End to End: Part 1

Earl Wong

Subjects

- **Optics**
- Sensor
- ISP
- GPU
- NPU

Optics - Image Formation

- We begin with image formation.
- We explain the basics, using a pinhole camera.
- A pinhole camera uses the following constructs to create an image:

1) ray tracing

2) a small “pinhole” aperture

3) an underlying scene

Optics - Image Formation

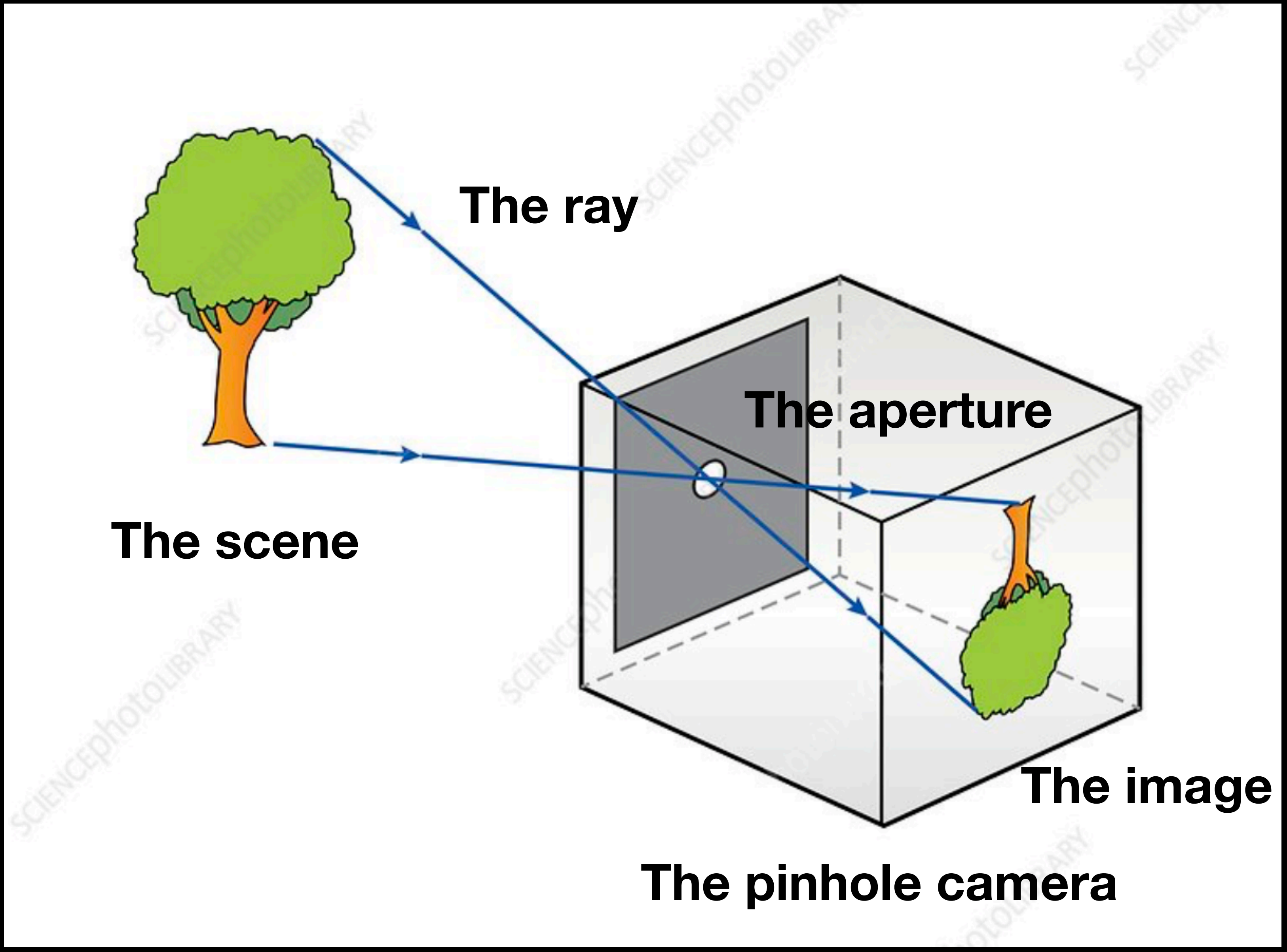
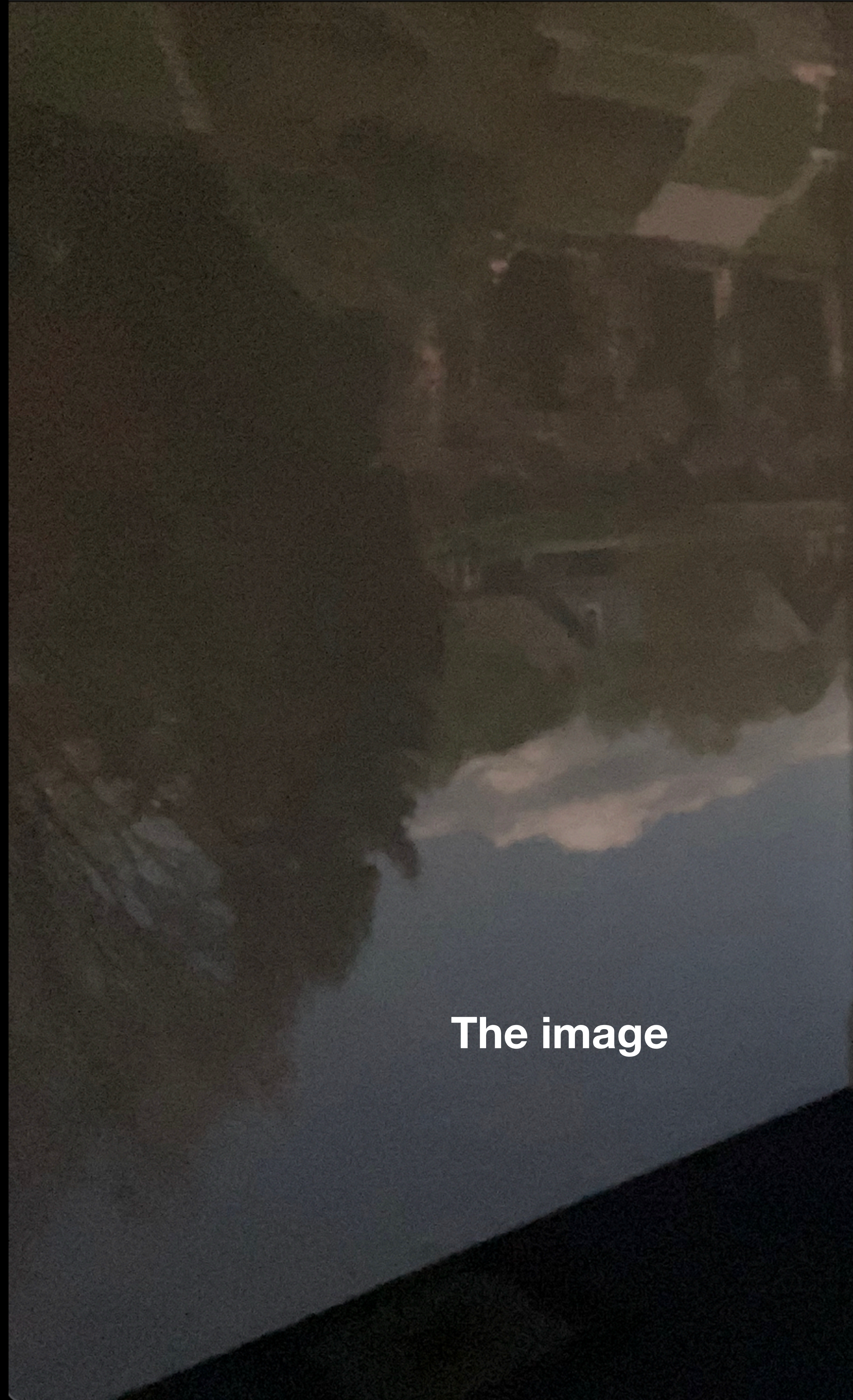


Image courtesy of Science Photo Library

Optics - Image Formation

- Using ray tracing, we can map illumination sources from the scene into the image plane.
- The next two slides show a pinhole camera in operation, taken by the author at the Eastman Kodak Museum, in Rochester, New York.

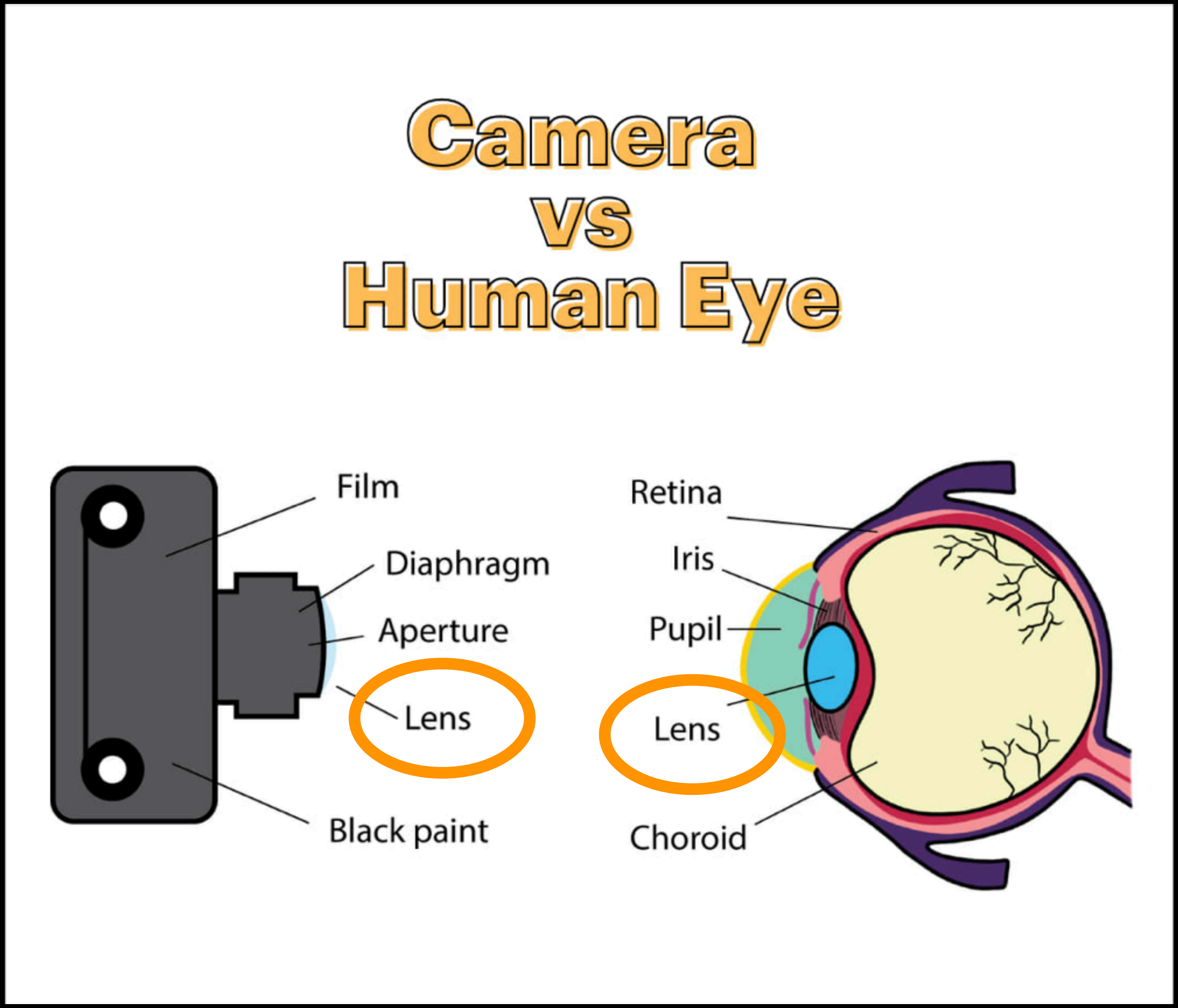

The aperture


The image

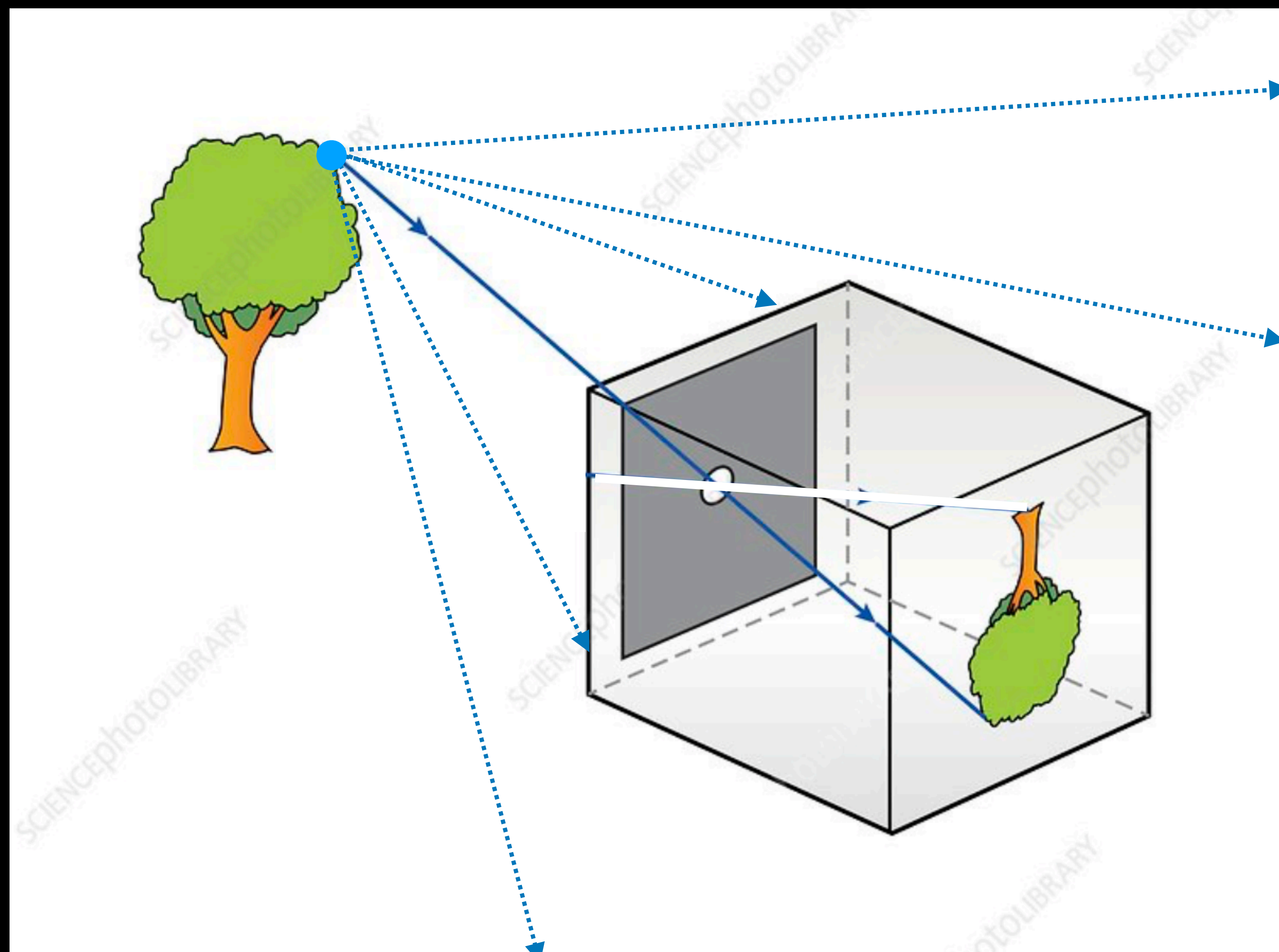
Optics - Image Formation

- The human visual system is an advanced pinhole camera - a camera with additional “features”.
- The human visual system has a lens for gathering and focusing light.
- The human visual system has a retina (containing photoreceptors) for capturing the image formed by the light.
- In this presentation, we will focus on the lens, and show how lens design is closely tied to the image sensor.
- In the next presentation, we will talk about image sensors in more detail.

Optics - Image Formation



Optics - Image Formation



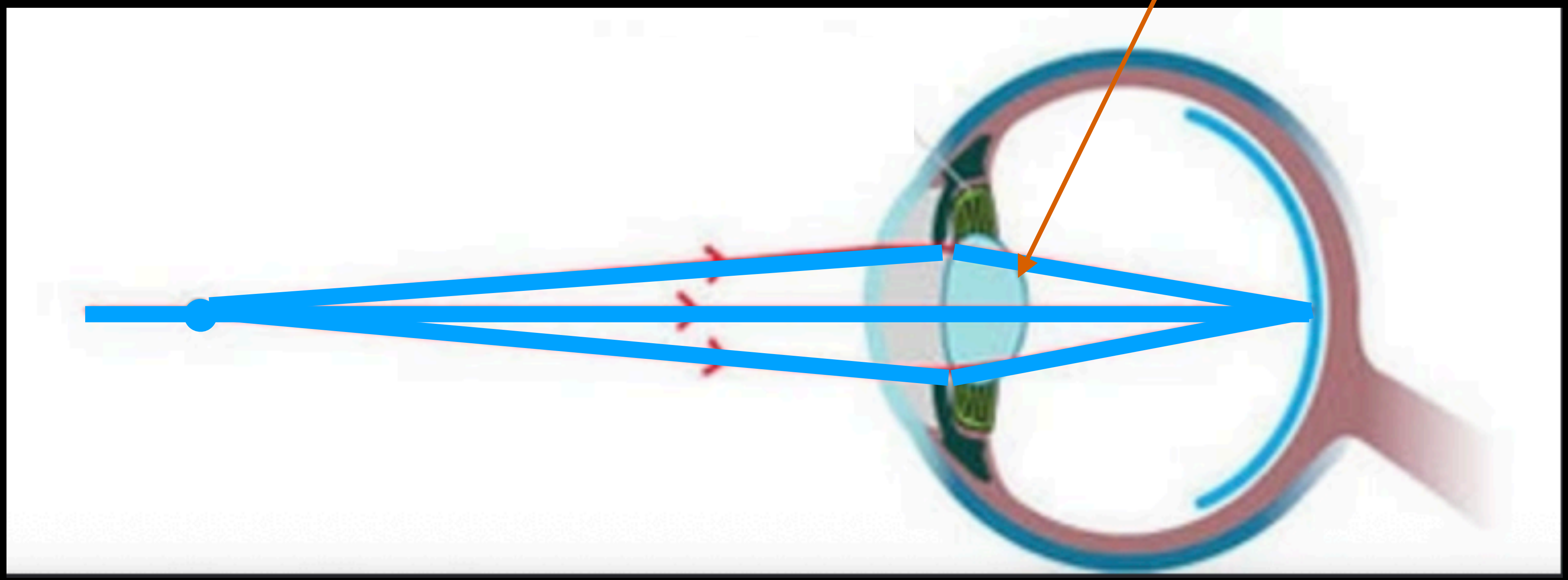
By construction, every point in the scene emits numerous rays.
However, a pinhole camera only allows a small number of these rays to be imaged.

Once a lens is substituted for the pinhole aperture, 1) additional rays can be collected and 2) focused.

Optics - Image Formation

- To increase the light gathering capabilities, a lens is introduced.
- The lens “collects” rays emitted from the scene location / point source, redirecting them to the image plane.
- In addition, the lens also focuses the rays onto the image plane.
- More will be said on the latter, during the discussion of depth of field (DoF).

Optics - Image Formation



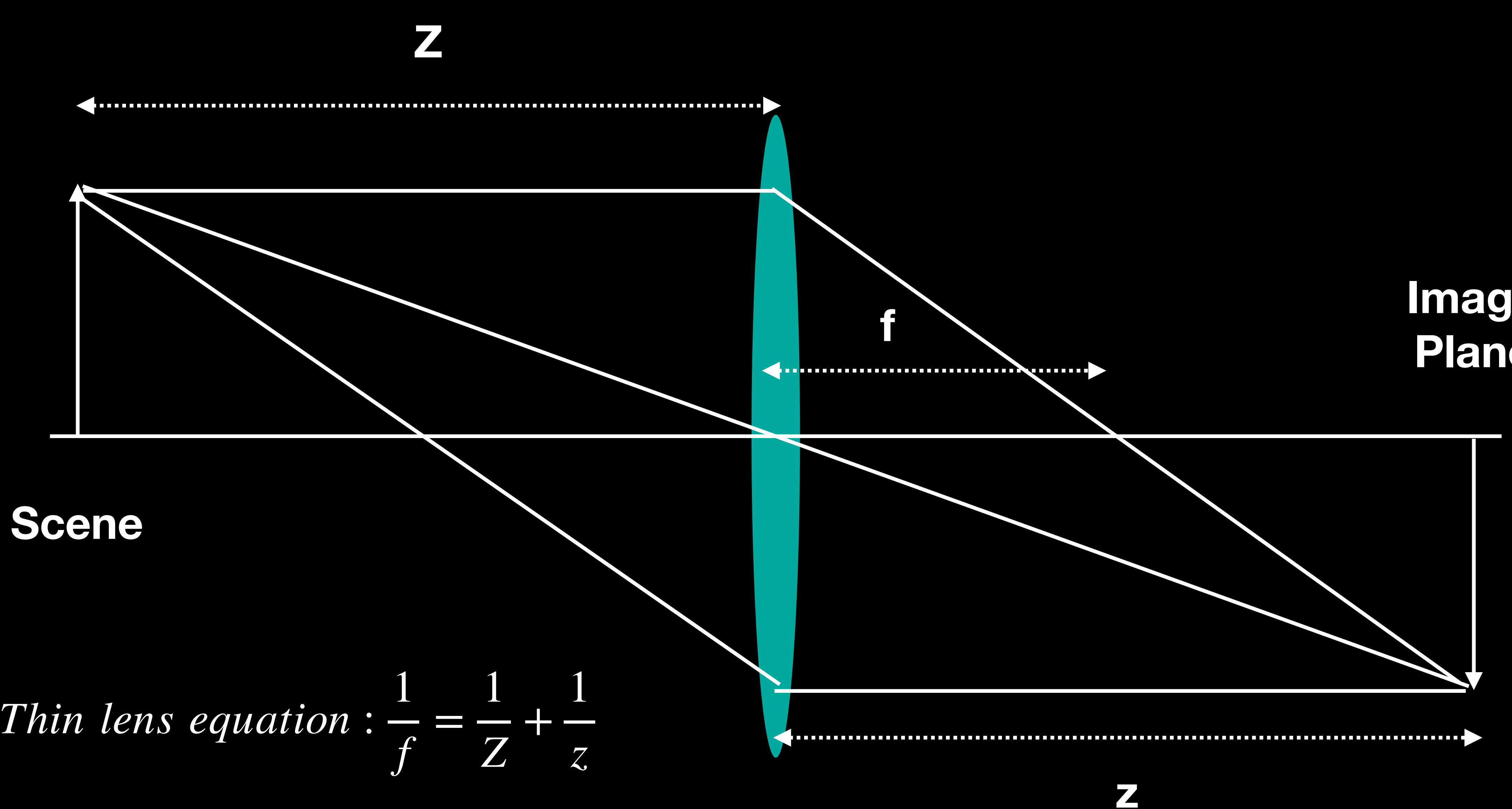
The human visual system collects multiple rays emitted by the source, focusing the rays onto the retina.

Optics - Image Formation

- The simplest lens model is the thin lens model.
- The thin lens model is governed by the following

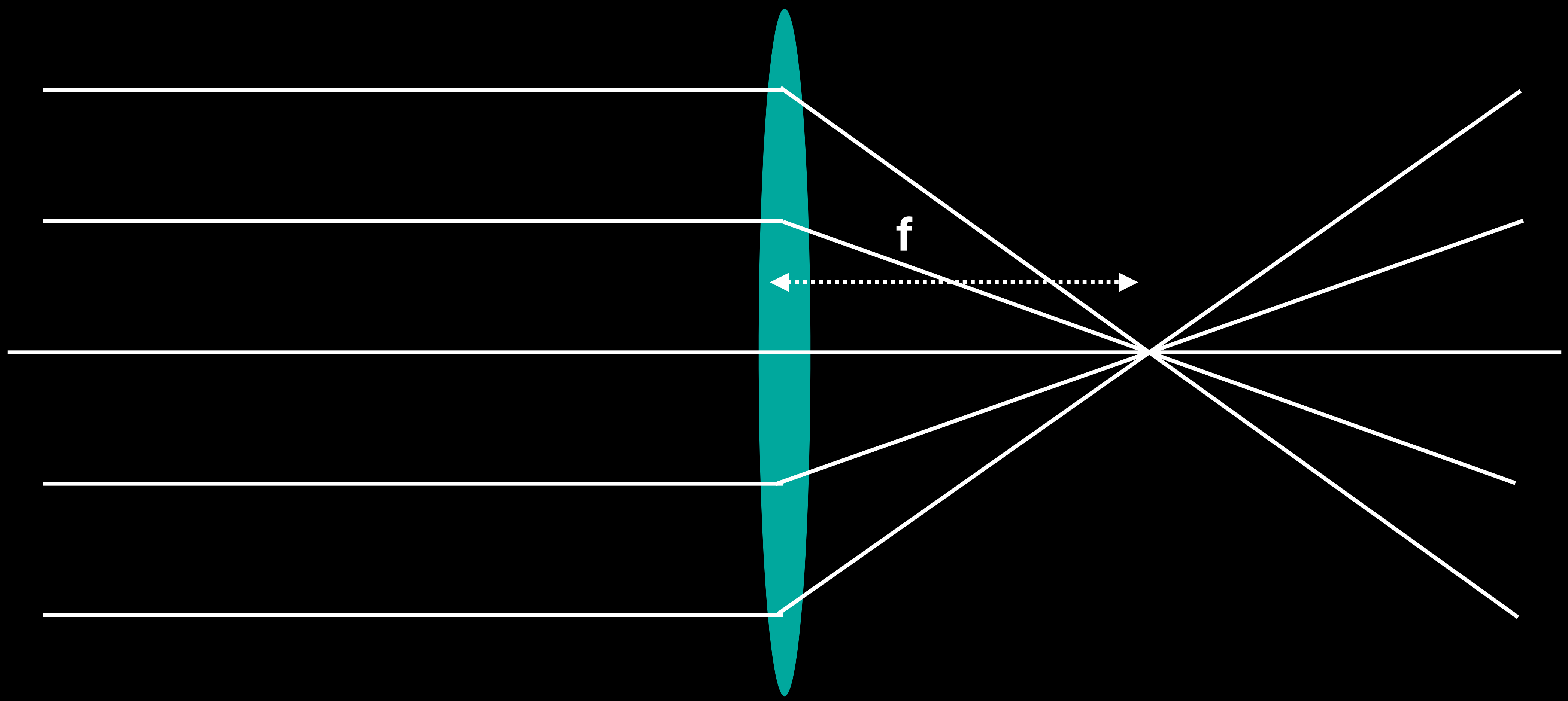
  equation: $\frac{1}{Z} + \frac{1}{z} = \frac{1}{f}$
- The thin lens equation says the following: For an object at a distance Z from the lens, an image will form at a distance z from the lens, for a lens with focal length f.
- The thin lens equation can be derived using ray tracing and triangle equalities.
- The lens focal length f is defined to be the location on the optical axis where parallel rays from the scene (=very distant object point source) and perpendicular to the lens, are focused.

Optics - Image Formation



$$Thin\ lens\ equation: \frac{1}{f} = \frac{1}{Z} + \frac{1}{z}$$

Optics - Image Formation



focal length f of a lens

Optics - Image Formation

- But what happens, if various objects in the scene are at different distances, like in real life?
- Objects at one distance will be in focus, while objects at other distances may or may not be.
- To determine what is and is not in focus, the depth of field (DoF) needs to be determined.
- The DoF determines (for a given lens focus position), the “range” of distances in the scene that will be imaged “in focus” on the image plane.

Optics - Image Formation

The DoF is a function of four different parameters.

- 1) Lens focus position.
- 2) Lens focal length.
- 3) Lens aperture size.
- 4) Sensor pixel size.

Optics - Image Formation

Lens focus position

- i.e. The DoF is larger for distant lens focus positions and smaller for closer lens focus positions.
- Hence, it is easier to focus your camera for landscape photographs versus macro photographs, since the DoF is significantly large in landscape use case.

Optics - Image Formation

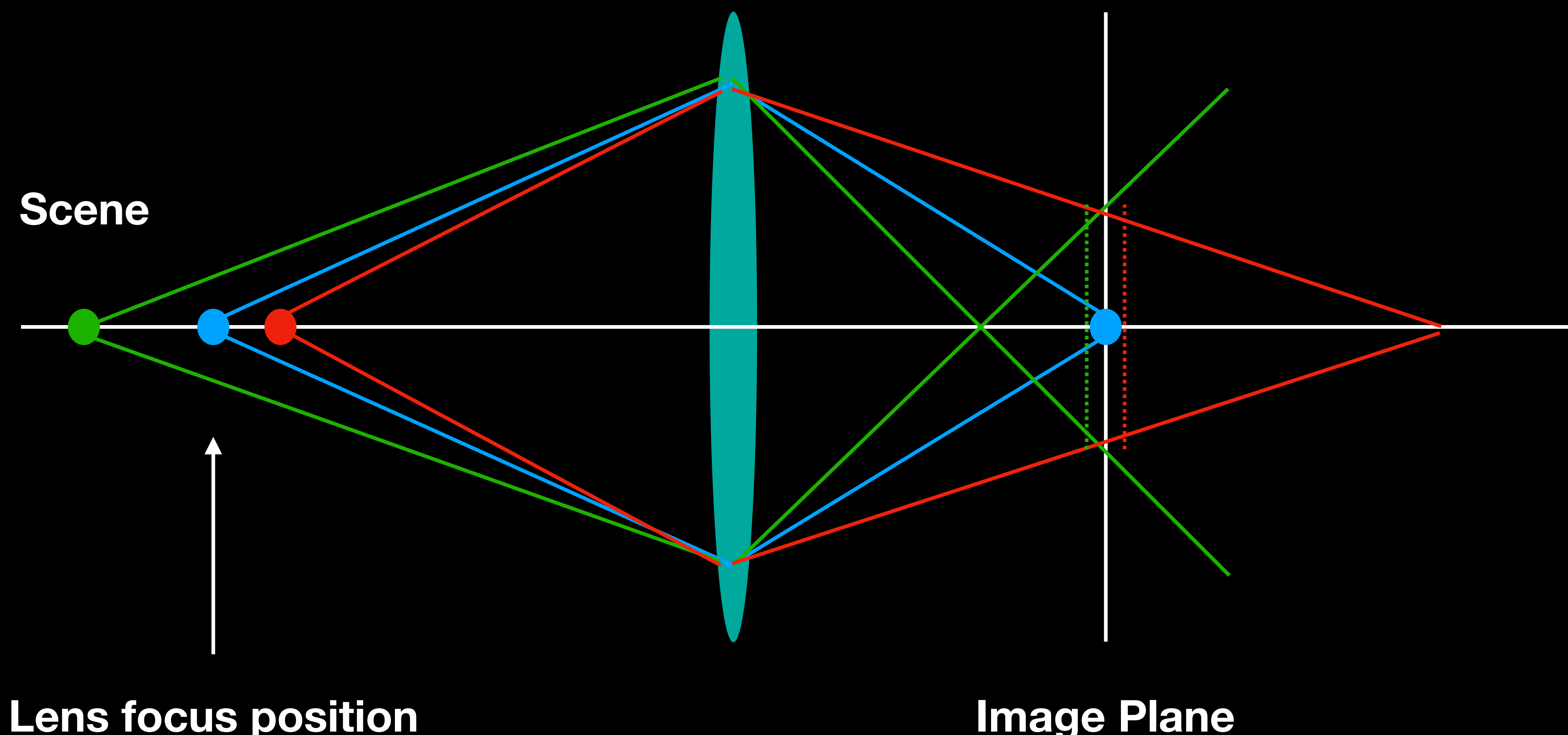
Lens focal length f

- The DoF is smaller for lenses with large focal lengths and larger for lenses with small focal lengths.
- Hence, tele photo lenses have smaller DoF’s than wide angle lens.
- This is why auto focus is a more challenging problem for tele photo lenses.

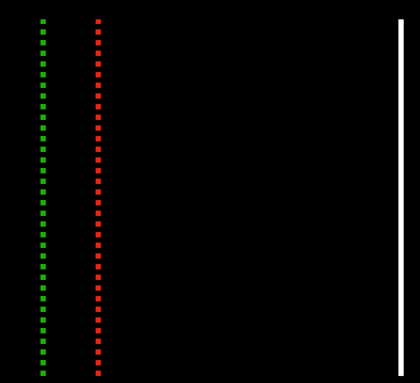
Optics - Image Formation

- In addition, the DoF is also a function of the sensor pixel size and the lens aperture size.
- We will now illustrate these two cases, using ray diagrams.

Optics - Image Formation

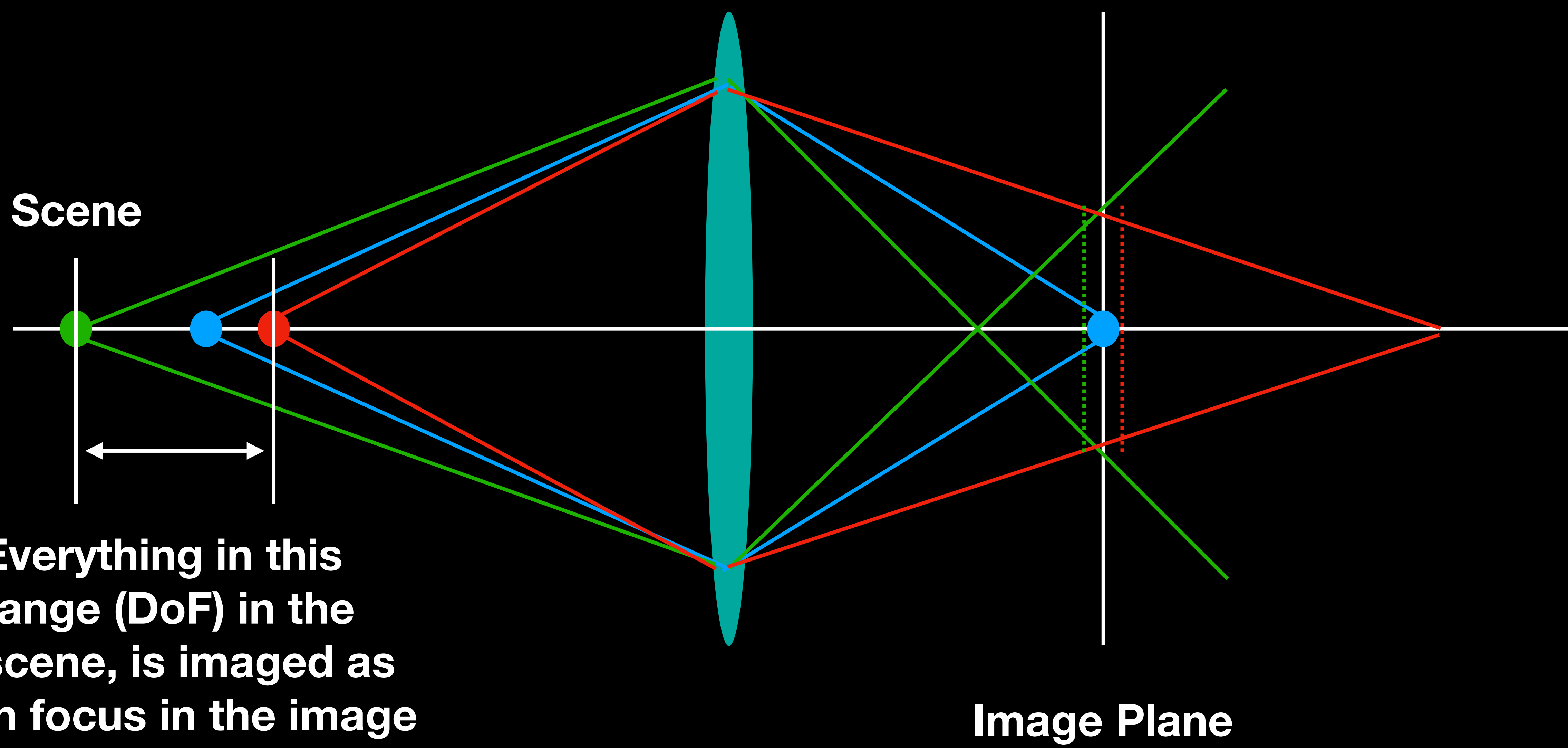


Optics - Image Formation



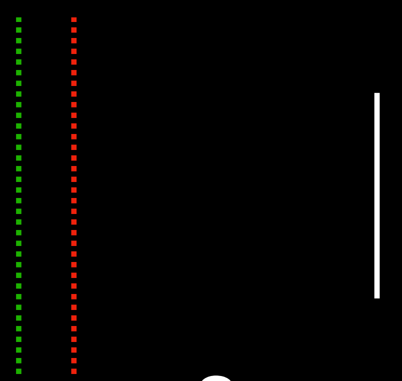
Suppose the image sensor pixel was square and had the following height

Optics - Image Formation



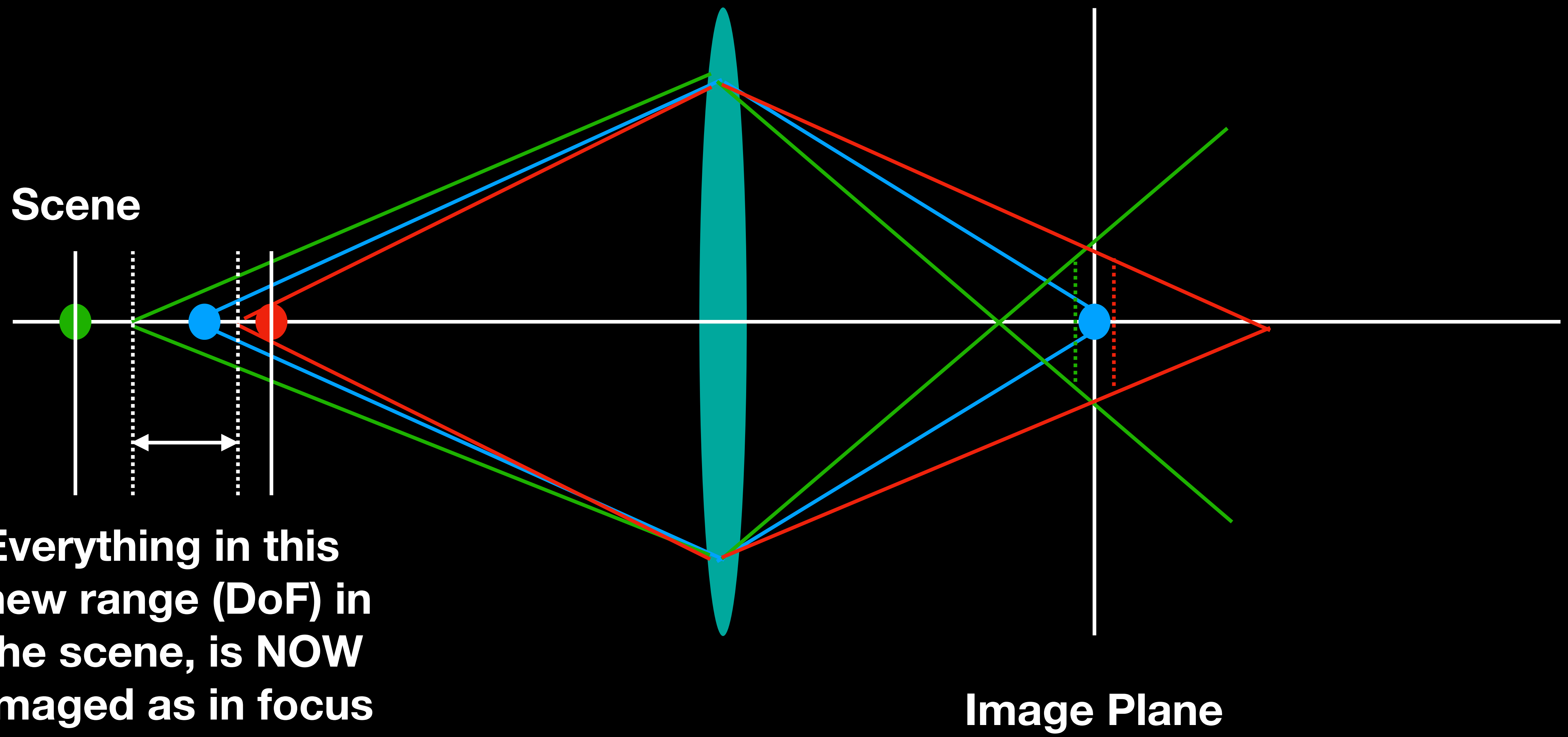
Everything in this range (DoF) in the scene, is imaged as in focus in the image plane

Optics - Image Formation



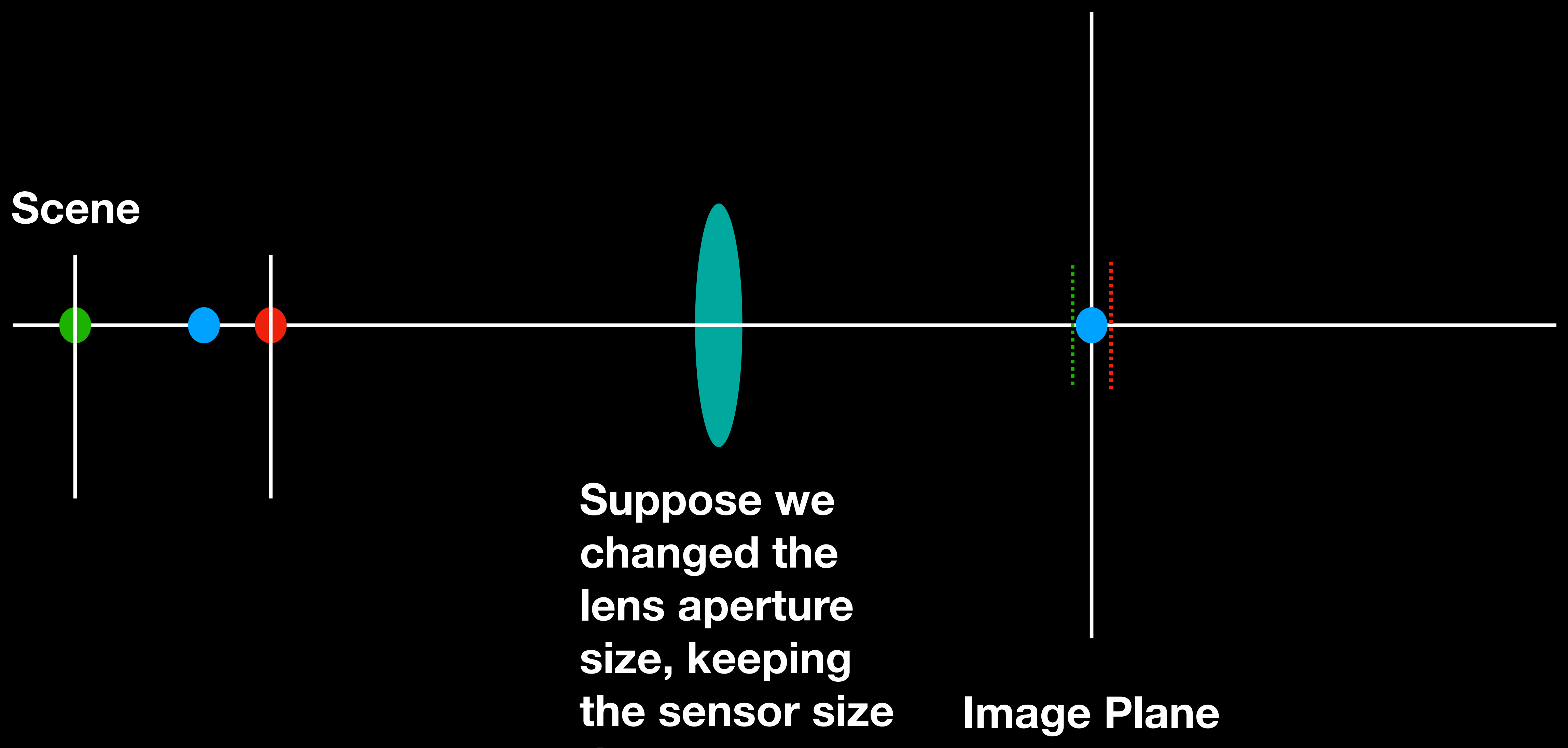
Suppose the image sensor pixel was square and NOW had the following reduced height

Optics - Image Formation



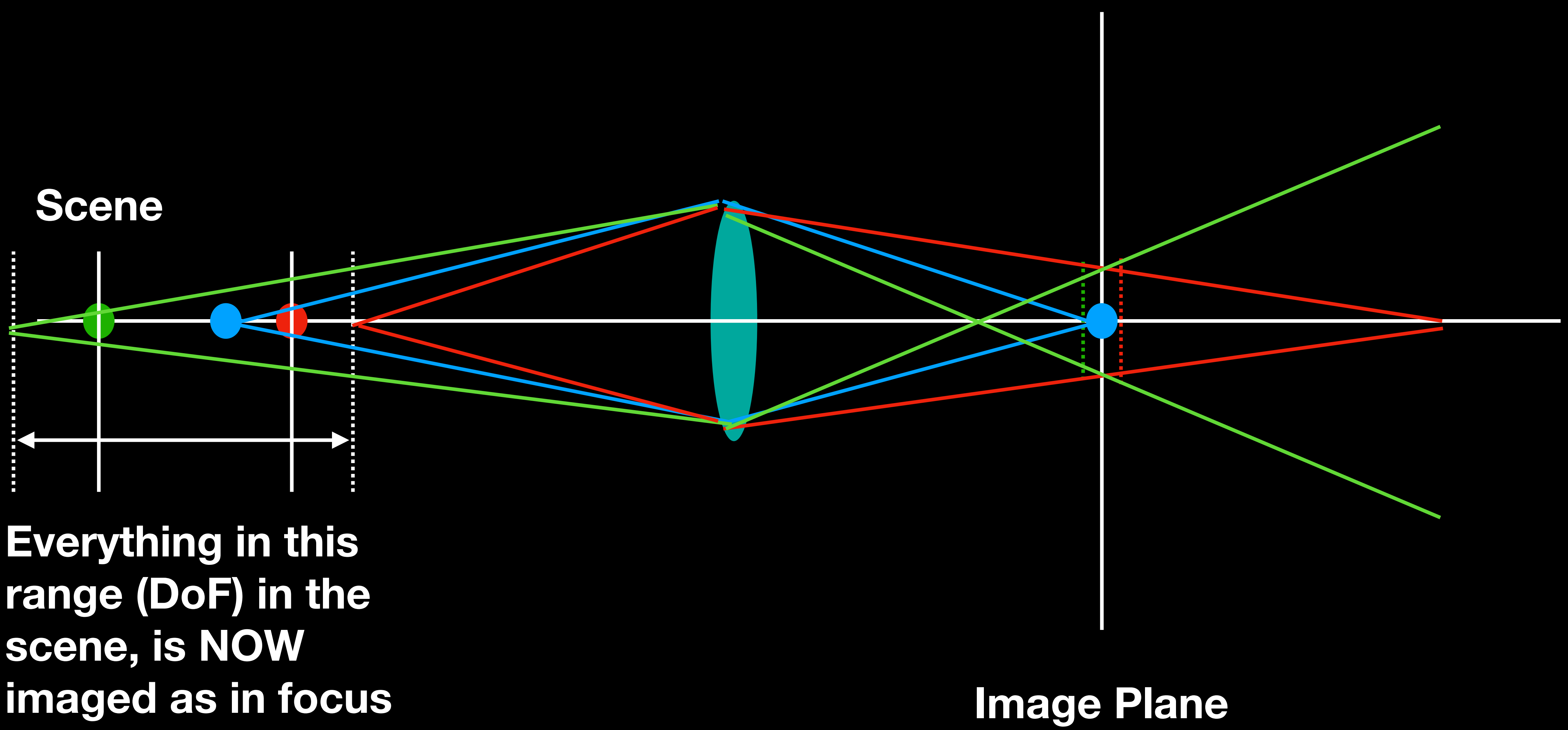
Everything in this new range (DoF) in the scene, is NOW imaged as in focus in the image plane

Optics - Image Formation



Suppose we changed the lens aperture size, keeping the sensor size the same

Optics - Image Formation



Everything in this range (DoF) in the scene, is NOW imaged as in focus in the image plane

Optics - Image Formation

- By construction, there will always be a distance in the scene plane (=focus position) that will be in focus in the image plane.
- The size of the region around the focus position that will be in focus increases as the aperture decreases and / or the pixel size increases.
- The size of the region around the focus position that will be in focus decreases as the aperture increases and / or the pixel size decreases.
- Different images result, as the DoF changes.

Optics - Image Formation

f/3.2, dof = 0.024m　　f/5.6, dof = 0.043m　　f/11, dof = 0.086m　　f/22, dof = 0.172m

Focal length 105mm, Focus Distance 1.5m

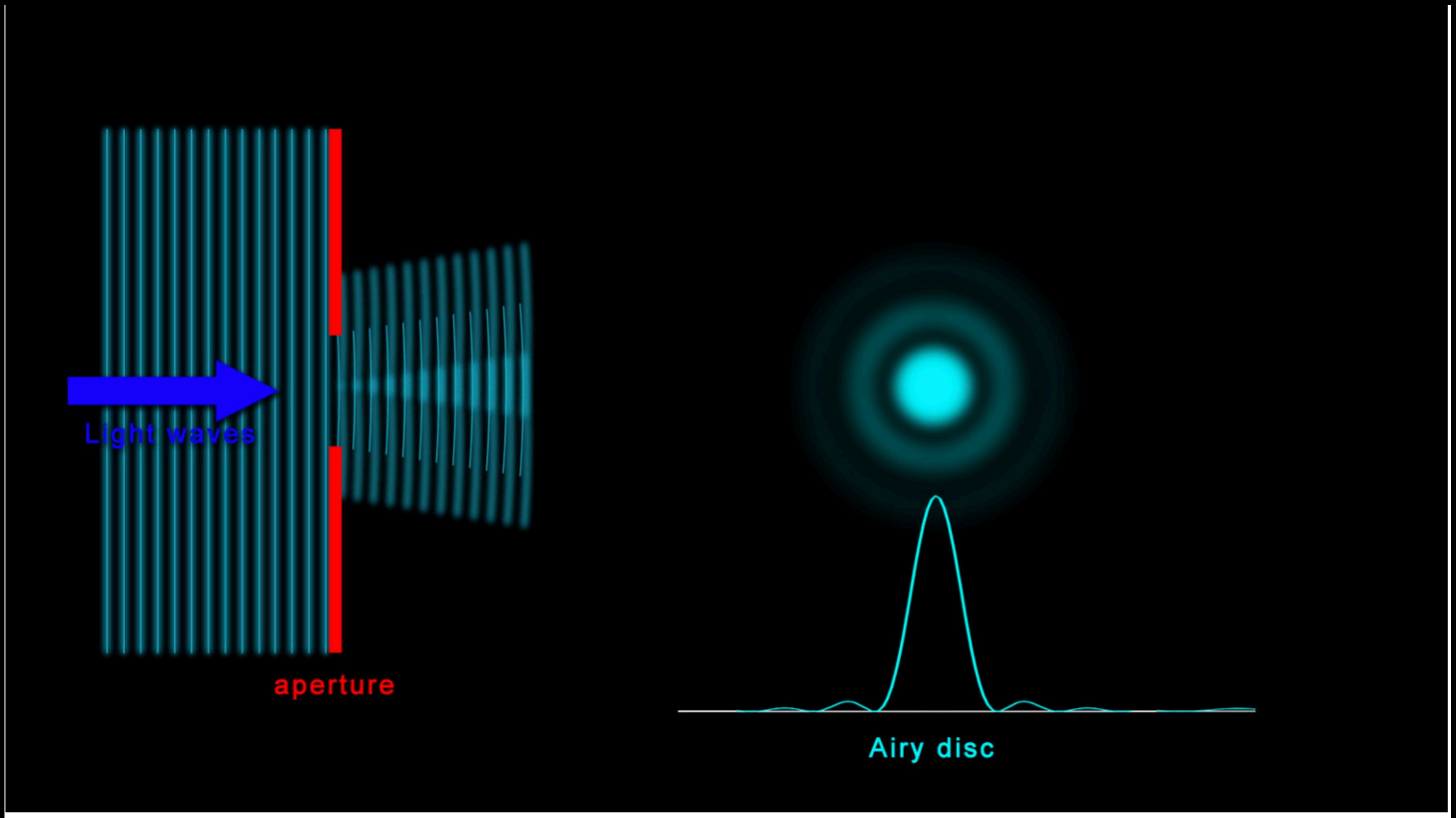
Image courtesy of Photography Life

Here, the lens aperture is changed.
The left most image results from a small / shallow DoF.
The right most image results from a large / deep DoF.

Optics - Design

- Because every lens has a finite aperture (diameter), diffraction is an unavoidable physical property.
- From the wave theory of physics, diffraction produces (downstream) constructive and destructive interference.
- For circular apertures, this interference is captured (by an imaging sensor) as an Airy disk.
- The diameter of the Airy disk is given by: $d_{Airy} = 2.44\lambda N$, where N is the f-number of the lens.
- $N = \frac{f}{D}$, f *is the focal length of the lens and* D *is the diameter*

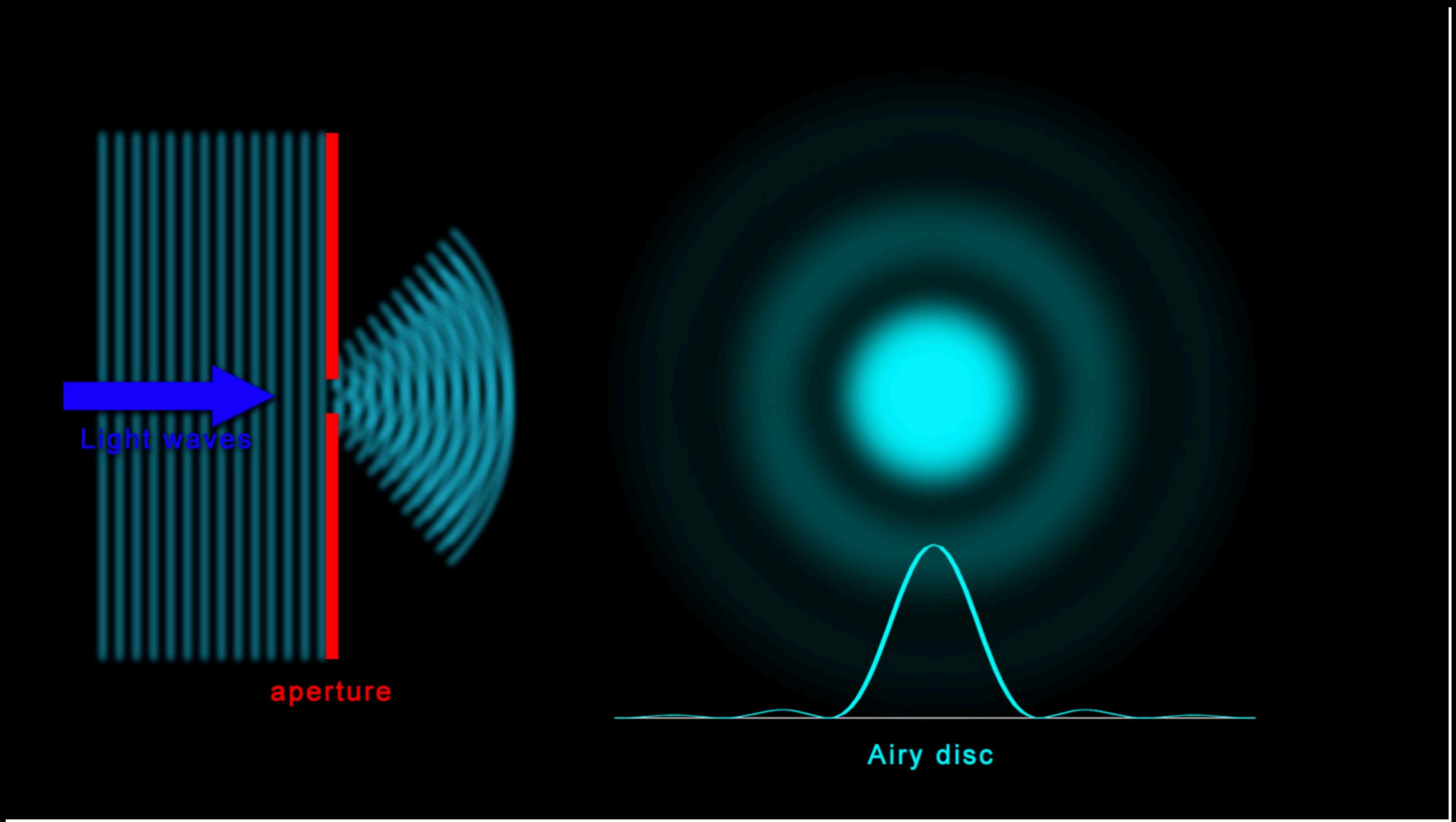
Optics -Design



A large lens opening will make the Airy disk small.

Image courtesy of FStoppers

Optics - Design



If the lens opening becomes smaller, the Airy disk will grow larger.

Image courtesy of FStoppers

Optics - Design

- The lens aperture is created by introducing a mechanical iris immediately in front of the lens.

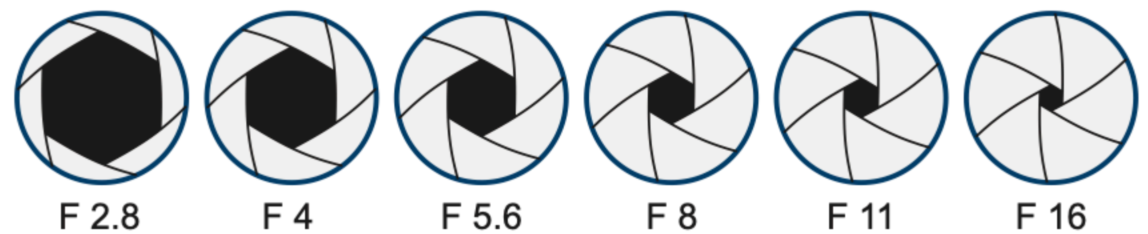


Fig. 15: Change in collected light with increasing f-number.

Image courtesy of Basler Optics

- For the human eye, the pupil functions as a mechanical iris.

Optics - Design

- To minimize diffraction (and subsequently, the diameter of the airy disk) a very large aperture is needed.

Technically:

- A Bessel function (shaped) output is produced, when light is diffracted through a circular aperture.

- The square of the Bessel function (=airy disk) is then captured by the image sensor (=point spread function (PSF))

- When the fourier transform of the PSF (=airy disk) is computed, the optical transfer function (OTF) is obtained.

- When we the magnitude of the (OTF) is computed, the modulation transfer function (MTF) is obtained.

The MTF describes how contrast decreases as a function of spatial frequency (line pairs /mm).

The perfect MTF curve would be a straight line with contrast of 1 (or modulation = 100 percent) for the full range of spatial frequencies under evaluation.

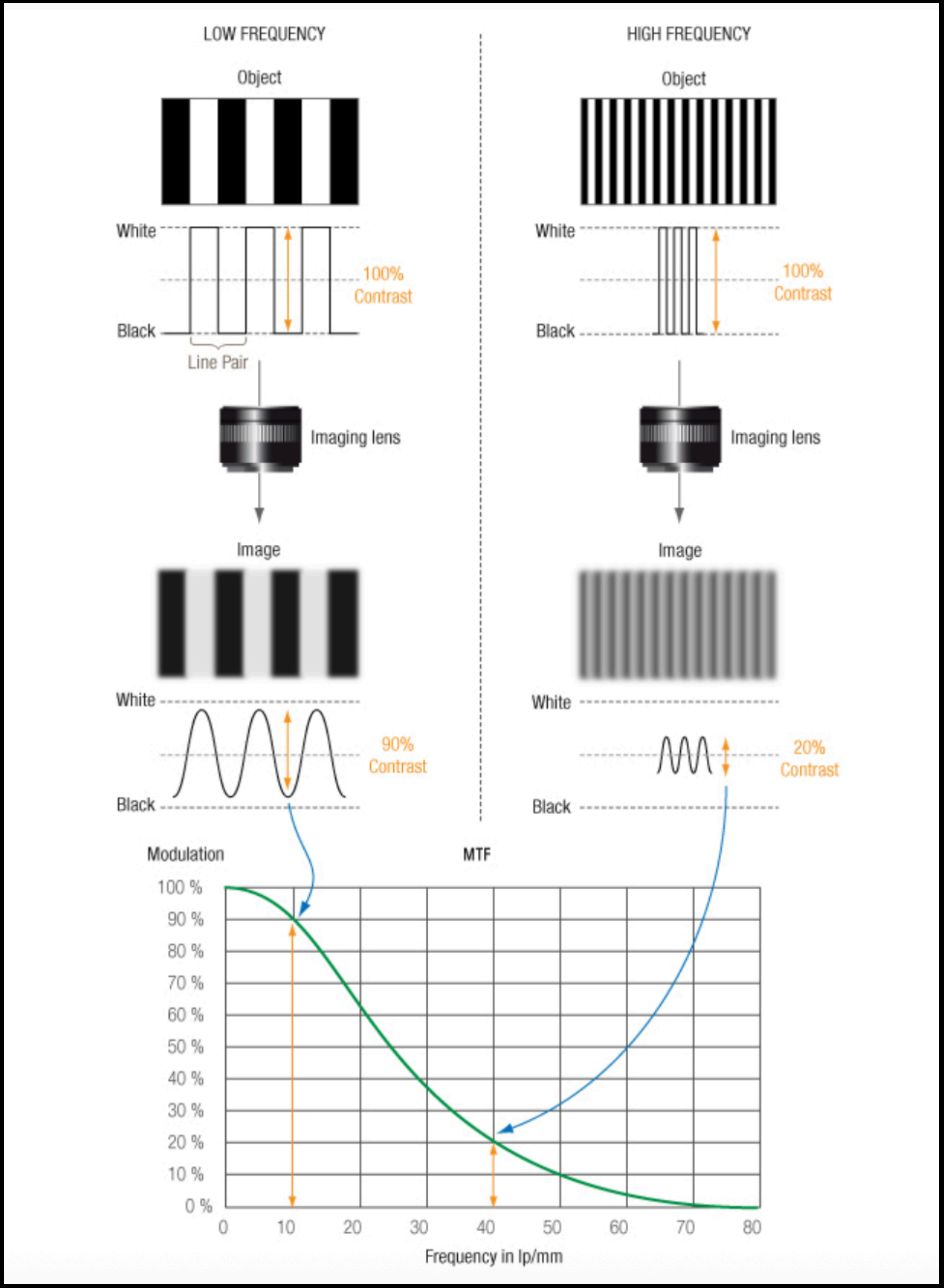


Image courtesy of DXOMark

Optics - Design

- In practice, the MTF of a lens is determined by diffraction and lens aberrations.
- i.e. $MTF_{lens=diffraction+aberration}$
- Diffraction sets the upper bound of the MTF, and is dictated by physics.
- Lens aberrations subsequently degrade this “upper bound” MTF.
- Manufacturing tolerance and focus inaccuracies then affect the measured repeatability of the MTF curve.

Optics - Design

- The strehl ratio is defined as: $S = \dfrac{Peak\ intensity\ of\ real\ PSF}{Peak\ intensity\ of\ diffraction\ limited\ PSF}$
- If the strehl ratio is approximately > 0.9, no significant MTF degradation (relative to the diffraction upper bound) is visible.
- If the strehl ratio is between 0.8 and 0.9, the MTF degradation due to lens aberrations starts to become noticeable.
- If the strehl ratio is approximately < 0.8, the MTF degradation due to lens aberrations becomes increasingly significant - with approximately equal contribution between diffraction and lens aberration at S ~ 0.5.
- Definitions: S > 0.8 = diffraction limited; S << 0.8 aberration limited

Optics - Design

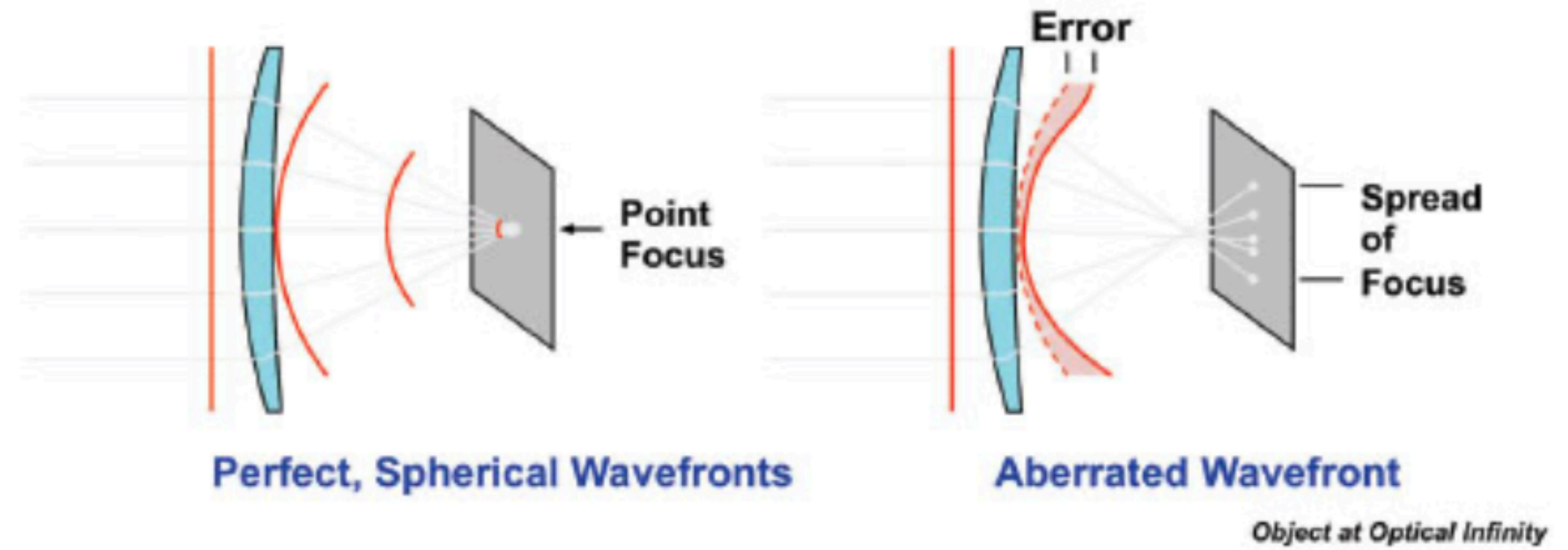
There are many different types of lens aberrations:

- Spherical aberration
- Coma
- Astigmatism
- Field curvature
- Chromatic aberration

Optics - Design

- Lens aberrations result, when all of the rays from a point source in the scene cannot arrive at a single point in the image plane in the same phase.
- i.e. The light spreads on the image plane due to imperfections in the lens.
- i.e. There is wavefront error / imperfect wavefront shaping.
- End result: The resulting PSF gets broader, smeared and / or asymmetric.
- The goal of the lens designer is to minimize lens aberrations by introducing aspherical lens elements, multiple lens elements, lens coatings, etc.

Optics - Design



In a perfect optical system, spherical waves of light converge to a single point focus at the desired location (e.g., the retina of the eye). If these wavefronts are too steep, too flat or distorted from their ideal shape, they produce wavefront errors that cause the focus of light rays to spread out or smear.

Wavefront Error

Optics - Design

- Certain lens aberrations can be significantly reduced by stopping down the lens aperture.
- These include spherical, coma and field curvature.
- The first two are reduced, by preventing rays from entering from the periphery of the lens.
- Field curvature is reduced because a stopped down lens creates a larger DoF.
- However, the more a lens is stopped down, the greater MTF degradation due to diffraction.

Optics - Design

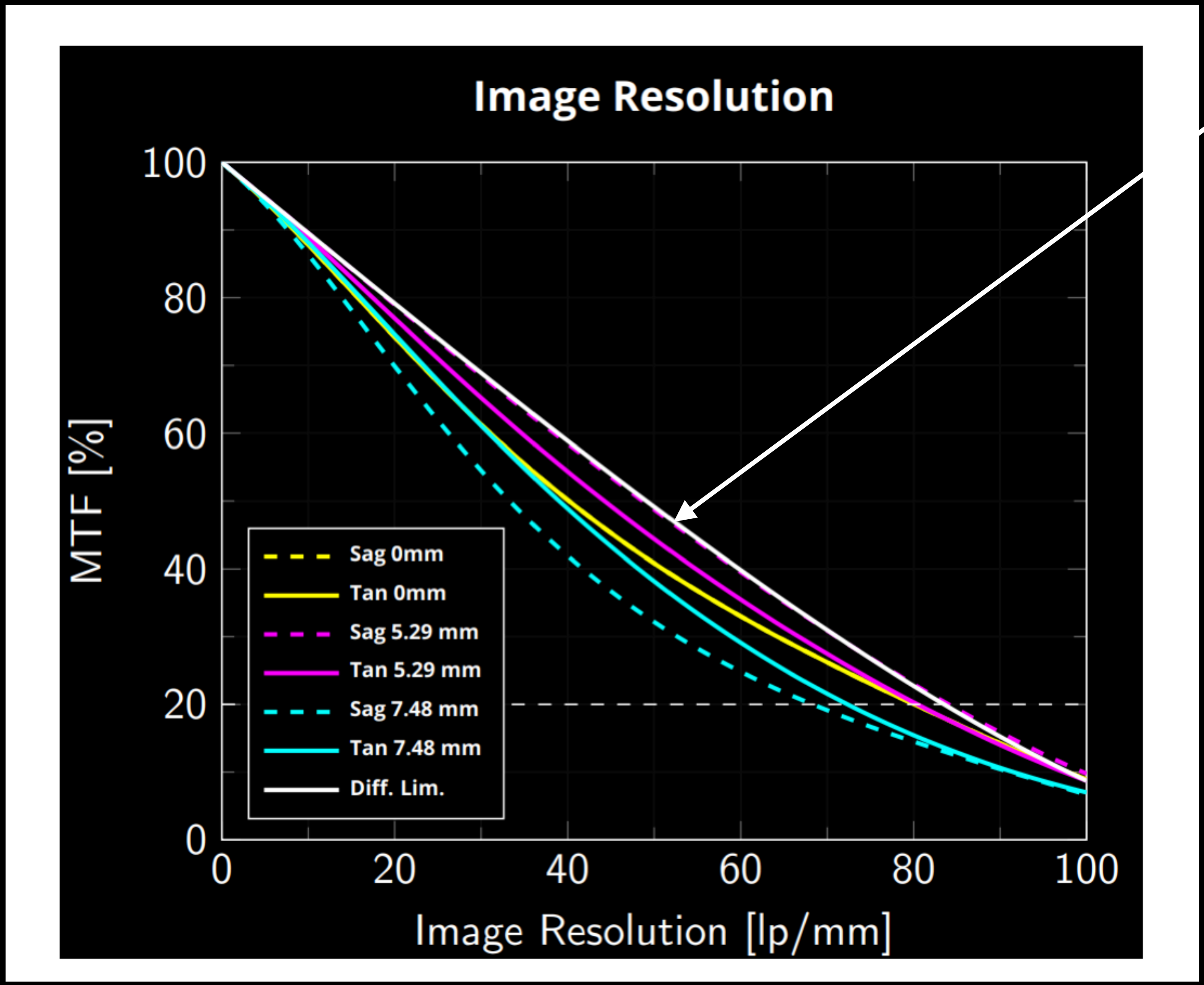
General truth

- Cheaper lenses are usually aberration limited
- Expensive lenses are usually diffraction limited

This makes logical sense.

Cost naturally increases as more “effort” is made to eliminate lens aberrations.

Optics - Design



The upper black curve sets the upper bound performance of the lens, based on diffraction.

When a lens is stopped down, this curve will rotate clockwise.

The curves below the black curve are the cumulative MTF results, when varying amounts of lens aberrations are introduced / not mitigated.

Sag and Tan are the two directions used to measure this degradation.
Sag can be viewed as line pairs corresponding to the spokes emerging from a bicycle wheel.
Tan are the the line pairs perpendicular to Sag.

Optics - Design

- As previously stated, the MTF measures contrast versus resolution.
- The lens dictates MTF performance though diffraction and lens aberrations.
- Both have associated PSF’s that can smear nearby pixel information.
- At the same time, the sensor also plays a significant role.
- For a given sensor size, a sensor with a denser number of pixels will result in a higher resolution image.

Optics - Design

- i.e. The pixel dimension / size also has an associated MTF.

- Consider a square pixel.

- The square pixel integrates light over a given area.

- Mathematically, the pixel performs a convolution operation with the scene.

- $g(x, y) = I(x, y) * h_{pix}(x, y)$ where $h_{pix}(x, y) = \frac{1}{a * a} rect(\frac{x}{a}) rect(\frac{y}{a})$

Optics - Design

- We compute the Fourier transform of $h_{pix}(x, y) = \frac{1}{a * a} rect(\frac{x}{a}) rect(\frac{y}{a})$ obtaining the OTF: $H_{pix}(f_x, f_y) = sinc(af_x)sinc(af_y)$

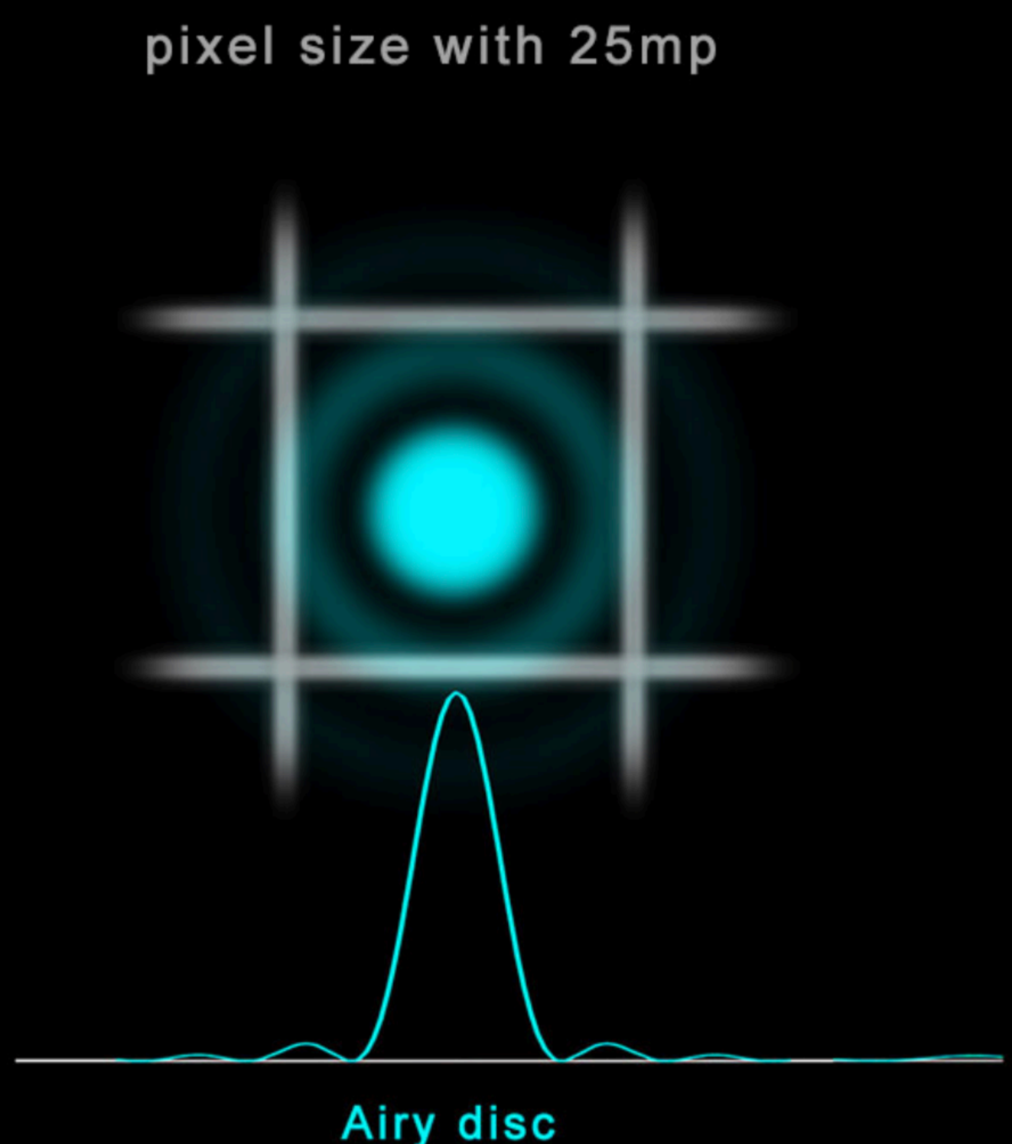
- The MTF is given by: $|H_{pix}(f_x, f_y)| = |sinc(af_x)sinc(af_y)|$

- The system MTF is now given by:

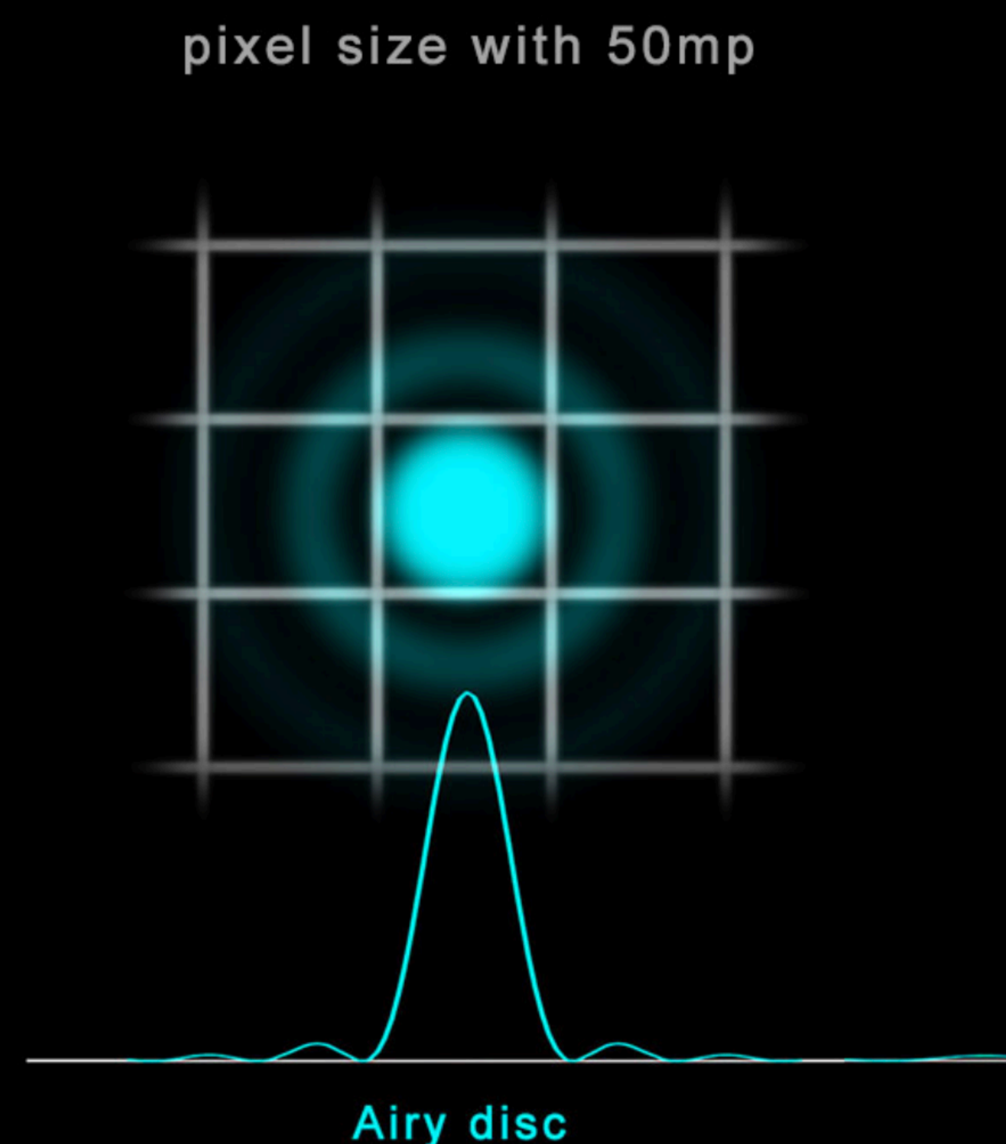
$$MTF_{system} = MTF_{lens=diffraction+aberration} MTF_{pixel}$$

- The picture on the next slide, illustrate the physical interaction between the two quantities in the spatial domain, assuming the lens is diffraction limited.

Optics - Design



Resolving power limited by pixel size.



Resolving power limited by aperture size (diffraction).

Image courtesy of FStoppers

Optics - Design

- In terms of the system MTF, we observe that we have a lens MTF curve and a pixel MTF curve.

- Both curves decrease, as the spatial frequency increases.

- To obtain the best system MTF curve (=the product of the two curves), we would need to:

1) Use the largest aperture possible (to minimize the airy disk diameter / minimize the PSF due to diffraction)

2) Use all of the lens aberration reduction techniques at our disposal

3) Use the smallest pixel dimension possible (to maximize the pixel MTF)

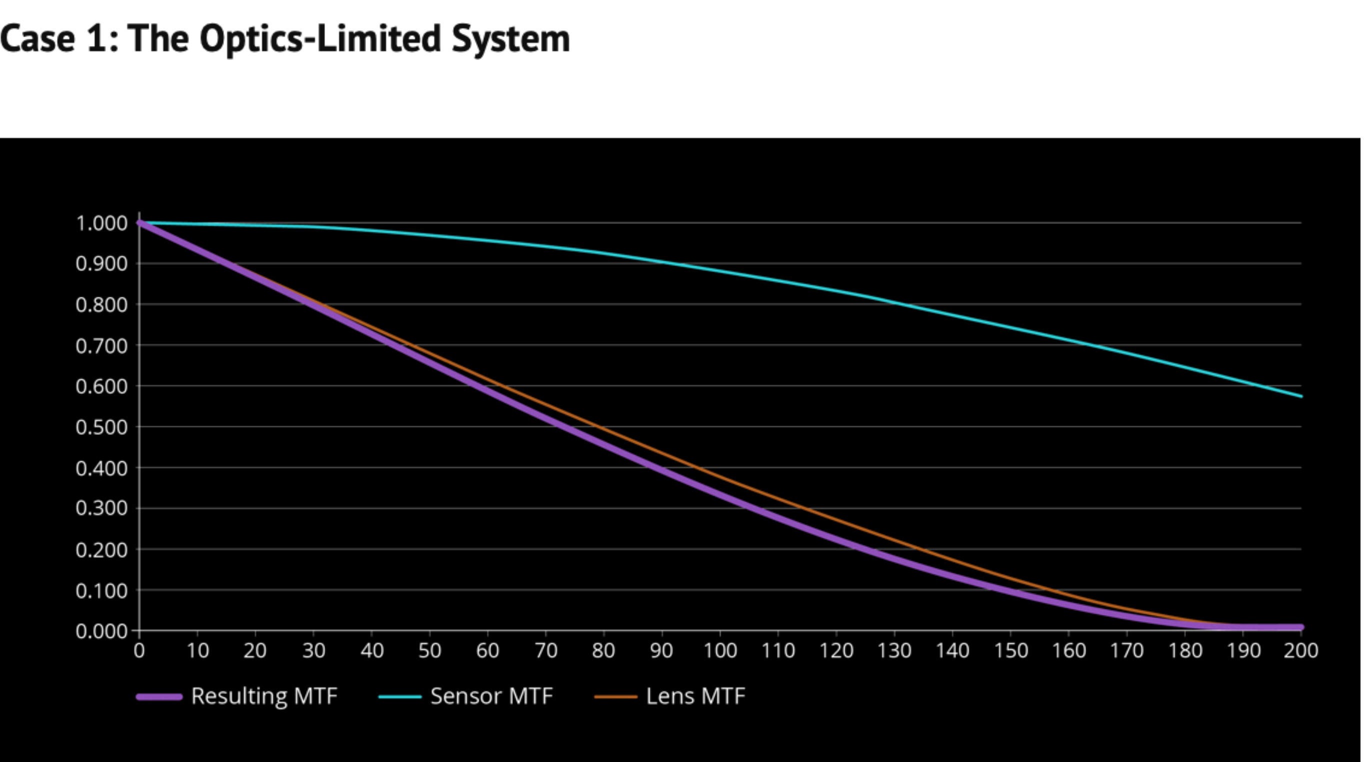
Optics - Design

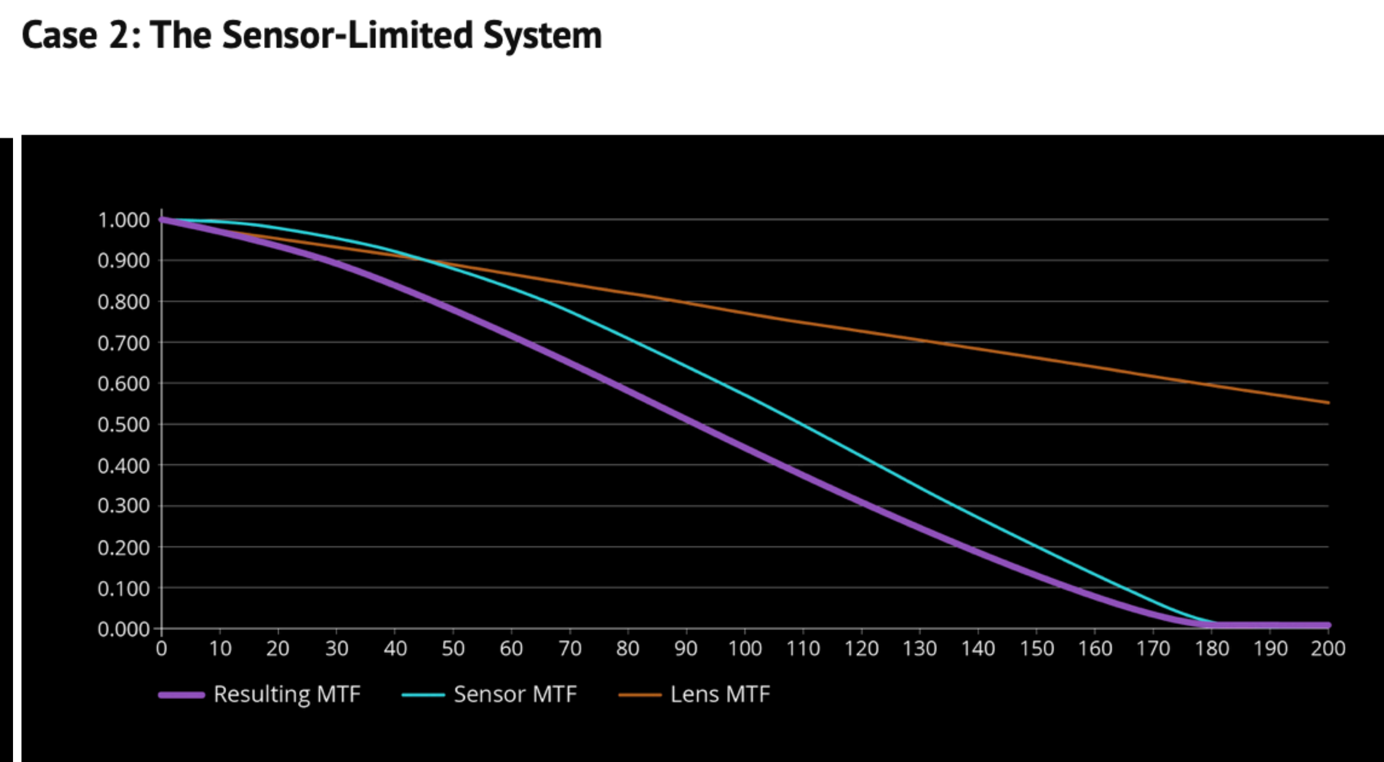
Things to consider:

- Larger apertures require large lenses.
- Large lenses weigh more, cost more, and are more difficult to manufacture.
- Adding more lens elements to reduce lens aberrations exacerbates the previously mentioned problem.
- Smaller pixels mean less light gathering capabilities.
- Less light gathering capabilities means greater image noise.

Optics - Design

- Different scenarios result, depending on the goodness of the lens MTF and the pixel MTF.





Images courtesy of Opto E

- Ideally, the design goal is to balance the two MTF's, so that neither is dominant.

Optics - Design

- Before trying to optimize the system MTF (=resulting MTF on the previous slide), there is one additional factor to consider.

- This factor is sample frequency, and the concept of aliasing.

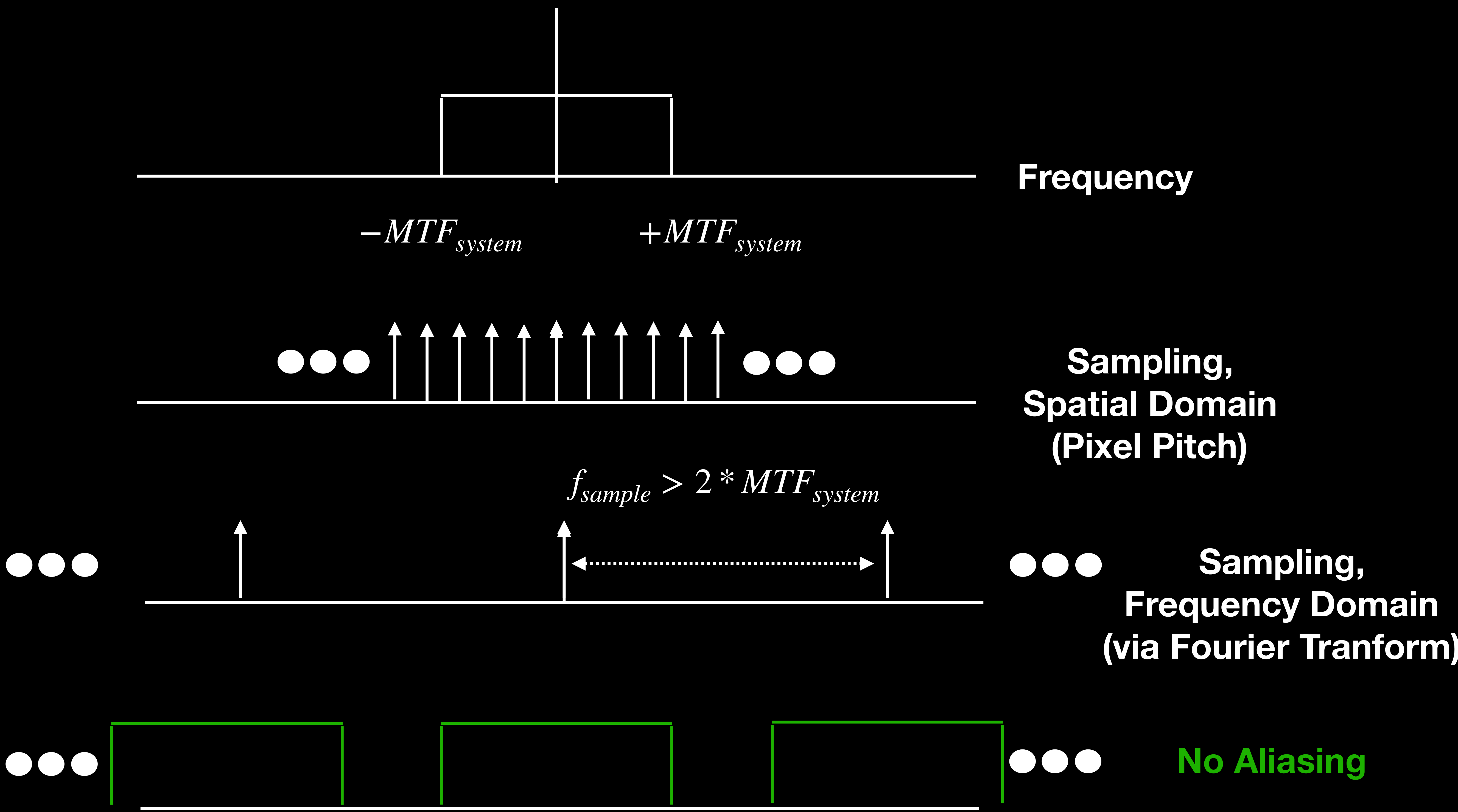
- From the Nyquist Sampling Theorem: To prevent aliasing artifacts, the sample frequency needs to be twice the frequency bandwidth of the baseband signal.

- But where is the sampling coming from, in our current system?

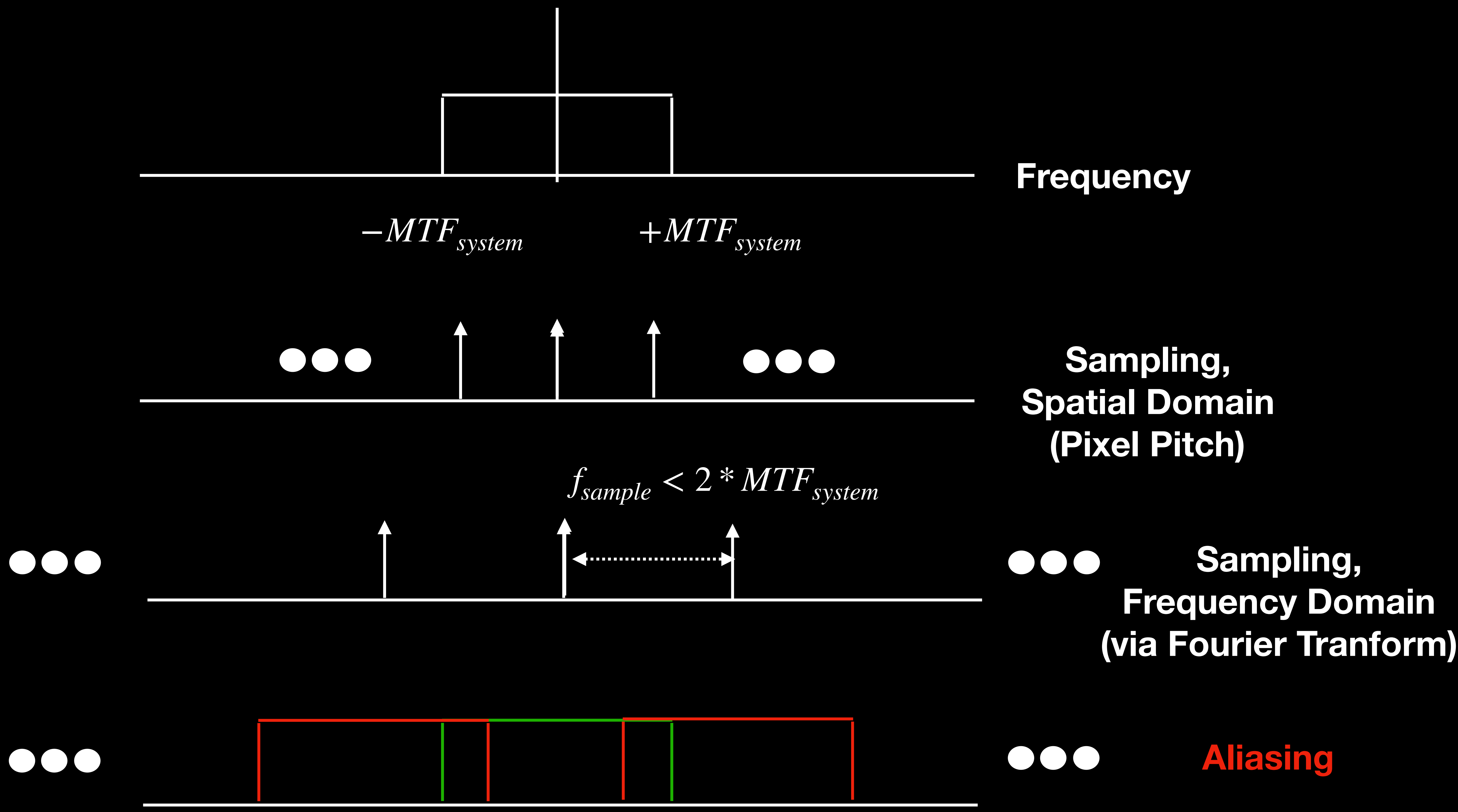
Optics - Design

- The spatial sampling comes from the pixel pitch.
- The fourier transform of the spatial sampling (then) defines the sampling used in the Nyquist Sampling Theorem.
- In practice, the pixel pitch is actually 2x larger than the distance between adjacent pixels, because a bayer color filter array (CFA) is used to perform the spatial sampling.
- This coarser sampling further exacerbates the aliasing problem.
- With this in mind, let’s illustrate the Nyquist Sampling Theorem when there is 1) no aliasing, 2) aliasing, and 3) “barely” no aliasing.

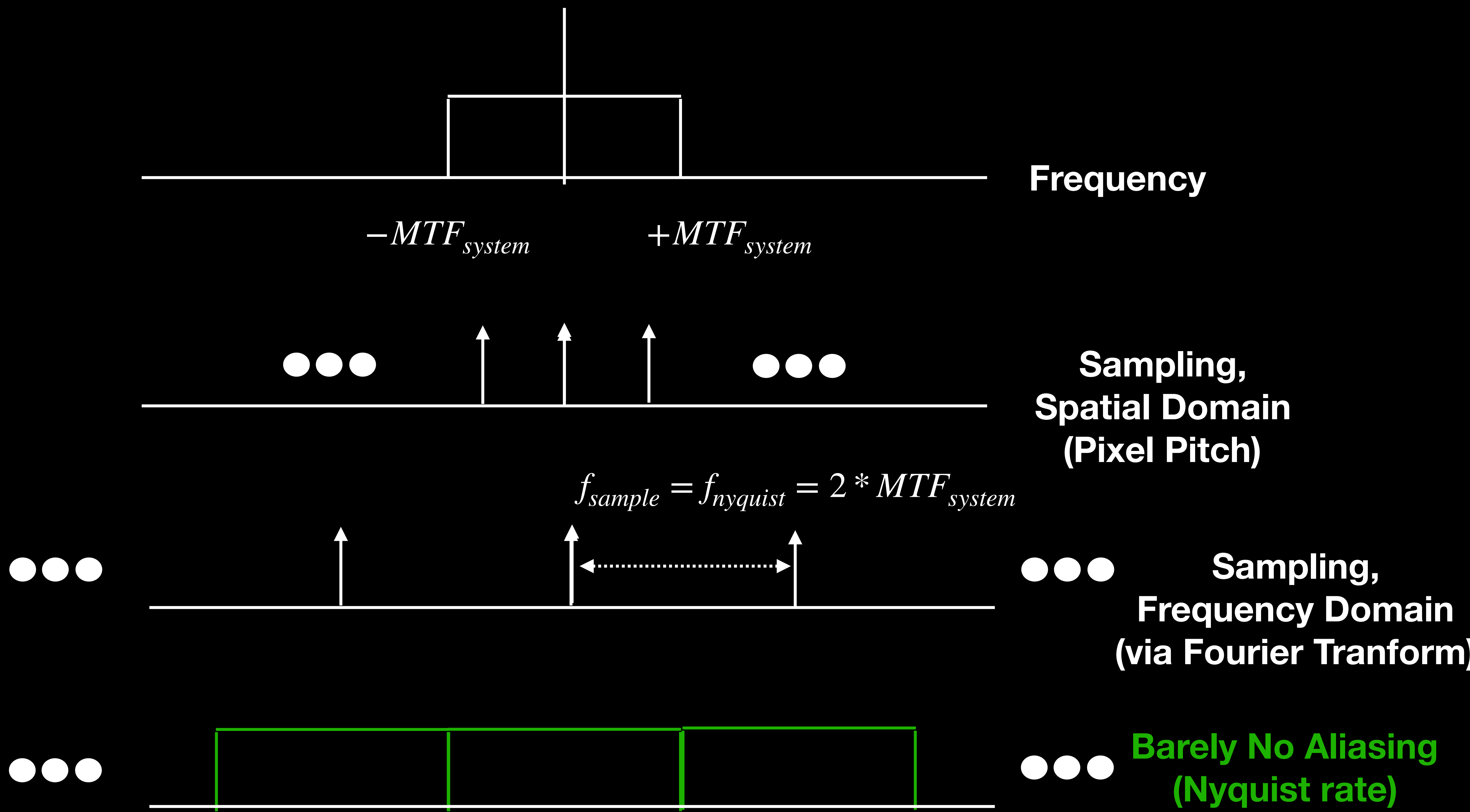
Optics - Design



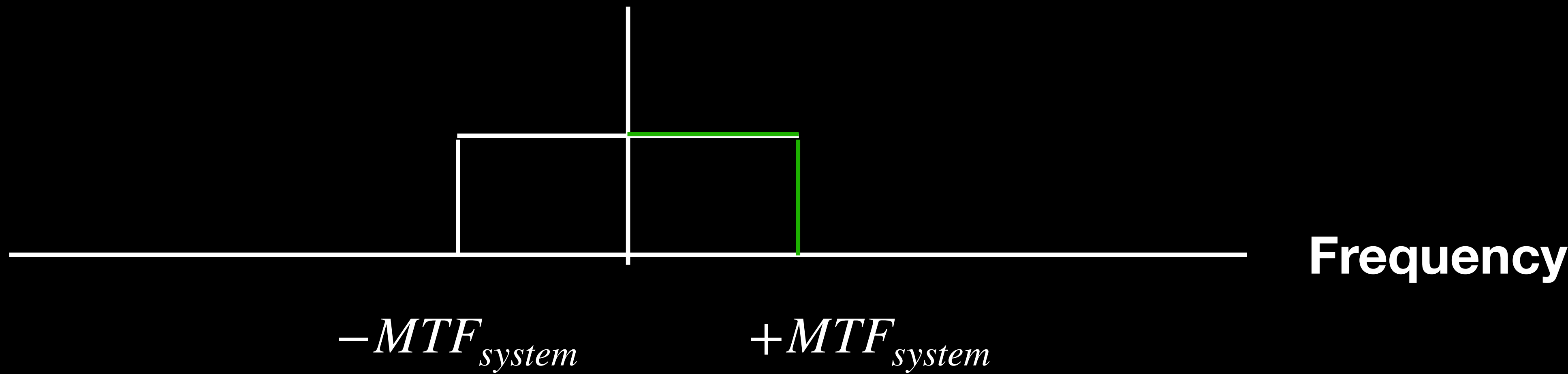
Optics - Design



Optics - Design



Optics - Design

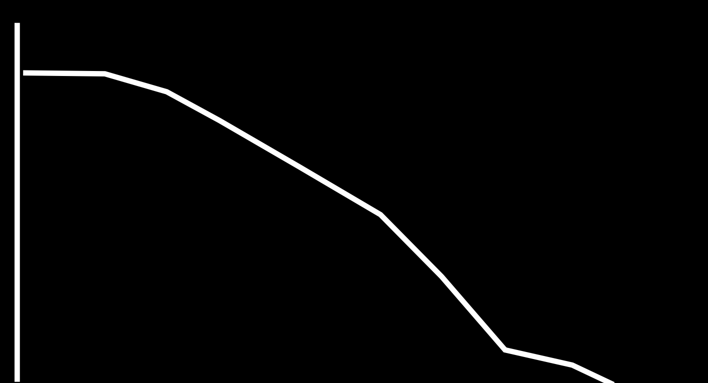


The green square represents an idealization of the frequency response.

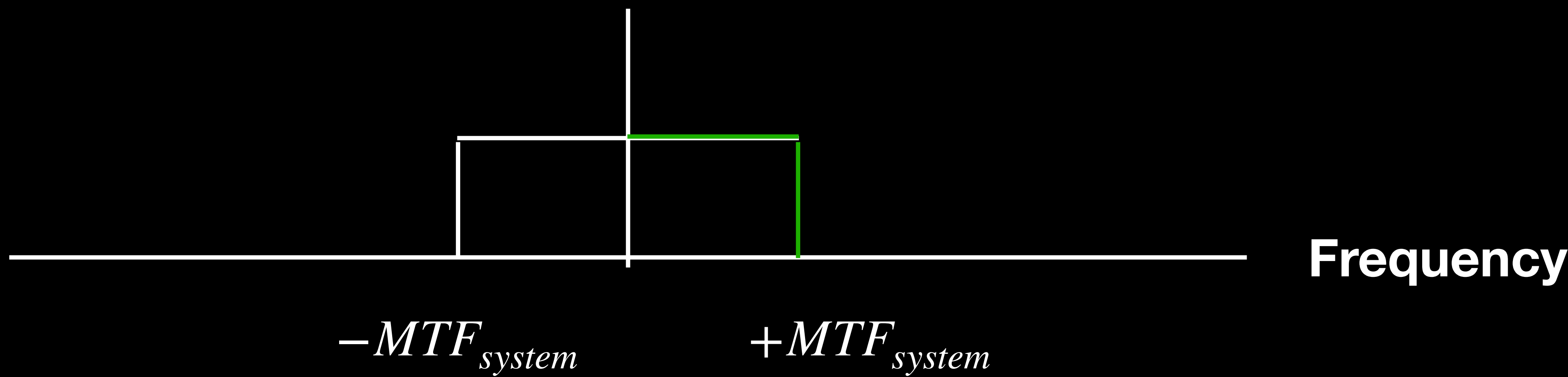
In practice, a decaying curve results.

To prevent aliasing, an optical low pass filter can be introduced, to eliminate the frequency content beyond a certain frequency.

i.e.



Optics - Design

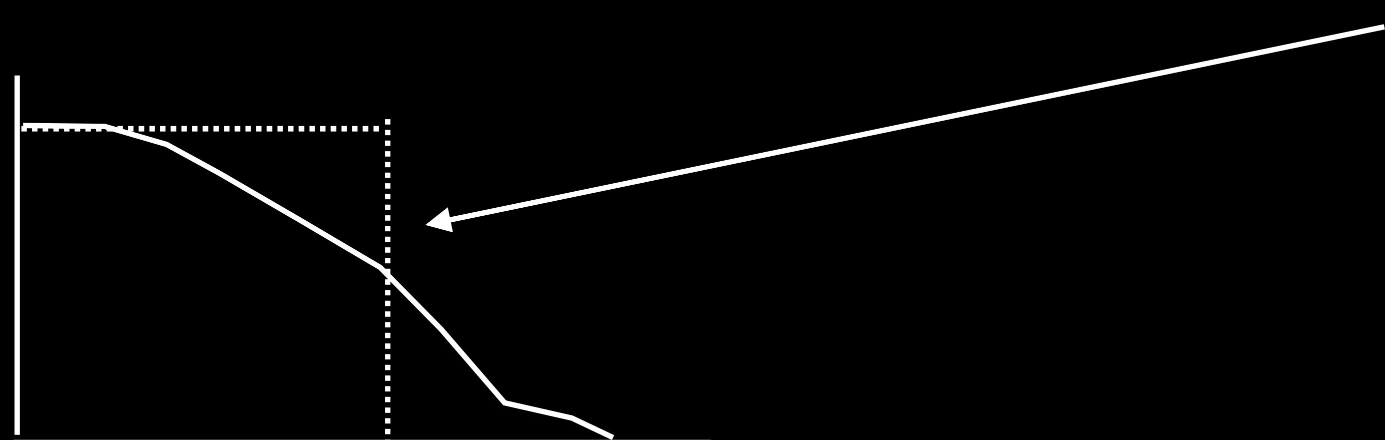


The green square represents an idealization of the frequency spectrum response.

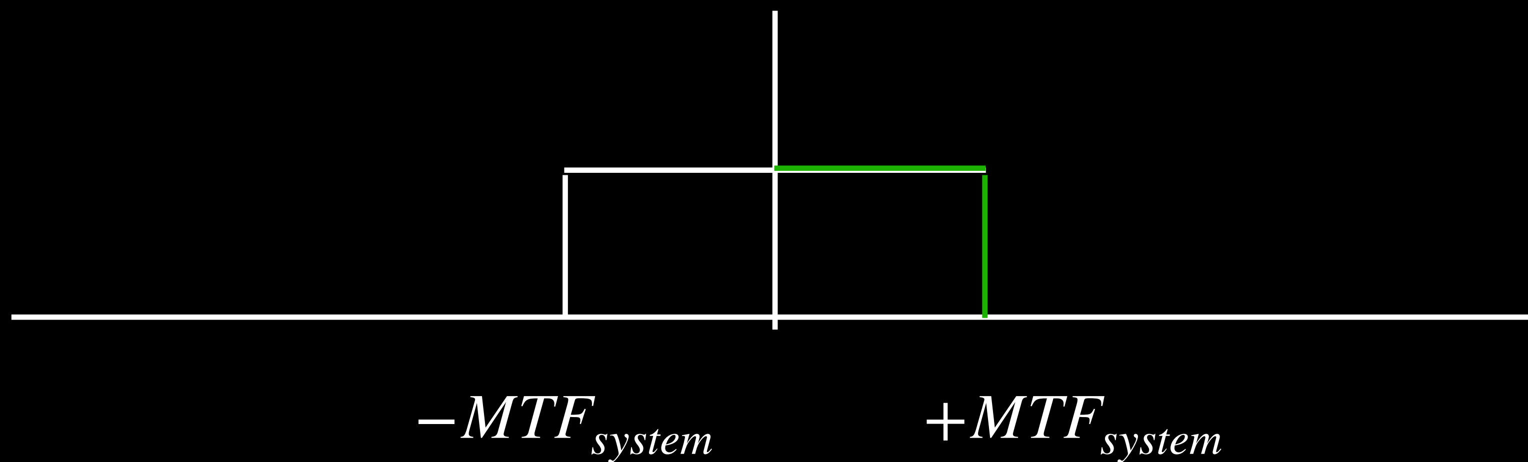
In practice, a decaying curve results.

To prevent aliasing, an optical low pass filter if often introduced, to eliminate the frequency content, beyond a certain frequency. Here, we add a theoretical brick wall filter.

i.e.



Optics - Design

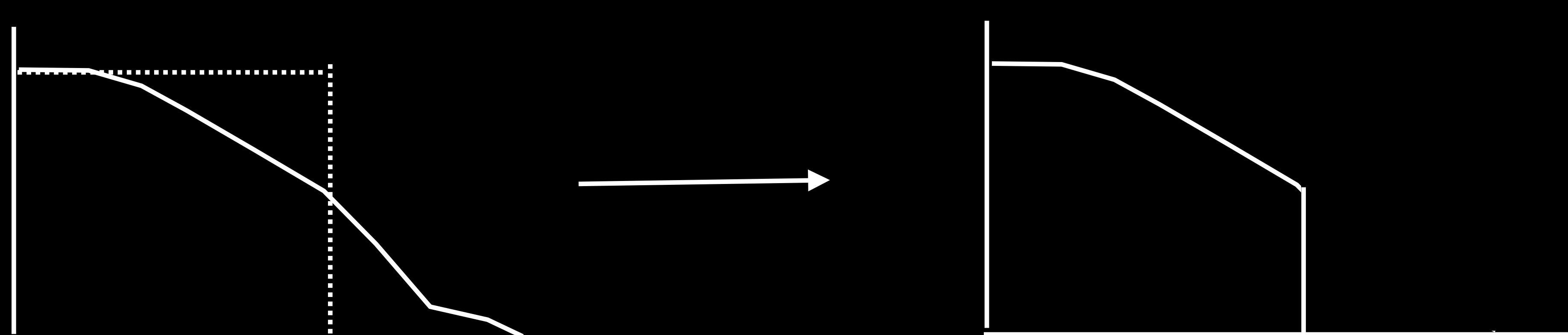


The green square represents an idealization of the frequency spectrum response.

In practice, a decaying curve results.

To prevent aliasing, an optical low pass filter is often introduced, to eliminate the frequency content, beyond a certain frequency. Here, we add a theoretical brick wall filter.

i.e.



To obtain the resulting output spectrum.

Optics - Design

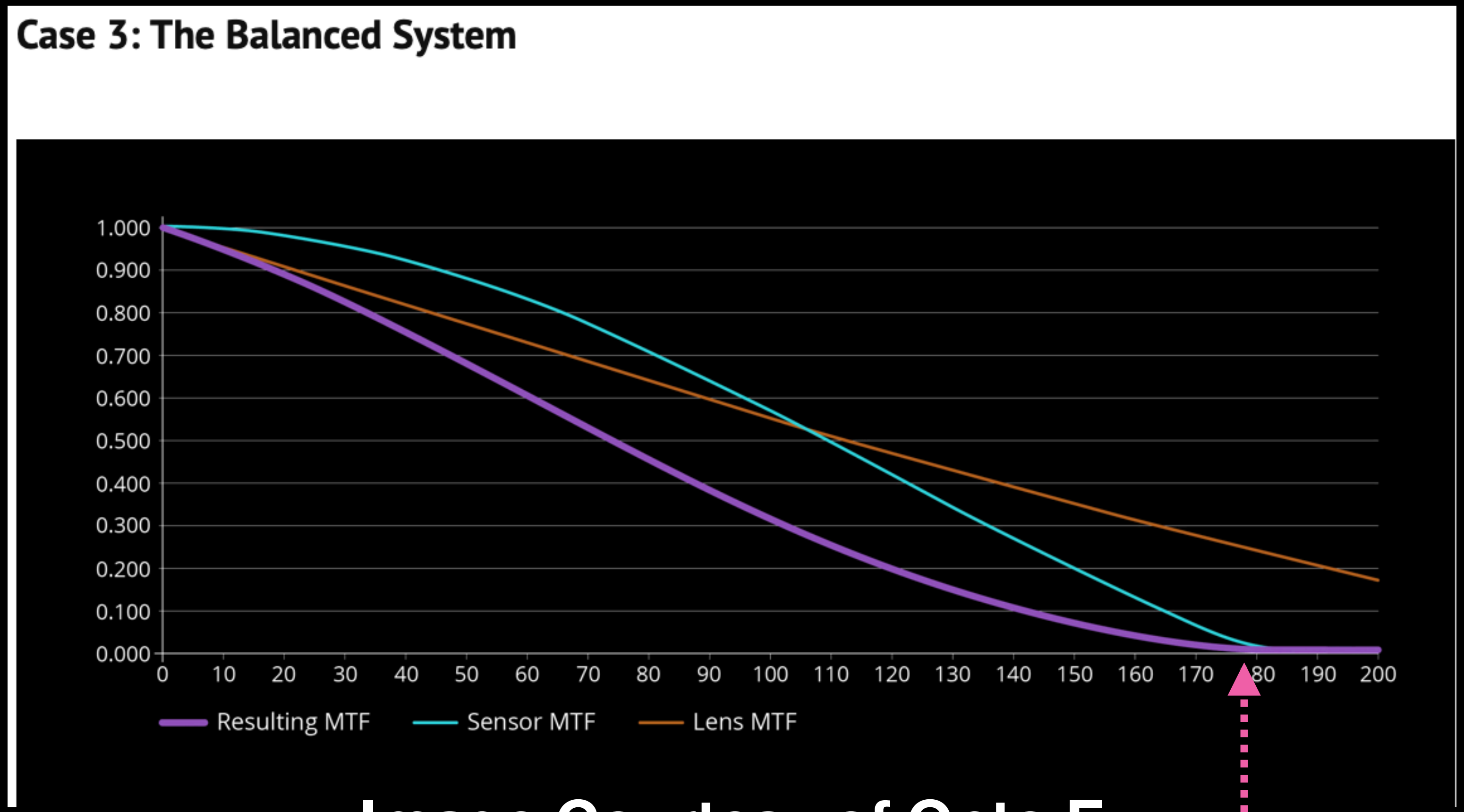


Image Courtesy of Opto E

In practice, we would like to design a system where contributions from the sensor MTF and lens MTF are balanced = neither is dominant.

In addition, the frequency response would ideally go to zero at a frequency that satisfies the Nyquist Sampling Theorem, eliminating the need for an optical low pass filter.

Optics - Design

- There are several important “numbers” associated with $MTF_{lens=diffraction+aberration}$ and MTF_{pixel}
- For $MTF_{diffraction}$ the spatial frequency at which the MTF = 0 is given as $f_{cutoff} = \dfrac{1}{\lambda N}$
- When we stop down a lens, this cutoff frequency decrease.
- The Nyquist frequency for MTF_{pixel} is given as $f_N = \dfrac{1}{2a}$
- At the Nyquist frequency, $MTF_{pixel} \sim 0.64$

Summary

- Lenses are used, to gather / collect light.

- With every lens, different DoF’s will result, based on the focal length of the lens, the focus position of the lens, the pixel size of the sensor, and the size of the lens aperture.

- As a result, a wide variety of images are created, with varying degrees of blur.

- Diffraction will always occur in every lens, due a finite lens aperture.

- Lens aberrations will also exist, due to lens imperfections.

Summary

Different lenses are designed / exist, to address different use cases = there is a wide range of lens options:

- Macro, wide angle, telephoto
- Apertures: f1.4, f1.8, f2.0, f2.8 … f16
- Single element to N lens elements
- Concave and convex lenses
- Plastic, glass and associated lens coatings, etc.
- Telescopic / folded optics, etc.

Summary

- The MTF curve captures a contrast versus spatial frequency relationship = a goodness of quality.
- MTF curves exist for the lens and the sensor.
- The Nyquist sampling frequency is defined as the fourier transform of the sampling performed by the sensor pixel pitch.
- To prevent aliasing, the Nyquist sampling frequency needs to be 2x the bandwidth of the system MTF curve.