Image Formation

Spatial Frequency and Transfer Function

- consider thin TEM specimen
  - columns of atoms, where the electrostatic potential is higher than in vacuum
  - electrons accelerate when entering the specimen
  - wavelength $\lambda$ decreases

- interaction of a plane wave $\psi_0$ with the specimen merely introduces a (spatially varying!) phase shift $\varphi_s[r]$, such that
  \[ \psi_e = \psi_0 \exp \left( i \varphi_s[r] \right) \]

- imaging with an ideal lens yields an image with the intensity distribution
  \[ I[r] = |F^{-1}\left[F\left[\psi_0 \exp \left( i \varphi_s[r] \right)\right]\right]|^2 = |\psi_0 \exp \left( i \varphi_s[r] \right)|^2 = |\psi_0|^2 \]
→ no contrast!
     (compare explanation of “phase contrast”)

• “real” lens, however:
  – finite aperture → truncates spectrum in the focal plane
  – lens aberrations cause additional phase shifts
  → “distortion” of the Fourier spectrum prior to inverse Fourier transformation
  → “distortion” of image wave function $\psi_i$

• analogy: audio amplifier
  – amplifier has particular transfer characteristics
  – different audio frequencies (basses, trebles) are transferred with different quality

• HRTEM:
  – consider crystal:
    periodic repetition of a motive in space
  – period: $a$

The repeat rate $a^{-1}$ of a motive that occurs with a period $a$ in space is denoted as

spatial frequency.

Dimension: m$^{-1}$.

– special case: lattice planes of a crystal, parallel to the primary beam, plane spacing $d$
Bragg condition for small diffraction angles:

\[ \lambda = 2d \sin[\theta_B] \approx 2d \cdot \theta_B = d \cdot \theta \]

\[ \Rightarrow \text{diffraction angle } \theta \text{ is proportional to the spatial frequency } d^{-1}: \]

\[ \theta = \lambda \cdot d^{-1} \]

- the spatial frequency \( d^{-1} \) corresponds to the length of the scattering vector \( q = k - k_0 \), which lies approximately normal to the primary beam:

\[ d^{-1} = \frac{\theta}{\lambda} = ||k - k_0|| = q \]

- “spatial frequency” corresponds to audio frequency in the example of an audio amplifier
- microscope (objective lens) has particular “transfer properties”
- different spatial frequencies (“basses”, “trebles”) are transferred with different quality
- note: contrast is obtained only because the transfer is not ideal

• mathematical description of the transfer properties: transfer function

Example: Effect of a Limited Aperture

• circular aperture in the focal plane limits the spectrum of plane waves with wave vectors \( k \) that can contribute to the image

• aperture function:

\[ A[k] = \begin{cases} 1 & k_x, k_y \leq k_a \\ 0 & \text{otherwise} \end{cases} \]

\( k_x, k_y \): components of the wave vector normal to the primary beam.
• the wave at the exit surface of the specimen is

\[ \psi_e = \psi_0 \exp(i \varphi_s[r]) \]

• instead of

\[ \psi_f[k] = F[\psi_0 \exp(i \varphi_s[r])] \]

however, the wave function in the focal plane is now

\[ \psi_f[k] = A[k] \cdot F[\psi_0 \exp(i \varphi_s[r])] \]

• under these conditions, inverse Fourier transformation yields

\[ \psi_i[r] = F^{-1}[A[k] \cdot F[\psi_0 \exp(i \varphi_s[r])]] \]

\[ = F^{-1}[A[k]] \ast F^{-1}[F[\psi_0 \exp(i \varphi_s[r])]] \]

→ the wave function \( \psi_i \) in the image plane is convoluted (“spread”) with the Fourier transform of the aperture function

• concrete example: \( A[k] \) and \( F^{-1}[A[k]] \) for different aperture sizes \( k_a \):

![Diagram showing convolution of wave functions](image-url)
the “narrower” $A[k]$, the “broader” $F^{-1}[A[k]]$ and the larger the discrepancy between $\psi_i$ (image wave function) and $\psi_c$ (object wave function)

Phase Shifts Introduced by the Objective Lens

Spherical Aberration

• spherical aberration: the lens refracts rays far from the optic axis too strongly (compared with “paraxial” rays)

⇒ in the image plane the ray far from the axis arrive at some distance $\Delta r$ away from the ideal image point

• relation between $\Delta r$ and angle $\theta$, at which the rays travel with respect to the optic axis on the object side:

$$\Delta r = M \cdot C_s \theta^3$$
→ the constant proportionality factor $C_s$ characterizes the degree of spherical aberration

**Defocusing**

- two possibilities:
  - change object distance $g$ by $\Delta g$
  - change focal length $f$ of the lens by $\Delta f$
Transfer Function

- spherical aberration and defocusing of a lens generate a phase shift of

\[
\chi[\theta] = \chi_s[\theta] + \chi_g[\theta] + \chi_f[\theta]
\]

\[
= \frac{2\pi}{\lambda} \cdot \left[ \frac{1}{4} C_s \frac{R^4}{f^4} - \frac{1}{2} (\Delta f - \Delta g) \frac{R^2}{f^2} \right]
\]

- \(\Delta f\) and \(-\Delta g\) have the same effect, thus we always consider them together and denote \(\Delta f - \Delta g\) as “underfocus” \(\xi\)

- recalling \(R/f = \theta\) we obtain for the phase shift introduced by the lens

\[
\chi[\theta] = \frac{\pi}{2\lambda} C_s \theta^4 - \frac{\pi}{\lambda} \xi \theta^2
\]

- expressing \(\theta\) by the spatial frequency \(q = \theta/\lambda\) that diffracts at \(\theta\) (corresponds to the length of the scattering vector) yields the

\[
\text{Phase shift introduced by the objective lens:}
\]

\[
\chi[q] = \pi C_s \lambda^3 q^4 / 2 - \pi \xi \lambda q^2
\]

Example

- JEM 4000 EX (JEOL)
  - accelerating voltage: 400 kV
  - de Broglie wavelength: 0.001644 nm
  - spherical aberration: \(C_s = 1.0 \text{ mm} = 10^6 \text{ nm}\)
phase shift $\chi[q]$ for spatial frequencies up to 10 nm$^{-1}$ (spacings down to 0.1 nm) and different defoci $\zeta$:

- the minimum of $\chi[q]$ occurs where the first derivative of $\chi[q]$ with respect to $q$ vanishes:
  \[ 2\pi C_s \lambda^3 q^3 - 2\pi \zeta \lambda q = 0 \]

- the positive, non-vanishing solution for $q$ is
\[ q = \sqrt{\frac{\xi}{C_s \lambda^2}} \]

- the phase shift at the minimum amounts to

\[
\frac{\pi C_s \lambda^3}{2C_s^2 \lambda^4} \frac{\xi^2}{2C_s \lambda^2} - \pi \xi \frac{\xi}{C_s \lambda^2} = -\frac{\pi}{2} \frac{\xi^2}{C_s \lambda}
\]

\( \Rightarrow \xi = \sqrt{C_s \lambda} \approx 40 \text{ nm} \) approximately yields the “ideal” phase shift of \(-\pi/2\) over an extended interval of spatial frequencies (“Scherzer focus”)

- Ansatz for the mathematical describing the effect of limited aperture and phase shifts introduced by spherical aberration and defocusing:

  \[
  T[q] = A[q] \cdot \exp\left[i \chi[q]\right]
  \]

  \( q = |q| = \sqrt{u^2 + v^2} \): length of the scattering vector;
  \( A[q] = A[k] \): aperture function, \( A[q] = 1 \) for \((u^2 + v^2) < r^2\), 0 otherwise;
  \( \chi[q] \): phase shift introduced by the lens.

- with this definition the wave function in the image plane is

  \[
  \psi_i[x, y] = F^{-1}[T[q] \cdot F[\psi_e[x, y]]]
  \]

- assumption:
  completely coherent illumination
  (coherence \(\leftrightarrow\) potential for interference)
Image of a “Weak Phase Object”

- assumption: object (specimen) so thin that no absorption occurs – pure “phase object”

\[
\begin{align*}
V_t[x, y] := \int_{-t/2}^{t/2} V[x, y, z] \, dz
\end{align*}
\]

- phase shift after transmission through a layer of infinitesimal thickness \(dz\) (compared to wave through vacuum):

\[
d \varphi = \sigma \, \Delta \phi[x, y, z] \, dz
\]

\(\Delta \phi[x, y, z]\): variation of the electrostatic potential;
\(\sigma\): “interaction” constant \(2\pi(E_0 + eU)/(\lambda U(2E_0 + eU))^{-1}\).

- phase shift at location \((x, y)\) at the exit surface of the specimen:

\[
\varphi[x, y] = \sigma \int_{-t/2}^{t/2} \Delta \phi[x, y, z] \, dz =: \sigma \, \phi_t[x, y]
\]

\(\phi_t\): projected variation of the electrostatic potential – “projected potential”.

- thus the (pure) phase object does nothing to a plane wave \(\psi_0 = 1\) but advance its phase (the phase shift depends on the spatial coordinates):

\[
\psi_e[x, y] = \psi_0 \, \text{Exp}[-i \sigma \phi_t[x, y]] = \text{Exp}[-i \sigma \phi_t[x, y]]
\]
• additional simplifying assumptions:
  
  – very thin object
    (a single atom of uranium would be too thick!)

  ⇒ at every location \((x,y)\) the phase shift is small compared to unity

  ⇒ approximate exponential function linearly:

\[
\text{Exp}[-i \sigma \phi_l(x, y)] = 1 - i \sigma \phi_l(x, y)
\]

• image wave function of this (unrealistic!) object:

\[
\psi_i(x, y) = F^{-1}[T[q] \cdot F[1 - i \sigma \phi_l(x, y)]]
\]

\[
= F^{-1}[T[q] \cdot (\delta[q] - i\sigma F[\phi_l(x, y)])]
\]

• applying the convolution theorem

\[
F^{-1}[f[u] \cdot g[u]] = F^{-1}[f[u]] \ast F^{-1}[g[u]]
\]

yields

\[
\psi_i(x, y) = T[0] - i\sigma F^{-1}[T[q] \cdot F[\phi_l(x, y)]]
\]

\[
= 1 - i\sigma \phi_l(x, y) \ast F^{-1}[T[q]]
\]

• the image of every object point is spread with \(t[x,y] := F^{-1}[T[q]]\)

• \(t[x,y]\) is denoted as “point spread” function or “impulse response” function

• intensity distribution in the image plane?

• split \(t[x,y]\) into real and imaginary part:
\[ t[x, y] = c[x, y] + i s[x, y] \]

- since convolution is a linear operation, the wave function in the image plane becomes
\[ \psi_1[x, y] = 1 - i \alpha \phi_t[x, y] * c[x, y] + \alpha \phi_t[x, y] * s[x, y] \]

- intensity:
\[ I[x, y] = |\psi_1[x, y]|^2 = \psi_1[x, y] \cdot \psi_1^*[x, y] \]
\[ = 1 + 2 \alpha \phi_t[x, y] * s[x, y] + \sigma^2 (\phi_t[x, y] * c[x, y])^2 + \sigma^2 (\phi_t[x, y] * s[x, y])^2 \]

- “weak” phase object: neglect \( \phi^2 \) terms, \( \Rightarrow \)

<table>
<thead>
<tr>
<th>Image intensity of a weak phase object:</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ I[x, y] = 1 + 2 \alpha \phi_t[x, y] * s[x, y] ]</td>
</tr>
</tbody>
</table>

- interpretation:
  - intensity varies proportional to the projected potential \( \phi_t \)
  - the image of every object point is convoluted (“spread”) by \( s[x, y] \)

- by definition, \( s[x, y] \) is the imaginary part of the point spread function:
\[ F^{-1}[T[q]] = F^{-1}[A[q] \exp[i \chi[q]]] \]

- \( A[q] \exp[i \chi[q]] \) has the form
\[ A[q] \exp[i \chi[q]] = A[q] \cos[\chi[q]] + i A[q] \sin[\chi[q]] \]

- \( A[q] \) and \( \chi[q] \) are real-valued
⇒ the imaginary part corresponds to the right term:

\[
\text{Im}[A[q] \text{Exp}[i \chi[q]]] = A[q] \text{Sin}[\chi[q]]
\]

- linearity of the Fourier transformation ⇒ imaginary part of the Fourier transform corresponds to the Fourier transform of the imaginary part:

\[
s[x, y] = \mathcal{F}^{-1}[A[q] \text{Sin}[\chi[q]]]
\]

- contrary to the case of the general phase object, the intensity distribution in the image of the weak phase object is entirely determined by the imaginary part

\[
A[q] \text{Sin}[\chi[q]] = \text{Im}[A[q] \text{Exp}[i \chi[q]]]
\]

of the transfer function

- \(A[q] \text{Sin}[\chi[q]]\) is denoted as contrast transfer function

(⇒ do not confuse “contrast transfer function” with “transfer function”)

- applying the convolution theorem yields

\[
s[x, y] = \mathcal{F}^{-1}[A[q]] * \mathcal{F}^{-1}[\text{Sin}[\chi[q]]]
\]

⇒ for increasing aperture size \(A[q]\), \(\mathcal{F}^{-1}[A[q]]\) approaches a delta function and \(\mathcal{F}^{-1}[\text{Sin}[\chi[q]]]\) determines the point spread function:

\[
s[x, y] \approx \mathcal{F}^{-1}[\text{Sin}[\chi[q]]]
\]

⇒ intensity distribution in the image of the weak phase object:

\[
I[x, y] = 1 + 2\alpha \phi_t[x, y] * \mathcal{F}^{-1}[\text{Sin}[\chi[q]]]
\]
Discussion of the Contrast Transfer Function

- $\sin[\chi(q)] < 0$ implies “positive” contrast:
  - atom columns appear dark (in the print, not the negative!)

- $\sin[\chi(q)] > 0$ implies “negative” contrast:
  - atom columns appear bright

- $\sin[\chi(q)] = 0$ implies no transfer of the respective spatial frequency at all!

$\Rightarrow$ even under the unrealistic assumption of imaging a weak phase object, the complex properties of $\sin[\chi(q)]$ hinder a straight-forward interpretation of HRTEM images

- example: hypothetical crystal with four different sets of planes parallel to the viewing direction
  - plane spacings: $d_1 > d_2 > d_3 > d_4$
  - corresponding spatial frequencies: $\frac{1}{d_1} < \frac{1}{d_2} < \frac{1}{d_3} < \frac{1}{d_4}$
– the planes with spacing $d_1$ appear with positive contrast
– the planes with spacing $d_2$ appear with negative contrast
– the planes with spacing $d_3$ do not appear at all
– it is difficult to predict the contrast of the planes with spacing $d_4$:
  · fast oscillation of the transfer function at $(d_4)^4$
  · alignment (defocusing and direction of the primary beam) is not infinitely precise $\Rightarrow$ focal plane coordinate $q$ is not exactly known
  $\Rightarrow$ avoid these problems by introducing an objective aperture

• at which spatial frequencies does the contrast transfer function $\sin[\chi[q]]$ exhibit minima, maxima, or a value of zero?
  – at the exit surface of the specimen, the plane waves that represent the respective spatial frequency exhibit a phase shift of $-\pi/2$ versus the incident plane wave
  – in order for regions of high projected potential to appear dark in the image, the electron optics needs to introduce an additional phase shift of $-\pi/2$ (modulo $2\pi$)

$\Rightarrow$ minima in $\sin[\chi[q]]$ occur at spatial frequencies where the electron optics shifts the phase by $-\pi/2$ (modulo $2\pi$)

$\Rightarrow$ accordingly, maxima occur where the electron optics shifts the phase by $+\pi/2$ (modulo $2\pi$)

$\Rightarrow$ the contrast transfer vanishes where the phase shift amounts to an integer multiple of $\pi$ – in this case the phase difference of $-\pi/2$ with respect to the incident wave remains unaltered and no contrast results

• an ideal transfer of all spatial frequencies, in contrast, would require a transfer function of the following shape:
• actually, this ideal shape can be realized approximately over at least some extended intervals of spatial frequencies

→ passbands

• under which conditions does $\text{Sin}[\chi[q]]$ exhibit a passband?
  
  – $\chi[q]$ has a shallow minimum $\chi_0$ at some spatial frequency $q_0 > 0$
  
  – a passband for positive phase contrast occurs if
    
    • the minimum $\chi_0$ lies at $-\pi/2$ (mod 2\pi)
    • in this case the required phase shift of $-\pi/2$ extends over a certain interval of spatial frequencies

• in general, the minimum $\chi_0$ of $\chi[q]$ occurs at the spatial frequency $q_0$ at which the first derivative of the phase shift with respect to the spatial frequency vanishes

• first derivation of the phase shift with respect to the spatial frequency (see above example of the JEM 4000 EX):
\[
\frac{d}{dq} \chi[q] = 2\pi C_s \lambda^3 q^3 - 2\pi \xi \lambda q = 2\pi \lambda q \left( C_s \lambda^2 q^2 - \xi \right)
\]

- since \( \lambda > 0 \) and \( q_0 > 0 \), the condition for the spatial frequency \( q_0 \) at which \( \chi[q] \) adopts its minimum reduces to:

\[
C_s \lambda^2 q_0^2 - \xi = 0
\]

- thus

\[
q_0 = \sqrt{\frac{\xi}{C_s \lambda^2}}
\]

- therefore, a first passband occurs if

\[
\chi \left[ \sqrt{\frac{\xi}{C_s \lambda^2}} \right] = -\frac{\pi}{2}
\]

- the phase shift \( \chi_0 = \chi[q_0] \) at this spatial frequency amounts to

\[
\chi_0 = \pi C_s \lambda^3 \left( \sqrt{\frac{\xi}{C_s \lambda^2}} \right)^4 / 2 - \pi \xi \lambda \left( \sqrt{\frac{\xi}{C_s \lambda^2}} \right)^2
\]

\[
= \pi \frac{\xi^2}{2C_s \lambda} - \pi \frac{\xi^2}{C_s \lambda}
\]

\[
= -\pi \frac{\xi^2}{2C_s \lambda}
\]

- phase shift and contrast transfer function for \( \xi = \sqrt{C_s \lambda} \) (most important case, see above example of the JEM 4000 EX and section “Scherzer focus”):
• the first passband for negative phase contrast needs to fulfill the condition

\[ \chi\left[\frac{\xi}{\sqrt{C_s \lambda^2}}\right] = -\frac{3\pi}{2} \]

(this corresponds to a phase shift of \( +\pi/2 \))

• in analogy to the first passband for positive phase contrast this condition implies an underfocus of

\[ \xi = \sqrt{3} \cdot \sqrt{C_s \lambda} \]
• phase shift and contrast transfer function for $\xi = \sqrt{3} \sqrt{C_s \lambda}$:

\[ \chi[q] \]

\[ \sin[\chi[q]] \]

• in general, one obtains passbands in the contrast transfer function at defoci of the form

$$\xi = \sqrt{n} \cdot \sqrt{C_s \lambda}$$

\[ n = 1, 3, 5, \ldots \]

- $n = 1, 5, 9, \ldots$ yields passbands for positive phase contrast (atom columns dark in the image)
- $n = 3, 7, 11, \ldots$ yields passbands for negative phase contrast (atom columns bright in the image)
- even $n$ do not yield passbands at all

**Scherzer Defocus and Point Resolution**

- the defocus at which the first passband occurs in the contrast transfer function $\sin[\chi(q)]$ plays a particularly important role

- at this defocus the contrast transfer function most closely resembles the contrast transfer function of an ideal microscope

- with increasing underfocus, the first passband occurs when the phase shift $\chi(q)$ amounts to about $-\pi/2$ for an extended interval of spatial frequencies

- in particular, a passband occurs if the minimum $\chi_0$ of $\chi(q)$ adopts the value of $\pi/2$ (see above)

- the corresponding defocus is known as

<table>
<thead>
<tr>
<th>Scherzer (de)focus:</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\zeta_S = \sqrt{C_s \lambda}$</td>
</tr>
</tbody>
</table>

- the larger the width of the passband, the larger the interval of spatial frequencies that the microscope can transfer in an “intuitively interpretable” way

$\rightarrow$ which defocus $\zeta_0$ yields a passband of maximum width?
Ansatz for the allowed interval of phase shifts

\[-\frac{2}{3}\pi \leq \chi[q] \leq -\frac{1}{3}\pi\]

• this yields a minimum phase shift of

\[\chi_0 = -\frac{2}{3}\pi\]

• insert in the expression for the phase shift \(\chi_0 \equiv \chi[q_0]\) (see above):

\[-\pi \frac{\xi_0^2}{2C_s\lambda} = -\frac{2}{3}\pi\]

• solving for the defocus yields:

\[\xi_0 = \sqrt[3]{\frac{4}{3}C_s\lambda}\]

• result: a first passband of maximum width results from a defocus \(\sqrt{4/3} \approx 1.2\) larger than Scherzer defocus

• since this defocus optimizes the resolution (when imaging a weak-phase object), it is designated as

Optimum defocus:

\[\xi_0 \approx 1.2\sqrt{C_s\lambda} \equiv 1.2\text{ Scherzer}\]

• comparison between Scherzer defocus and optimum defocus:
• up to the first zero of the contrast transfer function the microscope transfers all spatial frequencies with similar phase shifts and contrast

• spatial frequencies within the first passband, therefore, are “directly” resolved in the image (intuitively interpretable)

• the maximum width of the first passband determines the point resolution – a broader passband implies intuitively interpretable transfer of higher spatial frequencies and thus smaller spacings
• the **point resolution** \( d_P \) of the microscope corresponds to the inverse of the spatial frequency at which the first zero occurs in the contrast transfer function

\[
\chi[q] = 0; \quad q > 0, \quad \zeta = \frac{4}{3} C_s \lambda
\]

• the corresponding zero of \( q_P \) of the contrast transfer function is obtained from

\[
\begin{align*}
\pi C_s \lambda^3 q_P^4 / 2 - \pi \left( \frac{4}{3} C_s \lambda \right) \lambda q_P^2 &= 0 \\
\Rightarrow \sqrt{C_s} \lambda^{3/2} q_P^2 &= 2 \frac{4}{3} \\
\Rightarrow q_P &= \frac{2}{3^{1/4}} C_s^{-1/4} \lambda^{-3/4} \approx 1.52 C_s^{-1/4} \lambda^{-3/4}
\end{align*}
\]

• the inverse of \( q_P \) constitutes the

\[
\text{Point resolution:}
\]

\[
d_P = 0.66 C_s^{1/4} \lambda^{3/4}
\]

\[
\Rightarrow \text{the point resolution improves more rapidly with decreasing the wavelength (increasing the accelerating voltage) than with decreasing the spherical aberration}
\]

• typical order of magnitude:
  - JEM 4000 EX (JEOL)
  - \( C_s = 1.0 \text{ mm}, \lambda = 1.644 \text{ pm} \Rightarrow d_P = 0.170 \text{ nm (experimental: 0.175 nm)} \)