the dominant force.

kaolinite
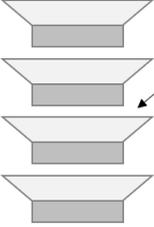
Kaolinite clays (part of the so-called non-swelling clays) on the other hand also disperse in water, but as their layers are tightly bound together by hydrogen bonds, they disperse in the form of stacks of platelets. These stacks are usually not so anisotropic in shape, and can be quite large compared to montmorillonite clays: something between 1 and 10 micrometers in diameter.

Until now, we illustrated the flocs by represented clay particles bound together. This type of flocs can be produced in an electrolyte solution. Another type of flocculation is possible which involves the presence of polyelectrolytes or polymer and microorganisms, see **Chapter 4**. In an electrolyte, different effects can occur, called

delamination and swelling, which lead to the breakage of the aggregates. Aggregates can also break and re-conform due to shear as discussed below.

Unstable clay suspensions: influence of salt

The flocculation by salt (coagulation) of kaolinite and montmorillonite has been studied for decades, given their importance in practical applications. Flocculation of mixtures of clays have also been studied[61]. As we have seen in **Chapter 3**, the amount of ions present in the water is of extreme importance: at moderate ionic strength[62], the particles can undergo a "secondary minimum" aggregation, whereas at high ionic strength the particles can be strongly aggregated, thanks to van der Waals forces.

In the case of montmorillonite clays, there is a complication in predicting the role of ions. As montmorillonite particles can swell, when ions are added to the suspension, some will penetrate the montmorillonite interlayers. An important question is then: how does these ions alter the structure of the montmorillonite particle? Will the particle remain a (swollen) stack of platelets, will the stack be **delaminated**?
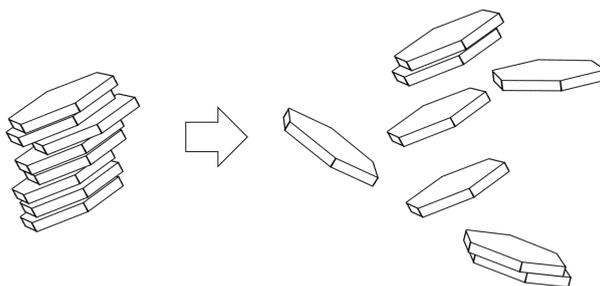
Delamination of a stack of montmorillonite

The answer is: usually the stack will delaminate for small amount of added ions, and this is why most montmorillonite particles are usually found as single platelets in water. The delamination process is related to the fact that by adding a small amount of ions, these ions (and related water) will penetrate the interlayer space. The platelets forming the stack will therefore be "pushed" away from the secondary minimum where they were residing, and undergo Coulombic repulsion. Note that this does not happen for non-swelling types of clays (like kaolinite): there, the platelets are more strongly bound thanks to hydrogen bonds. By adding substantially more ions in the system, a second important effect will take place: the screening of

---

[61] For extensive details, we refer to Bergaya, Faïza, and Gerhard Lagaly. Handbook of clay science. Vol. 5. Newnes, 2013.
[62] ionic strength is the term used in physical chemistry. In civil engineering one usually speaks of salinity or salt concentration.

the surface charge of each platelet. This will lead to a new aggregation (in the secondary or primary minimum, depending on ionic strength):
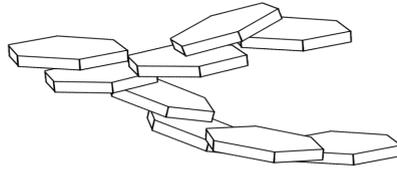
Band-type aggregation: small amount of calcium added to a suspension of delaminated montmorillonite leads to face-face aggregation. Usually the aggregates are then found in band-type.

An important parameter that should not be forgotten while studying this type of flocculation is pH (i.e. the concentration of H^+ or OH^- ions present in the water). Contrary to most ions, H^+ and OH^- ions can interact chemically with the clay's surface and hereby change its surface charge. An example has been given in **Chapter 1** in the case of kaolinite. When the electric charge of the faces of the clay particles is of different sign than their edges a strong Coulombic attraction between the positive edge of one particle and the negative face of another will occur. The pH, for most clays, where a change in surface charge sign occurs, is between 4 and 6. Depending on ionic strength and pH, different aggregate structures can therefore be obtained.

Some examples of clay aggregates
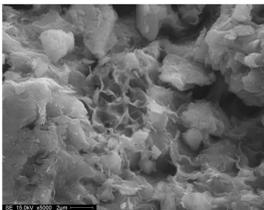
hectorite
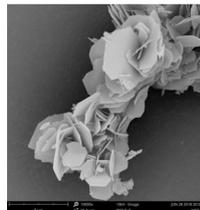
muscovite
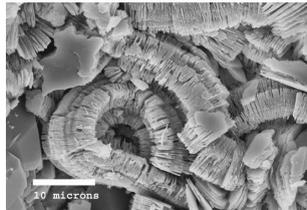
kaolinite
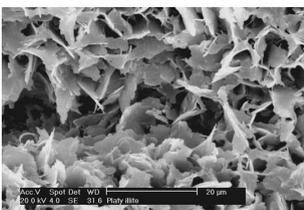
illite
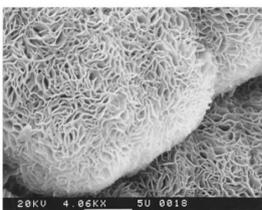
smectite

gibbsite

These TEM images have been obtained on samples for which the pH and salinity has not been specified. Several structures can usually be identified, depending on pH and ionic strength[63]:

	Delaminated particles with moderate/large addition of salt, at pH > 8: each particle is purely negatively charged. This charge is screened sufficiently to induced flocculation. A preferred order (face/face) is required, to maximize the van der Waals forces between particles.
	Delaminated particles with low/moderate addition of salt, at pH < 4: the positive edge with one particle will bind to the negative face to another. If the salt concentration would be high, face/face aggregation would also be possible, and the structure would be a mixture between this one and the one shown above (and hence be more compact).
	Non-delaminated particles with low/moderate addition of salt, at pH < 4: the positive edge with one particle will bind to the negative face to another. If the salt concentration would be high, face/face aggregation would also be possible, leading to a more compact structure.
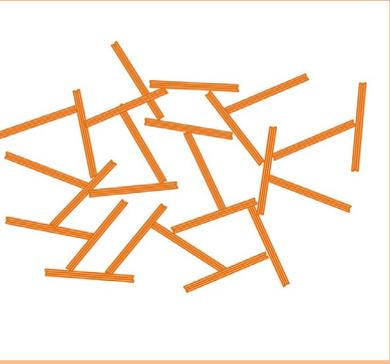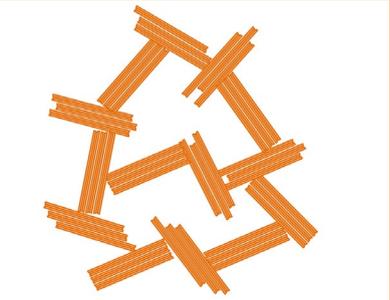
[63] O'Brien, Neal R. "Fabric of kaolinite and illite floccules." Clays and Clay Minerals 19.6 (1971): 353-359.

Swelling behaviour of clays

Swelling clays (like montmorillonite) do not always delaminate. One important reason is related to the concentration of particles (or in other word: how much water is added to the clay). If the water content is high, the montmorillonite becomes a **thixotropic**[64] gel, and at even higher water content it becomes a suspension (also called a **sol**), where the particles can delaminate. Here we are concerned with a water content such that the clay is a very thick paste, obtained from putting the dry clay in water. Na-montmorillonite (i.e. montmorillonite where the interlayer cations are Na$^+$), when placed in water, can take up 10 g of water per g of clay, and its volume can increase by about 20 times. In **Chapter 2**, we have already addressed the concept of **osmotic pressure**. This plays an important role is the swelling of clays.
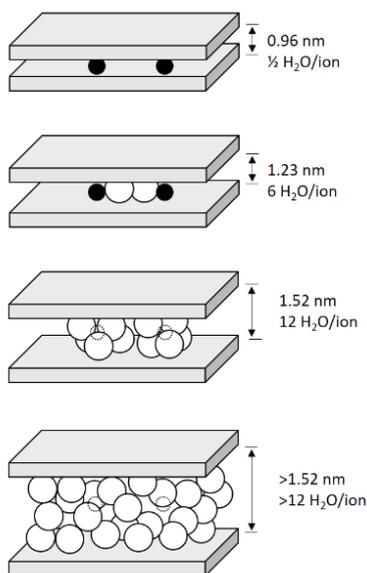
0.96 nm
½ H$_2$O/ion

1.23 nm
6 H$_2$O/ion

1.52 nm
12 H$_2$O/ion

>1.52 nm
>12 H$_2$O/ion

In montmorillonite and vermiculite the aluminosilicate sheets are separated by water layers whose thickness varies with the concentration and type of electrolyte. In the crystalline state, these extremely thin, negatively charged sheets are held together by electrostatic forces between alternate layers of bridging cations (like Na$^+$, K$^+$, Ca^{2+}). When the clay is placed in water or a dilute electrolyte suspension, some of the adsorbed counterions tend to take up water and may dissociate from the clay surface, creating a diffuse double layer. It is the repulsive interaction between these double layers that is at the origin of the swelling behaviour. On the schematic representation on the left, innercrystalline swelling of Na-montmorillonite is displayed. The layer distances and the maximum number of water molecules per sodium ion are given. The first water to enter the interlayer positions is the result of hydration of the ions, after which the water forms distinct layers which increase in number to four. The water molecules, of at least the first layers, are probably arranged in a hexagonal network whose order is determined by **hydrogen bonding** to the surface oxygens of the clay[65].

---

[64] See Chapter 7
[65] Norrish, K. (1954). The swelling of montmorillonite. Discussions of the Faraday society, 18, 120-134; Madsen, Fritz T., and Max Müller-Vonmoos. "The swelling behaviour of clays." Applied Clay Science 4.2 (1989): 143-156.

The swelling discussed so far is commonly referred to as **crystalline swelling**. In this range, the adsorbed water increases to about 0.5 g water per g of clay while the interlayer spacing increases from 0.96 nm, for dry material, to about 2 nm corresponding to 4 layers of water. This water content is the one that is adsorbed at 99% relative humidity and at this stage the montmorillonite has shown little physical swelling.

When this Na-montmorillonite is placed in contact with water, and increased about 20 times its original volume, one speaks of regular (**osmotic**) **swelling** of the clay (see last picture of the sketch above). Unlike innercrystalline swelling, which acts over small distances (up to 2 nm), osmotic swelling is based on the repulsion between overlapping electric double layers, and can act over much larger distances. (In Na-montmorillonite it can result in the complete separation of the layers, leading to delamination, if the water content allows it). The expansion is associated with the formation of these diffuse double layers, as more water penetrates the interlayer space: the electrostatic attractive force that existed between the cations and the negative surfaces of the interlayers is now changed into a (osmotic) repulsive one.
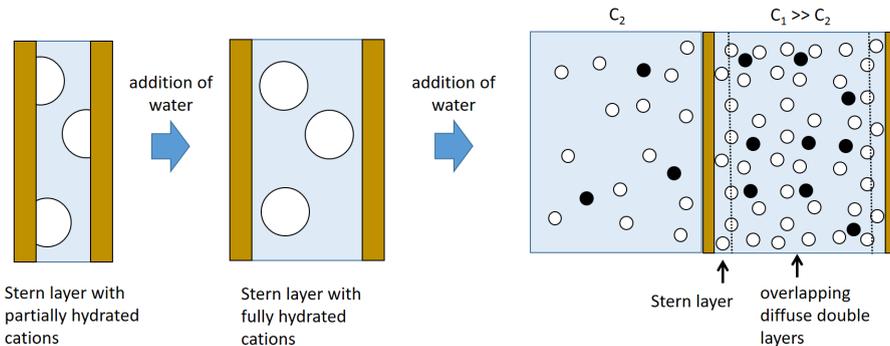
Before osmotic swelling, cations and their hydration shells (white balls) are electrostatically bound to the clay platelets. After water is added to the clay, double layers start to form with both anions and cations (white and black balls), leading rapidly to an electrostatic repulsion between the clay platelets, since the ion concentration C_1 between the layers is much higher than the ion concentration C_2 in the bulk water. An equilibration of the concentration can only be reached through the penetration of water into the space between clay layers (osmotic swelling). If the water content allows, this will eventually lead to delamination.

The repulsive force between the overlapping double layers of the clay particles also exists inside some geological formations. This force is in equilibrium with the overburden pressure coming from the mass above the double layers. If the load on them is removed and if water is available, a new equilibrium is sought: the water intrudes between the clay layers and pushes them apart. The swelling continues until the new balance between this inner force and outside, resisting forces is

reached. This phenomenon results in landslides and is further discussed in **Chapter 7**.

Norrish (1954) demonstrated that, when small crystals of montmorillonite were arranged in a parallel orientation and allowed to equilibrate with a dilute electrolyte suspension (10 mM < C < 250 mM), then the separation between the sheets (d) was given by:

$$d \ (nm) = 2.1 + \frac{1.14}{\sqrt{C(M)}}$$

[for 10 mM the separation is 13.5 nm and it is 4.38 nm for 250 mM]. It is logical that the separation decreases with increasing ionic strength, as the double layers are then more compressed and the concentration in the interlayer space, C_1, is then not too different from the concentration in the bulk, C_2.

Unstable clay suspensions: flocculation by polyelectrolytes

In the previous chapter, we have defined **polyelectrolytes**. Polyelectrolytes (flocculants) can attach themselves to clay particles and form large, elastic flocs. The structure of sediment-polyelectrolyte flocs is very complex and is still an on-going topic of research. In applied research, the open questions are related to the efficiency of polyelectrolytes as binders of colloidal particles, and to the strength of the obtained flocs. In particular, the floc size and density evolution as function of shear rate are important parameters to investigate. This is discussed in the next section.

Shear rate influence on floc size

The way the clay particles flocculate is also strongly dependent on the shear rate: at low shears, flocculation is usually promoted, whereas at high shear, flocs can be broken (or prevented to grow further). In **Chapter 6**, we are going to show how to model the growth and break-up of flocs with shear, and we will concentrate on the case where only aggregation is occurring. Here we simply give and discuss some measurements results, and we will pinpoint the limitations of the current models and measurement techniques.

The influence of shear rate on flocculation can best be studied using static light scattering (see **Chapter 2**). The set-up is schematized here:
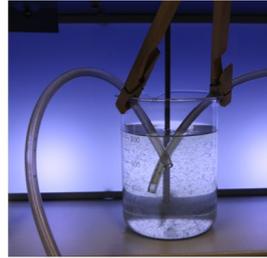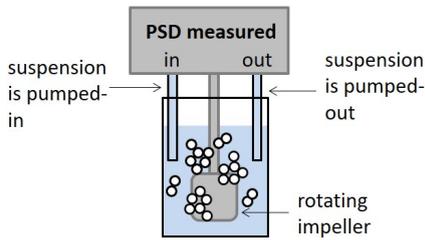
The set-up consists of a 1L jar connected to two tubes through which the suspension is pumped in and out the static light scattering device. To keep the suspension suspended it is gently stirred by a rotating impeller. The particle size distribution (PSD) is measured.

From the commercial instrument software, the raw data is converted into equivalent volumes, using the assumption that the particles are spherical. For each size (there are usually 100 given sizes, ranging logarithmically between 0.1 and 1000 µm in most commercial devices), a volume is evaluated that is equal to the volumes of all particles of that size added up. Each Particle Size Distribution (PSD in short) curves is normalized such that the integral over each curve gives 100%. This means that the volume corresponding to one size is a percentage volume. In other words, PSD measurements give the relative ratio between volumes, and changing the concentration of particles (adding more of the same) will not change the PSD. The PSD will however change if flocculation or break-up occur in the suspension as function of time, as each measured PSD will then not have the same volume ratios.

Underneath we give an example obtained from measuring the change in PSD over time of a suspension of clay in presence of a cationic polyelectrolyte.
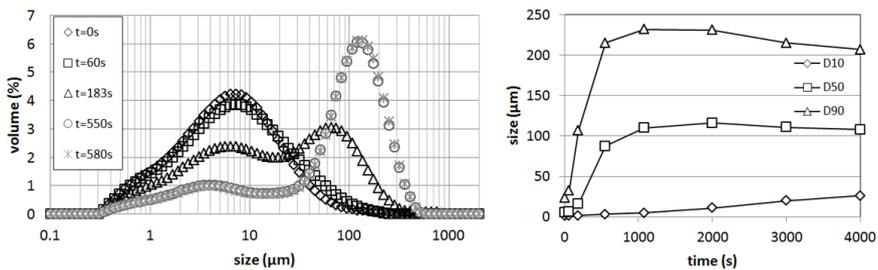
At t = 30 s, 0.7 g/L of clay is mixed with cationic flocculant. The ratio flocculant to clay is 0.5 mg/g. The full PSD is recorded every 30 s by static light scattering. The change in PSD over time is given in the left figure. The D10, D50 and D90 of the distributions are given in the right figure.

In the example given above we start at t = 0 with a clay suspension, which average mean size D50 is 6 µm. This mean size D50 is defined as the size for which 50% of

the particles are smaller (here in volume) than this D50 size. Similarly one defines the sizes D10 and D90 as the sizes for which respectively 10% and 90% of particles are smaller than these sizes.

At t = 30 s, cationic flocculant is added to the clay suspension. The PSD is evolving in time: as can be seen on the left figure, a new peak in size is appearing, corresponding to the creation of aggregates, and this peak is shifting to higher sizes over time. This is reflected in the changes in D50 over time (right figure). After 2000 s, an equilibrium size seems to have been reached. However, one can observed that after 2000 s there is a slight decrease in D50 over time. At the same time there is a larger decrease in D90 and increase in D10. This implies that the largest particles are eroding over time, creating small particles. This erosion of particles over time is usually not accounted for in traditional flocculation models (see **Chapter 6**).

It is important to note that there are large differences between flocs measured in-situ and the ones created in the jar. First of all, the shear rates generated in the jar and especially in the tubes leading to the measurement device are much higher generally than the ones encountered in-situ. As the tubes are about a few mm in diameter, this limits the growth of flocs and will also be a cause for erosion.

Clay in presence of salt

In the example given below[66], a large amount of NaCl salt has been added to a suspension of polystyrene latex particles (10 μm in diameter and with a density of 1055 g/L) and a suspension of silica particles (10 μm in diameter and with a density of 2000 g/L) to ensure that each suspension is unstable and the particles are aggregating. For each suspension and each shear, the mean equilibrium size of the suspension, D50, is plotted. This mean size is obtained after the PSD does not evolve any more in time.

The shear rate is varied by changing the rotational speed of the impeller. For this experiment it was verified that the flocs were not significantly broken/eroded in the tubes. The pumping speed was chosen as low as possible to prevent flocs settling in the tubes. The Kolmogorov microscale (defined in **Chapter 6**), reflecting the size at which the turbulence generated in the jar could disrupt a floc, is also plotted in the figure.

---

[66] Mietta, F., C. Chassagne, and J. C. Winterwerp. "Shear-induced flocculation of a suspension of kaolinite as function of pH and salt concentration." Journal of colloid and interface science 336.1 (2009): 134-141.
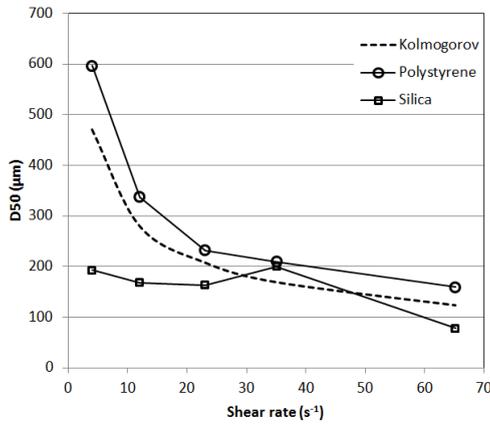
Mean floc size as function of shear rate, for two types of particles. Note that in-situ typical shear rates are less than 10 s^{-1}. Larger shear rates are however encountered in industry.

We see that this size is an important parameter, as the polystyrene particles follow this microscale, indicating that they cannot grow much larger than this size. Below the Kolmogorov microscale, the floc is simply transported (advected) with the flow whereas above the Kolmogorov microscale shear forces are tearing the floc apart. Shear forces are discussed in **Chapter 7**.

The silica particles behave differently. At low shear rate the silica particles seem to grow smaller than at moderate shear rates. This is however only an artefact. At low shears, large (heavy) silica particles are settling in the measuring jar, and are therefore not recorded by the static light scattering device used to measure their size: only the smallest particles remain in suspension. Polystyrene particles, having a density close to the one of water, are always fully suspended independently of their size.
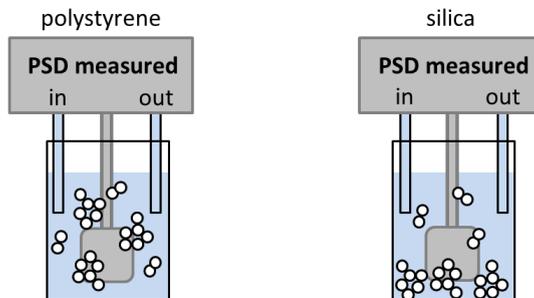
Each suspension is gently stirred in the jar, and the sample is pumped through the static light scattering device, where the D50 is recorded; at low shear the largest flocs made of silica particles are setting in the jar and are therefore not recorded.

Clay in presence of polyelectrolyte

Contrary to flocs created by salt, flocs created by polyelectrolytes can follow or not the Kolmogorov microscale. In the example underneath, the suspensions were made in demi-water (water containing a very low amount of ions). It was verified (not shown) that flocs created by anionic flocculant are always much smaller than flocs created by cationic flocculant, for all flocculant to clay ratios.
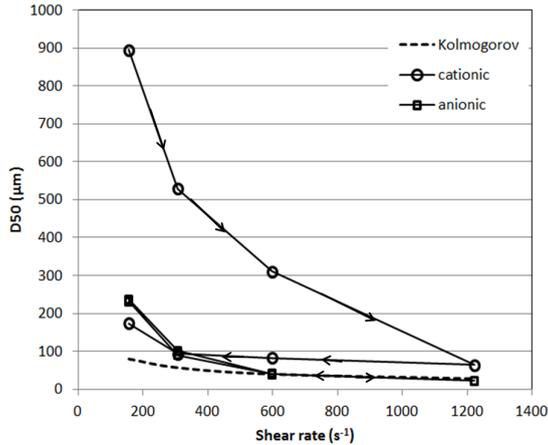
Equilibrium D50 for a suspension of clay mixed with (a) 10 mg/g Zetag 7587 (cationic flocculant) to clay ratio and (b) 0.5 mg/g Zetag 4110 (anionic flocculant) to clay ratio. The arrows indicate that each suspension was created at low shear, the shear was then increased and lowered again.

This is due to the fact that anionic flocculant needs cations to bridge with the clay (see **Chapter 4**) and there are not sufficient cations in the water to have an optimal flocculation. Flocs produced by anionic flocculant follow the Kolmogorov microscale quite well, for all shears, and the flocs that have been broken at high shear regrow to the size they had before the high shearing for all shear rates.
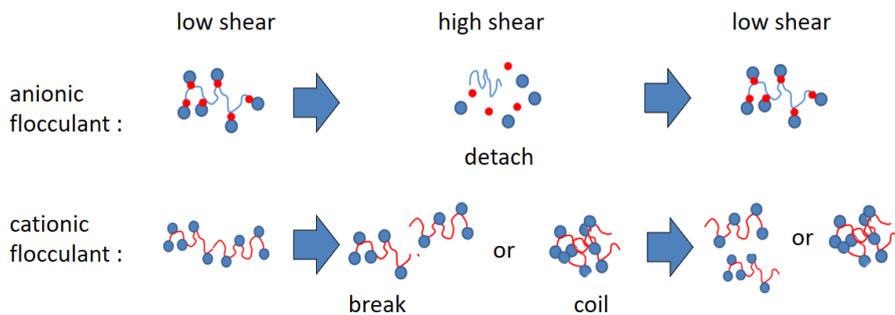
The interaction between the polymer and the clay is dependent on the bridging cation. Under the action of shear, when the Kolmogorov length is reached by the floc, the cation will not be able to bridge them anymore, and clay and polyelectrolyte will move independently. At sufficient low shear, the floc will be reformed.

Flocs produced by cationic flocculant, on the other hand are created thanks to attractive Coulombic interactions between the negatively charged clay and the positively charged polymer. Due to their low density, the large flocs (the cationic flocs grow nearly as large as 1 mm) remain in suspension even at low shears. When the shear is increasing the D50 is decreasing, but the D50 remain well above the Kolmogorov microscale. When the shear is lowered again, it is observed that the D50 barely increases in size. The kinetics of floc formation (at 150 s^{-1}) when the suspension is made at this low shear is very different from the small change in size when the shear is decreased from 300 s^{-1} to 150 s^{-1} after that the floc has been disrupted at higher shear:
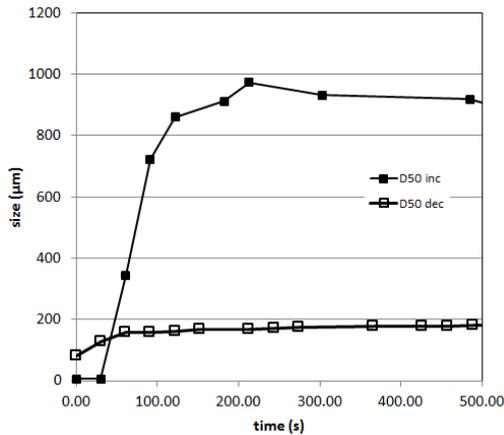
Change of D50 size over time at a shear rate of 150 s^{-1}. (D50 inc): formation of large flocs when clay and polyelectrolyte are mixed. The time to reach an equilibrium D50 is about 150-200 s. (D50 dec): flocs formed when the shear is decreased from the previous shear step at 300 s^{-1}. Only a small change in size is observed, and the change in D50 occurs is less than 100 s.

One see that flocs created at low shear will never regrow fully to their original large size after having experienced a higher shear. There can be two reasons for this: (a) the flocs broken at high shear are weakly positively charged and will experience a mutual repulsion and (b) due to the high shear the loops and tails (see **Chapter 4**) of the polyelectrolyte will be able to collapse and attach to any remaining exposed clay particle in their floc matrix. This will reduce the floc size and increase its density.

Only at low shear will there be a small re-flocculation, which might be due to a steric entanglement of flocs whose ends are able to decoil at low shears.

Density of flocs and fractal dimensions

Since flocs are usually non-spherical, permeable to water and can be composed of particles of different refractive indexes, this poses a problem for the in-situ recording of their size by **light scattering** (see **Chapter 2**), as the conversion of the raw data depend on a model which assumes that particles are spherical and have a given refractive index. Thanks to the development of camera's in the past years, it is nowadays possible to estimate the size of flocs and their settling velocity by video recording. From Stokes, it has then been found that the density of a floc is a decreasing function of its size, and this has led scientists to conveniently model them as if they were fractal objects. As flocs, especially the ones created in-situ have a very heterogeneous composition they are *not* "real" fractals as their structure is not **self-similar**.

Fractals and self-similarity

At the beginning of the chapter, we have evoked the term of **fractal**. This term stems from its inventor, B.B. **Mandelbrot** who introduced it in the late 1970s. In 1967 he wrote a famous article entitled "How Long Is the Coast of Britain?"[67] where he noted that the length of the coast depends on the scale at which it is measured. In fact, it was empirically found that the measured length of a coast could be estimated by:

$$L(\lambda) = M\lambda^{1-D}$$

where λ is the measurement scale, L(λ) the length of the coast, M a positive constant and D another constant, greater than or equal to 1. If D = 1, the length is independent of the measurement scale ($L = M$). It has been found that D is ranging from 1.02 for the coastline of South Africa to 1.25 for the West coast of Britain. Mandelbrot in the paper introduces the concept of statistical **self-similarity** and **fractional dimension** that will enable him to develop his concept of fractals in later publications.

Self-Similarity

Here we give an example showing how to construct a real self-similar fractal object. The one given here is called Koch snowflake:

---

[67] Mandelbrot, B. B. (1967). How long is the coast of Britain. Science, 156(3775), 636-638.
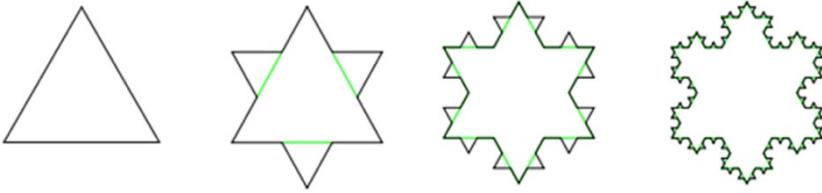
If one zoom-in on such a flake, one will realize that the structure is **self-similar**: at various scales, the shape of the interface will be the same, and indeed the perimeter of the fractal object will depend on the measurement scale used (the first triangle has obviously a smaller perimeter than the floc on the right).

L-system and the growth of floc-like structures

So-called L-systems were introduced and developed in 1968 by Aristid Lindenmayer, a Hungarian theoretical biologist and botanist at the University of Utrecht. A simple example of L-system is as follows, and was introduced to understand the growth of algae:

variables: A, B
rules: (A →AB), (B→A)

Starting with variable A and applying the rules, one obtains:

n = 0 : A

n = 1 : AB [as A →AB]

n = 2 : ABA [as A →AB and B→A]

n = 3 : ABAAB [as A →AB , B→A and A →AB]

n = 4 : ABAABABA

n = 5 : ABAABABAABAAB

n = 6 : ABAABABAABAABABAABABA

n = 7 : ABAABABAABAABABAABABAABAABABAABAAB
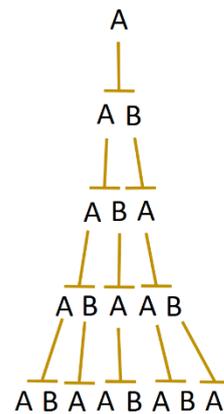
By using a computer it is nowadays easy to generate beautiful 2 or 3D images of plants grown using similar type of L-systems:
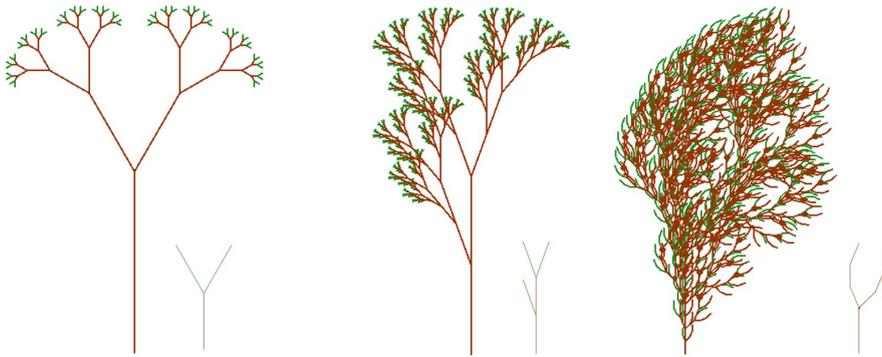
2D plant-like structures

Note from the examples that depending on the "grammar" used, the general symmetry can be broken. This also happens in natural systems.

Fractal dimension

Let us consider a floc consisting of N particles of radius a:
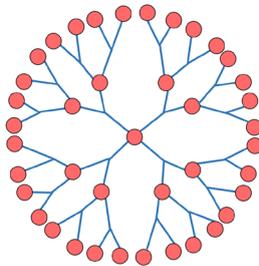
floc of radius R_N consisting of N particles (red balls) of radius a.

The number of particles inside the floc is given by:

$$N = \left(\frac{R_N}{a}\right)^D$$

where D is called the **fractal dimension**. If the floc would be non-fractal (imagine that all the red balls have fused to make a giant red floc), one can estimate that the volume of the giant floc is N times the volume of a red ball, i.e.:

$$\frac{4}{3}\pi R_N{}^3 = N\frac{4}{3}\pi a^3$$

from which follows that

$$N = \left(\frac{R_N}{a}\right)^3$$

implying D = 3. For natural flocs, one finds experimentally that 2 < D < 3, which is equivalent to say that not all space within a floc is occupied by clay particles.

When the fractal dimension of DLCA and RLCA flocs is analysed, it is found that the DLCA flocs have a fractal dimension that is lower than the RLCA flocs, in accordance with the fact that DLCA flocs have a loose structure compared to RLCA flocs. The fractal dimension is also depending on the environment of the flocs, in particular the shear stresses, that affects non only the aggregation but also the re-conformation of the floc:
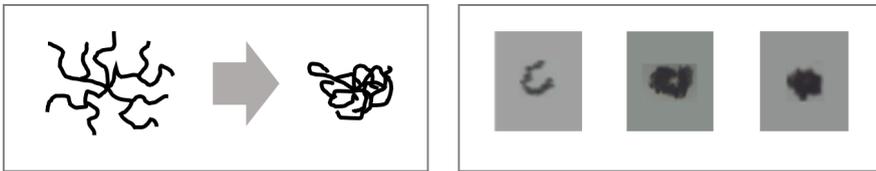
Re-conformation of a floc due to shear: the (pseudo) fractal dimension increases; on the right: flocs observed in the North Sea, close to the Port of Rotterdam, at different stages of coiling. The scale of each floc is approximately 125 μm.

Fractal dimension of flocs

A powerful technique to determine the fractal dimension of flocs is the use of video microscopy[68]. As stated above, flocs are not fractal objects as they are not self-similar but nonetheless a (pseudo) fractal dimension can be derived knowing their size and settling velocity.

The flocs are pipetted into a settling column and their settling is recorded. From the analysis of the video, both particle size R_N and velocity v_N can be determined. Here the settling velocity is plotted as function of the diameter $2R_N$. Each blue cross represents a measurement. The flocs were obtained by mixing river clay and cationic polyelectrolyte.

---

[68] Manning, A.J. and Dyer, K.R. (2002). The use of optics for the in-situ determination of flocculated mud characteristics. J. Optics A: Pure and Applied Optics, Institute of Physics Publishing, 4, S71-S81
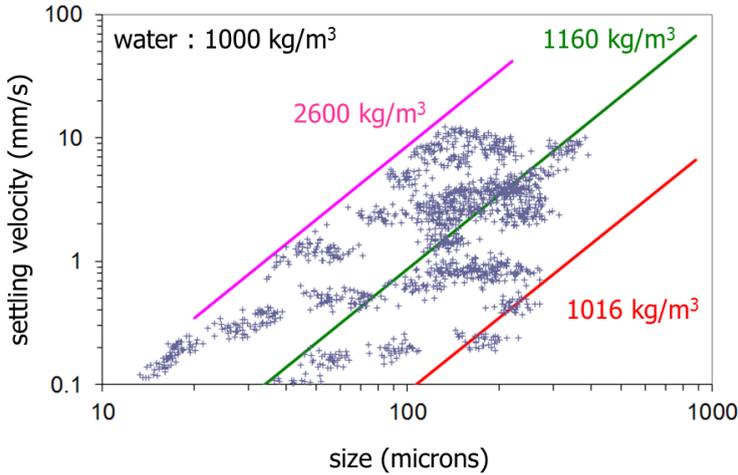
Note that for a given size, for example 100 microns, the velocity v_N can vary by a factor 100, implying that the flocs have very different densities. Using Stokes's law (see **Chapter 2**) the floc density ρ_N can then easily be determined by

$$\rho_N(R_N) = \rho_w + 9\eta v_N/(2R_N{}^2 g)$$

The *mean* density of flocs can be estimated by averaging the density for each size. Below the mean effective density $(\rho_N - \rho_w)$ is plotted as function of the floc diameters in red. The fractal dimension can then be calculated realising that

$$\frac{\rho_N - \rho_w}{\rho_p} = \frac{\phi_s(\rho_p - \rho_w)}{\rho_p}$$

where ϕ_s is the volume fraction of clay inside a floc, $\rho_{p/w}$ the density of the clay/water and that

$$\phi_s = \frac{N \cdot a^3}{R_N{}^3} = \left(\frac{R_N}{a}\right)^{D-3}$$

Combining these equations leads to:

$$\rho_N - \rho_w = (\rho_p - \rho_w)\left(\frac{R_N}{a}\right)^{D-3}$$

and by fitting the data (dashed line) a fractal dimension of D = 2.39 was obtained.
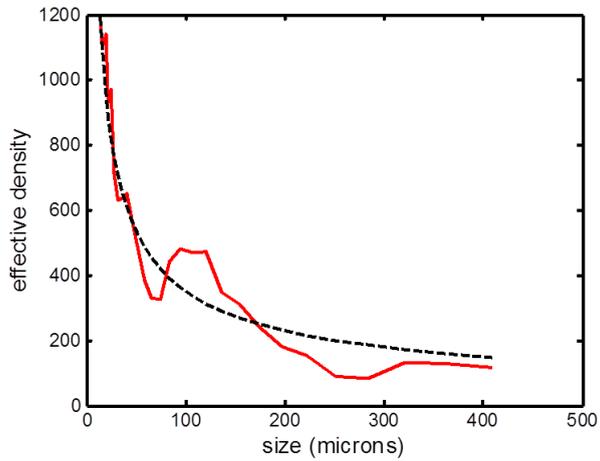
Illustrations

TEM of minerals
Image reproduced from the 'Images of Clay Archive' of the Mineralogical Society of Great Britain & Ireland and The Clay Minerals Society (https://www.minersoc.org/images-of-clay.html)

Fractal (creative commons)
https://en.wikipedia.org/wiki/Koch_snowflake

Fractal tree
Generated using MuPad (Matlab)

Chapter 6

Modelling and measuring the flocculation rate

From the examples given in the previous chapters, it is clear that modelling (and predicting) the floc structure is challenging, owing to the different modes of aggregation and the role of the shape of the particles. It is however easier to say something about the **rate of flocculation** as this depends mainly on the interactions between particles. This is what we are going to show in the present chapter.

The simplest model for predicting the flocculation rate has been developed by the Polish scientist **Marian von Smoluchowski** (1875 - 1917). Smoluchowski moved to Kraków (Poland) in 1913, to take over the chair in Experimental Physics Department, and died there in 1917, a victim of a dysentery epidemic, aged 45. His scientific output is however large: in 1904 he was the first who noted the existence of density fluctuations in the gas phase and in 1908 he became the first physicist to ascribe the phenomenon of critical opalescence to large density fluctuations. His investigations also concerned the blue colour of the sky as a consequence of light dispersion on fluctuations in the atmosphere. In 1906, independently of Albert Einstein, he described Brownian motion. Smoluchowski presented an equation which became an important basis of the theory of stochastic processes. In 1916, he proposed the equation of diffusion in an external potential field. This equation bears his name. In 1916 he also published the article in which he describes the rate of flocculation of colloidal particles[69].

Model of Smoluchowski: aggregation by Brownian motion

We consider the evolution of the number of particles of one **class**, say a class numbered k as function of time. A **class** is the ensemble of all the particles N_k which have as characteristic to be formed of k **primary particles** (particles that cannot break, i.e. the smallest in the system). For instance:

---

[69] Marian Smoluchowski, « Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen », Physik. Zeit., vol. 17, 1916, p. 557–571, 585–599

class 1	class 2	class 3	class 4
$N_1 = 7$	$N_2 = 3$	$N_3 = 1$	$N_4 = 2$
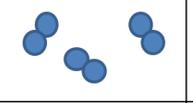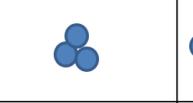
We assume that all primary particles are the same. Particles arrive in class *k* because:

- ➢ they are formed by aggregation of particles of smaller size. We call this gain by aggregation (GA)
- ➢ they are formed by the destruction of larger particles. We call this gain by break-up (GB)

Particles leave class *k* because:

- ➢ they are lost by aggregation of a particle in class *k* with another particle (of any class). We call this loss by aggregation (LA)
- ➢ they are lost by break-up of a particle in class *k*. We call this loss by break-up (LB)

GA	GB	LA	LB

Examples of GA, GB, LA and LB

The general formulation of the rate of change of the number of particle in class *k* is given by:

$$\frac{dN_k}{dt} = GA + GB - LA - LB$$

To simplify the model, we assume that there is no break-up in our system (all particles grow in time). The formulation then reduces to:

$$\frac{dN_k}{dt} = GA - LA$$

where:

$$GA = \frac{1}{2}\sum_{i=1}^{k-1} \alpha_{i,k-i}\, \beta_{i,k-i} N_i N_{k-i}$$

This term represents the rate of formation of a class *k* particle by aggregation of any particles from lower classes (class 1 until class (k-1)), such that $N_i + N_{k-i} \rightleftharpoons N_k$. The ½ accounts for the fact that the aggregates are counted twice: $N_{k-i} + N_i \rightleftharpoons N_k$ is the same as $N_i + N_{k-i} \rightleftharpoons N_k$. the coefficient $\alpha_{i,j}$ is called the collision efficiency (between particle of class i and particle of class j), and $\beta_{i,j}$ is the collision frequency. We also have:

$$LA = N_k \sum_{i=1}^{n-k} \alpha_{k,i}\, \beta_{k,i} N_i$$

This term represents the rate at which a class *k* particle disappears from class *k* by aggregation with any other particle. The summation is limited to $(n-k)$ where *n* is the total number of classes considered (usually taken to be extremely large). This means that we cannot make particles larger that class *n* particles. This is mainly done for numerical procedures reasons. In doing the numerical simulations, one always insures that class *n* is never populated. This is equivalent to say that the summation can be extended to infinity. Usually one defines:

$$K_{k,i} = \alpha_{k,i}\beta_{k,i}$$

In the simplest models, as the one derived by Smoluchowski, on usually assumes that $\alpha_{k,i} = 1$ implying that the particles always stick when they touch. In more realistic models, $\alpha_{k,i}$ is a value between 0 and 1. The collision frequency is linked to the way two particles approach each other. Smoluchowski considered the case where particles were moving because of Brownian motion, which we will discuss now. We have already introduced **Fick's first law** in a previous chapter. From this law, we may evaluate the flux of particles n_k (the number of particles of class k per unit of volume V, for example the volume of the settling column in which the experiment is performed) that enters a sphere of radius *r* centred in a given particle of radius a_i . It is given by:

$$J_{k,i} = 4\pi r^2 D\, \frac{\partial n_k}{\partial r}$$

The units of $J_{k,i}$ are in number of particles (of class) *k* entering the sphere per second. Assuming that $J_{k,i}$ is constant (steady-state condition), it is easy to integrate this equation, and one obtains:

$$n_k = n_{k,\infty} - \frac{J_{k,i}}{4\pi D r}$$

where $n_{k,\infty} = N_k/V$ is the (bulk) particle concentration far from the sphere. We now assume that when a particle *k* reaches $r = a_k + a_i$ it "disappears" from class *k*: it is

glued to the particle i which centre is located at $r = 0$. Thus $n_k(r = a_k + a_i) = 0$ from which follows that:

$$J_{k,i} = 4\pi D(a_k + a_i)n_{k,\infty}$$

This flux is now the rate (number/second) at which particles k meet particle i. The particle i which centre is located at $r = 0$ is not immobile: it also diffuses according to Brownian motion. This implies that the diffusion coefficient D is not the diffusion coefficient of a particle, but should be the **mutual diffusion coefficient** $D_{k,i}$ that accounts for the fact that both particle k and particle i are experiencing Brownian motion:

$$D = D_{k,i} = D_k + D_i = \frac{k_B T}{6\pi\eta} \frac{a_k + a_i}{a_k a_i}$$

The number of particles of class k that are leaving the class due to aggregation is then given by the rate at which particles k meet particle i multiplied by the amount of particles i and summed over all types of particles i :

$$\frac{dN_k}{dt} = \sum_{i=1}^{n-k} N_i J_{k,i} = \frac{N_k}{V} \sum_{i=1}^{n-k} \frac{2k_B T}{3\eta} \frac{(a_k + a_i)^2}{a_k a_i} N_i$$

By dividing this equation on both sides by V one can express it as a relation between concentrations:

$$\frac{dn_k}{dt} = n_k \sum_{i=1}^{n-k} \frac{2k_B T}{3\eta} \frac{(a_k + a_i)^2}{a_k a_i} n_i$$

and define a corresponding rate constant $k_{k,i}$ (m³/s) such that:

$$\frac{K_{k,i}}{V} = k_{k,i} = \frac{2k_B T}{3\eta} \frac{(a_k + a_i)^2}{a_k a_i}$$

We then find that:

$$k_{1,1} = \frac{2k_B T}{3\eta} \frac{(2a_1)^2}{a_1{}^2} = \frac{8k_B T}{3\eta}$$

Note that this rate constant is also equal to the rate $k_{k,k}$ for all colliding particles of equal sizes. (For particles of different sizes, one can easily verify that the collision rate will always be *smaller* than for equal particles).

The rate of change of the number of particle in class *1* at *the early stage of aggregation* is given by:

$$\frac{dn_1}{dt} = -k_{1,1}n_1^2$$

Primary particles cannot be formed, therefore their gain by aggregation (GA) is zero. At the early stage of aggregation, only **doublets** are formed from the aggregation of primary particles, so we do not have to consider the aggregation of class 1 particles with other particles than themselves.

One can solve the equation using the fact that $n_1(t = 0) = n_{1,0}$ and one obtains:

$$n_1(t) = \frac{n_{1,0}}{1 + k_{1,1}n_{1,0}t}$$

If we now consider the aggregation process at any given time, one should find the full equations:

$$\frac{dn_1}{dt} = -k_{1,1}n_1^2 - k_{1,2}n_1n_2 - k_{1,3}n_1n_3 \cdots$$

$$\frac{dn_2}{dt} = \frac{1}{2}k_{1,1}n_1^2 - k_{2,1}n_2n_1 - k_{2,2}n_2n_2 \cdots$$

$$\frac{dn_3}{dt} = \frac{1}{2}k_{1,2}n_1n_2 + \frac{1}{2}k_{2,1}n_2n_1 - k_{3,1}n_3n_1 \cdots$$

By assuming that all the rates $k_{i,j}$ are the same (which is reasonable if the particles are not too different in size) and using $k_a = k_{i,j}$ for all i and j, one obtains:

$$\frac{d(n_1 + n_2 + n_3 \cdots)}{dt} = -\frac{k_a}{2}(n_1 + n_2 + n_3 \cdots)^2$$

the factor ½ can be understood easily when one considers again the early stage of aggregation, when only doublets are formed. Then we have:

$$\frac{d(n_1 + n_2)}{dt} = -\frac{k_a}{2}(n_1 + n_2)^2$$

because:

$$\frac{dn_1}{dt} = -k_{1,1}n_1^2$$

$$\frac{dn_2}{dt} = \frac{1}{2}k_{1,1}n_1^2$$

From this, we realize that the aggregation (loss) of *two* primary particles leads to the creation of *one* doublet.

We define the total amount of particles in the system by:

$$n_T = n_1 + n_2 + n_3 \cdots$$

and the characteristic time:

$$\tau = \frac{2}{k_a n_{1,0}}$$

The solutions to the full equations can be shown to be:

$$n_1(t) = \frac{n_{1,0}}{(1 + t/\tau)^2}$$

$$n_2(t) = \frac{n_{1,0} t/\tau}{(1 + t/\tau)^3}$$

$$n_k(t) = \frac{n_{1,0}(t/\tau)^{k-1}}{(1 + t/\tau)^{k+1}}$$

$$n_T(t) = \frac{n_{1,0}}{1 + t/\tau}$$

Note that for $t \ll \tau$ one finds as expected:

$$n_1(t) = \frac{n_{1,0}}{1 + k_a n_{1,0} t}$$

$$n_{k>1}(t) = 0$$

The solutions are plotted here:
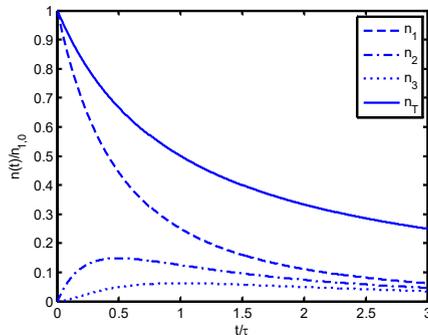
From the figure it can be seen that the concentration of primary particles (n_1) is decreasing as function of time, as these particles are aggregating. The other classes

start from a concentration equal to zero (we have assume that at t = 0 there are only primary particles), then their concentration increases (n_2 is formed by aggregation of n_1), and eventually decrease (n_2 aggregates with n_1 to form n_3 and aggregates with n_2 to form n_4 etc...).

The Brownian motion aggregation we have been reviewing is called **perikinetic aggregation** in textbooks. This type of aggregation does not lead to the rapid formation of very large aggregates, especially in dilute suspension. We can for instance estimate

$$k_a = \frac{8k_B T}{3\eta} \sim 1.23 \times 10^{-17} m^3/s$$

Using an initial concentration of $n_{1,0} = 58 \ g/L = \ 10^{16}$ particles/m^3 (particles of 1 μm)

$$\tau = \frac{2}{k_a n_{1,0}} \sim \frac{2}{1.23 \times 10^{-17} 10^{16}} = 16 \ s$$

Concentrations in natural environments (estuaries, rivers) are more of the order of mg/L leading to $\tau \sim 1.6 \times 10^4 \ s$ (approximatively 5 hours). Aggregation is natural environments, or in flocculation tanks, as the ones used in sanitary engineering are usually triggered by fluid motion. This type of aggregation is called **orthokinetic aggregation**.

Orthokinetic aggregation: aggregation by shear

As for the case of Brownian motion, we evaluate the flux of particles *k* that collide with a particle *i*, but because of the water shear rate $G(s^{-1})$[70], the particle *k* arrives laterally:

---
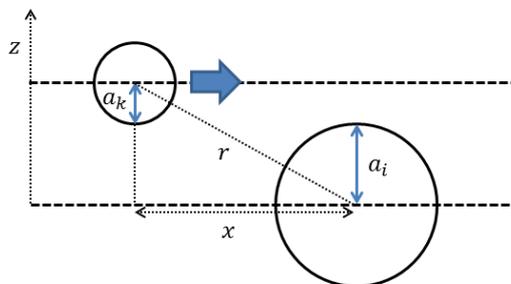
[70] In **Chapter 7** the shear rate will be defined by the symbol $\dot{\gamma}$. Both symbols are used in literature.

142

The flux of particles k arriving on i is given by:

$$J_{k,i} = 4n_k \int_0^{a_k+a_i} (Gz)x dz$$

where Gz is the fluid velocity (m/s) at position z (the velocity is along the x axis, as the blue arrow indicates in the illustration and depends on z only). There is contact between the two spheres when $r = a_k + a_i$ and therefore $x = \sqrt{(a_k + a_i)^2 - z^2}$. The 4 comes from the fact that one wants to know the number of particles whose center pass through the capture cross-section $\pi(a_k + a_i)^2$ accounting for the fact that the particles can come from any direction and thus the surface are of contact is represented by a sphere of area $4\pi(a_k + a_i)^2$. The units of $J_{k,i}$ are in number of particles (of class) k entering the sphere per second. It is easy to find that:

$$J_{k,i} = \frac{4}{3} Gn_k(a_k + a_i)^3$$

Similarly to what we have done before, we can construct:

$$\frac{dn_k}{dt} = \sum_{i=1}^{n-k} n_i J_{k,i}$$

From which we can evaluate the rate constant $k_{k,i}$ (m³/s) :

$$k_{k,i} = \frac{4}{3} G(a_k + a_i)^3$$

Comparing this rate of aggregation with the one obtained from Brownian motion, one observe that the rate we have found now depends significantly on the size of the particles (in fact it scales as the *volume* of the particles): a large particle will aggregate a lot of other particles. In the case of Brownian motion, a large particle would not capture so many other particles, because its (Brownian) motion would be very limited: large particles have a low Brownian diffusion coefficient.

In turbulent conditions, a mean shear stress can be obtained from:

$$G = \sqrt{\frac{\epsilon}{\nu}}$$

where ϵ is the power input per unit of mass of fluid and ν the kinematic viscosity (the ratio between the viscosity and the density of the fluid). The **Kolmogorov microscale**, which separates the inertial range (where the energy is transferred with very little dissipation) from the viscous subrange (where the energy is dissipated as heat) is given by:

$$L = \sqrt{\frac{\nu}{G}}$$

For typical values of shear rates (50-100 s^{-1}) in aqueous dispersions, L is of the order of 100-150 μm, meaning that flocs will be prevented to grow larger than these values, as shown in **Chapter 5**. If large flocs come in a region of high shear, break-up (GB and LB) should not be neglected. The general formulation of the rate of change of the number of particle is then generally calculated numerically for all the classes. These models are found under the name "**Population Balance Equation**" in literature.

Measuring the flocculation rate

The rate at which particles aggregate can be estimated by light scattering measurements (see **Chapter 2**), by evaluating the (mean) particle size as a function of time. In **Chapter 5**, we have introduced the concept of DLCA and RLCA. These can be linked to the mean aggregation diameter, which can be derived from the size of floc in the corresponding class. One finds the following behaviours[71]:
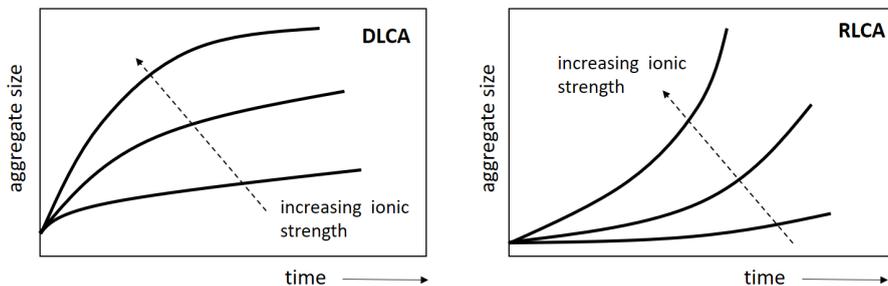
By adapting the Population Balance Equation, so as, in particular, to include the (pseudo) fractal size of the aggregates, it is possible to reasonably well fit measured data. The RLCA mode of aggregation is in particular verified for salt-induced aggregation (see Runkana et al. given in footnote). At high ionic strength however, it would seem that the process goes through a transition between RLCA (short times) and DLCA (longer times). This change in flocculation regime may be due to aggregate restructuring during aggregation.

Stability ratio

Despite the fact that shear-induced aggregation is probably the dominant mechanism for aggregation in engineering, Brownian-induced aggregation is quite

---

[71] Runkana, V., Somasundaran, P., & Kapur, P. C. (2005). Reaction-limited aggregation in presence of short-range structural forces. AIChE journal, 51(4), 1233-1245.

important to determine the **stability** of a suspension. When discussing the DLVO theory in an earlier chapter, we already made the link between zeta potential and stability. We are now going to explain how this link can be understood.

The force on a particle due to the energy barrier is given by $F_{rep} = -d\Phi/dr$ and the particle's velocity can be estimated from the relation derived in Chapter 2:

$$v_k = F \frac{D_k}{k_B T}$$

The flux of particles produced by the force field can be estimated to be:

$$J_{k,i}^{rep} = 4\pi r^2 n_k (2v_k)$$

The 2 comes from the fact that each particle experience the energy barrier (similarly to what we did when we introduced the mutual diffusion coefficient). We have here, for simplicity, assumed that all the particles have the same size. This implies that $D = 2D_k$. The total number of particles that hits the central particle i per second is given by:

$$J_{k,i} = 8\pi r^2 D_k \left[\frac{\partial n_k}{\partial r} + \frac{n_k}{k_B T} \frac{d\Phi}{dr} \right]$$

Assuming a steady-state, one has $J_{k,i}$ constant, and therefore:

$$\frac{J_{k,i}}{8\pi r^2 D_k} = \frac{\partial n_k}{\partial r} + \frac{n_k}{k_B T} \frac{d\Phi}{dr} = constant$$

In our case both n_k and Φ depend on r only (we can therefore replace the ∂ by d in the expression above). Using the relation:

$$\frac{d}{dr} \left[n_k \exp\left(\frac{\Phi}{k_B T} \right) \right] = \frac{dn_k}{dr} \exp\left(\frac{\Phi}{k_B T} \right) + \frac{n_k}{k_B T} \frac{d\Phi}{dr} \exp\left(\frac{\Phi}{k_B T} \right)$$

one gets:

$$\frac{J_{k,i}}{8\pi r^2 D_k} \exp\left(\frac{\Phi}{k_B T} \right) = \frac{d}{dr} \left[n_k \exp\left(\frac{\Phi}{k_B T} \right) \right]$$

which can be integrated from r to ∞ (where $\Phi = 0$ and $n_k = n_{k,\infty} = N_k/V$ is the (bulk) particle concentration far from the sphere)

We thus get:

$$n_{k,\infty} - n_k \exp\left(\frac{\Phi}{k_B T} \right) = \frac{J_{k,i}}{8\pi D_k} \int_r^\infty \frac{1}{r^2} \exp\left(\frac{\Phi}{k_B T} \right) dr$$

which can be rewritten as:

$$n_k = n_{k,\infty}\exp\left(\frac{-\Phi}{k_B T}\right) - \frac{J_{k,i}}{8\pi D_k}\exp\left(\frac{-\Phi}{k_B T}\right)\int_r^\infty \frac{1}{r^2}\exp\left(\frac{\Phi}{k_B T}\right)dr$$

When $r = 2a$ we have aggregation and therefore $n_k = 0$ leading to:

$$J_{k,i} = \frac{8\pi D_k n_{k,\infty}}{\int_{2a}^\infty \frac{1}{r^2}\exp\left(\frac{\Phi}{k_B T}\right)dr}$$

In the limiting case when there is no potential between the particles ($\Phi = 0$), except for an infinitely strong van der Waals attraction when they make contact, we get the so-called **fast aggregation rate**:

$$J_{fast} = J_{k,i}(\Phi = 0) = J_{k,i} = 16\pi a D_k n_{k,\infty}$$

which is indeed the one we found for Brownian motion ($= 4\pi D(a_k + a_i)n_{k,\infty}$).

The **stability ratio** is defined by:

$$W = \frac{J_{fast}}{J_{k,i}} = 2a\int_{2a}^\infty \frac{1}{r^2}\exp\left(\frac{\Phi}{k_B T}\right)dr = 2\int_2^\infty \frac{1}{s^2}\exp\left(\frac{\Phi}{k_B T}\right)ds$$

where $s = r/a$. This integral can be evaluated numerically, and Verwey and Overbeek showed in 1948 that W was determined almost entirely by the value of Φ at its maximum (Φ_{max}). A simple approximation, due to Reerink and Overbeek[72] in 1954, for unequal spheres, is:

$$W \sim \frac{1}{\kappa(a_k + a_i)}\exp\left(\frac{\Phi_{max}}{k_B T}\right)$$

For 1 µm particles, in 100 mM of a 1-1 electrolyte, $\kappa a \sim 500$. For a barrier of $10 k_B T$ one would get $W \sim 45$, whereas for $20 k_B T$ one would get $W \sim 10^6$ i.e. only one in every million collisions would occur between particles having sufficient energy to overcome the barrier.

A small change in electrolyte concentration or in particle's surface potential can have dramatic effects on the aggregation rate. In Chapter 3, when we derived the Schulze-Hardy rule, we have given an different form for Φ_{DLVO}. We will now use the following one:

---

[72] Reerink, H., & Overbeek, J. T. G. (1954). The rate of coagulation as a measure of the stability of silver iodide sols. Discussions of the Faraday Society, 18, 74-84.

$$\Phi_{DLVO} = \frac{-Aa}{12(r-2a)} + 32\pi\varepsilon_0\varepsilon_r\left(\frac{k_BT}{q}\right)^2 a\,\tanh^2\left(\frac{q\psi(a)}{4k_BT}\right)\exp\left(-\kappa(r-2a)\right)$$

which is valid for any $\psi(a)$ and $\kappa a \gg 1$ (as one can see, we have here simplified the expression for the van der Waals force by taking the one derived for plates, i.e. which is valid for spheres as well when $\kappa a \gg 1$). The position of the energy barrier when $\Phi_{max} = 0$ is represented by :

$$\Phi_{DLVO} = \frac{d\Phi_{DLVO}}{dr} = 0$$

From evaluating $\kappa\Phi_{DLVO} + d\Phi_{DLVO}/dr = 0$ one obtains:

$$\frac{-A\kappa a}{12(r-2a)} + \frac{Aa}{12(r-2a)^2} = 0$$

and therefore:

$$\kappa(r-2a) = 1$$

Using this expression in $\Phi_{DLVO} = 0$, and the definition of κ for a z-z electrolyte ($q = ze$):

$$\kappa^2 = \frac{2(ez)^2 N_A}{\varepsilon_0\varepsilon_r k_BT}C(\text{mM})$$

it follows that the critical concentration at which fast aggregation will occur is:

$$C_{\text{fast}}(\text{mM}) = \frac{\varepsilon_0\varepsilon_r k_BT}{2(ez)^2 N_A}\left(\frac{384\pi\varepsilon_0\varepsilon_r}{A}\right)^2\left(\frac{k_BT}{ez}\right)^4\tanh^4\left(\frac{ez\psi(a)}{4k_BT}\right)\exp(-2)$$

It is convenient to introduce here the **Bjerrum length** (after Danish chemist **Niels Bjerrum**) which is the separation at which the electrostatic interaction between two elementary charges is comparable in magnitude to the thermal energy scale, k_BT. The Bjerrum length reads:

$$l_B = \frac{e^2}{4\pi\varepsilon_0\varepsilon_r k_BT}$$

For water at room temperature ($T = 300\,K$), $\varepsilon_r \sim 80$ and $l_B \sim 0.7\,nm$.

We obtain:

$$C_{\text{fast}}(\text{mM}) = \frac{96^2\exp(-2)}{8\pi N_A}\left(\frac{k_BT}{A}\right)^2\frac{1}{l_B{}^3 z^6}\tanh^4\left(\frac{ez\psi(a)}{4k_BT}\right)$$

This expression scales with $1/z^6$ for large surface potentials but when when $ez\psi(a) \ll 4k_BT$ it scales with $1/z^2$ as we have found while deriving the expression for the **Schulze-Hardy rule** in **Chapter 3**.

We can also evaluate Φ_{max} : Φ_{max} occurs when $d\Phi_{DLVO}/dr = 0$ i.e. when

$$\frac{Aa}{12\kappa(r_{max} - 2a)^2} = 32\pi\varepsilon_0\varepsilon_r \left(\frac{k_BT}{q}\right)^2 a \tanh^2\left(\frac{q\psi(a)}{4k_BT}\right)\exp\left(-\kappa(r_{max} - 2a)\right)$$

For which we get:

$$\Phi_{max} = \frac{Aa}{12(r_{max} - 2a)}\left[\frac{1}{\kappa(r_{max} - 2a)} - 1\right]$$

Around the fast aggregation concentration, it is expected that $\kappa(r_{max} - 2a)$ should be close to one (see derivation just above). One may therefore write, using the fact that $\ln(x) \sim 1 + x$ for small x :

$$\Phi_{max} \sim \frac{-Aa}{12(r_{max} - 2a)}\ln[\kappa(r_{max} - 2a)]$$

This implies that:

$$\ln(W) \sim \frac{\Phi_{max}}{k_BT} \sim \frac{-Aa}{12k_BT(r_{max} - 2a)}\Big[\ln(\kappa) + \ln[(r_{max} - 2a)]\Big]$$

In view of the relation $\kappa \sim C^{1/2}$ one then gets:

$$\ln(W) \sim \frac{-Aa}{24k_BT(r_{max} - 2a)}\ln(C) + constant$$

Using (see above on this page):

$$\frac{Aa}{\kappa(r_{max} - 2a)^2} \sim \frac{a}{z^2}\tanh^2\left(\frac{ze\psi(a)}{4k_BT}\right)$$

One also gets:

$$\ln(W) \sim \frac{-a}{z^2}\tanh^2\left(\frac{ze\psi(a)}{4k_BT}\right)\ln(C) + constant$$

From this expression it is clear that when $\psi(a) \sim 0$ (and thus the zeta potential ζ close to zero), one has $\ln(W) \sim 0$ and thus $W \sim 1$ (fast aggregation). This is why, as a rule of thumb, one says that flocculation is predicted to happen at small zeta

potentials: typically fast flocculation occurs when $e\zeta/kT$ is of the order of 1 i.e. at room temperature, when ζ is of the order of 25 mV.

Niels Janniksen Bjerrum (1879 – 1958) had a sister Dr. Kirstine Meyer who was a prominent physicist[73]. His early interest was natural history, but this soon changed to chemistry. He carried out research stimulated by S. M. Jörgensen, one of the pioneers on coordination complexes, obtaining a Master's degree in 1902 and a Doctor's degree in 1908. He visited and worked with several of the great physical chemists of the time: in 1905 Luther in Leipzig, in I907 Werner in Zürich, in 1910 Perrin in Paris and in 1911 Nernst in Berlin. In 1912 he became docent at the University of Copenhagen. In 1914 he succeeded O.T. Christensen as Professor of chemistry at the Royal Agricultural College, Copenhagen where he remained until he retired in 1949. In the period 1906-1908 Bjerrum was determining the structure of various chromium complexes and measuring hydrogen ion concentrations before the expression pH had yet been invented by Sörensen. He also did fundamental work on the factors determining the acidity of soil.
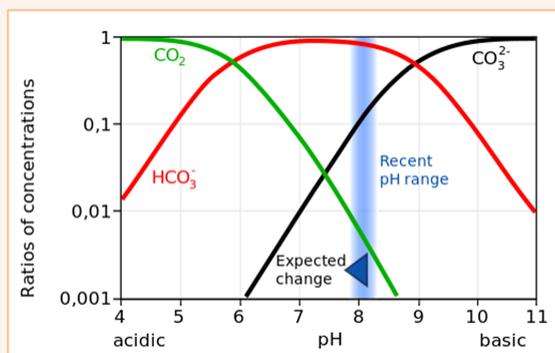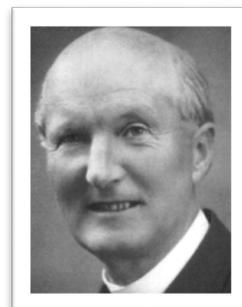
The Bjerrum plot is named after him. A Bjerrum plot is a graph of the concentrations (or ratio of concentrations) of the different species of a **polyprotic** acid (= which are able to donate more than one proton per acid molecule) in a solution, as functions of the solution's pH, when the solution is at equilibrium. Most often, the carbonate system is plotted, where the polyprotic acid is carbonic acid (a diprotic acid), and the different species are carbonic acid, carbon dioxide, bicarbonate, and carbonate. In

---

[73] She received her PhD in physics from the University of Copenhagen in 1909, becoming the first Danish woman to earn a doctorate in natural sciences. Her dissertation, *Temperaturbegrebets Udvikling gennem Tiderne* (The Development of the Temperature Concept through Time), was an in-depth treatment of the history of the concept of temperature. In 1902, Meyer founded Fysisk Tidsskrift, the Danish journal of physics. She was its editor until 1913.

acidic conditions, the dominant form is CO_2; in basic (alkalinic) conditions, the dominant form is CO_3^{2-}; and in between, the dominant form is HCO_3^-. At every pH, the concentration of carbonic acid (H_2CO_3) is assumed to be negligible compared to the concentration of CO_2, and so is often omitted from Bjerrum plots. These plots are typically used in ocean chemistry to track the response of an ocean to changes in both pH and of inputs in carbonate and CO2.

Experimental verification
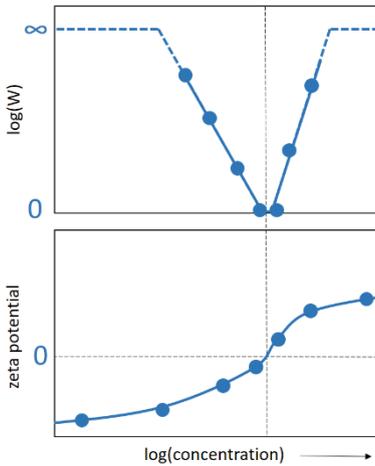
The stability ratio is not only depending on the zeta potential ζ, it is also highly dependent on the ionic strength, as we have seen in Chapter 3 that ζ is highly depending on it. Fast flocculation will occur when ζ is in good approximation zero. This fact is verified experimentally in most cases.

There are however still open questions related to the prediction of the stability ratio. Even though the DLVO theory can correctly predict the fast flocculation regime (based on zeta potential measurements in particular), it fails to describe correctly the slope between $\ln(W)$ and $\ln(C)$: from the DLVO theory, the slope should be much less steep than it is in reality for most systems. Sometimes, also, the fast flocculation regime is not exactly predicted from DLVO:
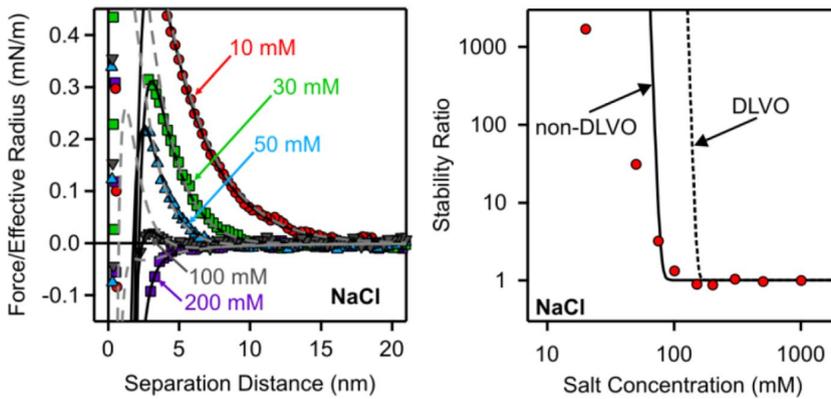
Left: Forces between pairs of latex amidine spheres (size 1μm) versus the surface. Right: measured stability ratios (symbols) with calculated ones based on the force measurements (solid lines). The results for DLVO theory are shown for comparison (dashed line)

From the force/separation graphs[74], the Hamaker constant could be estimated using the data at high salt concentration (where van der Waals force dominates). An additional attractive non-DLVO force could be estimated from the other curves. The result of DLVO theory only is shown for comparison (dashed lines). The non-DLVO behaviour was attributed to surface charge heterogeneities.

Measuring forces between particles

In order to estimate the forces between colloidal particles (as in the previous figure, on the left), one makes use of an **Atomic Force Microscope** (AFM) which is a very-high-resolution type of scanning probe microscopy with a resolution for the distances on the order of fractions of a nanometer (more than 1000 times better than a regular optical microscope) and for the forces of the order of 10^{-8} N. The AFM was invented by IBM Scientists in 1982. The precursor to the AFM, the scanning tunneling microscope (STM), was developed by Gerd Binnig and Heinrich Rohrer in the early 1980s at IBM Research in Zurich, a development that earned them the Nobel Prize for Physics in 1986. The first commercially available atomic-force microscope was introduced in 1989. Another technique to measure interparticle forces is the Surface Force Apparatus (SFA), developed in the early 1970s at Cambridge[75]. The SFA is more ideally suited than AFM to measuring surface-surface interactions, and can measure much longer-range forces more accurately. The SFA technique is however quite demanding and hence, only a handful of labs worldwide have these instruments.
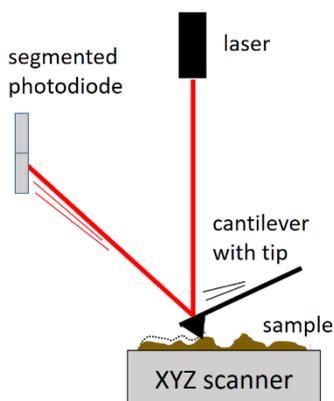
Both AFM and SFA techniques work on the principle of approaching two surfaces and measuring their interaction. Here we present the AFM only. To measure the force between two colloidal particles, a colloidal particle is glued to the tip of the cantilever and brought closer to a colloidal particle that is glued to the XYZ scanner. Depending on the interaction between the particles, the tip is attracted to or repelled from the other particle, and this deviation is measured through the deviation of the laser beam. This deviation can be linked to the force between the particles.

---

[74] Figures taken from Montes Ruiz-Cabello, F. Javier, et al. "Interaction forces and aggregation rates of colloidal latex particles in the presence of monovalent counterions." The Journal of Physical Chemistry B 119.25 (2015): 8184-8193.

[75] See the book by Israelachvili, Jacob N. (1992). *Intermolecular and surface forces*. Boston: Academic Press

Another AFM use is the scanning mode, where the tip is dragged along the surface and the change in elevation is recorded. 3D pictures can then be obtained from the scanned surface[76]:
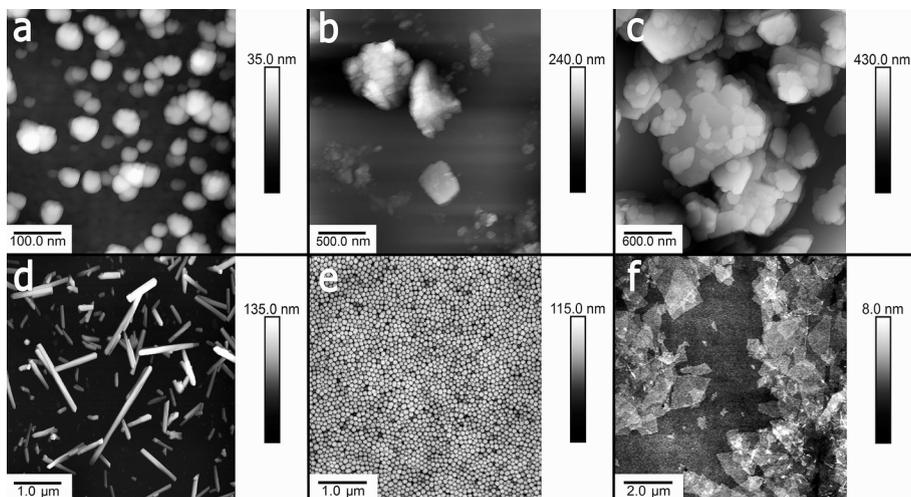
Typical AFM images of a) bentonite, b) montmorillonite, c) kaolin, d) halloysite, e) silica and f) graphene oxide

Settling velocity, optimal flocculation and zeta potential

Instead of measuring particle sizes via light scattering the optimal flocculation rate can also be determined by measuring settling velocities: at optimal settling rate, the zeta potential is minimum[77].

In this last example we will illustrate this fact and we will briefly show the experimental problems to overcome when one wants to study the flocculation of a natural clay with a commercial cationic polyelectrolyte[78]:

---

[76] Kryuchkova, Marina, et al. "Evaluation of toxicity of nanoclays and graphene oxide in vivo: a Paramecium caudatum study." Environmental Science: Nano 3.2 (2016): 442-452.

[77] Yu, X., & Somasundaran, P. (1996). Role of polymer conformation in interparticle-bridging dominated flocculation. Journal of Colloid and Interface Science, 177(2), 283-287.

[78] unpublished results

| 1.1 mg/g | 2.3 mg/g | 3.5 mg/g | 4.5 mg/g | 5.7 mg/g | 6.9 mg/g | 9.1 mg/g |
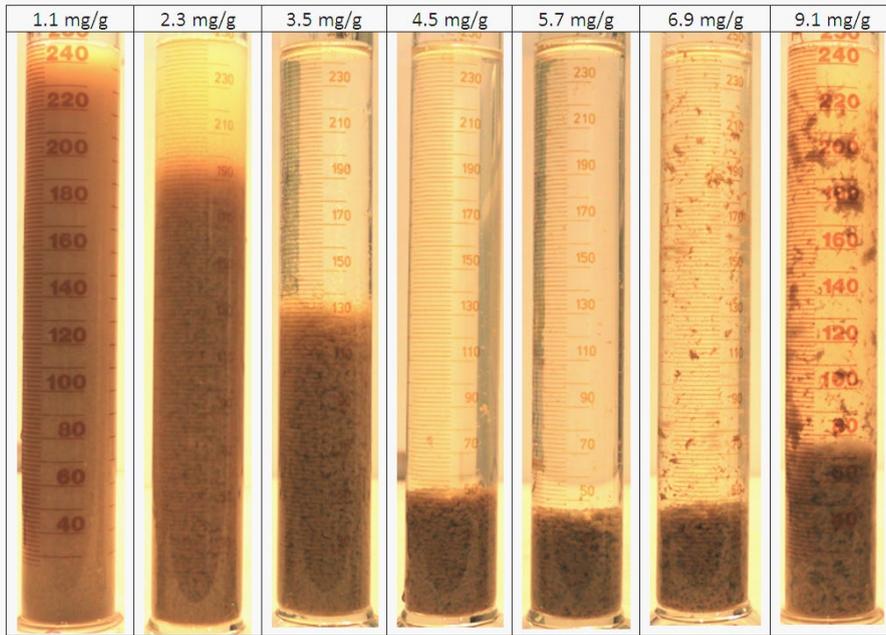
End of settling for a suspension made of 8.7 g/L clay in presence of various amount of cationic flocculant Zetag 7587; flocculant to clay ratios are indicated above each column.

From recording the suspension/water interface over time, the initial settling velocity (estimated from the first 30 s of settling) can be estimated:
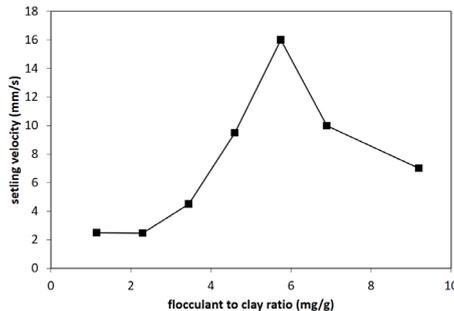
The settling velocities given here were measured in 250 mL columns, and each column was mixed by rotating the column upside down 10 times. Several parameters, like the volume and diameter of the column and the way to mix the suspension is found to influence the settling velocity values. There are several reasons for this: (1) the aggregation and break-up of clay and polyelectrolyte particles is highly dependent on the shear rate and the residence time in the column,

(2) the return flow of liquid in the column is influencing the settling of the flocs and (3) the flocs at higher concentrations of flocculant become positively charged and display an interaction with the wall. The walls of the settling columns are made of glass, which is slightly negatively charged and therefore flocs remain electrostatically attached to them.

From the protocol used to estimate the settling velocities it was found that the fastest settling velocity is obtained for 6 mg/g flocculant to clay ratio. Additional studies on the same system proved that the way of mixing the suspension did not change the ratio for which the fastest settling velocity is observed. The quantitative value of this velocity is however protocol-dependent, as some protocols lead to larger flocs (and larger settling velocities) than others.

The same system was studied by static light scattering (see **Chapters 2 and 5**) and electrophoresis (see **Chapter 3**). For these techniques, lower clay concentrations are required, but it was shown that the results nicely compare to the ones obtained for the settling tests.
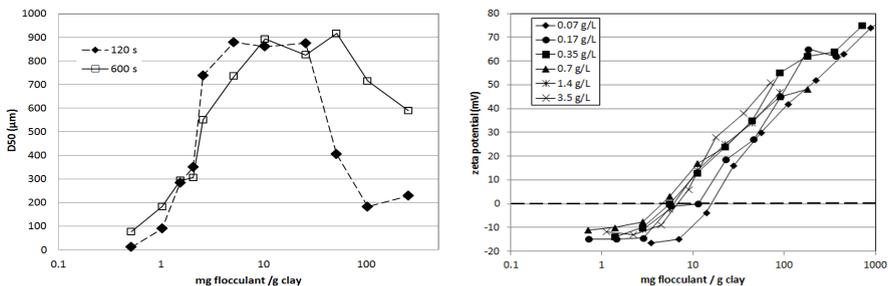
Left figure: mean particle (D50) measured by static light scattering as function of polyelectrolyte to clay ratio, for two times, indicated in the legend. The clay concentration is 0.7 g/L. Right figure: zeta potential estimated from the Smoluchowski formula and derived from electrophoretic mobility results. The zeta potential is given as function of polyelectrolyte to clay ratio for different clay concentrations, indicated in the legend.

As can be seen in the left figure, the largest flocs are obtained for a flocculant to clay ratio of about 5-25 mg/g when the flocs are measured 120 s after mixing. One can see that it corresponds to the ratio for which the zeta potential is zero (right figure). Zeta potential measurements are performed extremely rapidly after mixing, to prevent flocculation in the cell.

After 600 s, the largest sizes are found at ratios about 10-50 mg/g, and for higher ratio's the D50 does not decrease as fast as for flocs obtained at 120 s, as flocs keep growing slowly in time, since they are continuously stirred in the jar. For these high flocculant to clay ratio's the zeta potential of flocs is highly positive, implying the

aggregation mechanism between flocs is due to steric interaction and/or hydrogen bonding, but not to electrostatic attraction.

ZP = - 15 mV	ZP = 0 mV	ZP = + 40 mV
< 5 mg/g	**~ 5mg/g**	**> 5 mg/g**
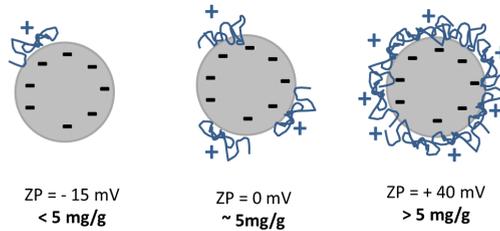
Interpretation of the electrophoretic mobility data: at low flocculant to clay ratio, the clay particle is not fully covered by polyelectrolyte and its charge (and zeta potential) is negative. At a specific flocculant to clay ratio the zeta potential of the system clay + polyelectrolyte is close to zero and rapid flocculation is expected. For higher flocculant to clay ratio, the system becomes positively charged

Illustrations

Smoluchowski (public domain)
https://en.wikipedia.org/wiki/Marian_Smoluchowski

Niels Janniksen Bjerrum (public domain)
https://pubs.rsc.org/en/content/articlelanding/1959/tf/tf959550x001/unauth#!div Abstract

Chapter 7

Rheological behaviour of colloidal suspensions

Rheology is the study of the flow of a liquid in response to an applied force. The word rheology has been introduced in 1928 by **Eugene Bingham**, professor at the university Lehigh in the United States, after a suggestion of a colleague refereeing to a famous quotation of Heraclitus: *Panta rhei* "everything flows" [the word *rhei* means in Greek "to stream"]. In this Chapter we are going to see how the flow of a colloidal suspension is different from that of a simple fluid (water for example).

Viscosity and yield stress

Viscosity

Intuitively, we know what **viscosity** is: it is a measure of "how well" a liquid flows. Honey is a highly viscous liquid, and water is flowing rather well and is therefore a low viscous fluid.
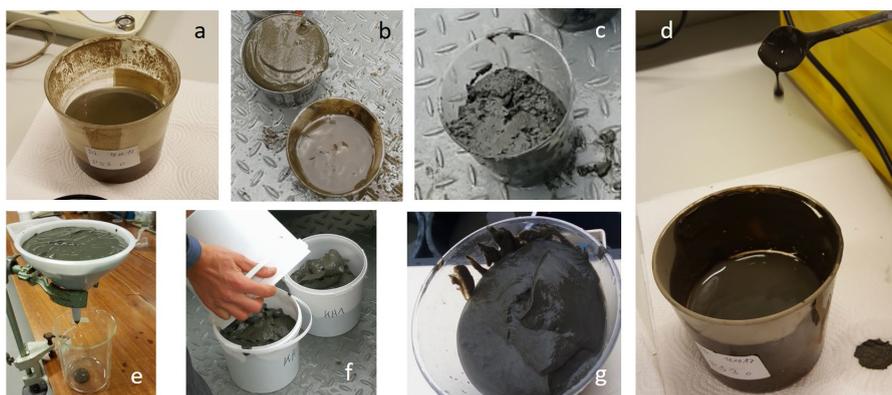
Different clay suspensions exhibiting different flow behaviours; (a) low viscous liquid (clay particles have settled and a clear water layer can be seen); (d) fluid mud; (e) high viscosity mud. The other photographs display mud is various stages of consolidation.

In order to quantify the way liquids flow one has first to define the conditions for the flow: is it a laminar flow or a turbulent flow?

Laminar flow occurs in layers without mixing. Viscosity causes drag between layers as well as with the fixed surface. Laminar flows occur at low water velocities, or low **Reynolds numbers**. An obstruction and high water velocities generate turbulence which mixes the fluid.

In rheological measurements in the laboratory mainly laminar flows are studied, as in laminar flows the shear rate (defined underneath) can properly be quantified. This is more difficult in turbulent flows, even though turbulent flows are the most encountered in the natural environment. The **Kolmogorov microscale** (see **Chapter 6**) becomes an important parameter in turbulent mixing.

laminar regime   turbulent regime
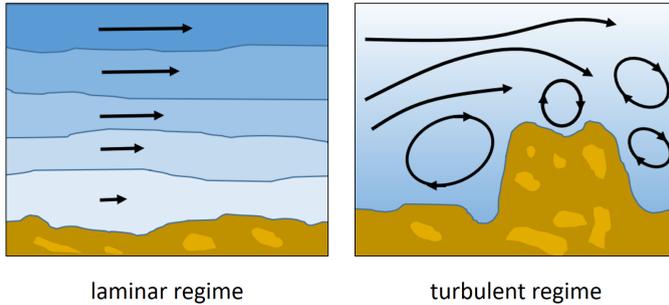
Let us consider a simple flow in the laminar regime, generated under controlled conditions in the lab. An upper plate is moved parallel with a velocity v with respect to another plate which is fixed (its velocity is zero in the frame of the laboratory). The distance between the two plates is L. The force applied to slide the upper plate is *F* and the cross-sectional area of the plate is *A*. The fluid will flow everywhere parallel to the plates, since there is no turbulence and the velocity is assumed to vary linearly across the gap. In most cases, the liquid layers near each plates have the same velocity as that plate ("no-slip" condition).
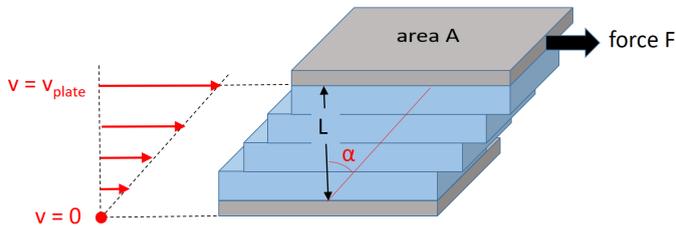
schematic representation of a laminar shear flow between a sliding and a fixed plate separated by a distance L; slices of liquid in blue and variation of the velocity in red

The fact that the fluid velocity varies linearly across the gap will now be discussed. One first estimates the force on one of the water slices. We call (2) the slice of water in contact with the upper plate:
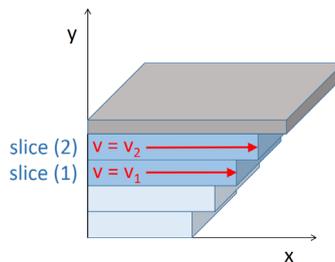
The force on the slice (2) per unit area is given by

$$\sigma = \frac{F}{A}$$

where σ (Pa) is called the **shear stress** acting on the surface (also often denoted τ). Due to stress materials **deform**. The deformation can be measured by the angle α in the top figure. The **shear strain** γ due to this stress is defined as the displacement dx of the top surface with respect to the bottom plate relative to the thickness of layer (2) which is dy :

$$\gamma = \frac{dx}{dy} = \tan(\alpha)$$

In the case we would consider a purely **elastic** material (like a rubber eraser for example) the deformation is reversible, and the material assumes its original shape once the force is removed. As the material we consider is a fluid, we cannot apply a static force, as in the case of a rubber eraser: because of its composition, the rubber will oppose a resistance to the deformation, and a balance of forces will establish between the imposed stress and the rubber resistance to the stress. A pure fluid with no elastic behaviour (like water) has an irreversible deformation, so one has to continuously shear the fluid to be able to measure its response. For fluids, one therefore defines the **shear rate** $\dot{\gamma}$ (s^{-1}) :

$$\dot{\gamma} = \frac{d\tan(\alpha)}{dt} = \frac{1}{dy}\frac{dx}{dt} = \frac{dv_x}{dy}$$

For a constant shear rate, *if the properties of the fluid do not change with height y,* one finds by integration that the velocity indeed varies linearly across the gap as:

$$v_x = \dot{\gamma}y$$

where $y = 0$ is defined at the bottom plate. The velocity $v_x = \dot{\gamma}L$ is the velocity of the upper plate. The relation between the applied shear stress and the induced shear rate (or the induced shear stress and measured shear rate – both are possible with the modern rheometers) is given by:

$$\sigma = \eta\dot{\gamma} = \eta\frac{dv_x}{dy}$$

where η is the **shear (dynamic) viscosity** (Pa · s) of the fluid. One can therefore estimate that the force per unit area exerted by slice (2) on slice (1) and in general slice (k+1) on slice (k) is given by:

$$\sigma_{k+1 \to k} = \eta\frac{dv_x}{dy}$$

This relation is true for so-called *Newtonian fluids*, the simplest fluids, which display a linear relation between shear stress and shear rate. Most concentrated suspensions, however, are non-Newtonian fluids. Here are the most common non-Newtonian behaviours:
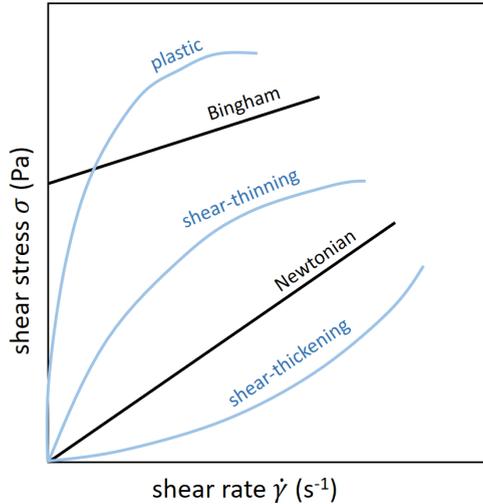
Modelling the shear stress as function of shear rate

For fluids with a zero **yield stress** ($\sigma(\dot{\gamma} = 0)$), which are the **shear-thinning**, **shear-thickening** and **Newtonian** fluids, the relation between shear stress and shear rate can be expressed as:

$$\sigma = k\dot{\gamma}^n$$

where n = 1 for a Newtonian fluid (and then k has the dimension of a viscosity: $k = \eta$), n < 1 for shear-thinning and n > 1 for shear-thickening. Note that the dimension of k depends on n.

For **Bingham** fluids, one uses:

$$\sigma = \sigma_0 + \eta\dot{\gamma}$$

where σ_0 is the **yield stress** (also often noted τ_c and defined underneath) . The non-Newtonian plastic fluids can be modelled with the so-called **Herschel-Bulkley** model:

$$\sigma = \sigma_0 + k\dot{\gamma}^n$$

Many more models exist, as obviously the rheological behaviours of suspensions can be complicated. A model used for the flow of blood is the **Casson** equation, given by:

$$\sigma^n = \sigma_0{}^n + k\dot{\gamma}^n$$

with $n = 0.5$.

Yield stress

The **yield stress** (Pa) can be, in a simplistic way, defined as the stress at which a material begins to flow. In the figure above, it is given by the point $\sigma(\dot{\gamma} = 0)$. For Newtonian fluids, as soon as a shear rate is imposed, the material starts to respond, therefore the yield stress is zero for a Newtonian fluid. Other types of liquids (see ideal Bingham fluid) needs a certain amount of stress in order to start to flow. One can define their yield stress by incrementally increasing the shear stress until their shear rate is changing to a non-zero value. One can also, in the case of a Bingham fluid, extrapolate the $\sigma(\dot{\gamma})$ line to find $\sigma(\dot{\gamma} = 0)$.

A question that arises is how to properly define the yield stress, as obviously it will depend on several factors:

1 – *the shear history*: has the sample been stirred before? One can imagine for example a mud that was undisturbed for a long time. Its yield stress will be higher than a mud that has just been stirred.

2 – *the rate at which the shear stress is applied*: if the shear stress increments are done slowly, the liquid would have time to adapt and its yield stress would be different than when the increments are done fast.

When yield stress experiments are performed, both these points should therefore be addressed. The change in fluid properties, like shear thinning and shear thickening, will be discussed later, as function of the colloidal interactions occurring in the suspensions.

In the case of suspensions which exhibit a yield stress for moderate volume fractions, one can use this yield stress σ_0 to estimate the "amount" of colloidal forces within the suspension:

$$\sigma_0 \sim N \times F_{bond}$$

where N is the number of bonds per unit area between the particles and F_{bond} the mean force of a bond. We make the hypothesis that the network breakage requires the rupture of almost every bond. The amount of hydrodynamic effects can be estimated through the stress of the equivalent suspension of non-interacting particles:

$$\sigma(\phi_S) = \eta(\phi_S)\dot{\gamma}$$

where ϕ_S is the volume fraction of particles, $\eta(\phi_S)$ is the viscosity of a force-free particle suspension and $\dot{\gamma}$ is the shear rate. Examples of $\eta(\phi_S)$ are given in the next sections. The ratio of the hydrodynamic to colloidal stresses is given by:

$$\Gamma = \frac{\eta\dot{\gamma}}{\sigma_0}$$

Formally Γ^{-1} is called a **Bingham number**. In general σ_0 rapidly increases with ϕ_S: for concentrated clay-water mixtures σ_0 is found to follow an exponential law :
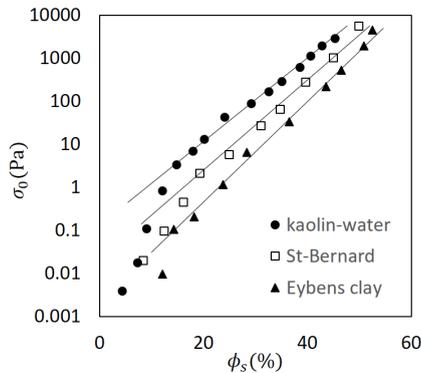
Yield stress of mud suspensions as a function of the volume fraction of particles[79]

For fine mud suspensions the range of Γ for the transition from a colloidal regime to a hydrodynamic one is found to be [0.3 - 50].

We will see in the last section how σ_0 can be linked to DLVO forces, which were introduced in **Chapter 3**.

Viscoelastic fluids

Viscoelastic fluids combine the properties of elastic solids and those of viscous fluids. This is in particular true for densely flocculated suspensions. Ideal plastic material always return to their non-deformed state when the stresses are released. These materials are said to have a perfect "memory" of their non-deformed reference configuration. On the other hand, liquids have no memory at all, and when the stress is released, they remain in their last position. Energetically, the work done in an elastic deformation is stored in the material as potential energy and can be totally recovered when the material returns to its non-deformed state.

---

[79] Coussot, P. (1995). Structural similarity and transition from Newtonian to non-Newtonian behavior for clay-water suspensions. Physical review letters, 74(20), 3971.

In order to study the viscoelastic properties of fluids, oscillatory shear flows are used. If, instead of having a constant speed, the upper plate executes a sinusoidal motion, this will induce a sinusoidal deformation or strain in the sample. Let the oscillatory motion of the upper plate be:

$$x(t) = x_0 \cos(\omega t)$$

where x_0 is the peak displacement and ω the frequency of the oscillation. The deformation (strain) of the sample can then be expressed as:

$$\gamma(t) = \gamma_0 \cos(\omega t)$$

with $\gamma_0 = x_0/L$ where L is the distance between the two plates. In the case of an ideal elastic sample, the shear stress should follow the shear strain deformation in agreement with **Hooke's law**:

$$\sigma(t) = G\gamma(t)$$

where the **shear modulus** G (Pa) is defined as the ratio between shear stress to shear strain. Stress σ and strain γ are then in phase. The stress for a viscous liquid, on the other hand, depends, as we have seen, on the shear rate $\dot{\gamma}$. We then have:

$$\dot{\gamma}(t) = -\omega\gamma_0\sin(\omega t) = \omega\gamma_0\cos(\omega t + \pi/2)$$

If we assume that the liquid is Newtonian we have $\sigma = \eta\dot{\gamma}$ and therefore stress σ and strain γ are $\pi/2$ out of phase. Viscoelastic fluids have a phase shift between stress σ and strain γ that is between 0 and $\pi/2$. A generalized Hooke's law can be defined for these kind of fluids:

$$\sigma(t) = \gamma_0[G'\cos(\omega t) - G''\sin(\omega t)]$$

The in-phase (real part) G' describes the elastic component of the stress. It is called the **storage modulus**. The out-of-phase (imaginary part) G'' represents the viscous part. It is called the **loss modulus**. Clearly, for pure elastic samples, $G'' = 0$ and $G' = G$ and the stress σ is given by $\sigma = G\gamma$. For pure Newtonian viscous samples $G' = 0$ and $\sigma = -\gamma_0 G''\sin(\omega t) = -\eta\omega\gamma_0\sin(\omega t)$, from which we deduce that $G'' = \eta\omega$. The phase angle between stress and strain determines how much mechanical energy will be dissipated in heat and is referred to as **loss angle** δ:

$$\tan\delta = G''/G'$$

Maxwell model

A viscoelastic fluid having a Newtonian loss modulus and a Hooke storage modulus can be represented by a pure viscous damper (symbol η) and a purely elastic spring (symbol G) connected in series[80]. This model is called after Maxwell, who proposed it in 1867:
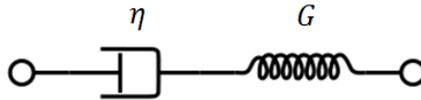
The total stress and total strain can be defined as:

$$\sigma = \sigma_\eta = \sigma_G$$

$$\gamma = \gamma_\eta + \gamma_G$$

[note the analogy with electronic (equivalent) circuits where σ would represent a current and γ a voltage]. One obtains:

$$\dot{\gamma} = \frac{\sigma}{\eta} + \frac{1}{G}\frac{d\sigma}{dt} = \frac{\sigma}{\eta} + \frac{1}{G}\dot{\sigma}$$

where we used $\sigma_\eta = \eta\dot{\gamma}_\eta$ as for Newton fluids for the pure viscous damper, and $\sigma_G = G\gamma_G$ for a purely elastic spring. From that equation a characteristic relaxation time can be estimated: $\tau = \eta/G$. If we adopt the complex representation such that $\gamma^* = \gamma + i\,\mathrm{Im}(\gamma^*) = \gamma_0\exp(-i\omega t)$ and $\sigma^* = \sigma + i\,\mathrm{Im}(\sigma^*) = \sigma_0\exp(-i\omega t)$, one can rewrite the previous equation as:

$$\dot{\gamma}^* = \frac{\sigma^*}{\eta} + \frac{1}{G}\dot{\sigma}^*$$

Taking the real part of both sides of this equation will give back the original equation. By substitution we find:

$$\dot{\gamma}^* = -i\omega\gamma^* = \left(\frac{1}{\eta} - \frac{i\omega}{G}\right)\sigma^* = \frac{1 - i\omega\tau}{\eta}\sigma^*$$

From which we deduce that:

---

[80] There exist also a model where the damper and the spring are in parallel. That model is called the **Kelvin–Voigt model.** This model is used to describe a simple viscoelastic solid.

$$\sigma^* = \frac{-i\omega\eta}{1 - i\omega\tau}\gamma^*$$

A complex **shear modulus** G^* can be defined such that:

$$\sigma^* = G^*\gamma^*$$

with

$$G^* = G' - iG''$$

We find from the previous relations:

$$G' - iG'' = \frac{\omega^2\eta\tau}{1 + \omega^2\tau^2} - i\frac{\omega\eta}{1 + \omega^2\tau^2}$$

Often, a **complex viscosity** η^* is also defined such that:

$$\eta^* = \eta' + i\eta'' = \frac{G''}{\omega} + i\frac{G'}{\omega}$$

which leads to:

$$\eta' = \frac{\eta}{1 + \omega^2\tau^2}$$

$$\eta'' = \frac{\eta\omega\tau}{1 + \omega^2\tau^2}$$

From which we can deduce that

$$G'' = \eta'\omega$$

For a pure viscous liquid $\tau = \eta/G \to 0$ and $G'' = \eta\omega$ as discussed above. We have now found for the complex stress/strain relation that

$$\sigma^* = \left(\frac{\omega^2\eta\tau}{1 + \omega^2\tau^2} - i\frac{\omega\eta}{1 + \omega^2\tau^2}\right)\gamma^*$$

Taking the real part gives:

$$\sigma(t) = \frac{\omega^2\eta\tau}{1 + \omega^2\tau^2}\gamma_0\cos(\omega t) - \frac{\omega\eta}{1 + \omega^2\tau^2}\gamma_0\sin(\omega t)$$

From $\sigma(t) = \gamma_0[G'\cos(\omega t) - G''\sin(\omega t)]$ we deduce that:

$$G' = G\frac{\omega^2\tau^2}{1 + \omega^2\tau^2}$$

$$G'' = G \frac{\omega\tau}{1 + \omega^2\tau^2}$$

These functions are plotted here:
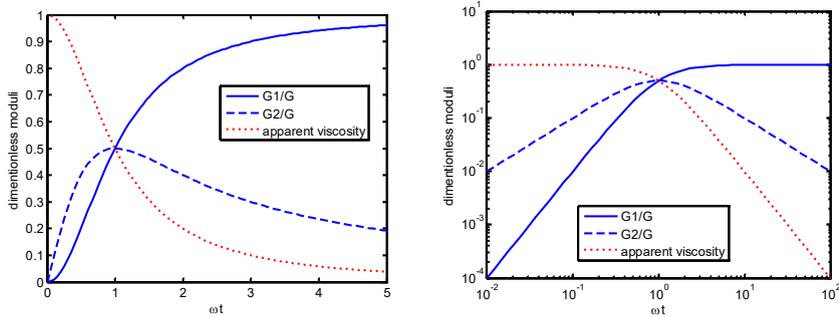
Left: Linear plot of the dimensionless storage modulus G'/G (blue line), loss modulus G''/G (blue dashed line) and apparent viscosity η'/η. Right: the same, but in log-log scale.

From the graphs it is clear that $\tau = \eta/G$ represents the time where one goes from a viscous-dominated fluid ($\eta' \sim \eta$) to an elastic-dominated fluid ($G' \sim G$). The loss modulus is highest at that transition.

Couette rheometer

A rheometer is a laboratory device used to measure how a liquid, suspension or slurry flows in response to applied forces. One of the most used rheometer is the concentric cylinder rheometer (also called Couette rheometer, after its inventor, the physicist Maurice Couette (1858-1943)). Other types of geometry (plate-plate, which was discussed at the beginning of the chapter or cone-plate,…) do also exist. In a Couette rheometer a torque M (thus the shear stress) is imposed and the rotation speed (thus the shear rate) is recorded. This type of rheometer is said to be stress-controlled. Strain-controlled rheometers also exist.

We assume that the outer cylinder remains fixed and that the inner cylinder rotates with an angular velocity $\omega = d\theta/dt$. The shear stress $\sigma(r)$ represents the friction between two layers located on either side of the distance r. The torque M on the inner cylinder is equal to the force due to friction at the distance r multiplied by r as:

$$\boldsymbol{M} = r\boldsymbol{e_r} \times F\boldsymbol{e_\theta}$$

$$M = rF \sin(\boldsymbol{e_r}, \boldsymbol{e_\theta})$$

$$M = rF(r)$$

The stress at the distance r is therefore linked to the torque by:

$$\sigma(r) = \frac{F(r)}{2\pi r h} = \frac{M}{2\pi r^2 h}$$

If we assume that R_{out}/R_{in} is close to one (which is the case in standard rheometers), then both σ and $\dot{\gamma}$ are independent of the position in the gap between the two cylinders. It is then easy to show that the shear rate is proportional to the angular velocity.

$$\dot{\gamma} = \frac{dv_\theta}{dr} = \frac{v_\theta}{(R_{out} - R_{in})} = \frac{R_{in}\omega}{(R_{out} - R_{in})}$$
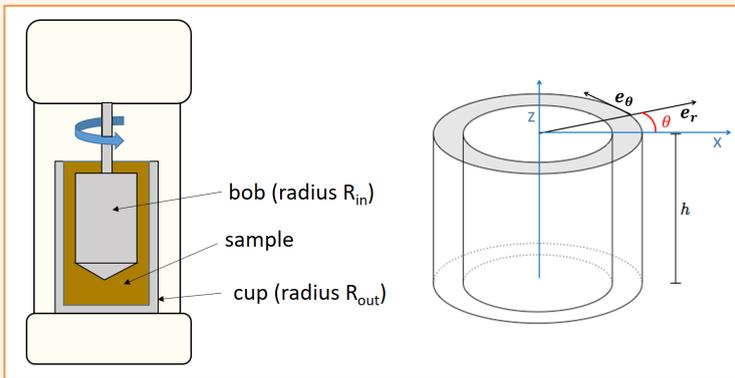
Left: illustration of a rheometer; usually the ratio R_{out}/R_{in} is very close to 1; Right: concentric inner and outer cylinders with cylindrical coordinates

Rheology of suspensions

Until now, we have reviewed general principles about rheology. We did not yet discuss the microscopic structure of these fluids, and considered the fluids as a continuum, having bulk properties. Fluids displaying "real" continuum properties are for instance water or oil, i.e. fluids composed of the same type of molecules. In the remainder of the chapter we will discuss the rheological properties of suspensions as function of the properties of the solvent and colloidal particles that form these fluids. To start, we give an example[81] of the flow behaviour of a diluted clay suspension in a 1-1 electrolyte:

---

[81] See Handbook of Clay Science Edited by F. Bergaya, B.K.G. Theng and G. Lagaly Developments in Clay Science, Vol. 1, Chapter 5
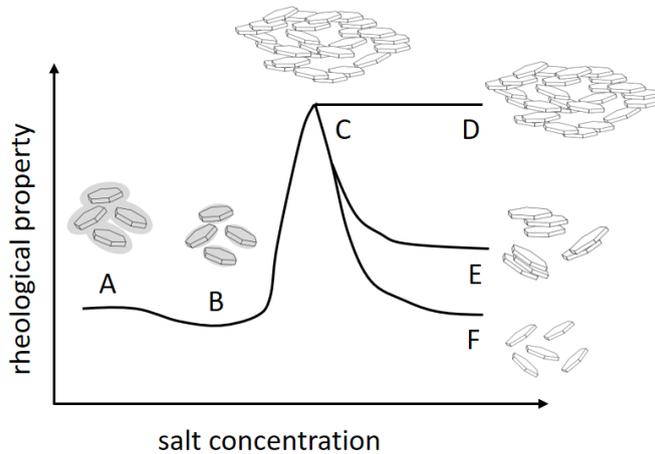
A: isolated particles; B: minimum of rheological properties (viscosity, yield stress) due to electroviscous effect; C,D: aggregation in the form of networks; E,F: fragmentation of the networks at high salt concentrations.

The effect of salt concentration can be understood as follows: both yield stress and viscosity are low at no added salt, as the clay platelets are just suspended in water, so the rheological properties measured should be close to (a bit larger than) the one of water (as we consider a dilute suspension). A minimum is observed at low salt concentration (in B) as a consequence of the **secondary electroviscous** effect: the double layers around the particles are sufficiently compressed so as to ensure that the particles do not "feel" each other over significant distances, so they can flow better. When the amount of added salt becomes significant, aggregation occurs and large aggregates are formed which hinder the flow, leading to an increase in viscosity and yield stress. At even higher salt concentration, several scenarios are possible: (D): the aggregates are very strong, and remain as they are, and the rheological properties do not change upon addition of salt, (E,F): some delamination occurs, and the size of the aggregates is reduced, leading to a reduction in rheological properties.

Dilute and semi-dilute suspensions

In his doctoral thesis Albert Einstein derived a relation between the viscosity of the suspension and the volume fraction of the *non-interacting* spherical colloidal particles inside the suspension. He found (after a colleague, Hopf, corrected a mistake[82]), that:

---

[82] Einstein had found $\eta = \eta_0(1 + \phi)$ which would be the correct expression if the suspension would be made of droplets or gas bubbles for which there is no friction at the particle's surface

$$\eta = \eta_0(1 + 2.5\phi_s)$$

where η_0 is the viscosity of the fluid when there is no particles and ϕ_s the volume fraction of particles. Even though the relation looks simple, it is quite difficult to derive. The result is obtained by calculating the energy dissipation in a sphere of radius R around the reference sphere and let R go to infinity. A rigorous mathematical derivation has been provided by **Batchelor** and Green, who even extended Einstein's formula for *interacting* particles. An alternative (simpler) derivation is outlined in the book of Theo van de Ven (*Colloidal Hydrodynamics*, cited in the reference list), based on the work of Happel and Brenner. The extended Einstein formula can be written:

$$\eta = \eta_0\left(1 + 2.5\phi_s + c_2\phi_s{}^2\right)$$

where c_2 is a numerical value. The term $\phi_s{}^2$ behind reflects the fact that c_2 should account for particle-particle interactions. Depending on the assumptions made, c_2 can therefore take different values.

Batchelor and the development of colloidal hydrodynamics

Batchelor and Green in an important article[83] have found the second order correction to Einstein's relation:

$$\eta = \eta_0\left(1 + 2.5\phi_s + 7.6\phi_s{}^2\right)$$

implying that $c_2 = 7.6$. This was in the case where they considered an **extensional flow**. For a **shear flow** of a semi-dilute suspensions of hard spheres with no Brownian motion (i.e. large particles), assuming a random particle distribution they found $c_2 = 5.2$.

shear flow :
adjacent particles moving over or past each other

extensional flow :
adjacent particles moving away from or towards each other
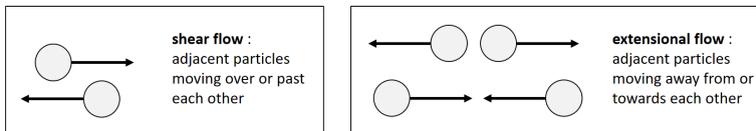
The calculation of the $\phi_s{}^2$ term requires evaluating diverging integrals which can be evaluated by using special mathematical techniques. Batchelor extended Einstein's original argument to show that Brownian motion in a suspension with two or more particles can be represented as a statistical thermodynamic force.

(free slip). The other contribution to the viscosity comes from the distortion of the flow lines, which exist in any case, due to the presence of the particles.

[83] Batchelor, G. K., & Green, J. T. (1972). The determination of the bulk stress in a suspension of spherical particles to order c^2. Journal of Fluid Mechanics, 56(03), 401-427.

George Keith Batchelor FRS (1920 – 2000) was an Australian applied mathematician and fluid dynamicist. He was for many years the Professor of Applied Mathematics in the University of Cambridge, and was founding head of the Department of Applied Mathematics and Theoretical Physics (DAMTP). In 1956 he founded the influential Journal of Fluid Mechanics which he edited for some forty years. Prior to Cambridge he studied in Melbourne High School.
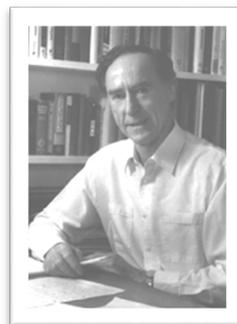
As an applied mathematician (and for some years at Cambridge a co-worker of Sir Geoffrey **Taylor** in the field of turbulent flow), he was a keen advocate of the need for physical understanding and sound experimental basis. His *An Introduction to Fluid Dynamics* (CUP, 1967) is still considered a classic of the subject.

Batchelor also contributed greatly to make Kolmogorov 's theory of turbulence understandable, by publishing 3 articles where he explains this theory in detail. In the 1950/60s even students in the Soviet Union used Batchelor's articles as an introduction to the subject. He worked in Cambridge with **G.I. Taylor**, a British physicist and mathematician and major figure in fluid dynamics and wave theory, whose interest moved from turbulence to other fields, so the work of Batchelor in turbulence was to a large extend independent. Realizing that he needed experiments to improve his theoretical work, Batchelor wrote to an Australian colleague and close friend A.A. Towsend " You will come to Cambridge, study turbulence and work with G.I. Taylor". The answer came immediately: " I agree, but I have two questions: what is turbulence and who is G.I. Taylor?". Townsend came and soon revealed himself as one of the most remarkable experimentalists working in turbulence.

Among the students of Batchelor, some of them made substantial contributions to colloidal science, Edward **Hinch** and Richard Wyngham **O'Brien** in particular for their work on electro-rheology. O' Brien was at the origin of the numerical code that enables to calculate the frequency-dependent electrophoretic mobility and electrokinetic response of charged colloids (for all κa and ζ potentials). He also developed the theory of electroacoustics for colloids, and is at the origin of the **Electrokinetic Sonic Amplitude** (**ESA**) measurement device, which enables to probe the response of concentrated slurries.

When Brownian motion is taken into account[84], then an extra term accounting for the stresses generated in the dispersion by the random movements of the particles

---

[84] Mendoza, C. I., & Santamaria-Holek, I. (2009). The rheology of hard sphere suspensions at arbitrary volume fractions: An improved differential viscosity model. The Journal of chemical physics, 130(4), 044904.

should be added. Numerical calculations have then showed that $c_2 = 6.17$. Depending on the assumptions made about the hydrodynamic flow and the particle-particle interactions, one finds that in general that $5.00 \leq c_2 \leq 6.17$.

The extended Einstein formula is valid primarily at low volume fractions and low shear frequencies (low ω). As many industrial applications concern non-diluted suspensions at any shear rate, many semi-empirical formulae have been proposed to fit experiments in the largest possible range of volume fractions.

Krieger – Dougherty model (for a large range of volume fraction)

This model was developed as an extension of Einstein's model for higher volume fractions. It is derived as follows. The addition of particles to a Newtonian medium will increases the viscosity:

$$\eta(\phi_s + d\phi_s) = \eta(\phi_s) + d\eta$$

By analogy with Einstein's formula, one would expect to have:

$$d\eta = 2.5\eta(\phi_s)d\phi_s$$

However, by adding particles, the space available is not the entire volume, but the volume reduced by a factor proportional to the current volume fraction. We define ϕ_c as the filling fraction at maximum packing, and therefore obtain that:

$$d\eta = \eta(\phi_s)\frac{2.5d\phi_s}{1 - \phi_s/\phi_c}$$

One can verify that for low volume fraction one has indeed $d\eta \sim 2.5\eta(\phi_s)d\phi_s$ whereas for $\phi_s \to \phi_c$ one gets $d\eta \to \infty$ which is expected, as at maximum packing the fluid cannot flow anymore.

We now have obtained:

$$\frac{\eta(\phi_s + d\phi_s) - \eta(\phi_s)}{\eta(\phi_s)} = \frac{d\eta}{\eta} = d\big(\ln(\eta)\big) = \frac{2.5d\phi_s}{1 - \phi_s/\phi_c} = -2.5\phi_c d\big(\ln(1 - \phi_s/\phi_c)\big)$$

which can be integrated:

$$\int_{\ln(\eta_0)}^{\ln(\eta(\phi_s))} d\big(\ln(\eta)\big) = -2.5\phi_c \int_{\ln(1)=0}^{\ln(1-\phi_s/\phi_c)} d\big(\ln(1 - \phi_s/\phi_c)\big)$$

This gives:

$$\ln\left(\frac{\eta(\phi_s)}{\eta_0}\right) = -2.5\phi_c\ln\left(\frac{1 - \phi_s/\phi_c}{1}\right)$$

from which we obtain the **Krieger-Dougherty relation**:

$$\eta(\phi_s) = \eta_0 \left(1 - \frac{\phi_s}{\phi_c}\right)^{-2.5\phi_c}$$

This relation agrees reasonably well with the experimental data. Moreover, it reduces to the Einstein's equation for low ϕ_s as one can verify easily using the relation $(1 + x)^n = 1 + nx$ which holds for small x.

Other models for a large range of volume fraction

Several authors have extended the Einstein relation by using an effective filling fraction:

$$\phi_{\text{eff}} = \frac{\phi_s}{1 - c\phi_s}$$

where the constant c takes into account the fact that the complete free volume cannot be filled by the spheres. They obtain:

$$\eta = \eta_0(1 + 2.5\phi_{\text{eff}})$$

using the same arguments as the ones presented in the derivation of the Krieger-Dougherty relation, it is possible to find:

$$\eta(\phi_s) = \eta_0(1 - \phi_{\text{eff}})^{-2.5}$$

This formula reduces to Einstein formula for low ϕ_s where $\phi_{\text{eff}} \sim \phi_s$ and for the critical packing ϕ_c one has $\phi_{\text{eff}} = 1$. This last equation enables to find a relation for the constant c:

$$c = \frac{1 - \phi_c}{\phi_c}$$

Here we give a plot of the theoretical behaviour of the functions we have described:
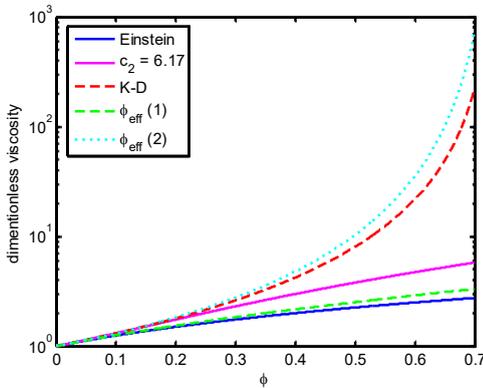
Dimensionless viscosity η/η_0 as function of volume fraction following the models:

$\eta = \eta_0\left(1 + 2.5\phi_s + c_2\phi_s^{\,2}\right)$ *where* $c_2 = 0$ *(Einstein),* $c_2 = 6.17$*, Krieger-Dougherty relation (K-D),* $\eta = \eta_0(1 + 2.5\phi_{\text{eff}})$ *(called* ϕ_{eff} *(1))* *and* $\eta(\phi) = \eta_0(1 - \phi_{\text{eff}})^{-2.5}$*(called* ϕ_{eff} *(2)). We have taken* $\phi_c = 0.74$*.*

In practice, it appears that relations as K-D and ϕ_{eff} (2) are close to the experimental data. Here we give an example with measurements on silica colloidal particles of about 75 nm in radius dispersed in cyclohexane (an oil that enables to nearly suppress all electrostatic and van der Waals interactions and ensures that the suspension remains stable). The data is taken from an old article[85] and was plotted with a linear-linear scale, which made it difficult to accurately estimate the points at low volume fraction. This is the reason of the scatter of the data points at low volume fraction.
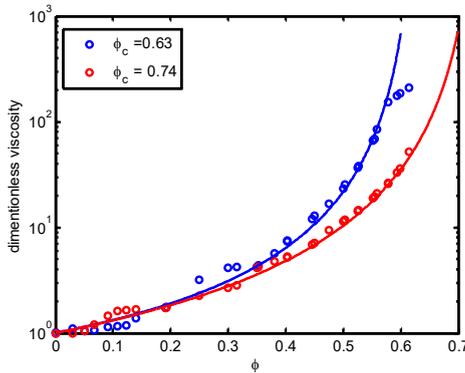
Dimensionless viscosity η/η_0 as function of volume fraction; symbols indicate measurements and lines are plotted according to the model: $\eta(\phi_s) = \eta_0(1 - \phi_{\text{eff}})^{-2.5}$.

An important factor to account for, which makes a universal theoretical (non-empirical) model for the whole volume fraction range quite impossible to realize is

[85] de Kruif, C. D., Van Iersel, E. M. F., Vrij, A., & Russel, W. B. (1985). Hard sphere colloidal dispersions: Viscosity as a function of shear rate and volume fraction. The Journal of chemical physics, 83(9), 4717-4725.

that phase transitions can occur in the suspension (without any aggregation). This has already been briefly discussed in Chapter 2, in the context of particle settling. In **Chapter 2** we have given the example of anisotropic particles, but phase transitions also occur for spherical (isotropic) particles. This is discussed next.

Phase transition

The phase transition for (spherical) colloidal particles can be schematized as follows:
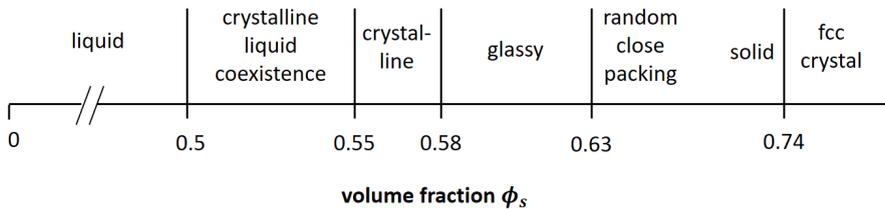
Hard-sphere phase diagram constructed from light diffraction measurements
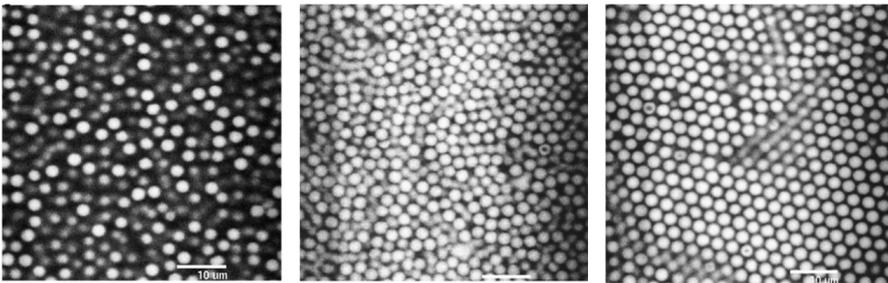
Confocal micrographs of different phases in a colloidal suspension with 5% polydispersity. Scale bar is 10 μm. From left to right: liquid, glass and crystal phase.

At low volume fraction the particles can diffuse freely and there is no long-range ordering in particle position implying that the dispersion is in a fluid state. Increasing concentration above $\phi_s = 0.5$ liquid and crystalline phases coexist in equilibrium and the fraction of crystalline phase increases until the sample is fully crystalline at $\phi_s = 0.55$. With increasing volume fraction, the particle mobility is dramatically reduced until the glassy state is reached. In a glassy state, the dispersion does not flow anymore. One can make the analogy with a (liquid) glass window that solidifies when the temperature is decreased[86], except that decreasing temperature is now

---

[86] The fact that glass panels do flow after hundreds of years (see the bottom of stained glass panels in old churches) cannot be explained from the simple statement that this is due to gravity. If only gravity would play a role, glass panels would have significantly changed their

equivalent with increasing volume fraction. Note that the existence of the glassy state requires some polydispersity (at least some %) otherwise the system will directly crystallise[87].
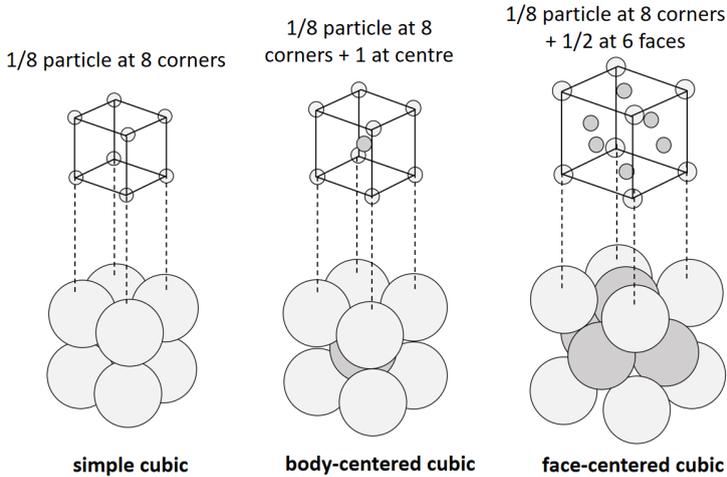
Some close packings happening in crystals and colloidal crystals

The upper limit for close packing (**hexagonal close packing** or **face-centered cubic, i.e. fcc**) is given by $\phi_c = \pi/(3\sqrt{2}) = 0.74$. Other close packings are the **body-centered cubic** packing where $\phi_c = \pi\sqrt{3}/8 = 0.68$, **simple cubic** packing where $\phi_c = 0.52$.
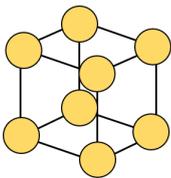
We here briefly recall how these values are found with taking the simple cubic as an example. In a simple cubic the particles are arranged as in the figure. For one element of fluid (the black cube – that we assume to have a length R), one can see that there is $1/8^{th}$ of a colloidal particle (of radius a) in each corner that is inside the cube. In total there is therefore $8 \times (1/8)$ = one particle's volume inside the cube. In close packing, the particles are touching, implying that $R = 2a$. The volume fraction can therefore be calculated by:

$$\phi_c = \frac{4\pi a^3/3}{R^3} = \frac{4\pi a^3/3}{(2a)^3} = \frac{\pi}{6} = 0.52$$

---

bottom thickness only after a period larger than the age of the universe. (http://engineering.mit.edu/ask/how-does-glass-change-over-time)

[87] Hunter, G. L., & Weeks, E. R. (2012). The physics of the colloidal glass transition. Reports on progress in physics, 75(6), 066501 and https://arxiv.org/pdf/1106.3581.pdf

The other close packings are found in the same way. Close packings are usually used to describe the geometry of crystals, made of atoms. As said in Chapter 2, this is why colloidal suspensions are often used as model for crystals and are termed **colloidal crystals**.

The transition from one packing order to another one can be observed experimentally. The measurements presented here (right figure) are done by torsional resonance spectroscopy, a method that enables to get the shear modulus at very low frequencies using acoustic waves[88].
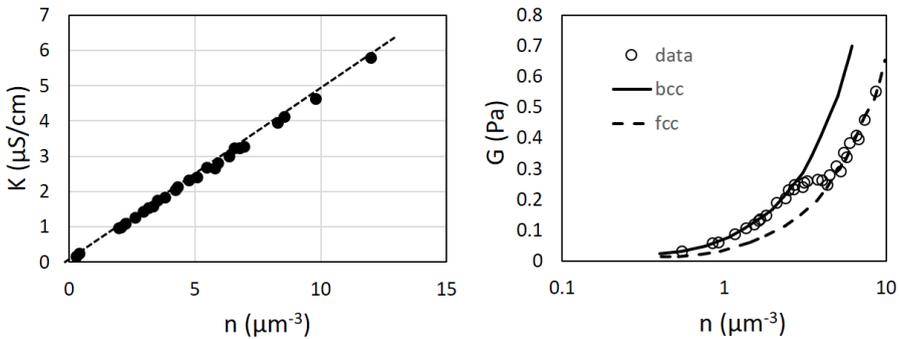
Conductivity (left) and shear modulus (right) as function of volume fraction, for spheres of a = 60 nm radius that are highly charged and dispersed in de-ionized (i.e. very pure) water. The x-axis represents the number of spheres per μm³

Given the size of the particles the number of particles can easily be converted in volume fraction. To give an idea, 1 particle/μm³ corresponds to $\phi_s = 10^{-3}$. The fact that these particles crystallise at such low volume fractions is linked to their charge.

Note that the conductivity measurements do not enable to give information about the packing transition. However, the slope of the line can be used to determine the charge of the spheres, according to the relation:

$$K = n(Ze)^2 \left(\mu_p + \mu_{H^+} \right) + K_B$$

where e is the elementary charge, μ_p is the independently measured particle mobility (from **electrophoresis** for example), μ_{H^+} the electrophoretic mobility of H⁺ that can be found in handbooks from the **limiting conductivity** (see Chapter 3) and Z is the valence of the particle. K_B represents the background conductivity from the self-dissociation of water and residual impurities. As the spheres are dispersed in

---

[88] Wette, P., Schöpe, H. J., & Palberg, T. (2003). Experimental determination of effective charges in aqueous suspensions of colloidal spheres. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 222(1), 311-321.

ultra-pure water, effectively only their counterions (H^+) are present in the water, and the total amount of counterion charges should be equal to the total amount of charges on the colloidal particles, because of electroneutrality. In the present case, one finds that $Z = 685$, which corresponds to a surface charge of $Ze/(4\pi a^2)$ = 0.24 C/m².

The pair correlation function

In the next section we will be using a function called pair correlation function $g(r)$ which is the radial distribution that describe positional correlations among particles in the equilibrium fluid. One of the experimental way to assess the pair correlation function $g(r)$ is to use (confocal) microscopic imaging. The positions of the particles are then determined (usually with an accuracy of 0.05 μm) from the image analysis. For small particles, neutron or x-ray scattering is used. The probability of finding particles separated by r is then obtained, which is directly linked to $g(r)$. The function $g(r)$ describe how density varies as a function of distance from a reference particle. It is a normalized function, implying that $g(r) = 1$ when the density is equal to the bulk density. Let us consider a reference particle (black) and its neighbours (in colour):

distance from the centre of the middle black sphere
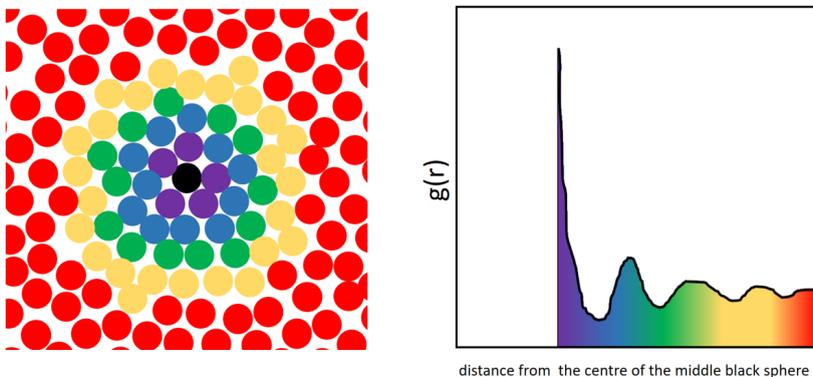
At position r/(2a) = 1 (the first peak), the probability to find neighbours is very high. These are the nearest neighbours that are touching the reference particle (r = 2a). Some minuscule distance away from these nearest neighbours there is statistically no particles (if there would be, they would overlap with the nearest neighbours, which is physically not possible), therefore one observes a dip between r/(2a) = 1 and r/(2a) = 2. At r/(2a) = 2 there is a next crowd of neighbours, and this continues until we are so far from the particles that the correlation becomes very weak and $g(r)$ = 1. In an ideal gas, there is no relevant correlation between neighbouring particles, which implies that $g(r) = 1$ all the time.

A typical example of $g(r)$ is given here:
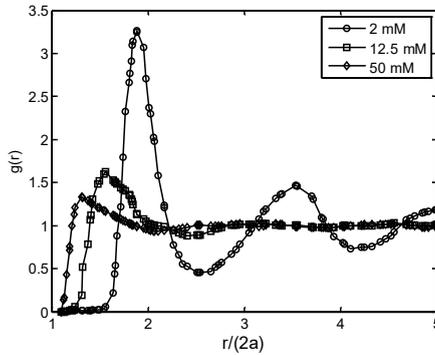
g(r) for 100 nm particles with a surface potential of 25 mV and dispersed in a KCl solution (concentrations are indicated in the legend). Volume fraction is 0.1.

One can note that increasing salt concentration decreases the separation between nearest neighbours and the amplitude of the peaks: in the limit of 50 mM added salt, the suspension behaves like a fluid (even nearly like a gas), and for 2 mM the structure of g(r) is close to the one of a crystal, where the position of nearest neighbours are extremely well defined.

It is also possible to perform computer simulation models to estimate g(r). One then has to use an interaction potential between particles and determine the particles' positions. One of the most classical function to describe the interaction between particles is the **Lennard –Jones potential**:

$$\Phi_{LJ}(r) = 4\Phi_0 \left[\left(\frac{a}{r}\right)^{12} - \left(\frac{a}{r}\right)^{6} \right]$$

which is used extensively to study molecules. Note the differences between this expression and the DLVO expression introduced in **Chapter 3**. The main reason for the differences lays in the fact that colloidal particles are much bigger than molecules, and many more effects are therefore incorporated in the DLVO theory.
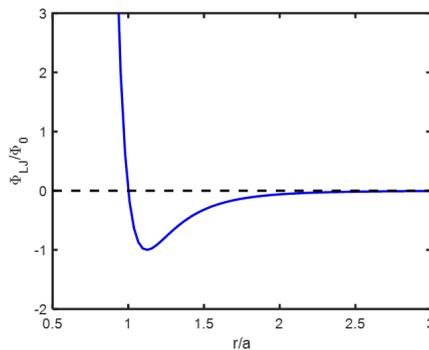
Colloidal charge determination from high-frequency shear measurements

For concentrated suspensions of (charged) spheres, it is possible to derive relationships between the high-frequency elastic modulus G'_∞ and the interaction potential[89]. The derivation of this relationship is beyond the scope of the present book, but its dimensionless form is given by:

$$G'_\infty \frac{a^3}{k_B T} = \frac{3\phi_s}{4\pi} + \frac{3\phi_s^2}{40\pi} \int_0^\infty g(r) \frac{d}{dr}\left(r^4 \frac{d\Phi(r)/(k_B T)}{dr}\right) dr$$

where $g(r)$ is pair correlation function. From the presence of the electric interaction potential $\Phi(r)$ the formula can be linked to the surface charge of the particles and the effects of double layer changes upon changes in salt concentration.

Typical frequency-dependent measurements are given here for latex nanospheres with a = 38 nm, dispersed in 30 mM of KCl:
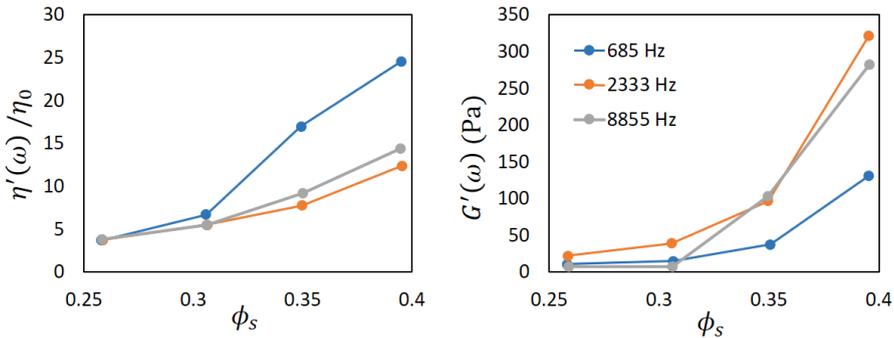
Left: normalized dynamic viscosity $\eta'(\omega)/\eta_0 = G''(\omega)/(\omega\eta_0)$ and right: dynamic storage modulus $G'(\omega)$, both as function of volume fraction

For the theoretical analysis to be valid, it is required that the frequency of the oscillation is high enough so that $\omega/(2\pi) \gg (D_S/d^2)$ where D_S is the short-time self-diffusion coefficient of a colloidal particle in the concentrated suspension and d is the mean interparticle separation. If we assume, to get an order of magnitude, that $D_S \sim D_0 = k_B T/(6\pi\eta a)$ (the theoretical diffusion coefficient of a colloidal particle in a dilute suspension) and $d \sim a$ we get $\omega/(2\pi) \gg 3$ kHz. Indeed, as can be seen in the figures, above that frequency both the storage modulus and the dynamic viscosity are fairly independent of ω.

---

[89] Bergenholtz, J., Willenbacher, N., Wagner, N. J., Morrison, B., Van den Ende, D., & Mellema, J. (1998). Colloidal charge determination in concentrated liquid dispersions using torsional resonance oscillation. Journal of colloid and interface science, 202(2), 430-440.

By performing a systematic study of G' at high frequency, as function of volume fraction and salt concentration, it is then possible to assess the surface charge of the particle: $g(r)$ can be measured, or approximated, as can $\Phi(r)$ (see **Chapter 3** about the theoretical evaluation of $\Phi(r)$ and **Chapter 6** for the AFM technique that enables to measure it). In the theoretical expression of $\Phi(r)$ the surface charge (or zeta potential) is the only unknown quantity. The theoretical expression for $\Phi(r)$ is inserted in the G'_∞ formula and the experimental data is fitted, giving access to the surface charge.

In the limiting case that the suspension is in a crystalline phase, it can be shown that the formula can be approximated by[90]:

$$G'_\infty \frac{a^3}{k_B T} \sim \frac{\phi_{max} N_{nn}}{10\pi} \left(\frac{\phi_{max}}{\phi_s}\right)^{1/3} \frac{a^2}{k_B T} \left[\frac{d^2 \Phi(r)}{dr^2}\right]_{r=2a_{eff}}$$

In the above, $\phi_{max} = 0.74$ if one assumes a **face-centered cubic** where each particle has $N_{nn} = 12$ nearest neighbours. Moreover

$$a_{eff} = a \left(\frac{\phi_{max}}{\phi_s}\right)^{1/3}$$

is a measure for the minimal separation of the particles (which is not $2a$ but $2a_{eff}$). In the limit of a fluid, one has the approximation:

$$G'_\infty \frac{a^3}{k_B T} \sim \frac{3\phi_s}{4\pi} - \frac{3\phi_s^2}{40\pi} g_{hs} \left[2\left(\frac{\phi_{max}}{\phi_s}\right)^{1/3}\right]^4 \frac{1}{k_B T} \left[\frac{d\Phi(r)}{dr}\right]_{r=2a_{eff}}$$

where for $0 < \phi_s < 0.5$:

$$g_{hs} = \frac{1 + \phi_s + \phi_s^2 - \phi_s^3}{(1 - \phi_s)^3}$$

(this formula is the **Carnahan-Starling** expression) and for $0.5 < \phi < 0.64$:

$$g_{hs} = \frac{1}{4\phi_s} \frac{1.21 + \phi_s}{0.64 - \phi_s}$$

Using this last expression for G'_∞ together with the DLVO expression developed in **Chapter 3**, one can show that it is possible to fit the measured data, and obtain the surface charge of particles:

---

[90] Buscall, R., Goodwin, J. W., Hawkins, M. W., & Ottewill, R. H. (1982). Viscoelastic properties of concentrated latices. Part 2.—Theoretical Analysis. J. Chem. Soc., Faraday Trans. 1, 1982,78, 2889-2899
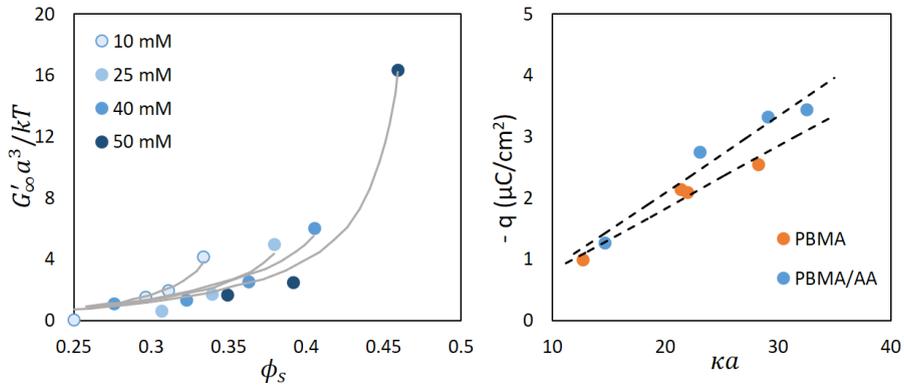
Left: Dimensionless high-frequency shear modulus as function of volume fraction for PBMA/AA dispersions. Right: surface charge density for two types of particles, evaluated following the method exposed above[91].

The PBMA/AA particles are latex spheres of size a = 43.5 nm, and the PBMA particles have a size of 38 nm. From their estimated surface charge, one can estimate that the PBMA/AA particles have a surface potential of -78 mV and the PBMA of -56 mV.

Shear thinning and shear thickening

A figure above showed the evolution of the rheological properties of a suspension in terms of flocs of particles that form or break as function of salt. We will here discuss a similar type of figure where the rheological property is given as function of shear rate[92].

[91] Bergenholtz, J., Willenbacher, N., Wagner, N. J., Morrison, B., Van den Ende, D., & Mellema, J. (1998). Colloidal charge determination in concentrated liquid dispersions using torsional resonance oscillation. Journal of colloid and interface science, 202(2), 430-440.

[92] See the book: Mewis, J., & Wagner, N. J. (2012). *Colloidal suspension rheology*. Cambridge University Press.
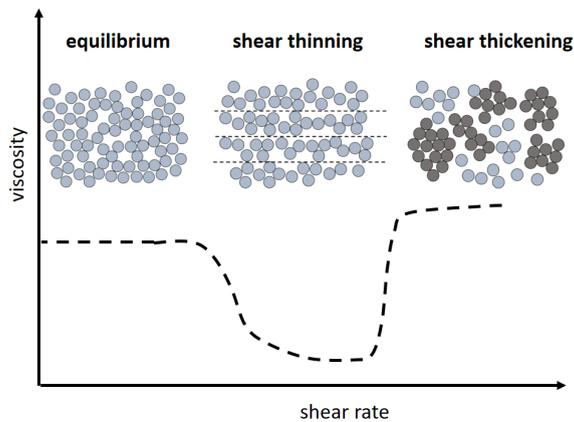
Sketch of the effect of shear rate on the viscosity of a stable (non-aggregating) concentrated colloidal suspension

At low particle concentrations, the suspension viscosity is nearly independent of shear. For higher particle concentrations, increasing the shear rate (and shear stress) leads to a marked *shear thinning* behaviour. This is of great importance in many industrial applications, as colloidal particles can be used made to flow, pour or spread with less effort at these higher shear rates. At even higher shear rates, the viscosity increases significantly. This *shear thickening* behaviour is often undesirable as it may damage processing equipment.

Shear thinning : the application of a shear flow distorts the equilibrium structure and leads to fewer particle interactions. The particles therefore re-arrange in the flow and the adopt an organization that permits flow with fewer particle encounters (the particles flow roughly in lanes, like cars on a highway). An example of shear thinning is given on the next page : instead of starting from a suspension, one has a gel made of clay and silt loosely aggregated in a fabric. By applying a pressure gradient (the famous Rissa landslide[93] was triggered by a small excavation and stockpiling along a lake-shore), fabric is locally broken and the particles rearrange in a (concentrated) suspension-like structure. The phenomenon propagates rapidly as at the front of the broken fabric inter-particle interactions are immediately changed.

Thixotropy : thixotropy is defined as *the continuous decrease of viscosity with time when flow is applied to a sample that has been previously at rest, and the subsequent recovery of viscosity when the flow is discontinued*. Thixotropy is therefore a time-

---

[93] L'Heureux, J. S., et al. (2012). The 1978 quick clay landslide at Rissa, mid Norway: subaqueous morphology and tsunami simulations. In Submarine mass movements and their consequences (pp. 507-516). Springer Netherlands. For a movie of the landslide, see: https://www.youtube.com/watch?v=3q-qfNlEP4A

dependent shear thinning property. Many gels and colloids are thixotropic materials, bentonite (montmorillonite) is a good example : when montmorillonite is dispersed in water, the platelets bind together electrostatically to form a house-of-cards structure and the liquid becomes viscous. When the structure develops further, the montmorillonite-dispersed liquid becomes a gel. However, shearing the gel returns it to a dispersed liquid. Landslides are generally thixotropic.
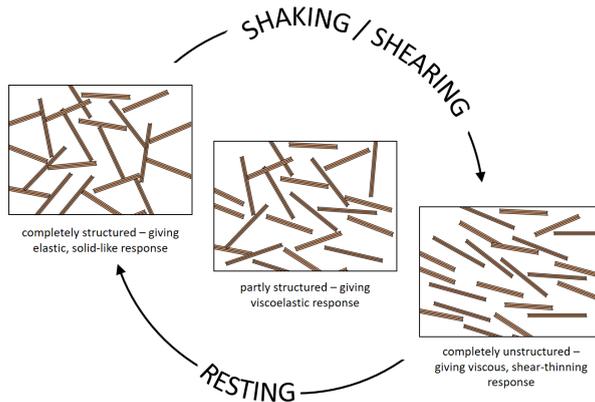
Breakdown of a 3D thixotropic structure[94]

Shear thickening : when the shear forces become sufficient, particle motions become hydrodynamically highly correlated. This induces the grouping of particles in clusters called **hydroclusters**. Hydroclusters are not necessarily aggregates : the coloured particles in the example are just in close contact, and if the shear rate would be set to zero the hydroclusters would disappear. The colour was used just to indicate the positions of the hydroclusters. The system would then return to its initial state thanks to Brownian motion and repulsion forces between particles. Of course, if the suspension would be unstable, particles could aggregate and the hydroclusters would then become "real" aggregates. When the shear would be reduced, these aggregates would remain and therefore there is no reversibility in a change of shear rate for this situation. Increasing flocculation due to shear rate in fact can lead to **rheopecty**.

---

[94] Barnes, H. A. (1997). Thixotropy—a review. *Journal of Non-Newtonian fluid mechanics*, 70(1), 1-33.

Rheopecty : rheopecty is defined as *the continuous increase of viscosity with time when flow is applied to a sample that has been previously at rest, and the subsequent recovery of viscosity when the flow is discontinued*. Rheopecty is therefore a time-dependent shear thickening property. Rheopectic fluids thicken or solidify when shaken. Examples of rheopectic fluids include gypsum pastes and printer inks. The synovial fluid in the joints that link our bones becomes thick the moment shear is applied in order to protect the joint and subsequently thins back to normal viscosity to resume its lubricating function.
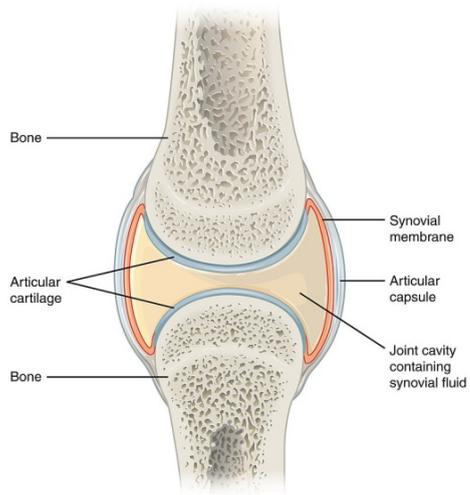
Marine clay, quick clay and landslides

Marine clay is clay that was deposited in a salty environment. The clay particles can self-assemble into different configurations (see **Chapter 5**) and the formed fabric can have very different properties. Construction in marine clays therefore presents a geotechnical engineering challenge. For example, swelling of marine clay can destroy building foundations in only a few years. Quick clay is a clay, which originally was a marine clay, with 'quick' properties. It is a fine-grained sediment where the grain structure may collapse even if the sediment is initially quite firm. Quick clay can be firm as long as it is undisturbed, but flows like liquid if it becomes overloaded or stirred, causing the loose grain structure to collapse. Quick clay landslides can developed rapidly when the firm clay liquefies.
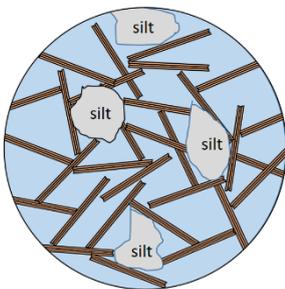
The clay and silt particles that constitute the quick clay were left behind during the retreat of the glaciers and deposited in a nearby sea. The particles are loosely packed and form a card house structure. When the sea retreats, the salt is washed out by fresh water (from rainfalls). The fabric's structure then becomes unstable, as the double layers have extended significantly, and repulsive forces started to act between the particles. Any mechanical stress can initiate liquefaction. The failure then rapidly propagates, leading to landslides. When a landslide encounters a river or a fjord, a tsunami can also be created as the quick clay flow pushes water. Marine clay is most widespread

in Norway, in the Trøndelag region and eastern Norway in particular. Quick clay are also found in parts of Canada and Sweden.

Quick clay landslide[95] at Lyngseidet (Norway) in 2010 (220.000 m³). The landslide was likely triggered by loading of fill along the shoreline. (See the white car in the upper part for scale)

Gels and hydrogels

We have so far implicitly spoken about the formation of a *gel* when the shear rate is zero in the case of thixotropic and rheopectic fluids. There is an important difference between a **concentrated suspension** and a **gel**. In general, gels are apparently solid, jelly-like materials:

---

[95]   https://www.ngu.no/en/topic/quick-clay-and-quick-clay-landslides;   Photo: Andrea Taurisano, NVE (with permission)

A **gel** is usually defined in colloid science as a substantially dilute polymer suspension, where the polymers are **cross-linked**. **Cross-linking** refers to the joining of polymer chains with covalent bonds. Cross-linking can occur during polymer synthesis or later with the addition of atoms or molecules which will share electrons with a part of the polymer chain.
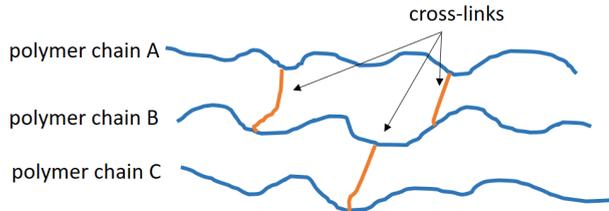
A **hydrogel** is a fabric made of polymeric chains that contains an enormous amount of water. There are two kinds of hydrogels:

1 – **chemical hydrogels**: these are formed by cross-linked polymers : the polymers/ polyelectrolytes are chemically bound to each other and form a fabric. Because of the presence of hydrophilic groups (-OH, -CONH, $-CONH_2$, $-SO_3H$) on the polymers water is "bound" to this fabric (more than 90% in weight, since there is a huge surface area available – the pores of the fabric can be as small as nanometers). The links between the polymers are permanent, and therefore, when agitated or stirred the hydrogel will break in pieces, and no water will be freed.

2 – **physical hydrogels** : the polymers are bound by reversible (non-permanent) links. These links can be due to hydrogen bonds, van der Waals forces, entanglements (when the polymer concentration is high the polymers are bound to touch each other and form a fabric). In this case, changing the properties, the gelation is reversible : one can think of adding water (in case of entanglement), changing the salinity (increase electrostatic repulsion between polymers and hence counteract the van der Waals attraction), changing pH, temperature… By agitating the gel, one generally observe a solid/liquid phase transition. This is the case for montmorillonite for example (see landslide example). Many of this type of hydrogels display thixotropy : they become fluid when agitated, but re-solidify when resting.

Hydrogels are studied a lot in biology (pharmacy) and food science, but their study is still limited in civil engineering : a flocculated clay for example can be seen as a complex hydrogel. Water is then trapped inside the floc structure like in a conventional hydrogel, but the flocs are forming a peculiar fabric of interacting macroscopic particles. The consolidation of this fabric is discussed in **Chapter 9 and 10**, but is still an on-going topic of research.

Sand/mud mixtures : fall velocity in shear flow

We now consider a colloidal mud *gel* in which sand particles are imbedded. (In a mud *suspension* the sand would simply settle to the bottom of the settling column.) If no shear is applied, the gel does not change over time and the sand particles stay imbedded. If shear is applied, one will observe that the sand particles start to settle[96]. This property causes critical problems: if settling occurs, the material loses its homogeneity, which can strongly affect its mechanical properties.

In slow flows, it is considered that the settling properties of suspended particles are not significantly affected by the material flow, and the sedimentation velocity is usually computed from the balance of gravity and drag forces. In order to avoid or slow down sedimentation, the only practical solution consists in inducing a sufficient agitation to the system which will induce some lift or dispersion forces to the particles. This principle is typically used in fluidization processes, in which a vertical flow of the interstitial fluid induces a drag force counterbalancing gravity force. For horizontal flows in conduits one may also rely on turbulence effects.
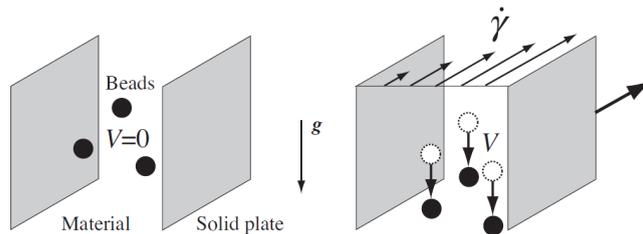
For many materials, the situation is different: the denser particles do not settle at rest because they are embedded in a yield stress fluid which is able to maintain the particles in their position. This situation is typically encountered with mortars or fresh concrete which are made of particles (sand or gravel) of density around 2.5 mixed with a cement-water paste of density around 1.5 (the density is here expressed as ρ/ρ_w). This is the same for toothpastes which contain silica particles of density 2.5 suspended in a paste of density close to 1. In that case the gravity force (weight of the particle) is counterbalanced by the elastic force from the fabric, as long as the yield stress obeys[97] :

---

[96] Talmon, A. M., & Huisman, M. (2005). Fall velocity of particles in shear flow of drilling fluids. Tunnelling and underground space technology, 20(2), 193-201.
[97] Ovarlez, G., Bertrand, F., Coussot, P., & Chateau, X. (2012). Shear-induced sedimentation in yield stress fluids. Journal of Non-Newtonian Fluid Mechanics, 177, 19-28.

$$\sigma_0 \geq \frac{\frac{4}{3}\pi(\rho - \rho_p)ga^3}{\pi a^2} \sim (\rho - \rho_p)ga$$

When the fluid is at rest, the weight of the sand particle in the mud fabric (gel) is compensated by the reaction of the gel underneath. When shear is applied, the reaction (static force) of the gel is changed in a friction (dynamic force) which is not enough to compensate for the weight anymore. The evaluation of the settling velocity of the sand particle in the sheared mud is a complex task, due to the fact that the variables become **tensors**.

Until now, we have simplified the definitions of the variables. The shear stress, for example (see beginning of the chapter) should properly be defined as:

$$\sigma_{xy} = \frac{F_{xy}}{S} = \eta \frac{dv_x}{dy}$$

where the subscript xy of the tensor σ_{xy} indicates that we are looking at the stress created by the the the $x-$ component of the force F_{iy} ($i = x, y, z$) on the slice y. This force is linked to the change in shear rate v_x over a thickness dy. Similarly, one can define σ_{xx}, σ_{xz}, σ_{yz}, etc...

Due the tensorial nature of the variables, the flow around the sand particle is complex. We therefore refer to Ovarlez et al. (see footnote) for a detailed derivation and explanations. In their experiments, two flows are present: the shear flow and the settling flow. The settling flow is considered to be a secondary flow as compared to the shear flow. This means that the shear flow is the major flow in the system. This has consequences on the estimation of the settling velocity, as the Stokes settling velocity has to be adapted. We assume a **Herschel-Bulkley** model for the stress:

$$\sigma = \sigma_0 + \eta_{HB}\dot{\gamma}^n$$

where σ is the macroscopic shear in the cell (and not the shear around the sand particle). The authors show that the settling of a sand particle is given by the following modified Stokes velocity, when $\sigma_0 \ll \eta_{HB}\dot{\gamma}^n$

$$\boldsymbol{v} = \alpha \frac{2}{9} a^2 \frac{\rho_p - \rho_w}{\eta_{HB}\dot{\gamma}^n} \boldsymbol{g}$$

where α is an adjustable parameter that accounts for the complexity of the flow around the sand particle.

Experiments, reported by Ovarlez et al. are given below. They are done on a suspension of glass beads in an emulsion (to control the viscosity). Experiments are performed in a **Couette** cell. There is a strong discrepancy between what is expected

from the analysis and the experimental measurements. While the experimental profiles show a rather narrow front with no broadening in time, the sedimentation front of the theoretical profiles gets broader and broader in time. This suggests that collective effects are at play, which tend to stabilize the front at a given speed. This point is discussed in **Chapter 8** where it is shown that sedimentation velocities are correlated in the horizontal plane over very long distances, of order of $20a\phi_s^{-1/3}$. For the particles in the study this lengthscale is of the order of 1 cm, i.e. of the order of the gap size (which is 1.9 cm). This may thus explain the observations.
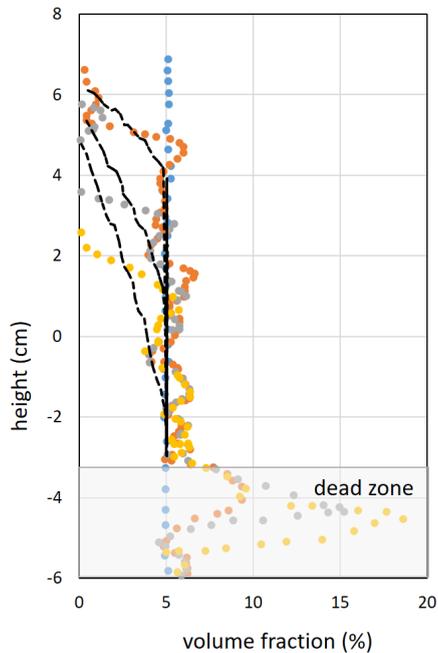
Vertical volume fraction profiles observed in the gap of a Couette geometry in a 5% suspension of 275 μm glass beads in a concentrated emulsion of yield stress σ_0 = 8.5 Pa after a 24 h rest (blue circles), after 15 min (orange circles), 30 min (grey circles) and 45 min (yellow circles) of shear $\dot{\gamma} = 4 \ s^{-1}$. The dotted lines are the theoretical profiles expected from the modified Stokes equation, taking into account the heterogeneity of the apparent viscosity in the sheared material, under the assumption that the sedimentation velocity of the suspension at a given radial position is set by the local viscosity $\eta(r) = \sigma(r)/\dot{\gamma}(r)$ of the sheared yield stress fluid only, independently of the sedimentation velocity in its surroundings.

A fifty cent rheometer for yield stress measurement[98]

Even though samples can be brought to the lab and analysed there, it is often convenient for an engineer to perform rapid and inexpensive tests in the field. The slump test schematized underneath is used extensively by civil engineers to estimate the ''workability'' of fresh concrete, but was shown to be useful for estimating the yield stress of mud slurries as well. "Workability" means that the slurry should have the proper yield stress and viscosity. For example, if fresh concrete is too stiff, the mixture will not flow into tight corners of moulding and if the concrete is too runny (it contains too much water), the concrete will flow better but the strength of the final hardened concrete will be reduced. Similarly, the flow properties of tailings in a waste disposal scheme need to be tailored for slope deposition: if the suspension is too thin, the material will result in little if any slope, while if it is too thick a tailing will result in the material being deposited around the discharge point and not flowing over the disposal area.
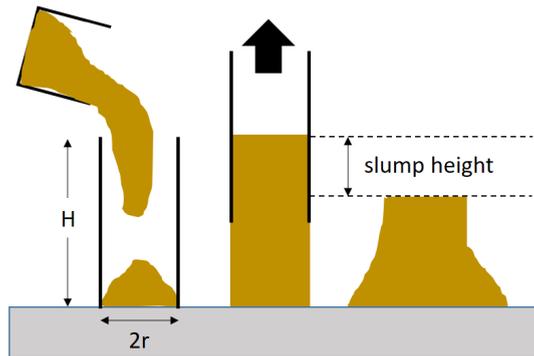
Schematic diagram of the slump test : a bottomless column is filled with the material to be tested. Lifting the column allows the material to collapse under its own weight. The height of the final deformed (or "slumped") material is measured. The difference between the initial and final heights is called the slump height.

The estimation of the yield stress is made as follows:

At t = 0 and at a given height z (where z = 0 is at the bottom of the column), the pressure is given by:

$$P(z) = \rho g (H - z) + P_0$$

---

[98] Pashias, N., Boger, D. V., Summers, J., & Glenister, D. J. (1996). A fifty cent rheometer for yield stress measurement. *Journal of Rheology (1978-present)*, 40(6), 1179-1189.

where $H = h + h_0$ and the atmospheric pressure is given by P_0. For simplicity we assume that $P_0 \ll \rho g H$ and therefore $P(z) = \rho g(H - z)$.

At t > 0 and at a certain height h_0 (see figure underneath) the material will experience a stress that is higher than the yield stress and the part of the material below that height will start to flow. The interface layer between the yielded and unyielded material is assumed to be a flat surface that moves down as the material beneath it flows. During the deformation stage it is assumed that all horizontal sections remain horizontal, and slumping is only due to radial flow. At the end of the slumping, the unyielded region represented by h_0 will have a stress distribution that is identical to that of the undeformed material (before slumping), while the stress in the remaining material is equal to the yield stress.

At t = 0, the sample is divided in slices with same thickness dz. Assuming incompressibility, conservation of volume ensures that at the end of the slumping the new thickness of each slice is given by

$$dz_1 = \frac{r^2}{r_1^2} dz$$
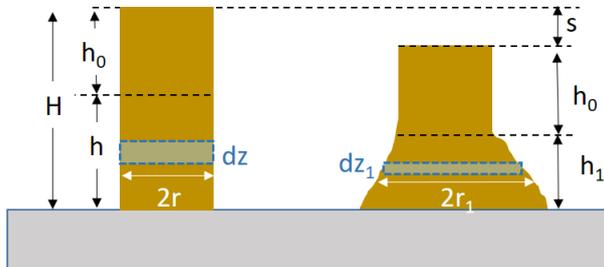
Schematic representation of the initial (left) and final (right) sediment distribution. The sample is divided in N different layer with same thickness dz at t = 0. At the end of the slumping, using conservation of volume, each layer thickness becomes dz₁

The amount of material above any given plane z will be the same before and after the slump (the flow only occurs in the cross-sectional area). This implies that the force due to gravity $m(z)g$ where $m(z)$ is the mass of the material above z and g is the gravity constant will be the same and lead to the following balance of forces at equilibrium:

$$\sigma r^2 = \sigma_0 r_1^2$$

where σ_0 is the yield stress. The shear stress that acts on a body when a pressure is applied to its normal direction is proportional to this pressure:

$$\sigma(z) = \alpha P(z) = \alpha \left[\rho g(H - z)\right]$$

From the equations above we get:

$$dz_1 = \frac{\sigma_0}{\sigma} dz = \frac{\sigma_0}{\alpha \rho g(H - z)} dz$$

By integration:

$$\int_0^{h_1} dz_1 = \int_0^h \frac{\sigma_0}{\alpha \rho g(H - z)} dz$$

$$h_1 = \frac{\sigma_0}{\alpha \rho g} \left[-\ln(h_0) + \ln(H)\right]$$

The height h_0 can be linked to the yield stress by:

$$\sigma(h) = \sigma_0 = \alpha \rho g h_0$$

As

$$s = H - h_0 - h_1$$

We obtain:

$$s = H - \frac{\sigma_0}{\alpha \rho g} \left[1 + \ln(H) - \ln\left(\frac{\sigma_0}{\alpha \rho g}\right)\right]$$

$$s = H - \frac{\sigma_0}{\alpha \rho g} \left[1 + \ln\left(\frac{\alpha \rho g H}{\sigma_0}\right)\right]$$

This expression is similar to the one found by Pashias and Boger cited above and the same as eq.(7.12a) in the book of Coussot[99]. Coussot uses a value of $\alpha = 1/\sqrt{3}$ which he derives in the book based on the **Von Mises criterion**, whereas Pashias and Boger uses $\alpha = 1/2$ which they say to hold for an ideal elastic solid. In fact, the rough assumption that is behind the result of $\alpha = 1/2$ is that a flow starts or stop in a layer when the maximum shear stress in the material reaches the yield stress value. One can rewrite the previous equation as:

$$s^* = 1 - \frac{\sigma_0^*}{\alpha} \left[1 - \ln\left(\frac{\sigma_0^*}{\alpha}\right)\right]$$

where we define the dimensionless coefficients:

---

[99] Coussot, P. (2005). Rheometry of pastes, suspensions, and granular materials: applications in industry and environment. John Wiley & Sons.

$$s^* = \frac{s}{H} \qquad ; \qquad \sigma_0^* = \frac{\sigma_0}{\rho g H}$$

We here give an example, in which "theory" is the plot for which $\alpha = 1/2$ and "theory2" for $\alpha = 1/\sqrt{3}$ (clearly a better fit) :
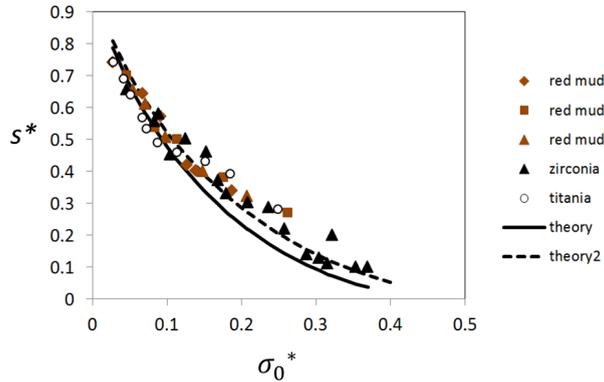
Dimensionless slump height as function of dimensionless yield stress for various slurries. The yield strength was obtained from vane tests. The different "red mud" (bauxite residue which was strongly flocculated with polyacrylamide) data corresponds to different structural states of the mud (the mud was allowed to be stirred for more or less long periods, causing structural decay, or process water was added which also lowered the yield stress). Experiments with titania and zirconia were done at the isoelectric point to maximize the yield stress (see next section).

Despite the fact that the "50 cent rheometer" test (also referred to as slump test) seems to provide some good qualitative (and quantitative) estimations of the yield stress for the given example, one should not forget that we did not address here the experimental problems of slurry/wall effects, inhomogeneous spreading etc... that can affect the results. The slump test is nonetheless an easy and fast test that can be used in the field. Coussot showed that for large slumps (where there is hardly an undeformed region) a spread test can be done to estimate the yield stress.

Maximum yield and zeta potential

In the legend of the previous figure, we have stated that at the **isoelectric point** the yield stress is maximum. The isoelectric point is defined as the pH value for which the zeta potential is minimum. At that point, van der Waals forces dominate, and the particles aggregate. For two of the previous samples, it was found that:
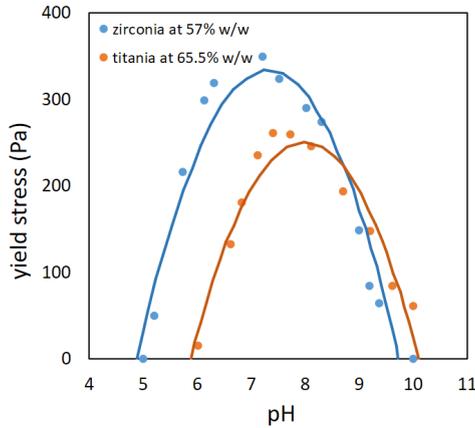
Yield stress versus pH : at the isoelectric point (pH = 7.2 and 7.6) the yield stress is maximum. At that point, the zeta potential is close to zero and the suspension is strongly flocculated.

A systematic study – theoretical and experimental - of the relation between yield stress and zeta potential has be performed by other authors[100].

The model introduced by Scales and co-workers (see footnote) evaluates the shear stress as a summation of all pair interactions calculated from the DLVO forces between the particles:

$$\sigma_0 = K_{struc}(\phi_s)\mathrm{F}_{DLVO}(r)$$

where K_{struc} is a network structural term dependent upon the particle size, the solids volume fraction ϕ_s and the mean coordination number. The term F_{DLVO} represents the DLVO force between two particles and r the distance between the centres of these particles. We note that:

$$\mathrm{F}_{DLVO}(r) = -\frac{d\Phi_{DLVO}}{dr}$$

where Φ_{DLVO} is the DLVO interaction potential, see **Chapter 3**. The expression for the DLVO used by Scales et al. is valid for identical spherical particles of radius a where $r \gg 2a$ and given by:

---

[100] Scales, P. J., Johnson, S. B., Healy, T. W., & Kapur, P. C. (1998). Shear yield stress of partially flocculated colloidal suspensions. AIChE Journal, 44(3), 538-544. and Zhou, Z., Scales, P. J., & Boger, D. V. (2001). Chemical and physical control of the rheology of concentrated metal oxide suspensions. Chemical Engineering Science, 56(9), 2901-2920.

$$F_{DLVO}(r) = \frac{Aa}{12(r-2a)^2} - 2\pi\varepsilon_0\varepsilon_r\zeta^2\kappa a \, \frac{\exp\big(-\kappa(r-2a)\big)}{1+\exp\big(-\kappa(r-2a)\big)}$$

where it was assumed that $\psi(a) = \zeta$ (the surface electric potential is the zeta potential, i.e. there is no Stern layer).

The expression for σ_0 implies that σ_0 will be maximum when $F_{DLVO}(r)$ is maximum. This will happen when the electrostatic repulsion is minimum, that is, when ζ is minimum (close to zero):
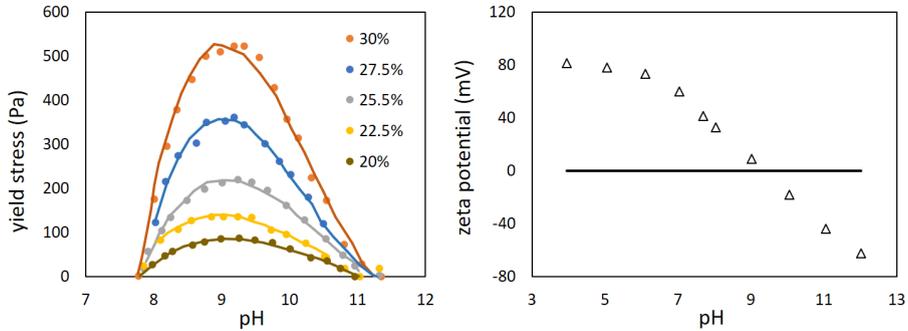
Yield stress (measured by the vane technique) and zeta potential (measured by electroacoustics) as function of pH and volume fraction for alumina particles suspended in 10 mM background electrolyte: note the relation between maximum yield stress and the isoelectric point (zero zeta potential).

For the same alumina suspension, one can plot the yield stress as function of the zeta potential squared (see underneath). The yield stress is maximum at the isoelectric point (where the zeta potential ζ is zero) and decreases in a parabolic manner as ζ increases. This is consistent with an increase in the electrical double layer repulsion, as the surface charge of the particles increases.

In addition to its dependence upon ζ and ϕ_S, a number of workers have demonstrated that the rheology of concentrated mineral suspensions is strongly dependent upon the mean particle size. For a given ϕ_S and microstructural arrangement, geometrical considerations dictate that the number of particles resisting an applied stress along a given shear plane is inversely related to the square of the particle size[101]. Note that the DLVO forces are proportional to the particle size. To complicate matters, the particle size can also impact upon microstructural factors

---

[101] Johnson, S. B., Franks, G. V., Scales, P. J., Boger, D. V., & Healy, T. W. (2000). Surface chemistry–rheology relationships in concentrated mineral suspensions. International Journal of Mineral Processing, 58(1), 267-304.

such as the **mean particle coordination number[102]** in a non-trivial manner. As a result, the theoretical particle size dependence of the yield stress properties of concentrated suspensions is not obvious although experimentally, systems consisting of small particles produce higher yield stresses than those containing larger colloids.
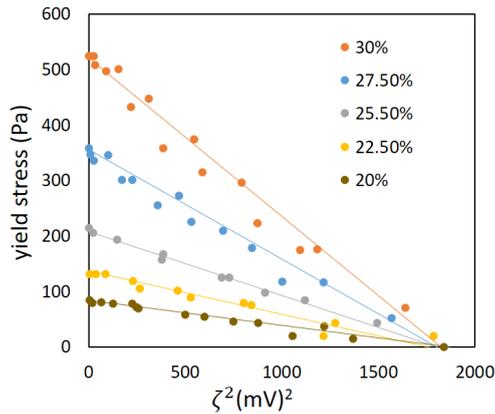
The shear yield stress of concentrated alumina as function of the zeta potential squared.

Illustrations

Batchelor (fair use)
http://www.damtp.cam.ac.uk/about/gkb/

Synovial joint (creative commons)
https://en.wikipedia.org/wiki/Synovial_joint

Gel (public domain)
https://fr.wikipedia.org/wiki/Dentifrice

---

[102] the mean number of particles touching a given one. See Jouannot-Chesney, P., Jernot, J. P., & Lantuéjoul, C. (2011). Practical determination of the coordination number in granular media. Image Analysis & Stereology, 25(1), 55-61.

Chapter 8

Settling of (concentrated) suspensions

In the previous chapters, we have discussed the properties of dilute suspensions in the context of settling (**Chapter 2**) and colloid interactions (**Chapter 3**). In **Chapter 7**, we investigated the rheological properties of concentrated suspensions. In this chapter, and in the next ones we are going to review the settling and consolidation properties of concentrated suspensions. Concentrated suspensions are very complex systems, owing to their complicated particle-particle interactions, and their influence on the hydrodynamics. Nevertheless concentrated suspensions are encountered in natural environments, and it is therefore important to have some of the keys to understand where lays the complexity of this type of systems and how to model them.

Concentrated suspensions and fluid mud

By definition, in a suspension, particles should be suspended (i.e. not resting on others). When all the particles are touching each other in some way, a large structure is formed which gets the name of **fabric** or **gel**. In colloid science, a gel is usually defined as a dilute cross-linked polymeric system, which exhibits no flow when in steady-state (see also **Chapter 7**). The name *gel* originates the word *gelatine*, which is an irreversible hydrolysed polymer (collagen). In civil engineering, when studying mud, the word gel is used to design a mud with a substantial yield stress.

Examples of concentrated mud. Left: concentration between 30 and 500 g/L. Right: concentration between 1 and 30 g/L

Fluid mud is a high concentration aqueous suspension of fine-grained sediment in which settling is substantially hindered by the proximity of sediment grains and flocs, but which has not formed an interconnected matrix of bonds strong enough to eliminate the potential for mobility[103]. Other terms used to denote fluid mud include **fluff (UK)**, **vloeibare sliblaag (NL)**, and **crème de vase (FR)**. Fluid mud typically exhibits concentrations of tens to hundreds of grams per liter and bulk densities between 1,080 and 1,200 kg m^{-3}.

---

[103] McAnally, W.H. et al. 2007. Management of fluid mud in estuaries, bays and lakes. Part I: Present state of understanding on character and behaviour. Journal of Hydraulic Engineering, Vol. 133, No. 1, 9-22

Fluid mud exists only because of its transient behaviour: if fluid mud is left at rest, it will eventually consolidate and form a gel (fabric). If, on the other hand, fluid mud can be picked-up by currents or waves, the local mud concentration will decrease as mud particles are dispersed and the fluid mud will become an usual mud suspension.

Modelling concentrated suspensions

A lot of work has been performed to understand the settling of particles in non-dilute conditions. Theoretical models for systems where more than two particles interact are usually very complicated, and this is why their study remains an open field of research.

Microscopic mechanics can be upscaled to give macroscopic (bulk) properties thanks to statistical physics. The difficulties are not just linked to solving the hydrodynamic equations with better, faster computers. The collective interactions between the particles can give rise to quite unexpected qualitative behaviour, actually often much simpler than the microscopic motions seem to suggest[104]. A litre of gas for instance could be studied by tracking the behaviour of its constituting molecules (about $N \sim 0.3 \times 10^{23}$ particles for one litre for usual pressure) – a huge task – however the macroscopic behaviour of a simple ideal gas is very well described by the simple relation: $PV = NkT$ where P is the pressure, V the volume, k the Boltzmann constant, T the temperature .

Statistical physics makes use of **probability distribution functions**. In Chapter 7, the **pair correlation function** $g(r)$ was introduced. The pair correlation function is used to establish the probability distribution function. For example, the probability density function $\Psi(r)$ for finding a particle at r given that there is a particle at $r = 0$ is written:

$$\Psi(r) = \frac{N}{V} g(r)$$

where N/V is the average number density of particles within the sample. As particles do not interpenetrate, the probability to find a particle between 0 and *2a* (where *a* is the radius of one particle) is zero. This volume, that cannot be occupied by other particles, is called the **excluded volume**. Moreover:

$$\int_{V(r>2a)}^{V_{tot}} \Psi(r) dV = N - 1$$

as the number of particles in the total volume considered V_{tot} is equal to $N - 1$ (the one at $r = 0$ is not counted). Because of the excluded volume, the reference particle

---

[104] Guazzelli, E., & Morris, J. F. (2011). A physical introduction to suspension dynamics (Vol. 45). Cambridge University Press.

influences the particle positioning near itself: this leads to interesting and physically significant microstructures in concentrated suspensions.

Settling of concentrated suspensions

Dilute suspensions

In Chapter 2, we have discussed the settling of individual particles according to **Stokes' law**. We recall that this settling velocity can be expressed as:

$$\boldsymbol{v_0} = \frac{2}{9} a^2 \frac{\rho_p - \rho_w}{\eta} \boldsymbol{g}$$

We here added the subscript *0* to indicate that we refer to an individual particle settling according to Stokes. In very dilute suspensions the particles can be assumed to settle according to Stokes.

Doublet and triplet of spheres

Using reversibility and symmetry principles one can show that two identical spheres close to each other fall at the same velocity and therefore do not change their orientation and separation. The pair has a sideways motion as there is a horizontal component of the fall motion (except when the angle between the centres of the spheres and the vertical, θ, is 0 or $\pi/2$) and they fall faster than when they are alone. The flow of the two spheres can be obtained by the **method of reflections**. One can show that:

$$\frac{\boldsymbol{v}_{\textbf{doublet}}}{\boldsymbol{v_0}} = 1 + \frac{3a}{2r} \quad \text{for} \quad \theta = 0$$

and:

$$\frac{\boldsymbol{v}_{\textbf{doublet}}}{\boldsymbol{v_0}} = 1 + \frac{3a}{4r} \quad \text{for} \quad \theta = \frac{\pi}{2}$$

From these equations we see that two spheres falling next to each other (in an horizontal plane) will fall slower than two spheres falling one above the other. In the limit of touching spheres, it is easy to verify that the particles will settle with a velocity $1.75\ v_0$ for $\theta = 0$ and $1.375\ v_0$ for $\theta = \pi/2$.
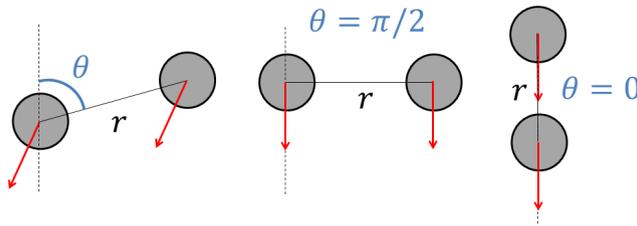
Schematic representation of the settling of two spheres separated by a distance r. Note that for an any angle θ between 0 and π/2 there is an horizontal component to the velocity (which is represented by the red arrows)

When a third sphere is introduced (the three-body problem) the behaviour is different as the particles do not usually maintain a constant separation, and the configuration is unstable. In a sedimenting triplet equally spaced on a vertical or an horizontal line one can then show that the middle particle will settle the fastest.

Concentrated suspensions

For the sedimentation of a suspension of many spheres, things become more complicated. One might try to compute the mean sedimentation velocity of particles by summing the effects between pairs of particles. The velocity of a particle can then be written:

$$v = v_0 + \Delta v$$

where Δv is the incremental velocity due to a second particle located at distance *r*. For accounting for more particles, one could average over all possible separations which occur with a probability $\Psi(r)$ (the probability of finding a sphere with its center at *r* given there is a particle at *r = 0*, see above). This would give:

$$v = v_0 + \int_{V(r>2a)}^{V_{tot}} \Delta v \cdot \Psi(r) dV$$

From a dimension analysis, one can infer that the flow field Δv decreases at first order as 1/r (see the doublet examples above), that $\Psi(r)$ scales at first order as a constant independent of *r* since for a reasonably dilute suspension $g(r) = 1$ and that:

$$\Psi(r) \approx \frac{N}{V}$$

The element of volume dV scales as r^3. This implies that

$$\int_{V(r>2a)}^{Vtot} \Delta\boldsymbol{v}\Psi(r)dV \sim L^2$$

where *L* is a characteristic length of the settling column. This dependence would imply a very strong divergence (imagine that the considered column would represent the sea…), and is due to the fact we omitted the long-range hydrodynamic interactions in our analysis. This divergence is not found experimentally, as, on the contrary, the velocity of particles in a concentrated suspension is even *lower* than the one of single particles. This is what is called **hindered settling**: the velocity of the particles is decreasing, relative to the Stokes velocity, with increasing particle concentration, as particles start to be in each other's way. The mean settling velocity of a particle in a concentrated suspension is defined as:

$$\boldsymbol{v} = f(\phi_S)\boldsymbol{v_0}$$

where $f(\phi_S)$ is the hindered settling factor which is function of the volume fraction ϕ_S. For a dilute suspension, one has $f(\phi_S) = 1$. A very widely used empirical correlation function is the one attributed to **Richardson and Zaki** in 1954[105]:

$$f(\phi_S) = (1 - \phi_S)^n$$

where n is a coefficient generally found to be close to 5 for the settling of monodisperse colloidal spheres (see figure below). The experimental data are from Nicolai et al.[106]
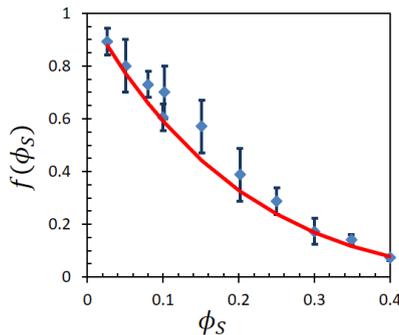
Hindered settling function: Richardson-Zaki (red line) with n = 5, and experimental data (symbols) for colloidal spheres

---

[105] Richardson and Zaki, Sedimentation and fluidization: Part I. Trans. Inst. Chem. Engrs., 32, 35-53 (1954)
[106] Nicolai, H., et al. "Particle velocity fluctuations and hydrodynamic self-diffusion of sedimenting non-Brownian spheres." Physics of Fluids 7.1 (1995): 12-23.

Experimental observations for the hindered settling velocity in the dilute regime ($\phi_S \ll 1$) have established that then $f(\phi_S) \approx 1 - 5\phi_S$, in accordance with $(1 - \phi_S)^n \approx 1 - n\phi_S$.

One cause for hindered settling is linked to the presence of the bottom of the settling column which imposes a back-flow due to volume conservation. The condition to express this constant volume flow is similar to the one that will be discussed in **Chapter 9** :

$$(1 - \phi_S)v_{w/lab} + \phi_S v_{s/lab} = 0$$

The velocity of the solid phase (particles) in the rest frame of the laboratory is given by $v_{s/lab}$ and the velocity of the water in the rest frame of the laboratory is given by $v_{w/lab}$. Newton's equation, with the gravity acting as only external field gives:

$$-6\pi\eta a(\boldsymbol{v}_{s/lab} - \boldsymbol{v}_{w/lab}) + \frac{4}{3}\pi a^3 \rho_s \left(1 - \frac{\rho}{\rho_s}\right)\boldsymbol{g} = 0$$

The first term represents the Stokes drag force and the second the force of gravity, compensated for Archimedes. The fluid density ρ is given by:

$$\rho = (1 - \phi_S)\rho_w + \phi_S \rho_s$$

Using the fact that:

$$1 - \frac{\rho}{\rho_s} = (1 - \phi_S)\frac{\rho_s - \rho_w}{\rho_s}$$

and taking into account the back-flow, one can rearrange the last equations to obtain[107]:

$$\boldsymbol{v}_{s/lab} = \frac{2}{9}a^2 \frac{(\rho_s - \rho_w)}{\eta}(1 - \phi_S)^2 \boldsymbol{g}$$

$$\boldsymbol{v}_{s/lab} = \boldsymbol{v_0}(1 - \phi_S)^2$$

Therefore, if one would neglect particle-particle interactions and hydrodynamic effects, we would find n = 2 in the Richardson and Zaki formulation. One of the $(1 - \phi_S)$ results from the return flow, and the other arises because a typical sphere

---

[107] Oliver, D. R. "The sedimentation of suspensions of closely-sized spherical particles." Chemical Engineering Science 15.3 (1961): 230-242. See also: Gourdin-Bertin, S., and C. Chassagne. "Onsager's reciprocal relations for electroacoustic and sedimentation: Application to (concentrated) colloidal suspensions." The Journal of chemical physics 142.19 (2015): 194706. Note the mistake in eq.(29) of that last article as can be found by inserting eq.(26) in eq.(28).

was considered to be settling in a fluid with density equal to that of the suspension, and not that of the water alone.

Other corrections can be done. One could in particular account for the change in viscosity. We have seen in **Chapter 7** that a possible expression for the viscosity is:

$$\eta(\phi_S) = \eta_0 \left(1 - \frac{\phi_S}{1 - c\phi_S}\right)^{-2.5}$$

where η_0 is the viscosity of water. If we assume that $1 - c\phi_S \approx 1$ (which is valid for not too concentrated suspensions), we get:

$$\eta(\phi_S) \approx \eta_0(1 - \phi_S)^{-2.5}$$

Therefore if we define the settling of an individual particle $\boldsymbol{v_0}$ by:

$$\boldsymbol{v_0} = \frac{2}{9}a^2 \frac{(\rho_s - \rho_w)}{\eta_0}\mathbf{g}$$

we now find:

$$\boldsymbol{v}_{s/lab} = \boldsymbol{v_0}(1 - \phi_S)^{4.5}$$

The exponent is now 4.5, quite close to the experimentally found one for spheres. This approach is however questionable, as it is not obvious that the viscosity experienced by a particle can correctly be approximated by the mean viscosity of the suspension.

Non-spherical particles

Needless is to say that for particles that are not monodisperse spheres, as assumed until here, the settling rate dependence on volume fraction is much more complex. In numerical simulations, it has been shown that suspensions that are homogeneous at t = 0, with random positions and orientations exhibit a peculiar settling behaviour. A the sedimentation proceeds, a broad suspension front forms at the interface and the clear fluid appears at the top. Careful observation of the bulk shows that the suspension does not remain homogeneous: particles tend to group into clusters.

Particles are also observed to rotate and to orient preferentially in the vertical direction, with a strong correlation between centre-of-mass positions and orientations; while the orientation remains random inside the dense clusters, the alignment in the vertical direction is much clearer in their periphery, where a strong vertical shear exists. For large ϕ_S the settling found experimentally could be described with a Richardson-Zaki law with n = 9.

Fluctuations in settling[108]

Large velocity fluctuations have been measured by tracking marked spheres in an otherwise transparent sedimenting suspension. The measurements were not collected immediately after the initial mixing of the suspension: the suspension was allowed to settle for some time and reach a steady behaviour before tracking started. The fluctuations did not appear to be affected by the size of the container in the large containers used.
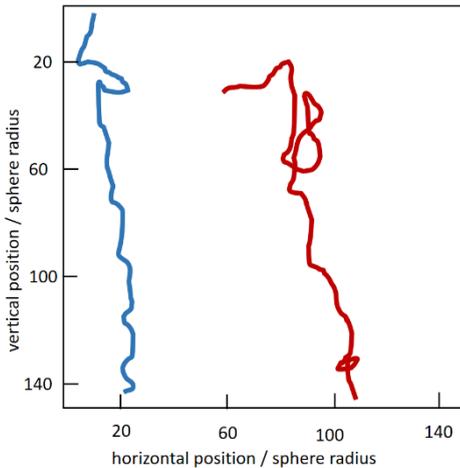
Tracking of two marked spheres in the midst of a 30% volume fraction sedimenting suspension of unmarked spheres made optically transparent by matching the index of refraction of the suspending fluid to that of the glass spheres. The particle trajectories are tortuous and exhibit large as well as small loops as the spheres sometimes moved upward against gravity.

For many years, theoretical estimates and numerical simulations predicted the steady-state fluctuations of the velocities of spheres to increase with the size of the container, whereas experiments found no such variation. In fact, the correlation length of the velocity fluctuations was found experimentally to be 20 interparticle separations, $20a\phi_S^{-1/3}$. This value of the correlation length was observed for volume fractions from 10^{-4} to 0.4. When the minimum dimension of the container L was less than $20a\phi_S^{-1/3}$ it was found that the velocity fluctuations Δv would scale as:

---

[108] Guazzelli, Élisabeth, and John Hinch. "Fluctuations and instability in sedimentation." *Annual review of fluid mechanics* 43 (2011): 97-116.

$$\Delta v \sim v_0 \sqrt{\phi_S \frac{L}{a}}$$

While the steady-state velocity fluctuations are independent of the size of the container if $L > 20 a \phi_S^{-1/3}$, the early fluctuations do depend on the container. The initial magnitude of the velocities again scales as $v_0 \sqrt{\phi_S L / a}$. Some experimental results are given below. These are obtained for suspensions of glass spheres of radius 150 microns, volume fraction 0.1% and density 4 g/cm^3. They were suspended in silicon oil to insure that the Reynolds number remain extremely small. The time t = 0 corresponds to the time just after mixing the column.
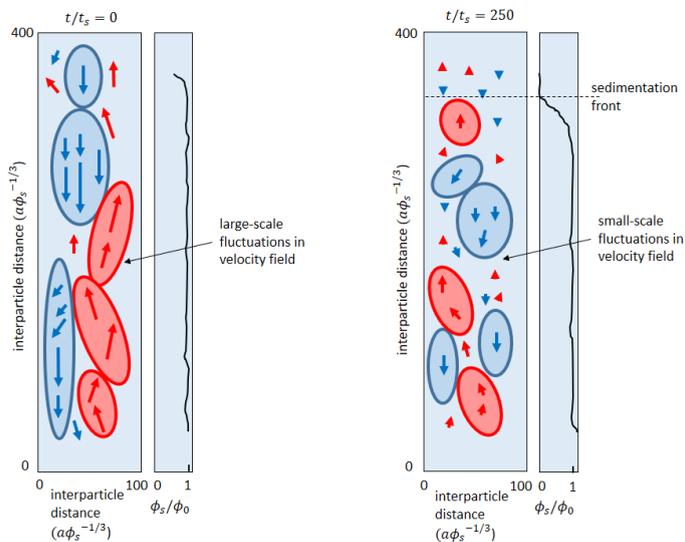
Relaxation of the large-scale fluctuations. The velocity field is from particle-image velocimetry sampling the whole-cell height and width within a laser sheet located in the middle plane, and the concentration profile ϕ_S / ϕ_0 is from light-attenuation measurements through the suspension. The timescale is the Stokes time $t_S = a/v_0$, i.e. the time for an isolated sphere to settle a distance equivalent to its radius.

The fluctuations magnitudes are strongly anisotropic with vertical velocities around four times the horizontal velocities. The initial strong large-scale fluctuations decay in time to weaker small-scale fluctuations. These small-scale fluctuations remain in a steady state until the sedimentation front arrives. This reduction of the initially large fluctuations to a smaller steady value independent of the size of the container, seen initially at low volume fractions ϕ_S, has been also observed at larger ϕ_S.

This behaviour can be partially understood by modelling the system by heavy blobs of particles falling to the bottom and light blobs rising to the top.

Falling blobs

Let us consider a blob of size d. This blob would contain on average nd^3 particles (where n is the number density of particles), but there would be statistical fluctuations of $\sqrt{nd^3}$ in the number if the particles were positioned randomly and independently. This fluctuation in the number gives a $m\sqrt{nd^3}$ fluctuation in the mass of the blob, where m is the mass of a particle compensated for Archimedes. Balancing the fluctuation in weight with a Stokes drag on the blob yields a fluctuation in velocity of:

$$\Delta v \sim \frac{mg\sqrt{nd^3}}{6\pi\eta d} = v_0\sqrt{\phi_S\frac{d}{a}}$$

As larger blobs give larger fluctuations in velocity, one would expect to see those corresponding to the largest spherical blob that can be fitted into the container. This largest size corresponds to the largest of the height, width, and depth of the column, which we have noted *L* above. We therefore find, in accordance with observations, that the velocity fluctuations Δv scale as:

$$\Delta v \sim v_0\sqrt{\phi_S\frac{L}{a}}$$

Falling clouds

We have just introduced the concept of a "blob" of particles. One other example of blob that is quite relevant for different studies is when the blob of particles is found in water. This is the case for example when a concentrated suspension is pipetted into a jar containing water. It is common to define as a **cloud** a blob containing a very large number of particles surrounded by clear water.

In **Chapter 2** we have illustrated what happens when regions of high particle concentration come into contact with regions of low particle concentration: the particles will diffuse to regions of low particle concentration. Implicitly, we have there assumed that the hydrodynamics could be neglected: the pipette is just put into contact with the water and the particles will start to diffuse thanks to Brownian motion (helped by gravity). This is also how the pipetting for particle size is performed in settling studies (see **Chapter 5**).

If a blob (or cloud) of particles would be deposited as such (for instance pushed with a syringe) in clear water, a different behaviour would be observed.

(a) t* = 0

(b) t* = 400

(c) t* = 700

(d) t* = 800
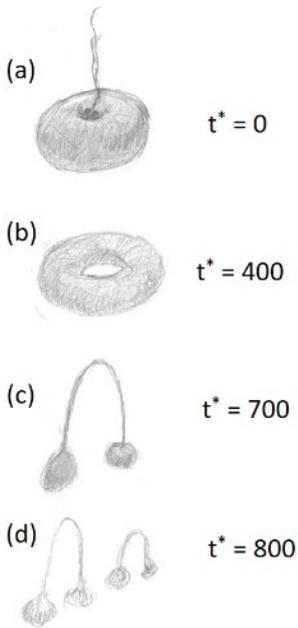
A cloud of particles sedimenting in a viscous fluid evolves into a torus that becomes unstable and breaks up into secondary droplets, which deform into tori themselves in a repeating cascade. This instability occurs even in the complete absence of inertia and without the need to perturb the initial shape. The particles circulate in closed toroidal streamlines as predicted by a continuum approach in which the cloud is modelled as an effective medium of higher density. Fluctuations arising from the multibody character of the hydrodynamic interactions cause particles to depart from these streamlines and to be carried into a downstream tail. Because the lost particles are those located in the circulation rim, this depletes the central region and leads to torus formation. The mechanism responsible for the further expansion of the torus remains unclear, but the breakup can be described as a change in the flow topology that occurs when the torus reaches a critical aspect ratio. Simulations using a point-particle approach containing the minimal physics of the long-range interactions capture this dynamics. Faster breakup is observed for clouds of fibers due to the self-motion of the anisotropic particles.

A cloud of particles can primarily be seen as a heavy fluid drop which has a different density and specific viscosity than the lighter water surrounding it. The problem was originally solved by **Hadamard** and **Rybczyński** in 1911[109]. The boundary conditions at the particle's surface are that the fluid velocities are continuous (contrary to the no-slip condition when one considers a hard sphere), and that there is no surface tension between the drop and the surrounding fluid. One finds that the terminal velocity of the spherical cloud is then given by the balance of the Hadamard-Rybczyński drag force and the weight (compensated for Archimedes):

$$U_S = \frac{N \frac{4}{3}\pi a^3 (\rho_s - \rho_w)g}{2\pi\eta \frac{2+3\lambda}{\lambda+1}R} = N \frac{6a}{2\frac{2+3\lambda}{\lambda+1}R}v_0$$

where N is the number of particles in the cloud of radius R and λ is the ratio between the cloud viscosity and the viscosity of water. The factor $2(2+3\lambda)/(\lambda+1)$ ranges from 5 for a cloud with a low volume fraction (where $\lambda \approx 1$) to 6 for a concentrated cloud (where $\lambda \gg 1$).

---

[109] Hadamard, J. S. (1911). "Mouvement permanent lent d'une sphere liquide et visqueux dans un liquide visqueux". CR Acad. Sci. (in French). 152: 1735–1738 and Rybczynski, W. (1911). "Über die fortschreitende Bewegung einer flüssigen Kugel in einem zähen Medium". Bull. Acad. Sci. Cracovie, A. (in German): 40–46.

The most remarkable feature observed during the cloud fall is the collective motion followed by the particles. While settling, the particles circulate in a toroidal vortex inside the cloud, similarly to the heavy fluid inside a drop sedimentating in a lighter fluid. As a result the cloud remains a cohesive entity for long times, maintaining a sharp boundary between its particle-filled interior and the clear fluid outside. It is the chaotic fluctuations arising from the many-body character of the hydrodynamic interactions that cause the particles to cross the boundary of the closed toroidal circulation which, unlike the drop, is not a material surface. Some of the particles may thus be carried by the outside flow into a vertical tail at the rear of the cloud (see figure (a) above).
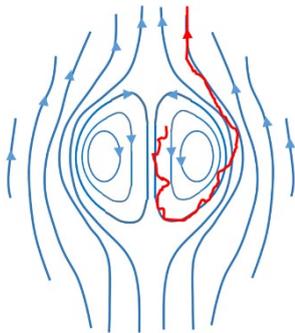
Sketch of the settling of a spherical cloud of particles showing the toroidal circulation of the particles inside the cloud and the particle leakage (in red) at the rear of the cloud (in a frame relative to the moving particle)

Clouds made of a small number of particles are found to keep their shape until they disintegrate owing to the constant loss of particles. Clouds having a large number of particles (N > 500) become unstable (see figures (b,c,d) above).

The evolution of a cloud of particles is a good example of how the long-range nature of many-body hydrodynamic interactions and the coupling between hydrodynamics and the microscopic arrangement of particles lead to a collective effect. While the suspension can be modelled as an effective medium with excess mass, the discrete nature of the suspension is a fundamental ingredient in understanding phenomena as leakage at the rear and destabilization of the cloud.

Wall effects
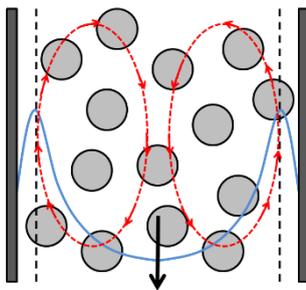
We have already discussed above the influence of the size of the column on the settling. Even though the relative sedimentation velocity is independent of the shape of the column (at steady-state) there exists a global convection of the suspension called **intrinsic convection** that was discovered by Prof. Peter Mazur of Leiden University and coworkers in 1985[110]. This convection effect exists even for dilute suspensions. The principle is sketched on the figure opposite: close to the column walls there is a small

---

[110] Beenakker, C. W. J., and P. Mazur. "Is sedimentation container-shape dependent?." Physics of Fluids (1958-1988) 28.11 (1985): 3203-3206.

depleted region (symbolized by the dashed lines) of size the radius of a particle. Between the wall and each dashed line the centre of particles cannot penetrate more than a distance corresponding to their radius. Conservation of volume and no-slip condition at the wall (= the velocity of fluid is zero at the walls) ensure that the hydrodynamic profile is given by the blue line on the figure[111]. This profile corresponds to a **Poiseuille** flow with a slip velocity $v_{slip} = 9\phi_S v_0/4$ at a particle radius distance from the walls, and a maximum downwards velocity in the centre of the column equal to $-v_{slip}/2$. This flow induces two vortices in the column with an ascending velocity near the wall and a descending in the middle (red lines in the figure). As this global convection is superimposed on the settling motion of the particles relative to the suspension (black arrow), the particles should settle faster in the centre of the cell than near the side walls. Experiments have shown the existence of this intrinsic convection, however the effect was observed to be much smaller than predicted. For a concentrated suspension, an ordering of particles near the wall was observed, as particles piled up against the side walls. Existing models do not take into account this effect.

Chemical potential, osmotic pressure and thermodynamics

Until now we have focussed on microscopic and mesoscopic descriptions of the settling of (concentrated) suspensions. Even though progress has been made in understanding the related observed features, it is clear that the complexity of problem prevent the use of microscopic and mesoscopic theories for large-scale engineering problems, such as predicting the settling velocity of a concentrated suspension. We have however already shown that even though the systems might be complex at a microscopic scale, it is possible to derive some simple relationships for their mesoscopic and macroscopic behaviour. Examples are the empirical Richardson-Zaki relationship and the ideal gas theory ($PV = NkT$). The mesoscopic/macroscopic description of a system can (in colloid science) be linked to statistical physics : the numerical simulations of the falling fibers given above or of interacting molecules in a gas are examples of it. There, the microscopic interactions between particles are simulated (one example of such interactions is given in **Chapter 7**, where we introduced the Lennard-Jones potential) and the macroscopic behaviour of the suspension can be inferred from the results. Another approach to macroscopic descriptions of a complex system is through thermodynamics.

In the previous chapters of this book, we have defined various concepts relevant for colloid science without explicitly referring to thermodynamics. In **Chapters 2** and **3**, we have discussed the **osmotic pressure** and the **Boltzmann distribution**. These concepts are both linked to **statistical physics** and **thermodynamics**. Statistical mechanics provides a framework for relating the microscopic properties of

---

[111] Bruneau, D., et al. "Intrinsic convection in a settling suspension." Physics of Fluids (1994-present) 8.8 (1996): 2236-2238.

individual atoms and molecules to the macroscopic or bulk properties. It enables to explain thermodynamics as a natural result of statistics, classical mechanics, and quantum mechanics at the microscopic level. Both statistical physics and thermodynamics originate from the work of **Gibbs** (see underneath). It is not the purpose of the present book to go too much into thermodynamic concepts, but it is useful to get some very basic understanding of it. Indeed, it is for example through thermodynamics that the osmotic pressure and the Boltzmann distribution can best be understood. Theoretical work in thermodynamics and statistical physics is still ongoing in order to understand the macroscopic behaviour of complex systems such as concentrated suspensions, slurries and transport in porous media .

Chemical potential

An important thermodynamic parameter is the **chemical potential** of component i (that can be an ion species, a colloid or the solvent), noted μ_i. The chemical potential is a form of **potential energy** that can be absorbed or released during a chemical reaction, and that can change during a phase transition. The chemical potential is defined as the partial derivative of the **free energy** with respect to the amount of the considered species (all other species' concentrations in the mixture remaining constant). At chemical equilibrium or in phase equilibrium the total sum of chemical potentials is zero, as the free energy is at a minimum. Particles tend to go from a higher chemical potential to a lower one : a simple example is the diffusion of particles from high to low concentrations that we have introduced when we discussed Fick's law in **Chapter 2**. The microscopic explanation for this is based on kinetic theory and the random motion of particles. However, it is simpler to describe the process in terms of chemical potentials: for a given temperature, a particle has a higher chemical potential in a higher-concentration area, and a lower chemical potential in a low concentration area. Movement of particles from higher chemical potential to lower chemical potential is accompanied by a release of free energy.

The total chemical potential can be split into an internal and an external chemical potential, where the external chemical potential originates from external gradients such as electric field and gravity.

For example let us consider an electrolyte solution in the presence of a charged wall or colloid. The electrochemical potential μ_i of ion species i is a function of the electric potential $\psi(r)$ originating from the presence of the charged wall or colloid. It is given by:

$$\mu_i(r) = q_i \psi(r) + k_B T \ln\left(\frac{n_i(r)}{n_0}\right) + \mu_i^0$$

where μ_i^0 is a reference value which only depends on temperature T, k_B is the Boltzmann constant and n_0 a reference density. The other variables are defined in **Chapter 3**.

**Josiah Willard Gibbs,
colloid science and thermodynamics**

Josiah Willard Gibbs (1839 – 1903) was an American scientist who made important theoretical contributions to physics, chemistry, and mathematics. His work on the applications of thermodynamics was instrumental in transforming physical chemistry into a rigorous deductive science. Together with **James Clerk Maxwell** and **Ludwig Boltzmann**, he created statistical mechanics (a term that he coined), explaining the laws of thermodynamics as consequences of the statistical properties of ensembles of the possible states of a physical system composed of many particles. Gibbs also worked on the application of Maxwell's equations to problems in physical optics. As a mathematician, he invented modern vector calculus (independently of the British scientist Oliver Heaviside, who carried out similar work during the same period).
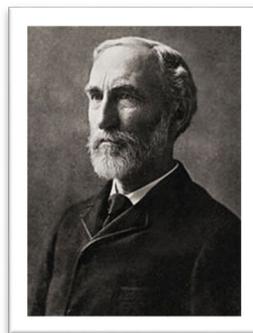
In 1863, Gibbs received the first Doctorate of Philosophy (Ph.D.) from Yale in engineering granted in the US, for a thesis entitled "On the Form of the Teeth of Wheels in Spur Gearing", in which he used geometrical techniques to investigate the optimum design for gears.

Gibbs traveled to Europe with his sisters where they spent the winter of 1866–67 in Paris. Gibbs attended lectures at the Sorbonne and the Collège de France, given by such distinguished mathematical scientists as **Joseph Liouville** and **Michel Chasles**. Moving to Berlin, Gibbs attended the lectures taught by mathematicians **Karl Weierstrass** and **Leopold Kronecker**, as well as by chemist **Heinrich Gustav Magnus**. In Heidelberg, Gibbs was exposed to the work of physicists **Gustav Kirchhoff** and **Hermann von Helmholtz**, and chemist **Robert Bunsen**. At the time, German academics were the leading authorities in the natural sciences, especially chemistry and thermodynamics.

After a three-year sojourn in Europe, Gibbs spent the rest of his career at Yale, where he was professor of mathematical physics from 1871 until his death. Working in relative isolation, he became the earliest theoretical scientist in the United States to earn an international reputation and was praised by Albert Einstein as "the greatest mind in American history".

Commentators and biographers have remarked on the contrast between Gibbs's quiet, solitary life in turn of the century New England and the great international impact of his ideas. Though his work was almost entirely theoretical, the practical value of Gibbs's contributions became evident with the development of industrial chemistry during the first half of the 20th century. According to Robert A. Millikan, in pure science Gibbs "did for statistical mechanics and for thermodynamics what

Laplace did for celestial mechanics and Maxwell did for electrodynamics, namely, made his field a well-nigh finished theoretical structure."

When Dutch physicist **J. D. van der Waals** received the 1910 Nobel Prize "for his work on the equation of state for gases and liquids" he acknowledged the great influence of Gibbs's work on that subject.

When no external electric field is applied there are no ionic fluxes, the system is in equilibrium and the electrochemical potential is constant in the whole system. This implies that:

$$\mu_i(r) = \mu_i(r \to \infty)$$

$$q_i \psi(r) + k_B T \ln\left(\frac{n_i(r)}{n_0}\right) + \mu_i^0 = k_B T \ln\left(\frac{n_i(\infty)}{n_0}\right) + \mu_i^0$$

This leads to:

$$n_i(r) = \text{n}_i(\infty) \exp\left(\frac{-q_i \psi(r)}{k_B T}\right)$$

which is the **Boltzmann distribution** given in Chapter 3.

Osmotic pressure

The derivation of the osmotic pressure from the chemical potential is due to the Dutch chemist **Jacobus Henricus van't Hoff** (1852 – 1911). Van't Hoff got the first Nobel prize in Chemistry in 1901.

His derivation is as follows:

Let us consider a suspension that is brought into contact with a semi-permeable membrane, as sketched in **Chapter 2**. The system is at equilibrium when the chemical potential of the solvent is equal on both sides of the membrane (there is then also no flux). The chemical potentials of the solvent for each side of the membrane are then equal:

$$\mu_w^0(P) = \mu_w(x_w, P + \Pi)$$

where μ_w^0 is the chemical potential of the pure solvent, P is the pressure at one side of the membrane and $P + \Pi$ is the pressure at the other side.

The variable x_w stands for the mole fraction of the solvent:

$$x_w = \frac{n_w}{n_{tot}}$$

where n_w is the number of moles of solvent and n_{tot} the total number of moles of the system. If x_s is the mole fraction of the solute, we have:

$$x_w + x_s = 1$$

For an ideal mixture, the chemical potential can be written:

$$\mu_w(x_w, P + \Pi) = \mu_w^0(P + \Pi) + RT\ln(x_w)$$

where $R = k_B N_A$ is the **gas constant**, with N_A being Avogadro's number. Moreover, the addition to the pressure can be seen as an energy of expansion:

$$\mu_w^0(P + \Pi) = \mu_w^0(P) + \int_P^{P+\Pi} V(P)dP$$

where V is the molar volume (m³/mol). Combining the above equations, we get:

$$-RT\ln(x_w) = \int_P^{P+\Pi} V(P)dP$$

If the liquid is incompressible the molar volume is constant, and the integral is then equal to ΠV. Thus, we get:

$$\Pi = \frac{-RT}{V}\ln(x_w) = \frac{-RT}{V}\ln(1 - x_s) \approx \frac{RT}{V}x_s = nk_BT$$

We have here used the fact that $x_s \ll 1$.

Kinetics of sedimentation in colloidal suspensions

We have seen in the previous sections the importance of the existence of a backflow and convection in the settling column, which is due to the hydrodynamic influence of the walls of the container. In the laboratory coordinate frame, the volume flux of colloidal material through a cross sectional surface area perpendicular to the sedimentation velocity is always compensated by fluid flowing in opposite direction. This is expressed as

$$(1 - \phi_S)v_{w/lab} + \phi_S v_{s/lab} = 0$$

Clearly, backflow tends to decrease sedimentation velocities, especially at larger volume fractions. Although the fluid backflow may be considered constant at a local scale, allowing statistical mechanical analysis for a uniform backflow, it certainly varies significantly from point to point over distances comparable to the size of the

container as we have discussed earlier. The backflow may be considered homogeneous over distances small compared to the size of the sample container and at the same time large compared to the average distance between Brownian particles. This assumption (called the **"chemical approximation"** by some authors[112]) allows the chemical potential and the drag on a particle to be seen as pre-averaged functions of the local mean volume fraction of particles in a region of size L around the particle, where

$$\text{particle radius} \ll L \ll \text{length over which } n(z) \text{ varies}$$

In volume elements of size L, the chemical potential and the drag are thus equal to the values they would have in a macroscopically uniform suspension of equivalent concentration $n_0 = n(z)$. This approximation is only valid if the rate of diffusion of the particle is greater than its rate of sedimentation. This is represented by the **Péclet** number Pe which is a measure of the ratio of the gravitational force on a particle to its Brownian motion:

$$Pe = \frac{4\pi(\rho_s - \rho_w)ga^3}{3k_BT}L$$

The chemical approximation is thus valid for slowly sedimenting dispersions.

Settled particles

Before we turn to the settling phase, let us first come back to the case where the particles have settled. In **Chapter 2**, we have found that for non-interacting particles a density profile could be established:

$$n(z) = n_0\exp\left(-\frac{4}{3}\pi a^3\left(\frac{\rho_p - \rho_w}{k_BT}\right)gz\right)$$

This relation can be written in the more general form:

$$n(z) = n_0\exp\left(\frac{\boldsymbol{F}_{ext}\cdot\boldsymbol{r}}{k_BT}\right) = n_0\exp\left(\frac{-F_{ext}z}{k_BT}\right)$$

where $\boldsymbol{F}_{ext}$ is the sum of the external forces applied to a colloidal (Brownian) particle. In **Chapter 2**, we have also already said that in the context of colloids, the pressure associated to the barometric profile is the osmotic pressure:

$$\Pi = nk_BT$$

---

[112] Buscall, Richard. "The sedimentation of concentrated colloidal suspensions." Colloids and surfaces 43.1 (1990): 33-53.

Note from the two last equations we find that:

$$F_{ext} = -\frac{d\Pi}{dz}\frac{1}{n} = -\frac{\partial \ln(n)}{\partial z}\frac{d\Pi}{dn}$$

which demonstrates that the external force is associated to the osmotic pressure gradient.

Osmotic pressure and external force

The link between the external force and osmotic pressure can be done using a thermodynamic approach. Again, we here only briefly sketch the derivation, and refer to the footnote – and complementary books / courses about thermodynamics - for more details[113].

In equilibrium the chemical potential is a constant, independent of position. The driving force for diffusion is equal to gradients in the chemical potential:

$$\boldsymbol{F}_w = -\boldsymbol{\nabla}\mu_w$$

$$\boldsymbol{F}_s = -\boldsymbol{\nabla}\mu_s$$

where the minus sign is introduced to indicate that the diffusion current is directed towards regions of lower chemical potential, so as to minimize the free energy. The two chemical potentials are not independent quantities: they are related by the **Gibbs-Duhem** relation (at constant mechanical pressure and temperature):

$$n(z)\boldsymbol{\nabla}\mu_s + n_w(z)\boldsymbol{\nabla}\mu_w = 0$$

where n_w is the local number density of solvent molecules. As $\phi_s + \phi_w = 1$ we get:

$$V_w n_w + V_s n = 1$$

where V_w and V_s are the volume of a solvent molecule and a colloidal particle.

We note that any force per unit of volume, acting on the solvent and the particles alike do not produce a relative velocity. The force per unit volume on the fluid is equal to $\boldsymbol{F}_w/V_w$. The force per unit volume of particle that generate a relative motion is therefore given by:

$$\frac{\boldsymbol{F}_{ext}}{V_s} = \frac{\boldsymbol{F}_s}{V_s} - \frac{\boldsymbol{F}_w}{V_w}$$

---

[113] See especially Dhont, J. KG. (1996). *An introduction to dynamics of colloids*. Vol. 2. Elsevier., Chapter 7

Using the equations given above:

$$\frac{\boldsymbol{F}_{ext}}{V_s} = \frac{n_w \boldsymbol{\nabla}\mu_w}{1 - V_w n_w} - \frac{-\boldsymbol{\nabla}\mu_w}{V_w} = \frac{1}{V_s n V_w} \boldsymbol{\nabla}\mu_w$$

which gives:

$$\boldsymbol{F}_{ext} = \frac{1}{n V_w} \boldsymbol{\nabla}\mu_w$$

Since the local osmotic pressure is by definition (see the definition above) equal to:

$$\Pi = -\frac{\mu_w - \mu_w^0}{V_w}$$

where μ_w^0 is the chemical potential of pure solvent, we find indeed that:

$$\boldsymbol{F}_{ext} = -\frac{1}{n} \boldsymbol{\nabla}\Pi$$

$$F_{ext} = -\frac{1}{n} \frac{d\Pi}{dz}$$

The force is therefore directly linked to the osmotic pressure. This force symbolizes the fact that there is an interaction between the colloidal particles, associated to the pair-correlation function g. The local osmotic pressure can be written (the derivation is beyond the scope of this book):

$$\Pi(\boldsymbol{r}) = n(\boldsymbol{r})k_B T - \frac{2\pi}{3} n^2(\boldsymbol{r}) \int_0^\infty R^3 \frac{d\varphi(R)}{dR} g(R) dR$$

where φ is the pair potential of interaction between two particles. The first term on the right-hand side is related to Brownian motion, the second to the interaction between particles. When $\varphi = 0$ (no interaction between particles), one finds:

$$\Pi = n k_B T$$

which corresponds to the barometric height distribution as given in **Chapter 2**.

Settling profiles

Let us now discuss the settling phase. The chemical approximation explained above implies that the thermodynamic and drag forces on a particle are only dependent on the local average concentration ϕ_S. This, in particular, is at the origin of the continuity equation:

$$\frac{\partial \phi_S}{\partial t} + \frac{\partial \left(v_{s/lab}\phi_S\right)}{\partial z} = 0$$

that will be used in **Chapter 9**. Note that the continuity equation for *one* particle reads:

$$\frac{\partial \Psi_m}{\partial t} + \sum_{n=1}^{N} \frac{\partial \left(u_{n/lab}\Psi_n\right)}{\partial x_n} = 0$$

where Ψ_m is the probability of finding a particle at a point x_m (and is linked to a pair-correlation function), $u_{n/lab}$ is the velocity of the nth particle.

The forces acting upon a settling particle in a concentrated suspension are: the gravitational force (compensated for Archimedes) and a force due to the interaction with other particles. This last one is expressed through the osmotic pressure. The viscous drag force exerted on the colloidal sphere is balanced by the sum of these forces:

$$\frac{6\pi\eta a}{\chi(\phi_S)}(v_s - v_w) = -(\rho_s - \rho_w)gV_p - \frac{V_p}{\phi_S}\frac{\partial \Pi}{\partial z}$$

The ratio $V_p/\phi_S = 1/n$ represents the mean volume element per particle (n is the mean number of particles per unit of volume). The factor $\chi(\phi_S)$ accounts for the hydrodynamic interaction effects on the particle mobility[114] and $V_p = 4\pi a^3/3$ is the volume of one particle. We have:

$$\chi(\phi_S) = 1 \ \ \text{for} \ \ \phi_S \to 0$$

$$\chi(\phi_S) = 0 \ \ \text{for} \ \ \phi_S \to 1$$

The last equation implies that for $\phi_S \to 1$ one has $v_s = v_w(= 0)$. Using $v_{s/lab} = -(1 - \phi_S)v_{w/s} = (1 - \phi_S)(v_s - v_w)$, the expression can be rewritten:

$$\frac{6\pi\eta a}{f(\phi_S)}v_{s/lab} = -(\rho_s - \rho_w)gV_p - \frac{V_p}{\phi_S}\frac{\partial \Pi}{\partial z}$$

where

$$f(\phi_S) = (1 - \phi_S)\chi(\phi_S)$$

One relation for $f(\phi_S)$ is for example the **Richardson-Zaki** expression given above.

---

[114] This implies that η is the viscosity of the medium.

The equation we have just proposed is not obtained from following a single particle in time. This would be too complex to do, as it would require to know, at each time, the interactions of all the surroundings particles with the particle we track, and moreover we would need the particle's initial position and velocity which is impossible to determine. In fact, even though the equation has been written as if it is for one particle, it is obtained from the pre-averaged functions of the local mean volume fraction of particles in a region of size L around the particle (see discussion about the chemical approximation above).

Note

At equilibrium, $v_{s/lab} = 0$ and we find:

$$\frac{\partial \Pi}{\partial z} = (\rho_s - \rho_w)g\phi_S$$

from which it can be inferred that it is the gradient in osmotic pressure (i.e. the interactions between particles) that keeps the colloidal particles suspended. For large particles (non-colloidal), the osmotic pressure term is negligible in comparison with the gravity force : this was discussed in **Chapter 2** in terms of volume and surface forces, see discussion for Case (C) and Case (D). We find in that case:

$$\frac{6\pi\eta a}{f(\phi_S)}v_{s/lab} = -(\rho_s - \rho_w)gV_p$$

This implies that the only way to get equilibrium ($v_{s/lab} = 0$) is to have $\rho_s = \rho_w$.

The particle flux along the vertical direction is given by:

$$J_{s/lab} = \phi_S v_{s/lab} = -\phi_S(1 - \phi_S)v_{w/s} = -\phi_S J_{w/s}$$

$$J_{s/lab} = -\phi_S \frac{f(\phi_S)}{6\pi\eta a}\left[(\rho_s - \rho_w)gV_p + \frac{V_p}{\phi_S}\frac{\partial \Pi}{\partial z}\right]$$

The osmotic pressure can be written:

$$\Pi = \frac{k_B T}{V_p}\phi_S Z(\phi_S) = Z(\phi_S)nk_B T$$

where $Z(\phi_S)$ is the compressibility factor that accounts for interparticle interactions. A widely used expression for a homogeneous hard-sphere suspension is given by the **Carnahan-Starling** equation of state[115]:

---

[115] Dhont, J. KG. (1996). *An introduction to dynamics of colloids*. Vol. 2. Elsevier., Chapter 7

$$Z(\phi_S) = \frac{1 + \phi_S + \phi_s^2 - \phi_s^3}{(1 - \phi_S)^3}$$

Incidentally, one can note that in Chapter 7 we introduced the **Carnahan-Starling** expression without referring explicitly to the osmotic pressure. We define the diffusion coefficient and velocity:

$$D_0 = \frac{k_B T}{6\pi\eta a} \quad ; \quad v_0 = \frac{(\rho_s - \rho_w)gV_p}{6\pi\eta a}$$

$[v_{s/lab} = v_0$ at infinite dilution], and we get:

$$J_{s/lab} = -\phi_S v_0 f(\phi_S) - f(\phi_S) D_0 \frac{\partial[\phi_S Z(\phi_S)]}{\partial z}$$

A characteristic length can be obtained from D_0 and v_0, namely:

$$l_g = D_0/v_0 = \frac{k_B T}{(\rho_s - \rho_w)gV_p}$$

which may be seen as the length a particle has to settle before the sedimentation drift becomes equal to the root mean square displacement due to Brownian motion.

The continuity equation imposes:

$$\frac{\partial\phi_S}{\partial t} + \frac{\partial J_{s/lab}}{\partial z} = 0$$

leading to the general equation that has to be solved in order to find the full hindered settling profile $\phi_S(z, t)$:

$$\frac{\partial\phi_S}{\partial t} - v_0 \frac{\partial}{\partial z}[\phi_S f(\phi_S)] = D_0 \frac{\partial}{\partial z}\left[f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial z}\right]$$

In general, the full solution of this equation can only be obtained numerically[116]. The equation has a Burger-like structure. **Burgers' equation** is from the family of conservation equations that can develop discontinuities and used to characterize **shock waves**. The "shock" in our context is the propagation of the suspension/bed interface through the fluid.

---

[116] Auzerais, F. M., Jackson, R., & Russel, W. B. (1988). The resolution of shocks and the effects of compressible sediments in transient settling. J. Fluid Mech, 195(1), 437-462.

Johannes Martinus Burgers (1895-1981) was a Dutch physicist, who studied in Leiden under **Paul Ehrenfest**. Three months before his PhD graduation he already got a position as professor at the *Technische Hogeschool Delft* (nowadays the Technical University Delft), where he worked in the laboratory for aero- and hydrodynamics. In particular, he designed the ventilation system for the Maastunnel in Rotterdam. In 1955 he emigrated to the United States for the facilities he was offered there. The Dutch Burgerscentrum[117] is named after him.

We will here only discuss some particular cases of the equation for which general behaviours can be defined. A (numerical) full solution is discussed at the end of the chapter.

Dilute regime

In the dilute limit $\phi_S \ll 1$ and $f(\phi_S) = Z(\phi_S) = 1$. At equilibrium, $\partial \phi_S / \partial t = 0$ and the equation reduces to:

$$-v_0 \frac{\partial \phi_S}{\partial z} = D_0 \frac{\partial^2 \phi_S}{\partial z^2}$$

This equation can be solved and yields:

$$\phi_S(z) = \phi_0 \, exp\big(-z/l_g\big)$$

where the integration constant ϕ_0 is obtained by knowing the initial conditions. In **Chapter 2**, we have found the same result when we analysed the settling of non-interacting particles, but expressed it as

$$n(z) = n_0 exp \left(-\frac{4}{3} \pi a^3 \left(\frac{\rho_p - \rho_w}{k_B T} \right) gz \right)$$

One can verify that $l_g = D_0/v_0 \ll 1$ for clay particles larger than 100 nm. In that case $\phi_S(z \neq 0) = 0$ and all the particles have settled at the layer $z = 0$.

Initial settling rate for an initial homogeneously mixed suspension

The initial settling rate of the water/suspension interface is now discussed for the special case of a suspension consisting of spherical colloidal (Brownian) particles that

---

[117] http://www.jmburgerscentrum.nl/

experience hindered settling. Let us assume that the suspension is well mixed at t = 0:

$$\phi_S(z, t = 0) = \phi_0$$

This implies that there are no gradients in concentration, and consequently:

$$\left(\frac{\partial \phi_S}{\partial z}\right)_{z,t=0} = 0$$

As the osmotic pressure depends on ϕ_S and:

$$\frac{\partial \Pi}{\partial z} = \frac{d\Pi}{d\phi_S} \frac{\partial \phi_S}{\partial z}$$

we find that:

$$\left(\frac{\partial \Pi}{\partial z}\right)_{z,t=0} = 0$$

We recall that:

$$J_{s/lab} = \phi_S v_s = -\phi_S \frac{f(\phi_S)}{6\pi\eta a}\left[(\rho_s - \rho_w)gV_p + \frac{V_p}{\phi_S}\frac{\partial \Pi}{\partial z}\right]$$

From which we deduce that:

$$(\boldsymbol{v_s})_{z,t=0} = \frac{f(\phi_0)}{6\pi\eta a}(\rho_s - \rho_w)V_p\boldsymbol{g}$$

This is the familiar Stokes's law for sedimentation rate of an isolated sphere in a medium of viscosity η modified through $f(\phi_0)$ to account for the increase in fluid drag due to the hydrodynamic interactions. Note that as soon as $t > 0$ we will have $\partial\Pi/\partial z \neq 0$ at the bottom of the settling column (as particles cannot penetrate the bottom). In time, the condition $\partial\Pi/\partial z \neq 0$ will propagate to the top of the column.

Boycott effect

We have (and will) assume that the suspensions we consider are settling vertically. In 1920 Boycott discovered that the sedimentation of blood components under gravity force occurred faster in inclined tubes than in vertical tubes. An analytical

model was later developped, called the PNK theory, after Ponder (1925), Nakamura and Kuroda (1937) who independently proposed it[118].
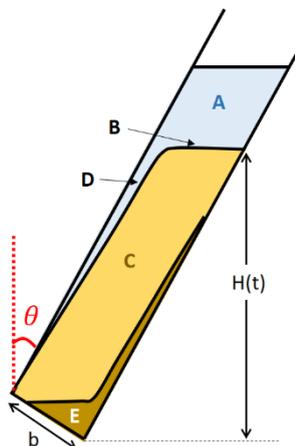
Figure opposite: (A) region of particle-free fluid above the suspension, (B) interface between the particle-free fluid and the suspension, (C) suspension, (D) thin particle-free fluid layer ·beneath the downward-facing surface, (E) concentrated sediment.

When the sediment particles fall onto the side wall of the column, they form a thin sediment layer that slides rapidly towards the bottom of the column, hereby increasing the settling speed. Because of conservation of volume, clear fluid will appear at the opposite wall and propagate to the top of the column. The PNK model states that:

$$\frac{dH}{dt} = -v_s \left(1 + \frac{H}{b}\sin(\theta)\right)$$

When a settling tube is inclined rather than vertical, the velocity at which colloidal particles sedimentate is enhanced by several orders of magnitude: this phenomenon is called the Boycott effect, named after **Arthur Edwin Boycott** (1877-1938), an eminent pathologist and naturalist. In childhood, Boycott displayed a taste for natural history and was particularly interested in snails. When he was 15 years of age, he made a list of those of Herefordshire which was published in *Science Gossip* in 1892. His interest in the topic never faded and the last scientific papers he published were two memoirs on the habitats of the land and freshwater mollusca in Britain. In 1912 he was appointed professor at the University of Manchester. His colleagues expected the professor to display a practical interest in the application of pathology to the everyday problems of clinical work. Boycott would not accept this view. He mainained that a professor must devote all his energies to the advancement of science, and that the application of laboratory methods to clinical medicine was not

---
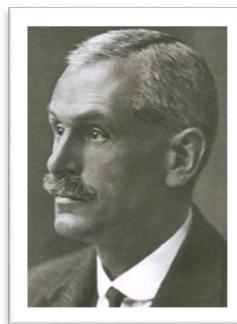
[118] Davis, R. H., & Acrivos, A. (1985). Sedimentation of noncolloidal particles at low Reynolds numbers. Annual Review of Fluid Mechanics, 17(1), 91-118 and Xu, Z. J., & Michaelides, E. E. (2005). A numerical simulation of the Boycott effect. Chem. Eng. Comm., 192(4), 532-549.

part of his duty. His uncompromised attitude was resented by his medical colleagues and rendered his tenure of the chair less happy than it should have been. As he grew older, Boycott arrived at the conclusion that it would be better for the progress of science if its workers spent less time piling up fresh facts and more in the contemplation of the implication of those already discovered.

When gravity forces are larger than thermodynamic forces

For hindered settling, assuming smooth initial conditions[119] i.e. $\phi_S(z, 0) = \phi_0$ and $\phi_S(h, t) = 0$ (where the characteristic length h represents the height of the column) and negligible thermodynamic forces, one can assume that the front's speed is constant. When the thermodynamic force is negligible compared to the gravitational force, one has $Pe = v_0 h / D_0 = h / l_g \gg 1$. Noting that the hindered settling relation can be written:

$$\frac{\partial \phi_S}{\partial (tv_0/h)} - \frac{\partial}{\partial(z/h)}[\phi_S f(\phi_S)] = \frac{D_0}{v_0 h}\frac{\partial}{\partial(z/h)}\left[f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial(z/h)}\right]$$

we introduce the dimensionless variables :

$$z^* = z/h \quad ; \quad t^* = t \cdot v_0/h$$

and therefore the equation becomes:

$$\frac{\partial \phi_S}{\partial t^*} - \frac{\partial}{\partial z^*}[\phi_S f(\phi_S)] = \frac{1}{Pe}\frac{\partial}{\partial z^*}\left[f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial z^*}\right]$$

For large *Pe* the hindered settling equation reduces to:

$$\frac{\partial \phi_S}{\partial t} - v_0 \frac{\partial}{\partial z}[\phi_S f(\phi_S)] = 0$$

$$\frac{\partial \phi_S}{\partial t} - v_0 \frac{d}{d\phi_S}[\phi_S f(\phi_S)]\frac{\partial \phi_S}{\partial z} = 0$$

This equation can be solved with the method of characteristics (see **Chapter 10** for details about this method) and leads to:

$$\phi_S(z, t) = \phi_S(z + vt)$$

---

[119] Buzzaccaro, S., Tripodi, A., Rusconi, R., Vigolo, D., & Piazza, R. (2008). Kinetics of sedimentation in colloidal suspensions. Journal of Physics: Condensed Matter, 20(49), 494219. See also Russel W B, Saville D A and Schowalter W R 1992 Colloidal Dispersions (Cambridge: Cambridge University Press), chapter 12.

with

$$v = v_0 \frac{d}{d\phi_S}[\phi_S f(\phi_S)]$$

Consequently, in the hindered settling region, ϕ_S is constant along the curves with slopes $dz/dt = v$.

Substituting $\phi_S(z,t) = \phi_S(z + vt)$ in the full hindered settling equation and integrating from z to infinity yields:

$$\int_z^\infty \frac{\partial \phi_S(z,t)}{\partial t} dz - v_0 \int_z^\infty \frac{\partial}{\partial z}[\phi_S f(\phi_S)] dz = D_0 \int_z^\infty \frac{\partial}{\partial z}\left[f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial z}\right] dz$$

$$\int_z^\infty \frac{\partial \phi_S(z + vt)}{\partial t} dz - v_0 \phi_S f(\phi_S) = D_0 f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial z}$$

We find that:

$$\int_z^\infty \frac{\partial \phi_S(z + vt)}{\partial t} dz = \int_{(z+vt)}^\infty \frac{\partial \phi_S(z + vt)}{\partial(z + vt)}\frac{\partial(z + vt)}{\partial t} d(z + vt) = v\phi_S(z + vt)$$

Therefore:

$$[v - v_0 f(\phi_S)]\phi_S = D_0 f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial z}$$

At the water/suspension interface, $\phi_S = \phi_0$ and the right-hand side of the equation is equal to zero. This implies that we then have:

$$v = v_0 f(\phi_0)$$

This steady-state, time-independent velocity is a direct consequence of the dependence of the settling velocity on particle concentration. As v is a monotonically decreasing function of ϕ_S the particles "left behind" by diffusion will sediment faster (with a velocity v_0) and catch up with the front region of the settling profile which settles with a velocity $v < v_0$. This front therefore self-sharpens. Eventually, this leads to a time-invariant shape, dictated by the balance between settling and diffusion and settling at the uniform speed v. In practice, depending on the size of the column, the time-invariant state might or not be reached as the gelling point might or not be reached before the steady-state is established. The initial concentration is also a variable: for a very dilute system for instance, the stationary profile is uniform in the whole column since then all particles settle with the same velocity v_0.

A similar reasoning enables to find the settling velocity of the suspension/bed interface: this time the integral should be done from 0 to z. We then also assume that the settled colloidal particles that form the bed ensure that the bed has a constant volume fraction ϕ_m. We get:

$$\int_0^z \frac{\partial \phi_S(z+vt)}{\partial t} dz - v_0[\phi_S f(\phi_S) - \phi_m f(\phi_m)] = D_0 f(\phi_S) \frac{\partial[\phi_S Z(\phi_S)]}{\partial z}$$

with

$$\int_0^z \frac{\partial \phi_S(z+vt)}{\partial t} dz = \int_{vt}^{(z+vt)} \frac{\partial \phi_S(z+vt)}{\partial(z+vt)} \frac{\partial(z+vt)}{\partial t} d(z+vt)$$
$$= v[\phi_S(z+vt) - \phi_m]$$

Noting that $f(\phi_m) = 0$ as the colloidal particles in the bed do not settle (they are hard spheres in contact) and that $\phi_S = \phi_0$ at the suspension/bed interface as in the hindered settling phase all the particles move with the same average speed, and therefore the volume fraction is equal to the initial volume fraction, we get the velocity of the suspension/bed interface:

$$v_{bed} = v_0 \frac{\phi_0 f(\phi_0)}{\phi_0 - \phi_m}$$

We here expressed the absolute values of the velocities. As we have taken z = 0 at the bottom of the column, v_{bed} is oriented along the z-axis unit vector, whereas v is antiparallel to it.

Finding the hindered settling function $f(\phi_S)$ from the equilibrium profile

The equation we found for the water/suspension interface velocity

$$[v - v_0 f(\phi_S)]\phi_S = D_0 f(\phi_S) \frac{\partial[\phi_S Z(\phi_S)]}{\partial z}$$

can be rearranged into:

$$\frac{f(\phi_S)}{f(\phi_0)} = \left[1 + l_g \frac{\partial[\phi_S Z(\phi_S)]}{\partial z} \frac{1}{\phi_S}\right]^{-1}$$

where we used the fact that $v = v_0 f(\phi_0)$ for the interface. Particularly interesting from an experimental point of view is that this equation enables to extract extensive *dynamic* information, i.e. the settling behaviour function $f(\phi_S)$ from a single *static* measurement of the equilibrium settling profile for a concentrated suspension. The alternative would be to measure the sedimentation velocity as function of ϕ_S which are measurements that require careful temperature control to avoid fluctuations in

the solvent viscosity. The right-hand-side of the equation contains variables that are directly measureable from the settling profile and the equilibrium measurements as we are going to discuss now.

First, we show the measurement results on the settling of a suspension (0.23 volume fraction) of spherical monodisperse (non-aggregated) colloidal particles of size 77 nm made of MFA (a copolymer). Their density is high, which enables a better accuracy in particle volume fraction determination by density measurements. Electrostatic interactions were minimized by the addition of salt. Strong electrostatic interactions would minimize the settling, as discussed in Chapter 2, case (B).
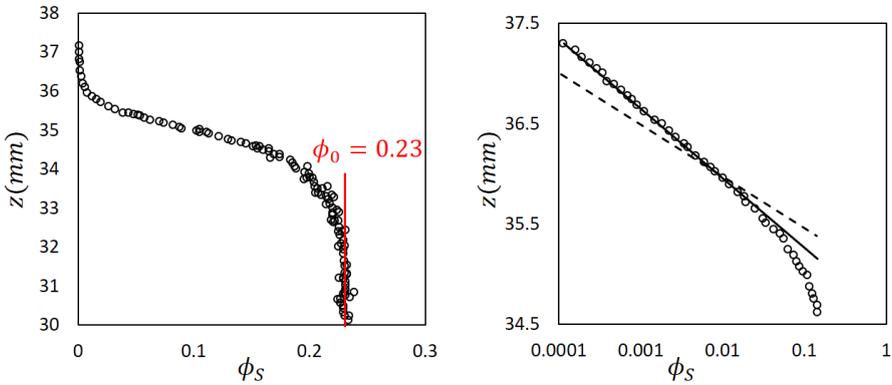
Left: stationary settling profile close to the water/suspension interface for a suspension of colloidal hard spheres with $\phi_0 = 0.23$. The measurement is done after the time-invariant shape is reached. (Note that the bulk volume fraction is still 0.23 close to the interface). Right: semilog plot of the low volume fraction region. The profile is barometric close to $\phi_S = 0$. The slope (full line) of the data points give l_g^{dyn} (defined underneath) which is sensibly larger than l_g (dashed line)
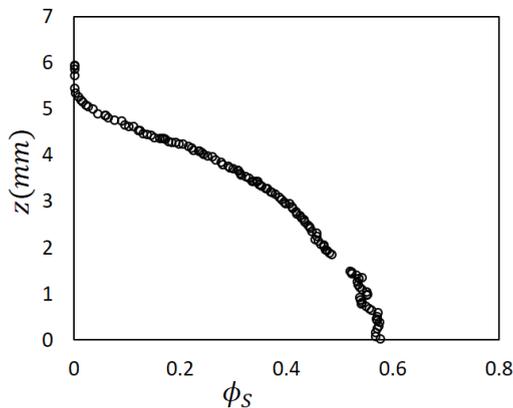
Equilibrium sedimentation profile for the same suspension of colloidal hard spheres. The barometric profile just at the sediment/water interface (between 5 and 6 mm) is similar to the previous figure.

Dynamic gravitational length

In the region around the water/suspension interface, see the stationary settling profile, one has $\phi_S \ll \phi_0$ and $Z(\phi_S) = f(\phi_S) = 1$. The hindered settling relation becomes:

$$\frac{\partial \phi_S}{\partial z} = \frac{v_0}{D_0}[f(\phi_0) - 1]\phi_S = \frac{-\phi_S}{l_g^{dyn}}$$

where we have defined a dynamic gravitational length:

$$l_g^{dyn} = \frac{l_g}{1 - f(\phi_0)}$$

Therefore the top part of the settling profile still has an exponential (barometric) shape but with a length $l_g^{dyn} > l_g$ since $f(\phi_0) < 1$. The smaller ϕ_0 the closer $f(\phi_0)$ comes to 1. In the limit of very dilute suspensions, l_g^{dyn} therefore goes to infinity, implying that all particles settle according to Stokes, as expected.

The osmotic pressure can be obtained from the equilibrium sedimentation profile. We have shown earlier that at equilibrium:

$$\frac{\partial \Pi}{\partial z} = (\rho_s - \rho_w)g\phi_S$$

Integrating this equation gives:

$$\Pi(z) = (\rho_s - \rho_w)g \int_z^\infty \phi_S(z)dz$$

From this integration, the compressibility factor $Z(\phi_S(z))$ can be evaluated using:

$$\Pi(z) = n(z)k_B T \cdot Z(\phi_S(z))$$

We recall that $n(z) = \phi_S(z)/V_p$. As we have seen, $f(\phi_S)$ is a function of $f(\phi_0)$, l_g, $\phi_S(z)$ and $Z(\phi_S)$. For each $\phi_S(z)$ one can estimate these required variables and therefore obtain $f(\phi_S)$ for $\phi_S \leq \phi_0$.
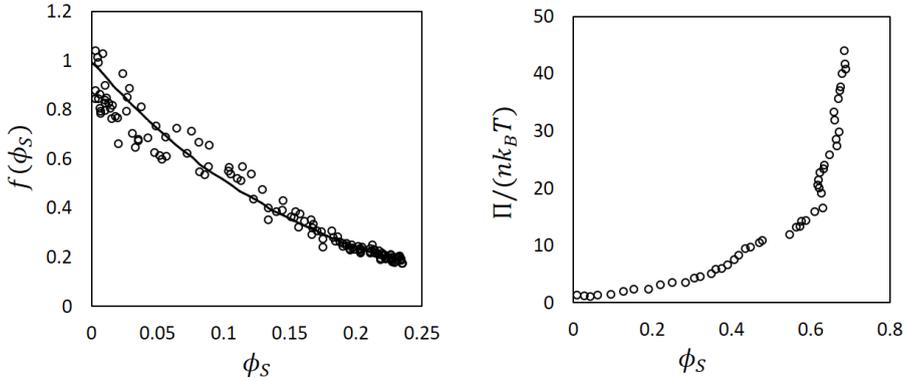
Left: hydrodynamic factor obtained from the stationary settling profile given above. Right: the osmotic pressure is obtained by integrating the equilibrium sedimentation profile.

Numerical solution for sedimentation of an initially uniform suspension

We here give a simple example of the numerical solution of

$$\frac{\partial \phi_S}{\partial t} - v_0 \frac{\partial}{\partial z}[\phi_S f(\phi_S)] = D_0 \frac{\partial}{\partial z}\left[f(\phi_S)\frac{\partial[\phi_S Z(\phi_S)]}{\partial z}\right]$$

where we use the Carnahan-Starling equation of state:

$$Z(\phi_S) = \frac{1 + \phi_S + \phi_S^2 - \phi_S^3}{(1 - \phi_S)^3}$$

and the Richardson and Zaki relation:

$$f(\phi_S) = (1 - \phi_S)^6$$

with the boundary conditions $\phi_s = \phi_0$ at t = 0. Note that the exponent is now taken to be 6 for a change (earlier it was shown that 5 was a good value for the data presented). The exponent depends on the type of particles in suspension. There is no flux at the bottom and top of the column of height h. The following dimensionless variables are used:

$$\tau = t\frac{D_0}{h^2}$$

$$Z = z/h$$

From numerical integration one obtains:

$(1)\, \tau = 10^{-5}$
$(2)\, \tau = 10^{-3}$
$(3)\, \tau = 2 \cdot 10^{-3}$
$(4)\, \tau = 3 \cdot 10^{-3}$
$(5)\, \tau = 4.5 \cdot 10^{-3}$
$(6)\, \tau = 10^{-2}$
$(7)\, \tau = 4 \cdot 10^{-2}$
$(8)\, \tau = 4 \cdot 10^{-1}$
$(9)\, \tau = 4 \cdot 10^{1}$
$(10)\, \tau = 2 \cdot 10^{4}$
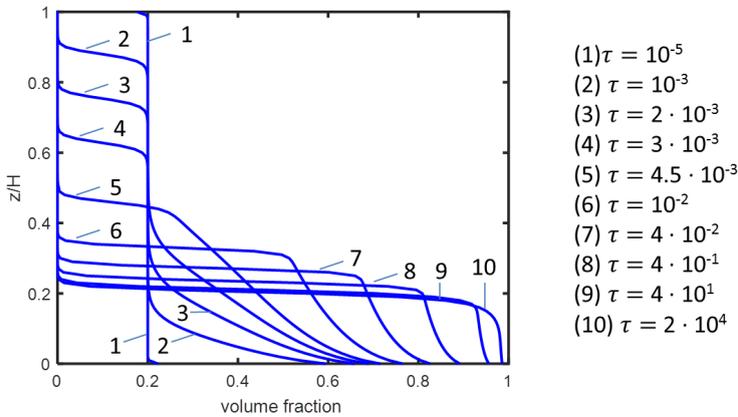
The volume fraction versus the distance from the bottom of the column (Z = 0)

As can be seen from that figure, after a short time a clear fluid layer develops at the top of the container, while particles accumulate at the bottom. An interface at the top of the container is then seen to move downwards with a constant velocity until it meets the sediment that is formed at the bottom. From then on the sediment compacts relatively slowly until the sedimentation-diffusion equilibrium is reached.

Illustrations

Josiah Willard Gibbs
https://en.wikipedia.org/wiki/Josiah_Willard_Gibbs

Jacobus H. van 't Hoff
https://www.nobelprize.org/prizes/chemistry/1901/hoff/biographical/
https://en.wikipedia.org/wiki/Jacobus_Henricus_van_%27t_Hoff

Burgers
http://www.burgers.umd.edu/burgers.html

Arthur Boycott
*https://www.npg.org.uk/collections/search/portrait/mw97885/Arthur-Edwin-Boycott*

Chapter 9
Permeability of slurries

One important parameter in soil science is **permeability**, which quantifies the ability of a soil to have water (or any fluid) permeate (diffuse through) it. This property is not only important for agriculture but also for construction: the bottom of ponds for example should be impermeable to water and so are dikes which should be built with a soil that will ensure a good water retention. Soils are also generally made up of different sediment layers and soil quality varies greatly from one layer to another. An important variable for permeability is grain size: the coarser the grain, the more permeable a sediment layer is. This led to the following classification:

	1	10^{-3}	10^{-5}	10^{-8}	10^{-12}
Permeability (m/s)	very good	average	semi-permeable	impermeable	
Type of soil	gravel	sand	silt	clay	

Some estimations[120] of the permeability K (m/s) and related properties

The measurement of permeability in soils is studied in geoengineering, and then concerns soils where particles have formed a fabric, also called a skeleton, implying that the particles are touching each other in a kind of network. We will here discuss a part of consolidation that lays at the frontier between hydraulic and geo-engineering: the very early part of consolidation, which happens when a bed is formed by the settling of (flocculated) particles. We will show that from the time evolution of the interfaces (blue for the suspension/bed interface and red for the water/suspension interface in the figure below) important information can be obtained about the structure of the fabric and the permeability in particular.
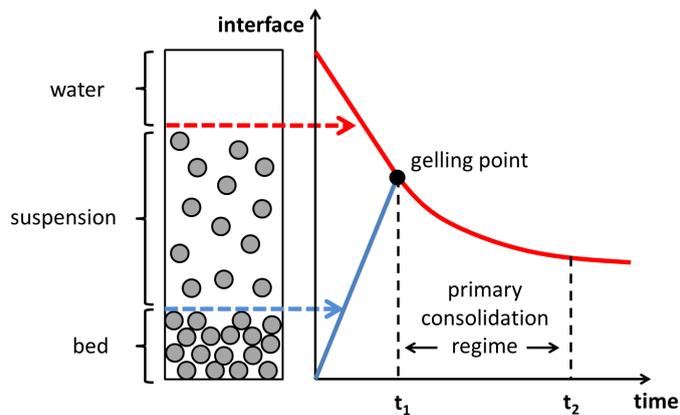
[120] Sobolewski, M. (2005). Various methods of the measurement of the permeability coefficient in soils-possibilities and application. Electronic Journal of Polish Agricultural Universities. Series Civil Engineering, 8(2).

For illustration, we will, in this chapter, study the behaviour of a mud suspension in an **undrained** settling column. *Undrained* means that there is conservation of the total mass (water + sediment), as the fluid cannot escape the column. This condition can be relaxed for describing real systems by adapting the boundary conditions.

To describe this initial stage of consolidation of soils, we will distinguish three different regimes:

$t < t_1$ **: the settling regime**. The particles are non-touching, and due to gravity, move to the bottom of the settling column. Both a suspension phase and a bed are observed.

$t_1 < t < t_2$ **: the primary consolidation regime.** The suspension phase has disappeared. The particles are in very close vicinity of each other and as they adjust their position with respect to each other, a significant amount of water is expelled from the fabric.

$t > t_2$: **the secondary compression regime.** The particles are trapped in their positions, and significant internal stresses develop at the points where the particles are touching. Over time, some links between particles can fail or change, leading to a new reordering of particles and a slow reduction in bed height.


The terms primary consolidation and secondary compression come from soil mechanics, where traditionally large loads are applied to soils, and their compression index is studied by measuring the deformation resulting from applying a load. Here the "load" applied to our settling mud is gravity. The dynamics of the system are solely governed by gravity and the (complex) interactions between the moving water and the particles. A full discussion about the settling regime is given in **Chapter 8**.

At the bed, the stresses developed between particles are extremely weak compared to those found in soil mechanics. The "load" felt by a particle in the bed is due to the weight of other particles and the water above it. To give an order of magnitude, a 1 m high mud suspension with a density of 1200 kg m^{-3} generates a pressure of approximately 2 kPa which is 4 or 5 orders of magnitude lower than the pressures applied in a standard **oedometer test** (this test is briefly discussed at the end of **Chapter 10**). Over time, the load felt by the particles in the bed increases as the density of the bed increases, but it is evident that the pressures will never come in range with those applied in soil mechanic tests.

Within the forming bed, the distribution of stresses will be depth-dependent, the particles closer to the bottom of the column experiencing higher stresses than particles higher up in the column. One can also argue that as soon as particles are touching and forming a fabric the problem becomes a 3 dimensional one as the stresses will be distributed in all directions. This in particular raises the question of dependence of the final bed height on the column diameter.

Observation of a settling suspension

The careful observation of a settling suspension shows that many factors should be accounted for in order to get a complete description of the system. A highly cited article[121] of Coe and Clevenger of 1916 gives an idea of the complexity of the consolidation process. We here give the description given by the authors in their article, in which they observe the settling of a pulp suspension.
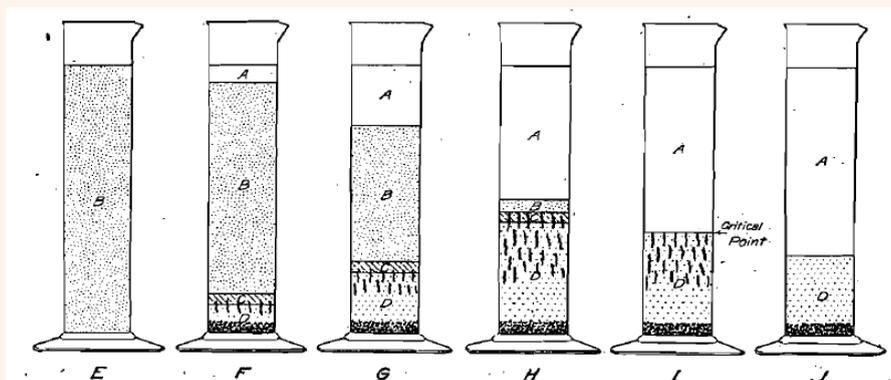

Settling experiment of Coe and Clevenger, done with a settling pulp

*"The first particles which reach the bottom of the cylinder are the coarser granular sand which may be present in the pulp. Immediately following this and somewhat contemporaneously with the settling of the sand, the slime flocs nearest the bottom settle, filling the interstitial spaces between the sand particles, and build up, one upon another, in a zone of increasing depth. This we term **zone D**, which may be defined as that portion of pulp wherein the flocs, considered as integral bodies, have settled to a point where they rest directly one upon another. After pulp enters **zone D**, further separation of liquid must come through liquid pressed out of the flocs and out of the interstitial spaces between the flocs. Immediately above **zone D** is a transition **zone C**. The pulp in **zone C** decreases in percentage of solids from the bottom, where the flocs enter **zone D,** to the top, where the consistency of the flocculated pulp is the same as the original pulp. In speaking of flocculated pulp, it is intended to eliminate from consideration the coarser portion of the contained sand which falls directly through the overlying zones into **zone D**. Above **zone C** is **zone B**, of constant consistency of flocculated pulp and of the same consistency as the flocculated pulp in the feed pulp. **Zone A**, overlying **zone B**, is clear water or solution. In the case of a very rapidly settling time, particularly with material which has been roasted, **zone A** in the earlier stages may be turbid, due to the finely divided matter remaining in suspension. Later, this very fine material settles and the liquid becomes*

---

[121] Coe, H. S. "Methods for determining the capacities of slime settling tanks." Transactions American Institute of Mining Engineerring 55 (1916).

clear, although there are cases, especially when the liquid contains very little electrolyte, where it remains turbid for a long time. […]
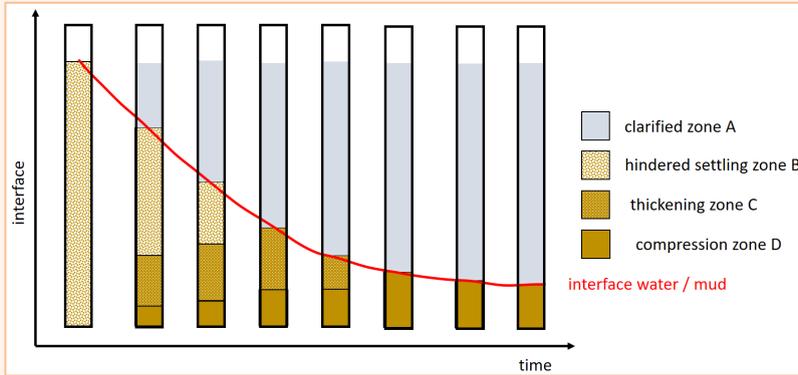
Schematic representation of a column filled with a suspension with high concentration (high volume fraction).

*We designate as free settling pulp all of the pulp in **zones B** and **C**, wherein the sand and flocs are falling freely through the liquid without pressing on the layers of flocs beneath, although it is evident, from the peculiar interlocking structure of flocculated pulp, that there are points of contact between the flocs even in these zones. We designate as **Critical Settling Point** the top of **zone D** just as **zone C** disappears. At this point the flocs at the surface just rest upon each other, but compression has not yet commenced in the surface layer. It is therefore obvious that any elimination of liquid from **zone D** cannot be accomplished by free settling but must be effected by compression of flocs. The water liberated by compression finds its way out of **zone D** through tubes or channels which form drainage systems upward through the zone. The trunk channel for any system has its outlet at the top of **zone D**. Since **zone B** is made of flocculated pulp of constant consistency, the flocs in this zone will settle at a constant rate so long as the zone exists. If **zone C** remains shallow and of constant depth, the liquid being expelled from **zone D** may be ejected through **zone C** with little admixture with this zone. This upward current of liquid which has diffused through **zone B** during the first stages of settling may also be ejected through **zone B** when B becomes very shallow, and thereby considerably increase the apparent settling rate of the pulp at this stage. This does not indicate that during this period more pulp passes into **zone D**. It is merely a surface phenomenon, indicating that **zone C** is very shallow. A curve plotted from the results of a settling test such as is illustrated in the figure will show a constant settling rate throughout the stages represented in **E, F and G**, an increased rate when the stage shown in **H** is reached, followed by a rapid retardation in the settling rate after **zone B** disappears, until **zone C** has also disappeared. Following this is a very slow rate of settling which gradually becomes slower as the water is compressed from **zone D**. The final stage of settling is reached when no further liquid is expelled from **zone D** by compression."*

We are now going to set-up the set of equations required to find the evolution of the water/suspension and the suspension/bed interfaces as function of time. We will consider the case where there is no sand inside the mud suspension and therefore zone D contains only flocs that are so compressed that water can barely escape from the fabric. We make no distinction between zone C and D.

Link between settling and consolidation

First, we have to understand the role of pressure and stresses in the observed behaviour. We will start by deriving a Darcy-like equation for settling particles. In the previous chapter, we have already found a set of equations to be solved for the settling of colloidal particles. In this chapter, we will make the link between the formulation found in **Chapter 8** and the formulation for settling and consolidation we are going to derive.

First, we will recall the traditional **Darcy equation** for porous media, and then we will show how this equation can be adapted for settling particles.

The Darcy equation

$$\nabla P$$

$$J_{w/lab} = (1 - \phi_s) v_{w/lab}$$

$$v_{w/lab} = v_{w/s}$$
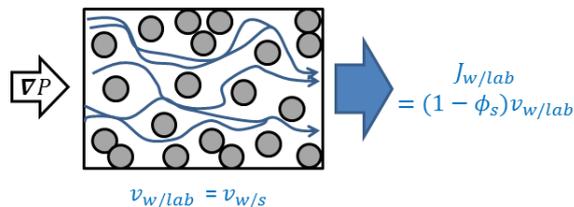
Water, under a pressure gradient ∇P, is forced into a porous medium (for instance a sandstone, which is used often as a reference stone in petroleum industry). The flux of water $J_{w/lab}$ (m/s) coming out of the stone is measured.

The Darcy equation links the flux of water that passes through a porous stone to the pressure gradient ∇P that is applied to push the water. It is given by:

$$J_{w/lab} = \frac{-k}{\eta} \nabla P$$

where k (m^2) is the permeability and η the viscosity. We have given here the general Darcy equation in terms of vectors (bold notation). The scalar $J_{w/lab}$ expresses the fact that the flux of the water (subscript w) is defined in the frame of reference of the laboratory (*lab*).

The gradient in pressure is a vector that is opposite to the vector of the water flow, as the flow of water goes from regions of high pressure to regions of low pressure.

To ensure that we have k > 0 we have a minus sign in front of k. In the following we will consider only unidirectional flows. For that reason, we will drop the vector notation for convenience.

The flux $J_{w/lab}$ represents a macroscopic flux and is defined as the volume of water exiting the porous media per unit of surface. Its units are therefore (m³/s)/m²= m/s. It can be linked to the microscopic flux of water $v_{w/lab}$ that flows inside the stone pores by realising that:

$$J_{w/lab} = \frac{dV_w}{dV} v_{w/lab} = \phi_w v_{w/lab} = (1 - \phi_S) v_{w/lab}$$

where dV_w is the small element of volume of water at a position z (we use z as we are going to discuss the case where the flow is vertical in the next section). The total small element of volume at position z is given by: $dV = dV_w + dV_s$ where dV_s is the small element of volume of soil (or solids) at z.

Definitions : volume fraction, porosity and void ratio

In this chapter we have chosen to use the volume fraction of soil as the relevant variable of the system. The reason is that this is the common variable in colloid science. In other fields of research, and engineering in particular, different variables are preferred which we give here. All the equations given in this chapter can therefore be re-written as function of any of these variables.

We define V_s as the total volume of the soil particles, V_p as the total volume of the pores (the voids filled with water) and V as the total volume of the matrix (soil + pores): $V = V_s + V_p$

The **volume fraction** ϕ_S of the soil is defined by:

$$\phi_S = \frac{V_s}{V}$$

The **porosity** ϕ (also called void fraction) is defined by:

$$\phi = \frac{V_p}{V}$$

The **void ratio** *e* is defined by:

$$e = \frac{V_p}{V_s}$$

The relation between these variables is given by:

$$\phi_S + \phi = 1 \quad ; \quad \phi = e/(1+e) \quad ; \quad \phi_S = 1/(1+e) \quad ; \quad e = \phi/(1-\phi)$$

Theoretically the porosity and volume fractions can vary between 0 and 1. The void ratio, on the other hand, can vary between zero (for $\phi = 0$) and infinity (for $\phi = 1$).

The volume fraction is usually determined from the mass fraction (amount of dry mass of soil present in the water). The dry mass of soil of a sample can be estimated from drying the sample in an oven at 105°C for several hours. The density of a clay gives the relation between mass and volume.

The size of the voids in a soil can be estimated by **mercury intrusion** and **nitrogen desorption** techniques.

Mercury is a non-wetting fluid and when it is injected at a measured pressure into a material the pore size can be estimated based on the relation (**Washburn's equation**) between pressure and pore size: the pressure difference is inversely proportional to pore size.

Nitrogen is a non-corrosive gas that can be used to determine the specific surface area of the pores, where the **BET** (Brunauer – Emmett – Teller) theory is used to quantify the adsorption of the gas molecules on the solid surfaces. The BET theory is an extension for polylayers of the monolayer **Langmuir theory** for the adsorption of molecules.

In a similar way, we can define the volume fraction of water by:

$$\phi_w = \frac{V_w}{V}$$

We recall that in the relation above we have defined:

$$\phi_w = \frac{dV_w}{dV}$$

There is no contradiction in these definitions. The first one is a general one: the volume fraction of water in a given system is the ratio between the volume of water present and the total volume. The second definition is more specific: the volume ratio at a given position z is the ratio between the small volume of water at z and the small total volume at z. To emphasize that the second relation is at a given z, we could write:

$$\phi_w(z) = \frac{dV_w}{dV}$$

The flux of water can also be expressed in the rest frame of the particles. If the particles are not moving, they are at rest in the laboratory frame their velocity is zero and therefore:

$$v_{s/lab} = 0$$

From which we deduce:
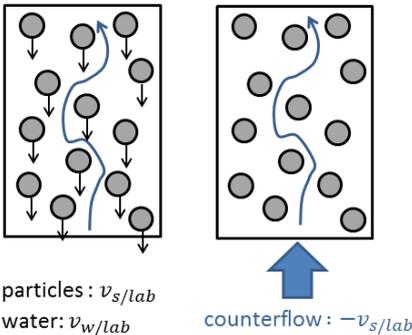
$$v_{w/lab} = v_{w/s} + v_{s/lab} = v_{w/s}$$

where $v_{w/s}$ is the velocity of water inside the pores in the frame of reference of the soil (s). The relation between the macroscopic water flux coming out of the porous media (measured in the reference frame of the laboratory) can now be expressed as function of the water flux in the reference frame of the particles as:

$$J_{w/lab} = (1 - \phi_S)v_{w/s} = \frac{-k}{\eta}\nabla P$$

where $J_{w/lab}$ is the macroscopic velocity of the fluid (m/s), k the permeability (m²).

A Darcy equation for settling particles

We want to write a similar relation, but now in the case that the particles are moving (settling) in a column. We do this by setting-up a thought experiment. This is illustrated by the picture opposite. The left figure depicts, in the frame of the laboratory, what is happening during settling: both water and particles are moving with velocities that we define as $v_{w/lab}$ and $v_{s/lab}$ respectively. On the right figure (our thought experiment), we have applied a (fictive) macroscopic counter flow $-v_{s/lab}$ generated by a (not yet defined) pressure gradient ∇P^*. In that case, the new velocities of water and particles $v^*_{w/lab}$ and $v^*_{s/lab}$ become:

particles : $v_{s/lab}$
water: $v_{w/lab}$

counterflow : $-v_{s/lab}$

$$v^*_{w/lab} = v_{w/lab} - v_{s/lab} = v_{w/s}$$

$$v^*_{s/lab} = v_{s/lab} - v_{s/lab} = 0$$

We are now satisfying Darcy conditions, as only water is moving and particles are immobile. We can therefore write:

$$J^*_{w/lab} = -v_{s/lab} = \frac{-k}{\eta} \nabla P^*$$

In a next step, we would like to express $v_{s/lab}$ (which is both the microscopic velocity of the settling particles and the minus the macroscopic flow $J^*_{w/lab}$ applied in the mind experiment) as function of $v_{w/s}$, i.e. the microscopic velocity of the water in the frame of the particles. We make the hypothesis that there is not net volume flux in the settling column (this implies that no water or soil is entering or leaving the column during the experiment[122]). This condition imposes that the volume flux of water should be minus the volume flux of soil at any height z:

$$J_{w/lab} = -J_{s/lab}$$

Using the same lines of derivation as above, we have:

$$J_{s/lab} = \phi_S v_{s/lab}$$

which leads to:

$$(1 - \phi_S)v_{w/lab} + \phi_S v_{s/lab} = 0$$

This equation can further be developed into:

$$(1 - \phi_S)(v_{w/s} + v_{s/lab}) + \phi_S v_{s/lab} = 0$$

$$(1 - \phi_S)v_{w/s} = -v_{s/lab}$$

The relation between the macroscopic water flux inside the column in the mind experiment and the macroscopic one in the real experiment is given by:

$$J_{w/s} = (1 - \phi_S)v_{w/s} = J^*_{w/lab}$$

Note the difference between the fluxes: one is in the frame of reference of the laboratory, the other in the frame of reference of the soil. Using the Darcy equation obtained for $J^*_{w/lab}$ we now obtain a **modified Darcy equation for settling particles in an undrained column**:

---

[122] Usually one speaks of an undrained system: in most experimental cases (also in-situ) the question is whereas water is leaving or not the system, either by seepage at the bottom of the column, or by evaporation at the top. The soil concentration is not varying in typical experiments.

$$J_{w/s} = \frac{-k}{\eta} \nabla P^*$$

The question remains so as to properly define ∇P^*. We have said that it is the pressure gradient necessary to generate the (fictive) counter flow $-v_{s/lab}$ in the case of our mind experiment. Conversely, it can be seen as *minus* the pressure gradient created by the settling of the particles in the real experiment. Note that ∇P^* is a macroscopic pressure difference, generating the macroscopic fluid velocity $J_{w/s}$. Other relations could be set-up to express the microscopic fluid flow $v_{w/s}$ and microscopic pressure gradients could then be defined, acting within the pores of the fabric. These relations are beyond our present scope. We are now going to discuss more into details the form of ∇P^*.

Pressure gradient

In Darcy experiments, a pressure difference is applied across the sample and the water flux is measured. The pressure gradient can then be defined as:

$$\nabla P = \frac{\Delta P}{L}$$

where ΔP is the pressure difference between one end and the other of the sample, in the direction of the water flow, and L is the length of that section of sample. In the case of settling particles in an undrained column, no external pressure is applied. There is the **hydrostatic pressure**, due to gravity:

$$P_{hyd}(z) = \rho_w g(h - z) + P_0$$

where ρ_w is the density of water, g is the gravitational acceleration, h is the height of the liquid in the column, $z = 0$ is the position at the bottom of the column and P_0 is the atmospheric pressure. This hydrostatic pressure does not contribute to $J_{w/s}$. It is easy to convince oneself of this fact: let's assume that the column is filled by a porous media such as a rock or a pile of glass beads. The fabric is not moving, and neither is the water, even though a hydrostatic pressure is present.

There is however another pressure borne by the water and caused by the settling of the particles. We define P_w as the total pressure borne by the water. In order to express P_w we will use the concept of **stress**. We recall that:

Pressure is a force applied perpendicular to a surface per unit area.

Stress is a force per unit of surface, not necessarily perpendicular to that surface. Stress can arise from the internal forces that neighbouring particles exert on each other: stress might therefore exist in the absence of external forces.

Strain is the measure of the deformation of a material related to a stress. It has no dimension. For instance it can be the difference between the length of a bar before and after applying a stress, divided by the length of the bar (before or after applying the stress). Stress can exist in a system without noticeable strain.

In the case of a column of soil, fully saturated with water, the stress in the vertical direction is defined by:

$$\sigma_{zz} = P_w + \sigma_{sk}$$

where σ_{sk} is the stress associated to the skeleton. In order to get a better understanding of σ_{zz} let us first consider the case where the particles are not settling (a rock saturated with water for instance) with everywhere the same volume fraction ϕ_S. We then have hydrostatic conditions (the water is not moving) and hence:

$$P_w(z) = P_{hyd}(z)$$

We then also have:

$$\sigma_{zz}(z) = \rho g(h - z) + P_0$$

with

$$\rho = (1 - \phi_S)\rho_w + \phi_S \rho_s$$

where ρ is the density of the sample (water + soil) and ρ_s is the density of the soil itself. Recombining the equations, we find:

$$\sigma_{sk}(z) = (\rho_s - \rho_w)g(h - z)\phi_S$$

If the soil and the water have the same density ($\rho_s = \rho_w$), we get $\sigma_{sk} = 0$ and the total vertical stress is the same as the one given by a column full of water:

$$\sigma_{zz}(z) = \rho_w g(h - z) + P_0$$

which is expected. In the other limit, when the column is only filled with soil (no water), we have $\phi_S = 1$ and:

$$\sigma_{zz}(z) = \rho_s g(h - z) + P_0$$

In that case $\sigma_{zz} = \sigma_{sk}$ as one can consider that $P_w = 0$ since there is no water. In the case of settling particles in a column (the water is moving), we now have $P_w(z) \neq P_{hyd}(z)$ and the pressure P^* we are looking for in order to define the pressure gradient can be defined as:

$$P^*(z) = P_e(z) = P_w(z) - P_{hyd}(z)$$

This pressure is called the **excess pore pressure** and is usually noted P_e. In hydrostatic conditions, we have already discussed that $P_w = P_{hyd}$ and therefore $P_e = 0$.

It is important to realize that the excess pore pressure is solely generated by the settling *velocity* of the particles. Due to their slow motion, stresses caused by vertical *accelerations* are not considered. The excess pore pressure is the pressure leading to the water flow $v_{w/s}$.

We have:

$$\nabla P_e = \nabla P_w - \nabla P_{hyd}$$

Using the definition $\sigma_{zz} = P_w + \sigma_{sk}$ we get:

$$\nabla P_e = \nabla(\sigma_{zz} - \sigma_{sk}) - \nabla P_{hyd}$$

In order to estimate σ_{zz} we make the hypothesis that the consolidation is slow and that we have nearly hydrostatic conditions (fluid accelerations are neglected):

$$d\sigma_{zz}(z) = -\left[\left(1 - \phi_S(z)\right)\rho_w + \phi_S(z)\rho_s\right]g\,dz$$

We find:

$$\frac{\partial P_e}{\partial z} = \rho_w g + \frac{\partial \sigma_{zz}}{\partial z} - \frac{\partial \sigma_{sk}}{\partial z}$$

Leading to:

$$-\frac{\partial P_e}{\partial z} = \frac{\partial \sigma_{sk}}{\partial z} + (\rho_s - \rho_w)g\phi_S$$

The Gibson equation

The Gibson equation gives the changes of volume fraction ϕ_S as function of time and space. The solution of this equation is used to fit the consolidation data and obtain relevant parameters for the system, one of which being the permeability. We will now set-up this equation by combining the Darcy equation for settling particles and the continuity equation.

The **continuity equation** imposes that mass is conserved. This can be written like:

$$\frac{\partial \phi_S}{\partial t} + \frac{\partial\left(v_{s/lab}\phi_S\right)}{\partial z} = 0$$

**General derivation of the continuity equation**

In order to derive the continuity equation, one has to put in equation the following observation: conservation of mass (a very fundamental principle, that only breaks down when one studies nuclear reactions) requires that no mass can be lost or gained from a small volume element. Let us first define mass as:

$$m = \rho V$$

where ρ is the density of the medium and V symbolizes the (small) volume considered which has a (small) mass m. We now consider a small volume element V (symbolized by the red cylinder underneath), through which a fluid is moving in the x-direction and we define the change of mass inside this volume as function of time as:

$$dm = m_{in} - m_{out}$$

Both the entering ("in") and exiting ("out") mass are defined as $m_k = \rho \cdot V_k$ where the volumes V_k are defined in grey in the figure and k stands for "in" or "out".
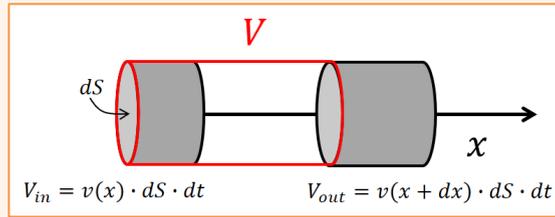
The velocity of the fluid is given by v. We get:

$$dm = \rho(x) \cdot v(x) \cdot dS \cdot dt - \rho(x + dx) \cdot v(x + dx) \cdot dS \cdot dt$$

$$dm = -\frac{d[\rho(x) \cdot v(x)]}{dx} \cdot dS \cdot dx \cdot dt$$

$$\frac{dm}{dt} = -\frac{d[\rho(x) \cdot v(x)]}{dx} \cdot dV$$

The same reasoning can be made if the fluid would move in the other directions (y, z) as well. If we add all the mass differences from all directions, we get:

$$\frac{d(m_x + m_y + m_z)}{dt} = -\left[\frac{d(\rho \cdot v_x)}{dx} + \frac{d(\rho \cdot v_y)}{dy} + \frac{d(\rho \cdot v_z)}{dz}\right] \cdot dV$$

In the illustration we have taken dS to be circular. This is not a requirement. The only assumption that is needed is that the velocity vector should always be taken

perpendicular to the surface (if it is not, only the component perpendicular to the surface should be used). Mathematically, this means that one can write the previous equation in the general form:

$$\frac{dm}{dt} = -\boldsymbol{\nabla} \cdot (\rho \cdot \boldsymbol{v}) \cdot dV$$

where dm is the total mass change inside the volume element considered (which can be of any shape). This equation can be expressed in any coordinate system. In Cartesian coordinates, one will find the equation we derived above.

As we are looking at a very general case, we may assume that inside the volume element considered there exist sources and sinks, where matter (and mass) can be appearing or disappearing: for instance roots that can take up water, or a leaking pipe. This leads to a change in mass that can be quantified by fluxes:

$$dm_{source\ or\ sink} = \rho \cdot (q_{source} - q_{sink}) \cdot dV \cdot dt$$

where q represent the flux in (s^{-1}). We therefore get:

$$\frac{dm_{tot}}{dt} = -\boldsymbol{\nabla} \cdot (\rho \cdot \boldsymbol{v}) \cdot dV + \rho \cdot (q_{source} - q_{sink}) \cdot dV$$

Realizing that $dm_{tot} = \rho dV$ and

$$dm_{tot} = \left(\frac{\partial m_{tot}}{\partial t}\right)_{\text{at V fixed}} dt = \frac{\partial \rho}{\partial t} \cdot dV \cdot dt$$

we get:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \cdot \mathbf{v}) = \rho \cdot (q_{\text{source}} - q_{\text{sink}})$$

This is the general form of the continuity equation. In the specific case that the particle are incompressible (ρ_s is constant), and that there are no sink and source terms, and that the movement is along the z-axis, we may use $\rho = \rho_s \phi_s$ and we find

$$\frac{\partial \phi_S}{\partial t} + \frac{\partial \left(v_{s/lab}\phi_S\right)}{\partial z} = 0$$

One can show that the continuity equation can also be expressed as:

$$-\frac{\partial e}{\partial t} + (1 + e)^2 \frac{\partial}{\partial z}\left(\frac{v_{s/lab}}{1 + e}\right) = 0$$

This is the form most used in geoengineering. In hydraulic engineering, where the free settling phase of the particles is also considered, it is preferable to work with volume fraction. The reasons are expressed by Eric Toorman in 1996[123]:

"In this text the solids volume fraction ϕ_S is used as the dependent variable to describe the solids content. There are several reasons for this. The solids concentration is preferred to the void ratio because in the sedimentation phase e becomes infinite when the concentration decreases to zero. Furthermore, if one has to consider the behaviour of a polydisperse sediment, the particle concentration must be taken as the dependent variable, since the mass balance must be solved now for each fraction separately. The volumetric solids concentration is also the best variable to relate other suspension properties (e.g. rheological properties) with."

We recall the modified Darcy equation obtained above:

$$-v_{s/lab} = \frac{-k}{\eta}\frac{\partial P_e}{\partial z}$$

Combining these equations, one obtains:

$$\frac{\partial \phi_S}{\partial t} = -\frac{\partial}{\partial z}\left(\frac{k}{\eta}\frac{\partial P_e}{\partial z}\phi_S\right)$$

which can further be developed into:

$$\frac{\partial \phi_S}{\partial t} = \frac{\partial}{\partial z}\left(\frac{k}{\eta}(\rho_s - \rho_w)g\phi_s^2 + \frac{k}{\eta}\phi_S\frac{\partial \sigma_{sk}}{\partial z}\right)$$

This equation is called the **Gibson equation**. Usually a new permeability is defined as:

$$K\,(\text{m/s}) = k(\text{m}^2)\frac{g\rho_w}{\eta}$$

where $K\,(\text{m/s})$ is usually called **hydraulic conductivity** and $k(\text{m}^2)$ **permeability**. At the beginning of the chapter we have seen that $K\,(\text{m/s})$ is also called coefficient of permeability by some authors.

The Gibson equation[124] can be written:

---

[123] Toorman, E.A. (1996) "Sedimentation and self-weight consolidation: general unifying theory", Géotechnique 46, No 1, 103-113.
[124] Gibson, R. E., G. L. England, and M. J. L. Hussey. "The Theory of one-dimensional consolidation of saturated clays: 1. finite non-Linear consildation of thin homogeneous layers." Geotechnique 17.3 (1967): 261-273.

$$\frac{\partial \phi_S}{\partial t} = \frac{\partial}{\partial z}\left(K\frac{\rho_s - \rho_w}{\rho_w}\phi_s^2 + \frac{K}{g\rho_w}\phi_S \frac{\partial \sigma_{sk}}{\partial z}\right)$$

Consolidation when gas is present in the soil

We now consider the consolidation of a soil in which gas bubbles are trapped[125]. We will only study the case where t > t₁ (the interface water/suspension has disappeared). If the gas bubbles are fixed to the soil skeleton one simply gets:

$$J_{w/s} = S_r(1 - \phi_S)v_{w/s}$$

where S_r is the degree of saturation : $S_r = 1$ is for the case there is no gas inside the soil.

We will now consider the general case, where gas bubbles are distributed within the soil in an undetermined way. One defines the porosity as:

$$\phi = \phi_g + \phi_w$$

where:

$$\phi_g = (1 - S_r)\phi \quad ; \quad \phi_w = S_r\phi = S_r(1 - \phi_S)$$

The total vertical stress can be estimated, for t > t₁, by:

$$\sigma_{zz}(z) = \int_z^{h_b} \rho(z)g dz + \int_{h_b}^{L} \rho(z)g dz + P_0$$

Between z and h_b we have :

$$\rho(z) = \phi_w\rho_w + \phi_S\rho_s + \phi_g\rho_g$$

If we assume that we can neglect the weight of the gas ($\rho_g \ll \rho_s, \rho_w$), we get:

$$\rho(z) = \phi_w\rho_w + \phi_S\rho_s$$

Therefore:

$$\sigma_{zz}(z) = \int_z^{h_b(z)} [\phi_w\rho_w + \phi_S\rho_s]g dz + \rho_w g(L - h_b) + P_0$$

---

[125] For more details, see " Consolidation behaviour of gassy mud: theory and experimental validation", B. Wichman, PhD thesis (1999), TU Delft

For numerical reasons, one likes to avoid having h_b , which is depending on z, as an upper bound of an integral. To get a fixed upper bound, the following change in variable is introduced:

$$dz_s = \phi_s dz$$

The variable z_s is called the **material coordinate** (for the soil).

Material coordinates
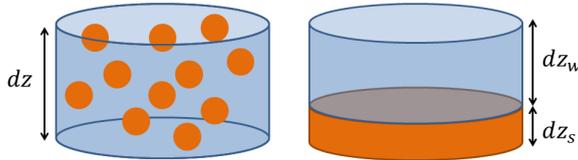
A small volume $dV = S\,dz$ where S is the cross section area of the column. The volume can be expressed as : $dV = dV_w + dV_s$ where dV_w represents the volume of water and dV_s the volume of soil Left figure: soil particles in suspension in dV. Right figure: the height dz_w and dz_s are defined by: $dV_w = S\,dz_w$ and $dV_s = S\,dz_s$.

Let us consider a slice dz of the column of cross area S. This slice is composed of water and soil, therefore we can define two (fictive) heights such that Sdz_w corresponds to the volume of water inside the slice and Sdz_s to the volume of soil. We have:

$$Sdz = Sdz_w + Sdz_s$$

Dividing both sides of this equation by Sdz we get:

$$1 = \frac{Sdz_w}{Sdz} + \frac{Sdz_s}{Sdz}$$

realising that:

$$\frac{Sdz_s}{Sdz} = \frac{V_s}{V} = \phi_s \text{ and } \frac{dz_w}{dz} = \phi_w$$

we get:

$$\phi_s + \phi_w = 1$$

The **material coordinate** is defined, is given by:

$$z_s(z) = \int_0^z \phi_s dz = \int_0^{z_s} dz_s$$

It is the height of solid there would be theoretically found in the column if all the solids between 0 and z would be squeezed at the bottom. Note that:

$$z_s(h) = L_G$$

where h is the height of the suspended particles + bed. (we assume to be in the special case where $h = h_b$). Following a similar reasoning one can define the material coordinate for water:

$$z_w(z) = \int_0^z \phi_w dz = \int_0^{z_w} dz_w$$

The material coordinate z_s is linked to the **Gibson height** (also called **material height**) that we will now define.
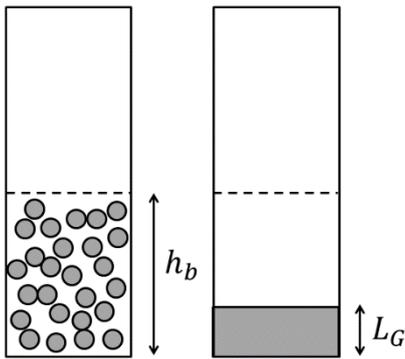
If all the particles of the left could be squeezed at the bottom of the column, a height L_G would be obtained. Because of conservation of mass, this is a constant whereas h_b varies in time as the particles settle.

The Gibson height L_G is a (constant) length obtained from the conservation of mass within the undrained column. It can be seen as the average volume fraction of soil per unit of area.

The **Gibson height** can easily be obtained using the fact that at t = 0:

$$L_G = \int_0^h \phi_0 dz = L\phi_0$$

where L is the height of the fluid in the column.

We get:

$$\sigma_{zz}(z_s) = \int_{z_s}^{L_G} \left[\frac{\phi_w \rho_w + \phi_s \rho_s}{\phi_s} \right] g dz_s + \rho_w g(L - h_b) + P_0$$

As we have $z_s(h_b) = L_G$ (the Gibson height) which is a constant (it represents the soil height that would be obtained in the column if no voids were present) the upper bound now does not anymore depend on z.

The material coordinates can conveniently be used to find again some of the results we have derived earlier, when setting-up the Darcy equation for settling particles. For a water/soil/gas system, one gets:

$$dz = dz_w + dz_s + dz_g$$

Using the fact that the volume is conserved (implying that each volume element is conserved as function of time i.e. $dz/dt = 0$) we obtain

$$\left(\frac{dz_w}{dt}\right)_z + \left(\frac{dz_g}{dt}\right)_z = -\left(\frac{dz_s}{dt}\right)_z$$

which should be read as: at a given position z in the column, the flux of water and gas is minus the flux of soil. In the notations adopted above, in the case there is no gas, we find as previously:

$$J_{w/lab} = -J_{s/lab}$$

The change in time of z_w at a constant z_s is the velocity of water for a fixed soil skeleton. This is what we can define as $J_{w/s}(z)$:

$$J_{w/s}(z) = \left(\frac{\partial z_w}{\partial t}\right)_{z_s}$$

Note that this equation is valid independent of the fact that gas is present or not in the soil. Using

$$dz_w = d(\phi_w z) = d\left(\frac{\phi_w}{\phi_s} z_s\right) = d(e_w z_s)$$

we get:

$$J_{w/s}(z) = z_s \left(\frac{\partial e_w}{\partial t}\right)_{z_s}$$

Leading to:

$$\left(\frac{\partial J_{w/s}}{\partial z_s}\right)_t = \left(\frac{\partial e_w}{\partial t}\right)_{z_s}$$

Using the Darcy equation for settling particles we get:

$$\left(\frac{\partial e_w}{\partial t}\right)_{z_s} = \frac{\partial}{\partial z_s}\left[\frac{k}{\eta}\frac{\partial P_e}{\partial z}\right] = \frac{\partial}{\partial z_s}\left[\phi_s \frac{k}{\eta}\frac{\partial P_e}{\partial z_s}\right]$$

We recall that the derivative of the excess pore pressure is given by:

$$\frac{\partial P_e}{\partial z} = \rho_w g + \frac{\partial \sigma_{zz}}{\partial z} - \frac{\partial \sigma_{sk}}{\partial z}$$

From which we get:

$$\frac{\partial P_e}{\partial z} = \left[\left(\phi_s(z) + \phi_g(z) \right) \rho_w - \phi_s(z)\rho_s \right] g - \frac{\partial \sigma_{sk}}{\partial z}$$

$$\frac{\partial P_e}{\partial z} = -\phi_s(z) \left[\rho_s - \left(1 + e_g(z) \right) \rho_w \right] g - \frac{\partial \sigma_{sk}}{\partial z}$$

$$\frac{\partial P_e}{\partial z_s} = -[\rho_s - (1 + e_g)\rho_w]g - \frac{\partial \sigma_{sk}}{\partial z_s}$$

This leads to:

$$\left(\frac{\partial e_w}{\partial t} \right)_{z_s} = -\frac{\partial}{\partial z_s} \left[\phi_s \frac{k}{\eta} \left(\frac{\partial \sigma_{sk}}{\partial z_s} + (\rho_s - (1 + e_g)\rho_w)g\phi_s \right) \right]$$

This is the general formulation of the Gibson equation, in the case that gas bubbles are distributed within the soil matrix. We also deduce that:

$$J_{w/s} = -\phi_s \frac{k}{\eta} \left(\frac{\partial \sigma_{sk}}{\partial z_s} + (\rho_s - (1 + e_g)\rho_w)g\phi_s \right)$$

If $e_g = 0$ (there is no gas in the soil), then:

$$\left(\frac{\partial e_w}{\partial t} \right)_{z_s} = -\frac{\partial}{\partial z_s} \left[\phi_s \frac{k}{\eta} \left(\frac{\partial \sigma_{sk}}{\partial z_s} + (\rho_s - \rho_w)g\phi_s \right) \right]$$

$$\left(\frac{\partial \phi_s}{\partial t} \right)_{z_s} = \phi_s^2 \frac{\partial}{\partial z_s} \left[\phi_s \frac{k}{\eta} \left(\frac{\partial \sigma_{sk}}{\partial z_s} + (\rho_s - \rho_w)g\phi_s \right) \right]$$

Using once more $dz_s = \phi_s dz$ and realizing that :

$$\left(\frac{\partial \phi_s}{\partial t} \right)_{z_s} = \frac{d\phi_s}{dt}$$

we get:

$$\frac{d\phi_s}{dt} = \phi_s \frac{\partial}{\partial z} \left[\frac{k}{\eta} \left(\frac{\partial \sigma_{sk}}{\partial z} + (\rho_s - \rho_w)g\phi_s \right) \right]$$

which is the original Gibson equation (when no gas is present).

We found earlier for the Gibson equation:

$$\frac{\partial \phi_S}{\partial t} = \frac{\partial}{\partial z}\left[\frac{k}{\eta}\left(\phi_S \frac{\partial \sigma_{sk}}{\partial z} + (\rho_s - \rho_w)g\phi_s^2\right)\right]$$

Where does the difference come from? To find the answer, we have to look back to one of the equation we have used to set-up the Gibson equation when we introduced it the first time: the continuity equation. We have expressed this equation as:

$$\frac{\partial \phi_S}{\partial t} + \frac{\partial\left(v_{s/lab}\phi_S\right)}{\partial z} = 0$$

This equation is the **Eulerian** form of the continuity equation. This implies that we follow the changes in ϕ_S at a *given* position z as function of time. Doing so, we say that the changes in ϕ_S in time at a position z are caused by the inflow and outflow of matter at position z. The rate of inflow/outflow is in our case the velocity $v_{s/lab}$. The **Lagrangian** form of the same equation gives the changes in ϕ_S for a given volume element within the laboratory frame of reference *that we follow in space and time*. The link between the Eulerian and Lagrangian descriptions is given by:

$$\frac{d\phi_S}{dt} = \frac{\partial \phi_S}{\partial t} + v_{s/lab}\frac{\partial \phi_S}{\partial z}$$

where $d\phi_S/dt$ gives the Lagrangian derivative and $\partial \phi_S/\partial t$ is the Eulerian one. This equation tells us that the changes in ϕ_S of the volume element we follow and currently is located at (z, t) is given by the change of ϕ_S (term $\partial \phi_S/\partial t$) at position z plus the change associated to the fact that the volume element we follow is leaving position z at the instant t (term $v_{s/lab}(\partial \phi_S/\partial z)$). In Lagrangian form, the continuity equation is therefore:

$$\frac{d\phi_S}{dt} - v_{s/lab}\frac{\partial \phi_S}{\partial z} + \frac{\partial\left(v_{s/lab}\phi_S\right)}{\partial z} = 0$$

$$\frac{d\phi_S}{dt} + \phi_S\frac{\partial v_{s/lab}}{\partial z} = 0$$

We can now combine the Lagrangian form of the continuity equation with the Darcy equation

$$v_{s/lab} = \frac{k}{\eta}\frac{\partial P_e}{\partial z}$$

and we get:

$$\frac{d\phi_S}{dt} = -\phi_S \frac{\partial}{\partial z}\left(\frac{k}{\eta}\frac{\partial P_e}{\partial z}\right)$$

which can further be developed into:

$$\frac{d\phi_S}{dt} = \phi_S \frac{\partial}{\partial z}\left(\frac{k}{\eta}(\rho_s - \rho_w)g\phi_S + \frac{k}{\eta}\frac{\partial \sigma_{sk}}{\partial z}\right)$$

This is the original form of the Gibson equation – and the one we have found above.

There exists numerical schemes that enable to solve the Gibson equation, in the presence or not of gas. In order to solve these equations, hypothesis have to be made for the dependence of k and σ_{sk} on z (and for the dependence of e_g on z is gas is considered).

In the next chapter (**Chapter 10**) we are going to present analytical and numerical solutions for the Gibson equation, in the absence of gas.

Lagrange and Euler derivatives

To illustrate the difference between the Lagrangian and Eulerian derivatives, let us consider a variable that occupies the space we study. For instance, in a fluid, we could measure the temperature T at any position in time and space.
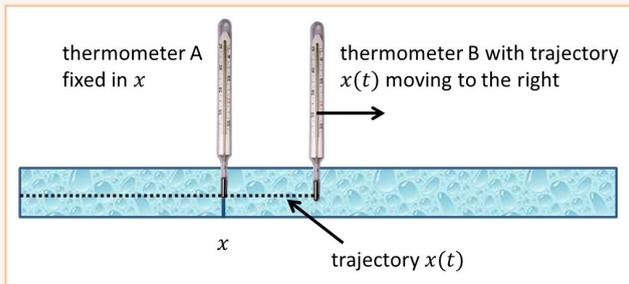
For simplicity, we will consider a stripe of fluid in direction x. We are measuring the temperature with a thermometer A that we place at a fixed position x and a thermometer B that we are moving along x. With thermometer B we get the temperature as a function of $x(t)$, i.e. the trajectory of the thermometer. This implies that for a given time t we are at a specific location x, i.e. x and t are correlated. In compact notation, we say that we get $T(x(t))$, that is the temperature along the trajectory $x(t)$. With thermometer A, we get the temperature at a fixed position x, where this x is therefore not dependent on time. In that case, in compact notation, we say that we get $T(x,t)$, that is the temperature for a fixed position x as a function of time. Of course, when thermometer B is at the same position as thermometer A, the temperatures are the same:

$$T\big(x(t)\big) = T(x,t)$$

However, the variations in temperature measured with thermometers A and B can be different. From mathematics, we get:

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{\partial T}{\partial x}\frac{dx}{dt}$$

where dT/dt is the change in temperature measured with thermometer B and $\partial T/\partial t$ the change in temperature measured with thermometer A. $\partial T/\partial x$ represents the change in temperature with displacement dx around time t of the measurement and $dx/dt = v_B$ is the velocity of thermometer B. From the relation above, we see that if the temperature does not dependent on time (but is not the same at every position x) we get

$$\frac{dT}{dt} = \frac{\partial T}{\partial x} v_B$$

which means that even though for a given x the temperature does not change ($\partial T/\partial t = 0$), the thermometer B will indicate a change of temperature in time as it is moving. Only if the thermometer B is not moving or if the temperature is the same at any x do we get $dT/dt = 0$.

Finding the permeability using colloid science

So far we have discussed the consolidation of soft soils so as to determine the permeability from the physical compaction (and in particular the water/mud interface as function of time) of sediment beds. Quite some research has been performed over the years to devise other methods for finding information about the permeability and porosity of consolidated soils, also as function of environmental parameters such as salinity, temperature and clay surface properties.

A simple and extremely popular relation is provided by the empirical equation found by **Gustavus E. Archie** (1907-1978) in 1942. Archie's law states that the electric conductivity σ of a porous media[126] saturated by an electrolyte is proportional to its porosity such that:

$$\sigma = \phi^m \sigma_e$$

where σ_e is the conductivity of the electrolyte, $\phi = (1 - \phi_S)$ is the porosity and m is an exponent to be fitted which varies usually between 1 and 4.

---

[126] In this chapter we will use the symbol σ for the conductivity to avoid any confusion with k and K that are used for the hydraulic conductivity. In **Chapter 3** we used K for expressing the conductivity and σ for designing a surface charge.

Quite some work has also been devoted to relate the electric conductivity σ to the hydraulic permeability k. In a first approach, one makes use of the relation:

$$k \approx A\, \phi r^2$$

where A is an unknown shape factor (that might or not depend on ϕ) and r a representative length for a pore radius (that might or not depend on ϕ). The relation originates from the **Hagen-Poiseuille** relation we have already mentioned above. Combining the two equations, one finds the link between conductivity and permeability to be:

$$k \approx A \left(\frac{\sigma}{\sigma_e}\right)^{1/m} r^2$$

More elaborate models have been created, for instance by assuming that the porous media is formed by a collection of colloidal particles. The parameter m can then be calculated[127].

If a surface is charged, as we have seen in **Chapter 3**, applying an electric field will provide information about this charge. This information can then be correlated to the permeability as we will see now for the simple example of a bundle of pores.

Bundle of charged pores in an electric field

If one applies and electric field $\boldsymbol{E} = E\, \boldsymbol{e_y}$ in the direction parallel to a charged pore filled with an electrolyte, the ions inside the pore will start to move. In the stationary state, the ions will have a constant velocity, implying that the total sum of forces on each ion must be zero. The electric force is compensated by the drag force exerted by the liquid.

If we now consider the forces exerted on the liquid, one finds that:

- in the stationary state, each layer of fluid dx moves with a uniform velocity parallel to the walls because the total force on such layer is zero.

- In virtue of Newton's third law (action = reaction) the force exerted by the ions on the liquid is equal in magnitude to the (electric) force exerted directly on the ions and which produce the movement.

---

[127] P.N. Sen et al. "A self-similar model for sedimentary rocks with application to the dielectric constant of fused glass beads", Geophysics, 46,5,781-795 (1981); S. Kostek et al. "Fluid permeability in porous media: Comparison of electrical estimates with hydrodynamical calculation", Physical Review B, 45, 1 (1992); Kirichek, A., C. Chassagne, and R. Ghose. "Dielectric spectroscopy of granular material in an electrolyte solution of any ionic strength." Colloids and Surfaces A: Physicochemical and Engineering Aspects 533 (2017): 356-370.

- A frictional force is exerted on a layer by the neighbouring layers which move with different velocities.
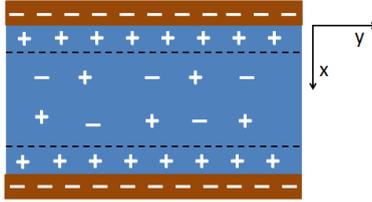
Schematic representation of a section of a capillary. The capillary walls are negatively charged. Double layers are represented by black dashed lines. Inside the double layers, one finds a majority of couterions (+), even though co-ions (-) are also present (but not represented for simplification). In the bulk of the capillary, far away of the double layers, the ionic concentration is independent of the surface charge of the pores.

This enables to write:

$$E\rho dx = \eta \left(\frac{du}{dx}\right)_{x+dx} - \eta \left(\frac{du}{dx}\right)_x = \eta \left(\frac{d^2u}{dx^2}\right) dx$$

where ρ is the ionic concentration.

Using **Poisson's equation** introduced in Chapter 3, here in Cartesian coordinates, one finds:

$$-\varepsilon_0 \varepsilon_r E \left(\frac{d^2\psi}{dx^2}\right) = \eta \left(\frac{d^2u}{dx^2}\right)$$

This equation can be integrated a first time:

$$-\varepsilon_0 \varepsilon_r E \int_{r/2}^{x} \left(\frac{d^2\psi}{dx^2}\right) dx = \eta \int_{r/2}^{x} \left(\frac{d^2u}{dx^2}\right) dx$$

and $r/2$ is the middle of the capillary. Using the fact that, by symmetry, both the derivatives of ψ and u are zero at $r/2$ we get:

$$-\varepsilon_0 \varepsilon_r E \left(\frac{d\psi}{dx}\right) = \eta \left(\frac{du}{dx}\right)$$

By integrating a second time:

$$-\varepsilon_0 \varepsilon_r E \int_{x_s}^{x} \left(\frac{d\psi}{dx}\right) dx = \eta \int_{x_s}^{x} \left(\frac{du}{dx}\right) dx$$

Using the boundary conditions:

$$\psi(x_s) = \zeta \quad ; \quad u(x_s) = 0$$

where x_s is the position of the slip plane (which is very close to the capillary wall) and ζ the **zeta potential** of the capillary tube, we get:

$$-\varepsilon_0 \varepsilon_r E \, \zeta = \eta u(x)$$

Rearranging this equation, we can write:

$$u(x) = \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta} E \quad \text{for} \quad x > x_s$$

This formula is similar to the one we have found for the **Smoluchowski** electrophoretic mobility of a charged sphere in an electric field. Similarly to what we have seen for spheres, a double layer exists close to the pore walls. In our example, we have assumed that the double layers do not overlap, which is most certainly true for not too small pores and reasonable ionic strength. In case of overlapping double layers, numerical solutions exist.

One should also realize that the derivation presented here is implicitly done for very thin double layers κ^{-1} compared to r as we have used Cartesian coordinates. For arbitrary κr one should work in cylindrical coordinates, assuming of course that the pores are cylindrical, which is generally a valid assumption.

The flux of water exiting the capillary can be estimated by:

$$J = Su = S \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta} E$$

where S is the cross section of the capillary. For a bundle of capillaries (all parallel to each other), one simply gets:

$$J_{w/s} = \phi S \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta} E$$

Contrary to the Darcy equation, the relation we found is a function of the applied electric field, since no pressure gradient was applied in the experiment. We have thus found that it is also possible to create a flow of *water* by applying an *electric* field. This is due to the fact that the ions are set in movement by the electric field and, by friction, set the water in motion. If the capillaries are uncharged ($\zeta = 0$) there is no net water flux ($J_{w/s} = 0$). This does not mean that the ions are not moving or that the water is not moving, but that the water flux created by the negative ions compensate the flux created by the positive ions such that no net flux is created. When the pores are charged (here negatively), there is a dissymmetry between the

fluxes: there are more positive ions that are mobile, compared to negative ones (of which some are fixed on the pores' walls) and hence a net flux is created.

Onsager relations

Is it possible to link the water flux we have found to the permeability of the porous medium? In order to find out, we will have to discuss more into details about what happens when one apply either a pressure gradient or an electric field to the same porous media. As before, we will here only consider a bundle of capillaries, fully saturated with electrolyte.

When an electric field E is applied to this porous medium, we have just found that we could create a water flux. By applying a pressure gradient ∇P, we also will create a water flux $J_{w/s}$ (m/s) and the relation between water flux and pressure gradient is then simply the Darcy equation. In general we can therefore write:

$$J_{w/s} = \phi S \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta} E - \frac{k}{\eta} \nabla P$$

This equation could be expanded by considering that a water flux can also be created by applying a temperature gradient or a ionic concentration gradient, etc… We will limit ourselves to electric and pressure gradients.

Another flux is created by the application of an electric field: a macroscopic electric flux (electric current J_e in A/m^2) which is proportional to the electric field following Ohm's law:

$$J_e = \sigma E$$

where σ is the electric conductivity (of the porous medium saturated with electrolyte). One could then make the following reasoning: by applying a pressure gradient, water will be set in motion. As this water contains ions, these will also be set into motion. The movement of charges create an electric current, and therefore an electric current can be created by applying a pressure gradient. When both electric and pressure gradients are considered, one can therefore write:

$$J_e = \sigma E - L_e \nabla P$$

where L_e is a nameless parameter. For commodity, we will rewrite the water flux in a more general form as:

$$J_{w/s} = L_w E - \frac{k}{\eta} \nabla P$$

This formulation is in fact valid for any porous media, as we have not specified the values of L_e, L_w, k or σ.

In 1968 **Lars Onsager** got the Nobel prize in Chemistry for having theoretically found that:

$$L_e = L_w$$

The formulation of Onsager was of course more general than the one given here, which is just one of the many examples of the **Onsager reciprocal relations**.

Conductivity of a porous medium consisting of a bundle of charged pores

Let us assume that a porous medium is submitted to a pressure gradient ∇P. By the movement of ions, an electric field E is then created. An example is given right underneath.
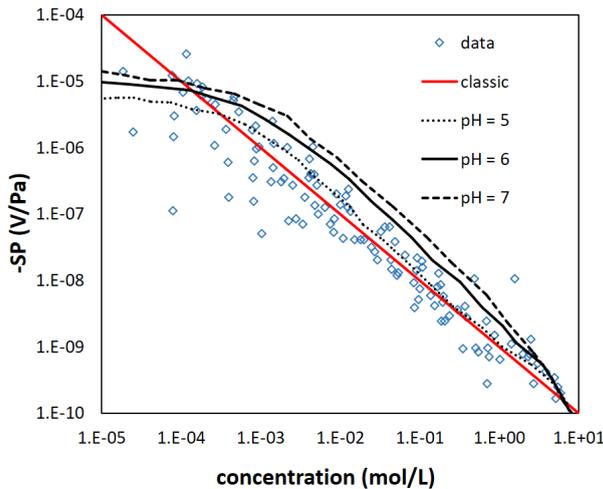
Streaming potential coupling coefficient in sandstones as function of NaCl concentration. Data from Jaafar et al.[128] The pH models are from Glover at al.[129] .The red line represents the classical dependence, explained in the text.

At steady state, assuming that no electrochemical reactions occur, and that no charges are created or destroyed, a bulk electric current will be created that opposes

---

[128] Jaafar, M. Z., J. Vinogradov, and M. D. Jackson, 2009, Measurement of streaming potential coupling coefficient in sandstones saturated with high salinity NaCl brine: Geophysical Research Letters, 36, L21306.

[129] P.W.J. Glover et al. "Streaming-potential coefficient of reservoir rock: A theoretical model", Geophysics, 77, 2 (2012).

the electric current along the walls (in the double layers). This leads to the fact that the (macroscopic) electric current is zero: $J_e = 0$. It follows that:

$$(E)_{J_e=0} = \frac{L_e}{\sigma} (\nabla P)_{J_e=0}$$

This relation defines the **streaming potential**: an electric field is created by the application of a pressure gradient in the absence of electric current.

The streaming potential coupling coefficient SP as given on the y-axis of the figure is defined by:

$$SP = \frac{-\Delta V}{\Delta P} = -\phi S \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta \sigma_e}$$

where σ_e is the conductivity of the electrolyte inside the pores. This expression can be found using the relation given above and realising that $E = -\nabla V$ and

$$L_e = L_w = \phi S \frac{\varepsilon_0 \varepsilon_r \zeta}{\eta}$$

Furthermore $\nabla V / \nabla P = \Delta V / \Delta P$ for a linear system ($\nabla P = \Delta P / L$ where L is the length of the bundle of pores). For a bundle of pores, the electric current at no pressure difference is given by:

$$J_e = \phi S \, \sigma_e E = \sigma E$$

One therefore finds:

$$SP = \frac{-\Delta V}{\Delta P} = -\frac{\varepsilon_0 \varepsilon_r \zeta}{\eta \sigma_e}$$

Recalling (see **Chapter 3**) that $\sigma_e \sim C$ where C is the salt concentration, one finds:

$$log(-SP) = -log(C) + \text{constant}$$

This implies that for the simple model of a bundle of pores, the logarithm of the streaming coupling coefficient should scale as minus the logarithm of the salt concentration. In the log-log plot above this corresponds to the red line of slope -1. The data presented on the figure has been measured on a sandstone saturated with high salinity NaCl. We can therefore conclude that the simple "bundle of pores" model is quite successful to explain the obtained results. From the data, we can also estimate the zeta potential: we find (in absolute value) 19 mV, which is a realistic value for such a system. Note that the models proposed by Glover et al., even though more elaborated, are less successful. Many models include so many adjustable parameters that they cannot provide an estimate for the zeta potential as they have

made the zeta potential dependent on parameters that are quite difficult to estimate.

Inserting the general streaming potential relation into the expression for $J_{w/s}$ and using the fact that $L_e = L_w$ leads to:

$$\left(J_{w/s}\right)_{J_e=0} = \left(\frac{L_e^2}{\sigma} - \frac{k}{\eta}\right)(\nabla P)_{J_e=0}$$

It is usually the case that

$$\frac{L_e^2}{\sigma} \ll \frac{k}{\eta}$$

which is why, in first approximation, Darcy's equation is generally used to express a flow of water. We note however, that the water flux exiting the porous medium could be smaller than the one expected on the basis of Darcy alone. The permeability can be correctly estimated by a series of experiments that accounts for both electric and hydraulic effects. One can easily show that:

$$L_w = \frac{k}{\eta}\left(\frac{\nabla P}{E}\right)_{J_{w/s}=0}$$

$$L_e = \sigma\left(\frac{E}{\nabla P}\right)_{J_e=0}$$

Combining these two equations using $L_e = L_w$ one finds:

$$k = \eta\sigma\frac{(E/\nabla P)_{J_e=0}}{(\nabla P/E)_{J_{w/s}=0}}$$

This means that two series of experiments (one at zero water flux and one at zero electric flux) should be performed, for which the conductivity, electric and hydraulic gradients should be measured. One then can find the hydraulic permeability k.

Lars Onsager and the reciprocal relations

Lars Onsager (1903 –1976) was a Norwegian-born American physical chemist and theoretical physicist. He held the Gibbs Professorship of Theoretical Chemistry at Yale University. He was awarded the Nobel Prize in Chemistry in 1968 for his work on the **reciprocal relations** he derived. After completing secondary school in Oslo, he attended the Norwegian Institute of Technology (NTH) in Trondheim (nowadays the NTNU – *Norges Teknisk-Naturvitenskapelige Universitet*, i.e. Norwegian University of Science and Technology), graduating as a chemical engineer in 1925.
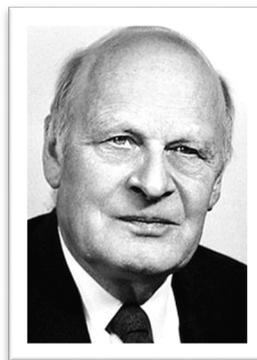
In 1925 he arrived at a correction to the Debye-Hückel theory of electrolytic solutions, to specify Brownian movement of ions in solution, and during 1926 published it. He travelled to Zürich, where **Peter Debye** was teaching, and told him his theory was wrong. He impressed Debye so much that he was invited to become Debye's assistant at the *Eidgenössische Technische Hochschule* (ETH), where he remained until 1928.

In 1928 he went to the United States and held a position at the Johns Hopkins University (JHU) in Baltimore and then at the Brown University in Providence. While clearly Onsager was extremely good at developing theories in physical chemistry, he was a very poor teacher. This is why he was dismissed at JHU after one semester and in 1933 at Brown. After a trip to Europe he was hired by Yale University, where he remained for most of the rest of his life, retiring in 1972. The only graduate student who could really understand his lectures on electrolyte systems, Raymond Fuoss, worked under him and eventually joined him on the Yale chemistry faculty. The 1932 paper that they wrote together on irreversible processes in electrolytes took up eighty-nine pages in the Journal of Physical Chemistry and remained the definitive treatment of the topic until it was taken up again by the two of them in the 1950s.
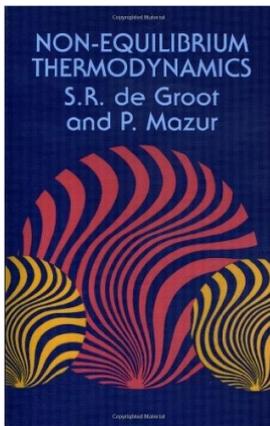
a classic book about irreversible thermodynamics[130]

The reciprocal relations of Onsager are derived in the frame of **irreversible thermodynamics**. First, relations between fluxes and forces are expressed by

---

[130] De Groot, Sybren Ruurds, and Peter Mazur. Non-equilibrium thermodynamics. Courier Corporation, 2013

considering the **entropy production.** Then the reciprocal relations can be given. These relations are a consequence of the **time reversibility of microscopic dynamics**. The idea of microreversibility was introduced when the kinetics of gases was studied. In 1872, Ludwig Boltzmann represented kinetics of gases as statistical ensemble of elementary collisions. Equations of mechanics are reversible in time, hence, the reverse collisions obey the same laws. According to Boltzmann, this microreversibility implies the principle of detailed balance for collisions: at the equilibrium ensemble each collision is equilibrated by its reverse collision. Another macroscopic consequence of microscopic reversibility is the symmetry of kinetic coefficients, the so-called reciprocal relations. The reciprocal relations were discovered in the 19th century by Thomson and Helmholtz for some phenomena but the general theory was proposed by Lars Onsager in 1931. He found also the connection between the reciprocal relations and detailed balance.

Illustrations

Lars Onsager
https://www.nobelprize.org/prizes/chemistry/1968/onsager/biographical/

Chapter 10

Modelling the consolidation of slurries

In this chapter, we are going to discuss some solutions of the Gibson equation for slurries that was set-up in **Chapter 9**. The Gibson equation reads:

$$\frac{\partial \phi_S}{\partial t} = \frac{\partial}{\partial z}\left[\frac{k}{\eta}\left(\phi_S \frac{\partial \sigma_{sk}}{\partial z} + (\rho_s - \rho_w)g\phi_s^2\right)\right]$$

We are first going to give analytical solutions, and the solutions will be critically analysed at the end of the chapter. We will then also make the link between **Chapter 8** (settling) and the consolidation processes discussed in **Chapter 9** and in the present chapter.

Consolidation of slurries : the fractal approach

The permeability k (and K) and the skeleton stress σ_{sk} depend both on ϕ_S. In order to solve the Gibson equation it is therefore primordial to first find a constitutive relation for both K and σ_{sk} as function of ϕ_S. An innovative approach has been introduced by Merckelbach and Kranenburg in 2000[131] to model the consolidation of a column filled with (flocculated) clay. They make use of the self-similarity properties of fractals, and hence assume that the column is filled with fractal flocs.

In their approach, at the initial time t = 0, the column is filled with fractal flocs that are touching one another and all the flocs have the same diameter. A schematic representation of a column filled by actual ("real") fractal flocs is given here for times larger than zero:
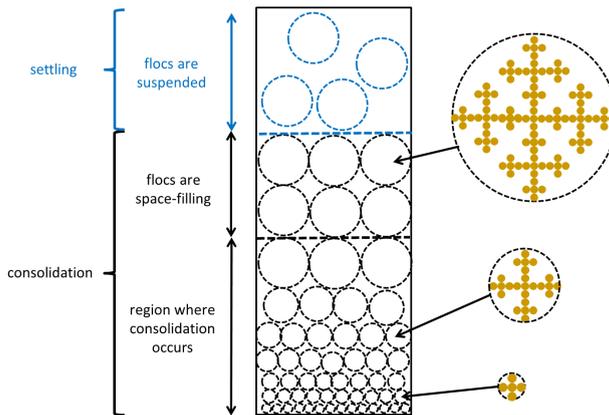
Schematic representation of the column filled with fractal flocs; the flocs are "real" flocs that are physically present in the column, with a defined (measurable) size and fractal dimension

---

[131] "Consolidation and strength evolution of soft mud layers", Lucas Merckelbach, PhD thesis TU Delft, 2000

On purpose, we have represented in the column a suspended phase: in the region the flocs are suspended, they are, by definition, not touching and hence will experience settling. Their settling velocity is in general different from the velocity of the flocs that are in the bed : a settling floc "has no clue" there is a bed underneath it, in other words: the settling velocity of a floc is independent of the settling velocity of the bed composed of the same flocs. Settling and consolidation should therefore in general be treated as two separate processes.

There is a special case, however, where settling and consolidation can be treated with a single model, which is the case we present here.

If we assume that the sediment concentration in the column is high, and homogeneous at t = 0, it is possible to *mathematically* define a fractal floc size, even when the sediment is *not* flocculated or the flocs (if any) are *not* fractal. In the case the sediment is flocculated, the *mathematical floc size* will always be larger or equal to the *actual floc size* and the fractal dimension of the mathematical floc size will not necessarily be the fractal dimension of the actual floc (if the actual floc is fractal). The reason for this is linked to an important requirement of the model: the *mathematical* flocs should be space-filling – also in the settling region. The proper illustration for the model is therefore the following:
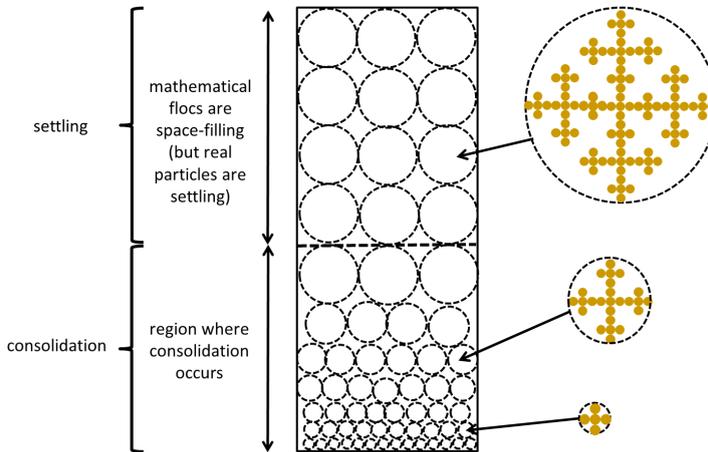
Schematic representation of the column filled by mathematical fractal flocs; contrary to the previous illustration, the flocs are here only mathematical objects. The real particles inside the column are not (necessarily) flocs or fractal and their size is usually smaller than the size of the mathematical fractal flocs, as the real particles do not have to be space-filling.

The advantage of this approach is that there is no mathematical discontinuity between the settling and the consolidation phases (the same parameters, i.e. the same permeability and effective stress are used for both phases). We will see with examples at the end of the chapter the limitations of this approach.

As we assume that at a position z in the column (also for times larger than zero) the whole slice is filled by flocs, we can say that the volume fraction of clay at height z is in good approximation equal to the fraction of clay within a floc:

$$\phi_{clay} = \frac{\text{volume clay in a floc}}{\text{volume of a floc}}$$

In **Chapter 5** we have introduced fractals. We have said that the number of primary particles inside the floc is given by:

$$N = \left(\frac{R_N}{a}\right)^D$$

where D is the **fractal dimension**, a is the size of a primary particle and R_N the size of the fractal floc. In the case that the primary particles are not spheres and have a radius of gyration R_p, the exact size one has to take for a primary particle is often unknown:
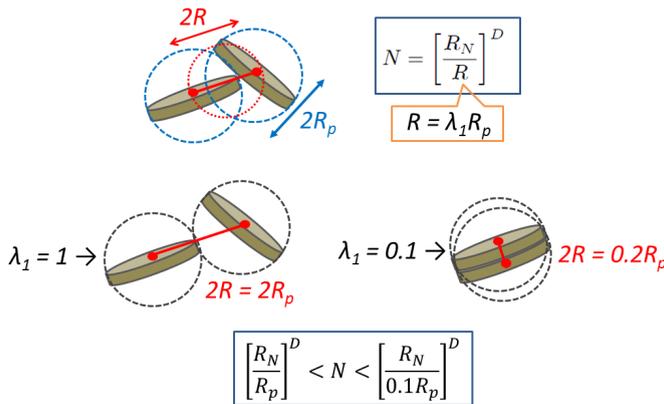
Depending on the way clay platelets aggregate, the characteristic size R of a primary particle (that can be approximated by the centre-to-centre distance between two platelets) can be quite different from the radius of gyration R_p of the platelet. An unknown factor λ_1 that ranges between a very small number and one can be introduced to reflect the fact that the number of primary particles in a floc is unknown:

$$N = \left(\frac{R_N}{\lambda_1 R_p}\right)^D$$

Similarly, the amount of clay contained in the radius of gyration R_p is depending on the shape of the clay particle, which leads to the introduction of a parameter λ_2:

$$\text{volume clay in a floc} = \lambda_2 \frac{4}{3}\pi R_p^3$$

One therefore obtains:

$$\phi_{clay} = \frac{\lambda_2 N R_p^3}{R_N^3} = \frac{\lambda_2}{\lambda_1^D}\left(\frac{R_p}{R_N}\right)^{3-D}$$

From this last equation, we find a relation between the volume fraction of clay and the size of a floc R_N. As on the right-hand side, only R_N is a function of z, we get:

$$\phi_{clay}(z,t) \sim \left(\frac{R_N(z,t)}{R_p}\right)^{D-3}$$

This small digression about non-spherical primary particles is only relevant for estimating the clay volume fraction of "real" flocs. For the "mathematical" flocs, it is simpler to assume the primary particles are spherical. The shape and size of flocs is irrelevant for the model.

A small discussion about the fractal dimension of "mathematical" flocs

Let us recapitulate briefly the way we connected the volume fraction ϕ_{clay} to the fractal dimension D (for simplicity we assume that there are no other particles present and hence $\phi_{clay} = \phi_s$):

The clay volume fraction $\phi_{\text{clay in a floc}}$ of a floc made of clay primary particles is equal to the number of primary particles times the volume of a primary particle divided by the volume of a floc:

$$\phi_{\text{clay in a floc}} \sim \frac{N \cdot R_p^3}{R_N^3}$$

As we have made the assumption that the flocs are space-filling, $\phi_{\text{clay in a floc}}$ is also the volume fraction of a slice of the column at a given height z:

$$\phi_{\text{clay in a floc}}(z) = \phi_s(z)$$

As, by definition of a fractal,

$$N \sim \left(\frac{R_N}{R_p}\right)^D$$

one gets

$$\phi_s \sim \left(\frac{R_p}{R_N}\right)^{3-D}$$

which is the relation we have been using so far.

There is, however, a little subtlety: we have attributed the value $\phi_s(z)$ to a slice of thickness dz of the column. This slice dz should of course be small for mathematical reasons, but as we have explicitly taken the fractal floc to be spherical, the slice should be larger than the radius R_N for all positions and all times, otherwise $\phi_s(z)$ would vary unrealistically:
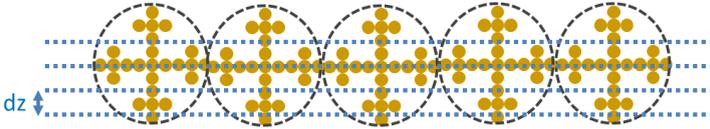
Improper thickness dz of a slice: as $dz < R_N$ the volume fraction $\phi_s(z)$ would vary from a low, high and again low number for three consecutive dz

In fact, as one assumes the volume fraction $\phi_s(z)$ to be constant for a small slice dz, one should preferably want to assume that the considered flocs are isotropic in that direction, i.e. implying that they should not be spherical but cylindrical, with translational symmetry along z for each thickness dz. This would put no restriction on how small dz could be. If we take such cylindrical flocs, with their base being a 2D fractal, one would get for the clay volume fraction:

$$\phi_{\text{clay in a floc}} \sim \frac{N \cdot R_p^2 \cdot dz}{R_N^2 \cdot dz}$$

Along the same derivation lines as above, this would give:

$$\phi_s \sim \left(\frac{R_p}{R_N}\right)^{2-D}$$

For primary particles that are space-filling, the fractal dimension would therefore be 2 in this case and not 3. A fractal dimension of 1 corresponds also to the case where the primary particles are space-filling, i.e. occupying the whole given line, curve or volume considered. In fact all fractal dimensions that are integers (1, 2 or 3) correspond to "primary particles that are space-filling":

Let us imagine that we have primary particles such that

$$N \sim \left(\frac{R_N}{R_p}\right)^1$$

This also implies that

$$N^D \sim \left(\frac{R_N}{R_p}\right)^D$$

If D = 2 or 3, this means that we have completely filled the area or volume $(R_N)^D$ with primary particles (which are therefore "space-filling"). If the dimension D is not an integer, then $(R_N)^D$ does not represent a curve, an area or a volume anymore, but simply relates the amount of primary particles to the size R_N:

$$R_N = N^{1/D} \cdot R_p$$

The important conclusion here is that the fractal dimension is simply a convenient tool to model the system: one may chose the mathematical flocs to be spherical or cylindrical. In the end, after fitting the data, one should find the same values for the fractal dimension with a 1 difference: if one finds, for example D = 2,63 with the spherical approach, one will find D = 1,63 with the cylindrical approach. A value of D = 2 (cylindrical approach) or D = 3 (spherical approach) implies that the primary particles are space-filling.

In order to be consistent with the work of Merckelbach, we will continue here to use the following definition:

$$\phi_{clay}(z,t) \sim \left(\frac{R_N(z,t)}{R_p}\right)^{D-3}$$

At time t = 0, we have $\phi_{clay} = \phi_0$ and therefore

$$\phi_0 \sim \left(\frac{R_N(z,t=0)}{R_p}\right)^{D-3}$$

In the limiting case that D = 3, one gets

$$\phi_0 = 1$$

and no fractal approach can be used, as there is no proper way to define the volume of a floc. In the case $\phi_0 = 1$, there is no settling and no consolidation, as all the primary particles (of radius R_p) are touching, and by hypothesis, cannot deform. The case $\phi_0 = 1$ is corresponding to a fully gelled state of primary particles. Depending on the shape of particles ϕ_0 will be smaller than 1 in most cases ($\phi_0 = 0.74$ for compacted spheres in their most compacted form).

If the characteristic size R_p of the primary particle can be evaluated (by static light scattering for instance), it is then possible to have an estimation of the mathematical fractal floc size at t = 0:

$$R_N \sim R_p (\phi_0)^{\frac{1}{D-3}}$$

If we assume that the fractal dimension is given by D = 2.63 (which is a reasonable value, close to what is generally found in experiments), we get:
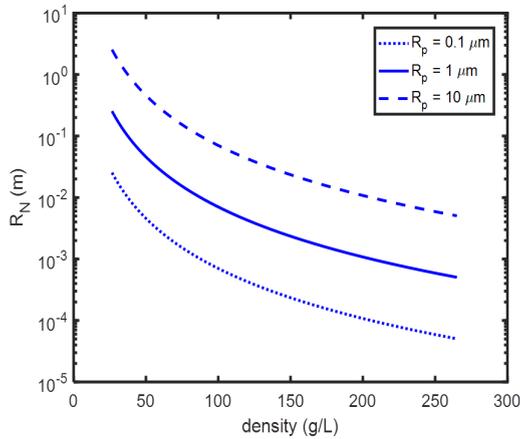
This implies that for an initial concentration of 150 g/L ($\phi_0 = 150/2650 = 0.0566$), having an homogeneous suspension of particles of size 0.1 microns would give an effective fractal floc radius of 0.23 mm whereas 10 microns particles would give 2.35 cm. For lower volume fractions, the floc radius becomes clearly unrealistic. For 50 g/L for example and 10 microns particles, one would get a radius of 45 cm, much larger than the settling column radius!

From above, we deduce that the fractal approach will only be appropriate for high concentrations of clay, as one of the assumption made in the model (see underneath) is that the size of a pore scales as the size of a floc.

Note that if the density of the suspension is known and equal to ρ and that we assume that the flocs are filling all the space, one finds, from the definition $\rho = \rho_s \phi_0 + \rho_w (1 - \phi_0)$ that:

$$\phi_0 = \frac{\rho - \rho_w}{\rho_s - \rho_w}$$

which is a convenient way to estimate the initial volume fraction.

The hypotheses required for the model are summarized here:

(a) the sample should be homogeneous at t = 0. This implies that if the sample is polydisperse and contains both silt and clay (for simplicity we assume here there is no sand fraction – which anyhow would settle quite quickly at the bottom of the column), that at any height in the column the ratio ϕ_{silt}/ϕ_{clay} does not depend on z.

(b) The portion of the column under consideration is filled by *mathematical* fractal flocs (so, no real flocs, see discussion above) that are touching one another. These flocs are made of clay particles. If silt particles are present, they act as filling material for the flocs:

at z_1 : N = $N(z_1)$ at z_2 : N = $N(z_2)$
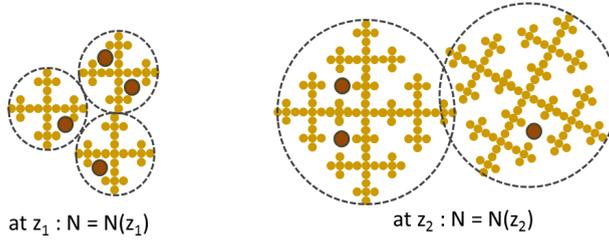
Schematic representation of flocs at two different heights in the columns. The flocs are made of clay particles (yellow) and some silt particles are imbedded inside the flocs (brown)

(c) at a same height, all flocs settle with the same velocity. As we also consider that all flocs have the same fractal dimension and are made of the same primary particles, this implies that all the flocs at a given height have the same size.

The volume fraction of clay is linked to the total solid volume fraction by:

$$\phi_s(z) = \phi_{clay}(z) + \phi_{silt}(z)$$

By dividing this relation by ϕ_{clay} one gets:

$$\frac{\phi_s(z)}{\phi_{clay}(z)} = 1 + \frac{\phi_{silt}(z)}{\phi_{clay}(z)}$$

As we have made the hypothesis that ϕ_{silt}/ϕ_{clay} does not depend on z, we deduce that the ratio ϕ_s/ϕ_{clay} does neither. This implies that at any given height the solid volume fraction is only proportional to the clay fraction:

$$\phi_s(z,t) = \left(1 + \frac{\phi_{silt}}{\phi_{clay}}\right)\phi_{clay}(z,t)$$

Relation for $K(\phi_S)$
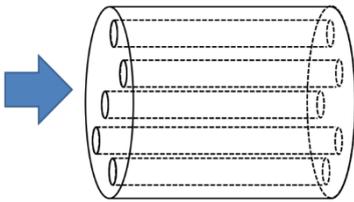
In order to find the dependence of the permeability on z we will use a dimension analysis. If one calculates the permeability of a system of pores as sketched opposite, one can show that the permeability is proportional to[132]:

$$K \sim (1 - \phi_S)r^2$$

where r is the size of a pore.

Using this equivalence, we obtain:

$$K(z,t) \sim \big(R_N(z,t)\big)^2$$

where we have used the assumptions that:

(a) $\phi_S \ll 1$. This hypothesis is certainly true at the beginning of consolidation, as we have seen that, to give an order of magnitude, an initial concentration of 150 g/L gives $\phi_0 = 150/2650 = 0.0566 \ll 1$. the hypothesis should however be checked at the end of consolidation, for the deepest layers in the column.

(b) The size of the largest connecting pores scales as R_N:
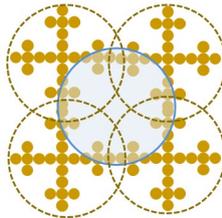
the blue circle is the middle (connecting pore) has the same size as the flocs (dashed circles)

This is a consequence of the scale invariance (a property of fractals). The fractal approach is therefore a convenient and mathematically elegant way to describe the change of volume fraction with position z and time t. As both R_p and D are assumed be constant, *mathematically* if ϕ_s is increasing, R_N is decreasing. This does not mean that the "real" flocs (if any) in the system are becoming smaller: physically, what is

---

[132] The equation that gives $K_{single} \sim r^2$ for one pore is the **Poiseuille equation**. For a bundle of pores as sketched here, one has: $K = (1 - \phi_S)K_{single}$ from simple geometrical consideration.

expected to happen is that real flocs becomes squeezed and their shape (and density) is changing. If there are no flocs but only suspended particles in the column what happens is that the distance between particles is decreasing until they are touching one another.
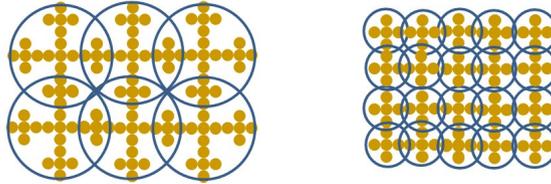
The fractal approach: the "flocs" are assumed to be space-filling, and the clay volume fraction is increasing when the size of the flocs is decreasing. Note that the size of the pores (voids between primary particles) indeed scales as R_N.

Linking R_N to ϕ_s, the permeability is defined by:

$$K(z,t) = K_k[\phi_s(z,t)]^{-n} \quad \text{with} \quad n = \frac{2}{3-D}$$

where K_k is a factor not depending on z or t.

Relation for $\sigma_{sk}(\phi_S)$

We now would like to find a similar type of relation for σ_{sk}. First we will define the isotropic stress in the skeleton as:

$$\sigma_{sk}^0 = F(n_{tot} - n_i)$$

where F is the force of one bond, n_{tot} is the total number of bonds per unit area:

$$n_{tot} = n_a + n_i$$

where n_a is the number of bonds per unit area on which a stress is exerted ("active bond") and n_i the number of bonds per unit area on which no stress is exerted ("inactive bond").
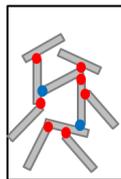
n_a = active bonds (with stress): ●

n_i = inactive bonds (no stress): ●

In the fractal approach we take here, with touching fractal flocs, the total number of bonds per unit area is given by:

$$n_{tot}(z,t) = \frac{n_0}{\left(R_N(z,t)\right)^2}$$

where n_0 is the number of bonds per floc: this is invariant for any floc size (scale invariance is a property of fractals):

whatever floc size:
n_0 = constant

$n_0 = 4$ $n_0 = 4$ $n_0 = 4$
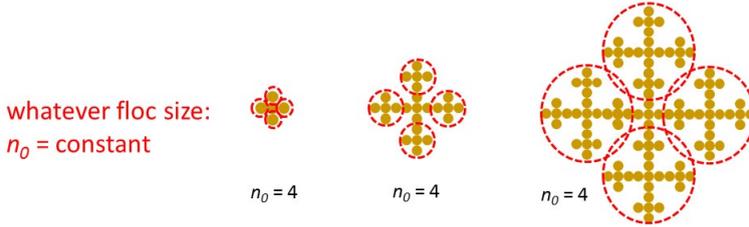
the number of bonds (in this case 4) does not change with the size of flocs

We deduce that:

$$\sigma_{sk}^0(z,t) = F\left(\frac{n_0}{(R_N(z,t))^2} - n_i\right)$$

Note that F, n_0 and n_i do not depend on z. From the dependence of R_N on ϕ_s we get:

$$\sigma_{sk}^0(z,t) \approx A[\phi_s(z,t)]^n - Fn_i \quad \text{with} \quad n = \frac{2}{3-D}$$

where *A* is a factor that does not depend on *z*. The total effective stress σ_{sk}^0 is related to the vertical effective (skeleton) stress σ_{sk} and the horizontal one σ_{sk}^h by:

$$\sigma_{sk}^0 = \frac{1}{3}\sigma_{sk} + \frac{2}{3}\sigma_{sk}^h$$

This relation can be re-written:

$$\sigma_{sk}^0 = \frac{1 + 2\sigma_{sk}^h/\sigma_{sk}}{3}\sigma_{sk}$$

Assuming that $\sigma_{sk}^h/\sigma_{sk}$ does not depend on ϕ_s the skeleton stress is defined as:

$$\sigma_{sk}(z,t) = K_\sigma[\phi_s(z,t)]^n - K_{\sigma 0}$$

where K_σ and $K_{\sigma 0}$ are factors that do not depend on z or t.

We have now found two constitutive equations $K\big(\phi_s(z,t)\big)$ and $\sigma_{sk}\big(\phi_s(z,t)\big)$ which enables us to solve (analytically) the Gibson equation. We will do this for three different cases. The first and second solutions given underneath correspond to the primary consolidation regime defined in **Chapter 9**. The first solution is given for the case there is both a settling and a consolidation phase whereas the second solution corresponds to the case the settling phase has disappeared and there is only consolidation. The transition between the first and second solution occurs at the gelling time t_1 that will be defined below. The last solution corresponds to the end of consolidation, at equilibrium. For this solution **creep** effects are neglected. Mathematically, this implies that K_σ and $K_{\sigma 0}$ are not time-dependent, i.e. that over time the fabric (the clay skeleton) does not deform. The secondary compression regime, as defined in **Chapter 9** cannot be solved analytically and will be discussed at the end of the chapter, where numerical solutions will be presented.

The setting regime (below the gelling time, t < t$_1$)

At the initial stage of consolidation, we make the hypothesis that the self-weight is almost entirely borne by the pore water, thus:

$$\frac{\partial \sigma_{sk}}{\partial z} \approx 0$$

for all heights. Intuitively, one can already predict that this hypothesis is questionable as time increases for the deepest layers close to the bottom of the column, which we define by the position z = 0. The validity of the assumption and its consequences will be discussed at the end of the chapter. From the previous subsection, we have:

$$K(z,t) = K_k[\phi_s(z,t)]^{-n}$$

The Gibson equation therefore becomes:

$$\frac{\partial \phi_S}{\partial t} = K_k \frac{\rho_s - \rho_w}{\rho_w}(2-n)\phi_s^{1-n}\frac{\partial \phi_s}{\partial z}$$

This equation is a nonlinear convection equation, similar to **Kynch's sedimentation equation**, and can be solved by the **method of characteristics**.

The method of characteristics

The method can be applied to solve equations of the type:

$$\frac{\partial u}{\partial t}a(x,t,u) + \frac{\partial u}{\partial x}b(x,t,u) = c(x,t,u)$$

This method relies on the property that the function $u(x,t)$ can be differentiated with respect to a variable s (not yet defined) as:

$$\frac{du}{ds} = \frac{\partial u}{\partial t}\frac{\partial t}{\partial s} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial s}$$

If we now define the variable s as the curvilinear coordinate along a curve in the (x,t) plane by:

$$\frac{\partial t}{\partial s} = a(x,t,u)$$

$$\frac{\partial x}{\partial s} = b(x,t,u)$$

Then we simply have to solve, along with the 2 previous equations:

$$\frac{du}{ds} = c(x,t,u)$$

In other words: we have transformed an equation with partial differentials in a series of 3 ordinary differential equations that are easier to solve.

In our case, we can apply the method of characteristics using the function ϕ_S instead of u and the variables z,t instead of x,t. We then find that we have to solve:

$$\frac{\partial t}{\partial s} = 1$$

$$\frac{\partial z}{\partial s} = -K_k \frac{\rho_s - \rho_w}{\rho_w}(2-n)\phi_s^{1-n}$$

$$\frac{d\phi_S}{ds} = 0$$

The last equation can be solved into:

$$\phi_S = \text{constant}$$

This implies that the lines of constant ϕ_S are representing the curves associated with the curvilinear coordinate s. These lines are represented by the black lines on the figure underneath.
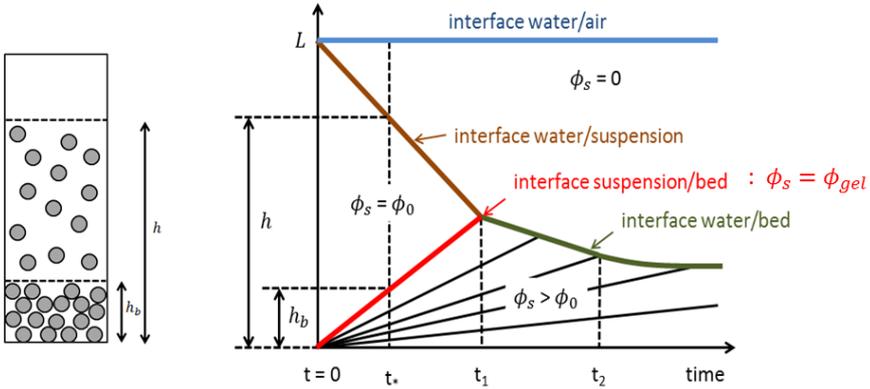
Variation of the interface with time for a suspension that at start is below the gelling concentration. ϕ_S represents the volume fraction of the particles, ϕ_0 is their initial volume fraction and $\phi_{gel} = \phi_0$ is the volume fraction at the suspension/bed interface, below which $\phi_S > \phi_{gel}$. Until the gelling time t_1 there is a suspension/bed interface. Above t_1 only the water/bed interface remains.

Above the suspension/bed line the volume fraction is by definition constant and equal to ϕ_0 so that the lines of constant ϕ_S occupy the whole area below the water/suspension interface.

The variable s can be eliminated by combining the two first equations from which $z(t)$ is obtained:

$$z(t) = K_k \frac{\rho_s - \rho_w}{\rho_w}(n-2)\phi_s^{1-n}t + z(0)$$

Solving this equation for ϕ_S with $z(0) = 0$, one finds for $0 < z \le h_b$:

$$\phi_S(z,t) = \left(K_k \frac{\rho_s - \rho_w}{\rho_w}(n-2)\frac{t}{z}\right)^{\frac{1}{n-1}}$$

An interesting case if for $z = h_b$ for which $\phi_S = \phi_0$. It enables to find the height of the bed with time:

$$h_b(t) = K_k \frac{\rho_s - \rho_w}{\rho_w}(n-2)\phi_0^{1-n}t$$

from which we deduce that:

$$\phi_S(z,t) = \phi_0 \left(\frac{h_b(t)}{z}\right)^{\frac{1}{n-1}}$$

As drawn at the beginning of the chapter, $h_b(t)$ is a linearly increasing function of time and hence v_{bed} is constant as function of time. Using conservation of mass (by evaluating the Gibson height L_G defined in **Chapter 9**) allows to determine $h(t)$, i.e. the water/suspension interface height:

$$L_G = \int_0^h \phi_S dz = \int_0^{h_b} \phi_S(z,t)dz + \int_{h_b}^h \phi_0 dz = \phi_0 H$$

where H is the height of the suspension in the column at $t = 0$. After integration one finds:

$$L_G = \left(K_k \frac{\rho_s - \rho_w}{\rho_w}(n-2)t \right)^{\frac{1}{n-1}} \frac{1-n}{2-n}(h_b)^{\frac{n-2}{n-1}} + \phi_0(h - h_b)$$

Combining the expressions for L_G and h_b one finds:

$$L_G = \left(\frac{1}{n-2} \right)\phi_0 h_b + \phi_0 h$$

$$h(t) = H - K_k \frac{\rho_s - \rho_w}{\rho_w}\phi_0^{1-n}t$$

The gelling time

The time t_1 is called the "gelling time" as the settling phase has now disappeared. At $t = t_1$ one has $h = h_b$ (the interface water/suspension has disappeared) and from the relation found above we get:

$$t_1 = \frac{\rho_w}{\rho_s - \rho_w} \frac{L_G \phi_0^{n-2}}{(n-1)K_k}$$

The height at the gelling time is given by

$$h_b(t_1) = h(t_1) = H \frac{n-2}{n-1}$$

Even though this height is independent on K_k and could be used to determine n, in practice, it is very difficult to estimate. A better way to estimate n is given in the next section.

The primary compression regime (above the gelling time, t > t₁)

When t $> t_1$ the interface water/suspension has disappeared. The density profile is then independent on the initial condition. In that case, we have:

$$L_G = \int_0^h \phi_S(z,t)dz$$

One then finds:

$$L_G = \left(K_k \frac{\rho_s - \rho_w}{\rho_w}(2-n)t\right)^{\frac{1}{n-1}} \frac{1-n}{2-n}(h)^{\frac{n-2}{n-1}}$$

This leads to:

$$h(t) = \left(L_G \frac{2-n}{1-n}\right)^{\frac{1-n}{2-n}} \left(K_k \frac{\rho_s - \rho_w}{\rho_w}(n-2)t\right)^{\frac{1}{2-n}}$$

We have now found that

$$h(t) \sim t^{\frac{1}{2-n}} = t^{\frac{3-D}{4-2D}}$$

The fractal dimension D can therefore be determined from the water/gel interface as function of time. When plotting the data on a double logarithmic scale, one gets $(3-D)/(4-2D)$ from the slope of the obtained line.

The unknown parameter K_k can subsequently be found by fitting the data for any time. In practice it is often found that the values for K_k differ when the data is plotted for $t < t_1$ or for $t > t_1$ as settling and consolidation are unrelated. From K_k the permeability $K(z)$ can be estimated.

Examples

The analytical solution for t < t₂

A typical interface profile is given underneath where we used the found analytical expressions for $h(t)$ and $h_b(t)$ and the following representative parameters:

$\rho_w = 10^3$ kg/m³	$\rho_s = 2600$ kg/m³	H = 1 m	g = 9.81 m²/s
D = 2.63	$\phi_0 = 2.5$ %	$K_k = 1.5 \times 10^{-9}$ m/s	

From the figure, one can see that the analytical solution gives unrealistic values for the interface at long times. In fact, at infinite time the size of the interface will be zero, as can be easily verified from the expression for $h(t)$ given above. At long times, it is not correct to assume that the effects of the skeleton stress can be omitted. This is discussed further in the next example, where we compare the analytical solution to the numerical solution, found by solving the full Gibson equation.

The final size of the bed (when the consolidation has stopped and $\phi_s(z)$ is independent of time) can also be evaluated analytically, as shown in the next section.
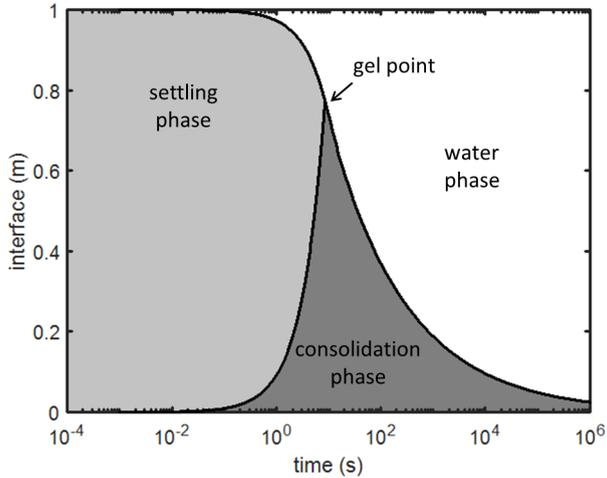
Water / sediment interface with time. In the settling phase the volume fraction at any height is given by the initial volume fraction. Darker region: build-up of the bed-sol interface. Note that we used a semilog scale. The two lines join at t = t₁ (gel point). In this example t₁ = 8.2 s.

Comparison between the analytical and numerical solution of the Gibson equation for t < t₁

The analytical solutions given in the sections above are here compared with the numerical solution of the full Gibson equation. We recall that the analytical solution has been derived under the assumption that $\partial \sigma_{sk} / \partial z \approx 0$.
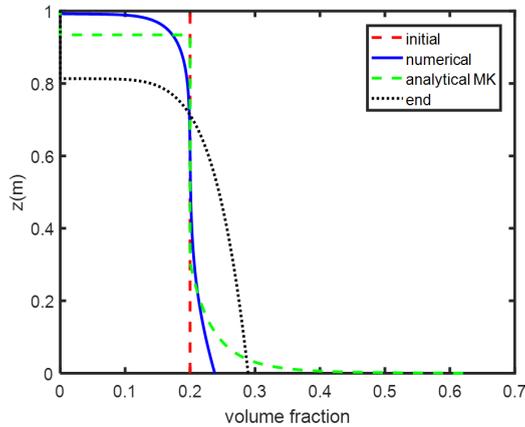
Volume fraction as function of height in the column after 290 days. Dashed red: initial volume fraction; Dotted black: volume fraction end profile; dashed green: analytical solution; blue line: numerical solution. $K_k = 1.74 \times 10^{-13}$ m/s, $K_\sigma = 1.26 \times 10^7$ Pa, D = 2.7, $\phi_0 = 0.2$; other parameters as in the table given above.

The volume fraction profile is for the situation we are below the gel(ling) point. This is confirmed on the figure by the fact that between 0.4 and 0.8 m the volume fraction is still equal to the initial volume fraction. The blue line, which corresponds to the numerical solution of Gibson equation, using both the permeability and effective stress terms, would reduce to the green dashed line, which corresponds to the analytical profile if the effective stress term would be set equal to zero.

Because of the effective stress term, which mathematically corresponds to a diffusion term (the permeability term is comparable to an advection term), the interfaces between water / suspension and suspension / bed are not clearly defined for the numerical solution. One can therefore already conclude that if one wants to use the analytical method detailed above in combination with measurements of interfaces against time this will only be possible for systems in which the effective stress terms are quite small – these kind of systems are for instance suspensions of concentrated hard (unflocculated) particles. Experimentally, as we have seen already in **Chapter 8** (see section "Finding the hindered settling function from the equilibrium profile") the change in volume fraction between 0 and ϕ_0 occurs over a distance of the order of 1 or 2 mm (in the numerical example we showed above, we used a very large K_σ to exaggerate the effect).

The discrepancy between analytical and numerical solution at the bottom of the column clearly demonstrate that even below the gelling point (in the primary consolidation regime) the assumption

$$\frac{\partial \sigma_{sk}}{\partial z} \approx 0$$

is not valid. In the absence of diffusion, there is no constraint to prevent all the particles reaching the bottom of the column as we have implicitly treated the particles as having no volume (point-like objects). Mathematically, this is symbolised by the fact that the analytical solution (in the absence of effective stress – i.e. a diffusion term), scales as

$$\phi_S(z,t) \sim \left(\frac{1}{z}\right)^{\frac{1}{n-1}}$$

At the bottom of the column one gets $\phi_S(z = 0, t) \to \infty$. This unphysical result implies that the water/suspension interface, which is calculated from mass conservation through the evaluation of the Gibson height, will become lower than the actual interface at longer times. In the case of the settling of hard particles, where only advection (settling) play a role, it is numerically easy to limit the filling of the settling column by imposing that each slice of the column should not exceed a given volume fraction, i.e. the maximum packing volume fraction of particles.

In conclusion, the fractal approach method explained in this chapter can be applied for $t < t_2$ in the case of concentrated hard particles that are settling without significant diffusion. From the procedures explained in the previous sections for below and just above the gelling time, K_k and the fractal parameter n dimension can be obtained from which the permeability can be deduced.

As discussed in the introduction, these fractal parameter n and K_k parameter are however not expected to be the same above and below t_1 as in the model we have made the hypothesis that the flocs are space-filling both in the settling and in the consolidation phase. In fact, the space-filling criterion is in most cases expected to be valid above t_1 and not below.

In a later section, we will show how this cross-over between settling and consolidation is found in experiments, and how we can model it.

Determination of the effective stress : end of consolidation (t > t$_2$)

From the results of this subsection, we will get an estimated of the final bed height when consolidation is finished.

In the final stage of consolidation, the structure is bearing and we can estimate that

$$v_{s/lab} \approx 0$$

From the modified Darcy equation (see **Chapter 9**), we had:

$$v_{s/lab} = \frac{k}{\eta} \frac{\partial P_e}{\partial z} = \frac{K}{g\rho_w}(\rho_s - \rho_w)g\phi_S + \frac{K}{g\rho_w}\frac{\partial \sigma_{sk}}{\partial z}$$

Which now gives:

$$\phi_S + \frac{1}{(\rho_s - \rho_w)g}\frac{\partial \sigma_{sk}}{\partial z} \approx 0$$

Using the equation

$$\sigma_{sk} = K_\sigma[\phi_s(z)]^n - K_{\sigma 0}$$

We substitute ϕ_s and find:

$$\phi_S + \frac{n}{(\rho_s - \rho_w)g}K_\sigma[\phi_s]^{n-1}\frac{\partial \phi_s}{\partial z} = 0$$

This yields

$$\frac{\partial[\phi_s]^{n-1}}{\partial z} = \frac{-(\rho_s - \rho_w)g}{nK_\sigma}(n-1)$$

At long times, the bed/water interface has reached a constant value ($z = h_\infty$) and the final volume fraction profile can be estimated to be, using $\sigma_{sk}(h_\infty) = 0$:

$$\phi_s(h_\infty) = \left(\frac{K_{\sigma 0}}{K_\sigma}\right)^{\frac{1}{n}}$$

We will assume that creep effects are neglected, hence $K_{\sigma 0} = 0$ and

$$\phi_s(h_\infty) = 0$$

We can now evaluate:

$$\int_z^{h_\infty} \frac{\partial[\phi_s]^{n-1}}{\partial z} dz = \int_z^{h_\infty} \frac{-(\rho_s - \rho_w)g}{nK_\sigma}(n-1)dz$$

which gives:

$$\phi_s(t \to \infty, z) = \left[\frac{(\rho_s - \rho_w)g}{nK_\sigma}(n-1)(h_\infty - z)\right]^{\frac{1}{n-1}}$$

$$\log[\phi_s(t \to \infty, z)] = \frac{1}{n-1}\left[\log[(h_\infty - z)] + \log\left[\frac{(\rho_s - \rho_w)g}{nK_\sigma}(n-1)\right]\right]$$

By plotting the particle volume fraction, obtained from a measured density profile in the final stage of consolidation, versus the distance below the interface $(h_\infty - z)$ on a double logarithmic plot the fractal dimension (via n) can be derived from the slope. Subsequently, the parameter K_σ can be obtained from any value of z. This method can lead to very large estimation errors however.

The final bed height can be obtained by using once more the Gibson height:

$$L_G = \phi_0 H = \int_0^{h_\infty} \phi_S(t \to \infty, z)dz$$

which gives:

$$h_\infty = \frac{n}{(n-1)}\phi_0 H \left[\frac{K_\sigma}{(\rho_s - \rho_w)g\phi_0 H}\right]^{\frac{1}{n}}$$

The volume fraction at the bottom of the settling column is given by

$$\phi_s(t \to \infty, z = 0) = \frac{n}{(n-1)}\phi_0 \frac{H}{h_\infty}$$

which also enables to find the fractal dimension in a simple way. The analytical expression is found to correctly predict the evolution of density as function of height, as shown in the following example:
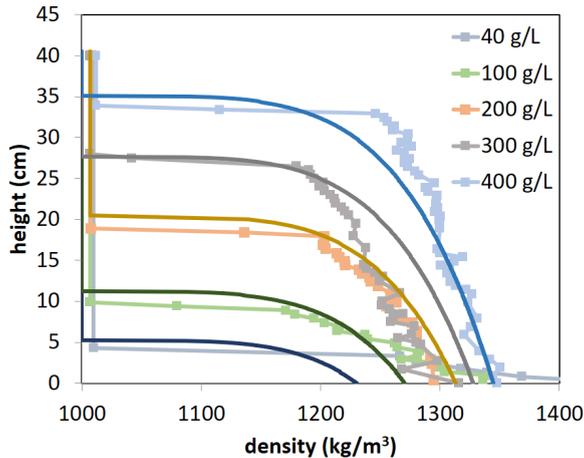
Final density profile of settled beds. Symbols: density measurements. Only the data for the bed of initial concentration 200 g/L (ϕ_0 = 0.077) has been fitted which gave the values $K_\sigma = 6.5 \times 10^6$ Pa and D = 2.65. All the other lines are plotted using these parameters.

The sediment used is from lake Markermeer[133]. The density measurements were done by Ultra-sonic High Concentration Meter (UHCM). We recall that the density is related to the volume fraction ϕ_s by: $\rho = \rho_s \phi_s + (1 - \phi_s)\rho_w$. By fitting one dataset, the one corresponding to the initial concentration of 200 g/L, it was possible to predict the bed height for all other initial concentrations.

It remains to be seen if the fractal dimension found by fitting the "end of consolidation" data is the same as the one found by fitting the data above t_1 . We will show that this is the case with the data and protocol we present.

Settling and consolidation of natural mud

 In the "end of consolidation" section given above, we have shown the fit of a density profile for a mud suspension of concentration 200 g/L, from which it was found that $K_\sigma = 6.5 \times 10^6$ Pa and D = 2.65. The parameter $K_k = 2.0 \cdot 10^{-12}$ m/s is found by fitting the time evolution of the interface for the 200 g/L mud sample using the full Gibson equation, keeping K_σ and D constant and equal to the values found from fitting the density profile. We give here the theoretical evolution of density and excess pore water pressure for this sample:
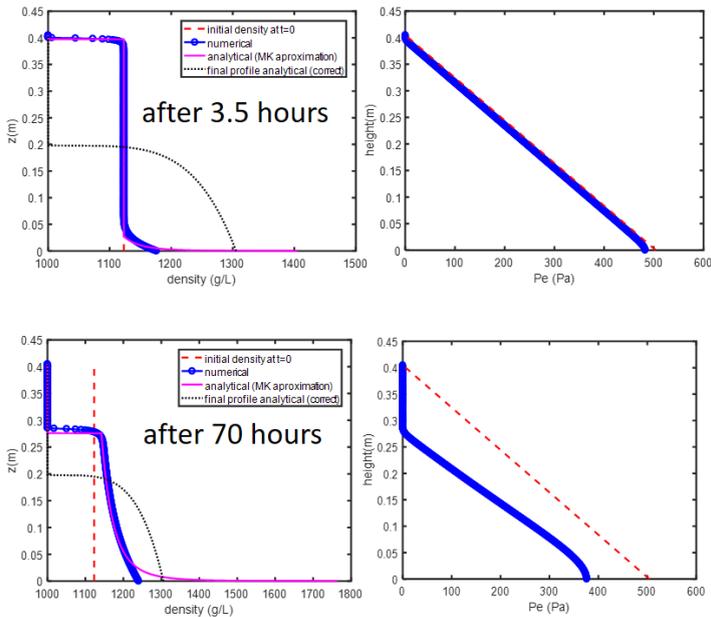
[133] van den Bosch, B.A.P. The effect of initial concentration on the consolidation behaviour of mud: A study on lake Markermeer sediment, Master thesis, TU Delft, 2016
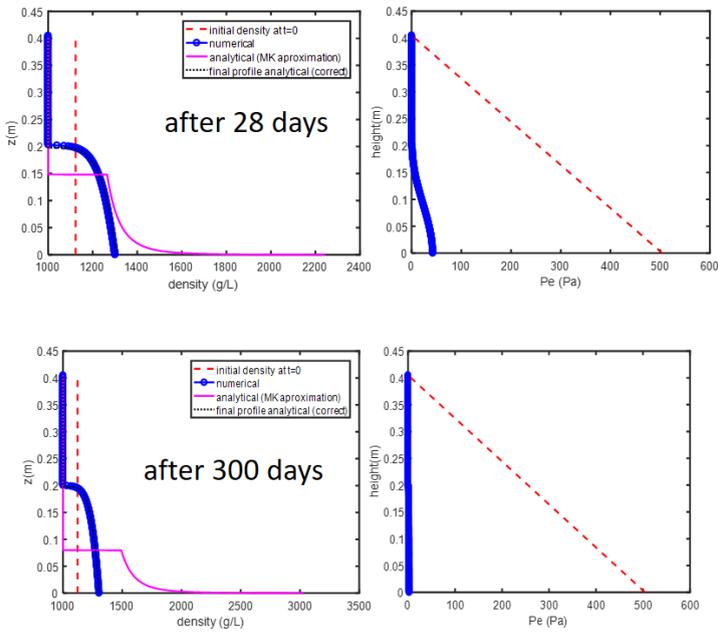
The full numerical solution is represented in blue. The analytical solution of Merkelbach and Kranenburg (MK) is in pink. In contrast to the example given earlier in this chapter, the analytical and numerical solutions agree with each other below the gelling point (in the period before 70 hours), as the K_σ is ten times smaller than in the previous example. At the end of consolidation the numerical solution is not to be distinguished from the exact analytical solution (dashed black line) and the excess pore water pressure is zero as $P_w = P_{hyd}$. The initial pore water pressure (red dashed line) is obtained by realizing that $\partial \sigma_{sk} / \partial z \sim 0$, which gives:

$$P_e(t=0) = (\rho_s - \rho_w) g \phi_s (H - z)$$

where H is the height of fluid in the column. The theoretical time evolution of the water/suspension and suspension/bed for 40 g/L and 200 g/L is given here:
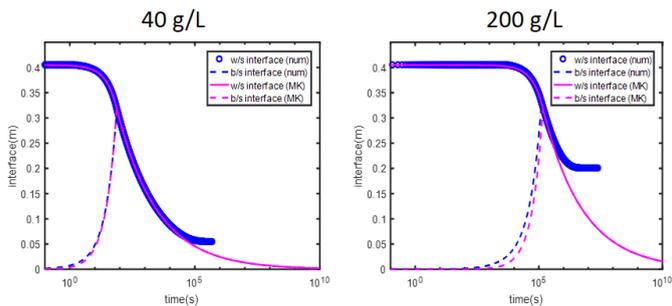
There is a good agreement between the numerical and analytical profiles until 10^4 s, but, as expected, the analytical solution becomes incorrect when diffusion (effective stress) plays a significant role. This happens much before the end of consolidation for highly concentrated mud suspensions.

In contrast to these theoretical profiles, when recording the time evolution of the interfaces of real suspensions, an observable transition is expected between the settling and the consolidation phases, in particular for the most diluted suspension (40 g/L). This is indeed what has been found, see figure below. For concentrations less than 200 g/L, there is a discontinuity in the profile at the gel point. The theoretical predictions, using the same parameters K_k, K_σ and D are given in full lines. For samples of higher concentrations (200 g/L, 300 g/L and 400 g/L) one can observe that data and predictions are in very good agreement.

We checked that the procedure described at the beginning of this chapter, i.e. using the analytical approximations just above t_1, works on these samples and that the numerical solution of the full Gibson equation reduces to the analytical solution in its range of validity (above t_1 but well below t_2).
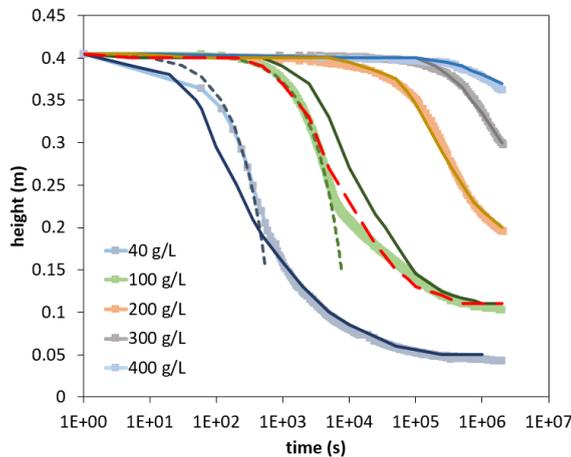
Time evolution of water/sediment interface for different sediment concentrations. Symbols are measurements. Lines are fits (see text). Full lines: $K_\sigma = 6.5 \cdot 10^6$ Pa, $K_k = 2.0 \cdot 10^{-12}$ m/s and $D = 2.65$. The sediment is from lake Markermeer.

For lower concentrations, the flocs are not interconnected at start (reflected in the change in slope around 10^3 s for the 40 g/L sample). The settling phase can then be fitted (dashed lines) using

$$h = h_0 - t \cdot \left(1 - \phi_{floc}\right)^m v_{Stokes}$$

where $h(m)$ is the position of the water / suspension, $h_0(m)$ the initial height. We used $v_{Stokes} = 1.5$ mm/s which is the Stokes settling velocity for 40 μm particles, in agreement with the fact that the clay was sieved through a 63 μm sieve prior dispersion. From the fit we found $\phi_{gel} = 0.075$ (195 g/L), which agrees with the fact that the 200 g/L sample is in a gel state.

Using the same ϕ_{gel} and v_{Stokes}, the settling phase of all samples below gelling concentrations (40 g/L and 100 g/L) could reasonably be predicted. For the lowest clay concentration (40 g/L) there is a good transition between the settling and the consolidation model. This is less the case for the 100 g/L sample. In order to improve the fit, it was necessary to double the value of K_k (red dashed line). The reason of the mismatch could be linked to the limitation of the model in the suspension/gel transition range or due to the fact that the 100 g/L sample was not well-mixed at the onset of the experiment and that a sediment layer was already present at the bottom of the column. To check this last hypothesis, one should then know the density profile at start (not known for the present set of experiments).

Link with Chapter 8 : how to couple settling and consolidation?

In **Chapter 8**, we found for settling that the velocity of the suspension/bed interface is given by

$$v_{bed} = v_0 \frac{\phi_0 f(\phi_0)}{\phi_m - \phi_0}$$

where ϕ_m is the solid volume fraction in the bed. Using the fractal approach we find that for the early stage of consolidation

$$v_{bed} = \frac{dh_b}{dt} = K_k \frac{\rho_s - \rho_w}{\rho_w}(n-2)\phi_0^{1-n}$$

for the settling of flocs. Depending on the sample and the interest of the researcher, one or the other of these formulations are used. Both formulations depend on a-priori unknown parameters: $f(\phi_0)$ or K_k and n, and both are decreasing functions of ϕ_0. These parameters can be found by fitting the function $v_{bed}(\phi_0)$.

In **Chapter 8**, we found for the settling phase that the velocity of the water/suspension interface is given by

$$v_s = f(\phi_0)v_0$$

We here find that (in absolute values):

$$v_s = \frac{dh}{dt} = K_k \frac{\rho_s - \rho_w}{\rho_w}\phi_0^{1-n}$$

By comparing the expressions for v_s and v_{bed} found in **Chapter 8** and the fractal approach used here, we find the following equivalence:

$$n = \frac{2}{3-D} = 2 + \frac{\phi_0}{\phi_m - \phi_0}$$

In **Chapter 8** we only considered a settling column filled with hard spheres. This means that in the bed, these hard spheres are in contact (they do not consolidate) and hence the volume fraction ϕ_m is a constant. In the fractal approach, there is, on the contrary, a smooth variation in volume fraction ϕ_s in the bed as function of z and t. One can estimate the average volume fraction in the bed from:

$$\phi_m = \frac{1}{h_b} \int_0^{h_b} \phi_s dz$$

which yields:

$$\phi_m = \phi_0 \frac{n-1}{n-2}$$

(note that ϕ_m does not depend on time). Inserting this equation in the relation above gives the consistent relation $n = n$.

In the early stage of consolidation, in the case the volume fractions at the bottom of the column remain realistic, one can make the assumption that

$$\frac{\partial \sigma_{sk}}{\partial z} \ll \frac{\partial P_e}{\partial z}$$

and this leads to

$$(1 - \phi_s)v_{w/s} = k \frac{g\rho_w}{\eta} \frac{(\rho_s - \rho_w)}{\rho_w} \phi_s$$

We recall that $v_{s/lab}$ is the velocity of the settling particles measured in the frame of the laboratory (see **Chapter 9**) and that

$$(1 - \phi_s)v_{w/s} = -v_{s/lab}$$

$$K(m/s) = k(m^2) \frac{g\rho_w}{\eta}$$

The settling velocity $v_{s/lab}$ is pointing downwards in the z direction, and therefore a minus sign appears while doing the substitution. We here gave the absolute value of $v_{s/lab}$:

$$v_{s/lab} = K \frac{(\rho_s - \rho_w)}{\rho_w} \phi_S$$

This equation is independent of any formulation for the permeability. In the present chapter, so far, we have analysed the settling and consolidation phases with the same model. One can however wonder if this is the most suited approach. Intuitively, one can see that in the settling phase (where $v_{s/lab}$ can be measured by recording the water/suspension interface) there will be other forces acting on the particles than in the consolidation phase, and hence $v_{s/lab}$ will be different in the settling and in the consolidation regime. The (hindered) settling regime was analysed in **Chapter 8** whereas the consolidation regime was defined in **Chapter 9**. We will now discuss about how to link these two regimes.

In **Chapter 8** we have discussed the settling of hard spheres that were experiencing repulsion (which lead to the introduction of an osmotic pressure term Π). We found the following equation to describe the settling behaviour:

$$\frac{6\pi\eta a}{\chi(\phi_S)}(v_s - v_w) = -(\rho_s - \rho_w)gV_p - \frac{V_p}{\phi_S}\frac{\partial\Pi}{\partial z}$$

This equation can be rewritten:

$$\frac{\partial\Pi}{\partial z} = \frac{\phi_S}{V_p}\frac{6\pi\eta a}{\chi(\phi_S)}v_{w/s} - (\rho_s - \rho_w)g\phi_S$$

We recall the definition of the excess pore water pressure we have defined in **Chapter 9**, and the associated Darcy-like equation:

$$-\frac{\partial P_e}{\partial z} = \frac{\partial\sigma_{sk}}{\partial z} + (\rho_s - \rho_w)g\phi_S$$

$$(1 - \phi_S)v_{w/s} = \frac{-k}{\eta}\frac{\partial P_e}{\partial z}$$

Combining these two last equations leads to:

$$\frac{\partial\sigma_{sk}}{\partial z} = \frac{\eta}{k}(1 - \phi_S)v_{w/s} - (\rho_s - \rho_w)g\phi_S$$

Comparing the equation found in **Chapter 8** given above and the last equation, it is tempting to identify:

$$k(\phi_S) = \frac{2}{9}a^2\left(\frac{1 - \phi_S}{\phi_S}\right)\chi(\phi_S) = \frac{2}{9}a^2\frac{f(\phi_S)}{\phi_S}$$

where we used $V_p = 4\pi a^3/3$ and $f(\phi_S) = (1 - \phi_S)\chi(\phi_S)$. We then can set-up a general force balance for a water-sediment mixture:

$$\frac{\partial \sigma_{sk}}{\partial z} + \frac{\partial \Pi}{\partial z} = \frac{\eta}{k}(1 - \phi_S)v_{w/s} - (\rho_s - \rho_w)g\phi_S$$

In the soil phase, one has $\Pi = 0$ and in the suspension $\sigma_{sk} = 0$. This last equation was proposed by Toorman in 1996[134].

This equation simply tells that the initial stage of consolidation of soils, as defined in **Chapter 9** (with a settling regime, a primary consolidation regime and a secondary compression regime) can be modelled by an advection-diffusion equation, but that the advective and diffuse terms are different in each regime.

Depending on the composition of the suspension (hard or soft particles) and the initial volume fraction (under or above the gelling state), different advective and diffusive terms will be chosen, and the transition between the settling regime and the primary consolidation regime has to be defined carefully. In the extreme case of the settling of hard heavy spheres for example, the transition between the settling phase and consolidated bed is extremely abrupt, as the velocity of a hard sphere goes from a finite value to zero over an extremely small distance when it is hitting the bed. This can be challenging to program numerically.

Hindered settling of flocculated clay

The derivations in **Chapter 8** are done assuming that the settling particles are hard spheres. If Brownian motion would be neglected, this would imply that at infinite times the profile would be $\phi_S = \phi_{max}$ (the maximum packing volume fraction) until the interface where ϕ_S would jump to zero. Setting-up the Richardson-Zaki hindered settling function, we have, for simplicity, assumed that $\phi_{max} = 1$ which leads to $f(\phi_S) = 0$ for $\phi_S = 1$. A more realistic Richarson-Zaki profile[135] for hard spheres is

$$f(\phi_S) = (1 - \phi_S/\phi_{max})^m$$

where ϕ_{max} is close to 0.74 for packed spheres. For flocculated clay or mud suspensions, at $\phi_S = \phi_{max}$ or more precisely at $\phi_S = \phi_{gel}$ the settling is certainly not zero as consolidation occurs: the particles (flocs) are squeezed at the bottom of the column, which makes the water/bed interface move down. In this case, one usually speaks of *settlement* (a term coming from soil science) as all the particles are

---

[134] Toorman, E.A. (1996) "Sedimentation and self-weight consolidation: general unifying theory", Géotechnique 46, No 1, 103-113.
[135] In Chapter 8, the exponent, following standard notations, was defined as n. To avoid confusions with the variable n $= 2/(3 - D)$ defined in the present chapter, we have chosen to rename the exponent m.

touching and form a fabric moving downwards under the action of gravity instead of *settling* which usually refers to individual particles.

The first illustration at the beginning of the chapter represents a column filled with real flocs that are settling and depositing at the bottom of the column, where they are squeezed. In the illustration we have assumed that the flocs are fractal, and to represent the squeezing at the bottom, we have drawn flocs of smaller dimension as this amounts to increase the local volume fraction (see the fractal approach detailed above). Instead of assuming that the flocs are mathematical space-filling objects, we will in this subsection assume that the particles are *real* fractal flocs of radius R_N and that the clay concentration (volume fraction) ϕ_s is below the gelling concentration, i.e. that the flocs are not touching. (There will therefore be a jump in velocity at the suspension/bed interface.)

In order to evaluate the settling velocity of a fractal floc, we first need to define some specific volume fractions:

Definitions regarding suspensions of fractal flocs

flocs are suspended   flocs are space-filling
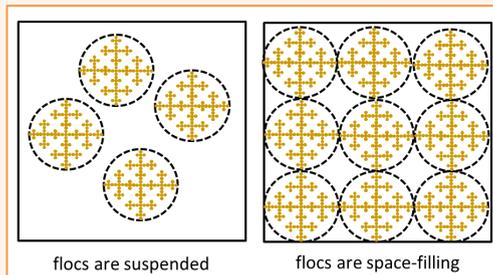
The suspended particles, made of primary particles of radius R_p, are fractal flocs of radius R_N with a fractal dimension D. The flocs are *not* necessarily assumed to be space-filling as was assumed in the first part of the chapter. Their fractal dimension is therefore in general different from the space-filling case.

We define:

$$\phi_{floc} = \frac{\text{volume of flocs}}{\text{volume of clay}} \times \frac{\text{volume of clay}}{\text{total volume}}$$

If we assume that all the clay is contained in the flocs:

$$\phi_{floc} = \frac{\text{volume of a floc}}{\text{volume of clay in a floc}} \times \phi_s$$

We recall that

$$\phi_{\text{clay in a floc}} = \left(\frac{R_p}{R_N}\right)^{3-D}$$

which leads to

$$\phi_{floc} = \left(\frac{R_N}{R_p}\right)^{3-D} \times \phi_s$$

If the flocs are space-filling $\phi_s = \phi_{\text{clay in a floc}}$ and one gets $\phi_{floc} = 1$. This implies that we can define a gelling concentration (= concentration of clay) for which the flocs are space-filling, which is given by

$$\phi_{gel} = \phi_{\text{clay in a floc}} = \left(\frac{R_p}{R_N}\right)^{3-D}$$

When the flocs are space-filling, it is clear that the fractal dimension is the same as the one discussed in the first part of this chapter.

We can now evaluate the Stokes settling velocity of a fractal sphere:

$$v_{Stokes(floc)} = \frac{2(\rho_{floc} - \rho_w)R_N{}^2 g}{9\eta}$$

The density of a fractal floc can be calculated as follows:

$$\rho_{floc} = \frac{\text{mass clay and mass water in a floc}}{\text{volume of a floc}}$$

$$\rho_{floc} = \rho_s \frac{\text{volume clay in a floc}}{\text{volume of a floc}} + \rho_w \left(1 - \frac{\text{volume clay in a floc}}{\text{volume of a floc}}\right)$$

Using the definition of ϕ_{gel} introduced above, we find that

$$\rho_{floc} - \rho_w = (\rho_s - \rho_w)\phi_{gel}$$

We get

$$v_{Stokes(floc)} = \frac{2(\rho_s - \rho_w)R_N{}^2 g}{9\eta}\phi_{gel}$$

This relation for the Stokes settling of a floc has been found by Kranenburg[136].

---

[136] Kranenburg, C. (1994). The fractal structure of cohesive sediment aggregates. Estuarine, Coastal and Shelf Science, 39(5), 451-460.

In order to find a hindered settling velocity of a floc, v_{hind}, we can adapt the Richardson-Zaki formula:

$$v_{hind} = \left(1 - \phi_{floc}\right)^m v_{Stokes(floc)}$$

where ϕ_{floc} is the volume fraction of flocs in the hindered settling zone. In that zone, we have everywhere

$$\phi_{floc} = \phi_s/\phi_{gel} = \phi_s \left(\frac{R_N}{R_p}\right)^{3-D}$$

One can verify that for D = 3 (the particle is a solid particle) one gets $\phi_{floc} = \phi_s$. The adapted Richardson-Zaki expression then reduces to the standard Richardson-Zaki expression. In the limit of space-filling flocs, $\phi_s = \phi_{gel}$ and the adapted Richardson-Zaki formula gives the result $v_{hind} = 0$. There is no settling velocity as the Richardson-Zaki formulation does not account for compression.

At volume fractions such that $\phi_{floc} < 1$, one can make the approximation

$$v_{hind} \sim \left(\left(\frac{R_N}{R_p}\right)^{D-3} - m\phi_0\right) \frac{2(\rho_s - \rho_w)R_N{}^2 g}{9\eta}$$

This implies that the function v_{hind} as function of ϕ_0 is a straight line of slope

$$m \frac{2(\rho_s - \rho_w)R_N{}^2 g}{9\eta}$$

from which m can theoretically be deduced if the average size of flocs are known.

Beyond self-weight consolidation

The Gibson equation, which is the topic of **Chapter 9** and the present chapter was originally created to study the **large strains** occurring during the self-weight consolidation of slurries. In **Chapter 9**, we have already stated that the stresses felt by particles during self-weight consolidation are very small compared to the pressures applied in an **oedometer test**.

The oedometer test is used to investigate the 1D consolidation of fine-grained soils. From the oedometer test one gets the relation between **vertical effective stress** (the load applied) and **vertical strain** (the deformation), from which the void ratio of the sample can be determined.

The sample is placed between two porous disks at the top and bottom. The disks are in contact with the same bath of water. This implies that at the end of the test, the water pressure will be hydrostatic through the sample. A compressive stress is applied from the top, by a vertical load, which is assumed to act uniformly over the area of the soil sample:
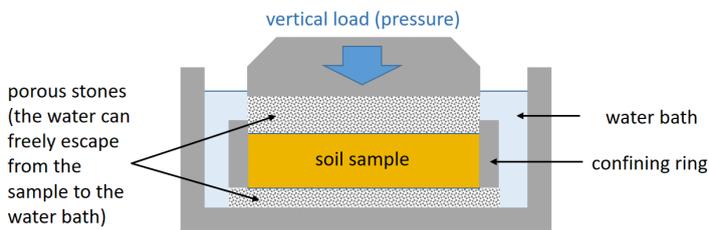
Initially, all the vertical load is taken by pore water, because, due to the low permeability of the soil, the pore water is unable to flow out of the voids quickly. Therefore, there is very little compression of the soil sample immediately after placing the load. After a few seconds, the pore water begins to flow out. This results in a decrease in pore water pressure. At the same time, the effective stress increases. As a result, the sample settles. Several increments of vertical stress are applied in an oedometer test, usually by doubling the previous increment. For each load, the final settlement of the soil sample and the time taken to reach this final settlement are recorded.

When analysing these types of experiments, several hypothesis are made:

1 – the sample does not undergo self-weight consolidation; the settlement is solely due to the applied stress.

2 – the volume fraction is constant over the height of the sample.

3 – The particles are incompressible

4 – the water is incompressible

Using hypothesis 1 (and hypothesis 3 and 4 that have been used implicitly until now, also in **Chapters 9** and **10**), the full Gibson equation reduces to

$$\frac{\partial \phi_S}{\partial t} = \frac{\partial}{\partial z}\left[\frac{K}{g\rho_w}\phi_S \frac{\partial \sigma_{sk}}{\partial z}\right]$$

where, to remove the effect of self-weight consolidation we have taken $\rho_s = \rho_w$ (which mathematically amounts to the same). This equation can further be developed into

$$\frac{\partial \phi_S}{\partial t} = \frac{\partial}{\partial z}\left[\frac{K}{g\rho_w}\phi_S \frac{\partial \sigma_{sk}}{\partial \phi_S}\frac{\partial \phi_S}{\partial z}\right]$$

The so-called **consolidation coefficient** is defined by

$$C_v = \frac{K}{g\rho_w}\phi_S \frac{\partial \sigma_{sk}}{\partial \phi_S}$$

(the curl d's are here equivalent to the straight d' as σ_{sk} only depends on ϕ_S). Using the fractal approach given in the present chapter, one finds that

$$C_v = n\frac{K_k K_\sigma}{g\rho_w}$$

which is a constant. The fact it is a constant depends on the peculiar expressions that were chosen for $K(\sim \phi_S^{-n})$ and $\sigma_{sk}(\sim \phi_S^{n})$. In soil science, it is generally assumed that the consolidation coefficient is constant.

With C_v constant, the equation reduces to

$$\frac{\partial \phi_S}{\partial t} = C_v \frac{\partial^2 \phi_S}{\partial z^2}$$

This equation is mathematically equivalent to the one called **Fick's second law** in **Chapter 2** and can be solved analytically.

The hypothesis used to derive this equation were (1,3,4). We did *not* assume hypothesis 2, which implies that this equation is valid for samples exhibiting a change in volume fraction over height. This equation is useless in the case we assume hypothesis 2, as it then reduces to

$$\frac{\partial \phi_S}{\partial t} = 0$$

This equation tells us that there is no local (at a position z) variation of ϕ_S in time, which is compatible with hypothesis 2 : there is only a variation in time of the *total* volume fraction of the sample between before and after the loading (but no variation within the sample).

Usually, in an oedometer test, the vertical strain is recorded as function of the vertical effective (skeleton) stress. The strain is defined by

$$d\varepsilon = \frac{h - h_0}{h_0}$$

where h is the height of the sample after the test and h_0 the height before. We used the symbol $d\varepsilon$ to indicate that the change is height is assumed to be quite small. As the sample does not deform laterally, we also have

$$d\varepsilon = \frac{V - V_0}{V_0} = \frac{dV}{V_0}$$

where the volumes are related to the heights by $V = Ah$ and $V_0 = Ah_0$ with A being the cross-sectional area of the sample. By definition

$$V = V_w + V_s$$

where V_w is the volume of water and V_s the volume of solids.

The volume V is changing because water is flowing out of the sample. Therefore

$$d\varepsilon = \frac{dV_w}{V_0}$$

and the volume V_s of solids is not changing. From the definition of volume fraction, we have

$$d\phi_S = \frac{V_s}{V} - \frac{V_s}{V_0} = -\frac{V_s}{V} \frac{V - V_0}{V_0} = -\phi_{S,\text{end}} d\varepsilon$$

where $\phi_{S,\text{end}}$ is the volume fraction after loading. As the incremental change $d\varepsilon$ is very small, we have

$$h = h_0(1 + d\varepsilon) \sim h_0$$

and therefore

$$V \sim V_0$$

$$\phi_{S,\text{end}} \sim \phi_{S,\text{begin}} \sim \phi_S$$

where $\phi_{S,\text{begin}}$ is the volume fraction before loading. One has to realize that by defining the strain, we have implicitly defined it as a macroscopic quantity, a property affecting the whole sample. Therefore the volume fraction ϕ_S specified here is representing the volume fraction of the sample (not a slice of the sample). This implies that we fulfil hypothesis 2 (we assume that ϕ_S is constant over the height). We get

$$d\phi_S = -\phi_S d\varepsilon$$

$$\frac{d\phi_S}{dt} = -\phi_S \frac{d\varepsilon}{dt}$$

This last equation indicates that there is a (small) change in the total volume fraction of the sample between before and after applying the loading. Note that we have here used the Lagrangian derivatives, as ϕ_S and ε only depend on time only. Using the Lagrangian form of Gibson equation, which we simplify using $\rho_s = \rho_w$ for the reason given above, we get

$$\frac{d\phi_S}{dt} = -\phi_S \frac{d\varepsilon}{dt} = \phi_S \frac{\partial}{\partial z}\left[\frac{K}{g\rho_w}\frac{\partial \sigma_{sk}}{\partial z}\right]$$

$$\frac{d\varepsilon}{dt} = \frac{\partial}{\partial z}\left[\frac{K}{g\rho_w}\frac{\partial \sigma_{sk}}{\partial z}\right]$$

This equation links the change in strain to the change in stress. We also get

$$\frac{d\phi_S}{dt} = \frac{d\sigma_{sk}}{dt}\frac{d\phi_S}{d\sigma_{sk}} = \phi_S \frac{\partial}{\partial z}\left[\frac{K}{g\rho_w}\frac{\partial \sigma_{sk}}{\partial z}\right]$$

$$\frac{d\sigma_{sk}}{dt} = \frac{d\sigma_{sk}}{d\phi_S}\phi_S \frac{\partial}{\partial z}\left[\frac{K}{g\rho_w}\frac{\partial \sigma_{sk}}{\partial z}\right]$$

In order to define the consolidation coefficient, it is now necessary to assume that the permeability K is constant, so it can get out of the bracket term. This condition is fulfilled if we assume that ϕ_S is constant (our hypothesis 2). We then get

$$\frac{d\sigma_{sk}}{dt} = C_v \frac{\partial^2 \sigma_{sk}}{\partial z^2}$$

This equation enables to find the change in stress, from which the change in strain can be obtained. The curl d's indicate that the derivative in z should be taken at t constant. We see that even if ϕ_S is constant over the whole sample there is a change in σ_{sk} over height, which originates from the fact that we apply a load at the top.

This implies that the stress between particles is higher at the bottom of the sample, but that this does not influence their local (= at given z) compaction.

We also have

$$\frac{d\sigma_{sk}}{dt} = \frac{\partial \sigma_{sk}}{\partial t} + \frac{dz}{dt}\frac{\partial \sigma_{sk}}{\partial z}$$

where dz/dt is the velocity of the settlement. If that settlement is slow, we have in good approximation

$$\frac{d\sigma_{sk}}{dt} = \frac{\partial \sigma_{sk}}{\partial t}$$

The change in effective stress over time is then the same at any z.

The load applied to the sample in a test is constant. From the definitions of stresses (see **Chapter 9**), we get

$$\sigma_{zz} = \sigma_{sk} + P_e + P_{hyd}$$

As inertia is neglected, one can say that the total stress σ_{zz} of the sample is equal to the stress from the load. As neither σ_{zz} (which is constant as the load is constant) nor P_{hyd} depend on time, and that P_{hyd} is linear in z we get

$$\frac{d\sigma_{sk}}{dt} = -\frac{dP_e}{dt}$$

$$\frac{\partial^2 \sigma_{sk}}{\partial z^2} = -\frac{\partial^2 P_e}{\partial z^2}$$

which gives

$$\frac{dP_e}{dt} = C_v \frac{\partial^2 P_e}{\partial z^2}$$

This equation is only fulfilled when all hypothesis (1-4) are obeyed.

Numerical solution

The differential equation

$$\frac{dP_e}{dt} = C_v \frac{\partial^2 P_e}{\partial z^2}$$

can be solved numerically using the so-called finite differences method[137]. Originally it was solved analytically by **Terzaghi[138]** in 1923. For numerical stability it is always better to convert the equation to a dimensionless one:

$$\frac{dP_e{}^*}{dt^*} = C_v{}^* \frac{\partial^2 P_e{}^*}{\partial z^{*2}}$$

with

$$t^* = t/\tau$$

$$z^* = z/H$$

$$C_v{}^* = \frac{C_v}{H^2/\tau}$$

$$P_e{}^* = P_e/P_0$$

where τ and H are characteristic time and height and P_0 a reference pressure. Writing the equation in dimensionless form also enables to find that the final deformation has been reached when

$$\frac{dP_e{}^*}{dt^*} \sim 0$$

which occurs when

$$\frac{C_v}{H^2/\tau} \ll 1$$

In other words, the consolidation time τ will take 4 times as long if the layer (H) is twice as thick.

For ease of notation, we drop the *, but in the following we only refer to the dimensionless equation.

The equation is discretized as follows:

$$\frac{dP_e}{dt} = \frac{P_e(t + dt) - P_e(t)}{dt}$$

---

[137] Verruijt, Arnold. *An Introduction to Soil Mechanics*. Springer, Theory and Applications of Transport in Porous Media, Vol.30 (2018)

[138] Karl von Terzaghi (1883 - 1963) was an Austrian geotechnical engineer known as "the father of soil mechanics"

$$\frac{\partial^2 P_e}{\partial z^2} = \frac{P_e(z + dz, t) - 2P_e(z, t) + P_e(z - dz, t)}{(dz)^2}$$

This leads to

$$P_e(z, t + dt) = P_e(z, t) + C_v \frac{dt}{(dz)^2}[P_e(z + dz, t) - 2P_e(z, t) + P_e(z - dz, t)]$$

In an oedometer test the sample is usually drained at the top, using a thin sheet of filter paper and a porous stone. In the container in which the sample and its surrounding ring are placed, the water level is kept constant. This implies that at the top of the sample the excess pore pressure is zero, which results in the boundary conditions:

$$P_e(z = 1, t) = 0$$

where $z = 1$ represents the top of the sample.

The sample is usually undrained at the bottom where an impermeable plate prevents the water flow. The boundary condition at the bottom of the sample is therefore

$$\frac{\partial P_e(z = 0, t)}{\partial z} = 0$$

where the curl d represents the fact that the derivative is taken at a given time: for all times, the boundary condition is respected.

Numerically, this last boundary condition is usually implemented by defining a (fictive) pressure $P_e(-dz, t)$ located at a position dz below the sample and imposing the condition

$$P_e(-dz, t) = P_e(+dz, t)$$

This condition ensures that whatever the value of $P_e(0, t)$ the boundary condition $\partial P_e(z = 0, t)/\partial z = 0$ is satisfied. Numerically, this means that the value of $P_e(0, t)$ will be determined by

$$P_e(0, t + dt) = P_e(0, t) + C_v \frac{dt}{(dz)^2}[P_e(dz, t) - 2P_e(0, t) + P_e(dz, t)]$$

At $t = 0$ we also assume that the pressure within the sample is known:

$$P_e(z < 1, t = 0) = 1$$

In order to get a stable numerical solution it is important to ensure that

$$dt \ll \frac{(dz)^2}{C_v}$$

Using $C_v = 0.1$ and $dz = 0.1$, this implies that $dt \ll 0.1$. In the Matlab code given underneath we used dt = 0.01.

The dissipation of pore water pressure over time (in dimensionless units) is given by:
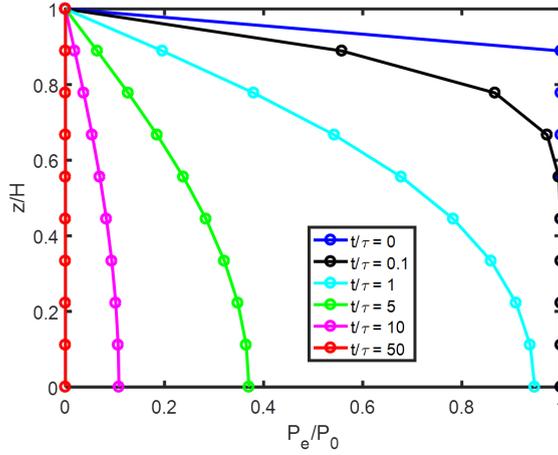
From the definition of the consolidation coefficient and the strain, we get

$$C_v = \frac{-K}{g\rho_w} \frac{d\sigma_{sk}}{d\varepsilon}$$

The **compressibility** m_v is defined as

$$m_v = \frac{K}{g\rho_w C_v}$$

and this gives

$$d\varepsilon = -m_v d\sigma_{sk}$$

The minus sign indicate the fact that an increase in effective stress is coupled with a decrease in strain (the sample is compressed). As we fulfil hypothesis 1- 4, m_v is a constant.

As discussed above, the strain (and the volume fraction) is defined as a property applied to the *whole* sample. We have found that the effective stress (as the excess pore water pressure) is changing over the height. The equation we have derived for the strain should therefore be written

$$d\varepsilon = \frac{-m_v}{h} \int_0^h d\sigma_{sk} dz$$

which indicates that the strain is a function of the *average* skeleton stress. We get from the definition of the strain:

$$\frac{h(t+dt) - h(t)}{h(t)} = \frac{-m_v}{h(t)} \int_0^{h(t)} \sigma_{sk} dz$$

$$dh = h(t+dt) - h(t) = -m_v \int_0^{h(t)} \sigma_{sk}(z,t) dz$$

We define

$$m_v{}^* = m_v P_0$$

$$\sigma_{sk}{}^* = \sigma_{sk}/P_0$$

where P_0 is, as above, a reference pressure. We get

$$dh = -m_v{}^* \left[\int_0^{h(t)} \left(\sigma_{zz}{}^* - P_{hyd}{}^* \right) dz - \int_0^{h(t)} P_e{}^* dz \right]$$

The first integral on the right-hand side is a constant as the load and average hydrostatic pressure are both constant in time. At the end of consolidation the excess pore water pressure is zero, as demonstrated above and therefore we define

$$dh_\infty = -m_v{}^* \int_0^{h(t=\infty)} \left(\sigma_{zz}{}^* - P_{hyd}{}^* \right) dz$$

We get

$$dh - dh_\infty = -m_v{}^* \int_0^{h(t)} P_e{}^* dz$$

We also have, with $dh_0 = h(dt) - h(0) = h(dt) - h_0$

$$dh_0 - dh_\infty = -m_v{}^* h_0$$

as by definition $P_e{}^*(t=0) = 1$, which implies that

$$\frac{dh(t) - dh_\infty}{dh_0 - dh_\infty} = \frac{1}{h_0} \int_0^{h(t)} P_e{}^* dz = \int_0^{h(t)/h_0} P_e{}^* dz$$

The **degree of consolidation** U is defined as

$$U = \frac{dh_0 - dh(t)}{dh_0 - dh_\infty} = 1 - \frac{dh(t) - dh_\infty}{dh_0 - dh_\infty}$$

yielding

$$U = 1 - \int_0^{h(t)/h_0} P_e^* dz$$

The degree of consolidation U as function of time is given by:
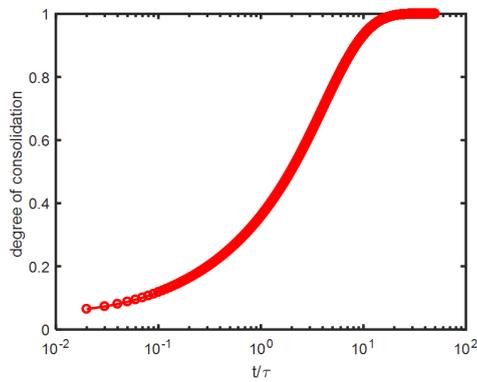
The Matlab code is given by:

```matlab
clear all
close all

nh=10; % number of steps in z direction
zbin(1) = 0.0d0; % bottom of the sample
dz = 1/(nh-1); % dimensionless increment in z direction
%zbin(1) = 0 ; zbin(2) = dz ; … ; zbin(nh) = (nh-1)*dz = 1

for m = 2:nh
    zbin(m)  = zbin(m-1)+dz;
    Pe(m)  = 1; % dimensionless excess pore water pressure
end

Pe(1)=1;
Pe(nh)=0;

Cv = 0.1; % dimensionless consolidation coefficient

tnum(1) = 0.01;
Pestart = Pe; % Pestart is Pe(t=0)

for kt = 2:5000
    dt = 0.01;
    tnum(kt) = tnum(kt-1)+dt;
    tnumplot(kt) = tnum(kt);
    a = Cv*dt/dz.^2;
    for m = 2:nh-1
        Penew(m) = Pe(m)+a*(Pe(m+1)-2*Pe(m)+Pe(m-1));
        % Penew(m) is Pe(t+dt)
        % Pe(m) is Pe(t)
    end
    Penew(nh) = Pe(nh); % Pe(H) = 0
    Penew(1) = Pe(1)+a*(Pe(2)-2*Pe(1)+Pe(2));
    Pe = Penew;

    height(kt) = trapz(zbin,Pe);

    if kt == 10 % t=0.1
        Pe1 = Pe;
    end

    if kt == 100 % t=1
        Pe2 = Pe;
    end

    if kt == 500 % t=5
        Pe3 = Pe;
    end

    if kt == 1000 % t=10
        Pe4 = Pe;
    end

    if kt == 5000 % t= 50
        Pe5 = Pe;
    end
```

```matlab
end

plot(Pestart,zbin,'-ob','LineWidth',2)
hold on
plot(Pe1,zbin,'-ok','LineWidth',2)
hold on
plot(Pe2,zbin,'-oc','LineWidth',2)
hold on
plot(Pe3,zbin,'-og','LineWidth',2)
hold on
plot(Pe4,zbin,'-om','LineWidth',2)
hold on
plot(Pe5,zbin,'-or','LineWidth',2)
hold on

legend('t/\tau = 0','t/\tau = 0.1','t/\tau = 1','t/\tau = 5','t/\tau =
10','t/\tau = 50');
ylim([0 1])
xlabel('P_e/P_0','FontSize',12);
ylabel('z/H','FontSize',12);
set(gca,'LineWidth',2,'FontSize',12);
set(gcf,'Color',[1 1 1])

figure

semilogx(tnumplot,1-height,'-or','LineWidth',2)
hold on

xlabel('t/\tau','FontSize',12);
ylabel('degree of consolidation','FontSize',12);
set(gca,'LineWidth',2,'FontSize',12);
set(gcf,'Color',[1 1 1])
```

Conclusion

from colloid science
to large-scale applications

As stated in the introduction, mud and clayey systems are studied in different fields of research. We mentioned soil science and sediment dynamics, but these generic names are associated to specific expertise such as agronomy, chemistry, geology, physical geography, biology, civil engineering… In this book, we focussed primarily on colloid science.
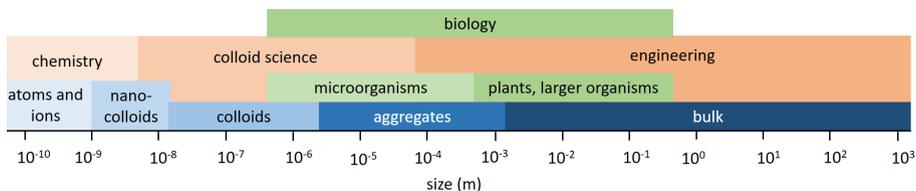
Length scales and associated particle sizes and field of science in the context of sediment characterization. The term "bulk" refers to the fact that above a given length scale no distinction is made between the carrying fluid and the sediment: water and sediment is then seen as a continuum, with bulk properties (density, viscosity, …)

Despite the promise of the subtitle of the book (Applications to sediment characterization), which hints to the study of natural muddy suspensions, the reader will have noticed that many findings described in the book apply to well-characterized systems only, i.e. particles with a defined size, shape and surface charge. Even though colloid scientists are working on more complex systems, for instance mixtures or anisotropically shaped particles (see articles cited and books in the reference list), it is clear that fundamental research is at present not able to bridge completely the knowledge between well-characterized and natural systems. Even if some models, like the flocculation models presented in **Chapter 6** can – in some cases – be applied quite successfully to the flocculation of natural systems, these models suffer from a severe restriction: their validity is limited to a scale that is far below the scale usually desired from an engineering point of view.

The primary reason elaborated flocculation models cannot be applied in large-scale transport models is due to the fact that, when coupled to a large-scale hydrodynamic model, the resulting sediment transport model would be too CPU intensive to run. It would be impossible to make long-term predictions with it. It is therefore necessary to use less elaborated models[139,140,141], that will be able to reproduce the

---

[139] Verney, Romaric, et al. "Behaviour of a floc population during a tidal cycle: laboratory experiments and numerical modelling." *Continental Shelf Research* 31.10 (2011): S64-S83.

[140] Lee, Byung Joon, et al. "A two-class population balance equation yielding bimodal flocculation of marine or estuarine sediments." *Water research* 45.5 (2011): 2131-2145.

[141] Manning, Andrew J., et al. "Flocculation settling characteristics of mud: sand mixtures." *Ocean dynamics* 60.2 (2010): 237-253.

in-situ measured data (mean floc size, settling velocities) to a reasonable accuracy, with a limited number of adjustable parameters.

Doing so, one is forced to leave the bottom-up approach usually adopted in fundamental (colloid) science and presented in this book in order to use a top-down approach. The bottom-up approach consists in starting with simplified models and add layers of complexity, whereas the top-down approach consists in taking a system as a whole and breaking it down in sub-systems in a reverse engineering fashion. Both approaches are not mutually exclusive, and there is often much to gain to combine them. A lot of work is currently devoted to bridge the gap between colloid science and engineering.
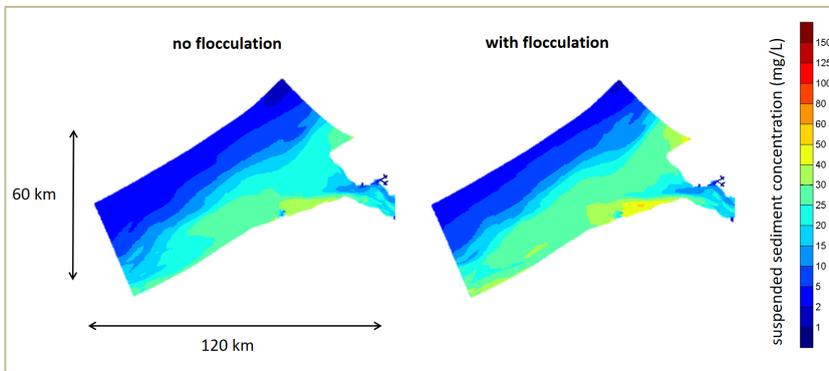
Illustration of the impact of a microscopic process (flocculation) on large-scale (km) sediment transport. The area represents the mouth of the Western Scheldt estuary and the colour indicates a three-month average suspended sediment concentration (SSC) with and without taking flocculation into account. The differences in SSC are particularly obvious close to the mouth. The flocculation model is taken from Manning et al.[142]

---

[142] Manning, A. J., and K. R. Dyer. "Mass settling flux of fine sediments in Northern European estuaries: measurements and predictions." Marine Geology 245.1-4 (2007): 107-122.

References for colloid science books

Kruyt, H. R. editor. (1952)
 Colloid Science. Vol. 1. Irreversible Systems. Elsevier.
A classic reference, with large contributions of J. Th. Overbeek

Russel, W. B., Saville, D. A., & Schowalter, W. R. (1989)
Colloidal dispersions. Cambridge university press.
A complete reference, especially suited for the theoretical modelling of suspensions.

Van de Ven, T. G. (1989)
 Colloidal hydrodynamics. Academic Press.
A great reference to understand the principles of colloidal hydrodynamics from first principles.

Hunter, R. J. (1993)
 Introduction to modern colloid science. Oxford University Press.
A classic and easy to read reference to understand the fundamentals of colloid science, with limited derivations.

Hunter, R. J. (2000)
Foundations of Colloid Science. Oxford University Press.
A completely revised and updated edition of a classic textbook on colloid science

Shaw, D. J., & Costello, B. (1993)
Introduction to colloid and surface chemistry. Butterworth-Heinemann, Oxford.
A concise overall coverage of colloid and surface chemistry, well suited for getting a broad background on the topic.

Dhont, J. KG. (1996)
 An introduction to dynamics of colloids. Vol. 2. Elsevier.
A great reference for detailed mathematical derivations and interpretation thereof

Hiemenz, P. C., & Rajagopalan, R. (Eds.). (1997)
Principles of Colloid and Surface Chemistry, revised and expanded (Vol. 14). CRC press.
A very complete book that help to familiarize with the fundamentals of colloid and surface science, characterization and measurements techniques.

Ohshima, H
Theory of colloid and interfacial electric phenomena. Vol. 12. Academic Press, 2006.
A classic book to understand the fundamentals and current developments in colloidal electrokinetic phenomena

G.J.M. Koper (2011)
An introduction to Interfacial Engineering. VSSD.
A concise but broad overview of various interfacial properties, including colloidal aspects

Mewis, J., & Wagner, N. J. (2012)
Colloidal suspension rheology. Cambridge University Press.
A recent book focussing on the link between microstructure and rheology, with limited derivations but multiple references.

Hunter, R. J. (2013)
Zeta potential in colloid science: principles and applications (Vol. 2). Academic press.
A recent reprint of a classic book focusing on electrokinetics, both theoretically and experimentally

Index

S

Z