1 Development of Membrane Processes

K. Smith

1.1 Historical background

The ability of membranes to separate water from solutes has been known since 1748, when Abbé Nolet experimented with the movement of water through a semi-permeable membrane. Depending on the reference, either Abbé Nolet or Dutorchet coined the word osmosis to describe the process. Throughout the 18th and 19th centuries, membranes were used exclusively for laboratory applications, and often consisted of sausage casings made from animal intestines or the bladders of pigs, cattle or fish.

The first synthetic membranes were produced by Fick in 1855, and appear to have been made of nitrocellulose. Membranes based on cellulose were known as collodion and had the advantages of reproducible characteristics compared with the previously used animal-based membranes. Bechhold further advanced the process for manufacturing collodion membranes when he developed methods for controlling pore size and measuring pore diameters in 1907. He is generally credited with first using the term ultrafiltration (UF). In addition, Richard Zsigmondy at the University of Göttingen, Germany, patented a membrane filter in 1918 that was referred to as a cold ultrafilter. His work becomes the basis of the membrane filters produced by Sartorius GmbH.

Collodion membranes produced by the Sartorius GmgH of Germany became commercially available in 1927. The primary use of membranes until the 1940s was the removal of micro-organisms and particles from liquids and gases and research applications. There was a critical need to test drinking water in Europe for microbial content following the Second World War, and membranes were developed that could rapidly filter water and capture any micro-organisms on the membrane surface, where they could quickly be enumerated to determine the safety of the water for human consumption.

In addition to the separation of relatively large particles from water, there was interest in developing membranes that could desalinate sea or brackish water. The term reverse osmosis (RO) had been coined in 1931 when a patent was issued for desalting water; however, the available membranes could not withstand the pressures required.

Although many improvements were made in the following years, including the use of other polymers for constructing membranes, membranes were limited to laboratory and small specialised industrial applications. Factors limiting the use of membranes included a lack of reliability, being too slow, not sufficiently selective and cost.
A breakthrough came in the early 1960s when Sourirajan and Loeb developed a process for making high-flux, defect-free membranes capable of desalinating water. Researchers at the time believed the best approach to improving flux would be to reduce the thickness and thereby the resistance to flow of the membrane. Sourirajan and Loeb attempted to produce such membranes by taking existing cellulose acetate membranes and heating them while submersed in water in a process known as annealing. They expected the membrane pores would increase in size by such a process, but instead the pores became smaller and the membrane more dense. When they attempted the same process with cellulose acetate UF membranes, they discovered not only did the pores become smaller but the ability of the membrane to reject salt increased, as did flux. The flux improvement was such that the membranes could be a practical way to desalinate water.

The annealing process of Sourirajan and Loeb had created an anisotropic or asymmetric membrane. Anisotropic membranes have different behaviour depending on which side of the membrane is used for the separation. Although this type of membrane had been seen over 100 years earlier with natural membranes, it had not been reproduced with the synthetic variety.

The key to the anisotropic membrane of Sourirajan and Loeb was the thin ‘skin’ on one surface of the membrane. The skin typically was approximately 0.1–0.2 μm thick and had a dense structure whereas the remainder of the membrane had a very porous open structure. The thickness of the membrane essentially determined the flux and so by reducing the effective separating distance from 100–200 μm to 0.1–0.2 μm the rate of liquid crossing the membrane dramatically improved, but because of the small pores in the skin the rejection of salt remained high.

Many changes in the production of membranes occurred during the 1960s, 1970s and 1980s. By continuing the work of Sourirajan and Loeb, others were able to develop additional methods for producing membranes. Initial membrane modules were plate-and-frame (Danish Sugar Corporation) or hollow fibre (Amicon) designs, but membranes in formats, such as spiral-wound and tubular (Abcor), were introduced shortly afterwards. The thickness of the separating layer was further reduced to less than 0.1 μm. Large plants using RO, UF and microfiltration (MF) were operating around the world by 1980.

Cellulose acetate remained the material of choice until the mid-1970s, when methods of producing composite membranes for water desalination were developed. By combining polysulphone and polyamide, composite membranes had the advantage of high salt rejection combined with good water flux and increased resistance to temperature and chemicals. Nanofiltration (NF) or ‘loose RO’ membranes became available in the mid-1980s. The NF membranes operated at lower pressures than RO systems, and were able to permeate monovalent ions. They found immediate application in producing ultrapure water by permeating trace salts from water produced by RO.

In addition, membranes made from inorganic materials, such as zirconium and titanium dioxide, became commercially available in the mid-1980s. Membranes made from these materials are referred to as mineral or ceramic and are available in tubular form for UF and MF. Union Carbide (USA) and Société de Fabrication d’Elements (France) used carbon tubes covered with zirconium oxide for their inorganic membranes. Later Ceraver (France) used a ceramic base with aluminium oxide. Chemical and temperature resistance were the significant advantages of ceramic membranes. It was originally thought
that such membranes had an unlimited life, but subsequent experience has shown this is not the case.

Advancements in membrane composition and design along with operation of membrane systems have continued. A wide variety of membrane polymers and designs have been adapted for RO, NF, UF and MF, resulting in many commercial applications. The feasibility of membrane-based applications depends chiefly on the ability of the filtration process to economically produce an acceptable product. Membrane pore size distribution, selectivity, operating conditions, membrane life, capital and operating costs become important economic considerations. These parameters are in turn influenced by many factors, such as the membrane polymer, element configuration and system design.

1.2 Basic principles of membrane separations

Membrane filtration is a pressure-driven separation process using semi-permeable membranes. The size of membrane pores and the pressure used indicate whether the term RO, NF, UF or MF is used for a given separation. RO and NF systems use the highest pressures and membranes with the smallest pores, whereas MF has the lowest operating pressures and membranes with the largest pores. UF is intermediate in pressure used and membrane pore size.

1.2.1 Depth versus screen filters

In the past, filtration processes relied on depth filters. This type of filter has fibres or beads in a mesh-like structure. Particles in the feed solution become trapped or adsorbed within the filter network, which eventually clogs the filter, thereby resulting in replacement of the filter (Fig. 1.1). By contrast, screen type filters generally rely on pores, with the size and shape of the pores determining passage of particles. Pores are more rigid and uniform and have a more narrowly defined size than mesh openings in a depth filter. Components not able to pass through pores remain on the membrane surface and, therefore, do not typically become trapped within the membrane structure (Fig. 1.2). Because the fouling materials remain on the surface, internal fouling decreases and the membrane can be reused.
1.2.2 Isotropic versus anisotropic membranes

Membranes can have several types of internal structure. Terms, such as microporous, non-porous, isotropic and anisotropic, refer to the structure of the membrane. Typically, membranes are either isotropic or anisotropic. Microporous and non-porous refer to isotropic membrane structure. An isotropic membrane will have a relatively uniform structure (Fig. 1.3), i.e. the size of the pores is similar throughout the membrane. The membrane, therefore, does not have a top or bottom layer, rather the membrane properties are uniform in direction. Isotropic membranes generally act as depth filters and, therefore, retain particles within the internal structure resulting in plugging and reduced flux.

Microporous and non-porous membranes typically are isotropic. Microporous membrane structure can resemble a traditional filter; however, the microporous membrane has extremely small pores. Materials are rejected at the surface, trapped within the membrane or pass through pores unhindered, depending on particle size and size of the pores. A non-porous membrane will not have visible pores and materials move by diffusion through the membrane.

An anisotropic or asymmetric membrane has pores that differ in size depending on their location within the membrane (Fig. 1.4). Typically, anisotropic membranes will have a thin, dense skin supported by a thicker and a more porous substructure layer. The thin top layer provides high selectivity, whereas the porous bottom layer has good flux. Membranes used for commercial separations in the food industry are typically anisotropic.
1.2.3 Cross-flow filtration

RO, NF, UF and MF systems all involve cross-flow filtration, which can be compared to the traditional method of perpendicular filtration. In traditional filtration (Fig. 1.5), the entire feed stream passes through the filtering media, i.e. the incoming stream flows perpendicularly to the filter with the filter retaining any trapped solids. The result is a filtered stream with solids trapped on and within the filter.

In cross-flow filtration (Fig. 1.6), the feed stream passes parallel to the membrane. Some of the incoming feed stream and particles will cross the membrane into the permeate section, whereas the other portion with the concentrated solids is the retentate stream. At any time only some of the water and particles will cross the membrane into the permeate stream, unlike traditional filtration where most particles are trapped after one pass through the filter. Because the feed stream flows parallel to the membrane rather
than perpendicular to it, cross-flow filtration is self-cleaning by comparison. Solutes and particles are continually swept along and away from the membrane surface by the retentate stream, thereby allowing longer operating times without cleaning than is possible with traditional filtration.

The affect of cross-flow permeate flow and thickness of the fouling or cake layer can be seen in Fig. 1.7. In perpendicular filtration, the flow of permeate is reduced as the thickness of the material on the surface of the filter, i.e. the thickness of the cake layer, increases over time. With cross-flow filtration, however, the thickness of the material on the membrane is limited by action of the feed stream sweeping across the surface of the
membrane. Because the thickness of the deposited material is limited, permeate flow is maintained at a higher level throughout filtering.

1.2.4 Requirements of membrane processes

The shared characteristics of membrane processes are pressure-driven processes using semi-permeable membranes. Pressure is used to reverse the direction of the osmosis process, while differences in membrane permeability determine separation of molecules. The process of osmosis is illustrated in Fig. 1.8. Solutions containing two different concentrations of dissolved materials are separated by a membrane that will allow only water to cross

![Fig. 1.8 The processes of osmosis and reverse osmosis.](image-url)
(Fig. 1.8a). Nature will try to equalise the concentration of the two solutions. Since the dissolved material cannot cross the membrane, water must flow from the solution of lower concentration to the solution at the higher concentration (Fig. 1.8b). The flow of water will continue until the solutions are of equal concentration or no more water is available. The difference in the height of water in the corresponding tubes is a result of the movement of water from lower to higher concentration. The final water level in the more concentrated solution compared with the original level is equal to the apparent osmotic pressure.

In the process of ‘reverse’ osmosis, pressure is used to force water to flow in the opposite direction (Fig. 1.8c). Enough pressure must be applied to overcome the apparent osmotic pressure of the more concentrated solution before water can flow from the more concentrated to the less concentrated side. In doing so, the more concentrated side becomes even more concentrated through the loss of water. It is this ability to concentrate and separate that is taken advantage of in commercial membrane separations.

Another shared characteristic is the use of semi-permeable membranes. Membranes can be distinguished from filters by the size of the particulates that are separated. By convention, filters generally separate particulates that are greater than 1–10 μm in size, whereas membranes separate smaller particles. Semi-permeable refers to the ability to separate some particles from other particles.

1.3 Types of membrane separations

The classification of membranes as RO, NF, UF and MF is somewhat arbitrary, and has considerable overlap between categories. Generally, RO/NF membranes will retain molecules in the ionic size range, UF membranes will separate macromolecules, and MF will retain particles of micron size. Because RO, NF, UF and MF membranes differ in the size of molecules they separate, the osmotic pressure involved is considerably different between the processes. RO, which retains the smallest molecules, has the highest osmotic pressure to overcome and, therefore, requires the highest operating pressure. A range from 1.38 to 8.28 MPa is common for RO, 1.03 to 2.76 MPa for NF, 0.21 to 1.03 MPa for UF, and MF requires only from 0.07 to 0.69 MPa.

1.3.1 Reverse osmosis

RO membranes generally retain all compounds allowing only water to cross into the permeate. There are exceptions to this general statement and, at times, relatively large molecules may pass into the permeate. RO membranes can, therefore, either concentrate a feed stream (retentate stream) through removal of water or produce very pure water (permeate stream).

1.3.2 Nanofiltration

NF membranes are very similar to RO membranes with the exception that NF membranes will allow the passage of monovalent ions into the permeate. NF membranes are very effective at concentrating materials in the feed stream since only monovalent ions are
removed into the permeate. The loss of monovalent ions into the permeate allows NF systems to operate at lower pressures than RO systems.

1.3.3 Ultrafiltration

UF membranes both fractionate and concentrate materials in the feed stream. Larger components, such as micro-organisms, lipids and proteins, typically are retained by UF membranes, whereas smaller molecules, such as sugars and minerals, pass into the permeate. UF systems operate at much lower pressures than RO and NF systems because of the permeation of sugars and minerals into the permeate stream.

1.3.4 Microfiltration

MF membranes also fractionate materials in the feed stream. Because MF membranes allow the passage of many larger components, such as smaller proteins, they are not as good at concentrating the feed stream compared with UF membranes. Operating pressures are the lowest of the four types of membranes and, in fact, MF membrane systems operated at higher pressures often lose their ability to permeate larger molecules and behave like UF membranes.

1.4 Theory of membrane transport

Although each of these membrane processes separates molecules of different sizes, the method of separation varies considerably. Generally, UF and MF membranes separate molecules based on size, shape and charge, whereas separation during RO and NF is based neither on size nor on shape.

1.4.1 Transport models

Two models, pore flow and solution diffusion, have been proposed for describing the movement of water and solutes into the permeate. The most significant difference between the two models is the size and permanence of membrane pores. In the pore flow model (Fig. 1.9), pressure-driven convective flow drives solutes through small pores or openings in the membrane. Solutes that are too large to move through the pores remain behind thereby resulting in a separation of solution components.

Separations in the solution diffusion model (Fig. 1.10) are due to differences in the solubility of solution components in the membrane and the rate at which the components will move across the membrane. Components flow from high to low pressure. Movement of individual components is not related and is determined by mobility, concentration and pressure gradients.

The size of the pores within a membrane probably determines the model that applies. A pore diameter of 5–10 Å probably represents the transition zone between the two models. It is very difficult to directly measure the size of pores within a membrane. Pore size often is inferred from an indirect technique, such as the size of molecules that will permeate the membrane. Given the limitations in determining membrane pore size, it is possible to
assign RO, NF, UF and MF membranes into three general categories of separation models. RO membranes, which have no visible pores, are thought to follow the solution diffusion model. Separation in UF and MF membranes is probably based on the pore flow model. NF membranes are in the transition zone between the two models and, therefore, probably use a combination of the two models.

1.4.2 Reverse osmosis/nanofiltration membranes

Pores are not discernible in RO/NF membranes; however, if pores are present, they are thought to be 1–5 Å in diameter. It is the structure of water within the membrane that is
important. The RO/NF membrane can be considered a water-swollen gel. Water is thought to move through the membrane based on diffusion and the ability of several water molecules to form a tetrahedral or ice-like structure through hydrogen bonding. One theory holds that water is absorbed into the voids of the RO/NF membrane where the water molecules form this tetrahedral structure. When pressure is applied, water on the retentate side of the membrane joins the structured water within the membrane, while water in the membrane on the permeate side is released into the permeate stream. Only those molecules or ions able to fit into this tetrahedral structure can cross the RO/NF membrane. Because ions are surrounded by a water shell, they do not readily fit into the ice-like structure and, therefore, do not cross into the permeate stream; however, small molecules, such as methanol, urea and lactic acid that exhibit hydrogen bonding, may be able to cross the membrane.

Based on this theory of RO/NF systems, water and solute passage across the membrane is not connected. Water flow into the permeate is based on pressure, whereas solute passage is based on the concentration gradient. An increase in pressure, therefore, will increase the flow of water across the membrane while solute passage remains unaffected and results in less solute in the permeate. Because each solute has its own electrochemical potential, the flux of each component is not related to other solutes. NF membranes are sometimes referred to as 'loose RO' membranes. In addition to the passage of compounds that resemble the structure of water, NF membranes also permit the passage of monovalent ions while rejecting polyvalent ions. Additional factors affecting the separating ability of both RO and NF membranes are discussed in Section 1.5.

### 1.4.3 Ultrafiltration/microfiltration membranes

Both UF and MF membranes separate compounds largely on size. Fundamentally, the membrane acts as a sieve by rejecting those molecules too large to fit through pores, while permitting smaller molecules to cross into the permeate stream. The change in pressure across the membrane is the driving force. Although molecular weight often is used to indicate the size of molecules retained by the membrane, other factors are important and are discussed in Section 1.5.

### 1.5 Factors affecting membrane separations

The methods of separation for RO and UF differ significantly; therefore, the factors affecting these processes also will differ. However, NF will more closely resemble RO, whereas MF is more similar to UF.

#### 1.5.1 Factors affecting reverse osmosis/nanofiltration separations

Permeability of components during RO or NF is not based on size, but instead depends on the ability of a compound to cross the membrane using a pressure-driven solute diffusion process. Non-electrolytes and electrolytes, therefore, will be affected by different factors. Pressure, temperature, concentration of components and the type of compound affect RO/NF separations. Permeability of solution components during RO/NF also is affected
by membrane composition. In general, as pressure increases the rate of water permeation increases. In turn, the concentration of dissolved solids in the permeate increases. A similar trend occurs with temperature during RO/NF. Increasing the temperature of the process increases the rate of permeation; however, the concentration of dissolved solids in the permeate also increases.

Permeation of components that typically do not cross a membrane is less affected by the presence and concentration of other solutes than components in the form of electrolytes. Conversely, electrolyte retention is affected by concentration and other solutes, with less retention of these components as their concentration increases. An example of the effect of membrane composition on permeability is illustrated by the membrane polymer cellulose acetate, which has a strong sorption of higher aliphatic alcohols and flavour compounds. Cellulose acetate is hydrophobic as are many of these compounds which, therefore, are drawn to the membrane as a result of being repelled by the water phase. Interactions between cellulose acetate and water also have a large affect on salt permeability during RO/NF. Increased interactions between cellulose acetate and water result in less water available for hydration of ions. This causes the salts to be less soluble and, therefore, more energy is required to hydrate these salts. Different amounts of energy needed to hydrate salts account for differences in permeability.

The amount of charge on the ions present in solution is very important to their permeability with NF membranes. NF membranes typically reject polyvalent ions while permitting the passage of most monovalent ions. The concentration of ions also is a factor with rejection of ions increasing as the concentration of ions increases in the solution.

1.5.2 Factors affecting ultrafiltration/microfiltration separations

Size, shape, deformability and hydrodynamic radius of a molecule are very important in determining whether a component is retained during UF/MF. Generally, linear, easily deformed molecules are less likely to be retained than highly structured, rigid molecules of equivalent molecular weight. Under pressure, the more flexible molecules can deform to fit through pores generally considered too small to allow their passage. Globular proteins, therefore, are used to define the molecular weight cut-off of a UF membrane since they are less likely to deform under pressure.

Membrane–molecule interactions may be more important than physical factors in determining the ability of a molecule to cross a UF membrane. An example of two materials having opposite effects is styrene and erythrocytes. Styrene, although small enough to cross UF membrane pores, will interact with pore walls and form agglomerates effectively blocking the pore. Erythrocytes, in contrast, are too large to pass through pores; however, their flexibility and lack of interactions with pore walls allow them to cross into the permeate.

Ionic strength and pH of the solution also influence the apparent size of a molecule. For example, a highly charged polyelectrolyte is more easily retained when in solution of lower ionic strength since these conditions give the molecule a larger effective radius and, therefore, limit passage through membrane pores.

The type of compound also is important. Acids, for example, are retained less readily than corresponding salts. Small organic molecules (alcohols, esters, aldehydes and ketones)
and small non-ionic inorganic molecules, such as ammonia, are more likely to cross a membrane than ionic compounds.

Composition of the feed material also influences permeability. An example is small molecular weight proteins, which are more likely to cross a membrane when in a pure solution than when ultrafiltered in the presence of compounds such as larger molecular weight proteins.

1.5.3 System parameters

System parameters, such as operating conditions and membrane composition can influence passage of a molecule. Operating parameters including temperature, feed velocity, pressure and feed concentration can alter permeation of a molecule. For example, use of higher pressures and lower feed velocities can result in the rejection of proteins by a MF membrane that otherwise would permeate the protein. The affects of system parameters on separations will be discussed in a later section.

1.6 General characteristics of membrane processes

Certain characteristics of a particular membrane are especially important since they influence the economic viability and possible applications of the membrane system. Pore size and retention can be used as an indication of the ability of a membrane to retain a certain size molecule. Flux and membrane life, in contrast, affect the economics of processing with membranes.

1.6.1 Retention and rejection

The terms retention and rejection may be used interchangeably depending on whether the component is desired or undesired in the retentate stream. Retention \( R \) can be defined in several ways, and one common definition is:

\[
R = \frac{c_f - c_p}{c_f} \tag{1.1}
\]

where \( C_f \) is the concentration of a component in the feed stream and \( C_p \) is the concentration of a component in the permeate stream.

If a component freely permeates the membrane, \( R \) will be near zero, whereas a completely retained component has an \( R \) value of 1, and expressed as a percentage would be equivalent to 100% retention. Components having a rejection or retention value of zero will be found in the same concentration in the permeate and feed streams.

An alternative system that measures membrane performance according to permeability \( P \) is as follows:

\[
P = 1 - R \tag{1.2}
\]

or

\[
P = 100 - R(\text{percentage}) \tag{1.3}
\]
Although easily defined, rejection characteristics of a membrane are affected by many factors, and there is no uniform set of conditions used by manufacturers to determine this property. Pore diameter, chemical composition of the membrane and interactions between the membrane and feed material are major factors influencing rejection characteristics of a membrane. Rejection values also change depending on operating conditions and during operation of a membrane system.

Rejection is one method of classifying a membrane and, typically, RO and NF membranes are rated according to their ability to reject sodium chloride or other salts. Because RO membranes are used extensively for water desalination, the ability of the membrane to limit the passage of sodium chloride into the permeate is extremely important in the production of water for human consumption. Even small amounts of salt in drinking water can affect taste of the water; therefore, RO membranes often are rated on salt rejection.

1.6.2 Pore size

Pore size is used by some membrane manufacturers to indicate the separating ability of a membrane. Units of microns (μm) are often used. Reference to pore size is typical of membranes in the MF range or to specific categories of membranes, such as ceramic and track etched.

1.6.3 Molecular weight cut-off

Manufacturers often use molecular weight cut-off to indicate the separating ability of a given membrane. Membranes with the same molecular weight cut-off, however, may not have the same retention for a compound since manufacturers use a variety of methods to determine molecular weight cut-off. Generally for a given molecular weight cut-off, 80% of the molecules of that molecular weight will be rejected. Furthermore, molecular weight cut-off is typically used for membranes in the UF range.

1.6.4 Flux

Flux is the amount of permeate produced in a given time period, and the term generally is given as a volume or mass per unit membrane per unit time. A unit, such as L m\(^{-2}\) h\(^{-1}\), is a possible volume flux unit, whereas kg m\(^{-2}\) h\(^{-1}\) is a mass flux unit.

Flux determines the area of membrane required to process a given amount of product to a certain concentration in a specific time period. The lower the flux for a given membrane the greater the membrane area required to process the same amount of product within a certain time than with a higher flux membrane. Flux, therefore, affects the economics of an operation, and is used as an indication of membrane fouling and cleaning adequacy. Factors influencing flux include pressure, feed velocity, temperature, viscosity and turbulence.
1.6.5 Concentration factor

Concentration factor (CF) expresses the degree of concentration for a feed and can be defined as:

\[ CF = \frac{\text{volume of original feed}}{\text{volume of final feed}} \]  \hspace{1cm} (1.4)

Concentration factor is often expressed as 1×, 2×, 3× etc. Volumes or weights may be used. An example would be 100 L of product that is processed to a final volume of 33 L:

\[ \frac{100 \text{ L}}{33 \text{ L}} = 3 \times \]

The ability to concentrate a feed material generally is determined by feed constituents, osmotic pressure and feed viscosity.

1.6.6 Membrane life

Although the cross-flow design allows reuse of a membrane, the operating life of a membrane is not indefinite. As membranes are used for processing, their characteristics flux and retention change, and it is the decreasing flux and retention with time that result in membrane replacement. The life of a membrane ends when membrane performance no longer meets specific performance criteria. In the case of a protein concentration operation, the loss of valuable protein into the permeate might be a criteria. In a water desalination system the criteria might be the permeation of excessive amounts of salt into the permeate, which is to be as drinking water.

Membrane life is influenced by many factors. Membrane composition is important, but often the processing and cleaning conditions have a greater influence. Composition, however, affects the resistance of a membrane to processing and cleaning treatments, and determines acceptable operating conditions.

1.7 Conclusion and future development

Although membranes have been known about for many decades, it has not been until relatively recently that technological improvements have permitted their widespread use by the food industry. Thus, the use of membrane processes for food applications can be considered a relatively new tool for the food processor. Knowledge about the theory of membrane systems operate will continue to increase as new applications for membrane systems are found.

Suggested literature
