# Lecture 6: Structure in Liquids

In this lecture, we discuss how to understand the origin of structure of liquids and also mathematically formalize how to characterize the structure. In the end, a connection to scattering experiments is provided, which offers us a way to actually see the structure in liquids.

## 6.1 Can Liquids Have Structure?

### 6.1.1 Thermal Energy vs Interaction Potential

Liquids appear disordered in contrast to crystalline solids, yet experiments and simulations show that they possess well-defined *short-range* structure. Understanding this requires examining the competition between the interaction potential between atoms and the thermal energy available in the system.

Atoms interact through a pair potential $V(r)$, which typically has a steep repulsive core at small $r$, an attractive well at intermediate distances, and decays at infinity. Examples include the Lennard–Jones potential and the Morse potential.

Thermal motion introduces energy of order $k_B T$ (in the canonical ensemble), which tends to disrupt order.

**Crystals (Low $T$):**

- Thermal energy $\ll$ well depth.

- Atoms remain inside the harmonic region (Phonons) around equilibrium, with little anharmonicity in oscillations

- Long-range order persists.

**Liquids (Intermediate $T$):**

- Thermal energy comparable to well depth.

- Atoms escape the harmonic region and reorganize.

- Short-range order persists; long-range order disappears.

**Gases (High $T$):**

- Thermal energy dominates interactions.

- Motion is essentially free, except occasional collisions.
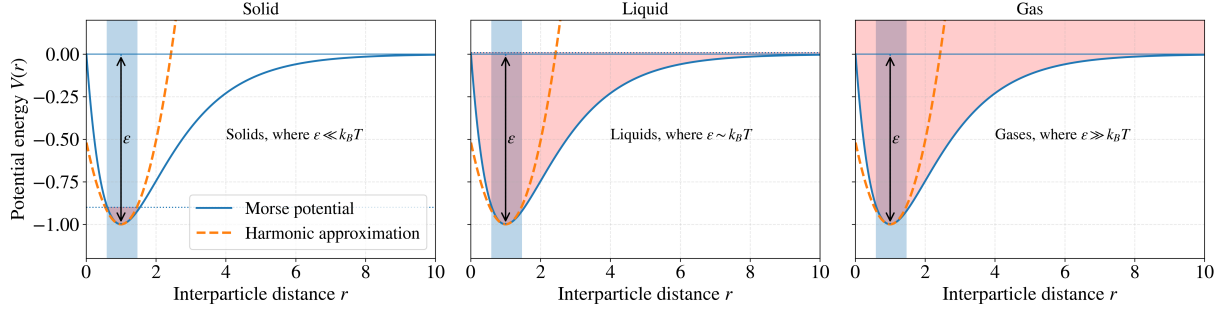
- No structure.

Figure 1: Interaction potential between particles (Morse potential) as a function of separation distance $r$ in different phases. The red shaded region corresponds to possible energy and separation distances accessible to particles given a thermal energy of $k_B T$. Here, $\epsilon$ is the typical depth of the potential well. The dotted lines indicate Harmonic approximation, with the blue shaded region highlighting the region valid with harmonic approximation.

### 6.1.2 Harmonic Approximation around Equilibrium

Near equilibrium distance $r_0$, we may expand

$$V(r) \approx V(r_0) + \frac{1}{2}k(r - r_0)^2.$$

In solids, $k_B T$ is too small to drive atoms outside this region; in liquids, thermal energy allows escape.

### 6.1.3 Short-Range Structure in Liquids

Even in the liquid, atoms still experience strong forces from their nearest neighbors. The repulsive and attractive parts of $V(r)$ continue to enforce a preferred separation $r_0$ on short length scales. As a result:

- Each atom typically has a "first-neighbor shell" of atoms located at distances close to $r_0$.

- There may be hints of second and third shells, although these are less sharply defined than in a crystal.

- The immediate environment of an atom resembles a *distorted, fluctuating* version of a local crystalline neighborhood.

At any given instant, if one could "freeze" the configuration of a liquid and examine a local region, one would often recognize fragments of crystalline-like motifs: near-regular polygons or polyhedra, characteristic angles, and preferred local distances. The difference from a crystal is that these motifs do not repeat periodically throughout the sample and are constantly changing in time.

Mathematically, one can say that the interatomic potential still imposes a characteristic length scale $r_0$ on the structure. This length scale remains visible in many observables (for example, in the position of the first peak of the pair correlation function or the static structure factor, which we will introduce in the next sections).

## 6.2 Characterizing Structure: Pair Distribution Function

### 6.2.1 Microscopic Number Density

In order to understand structure in liquids, we begin by mathematically defining a field called the microscopic number density field, $n(\mathbf{r})$.

$$n(\mathbf{r}) = \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j),$$

which provides the number density of particles at any arbitrary position $\mathbf{r}$ inside the system. In particular, the quantity $n(\mathbf{r})d\mathbf{r}$ gives the number of particles in a volume element $d\mathbf{r}$.

In a homogeneous uniform fluid, we have an average number density which can be calculated by considering an ensemble average

$$\langle n(\mathbf{r}) \rangle = \left\langle \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j) \right\rangle,$$

where $\langle \rangle$ represent averaging over all the thermal states of a canonical ensemble.

### 6.2.2 Translational Invariance from Pairwise Interactions

Before, we proceed to calculate the distribution function, we firstly consider the symmetry of the system we have. In most systems, an assumption that particles interact via an isotropic pair potential is reasonable,

$$V(\mathbf{r}_i - \mathbf{r}_j),$$

The potential depends only on the separation distances $\mathbf{r}_i - \mathbf{r}_j$ and don't depend explicitly on individual positions of the particles. A global translation $\mathbf{r}_i \rightarrow \mathbf{r}_i + \boldsymbol{\Delta R}$ leaves the potential unchanged:

$$V((\mathbf{r}_i + \boldsymbol{\Delta R}) - (\mathbf{r}_j + \boldsymbol{\Delta R})) = V(\mathbf{r}_i - \mathbf{r}_j).$$

Hence the equilibrium ensemble is **translationally invariant**, and the structure is essentially dependent on the potential, we can argue that the mathematical function describing structure depends only on separation vectors, not on absolute positions. This is a very important argument, as you will see later that, using this we can reduce our correlation functions into a function of separation distance, $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j$

### 6.2.3 Density–Density Correlation Function

Now, we formally define the two-point density correlation function

$$G(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{N} \langle n(\mathbf{r}_1)n(\mathbf{r}_2) \rangle.$$

which quantifies the correlation between number density at position $\mathbf{r}_1$ and number density at position $\mathbf{r}_2$. Substituting the microscopic density,

$$G(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} \delta(\mathbf{r}_1 - \mathbf{r}_i)\delta(\mathbf{r}_2 - \mathbf{r}_j) \right\rangle.$$

Translational invariance, as discussed in the previous subsection, implies

$$G(\mathbf{r}_1, \mathbf{r}_2) = G(\mathbf{r}_2 - \mathbf{r}_1) \equiv G(\mathbf{r}).$$

Since the origin of the system is arbitrary, it is possible to eliminate $\mathbf{r}_1$ completely by integrating over $\mathbf{r}_1$:

$$G(\mathbf{r}) = \int d\mathbf{r}_1 \, G(\mathbf{r}_1, \mathbf{r}_1 + \mathbf{r})$$

$$= \frac{1}{N} \int d\mathbf{r}_1 \left\langle \sum_{i,j} \delta(\mathbf{r}_1 - \mathbf{r}_i)\delta(\mathbf{r}_1 + \mathbf{r} - \mathbf{r}_j) \right\rangle$$

$$= \frac{1}{N} \left\langle \sum_{i,j} \delta(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)) \right\rangle.$$

Thus $G(\mathbf{r})$ is proportional to the probability of finding a pair of particles separated by $\mathbf{r}$. Intuitively, it picks up a value whenever $\mathbf{r}$ becomes equal to $\mathbf{r}_j - \mathbf{r}_i$ (i.e., whenever it finds a two particles, $i$ & $j$ separated by a distance $r$). The above equation considers correlation running over all particle pairs. However, a trivial case for the correlation of particle with itself (which is referred to as the self-correlation), can be ignored when trying to understand the structure. Therefore, we split the double sum into $i = j$ (self) and $i \neq j$ (distinct):

$$G(\mathbf{r}) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \delta(\mathbf{r} - (\mathbf{r}_i - \mathbf{r}_i)) + \sum_{i \neq j} \delta(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)) \right\rangle$$

$$= \frac{1}{N} \left\langle \sum_{i=1}^{N} \delta(\mathbf{r}) \right\rangle + \frac{1}{N} \left\langle \sum_{i \neq j} \delta(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)) \right\rangle$$

$$= \delta(\mathbf{r}) + \frac{1}{N} \left\langle \sum_{i \neq j} \delta(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_i)) \right\rangle.$$

The above equation can be further reduced for a homogeneous system. In this case, the local environment around each particle is statistically identical. Therefore, we may fix one particle (say $i = 1$) and multiply by $N$, do carry out the summation of $i$,

$$\sum_{i \neq j}(\cdots) \longrightarrow N \sum_{j=2}^{N}(\cdots \text{ with } i = 1).$$

Hence

$$G(\mathbf{r}) = \delta(\mathbf{r}) + \langle n \rangle \left\langle \sum_{j=2}^{N} \delta(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_1)) \right\rangle.$$

which is rewritten as,

$$G(\mathbf{r}) = \delta(\mathbf{r}) + ng(r),$$

which defines the **pair distribution function** $g(\mathbf{r})$ as

$$g(\mathbf{r}) = \frac{1}{\langle n \rangle} \left\langle \sum_{j \neq 1} \delta(\mathbf{r} - (\mathbf{r}_j - \mathbf{r}_1)) \right\rangle.$$

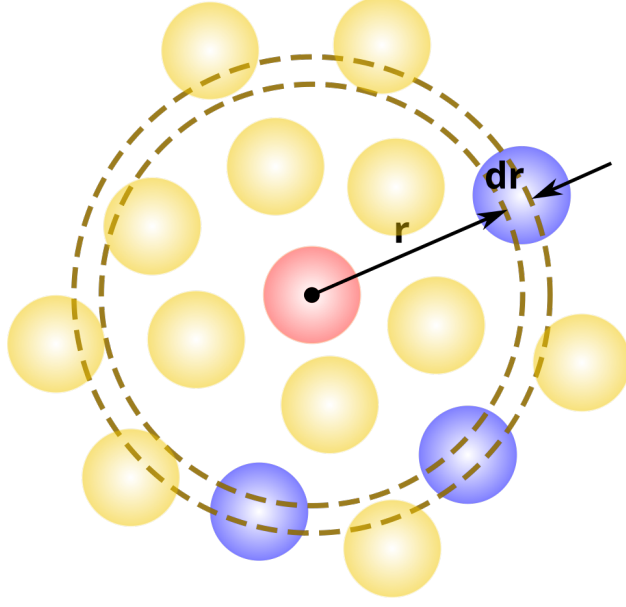Therefore, the pair distribution function can be interpreted as a ratio between

4

Figure 2: Schematic indicating how radial distribution function is calculated.

$$\text{number density of particles at position } \mathbf{r}$$
$$\text{and}$$
$$\text{average number density of particles in the system.}$$

### 6.2.4 Radial Distribution Function

Considering the spherical symmetry of interactions between particles, we can further reduce the dependence of pair-distribution function $g(\mathbf{r})$, to only depend on the magnitude of the separation between particles. This can be accomplished by integrating over the angular coordinates, $\theta$ and $\phi$, considering a spherical polar coordinate system for $\mathbf{r} = (r, \theta, \phi)$. Since the pair distribution function doesn't depend on $\theta$ and $\phi$, we have

$$g(r) = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin(\theta) \, r^2 g(\mathbf{r}) = 4\pi r^2 g(\mathbf{r})$$

where g(r) is referred to as the radial distribution function.

From here, we can calculate the average number of particles in a spherical shell of radius $r$ and thickness $dr$ around a reference particle is

$$dN(r) = n \, g(r) \, 4\pi r^2 \, dr.$$

Typically, for any system the follow the limits,

$$g(r \to 0) = 0 \quad \text{(no self-overlap)}, \qquad g(r \to \infty) = 1 \quad \text{(loss of correlations)}.$$
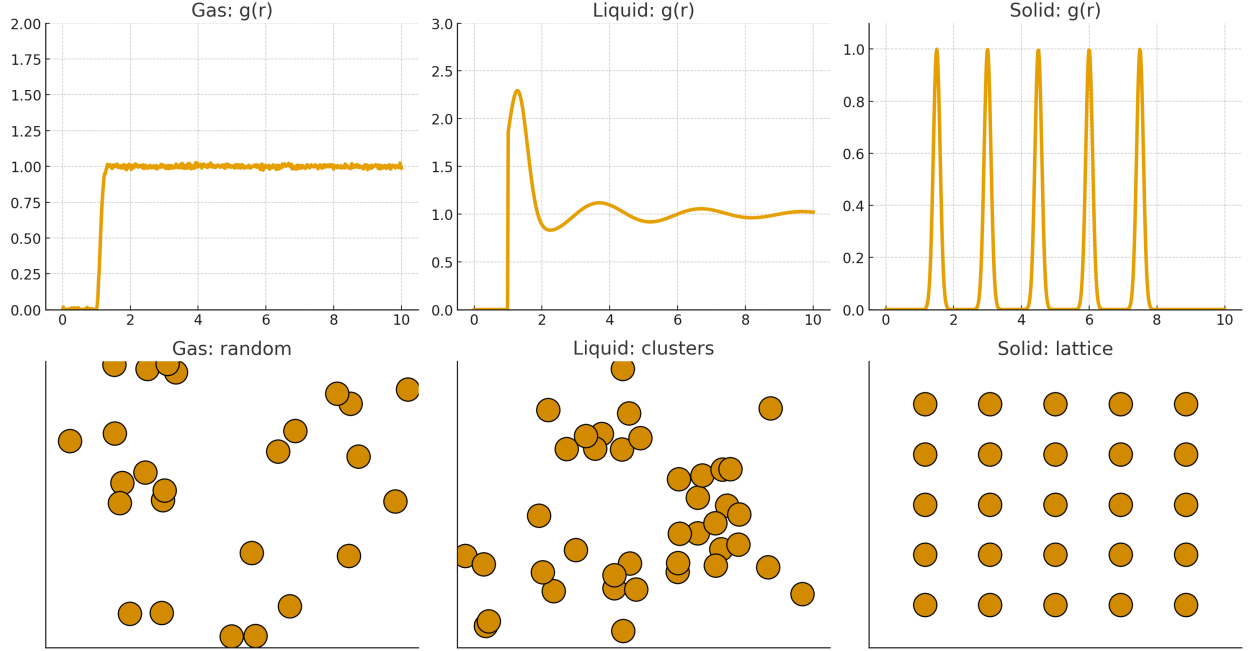
Figure 3: Typical radial distribution function (RDF) for different gases, liquids and solids along with representative arrangement of atoms in these states.

## 6.3 Structure Factor and Its Connection to $g(r)$

### 6.3.1 Microscopic Density in Fourier Space

We now move to reciprocal space. The Fourier transform of the microscopic density is defined as

$$n(\mathbf{q}) = \int d\mathbf{r} \, e^{-i\mathbf{q}\cdot\mathbf{r}} \, n(\mathbf{r}).$$

Inserting $n(\mathbf{r}) = \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j)$ gives

$$n(\mathbf{q}) = \sum_{j=1}^{N} e^{-i\mathbf{q}\cdot\mathbf{r}_j}.$$

### 6.3.2 Definition of the Structure Factor

The (static) structure factor is defined by

$$S(\mathbf{q}) = \frac{1}{N} \langle n(\mathbf{q}) \, n(-\mathbf{q}) \rangle.$$

Using $n(-\mathbf{q}) = \sum_{j=1}^{N} e^{i\mathbf{q}\cdot\mathbf{r}_j}$, we obtain

$$S(\mathbf{q}) = \frac{1}{N} \left\langle \sum_{i=1}^{N} e^{-i\mathbf{q}\cdot\mathbf{r}_i} \sum_{j=1}^{N} e^{i\mathbf{q}\cdot\mathbf{r}_j} \right\rangle$$

$$= \frac{1}{N} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N} e^{-i\mathbf{q}\cdot(\mathbf{r}_i - \mathbf{r}_j)} \right\rangle.$$

6

This is the starting point for connecting $S(\mathbf{q})$ to real-space correlations.

### 6.3.3   Introducing Delta Functions to Connect with $G(\mathbf{r}_1, \mathbf{r}_2)$

We now write the exponential factor using an integral representation with Dirac delta functions:

$$e^{-i\mathbf{q}\cdot(\mathbf{r}_i - \mathbf{r}_j)} = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \, e^{-i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \, \delta(\mathbf{r}_1 - \mathbf{r}_i) \, \delta(\mathbf{r}_2 - \mathbf{r}_j).$$

Substituting this back into $S(\mathbf{q})$:

$$S(\mathbf{q}) = \frac{1}{N} \left\langle \sum_{i,j} \int d\mathbf{r}_1 d\mathbf{r}_2 \, e^{-i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \, \delta(\mathbf{r}_1 - \mathbf{r}_i) \, \delta(\mathbf{r}_2 - \mathbf{r}_j) \right\rangle$$

$$= \int d\mathbf{r}_1 d\mathbf{r}_2 \, e^{-i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \, \frac{1}{N} \left\langle \sum_{i,j} \delta(\mathbf{r}_1 - \mathbf{r}_i) \, \delta(\mathbf{r}_2 - \mathbf{r}_j) \right\rangle.$$

But by definition of the density–density correlation function,

$$G(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{N} \langle n(\mathbf{r}_1) n(\mathbf{r}_2) \rangle = \frac{1}{N} \left\langle \sum_{i,j} \delta(\mathbf{r}_1 - \mathbf{r}_i) \, \delta(\mathbf{r}_2 - \mathbf{r}_j) \right\rangle.$$

Therefore,

$$S(\mathbf{q}) = \int d\mathbf{r}_1 d\mathbf{r}_2 \, e^{-i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \, G(\mathbf{r}_1, \mathbf{r}_2).$$

### 6.3.4   Using Translational Invariance: $S(\mathbf{q})$ as Fourier Transform of $G(\mathbf{r})$

For a translationally invariant system,

$$G(\mathbf{r}_1, \mathbf{r}_2) = G(\mathbf{r}_2 - \mathbf{r}_1) \equiv G(\mathbf{r}), \quad \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1.$$

We change variables from $(\mathbf{r}_1, \mathbf{r}_2)$ to $(\mathbf{r}, \mathbf{R})$, where

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1, \qquad \mathbf{R} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2},$$

with Jacobian equal to 1. Then

$$\int d\mathbf{r}_1 d\mathbf{r}_2 = \int d\mathbf{r} \, d\mathbf{R}.$$

Using $G(\mathbf{r}_1, \mathbf{r}_2) = G(\mathbf{r})$, the structure factor becomes

$$S(\mathbf{q}) = \int d\mathbf{r} \, d\mathbf{R} \, e^{-i\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} G(\mathbf{r})$$

$$= \int d\mathbf{r} \, d\mathbf{R} \, e^{-i\mathbf{q}\cdot(-\mathbf{r})} G(\mathbf{r})$$

$$= \int d\mathbf{r} \, e^{i\mathbf{q}\cdot\mathbf{r}} G(\mathbf{r}) \quad \underbrace{\int d\mathbf{R}}_{\text{system volume factor}}.$$
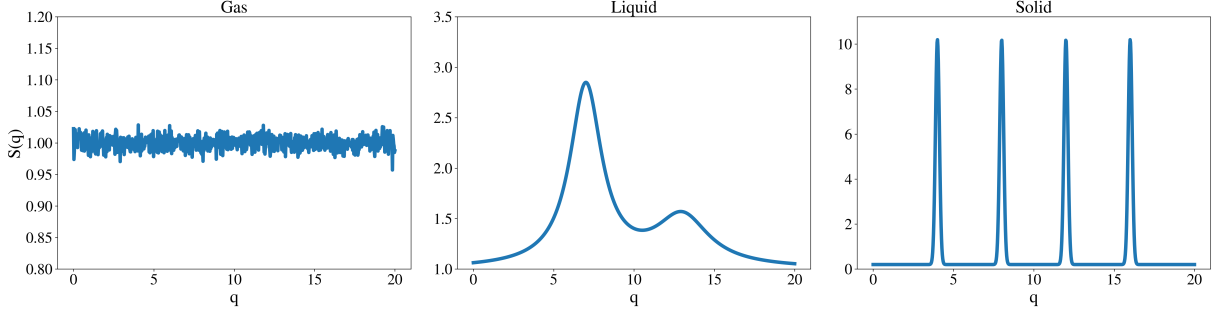
Figure 4: Typical structure factor, $S(q)$ for different systems, depending on their structure and ordering. These can be observed experimentally from X-Ray or neutron scattering.

The integration over $\mathbf{R}$ simply gives a factor proportional to the system volume; in practice, we absorb this into the definition of $G(\mathbf{r})$ when working with intensive quantities. Consistent with our earlier normalization (where $G(\mathbf{r})$ was already defined as an intensive correlator), we write simply

$$S(\mathbf{q}) = \int d\mathbf{r}\, e^{-i\mathbf{q}\cdot\mathbf{r}} G(\mathbf{r}),$$

up to the standard convention that $S(\mathbf{q})$ is intensive.

Using the decomposition

$$G(\mathbf{r}) = \delta(\mathbf{r}) + \langle n \rangle\, g(r),$$

we obtain

$$S(\mathbf{q}) = \int d\mathbf{r}\, e^{-i\mathbf{q}\cdot\mathbf{r}} \left[ \delta(\mathbf{r}) + n g(r) \right]$$

$$= 1 + \langle n \rangle \int d\mathbf{r}\, e^{-i\mathbf{q}\cdot\mathbf{r}} g(r).$$

This form is convenient, but often we prefer to express $S(\mathbf{q})$ in terms of $g(r) - 1$, which isolates *correlations* beyond the ideal gas.

For $\mathbf{q} \neq \mathbf{0}$, one writes

$$S(\mathbf{q}) = 1 + \langle n \rangle \int d\mathbf{r}\, e^{-i\mathbf{q}\cdot\mathbf{r}} [g(r) - 1].$$

(The contribution from $\int e^{-i\mathbf{q}\cdot\mathbf{r}}\, d\mathbf{r}$ is proportional to a Dirac delta at $\mathbf{q} = \mathbf{0}$ and does not affect $S(\mathbf{q})$ at finite $\mathbf{q}$.)

### 6.3.5   Isotropic systems

For isotropic fluids, $g(r)$ depends only on the magnitude $r$ and $S(\mathbf{q})$ depends only on $q = |\mathbf{q}|$. The angular integration can be performed explicitly:

$$\int d\mathbf{r}\, e^{-i\mathbf{q}\cdot\mathbf{r}} [g(r) - 1] = 4\pi \int_0^\infty r^2 [g(r) - 1] \frac{\sin(qr)}{qr}\, dr.$$

Thus the final relation between $S(q)$ and $g(r)$ is

$$S(q) = 1 + 4\pi \langle n \rangle \int_0^\infty r^2 [g(r) - 1] \frac{\sin(qr)}{qr}\, dr.$$

8

## 6.4   Summary

- $g(r)$ describes how particles are arranged around a reference particle in real space.

- $S(q)$ encodes the same information in reciprocal space and is directly accessible in scattering experiments.

- The connection between them is given by a Fourier transform:

$$S(q) = 1 + 4\pi \langle n \rangle \int_0^\infty r^2 [g(r) - 1] \frac{\sin(qr)}{qr} \, dr.$$

- Limiting cases:

  - Ideal gas: $g(r) = 1 \Rightarrow S(q) = 1$.
  - Liquid: damped oscillations in $g(r)$ and corresponding broad peaks in $S(q)$.
  - Crystal: Bragg peaks in $S(q)$ corresponding to periodic structure in real space.