The number of molecules in a microscopic probe volume within a dense liquid is often Gaussian distributed to a *very* good approximation, as you demonstrated for a hard sphere fluid in the previous assignment. We will eventually exploit this remarkable observation to construct a tractable theory for liquid structure. To do so, however, we need to switch attention to continuous variables that are defined everywhere in space, e.g., the microscopic density field

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

where $\delta(\mathbf{r})$ is the Dirac delta function and $\mathbf{r}_i$ denotes the position of particle i. In this assignment you will examine the statistics of these kinds of fields.

For a specific molecular configuration, the density field $\rho(\mathbf{r})$ is a scattered collection of delta functions; in other words, it is unpleasantly singular. Here we focus instead on the more politely behaved Fourier components,

$$\hat{\rho}(\mathbf{k}) = \int_V d\mathbf{r}\, \rho(\mathbf{r})\, e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{1}$$

Eq. 1 acknowledges explicitly the finite integration domain $V = L^3$ of systems accessible to simulation. For a periodically replicated system, the volume V refers to a central simulation cell containing one image of each particle.

When the spatial domain is finite, the rules of Fourier transforms change a little from their familiar form. For a function $f(\mathbf{r})$ and its transform $\hat{f}(\mathbf{k}) = \int_V d\mathbf{r}\, f(\mathbf{r})\, e^{i\mathbf{k}\cdot\mathbf{r}}$:

- The set of meaningful wavevectors $\mathbf{k}$ is *discrete*:

  $$\mathbf{k} = \frac{2\pi}{L}\mathbf{n},$$

  where $\mathbf{n}$ is a three-dimensional vector with integer components (i.e., $n_x = 0, \pm 1, \pm 2, \ldots$).

- The orthogonality condition

  $$\int_V d\mathbf{r}\, e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} = V\,\delta_{\mathbf{k},\mathbf{k}'}$$

  involves the Kronecker delta ($\delta_{\mathbf{a},\mathbf{b}} = 1$ if the vectors $\mathbf{a}$ and $\mathbf{b}$ are identical; otherwise it vanishes) rather than the Dirac delta function.

- The inverse transform involves a sum, rather than an integral, over the discrete set of wave vectors:

  $$f(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{k}} \hat{f}(\mathbf{k})\, e^{-i\mathbf{k}\cdot\mathbf{r}}$$

1. **General properties of density field Fourier components**

   Macroscopic liquids possess translational symmetry, a property preserved in simulations by the use of periodic boundary conditions. For such systems, show that:

   (i) The ensemble average of $\hat{\rho}(\mathbf{k})$ vanishes for all non-zero wavevectors:

   $$\langle \hat{\rho}(\mathbf{k}) \rangle = N \delta_{\mathbf{k},0}.$$

   [**Hint:** When calculating $\langle \rho(\mathbf{r}) \rangle$, note that the probability distribution $P(\mathbf{r}) = \langle \delta(\mathbf{r} - \mathbf{r}_i) \rangle$ of the position of particle i must be normalized and independent of $\mathbf{r}$.]

   (ii) The correlation function $\langle \hat{\rho}(\mathbf{k}) \hat{\rho}(\mathbf{k}') \rangle$ vanishes unless $\mathbf{k} + \mathbf{k}' = 0$:

   $$\langle \hat{\rho}(\mathbf{k}) \hat{\rho}(\mathbf{k}') \rangle = N \, \hat{S}(k) \, \delta_{\mathbf{k},-\mathbf{k}'},$$

   where $\hat{S}(k)$ is the Fourier transform of

   $$S(r) = \frac{\langle \rho(0) \rho(\mathbf{r}) \rangle}{\bar{\rho}},$$

   and $\bar{\rho} = N/V$ is the bulk density.

   [**Hint:** Note that $f(\mathbf{r} - \mathbf{r}') = \frac{1}{V} \sum_{\mathbf{k}} \hat{f}(k) e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}$.]

   Also explain why $S(r)$ depends only on the scalar distance $r = |\mathbf{r}|$ and not on the direction in which the vector $\mathbf{r}$ points.

   (iii) The real and imaginary parts of $\hat{\rho}(\mathbf{k})$ are statistically equivalent (i.e., they are distributed according to the same probability distribution). Begin by showing that

   $$\mathrm{Re}[\hat{\rho}(\mathbf{k})] = \sum_{j=1}^{N} \cos(\mathbf{k} \cdot \mathbf{r}_j), \qquad \mathrm{Im}[\hat{\rho}(\mathbf{k})] = \sum_{j=1}^{N} \sin(\mathbf{k} \cdot \mathbf{r}_j),$$

   and then consider the consequence of moving the system's arbitrarily chosen origin $\mathbf{r} = 0$ by a displacement $\Delta\mathbf{r}$.

   [**Hints:** When the origin is moved by $\Delta\mathbf{r}$, the position of each particle i changes from $\mathbf{r}_i$ to $\mathbf{r}_i - \Delta\mathbf{r}$. The displacement $\Delta\mathbf{r} = \frac{\pi}{2k}\hat{\mathbf{k}}$ should be particularly revealing.]

   (iv) The real and imaginary parts of $\hat{\rho}(\mathbf{k})$ are symmetrically distributed about zero.

   [**Hint:** Shifting the system's origin should again be helpful.]

   (v) $\hat{S}(k)$ can be determined from $\left\langle (\mathrm{Re}[\hat{\rho}(\mathbf{k})])^2 \right\rangle$ alone.

2. **Computing density field statistics by simulation**

   Extend the hard sphere fluid Monte Carlo simulation you developed for Problem Set 1 to collect statistics of $\hat{\rho}(\mathbf{k})$. As before, focus on a three-dimensional fluid at density $N/V = 0.5d^{-3}$. For this assignment, you can obtain useful results with much smaller systems, e.g., comprising $N = 125$ particles. You are welcome to build from the code provided in solutions to the first problem set.

   Given the properties described in question 1, it is sufficient to focus solely on the real part of $\hat{\rho}(\mathbf{k})$. Since the fluid is isotropic, it is also sufficient to focus on wavevectors pointing in a particular direction (let's pick a wavevector $k\hat{\mathbf{x}}$, with magnitude k and pointing in the x-direction). To simplify notation (and to work with intensive quantities), define

   $$\Re_k = \frac{\mathrm{Re}[\hat{\rho}(k\hat{\mathbf{x}})]}{\sqrt{N}}.$$

Note: While symmetry allows us to restrict attention to just the real part and a single direction, you would be wise to exploit the statistical equivalence of other quantities when collecting data. For example, since distributions of $\Re_k$ and its imaginary counterpart are identical, you can add their fluctuating values to the same histogram.

(i) Compute $\langle \Re_k \rangle$ for a range of wavevectors with $k = 2\pi/L, 4\pi/L, 6\pi/L, \ldots$ up to $k \approx 10\pi/d$ (where d is the hard sphere diameter). Demonstrate that $\langle \Re_k \rangle = 0$ at each wavevector to within some reasonable estimate of your statistical uncertainty.

(ii) For the same set of wavevectors, compute $\langle (\Re_k)^2 \rangle$. Plot $\hat{S}(k)$ as a function of k.

(iii) Your result for $\hat{S}(k)$ should exhibit a peak near $k = 2\pi/d$. What aspect of liquid structure does this peak reflect? (It might help to think about what $\hat{S}(k)$ should look like for a crystalline solid.)

(iv) For wavevectors with $k = \pi/d$, $k = 2\pi/d$, and $k = 3\pi/d$, make histograms $P(\Re_k)$ of values of $\Re_k$ (exploiting any symmetries you like). Plot your result for the logarithm of this distribution $\ln[P(\Re_k)]$ as a function of $\Re_k$. Alongside your simulation results, plot a Gaussian distribution with the same mean and variance. Comment on the quality of agreement.

3. **Little mathematical exercises**

(i) Calculate the characteristic function

$$\hat{p}(k) = \sum_{n=0}^{\infty} p(n)\, e^{ikn}$$

for a Poisson distribution

$$p(n) = \frac{\langle n \rangle^n \, e^{-\langle n \rangle}}{n!}.$$

By differentiating $\ln \hat{p}(k)$, show that *all* cumulants C_m of this distribution have the same value:

$$C_m = \langle n \rangle.$$

(ii) Show that the exponential average $\langle e^{ax} \rangle$ for a Gaussian random variable x evaluates to

$$\langle e^{ax} \rangle = (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{\infty} dx \, \exp\left[-\frac{(x - x_0)^2}{2\sigma^2} + ax \right] = e^{ax_0} e^{a^2\sigma^2/2}.$$

You may make use of the basic result of Gaussian integration:

$$\int_{-\infty}^{\infty} dx \, \exp\left[-\frac{x^2}{2\sigma^2} \right] = \sqrt{2\pi\sigma^2}$$