Optics à la mode - a new way of making, using and understanding optics

stanford.io/3NK9w8Z

David Miller

Stanford University

Optics à la mode?

Optics à la mode

served with ice cream?

No!

in the most modern style or fashion?

Sort of …

Noting too that fashions change

so our “optics à la mode” would have to be able to change to suit the “fashion”

programmable optics

and even self-configuring optics!

a pun on the word “mode”

Yes!

stanford.io/3NK9w8Z

Optical components

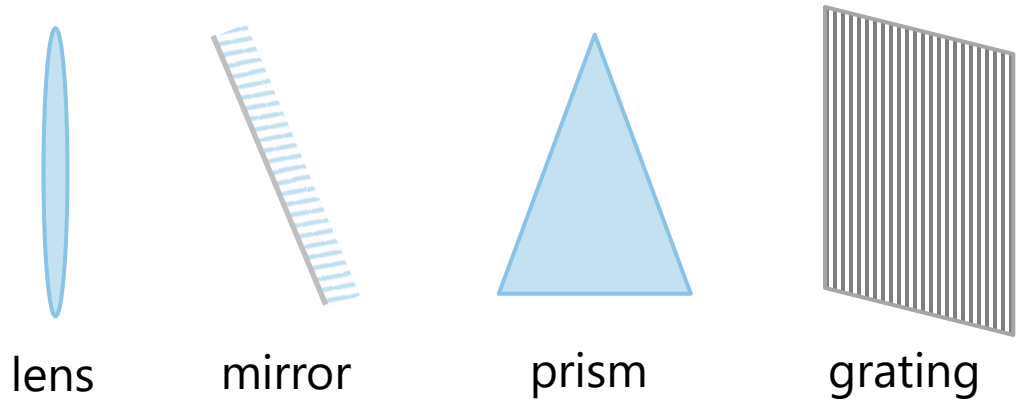
Classic optics

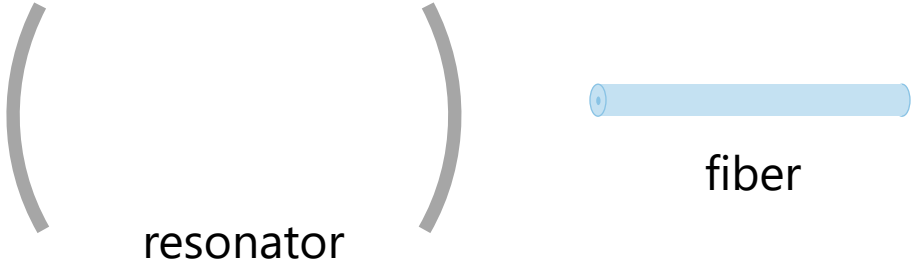
rays, imaging, plane waves

basic spectroscopy

Lasers and optical fiber communications

add “modes”





New optical circuits

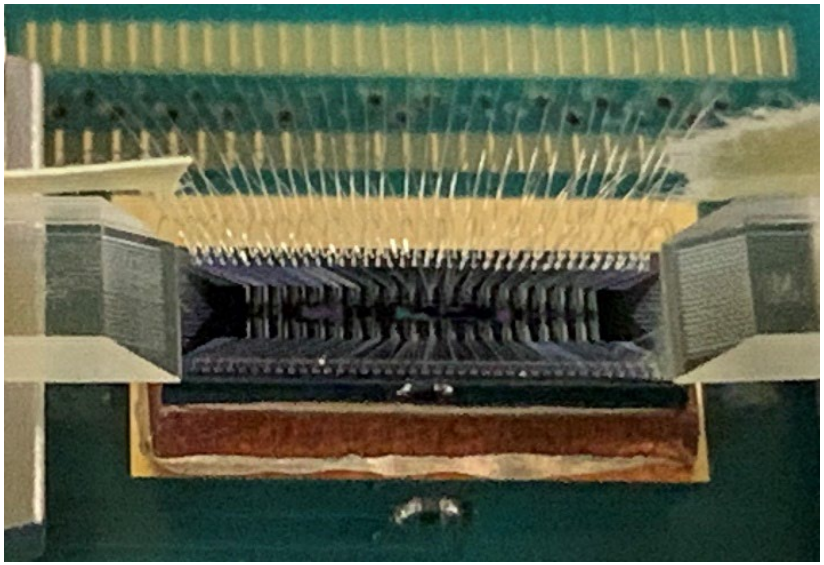
This silicon photonics circuit

is also a piece of optics

- with an array of fibers as its input
- and an array of fibers as its output
- and the ability to arbitrarily define the relation between input and output

How do we think about this as optics?

Is there a broader approach that includes this as well as existing optics and gives us new insights and tools?



Sunil Pai et al., Stanford

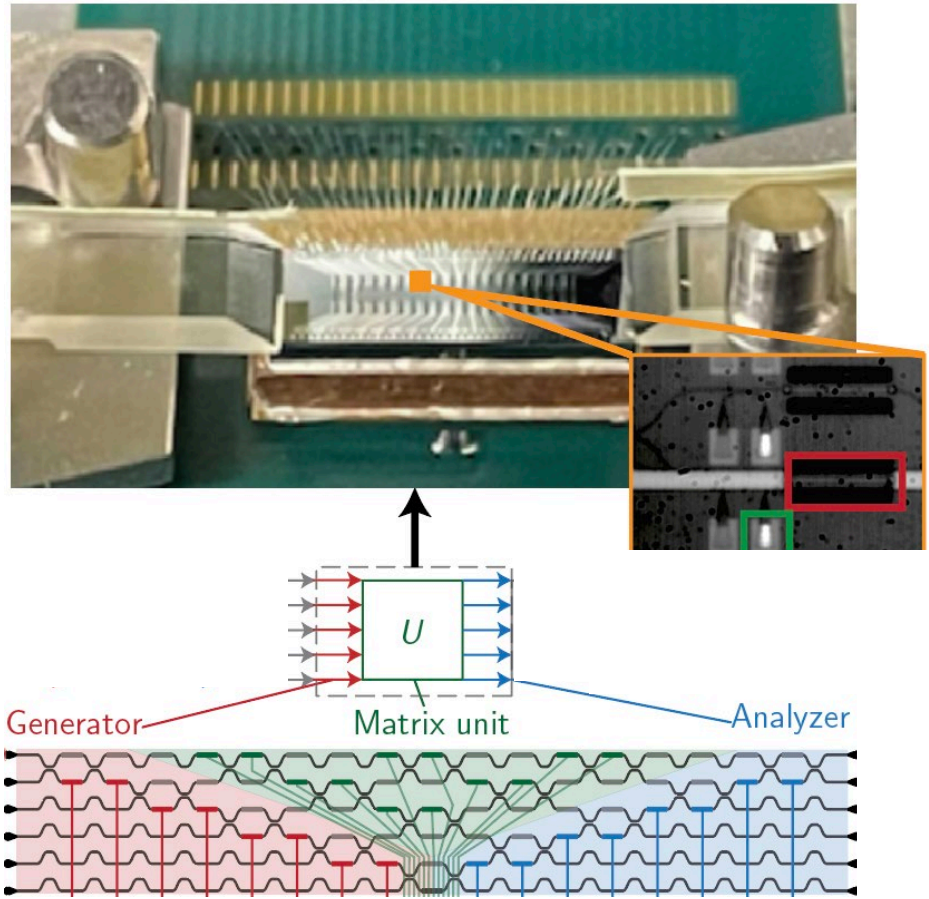
Universal matrix multiplier chip

Universal matrix multiplying chip

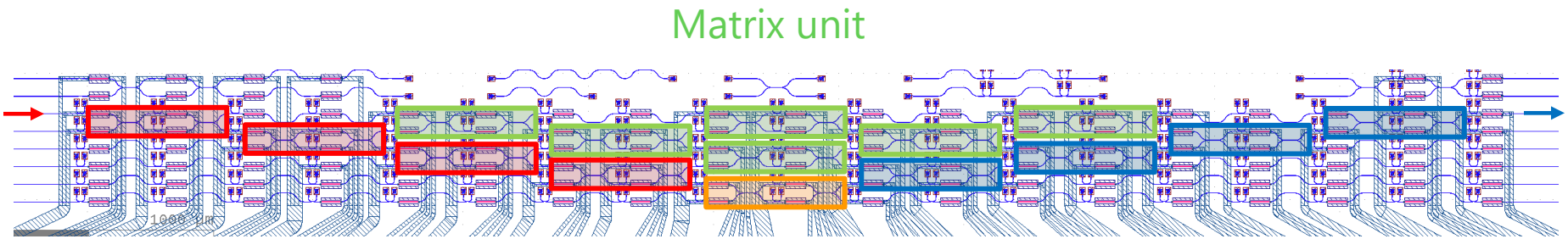
"4x4" unitary Mach-Zehnder mesh with

- a "generator" to create any complex input vector
- an "analyzer" to measure the complex output vector

This can be programmed to implement any "unitary" (loss-less) transformation from the inputs to the outputs



Mask layout and block diagram



Universal matrix multiplier chip

Full complex matrix multiplication

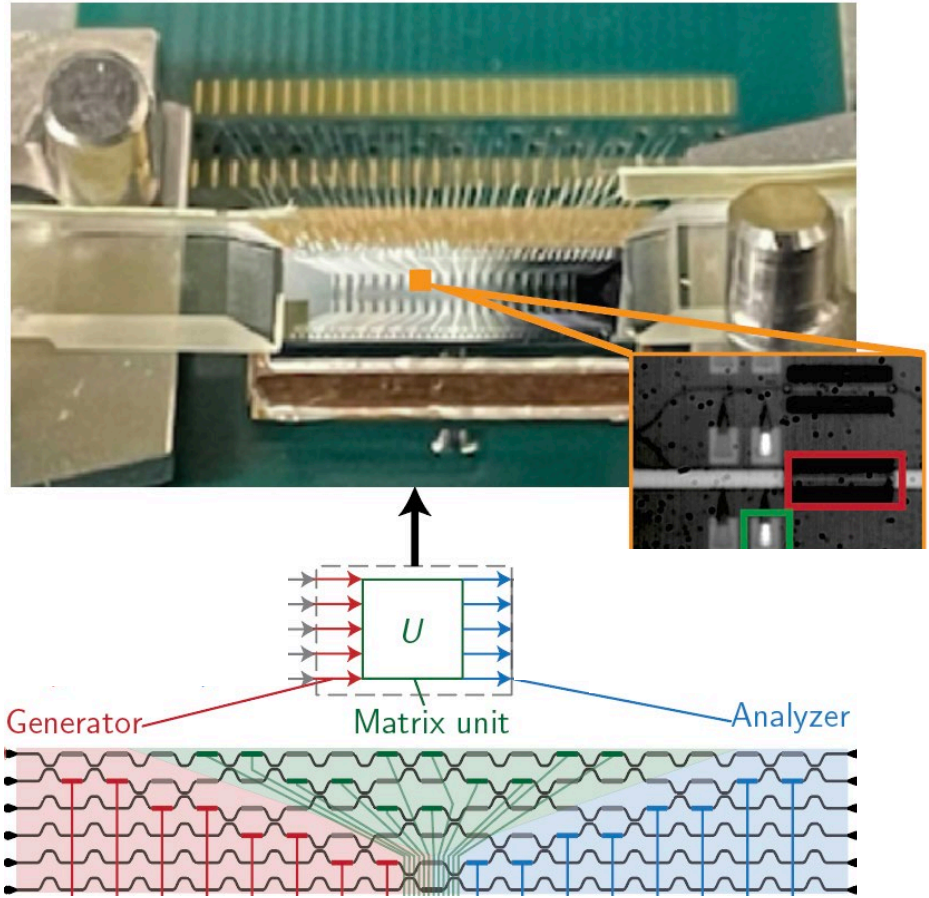
with vector generation and vector analysis

Photonic back-propagation neural net training

S. Pai, Z. Sun, T. W. Hughes, T. Park, B. Bartlett, I. A. D. Williamson, M. Minkov, M. Milanizadeh, N. Abebe, F. Morichetti, A Melloni, S. Fan, O. Solgaard, D. A. B. Miller, "Experimentally realized in situ backpropagation for deep learning in photonic neural networks," **Science** 380, 398-404 (2023)

Digital matrix multiplication for cryptography

S. Pai, T. Park, M. Ball, B. Penkovsky, M. Dubrovsky, N. Abebe, M. Milanizadeh, F. Morichetti, A. Melloni, S. Fan, O. Solgaard, and D. A. B. Miller, "Experimental evaluation of digitally verifiable photonic computing for blockchain and cryptocurrency," **Optica** 10, 552-560 (2023)



Self-configuring beam separator

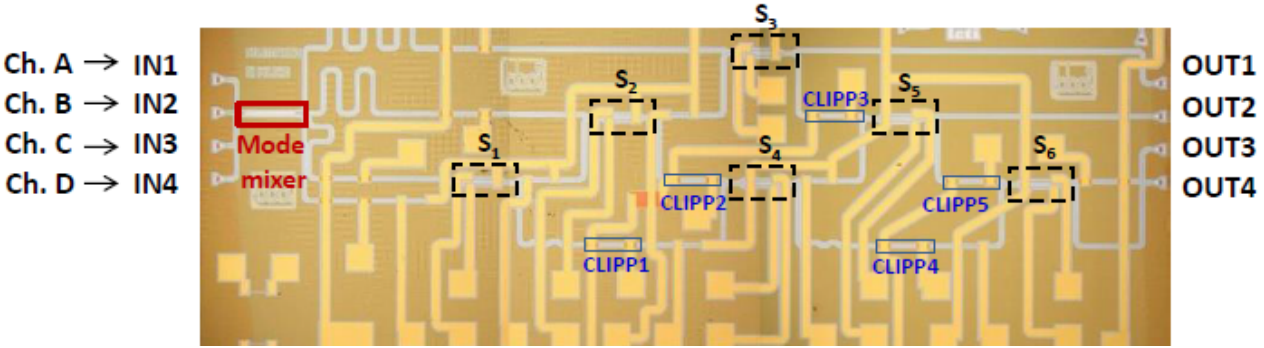
Light from four input fibers

deliberately mixed in a mode mixer

are automatically separated out again by a mesh of interferometers

by sequential power maximizations

without calculations



A. Annoni et al., "Unscrambling light – automatically undoing strong mixing between modes," Light Science & Applications 6, e17110 (2017)

See, e.g., review W. Bogaerts et al., "Programmable photonic circuits," Nature 586, 207 (2020)

Truly arbitrary linear optics

stanford.io/3NK9w8Z

Simple optical components – a mirror

We “design” a plane mirror

by choosing its angle

so it takes a beam of one angle

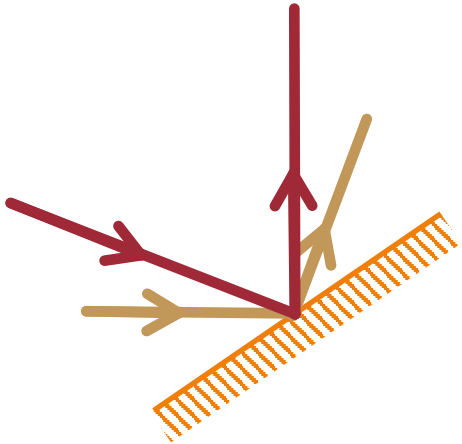
and changes it into a beam of another angle

For another beam at another angle

the mirror changes it to a beam of yet another angle

but we have no independent control

over what happens for the second beam



Simple optical components – a lens

We design a lens

by choosing its index and curvatures

so it takes a plane wave in one direction

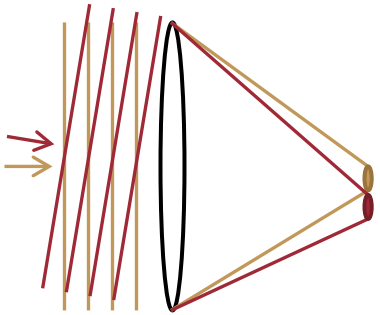
and focuses it to a spot

For another plane wave in another direction

the lens focuses it to another spot

but we have no independent control

over what happens for the second beam



Simple “thin” optical components

This kind of behavior is general for “thin” optical components

e.g., thin holograms, diffractive optical elements

spatial light modulators, adaptive optics,

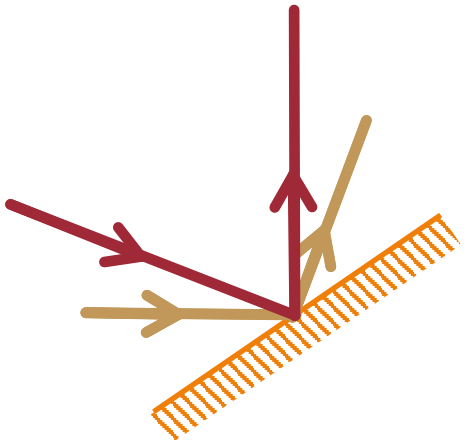
metasurfaces

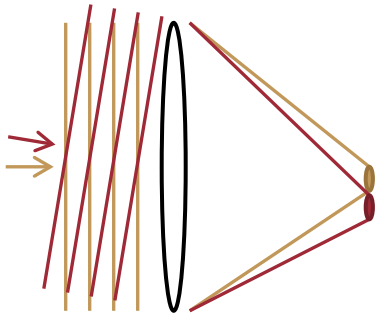
We design them to perform some useful function for one input beam

but we have no independent control

of what happens for other beams

So these are not “arbitrary” optical components





Example - Separating overlapping beams

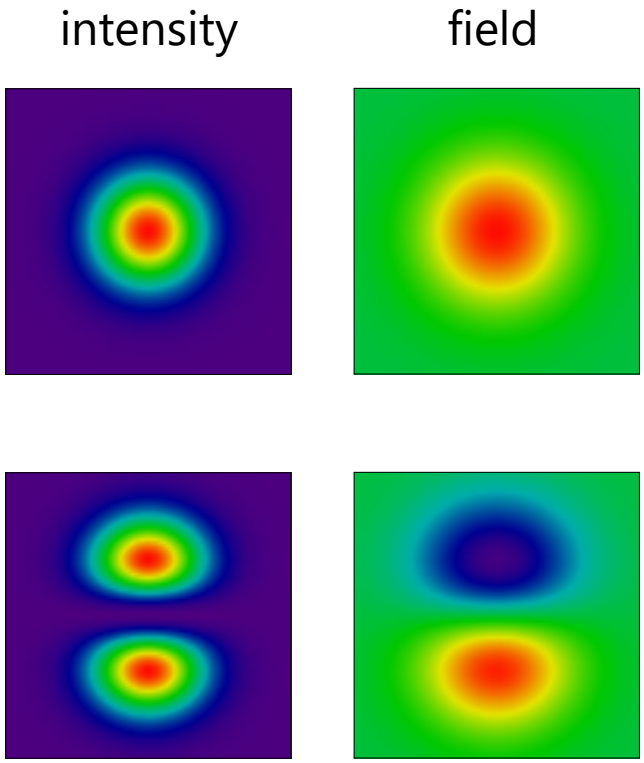
Suppose we have two different (orthogonal) beams

e.g., from an optical fiber such as

a “single bump” beam

and

a “two bump” beam



Example - Separating overlapping beams

Mathematically,

two (non-zero) beams are "orthogonal" if

$$\iint \mathbf{E}_1^*(x,y)\cdot\mathbf{E}_2(x,y)\,dxdy = 0$$

Here, the product of the

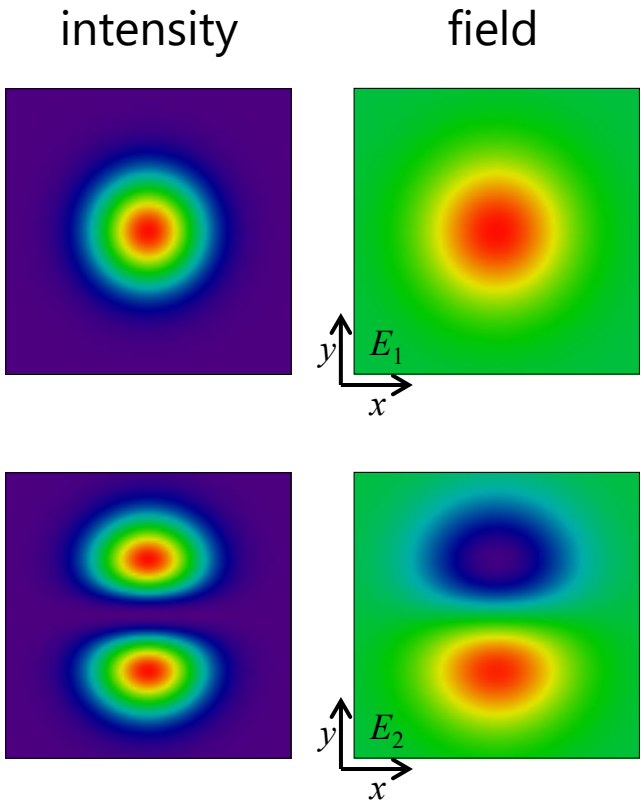
single-bump beam and the

two-bump beam

would be negative in the top half

but positive in the bottom half

so the resulting integral would be zero



Example - Separating overlapping beams

If both of these beams emerge simultaneously from the fiber

how can we separate them

for example to different fibers

without loss?

Example - Separating overlapping beams

If both of these beams emerge simultaneously from the fiber

how can we separate them

for example to different fibers

without loss?

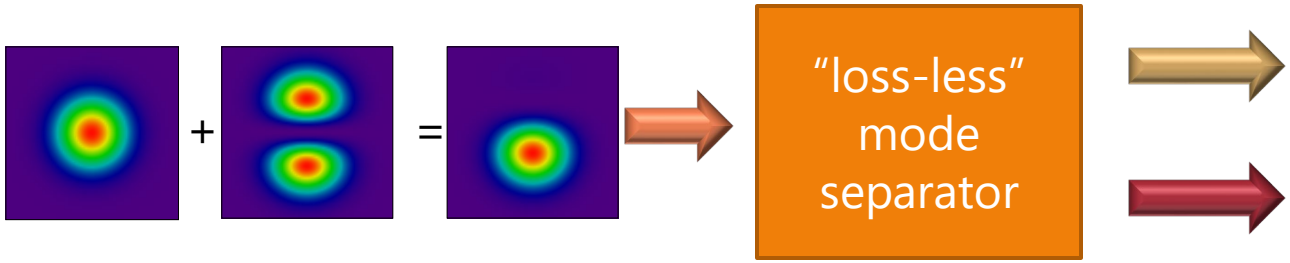
Example - Separating overlapping beams

If both of these beams emerge simultaneously from the fiber

how can we separate them

for example to different fibers

without loss?



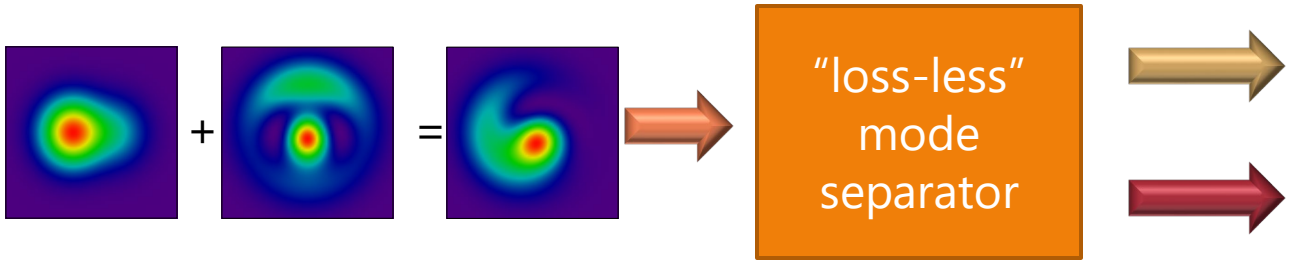
Example - Separating overlapping beams

In situations with

fixed

highly symmetric beams

good specific low-loss separation solutions are known



But for general cases

of lower symmetry and/or higher complexity

or where the beams change in time

general solutions have not been known

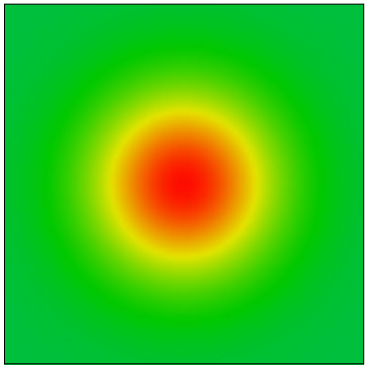
Dividing the beam into “patches”

We can approach the beam-separation problem

by presuming it will be good enough

to imagine that we can divide the beam

into a finite number of “patches”



Dividing the beam into “patches”

We can approach the beam-separation problem

by presuming it will be good enough

to imagine that we can divide the beam

into a finite number of “patches”

We treat each of these patches

as if it was approximately uniform

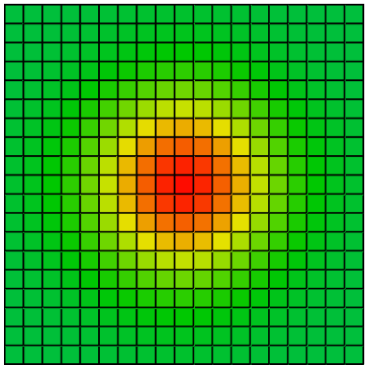
in intensity and

in phase

At least with a sufficiently large number of patches

this could be a good enough approximation

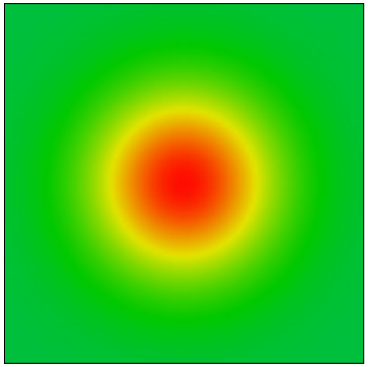
and “sampling loss” may be small

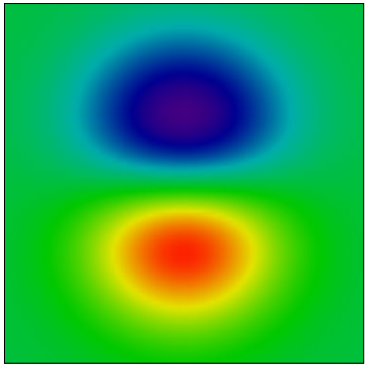


Dividing the beam into “patches”

Even relatively small numbers of patches

- are sufficient to distinguish beams of low or moderate complexity

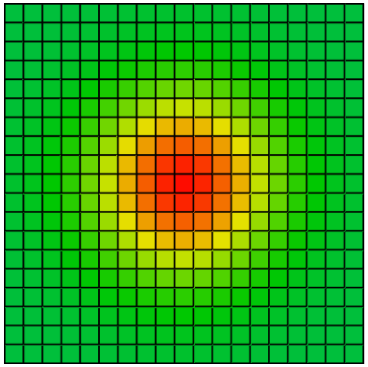


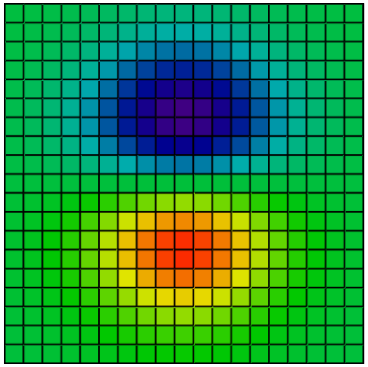


Dividing the beam into “patches”

Even relatively small numbers of patches

- are sufficient to distinguish beams of low or moderate complexity

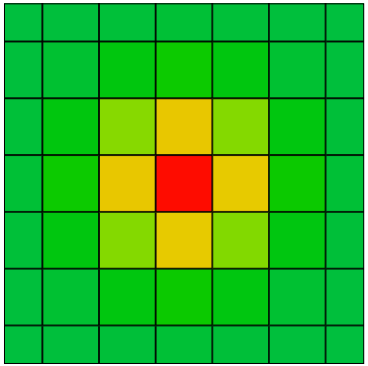


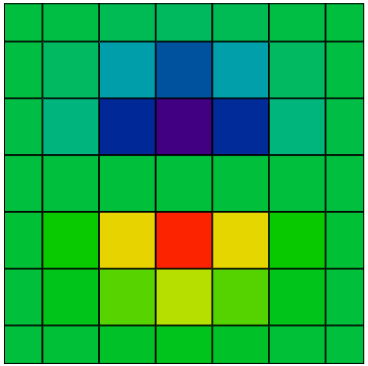


Dividing the beam into “patches”

Even relatively small numbers of patches

- are sufficient to distinguish beams of low or moderate complexity





Separating beams

So, by dividing the beam into patches

we may be able to approximate the problem

by one of finite dimensions

Indeed, the approach discussed here

will only be practical for problems of limited dimensionality

e.g., 10's or possibly 100's

But still, even in principle

how can we separate these beams?

Mach-Zehnder interferometer meshes

stanford.io/3NK9w8Z

Nulling a Mach-Zehnder output

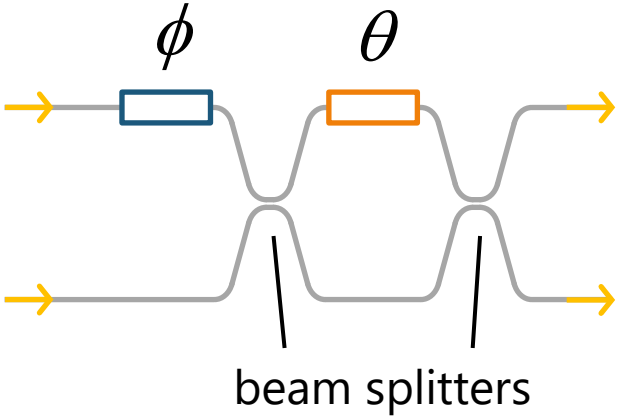
Consider a waveguide Mach-Zehnder interferometer (MZI)

formed from two "50:50" beam splitters

and at least two phase shifters

one, ϕ, to control the relative phase of the two inputs

a second, θ, to control the relative phase on the interferometer "arms"



Nulling a Mach-Zehnder output

In such an MZI with 50:50 beamsplitters

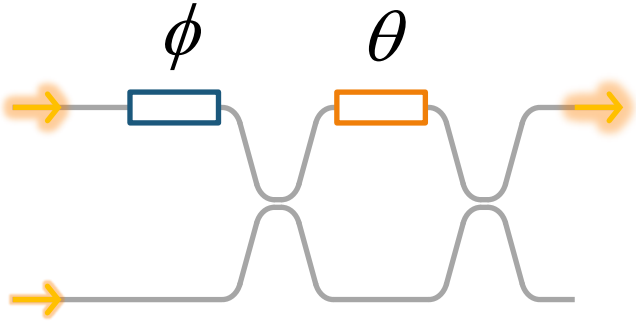
for any relative input amplitudes and phases

we can "null" out the power at the bottom output

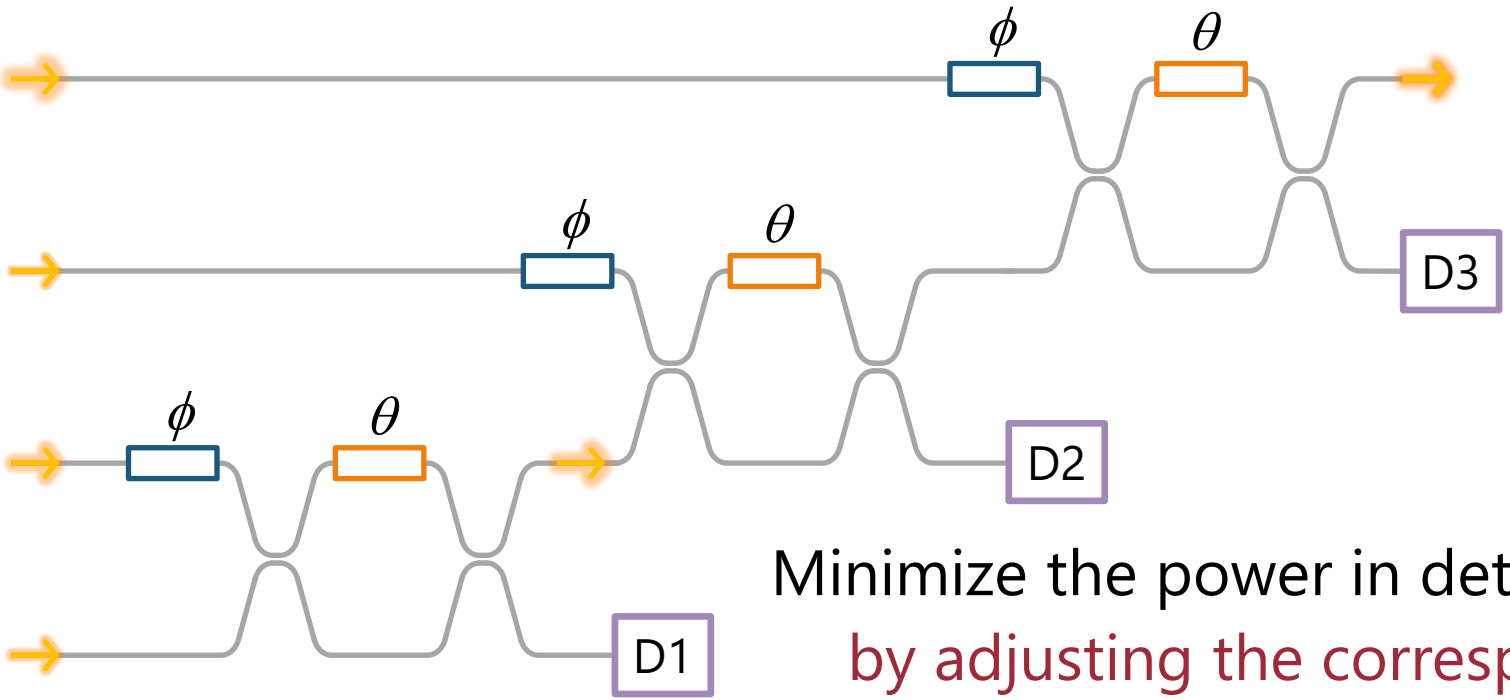
by two successive single-parameter power minimizations

first, using ϕ

second, using θ



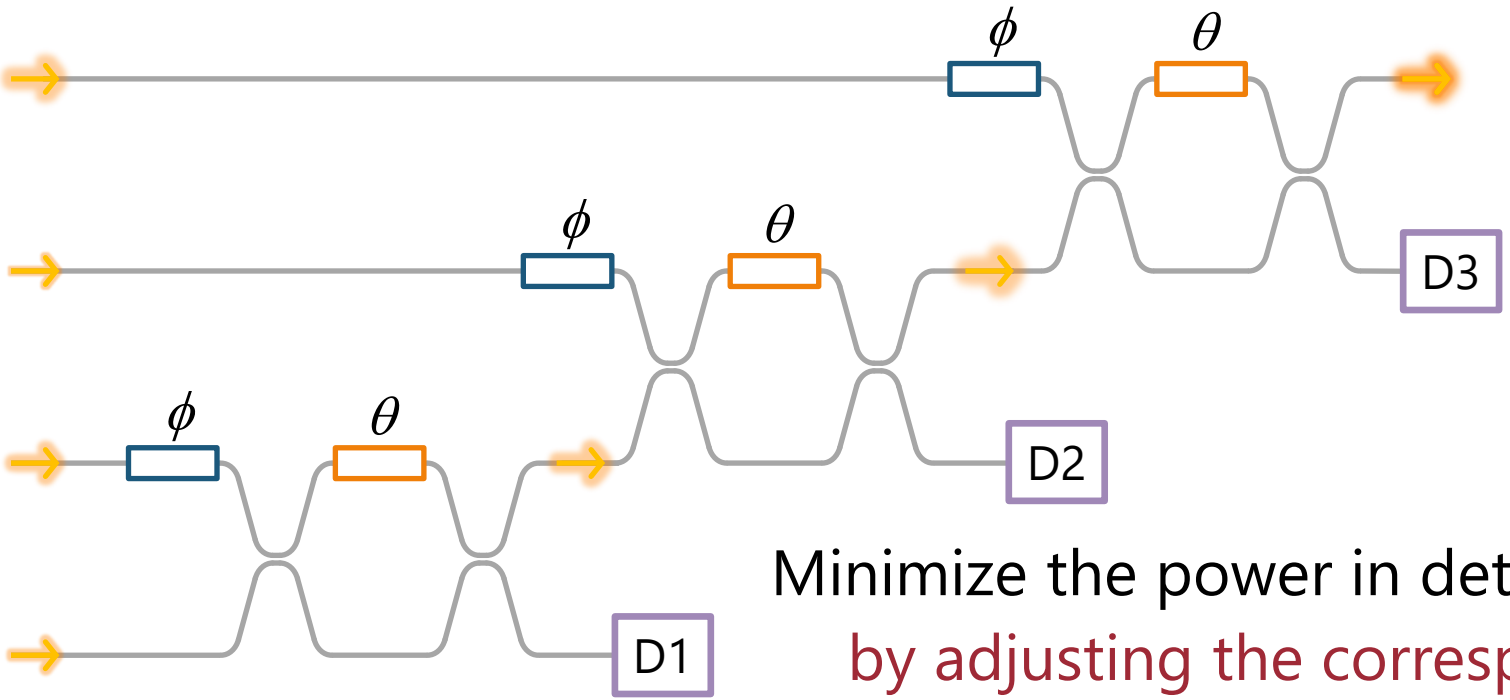
“Diagonal line” self-aligning coupler



"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Minimize the power in detector D1
by adjusting the corresponding ϕ
and then θ
putting all power in the upper output

“Diagonal line” self-aligning coupler



"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

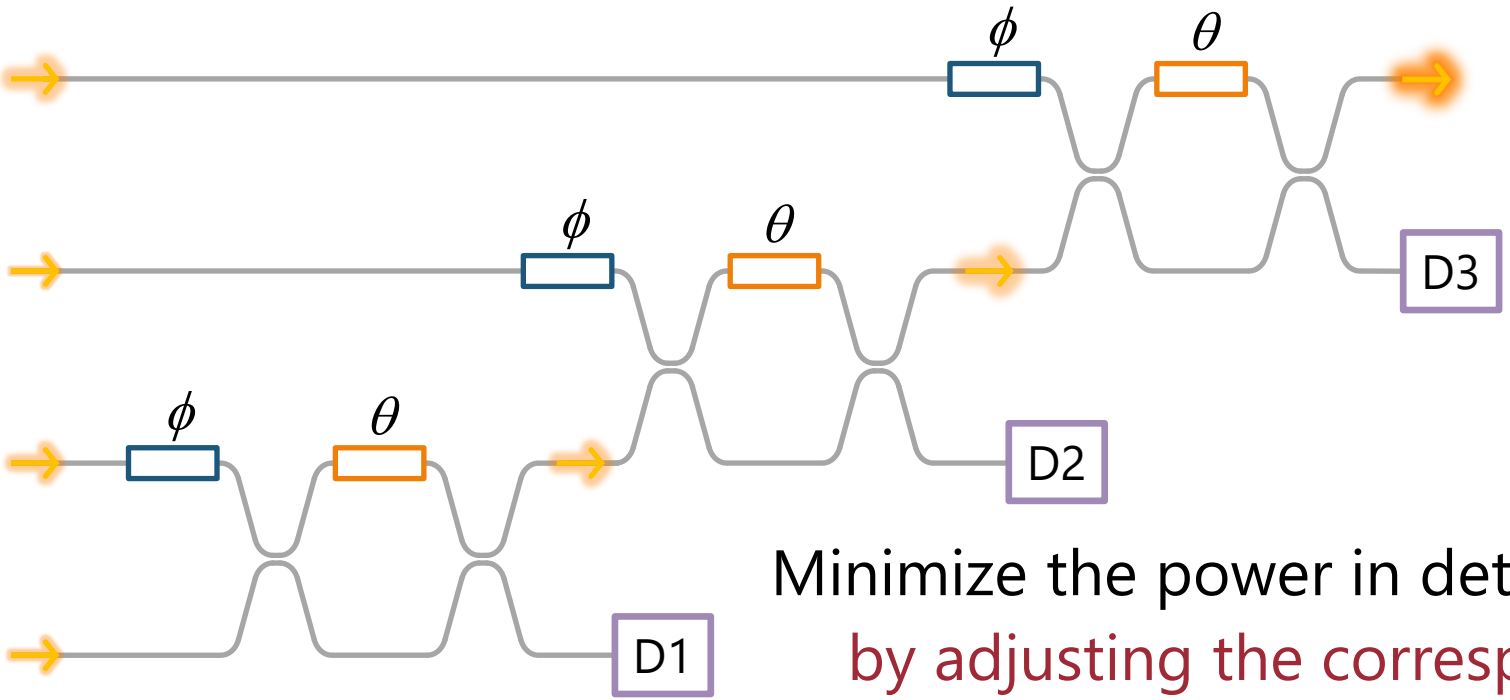
Minimize the power in detector D2

by adjusting the corresponding ϕ

and then θ

putting all power in the upper output

“Diagonal line” self-aligning coupler



"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Minimize the power in detector D3
by adjusting the corresponding ϕ
and then θ
putting all power in the upper output

Self-aligning beam coupler

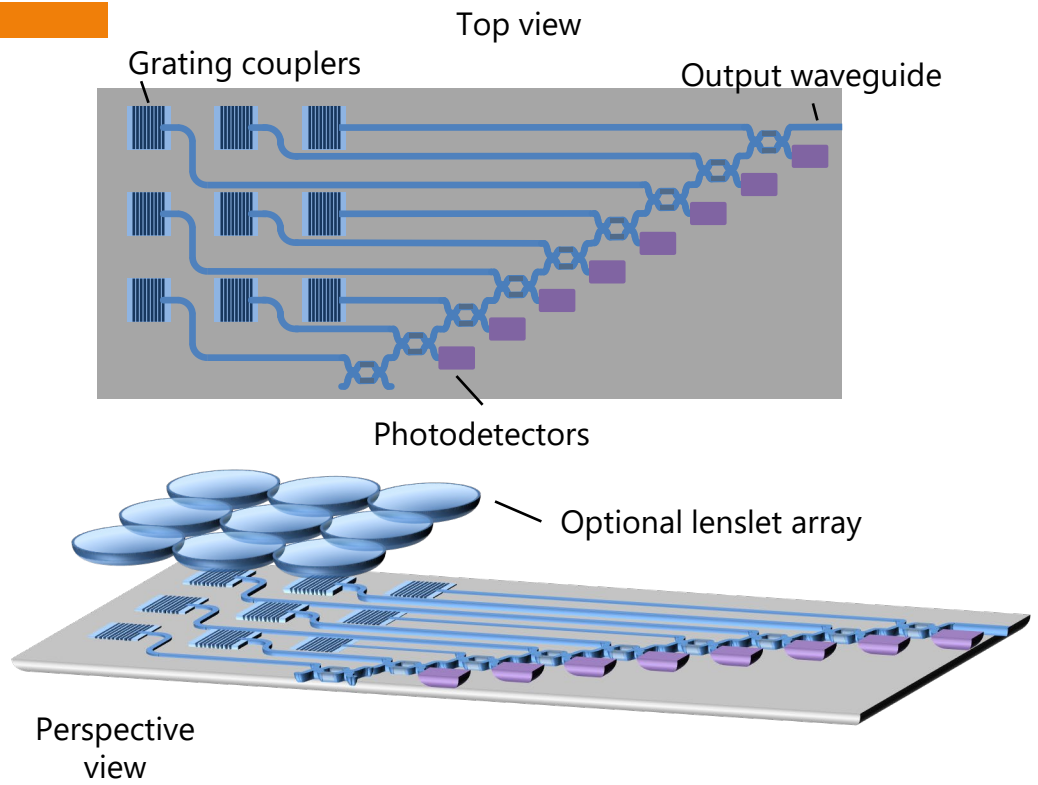
Grating couplers could couple a free-space beam to a set of waveguides

Then

we could automatically couple all the power to the one output guide

This could run continuously

tracking changes in the beam

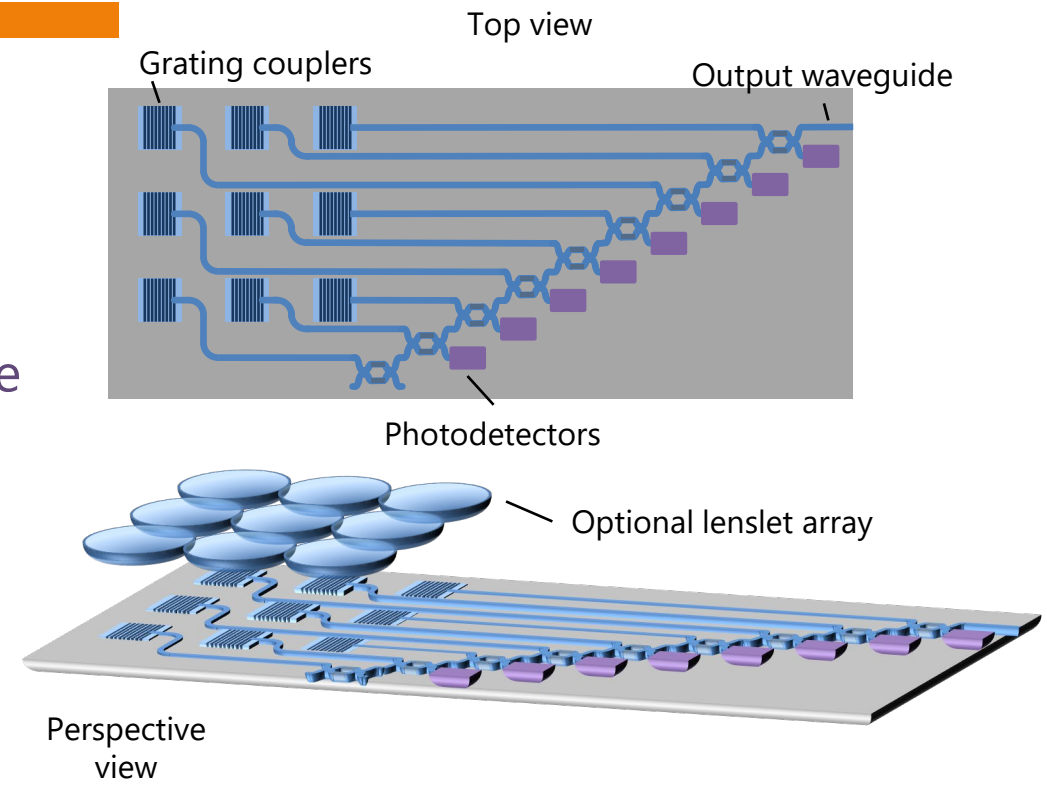


"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Self-aligning beam coupler

This has several different uses

- tracking an input source
  both in angle and focusing
- correcting for aberrations
- analyzing amplitude and phase of the components of a beam
  - …

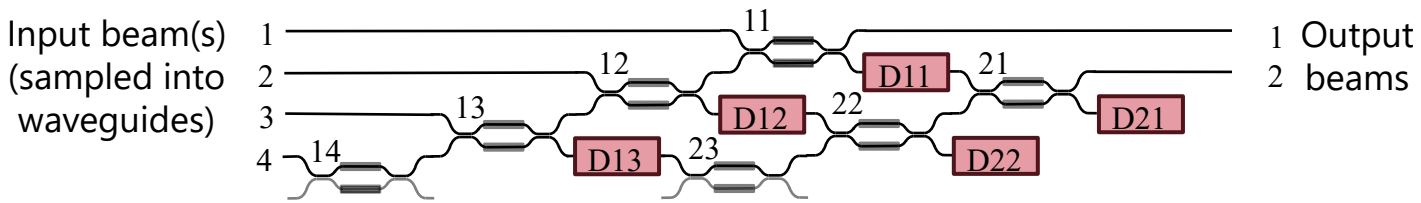


"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Separating beams with interferometer meshes

stanford.io/3NK9w8Z

Separating multiple orthogonal beams



"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Once we have aligned beam 1 to output 1 using detectors D11 – D13

an orthogonal input beam 2 would pass entirely into the detectors D11 – D13

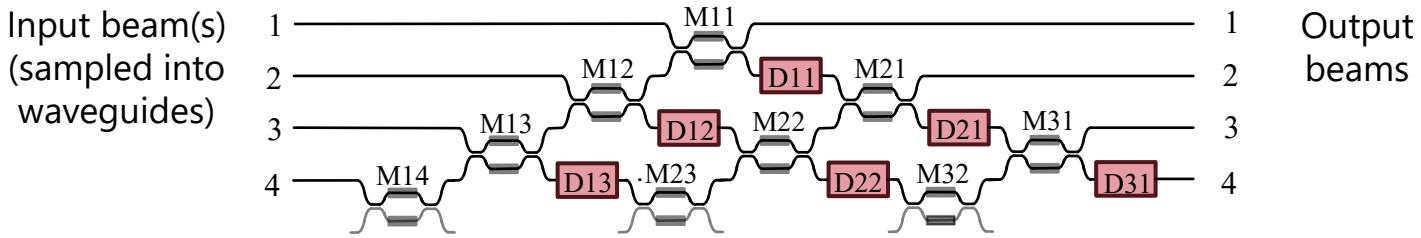
If we make these detectors mostly transparent

this second beam would pass into the second diagonal "row"

where we self-align it to output 2 using detectors D21 – D22

separating two overlapping orthogonal beams to separate outputs

Separating multiple orthogonal beams



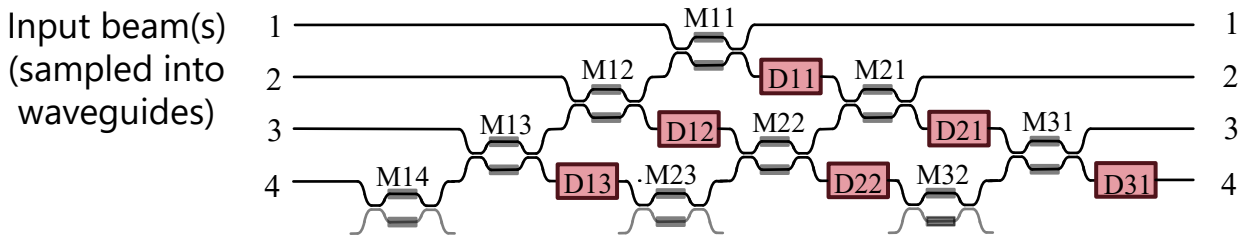
"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

Adding more rows and self-alignments

separates a number of orthogonal beams

equal to the number of beam "segments", here, 4

Separating multiple orthogonal beams



"Self-aligning universal beam coupler," Opt. Express **21**, 6360 (2013)

If we put identifying “tones” on each orthogonal input “beam”

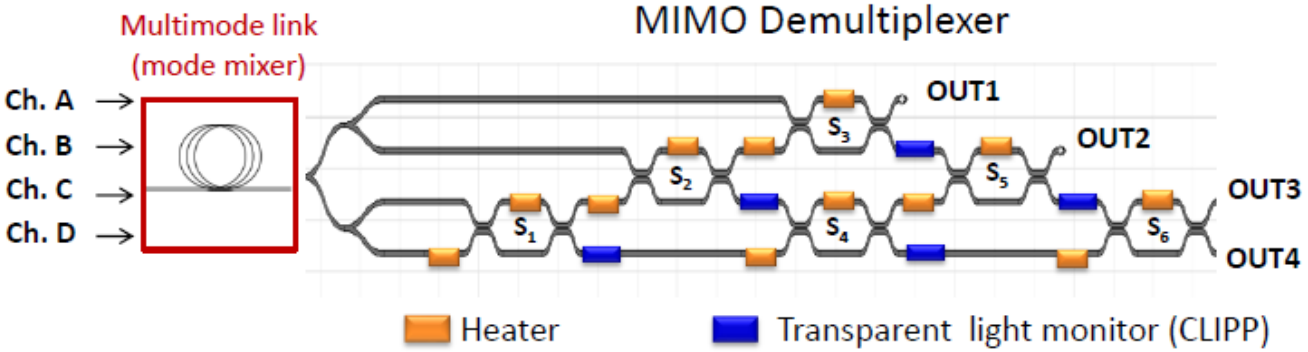
and have the corresponding diagonal row of detectors look for that tone

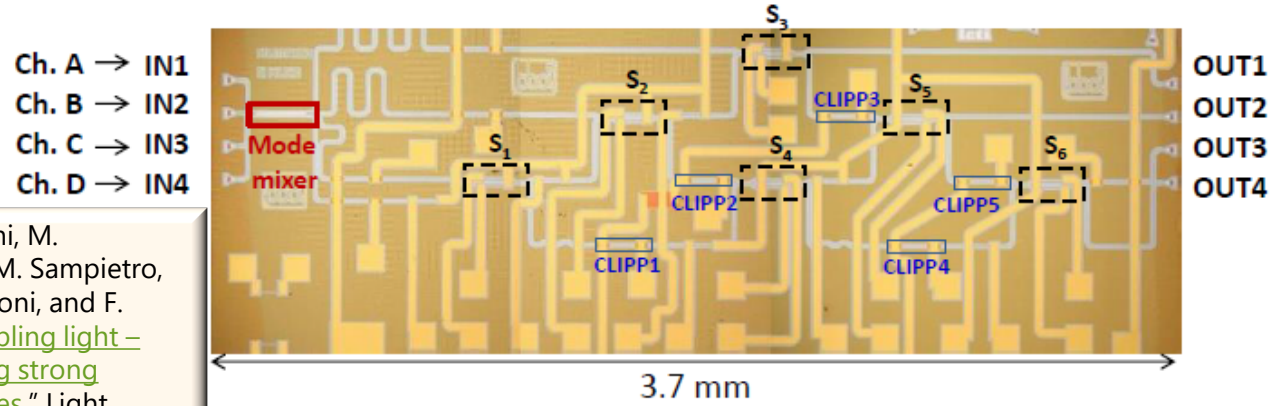
then the mesh can continually adapt to the orthogonal inputs

even when they are all present at the same time

and even if they change

Integrated MIMO demultiplexer: technology





- Transparent detectors required for sequential tuning
- CLIPP-assisted circuit reconfiguration & feedback control

A. Annoni, E. Guglielmi, M. Carminati, G. Ferrari, M. Sampietro, D. A. B. Miller, A. Melloni, and F. Morichetti, "Unscrambling light – automatically undoing strong mixing between modes," Light Science & Applications 6, e17110 (2017)

Si chip fabricated by CEA-Leti

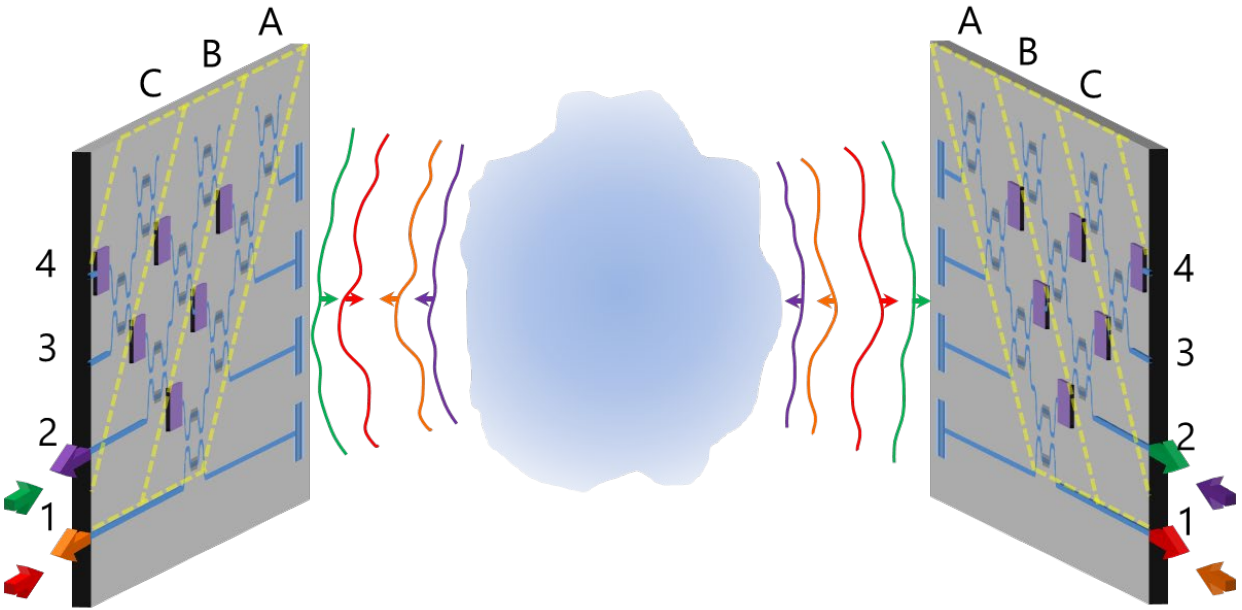
Establishing optimum orthogonal channels

In this architecture, using meshes on both sides

we proposed we could find optimal orthogonal channels through a scatterer

between waveguides on the left and waveguides on the right

by iterating back and forward between the two sides



"Establishing optimal wave communication channels automatically," J. Lightwave Technol. 31, 3987 (2013)

Establishing optimum orthogonal channels - experiment

We have now demonstrated this with two “facing” interferometer meshes with

- arbitrary optics between them

The optics can be misaligned

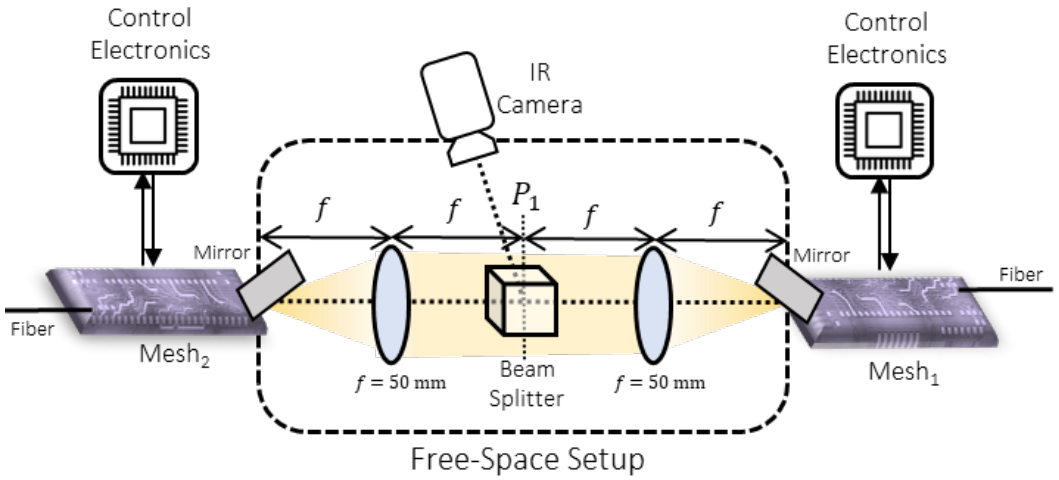
- and we can introduce
  - aberrations or partial blocking
- in the path

The system still self-aligns to find

- the best, orthogonal channels

This uses simple power optimizations

- on one “channel” at a time
  - at the output of each line of interferometers at the receiving end

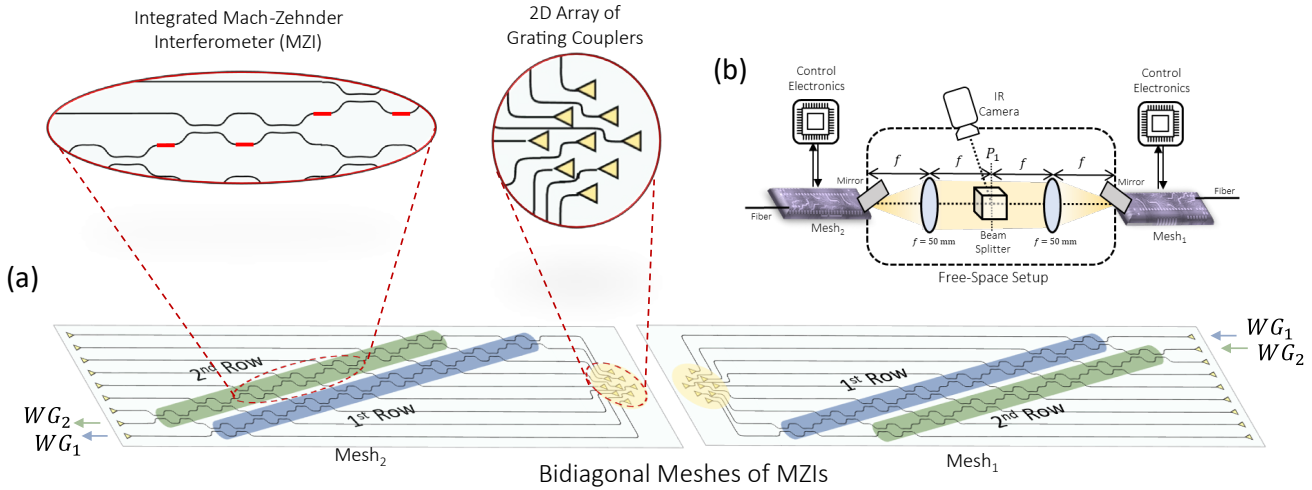


S. SeyedinNavadeh, M. Milanizadeh, F. Zanetto, G. Ferrari, M. Sampietro, M. Sorel, D. A. B. Miller, A. Melloni, and F. Morichetti, "Determining the optimal communication channels of arbitrary optical systems using integrated photonic processors," Nat. Photon. **18**, 149-155 (2024)

Establishing optimum orthogonal channels - experiment

Two “9x2” meshes allow automatic self-configuration

- signals in WG1 on the right can automatically be aligned to appear out of WG1 on the left, and, at the same time
- signals in WG2 on the right can automatically be aligned to appear out of WG2 on the left



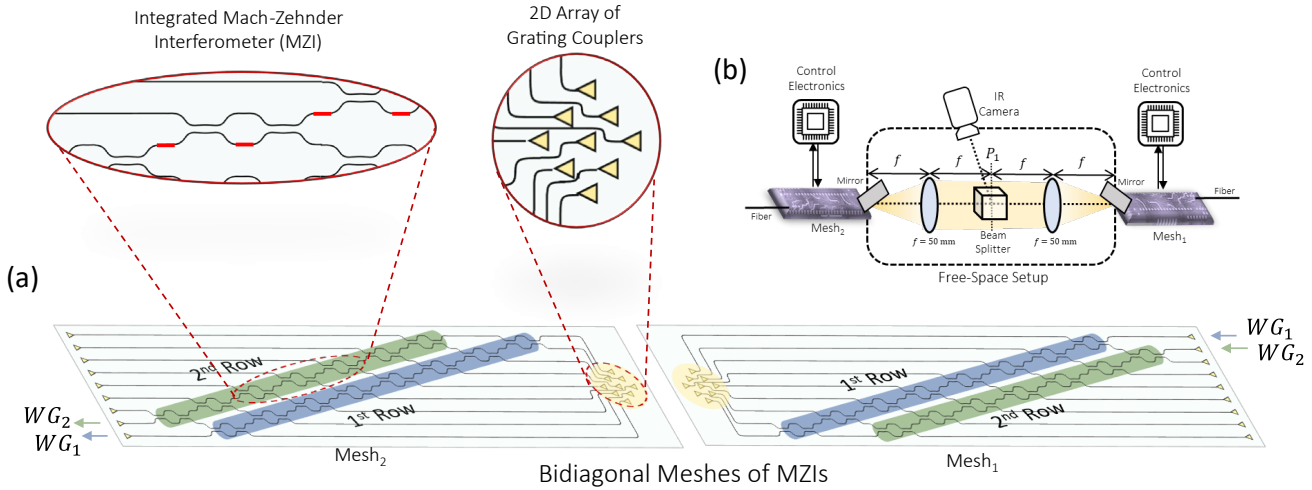
S. SeyedinNavadeh, M. Milanizadeh, F. Zanetto, G. Ferrari, M. Sampietro, M. Sorel, D. A. B. Miller, A. Melloni, and F. Morichetti, "Determining the optimal communication channels of arbitrary optical systems using integrated photonic processors," Nat. Photon. **18**, 149-155 (2024)

Establishing optimum orthogonal channels - experiment

Even after inserting a partially blocking mask in the optical path between the meshes

the system can re-establish orthogonal channels automatically

with > 30 dB rejection between the channels



S. SeyedinNavadeh, M. Milanizadeh, F. Zanetto, G. Ferrari, M. Sampietro, M. Sorel, D. A. B. Miller, A. Melloni, and F. Morichetti, "Determining the optimal communication channels of arbitrary optical systems using integrated photonic processors," Nat. Photon. **18**, 149-155 (2024)

An actual “optical computer”

This system solves this problem based on a simple algorithm of

adjusting phase shifters to maximize output powers

without explicitly measuring the optical system

and without any calculations in any external computer

The calculations are performed by the optics itself

with phase shifters adjusted based on applied voltages derived from

measuring optical powers at the outputs

This complex analog computer works stably and reliably

despite having at least 60 analog variables

It also does not require any calculations at the “data rate”, being

transparent to the modulation formats of any signals on the beams

S. SeyedinNavadeh, M. Milanizadeh, F. Zanetto, G. Ferrari, M. Sampietro, M. Sorel, D. A. B. Miller, A. Melloni, and F. Morichetti, "Determining the optimal communication channels of arbitrary optical systems using integrated photonic processors," Nat. Photon. **18**, 149-155 (2024)

A new way of looking at optics

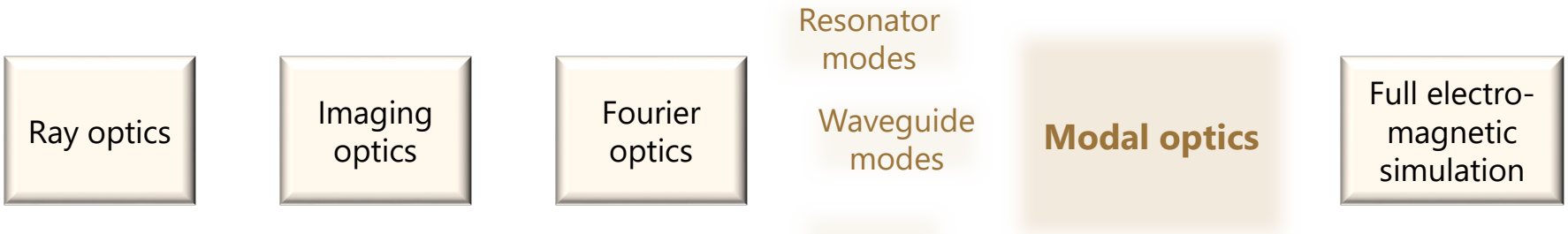
stanford.io/3NK9w8Z

A new way of looking at optics

This new way

- reproduces existing results
  - such as the limits from diffraction
- resolves paradoxes about the number of channels for communication
- clearly defines such channels
- gives us new physical laws
  - that only emerge from this view
    - e.g., new “Kirchhoff” radiation laws
    - new “modal” Einstein “A&B” law
- works to describe optical systems made with interferometer meshes

Modal optics



We want a "modal" optics

to give the "right" way to describe optical systems

The optimal sets of functions

which even have basic physical laws that apply only to them

that give the most economical way to describe systems

including the "right" number of the "right" functions

To do this properly

we need to move beyond "resonator" and "waveguide" modes

and even beyond standard "beams"

"Waves, modes, communications, and optics: a tutorial," Adv. Opt. Photon. **11**, 679-825 (2019)

A different way of thinking about modes and waves

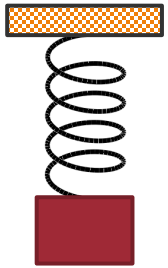
We are used to modes for

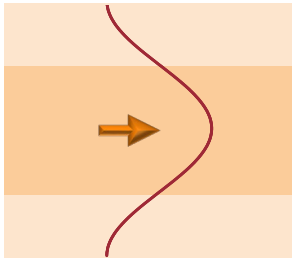
- resonators
- propagating modes in waveguides

We like “modes” because they are economical

- We can use a few mode amplitudes
  - not fields at every point

We can often “count” modes meaningfully





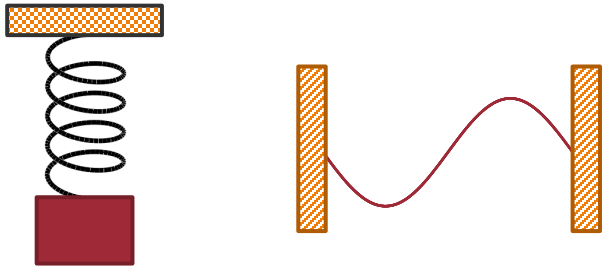
A different way of thinking about modes and waves

Modes have very useful mathematical properties, e.g.,

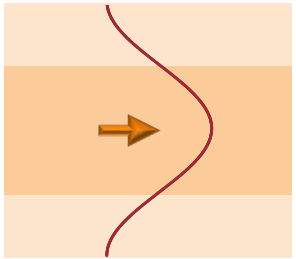
orthogonality

completeness

We can give a definition of a mode



A mode is an eigenfunction of an eigen problem describing a physical system



A different way of thinking about modes and waves

But when we look generally at

communications with waves

or scatterers, optical devices, or nanostructures

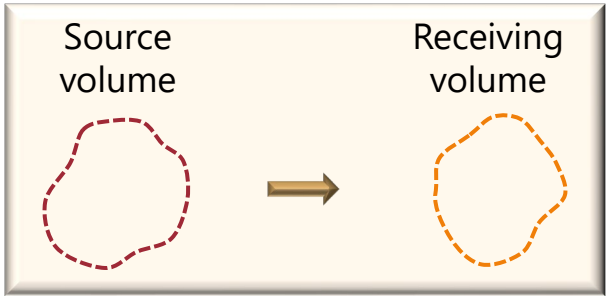
we need a different kind of "mode" that looks at

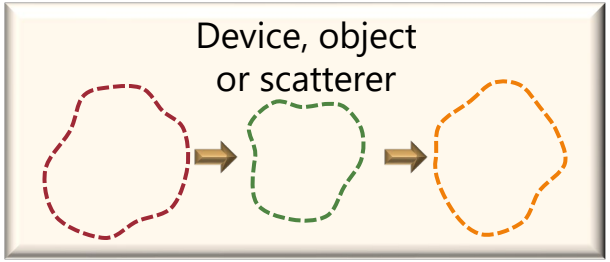
- "source" or input spaces
- and "receiving" or output spaces

They are "modes" in **two** spaces

not one space

They are **not** the "beams" between the spaces





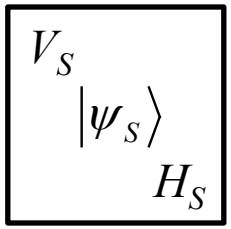
"Waves, modes, communications, and optics: a tutorial," Adv. Opt. Photon. **11**, 679-825 (2019)

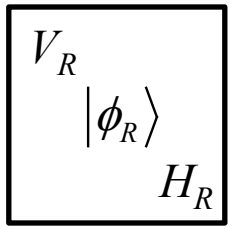
Communication modes

To set up the mathematics of this problem

we consider two spaces

Source or input volume or space





Receiving or output volume or space

A source or input volume V_S (rigorously, a Hilbert space H_S)

containing the possible source functions

written using a Dirac notation for convenience, e.g., $|\psi_S\rangle$

A receiving or output volume V_R (rigorously a Hilbert space H_R)

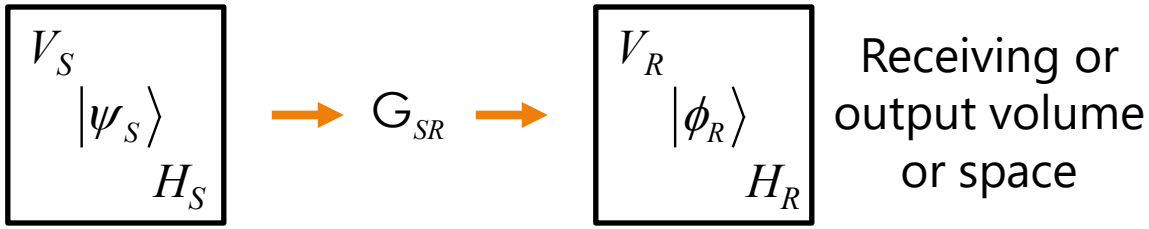
containing the possible wave functions

written using a Dirac notation for convenience, e.g., $|\phi_R\rangle$

Communication modes

The sources in the input space give waves in the receiving space through some coupling operator G_{SR}

Source or input volume or space



Receiving or output volume or space

For free space, this would be based on a free-space Green's function such as a scalar monochromatic Green's function

$$G_\omega(\mathbf{r}_R;\mathbf{r}_S) = -\frac{1}{4\pi}\frac{\exp\left(ik\left|\mathbf{r}_R-\mathbf{r}_S\right|\right)}{\left|\mathbf{r}_R-\mathbf{r}_S\right|}$$

giving the wave at point $\mathbf{r}_R$ in the receiving space from the point source at $\mathbf{r}_S$ in the source space

Choosing eigen problems

We want eigen problems to get modes

but we need *two* eigen problems because we have two different spaces

But these are *not* just the usual eigen problems of, say,

a resonator in each volume

There is, however, a key mathematical trick we can use instead

Finding mode pairs

With the coupling operator G_{SR} between the spaces

- for the source space, solve the eigen problem for the operator $G_{SR}^{\dagger} G_{SR}$

$$G_{SR}^{\dagger} G_{SR} \left| \psi_{Sj} \right\rangle = \left| s_j \right|^2 \left| \psi_{Sj} \right\rangle$$

which gives an orthogonal set of source functions (or mathematical vectors) $\left| \psi_{Sj} \right\rangle$ in H_S

Note: $G_{SR}^{\dagger}$ is the Hermitian adjoint of G_{SR}. As a matrix, it would be the complex conjugate of the transpose of the matrix. As a Green's function, it is the complex conjugate, with the roles of "source" and "receiver" points interchanged.

Finding mode pairs

With the coupling operator G_{SR} between the spaces

- for the receiving space, solve the eigen problem for the operator $G_{SR}G_{SR}^{\dagger}$

$$G_{SR}G_{SR}^{\dagger}\left|\phi_{Rj}\right\rangle = \left|s_j\right|^2\left|\phi_{Rj}\right\rangle$$

which gives an orthogonal set of wave functions (or mathematical vectors) $\left|\phi_{Rj}\right\rangle$ in H_R

(These problems have the same, positive eigenvalues $\left|s_j\right|^2$)

Finding mode pairs

When we have done this, we find that

$$G_{SR}\left|\psi_{Sj}\right\rangle = s_j\left|\phi_{Rj}\right\rangle$$

So, the source eigenfunction $\left|\psi_{Sj}\right\rangle$

generates the corresponding eigenfunction

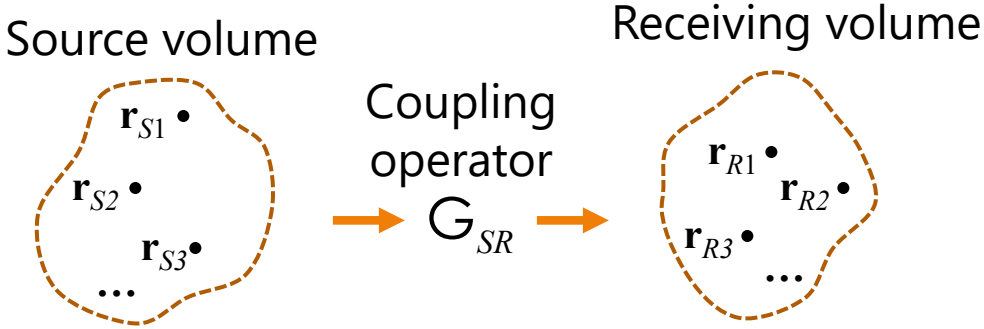
the wave $\left|\phi_{Rj}\right\rangle$ in the receiving space

with the coupling amplitude s_j

We have established the communication mode **pairs** of functions

This process is the **singular-value decomposition** of the coupling operator G_{SR}

Constructing examples with point sources and receivers



"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679 (2019)

We can see how this works first for a finite number of point sources and receivers

e.g., "loudspeakers" at positions $\mathbf{r}_{S1}$, $\mathbf{r}_{S2}$, $\mathbf{r}_{S3}$, etc., in the source volume and "microphones" at positions $\mathbf{r}_{R1}$, $\mathbf{r}_{R2}$, $\mathbf{r}_{R3}$, etc., in the receiving volume

Using the Green's function, we can construct the resulting matrix to represent G_{SR}

3 sources and receivers

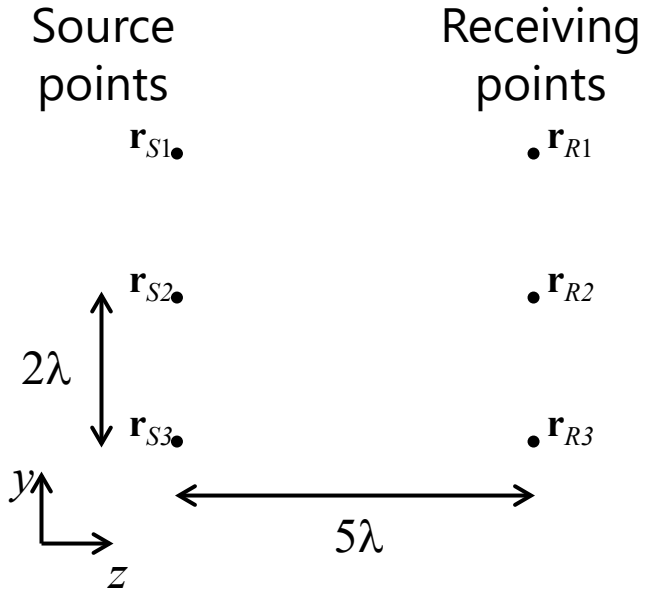
For these source and receiving points

using the Green's function

$$G_\omega(\mathbf{r}_R;\mathbf{r}_S) = -\frac{1}{4\pi}\frac{\exp\left(ik\left|\mathbf{r}_R-\mathbf{r}_S\right|\right)}{\left|\mathbf{r}_R-\mathbf{r}_S\right|}$$

gives the matrix (for unit wavelength λ)

$$G_{SR} \cong \frac{-1}{62.83}\begin{bmatrix} 1 & -0.7+0.6i & -0.64+0.45i \\ -0.7+0.6i & 1 & -0.7+0.6i \\ -0.64+0.45i & -0.7+0.6i & 1 \end{bmatrix}$$

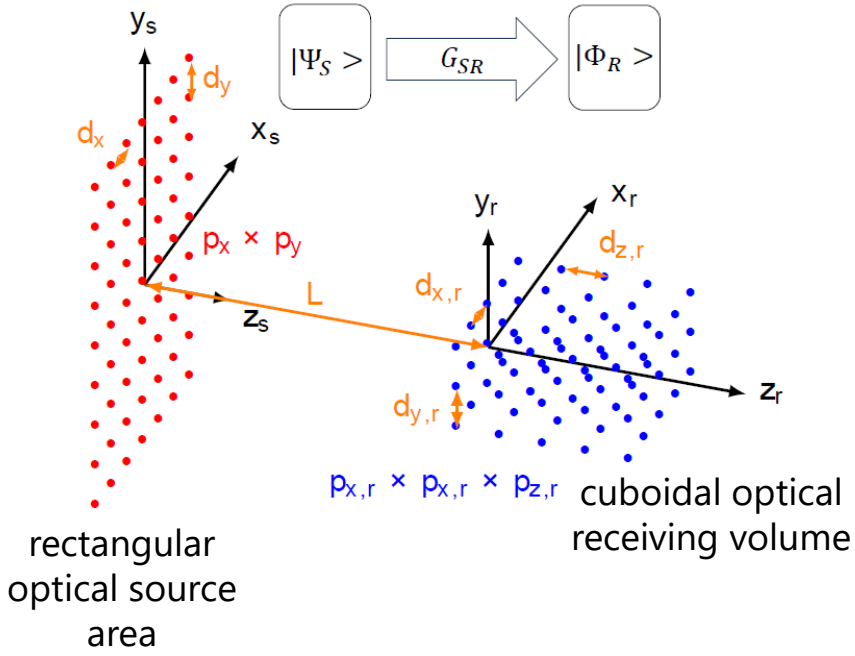


"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679 (2019)

Generating a volume field

We choose dense sets of points in
- a planar rectangular optical source "volume"
- a cuboidal optical receiving volume

We can similarly construct the matrix representing G_{SR}
- and perform the SVD of it
  - to get the source and receiving communication mode vectors
    - and coupling strengths (singular values)



V. S. de Angelis, A. H. Dorrah, L. A. Ambrosio, D. A. B. Miller, and F. Capasso, "3D holography using communication mode optics," in *Optica Imaging Congress 2024 (3D, AOMS, COSI, ISA, pcAOP)*, Technical Digest Series (Optica Publishing Group, 2024), paper DF4H.4. https://doi.org/10.1364/3D.2024.DF4H.4

Generating arbitrary waves with communication modes

Suppose we want a specific wave $|\phi_{Ro}\rangle$ in the receiving space

We expand it in the "receiving" communication modes as $|\phi_{Ro}\rangle = \sum_j a_j |\phi_{Rj}\rangle$

where $a_j = \langle \phi_{Rj} | \phi_{Ro} \rangle$ is the "inner product" or "overlap integral"

Since $G_{SR} |\psi_{Sj}\rangle = s_j |\phi_{Rj}\rangle$

to generate any specific component $a_q |\phi_{Rq}\rangle$ for this expansion

we need an amplitude a_q / s_q of the source function $|\psi_{Sq}\rangle$

So, the required source function $|\psi_{So}\rangle$ to generate $|\phi_{Ro}\rangle$ is

$$|\psi_{So}\rangle = \sum_j \frac{a_j}{s_j} |\psi_{Sj}\rangle \equiv \sum_j \frac{1}{s_j} \langle \phi_{Rj} | \phi_{Ro} \rangle |\psi_{Sj}\rangle$$

This lets us generate any desired wave in the receiving space

even if the coupling strengths s_j are not the same for every communication mode

"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679 (2019)

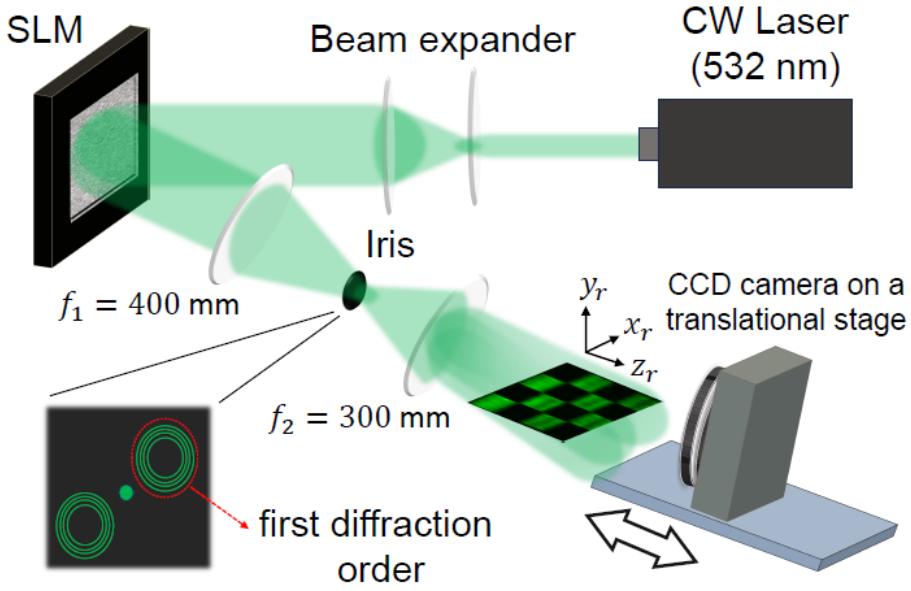
de Angelis et al., "3D holography using communication mode optics," in *Optica Imaging Congress 2024,* paper DF4H.4. https://doi.org/10.1364/3D.2024.DF4H.4

Creating a 3-D optical field

Using this approach, we can generate an arbitrary desired 3D field

calculating the necessary amplitudes for the pixel "point sources" on a spatial light modulator

to generate the field of interest in the 3D volume



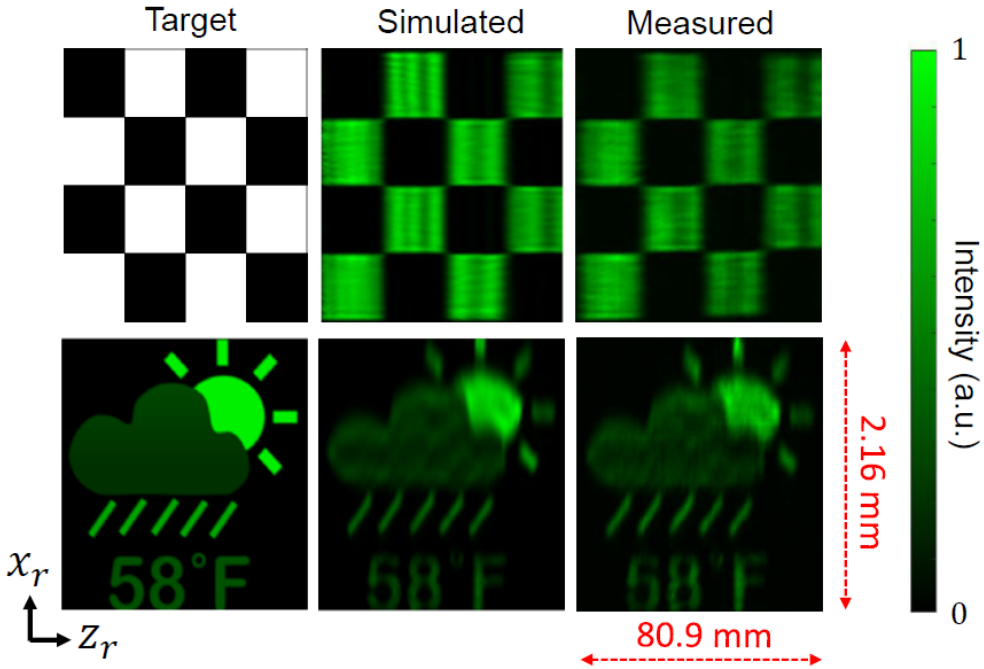
de Angelis et al., "3D holography using communication mode optics," in *Optica Imaging Congress 2024,* paper DF4H.4. https://doi.org/10.1364/3D.2024.DF4H.4

Designing and creating a 3-D optical field

Using this approach, we can generate an arbitrary desired 3D field

calculating the necessary amplitudes for the pixel "point sources" on a spatial light modulator

to generate the field of interest in the 3D volume



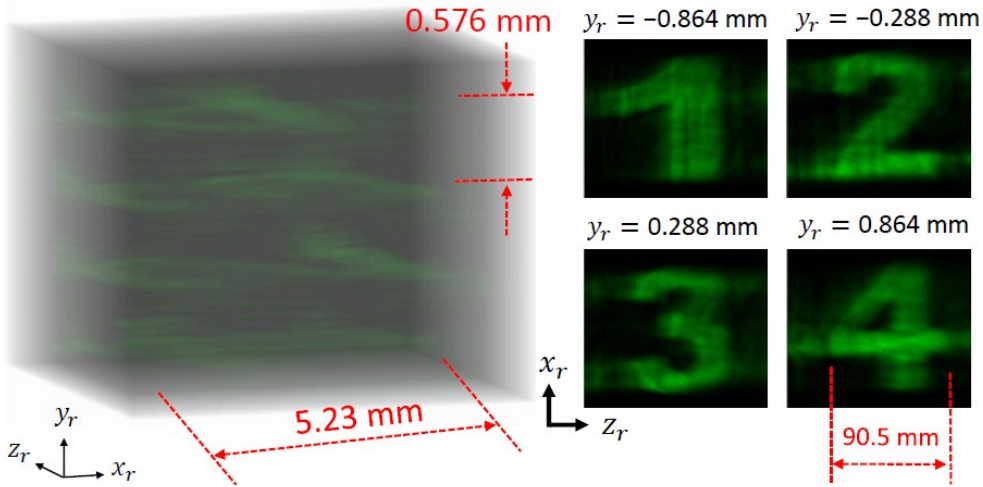
de Angelis et al., "3D holography using communication mode optics," in *Optica Imaging Congress 2024,* paper DF4H.4. https://doi.org/10.1364/3D.2024.DF4H.4

Designing and creating a 3-D optical field

Using this approach, we can generate an arbitrary desired 3D field

calculating the necessary amplitudes for the pixel “point sources” on a spatial light modulator

to generate the field of interest in the 3D volume



de Angelis et al., "3D holography using communication mode optics," in *Optica Imaging Congress 2024,* paper DF4H.4. https://doi.org/10.1364/3D.2024.DF4H.4

Waves, modes, and minimum thicknesses for optics

stanford.io/3NK9w8Z

DM, "Why optics needs thickness," Science 379, 41 (2023)

Why optics needs thickness

For metasurfaces and metastructures

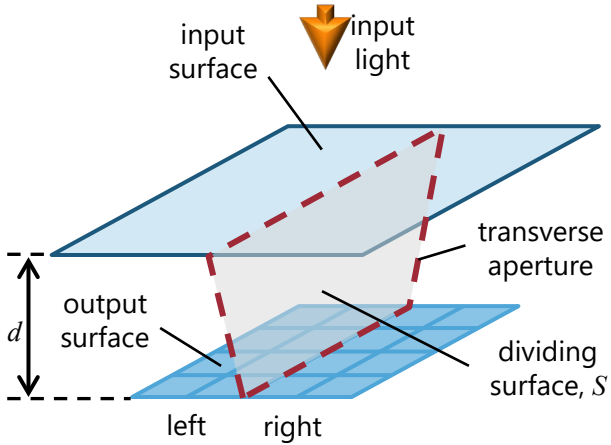
and for compact optics generally

we need to understand whether they need thickness

Can we make a given optical device in just one “layer”, for example?

Generally, no.

But why?

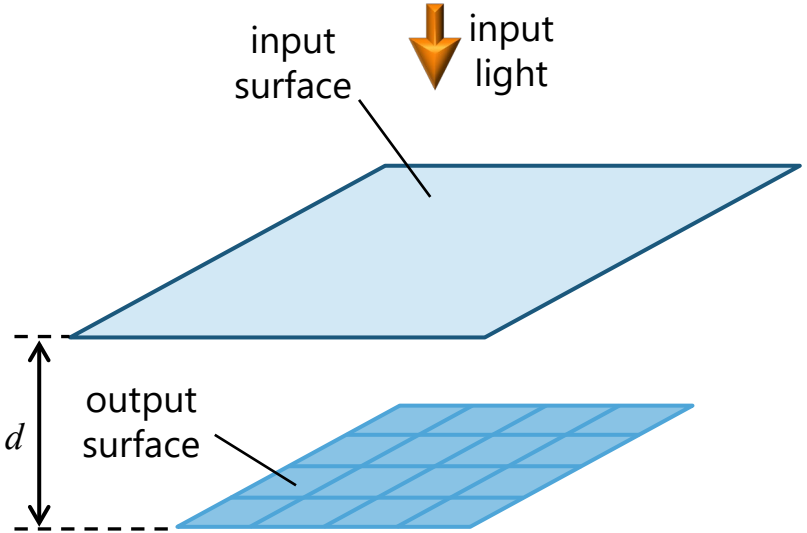


David Miller, "Why optics needs thickness," Science 379, 41 (2023)

Why optics needs thickness

Think of an optical system with
- an input surface
  - such as a lens surface or metasurface
- an output surface
  - such as an image sensor plane
- with a distance *d* between them

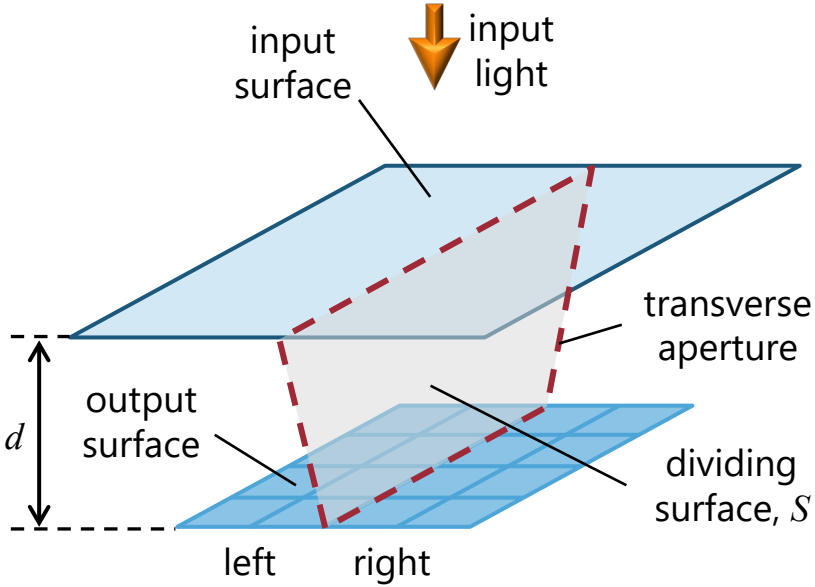
Note we are not yet specifying what is between these two surfaces
- and we will not need to do so



DM, Science 379, 41 (2023)

The key idea – channels through a transverse aperture

Now imagine we divide each surface in two parts
left and right
by passing an imaginary mathematical dividing surface *S* through them
This defines a "**transverse aperture**"
Because of what we want the system to do
some number *C* of channels must pass
from right to left (or left to right)
through this aperture
We call *C* the "**overlapping nonlocality**"
The transverse aperture must be large enough
for these channels to propagate through it
which requires minimum area and/or thickness



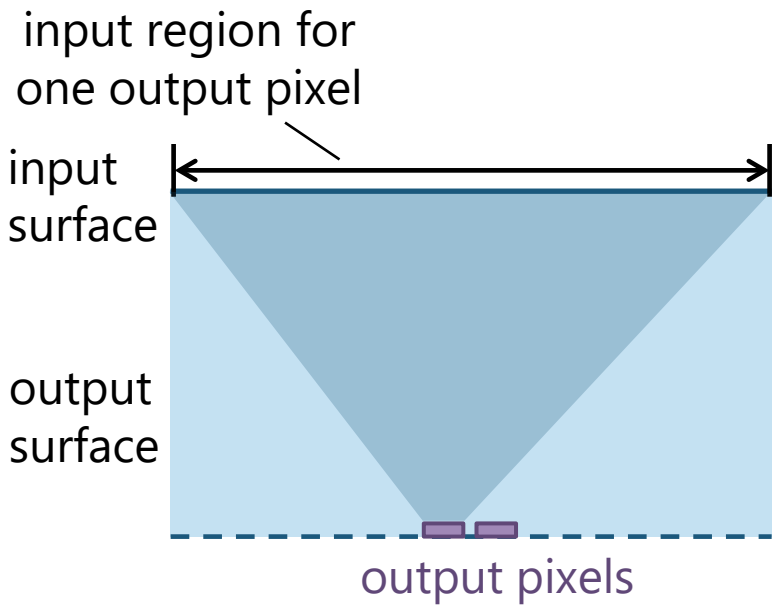
DM, Science 379, 41 (2023)

Nonlocality in optics

nonlocality

the output at one point depends on the input at many points

Imager example



Nonlocality in optics

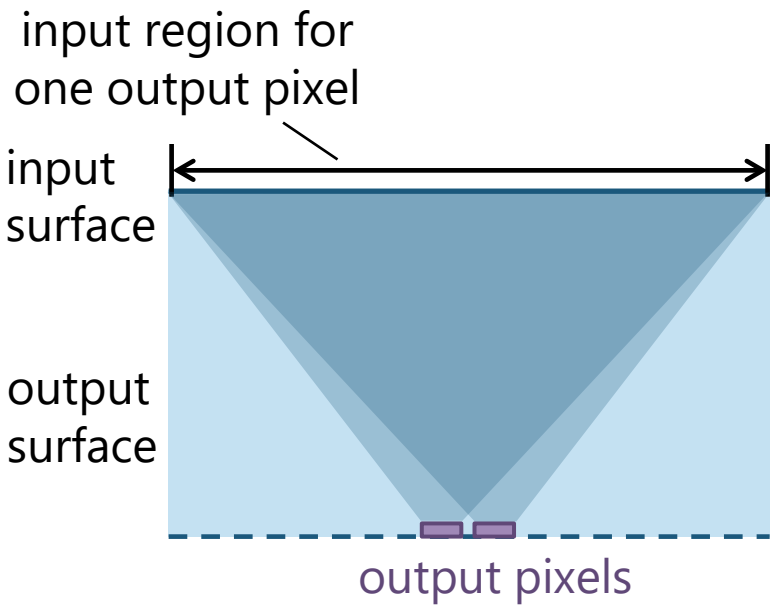
nonlocality

the output at one point depends on the input at many points

overlapping nonlocality

the input regions for different output points overlap with one another

Imager example



Nonlocality in optics

nonlocality

the output at one point depends on the input at many points

overlapping nonlocality

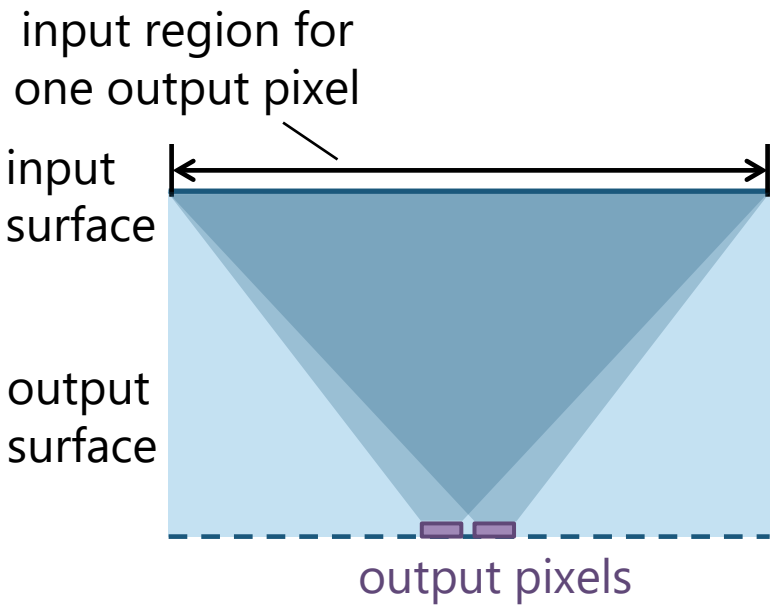
the input regions for different output points overlap with one another

overlapping nonlocality C

loosely, the number of such overlapping "channels"

For an imager, C ends up being half the number of pixels

Imager example



Nonlocality in optics

nonlocality

the output at one point depends on the input at many points

overlapping nonlocality

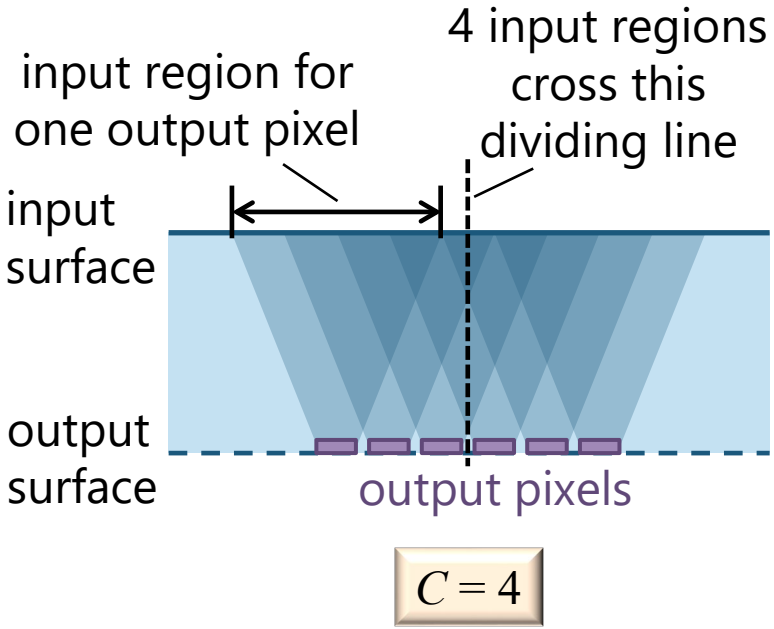
the input regions for different output points overlap with one another

overlapping nonlocality C

loosely, the number of such overlapping "channels"

For this example, C is 4

Space-invariant example
e.g., image differentiator



How big a transverse aperture for a given C?

For a 1D system with free-space wavelength λ_o and maximum refractive index n_{max}

we presume we need a thickness

$$\Delta d \geq \lambda_o / 2\alpha n_{max}$$

for each channel

where we allow for some practical factor

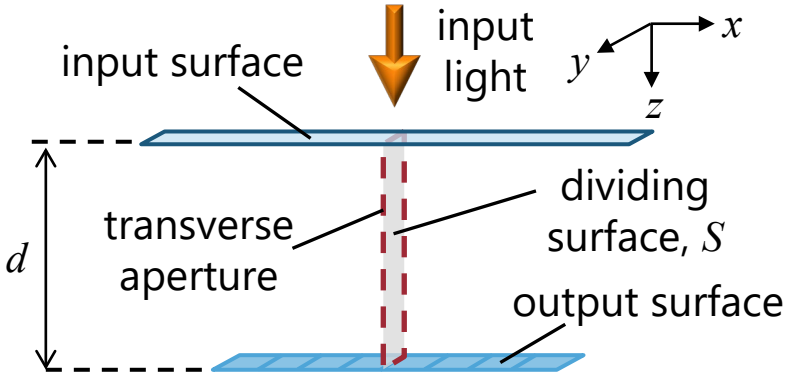
$\alpha > 1$

which comes from some practical restriction on the range of

usable angles

or usable k-space

inside the device



Thickness of a one-dimensional imager

Quite generally, for some value of C_x in a one-dimensional device

with $\Delta d \geq \lambda_o / 2\alpha n_{max}$ of thickness per channel

then $d \geq C_x \Delta d$

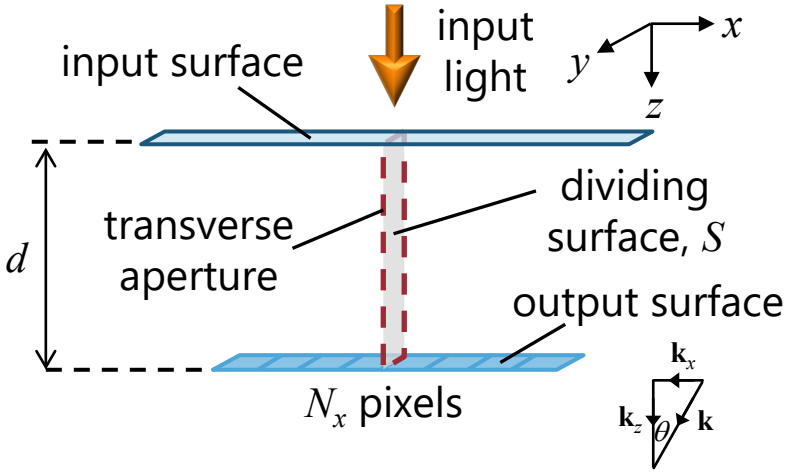
so $d \geq C_x \lambda_o / 2\alpha n_{max}$

For our one-dimensional imager with $C_x = N_x / 2$

$$d \geq N_x \lambda_o / 4\alpha n_{max}$$

For $\lambda_o = 700$ nm, $N_x = 4000$ (one "line" of a 12 MP smartphone camera), $n_{max} = 1.5$ and no "rays" past 45° angle, $d \geq 1.6$ mm

Note: this will also be the limit for a 2D imager based on conventional lenses



See V. Blahnik, O. Schindelbeck, Advanced Optical Technologies 10, 145 (2021) for general discussion of modern smartphone cameras

Why optics needs thickness

For a conventional cell phone camera

in practice, even if we took all the thickness out of the lenses themselves

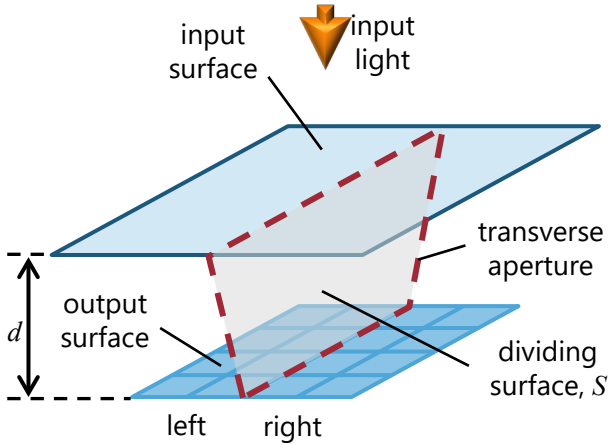
the camera would still need to be ~ 1.6 mm thick

The formal way to analyze this problem is to perform

the singular value decomposition of the coupling between

the left side of the input and the right side of the output

allowing us to count the number of modes we need



David Miller, "Why optics needs thickness," Science 379, 41 (2023)

A pixelated differentiator

Consider a 5th order finite difference derivative kernel

formed from a

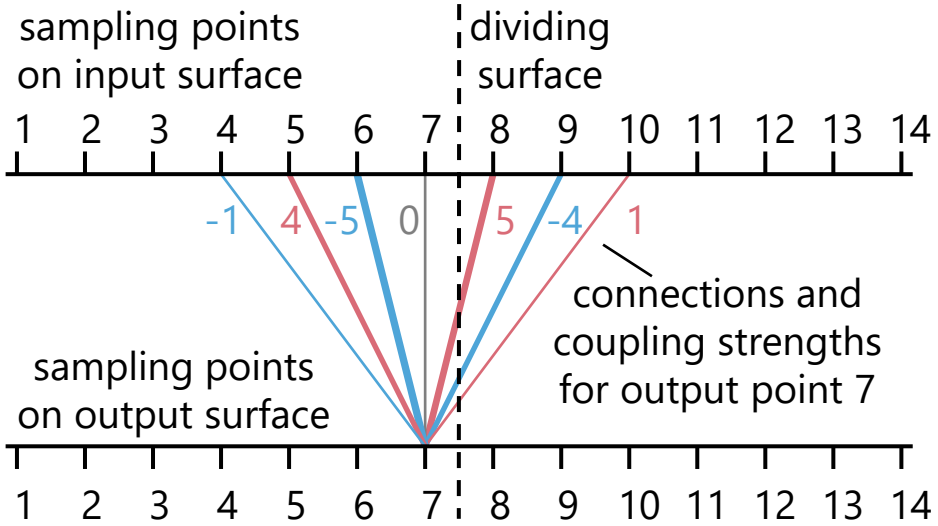
-1, 4, -5, 0, 5, -4, 1

weighting of adjacent input points

In this case, we can set up a matrix D

which gives all the connection strengths between inputs and outputs

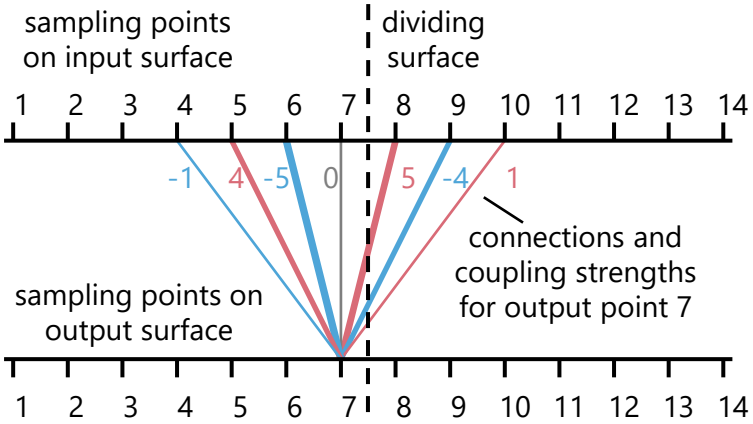
for the full “space-invariant” kernel

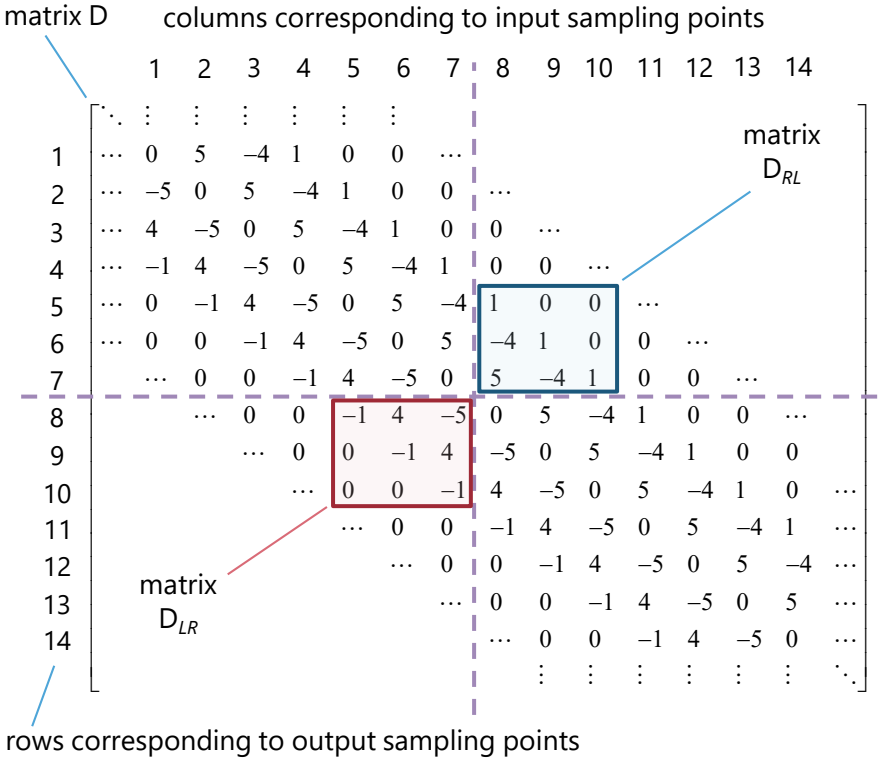


A pixelated differentiator

We can construct the full matrix D of the full “space-invariant” kernel

arbitrarily choosing one vertical position for the dividing surface between pixels 7 and 8

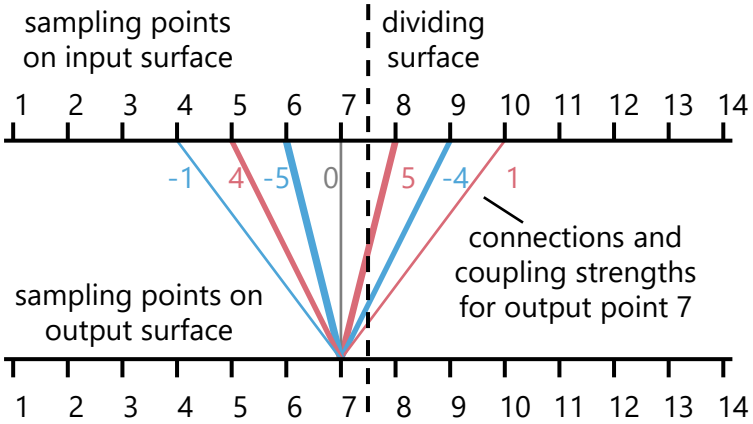


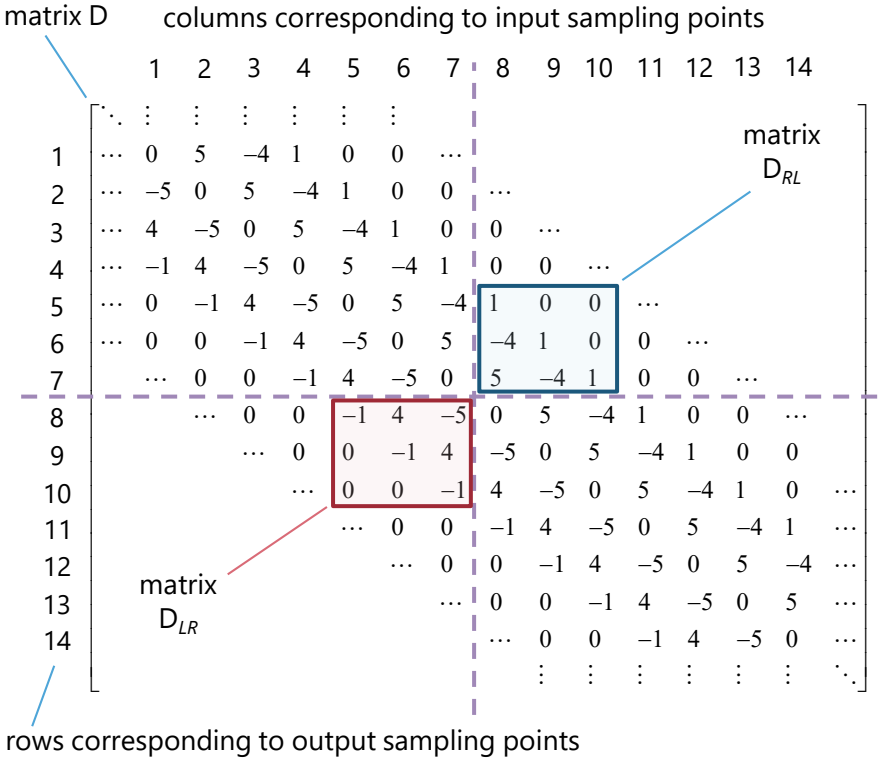


A pixelated differentiator

Sub-matrix D_{RL} gives all the connections
from the right inputs to the left outputs

Sub-matrix D_{LR} gives all the connections
from the left inputs to the right outputs





Singular-value decomposition approach

We can count directly as before

deducing $C = 6$

But with these matrices

we can take another formal approach -

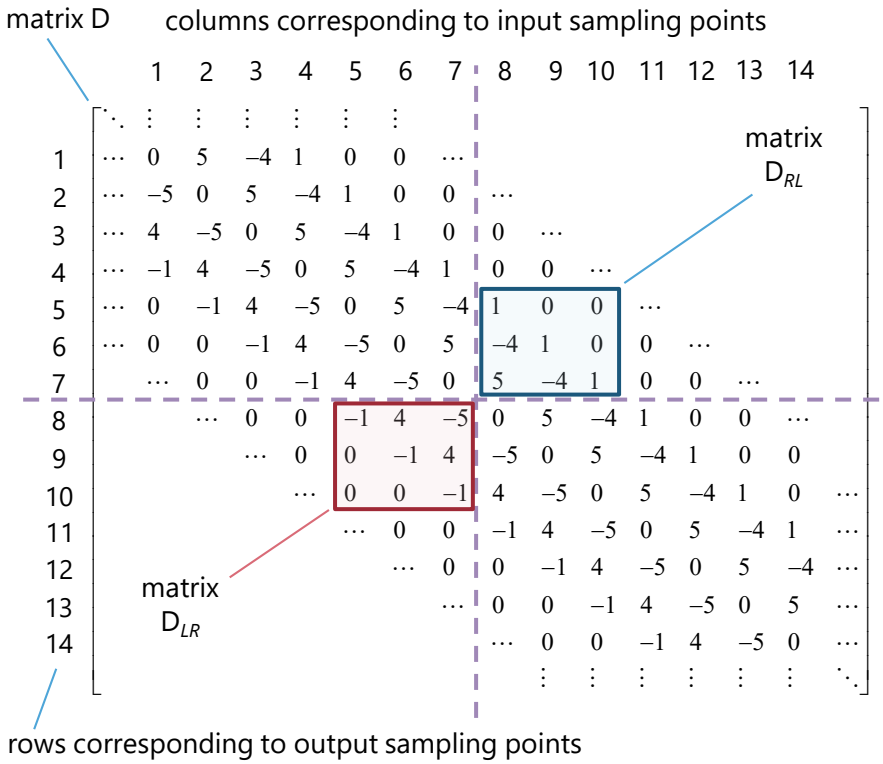
singular value decomposition (SVD) of the matrices D_{RL} and D_{LR}

which gives C_{RL} and C_{LR} as the numbers of singular values of these matrices

Though we don't need this approach here

we can use this approach for other problems where counting is not so clear

See "Waves, modes, communications, and optics: a tutorial," Adv. Opt. Photon. **11**, 679 (2019) for the SVD approach to optics



Example – metastructure for smoothed derivative

Wang et al. designed a "thick" 2D photonic crystal to perform a smoothed ("Gaussian") derivative with kernel

$$D(u;x) = \frac{(x-u)}{\beta}\exp\left(-\frac{(x-u)^2}{\beta^2\Delta_t^2}\right)$$

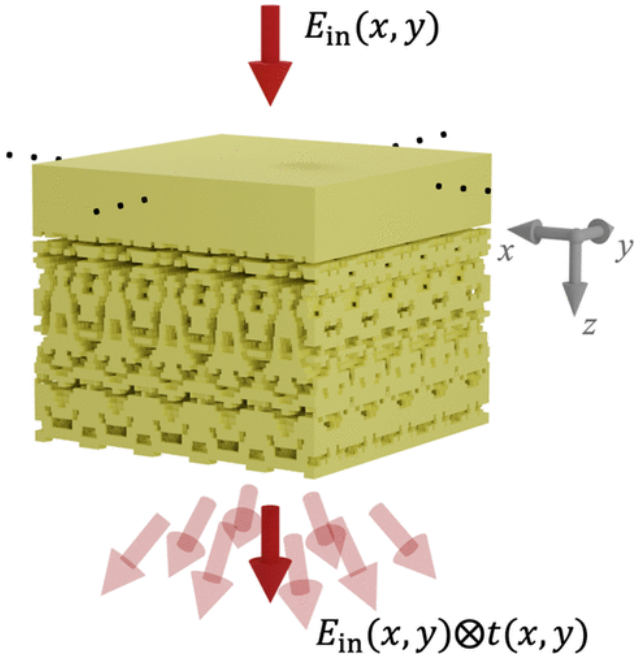
The "divided" kernel has ~ 6 significant singular values

so we should need ~ 6 physical channels through the "transverse aperture"

The thickness of the actual designed structure is ~ 6 wavelengths thick

so more than thick enough at half a wavelength thickness per channel

obeying the proposed (1D) limit here



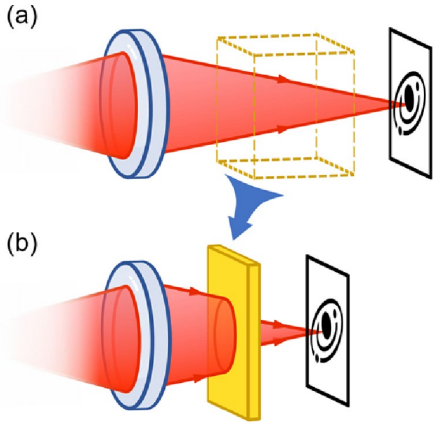
H. Wang, W. Jin, C. Guo, N. Zhao, S. P. Rodrigues, S. Fan, ACS Photonics 9, 1358–1365 (2022)

Why optics needs thickness

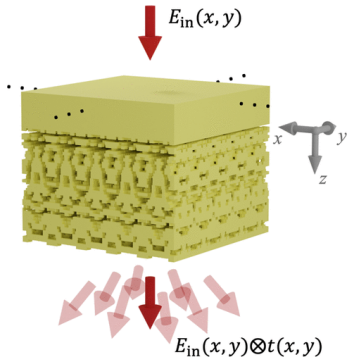
This explains the necessary thickness of, e.g.,

- smart phone cameras, which are within a factor of 3 of this limit
- "space plates" intended to make imagers thinner
- metasurface/metastructure devices as the "kernel" becomes more nonlocal

  e.g., as in image differentiation



Guo et al., Optica 7, 1133 (2020)



H. Wang et al., ACS Photonics 9, 1358 (2022)

David Miller, "Why optics needs thickness," Science 379, 41 (2023)

How many waves can get out of a volume?

stanford.io/3NK9w8Z

David Miller, *Stanford University*
Zeyu Kuang, Owen Miller, *Yale University*

How many waves can get out of a volume?

Suppose we have some arbitrary volume

which could contain

some optical source

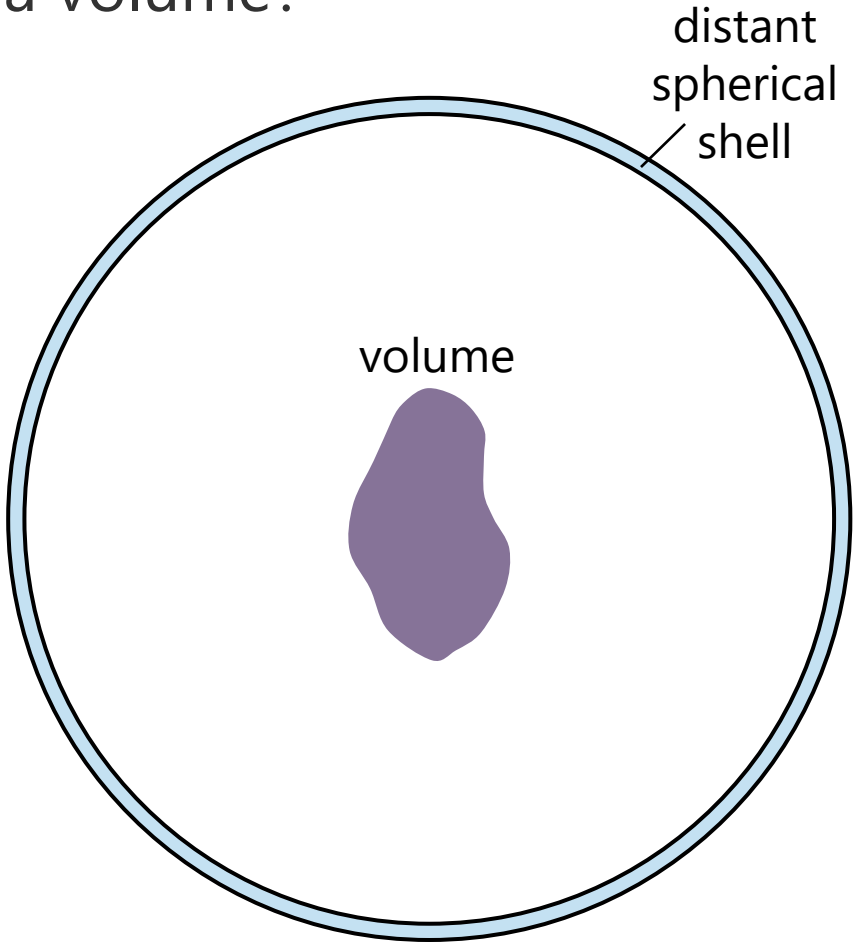
or some set of antenna elements

Can we deduce just how many waves or channels

can effectively get out of it

propagating into the far field

e.g., to a distant spherical shell?



The rigorous approach to channels between volumes

We return to the the singular-value decomposition of the coupling operator G_{SR}

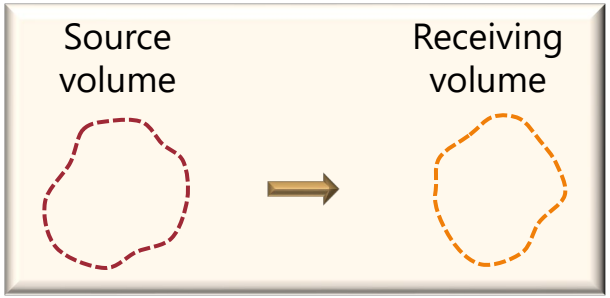
giving orthogonal source functions $|\psi_{Sj}\rangle$

that couple, one by one, to orthogonal received waves $|\phi_{Rj}\rangle$

with some coupling strength s_j

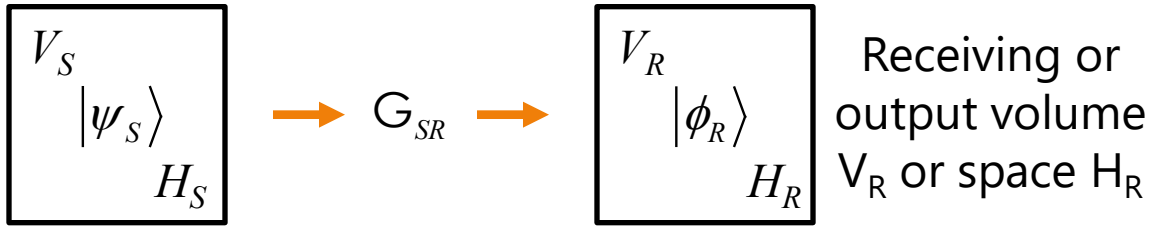
These pairs of functions $|\psi_{Sj}\rangle$ and $|\phi_{Rj}\rangle$

are the "communication modes"



"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679-825 (2019)

Source or input volume V_S or space H_S

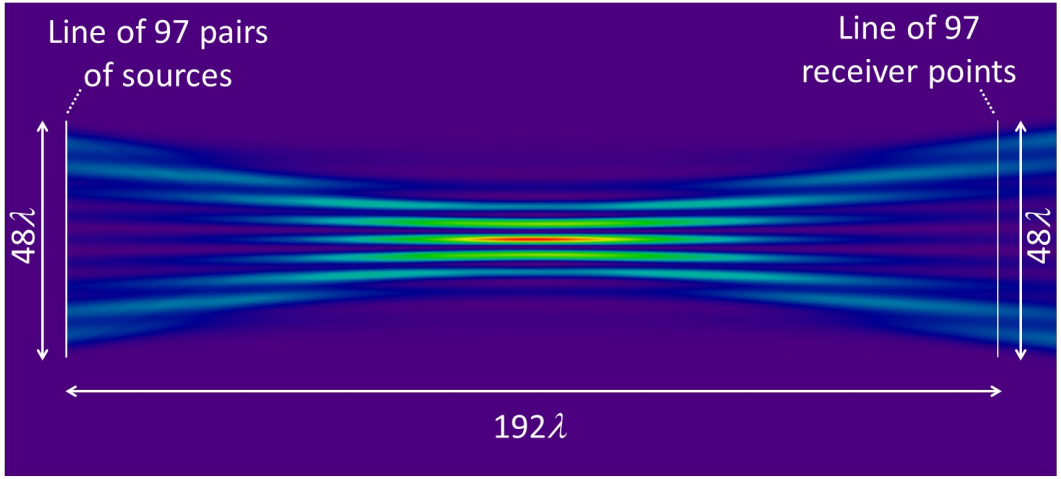


Receiving or output volume V_R or space H_R

Communication modes for a large paraxial example

Now we consider a large line of sources and a line of receiver points

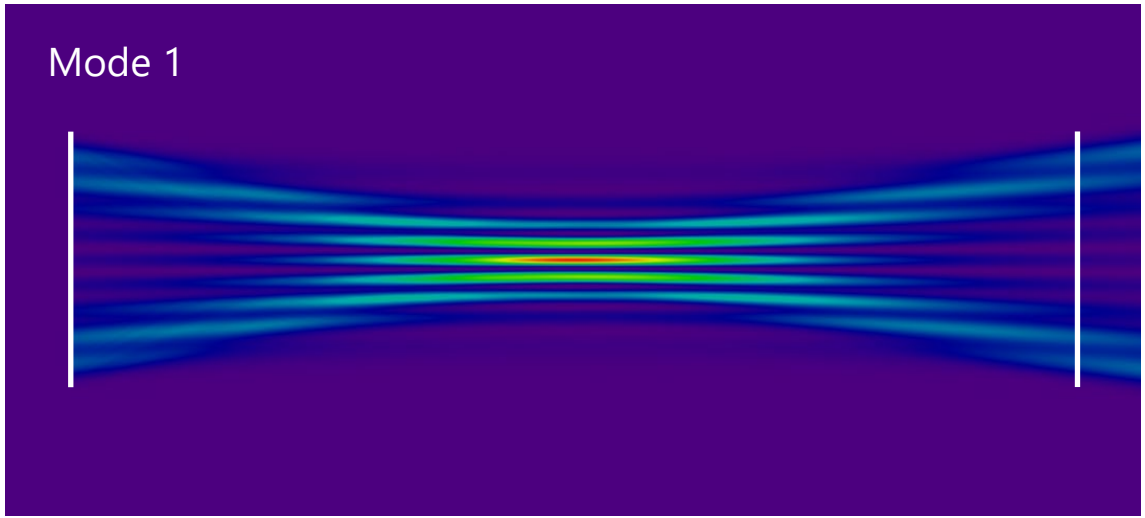
with an approximately “paraxial” set of dimensions

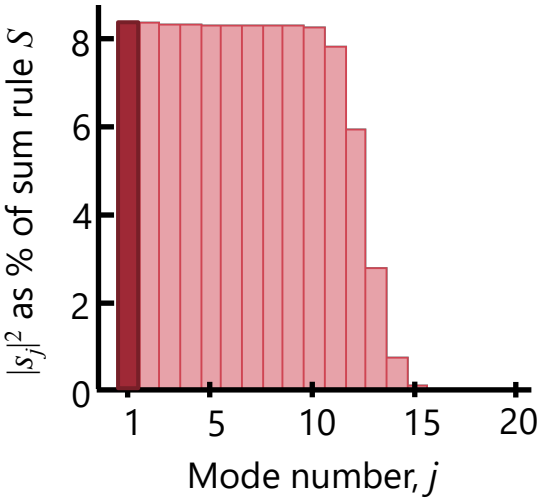


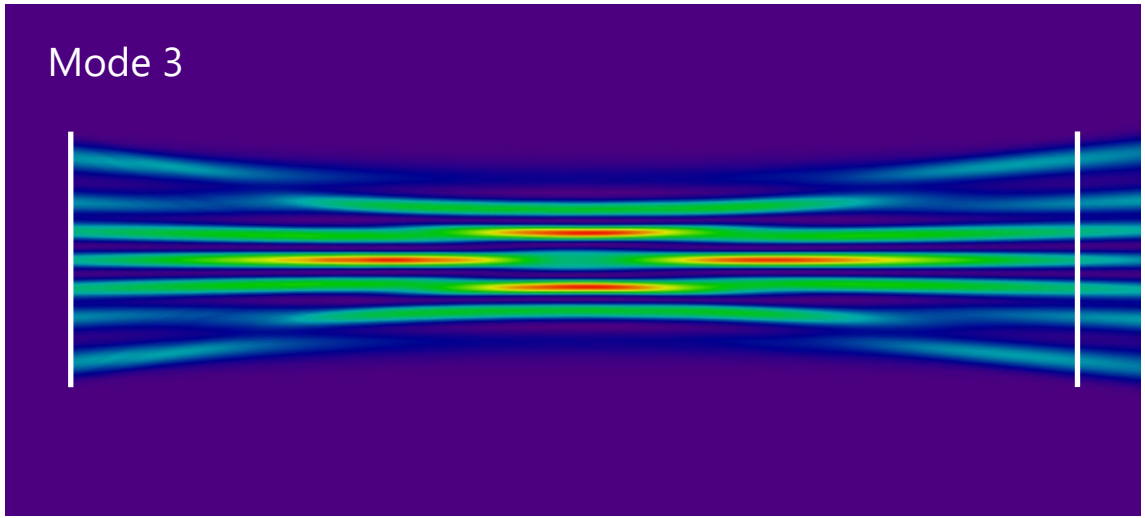
and we establish the communication modes between them

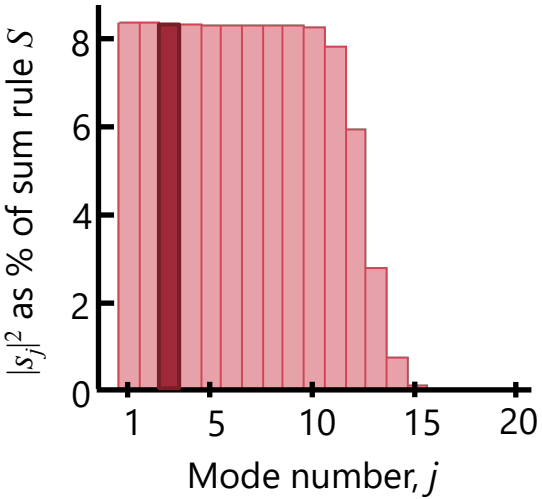
The picture shows the cross-section of the intensity in the plane

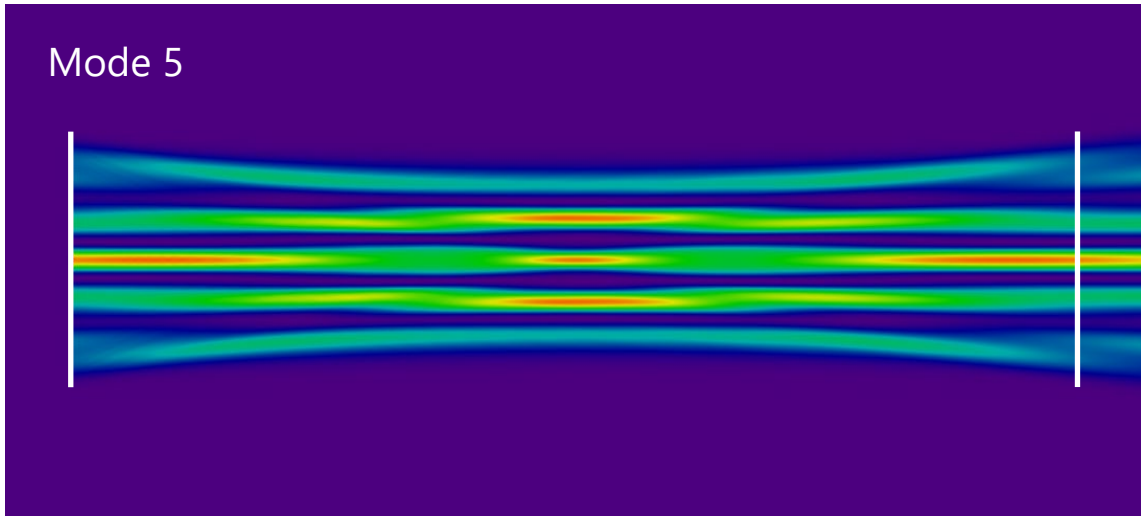
here for the most strongly coupled mode

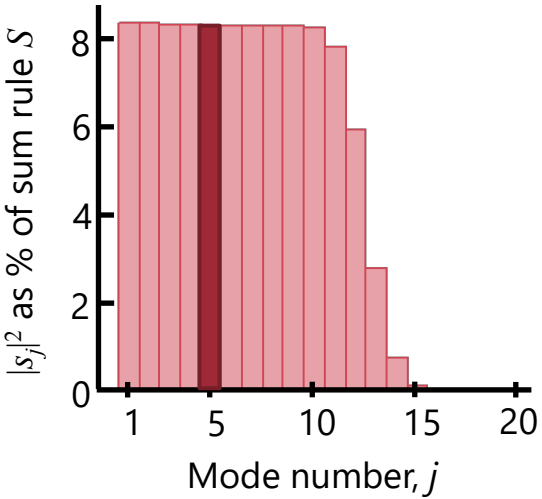

Mode 1

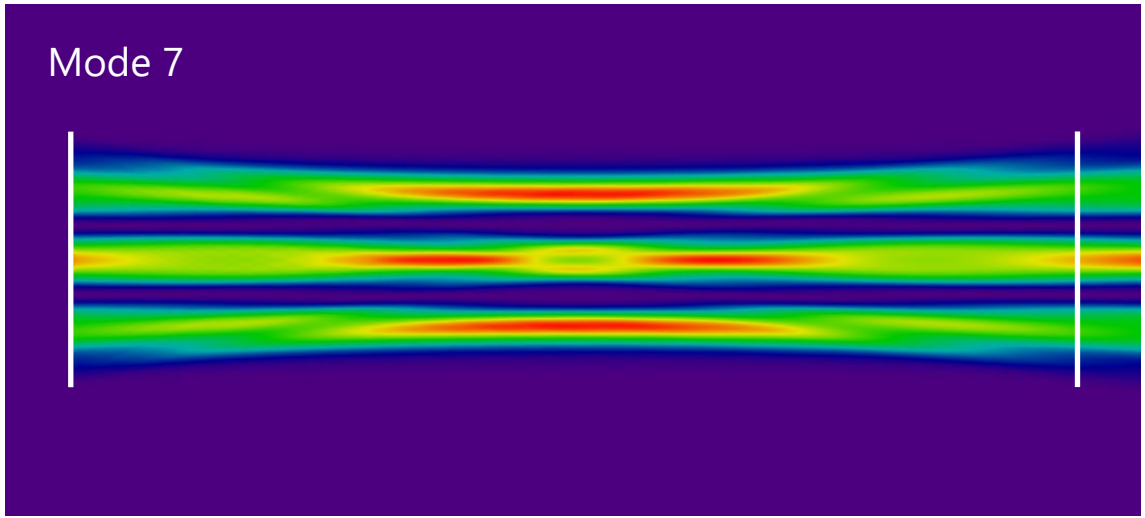

$|s_j|^2$ as % of sum rule S
0
2
4
6
8
1
5
10
15
20
Mode number, j

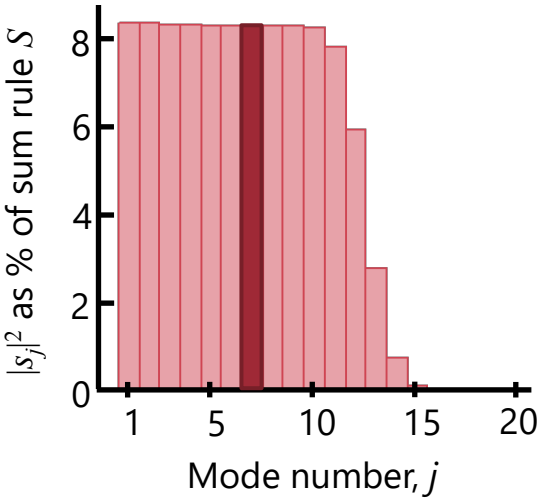

Mode 3

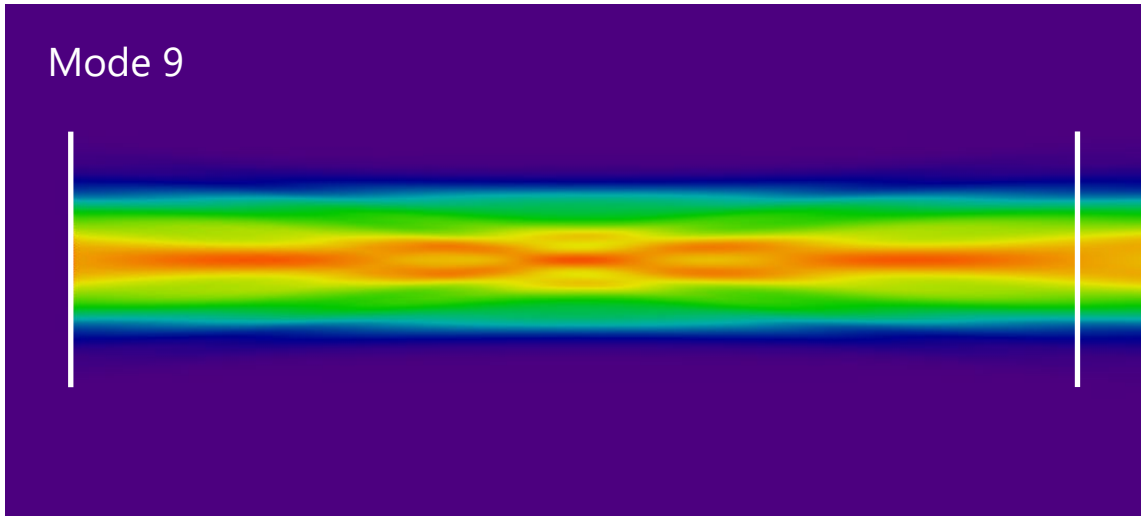

$|s_j|^2$ as % of sum rule $\mathcal{S}$
0
2
4
6
8
1
5
10
15
20
Mode number, j

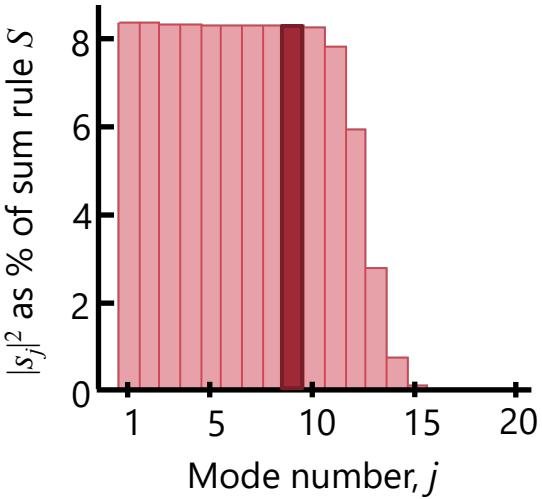

Mode 5

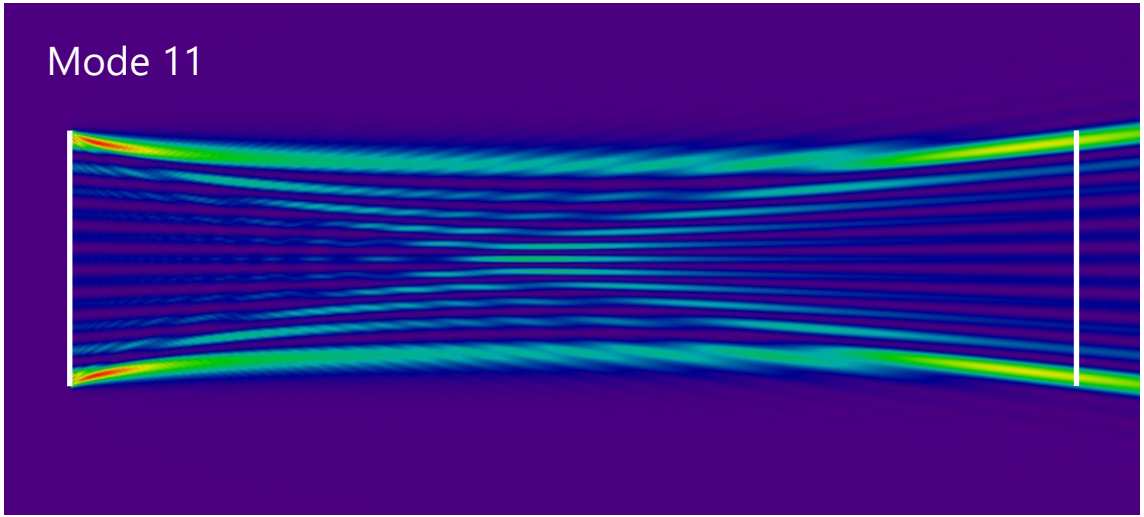

$|s_j|^2$ as % of sum rule $\mathcal{S}$
0
2
4
6
8
1
5
10
15
20
Mode number, j

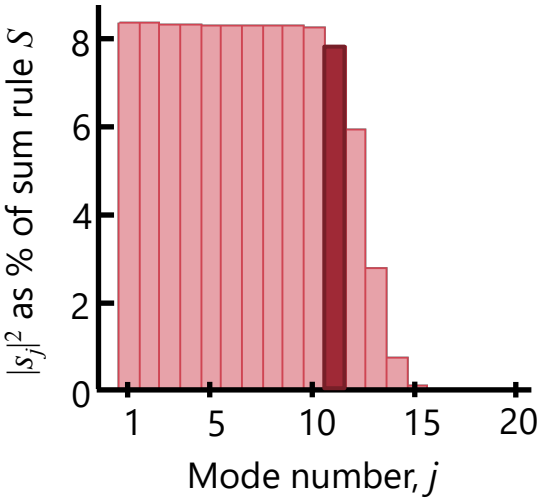

Mode 7

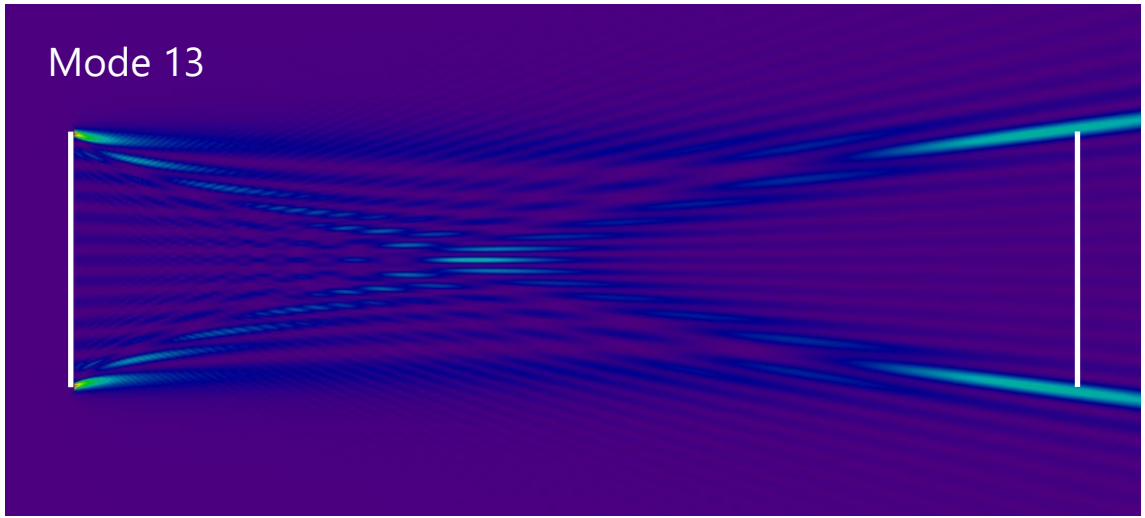

$|s_j|^2$ as % of sum rule S
0
2
4
6
8
1
5
10
15
20
Mode number, j

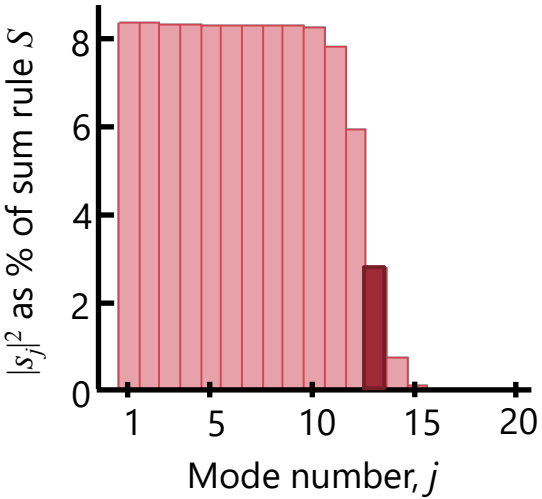

Mode 9

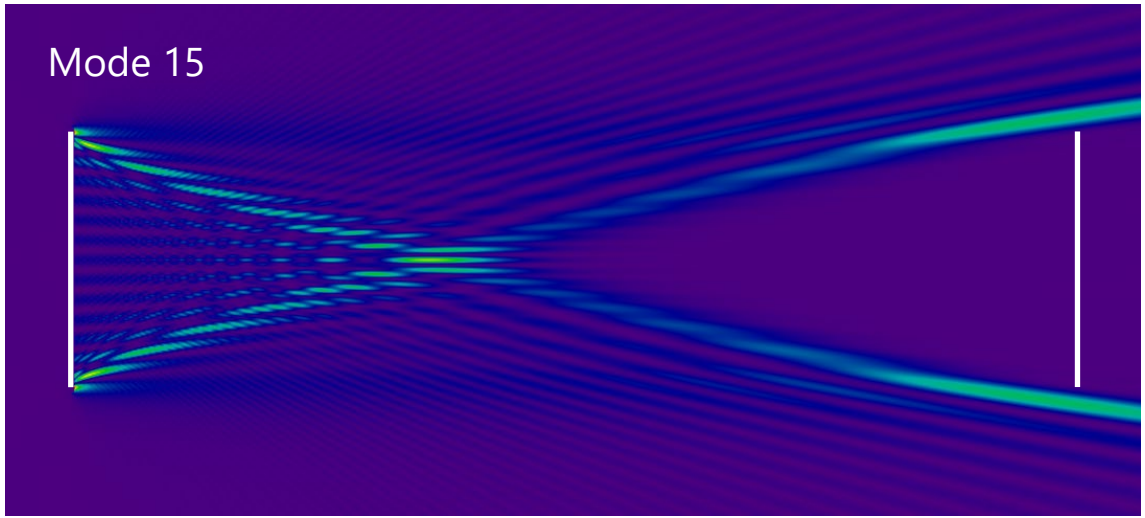

$|s_j|^2$ as % of sum rule $\mathcal{S}$
0
2
4
6
8
1
5
10
15
20
Mode number, j

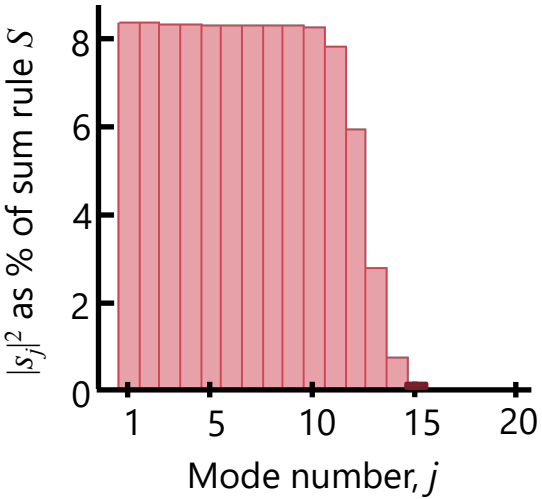

Mode 11

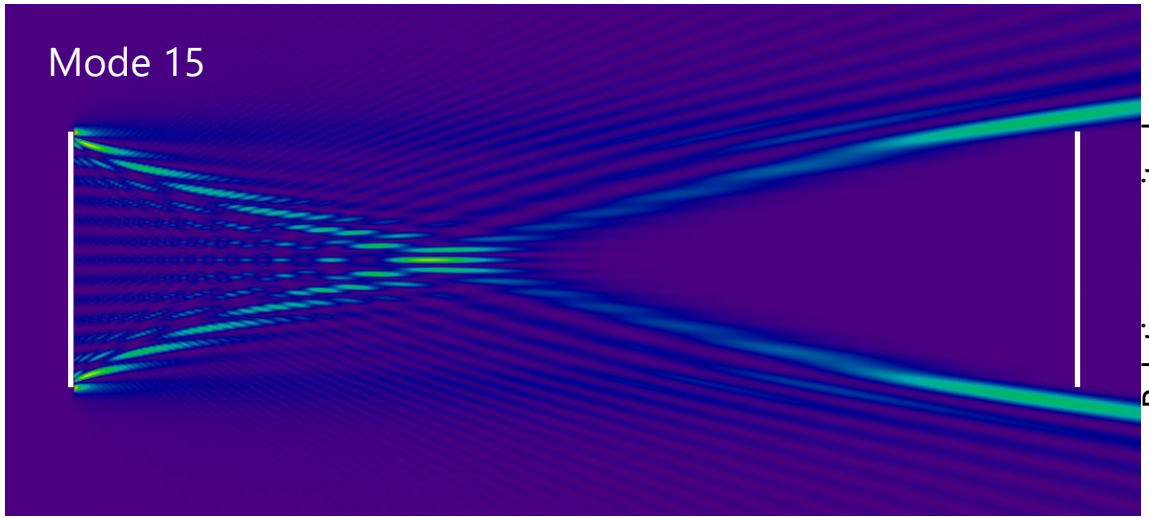

$|s_j|^2$ as % of sum rule $\mathcal{S}$
0
2
4
6
8
1
5
10
15
20
Mode number, j

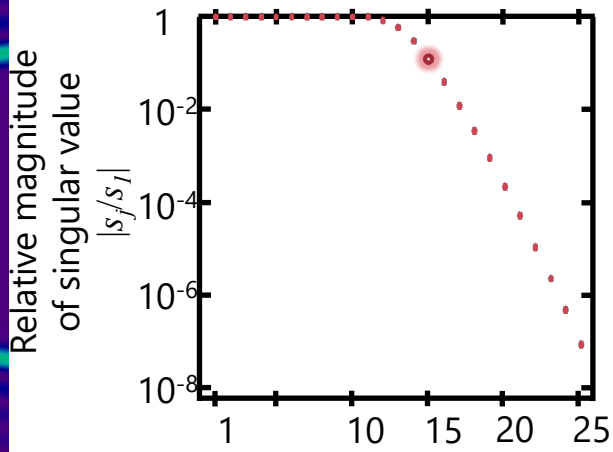

Mode 13

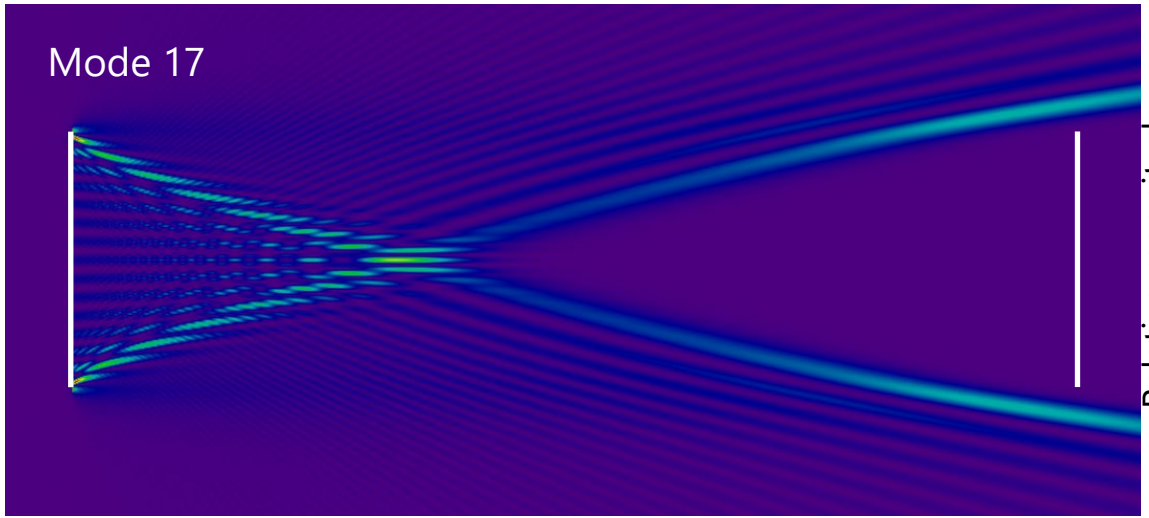

$|s_j|^2$ as % of sum rule S
0
2
4
6
8
1
5
10
15
20
Mode number, j

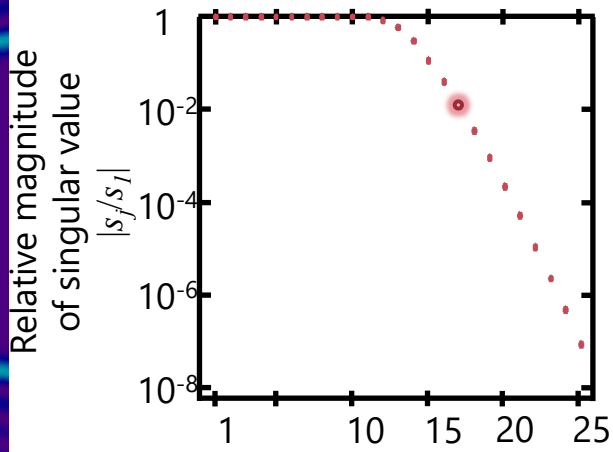

Mode 15

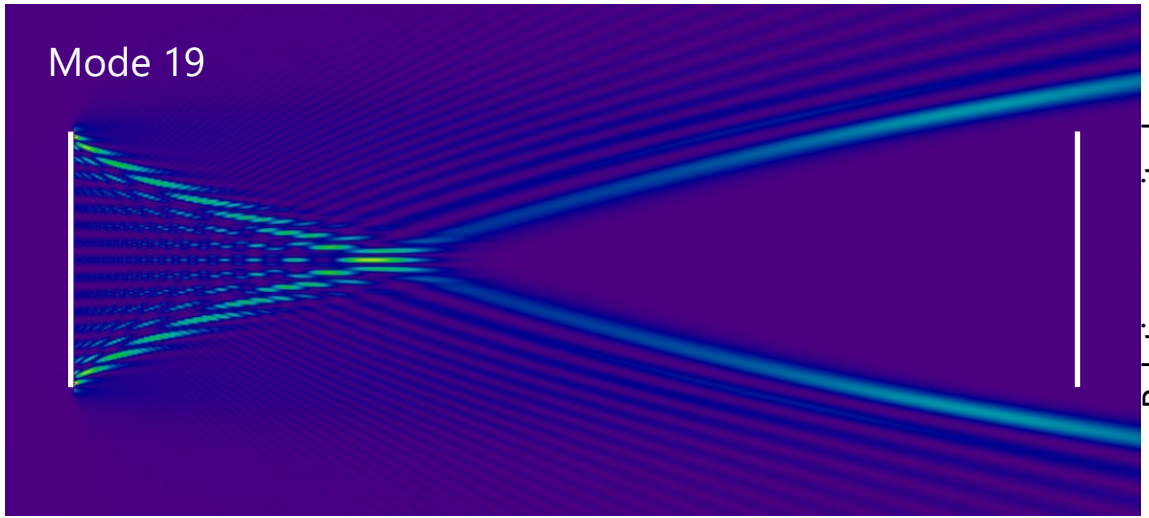

$|s_j|^2$ as % of sum rule $\mathcal{S}$
0
2
4
6
8
1
5
10
15
20
Mode number, j

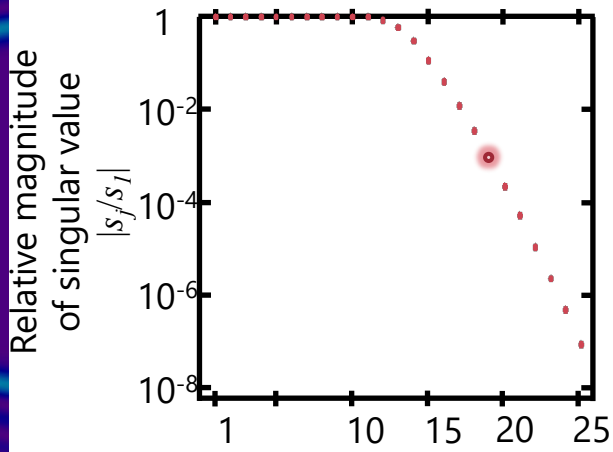

Mode 15


Relative magnitude of singular value
$|s_f/s_1|$
1
10^{-2}
10^{-4}
10^{-6}
10^{-8}
1
10
15
20
25

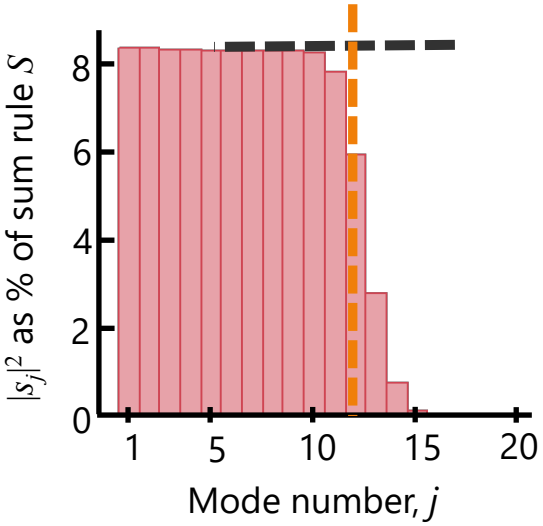

Mode 17

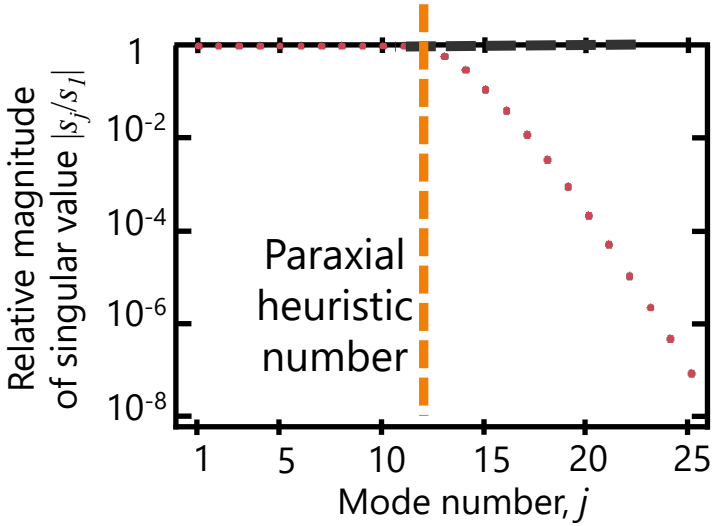

Relative magnitude of singular value
$|s_f/s_1|$
1
10^{-2}
10^{-4}
10^{-6}
10^{-8}
1
10
15
20
25


Mode 19


Relative magnitude of singular value
$|s_f/s_1|$
1
10^{-2}
10^{-4}
10^{-6}
10^{-8}
1
10
15
20
25

Paraxial heuristic number and paraxial degeneracy





"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679 (2019)

Paraxial heuristic number

$N_H \sim W_S W_R / \lambda L$

for source and receiver widths W_S, W_R

separation L

wavelength λ

Once we pass the number we expect from conventional "diffraction limits"

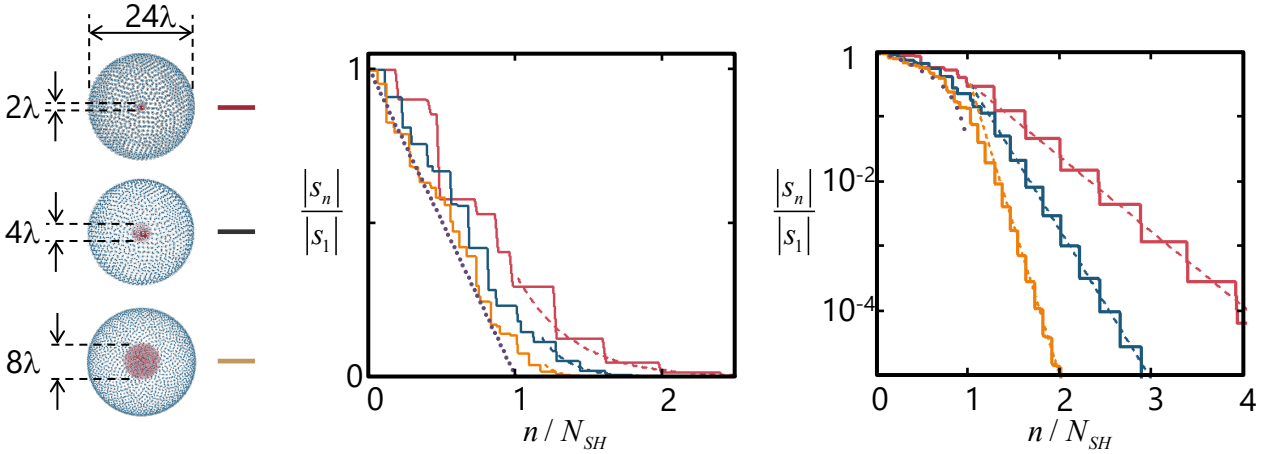
coupling strengths for further communication modes

fall off drastically and somewhat exponentially

We might think this is because the waves "miss" the receiving space

but that is not the general explanation

3D examples – concentric spherical shells



"Waves, modes, communications and optics," Adv. Opt. Photon. 11, 679 (2019)

Z. Kuang, D. A. B. Miller, and O. D. Miller, "Bounds on the Coupling Strengths of Communication Channels and Their Information Capacities," https://doi.org/10.48550/arXiv.2205.05150

Concentric spherical shell source and receiver spaces

are not easily analyzed by conventional "diffraction limit" theories

and do not show "paraxial degeneracy"

and the waves from the source space cannot "miss" the receiving space

but we still get some characteristic number of well-coupled communication modes

and a quasi-exponential fall-off of coupling beyond that

Why the abrupt fall-off past some number

Why do we *always* see

regardless of the shape of the source and receiving volumes or surfaces

some number of “well coupled” channels

followed by an abrupt, quasi-exponential fall-off past this number

and just what gives this number?

We might argue this is just “diffraction”

though that does not explain the concentric spheres case

where the waves cannot “miss” the receiving volume

Is there some underlying piece of physics we are missing?

Tunneling escape of waves

stanford.io/3NK9w8Z

David Miller, *Stanford University*
Zeyu Kuang, Owen Miller, *Yale University*

Waves from arbitrary volumes

How can we count the maximum number of well coupled waves (at a given frequency)

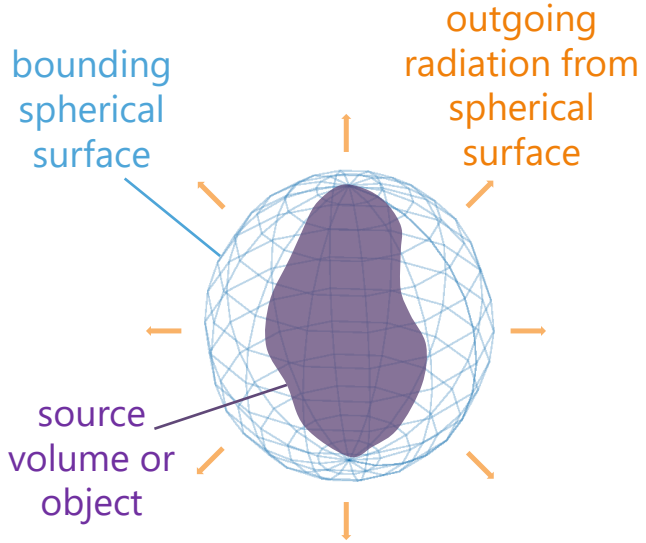
from some finite volume?

Our approach

Surround the volume with a mathematical "bounding" spherical surface

Count the number of well-coupled waves possible from this spherical surface

which then becomes the upper bound for waves from the source volume



D. A. B. Miller, Z. Kuang, O. D. Miller, "Tunneling escape of waves," http://arxiv.org/abs/2311.02744

Waves from arbitrary volumes

We show that, for spherical waves

with one key mathematical trick

there is a very simple and physical result

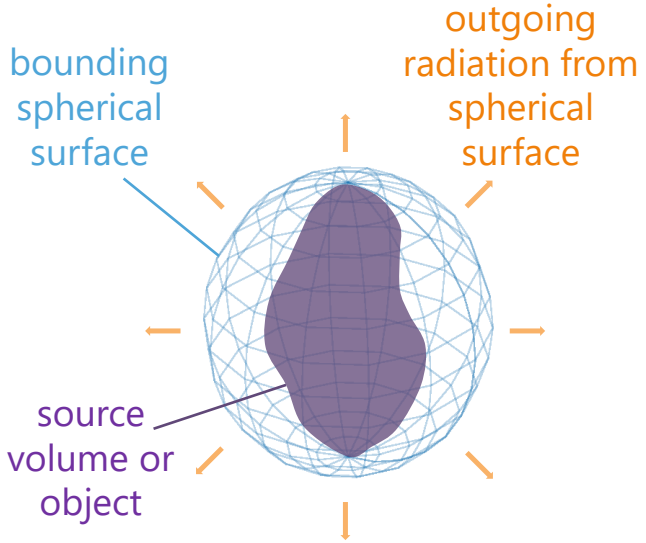
Beyond a certain simple threshold of "complexity" of spherical waves

they must "tunnel" to escape

Because the fall-off from tunneling is generally so rapid

this threshold effectively tells us the maximum number of well-coupled waves

and explains the quasi-exponential fall-off



D. A. B. Miller, Z. Kuang, O. D. Miller, "Tunneling escape of waves," http://arxiv.org/abs/2311.02744

Waves in spherical coordinates

In spherical coordinates r, θ, and ϕ

the solution to the wave equation separates to

$$U_{nm}(\mathbf{r}) = z_n(kr)Y_{nm}(\theta,\phi)$$

where $z_n(kr)$ is one of the spherical Bessel functions of order n, and

$Y_{nm}(\theta,\phi)$ is a spherical harmonic

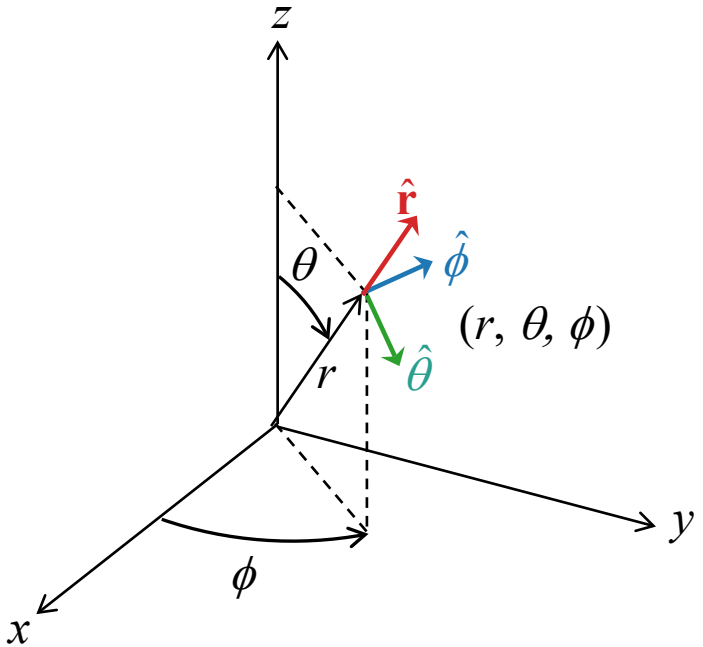
Here m and n are integers with

$n = 0,1,2,\ldots$ and $-n \le m \le n$

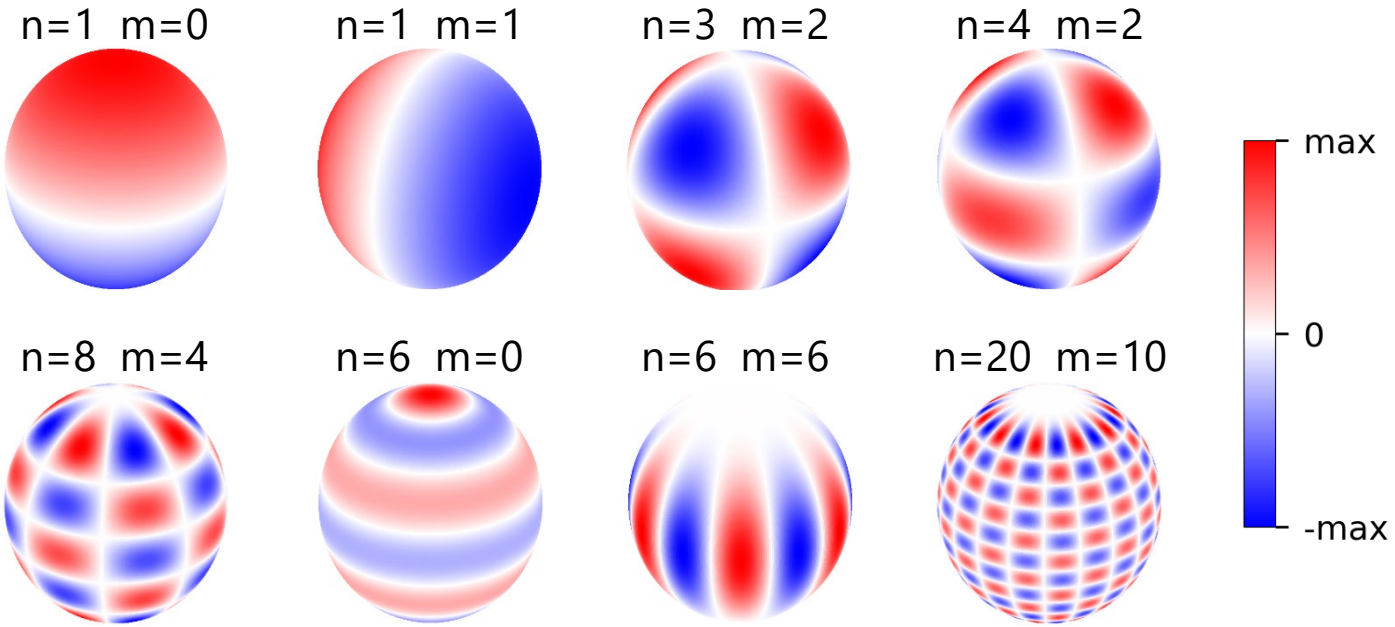
So, if we know the largest n for waves to propagate without tunneling

we can easily add up the total number of waves up to and including that n

$2n+1$ for each n



Spherical harmonics



Spherical harmonics are functions of angle only, and can be plotted on a spherical surface

They have *n* nodal circles altogether, with |*m*| through the poles (in their real form)

Escape radius

Specifically, for a given "order" n of spherical wave

there is an "escape radius"

$$r_{escn} = \frac{\sqrt{n(n+1)}}{k} \equiv \frac{\lambda_o}{2\pi}\sqrt{n(n+1)}$$

So, if the radius r_o of the spherical surface of interest

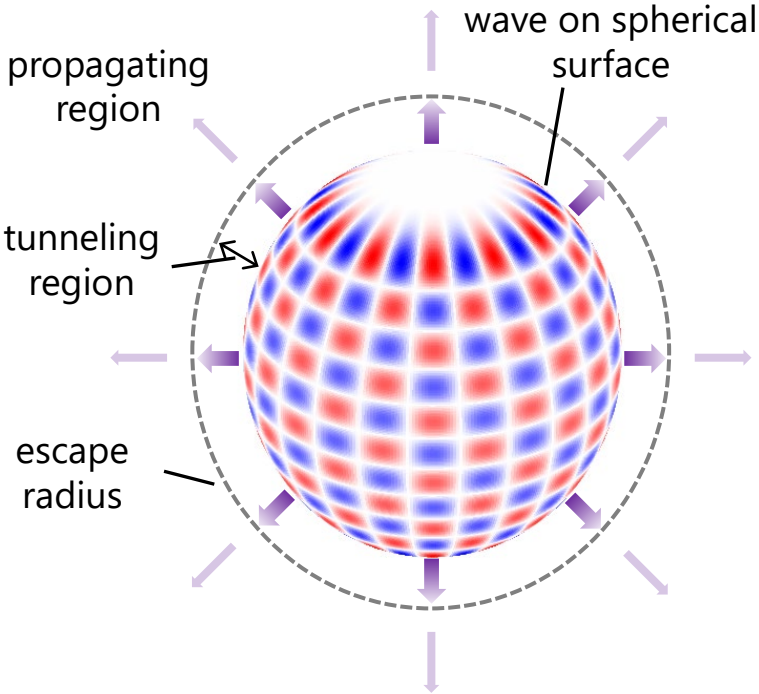
is smaller than the escape radius

for some order n of spherical wave

a wave with this n must tunnel

until it reaches the escape radius

after which it can propagate



D. A. B. Miller, Z. Kuang, O. D. Miller, "Tunneling escape of waves," http://arxiv.org/abs/2311.02744

Spherical Bessel functions and equation

Spherical Bessel functions obey

$$\rho^2 \frac{d^2 z_n(\rho)}{d\rho^2} + 2\rho \frac{dz_n(\rho)}{d\rho} + \left(\rho^2 - n(n+1)\right) z_n(\rho) = 0$$

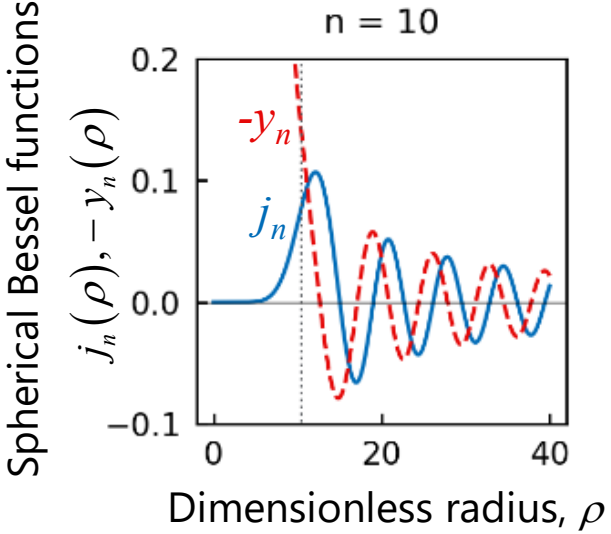
Classic radial standing wave solutions are

j_n which grows quasi-exponentially for small radii
and is quasi-oscillatory for larger radii

y_n which is singular at the origin
decaying quasi-exponentially for small radii
becoming quasi-oscillatory at large radii

Physically, ρ here is the dimensionless radius

$$\rho = kr \equiv 2\pi \frac{r}{\lambda}$$



Taking out the 1/radius dependence

Since the spherical Bessel functions have an underlying 1/radius dependence at large radius

as appropriate for what are ultimately spherically expanding waves

it could be useful to remove that dependence

multiplying by radius

which gives functions corresponding to power per unit solid angle

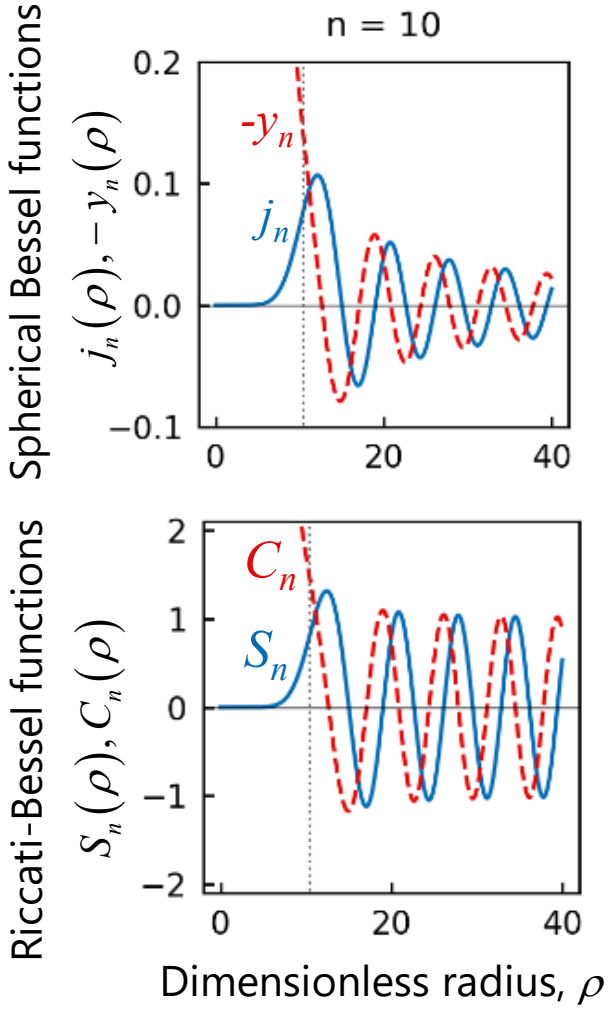
rather than power per unit area

So, we recast in terms of such functions

known as Riccati-Bessel functions

$$S_n(\rho) = \rho j_n(\rho) \qquad C_n(\rho) = -\rho y_n(\rho)$$

$$\xi_n(\rho) = \rho h_n^{(1)}(\rho) \equiv S_n(\rho) - iC_n(\rho)$$



Riccati-Bessel equation

Given that the spherical Bessel functions satisfy

$$\rho^2 \frac{d^2 z_n(\rho)}{d\rho^2} + 2\rho \frac{dz_n(\rho)}{d\rho} + \left(\rho^2 - n(n+1)\right) z_n(\rho) = 0$$

then we can easily check that all the Riccati-Bessel functions satisfy

$$\rho^2 \frac{d^2 \zeta_n}{d\rho^2} + \left(\rho^2 - n(n+1)\right) \zeta = 0$$

We can rearrange that to

$$-\frac{d^2 \zeta_n}{d\rho^2} + \frac{n(n+1)}{\rho^2} \zeta_n = \zeta_n$$

Riccati-Bessel "Schrödinger" equation

But wait!!!!!!

$$-\frac{d^2\zeta_n}{d\rho^2}+\frac{n(n+1)}{\rho^2}\zeta_n=\zeta_n$$

is in the form of a Schrödinger equation

$$-\frac{d^2\zeta_n}{d\rho^2}+V(\rho)\zeta_n=E_n\zeta_n$$

with effective radial potential

$$V(\rho)=\frac{n(n+1)}{\rho^2}$$

and the same "eigenenergy" $E_n=1$ for all n

Tunneling escape and escape radius

With the equation $$-\frac{d^2\zeta_n}{d\rho^2}+\frac{n(n+1)}{\rho^2}\zeta_n=\zeta_n$$

if the "potential energy" exceeds the "total energy", i.e., if

$$\frac{n(n+1)}{\rho^2}>1 \quad \text{or, equivalently} \quad n(n+1)>\rho^2$$

the wave will be tunneling rather than propagating

So, for each n, there is an "escape radius" $$\rho_{escn}=\sqrt{n(n+1)}$$

or, equivalently, in dimensioned form $$r_{escn}=\frac{\sqrt{n(n+1)}}{k}\equiv\frac{\lambda_o}{2\pi}\sqrt{n(n+1)}$$

Evanescent and spherical escaping waves

Both plane and spherical waves start with the same tunneling barrier height

and hence the same initial decay

but the falling barrier height for the spherical wave

means it eventually escapes

to being a propagating wave

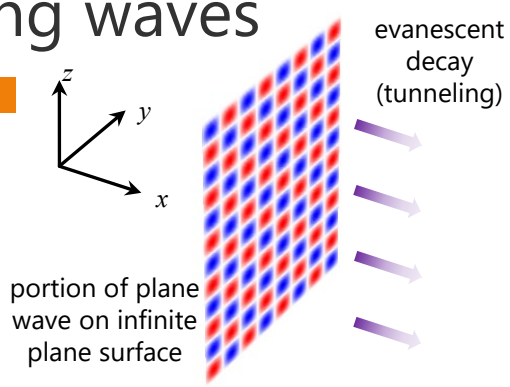
Note *all* such spherical waves

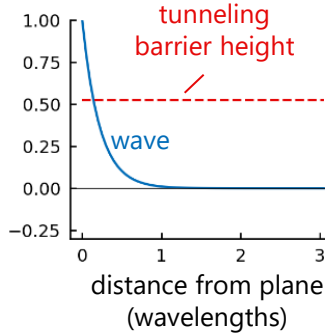
eventually escape to some degree

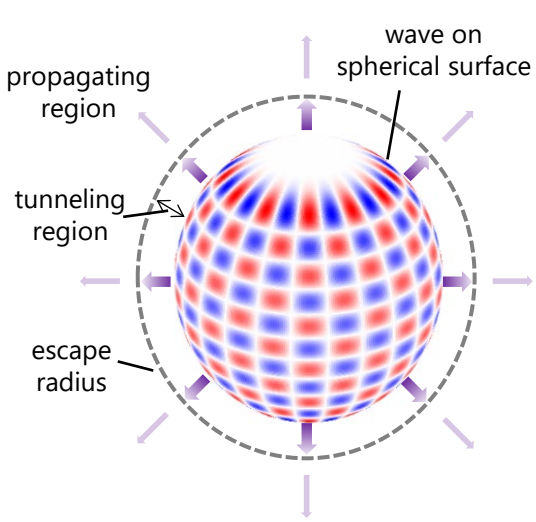
though the evanescent plane wave

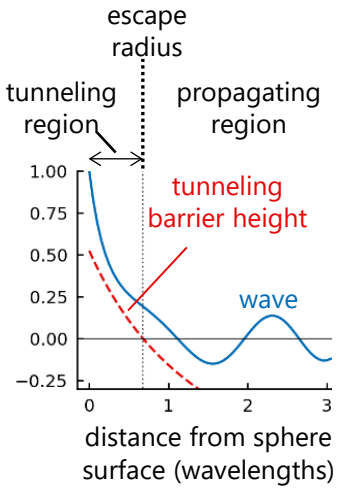
never does

This is an artifact of the "infinite" extent of the plane wave

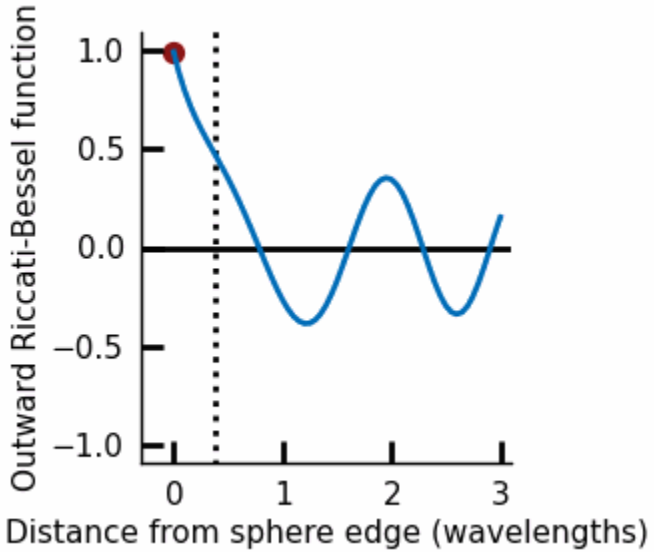


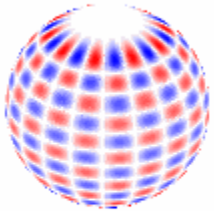






Snapshot in time of a spherical wave





Real part of outward (Riccati- Bessel) spherical wave

Starting spherical surface radius 2.9 wavelengths

Wave with n = 20, m = 10, escape radius 3.26 wavelengths

Note the angular shape is constant as the wave expands

Outward wave propagation

As time progresses

the wave beyond the escape radius

propagates outwards

We plot the outward Riccati-Bessel wave

as a function of time

technically the real part of

$$\xi_n(2\pi r)\exp(-i\omega t)$$

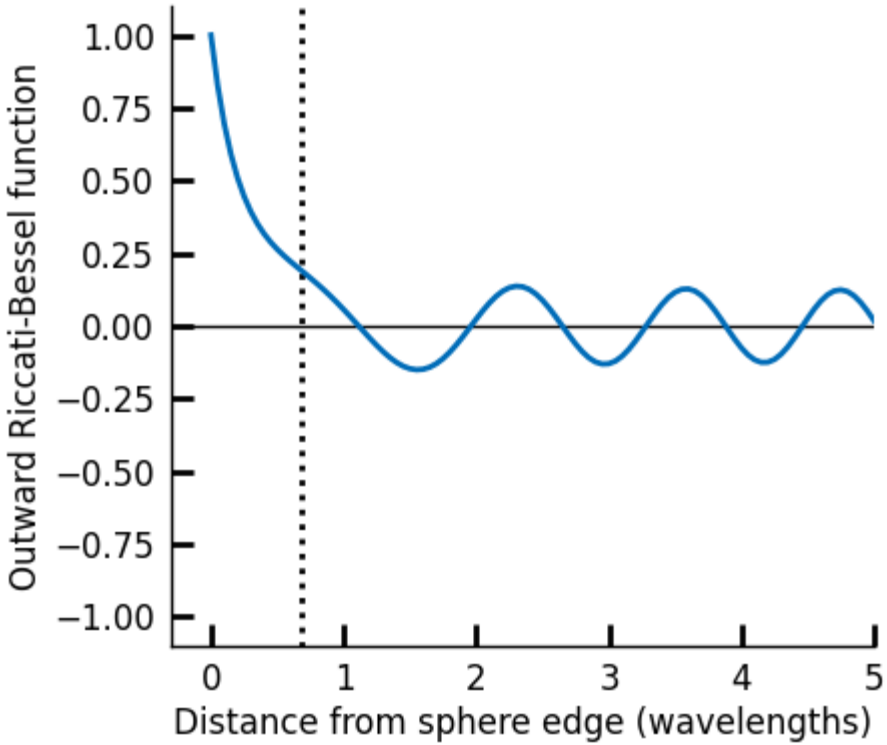
normalized to unit amplitude at the sphere edge

for a sphere of radius 2.9 wavelengths

with $n = 22$

which has an escape radius of

3.58 wavelengths



$r_o = 2.9 \quad n = 22 \quad r_{escn} = 3.58$

Outward wave propagation

As time progresses

the wave beyond the escape radius

propagates outwards

We plot the outward Riccati-Bessel wave

as a function of time

technically the real part of

$$\xi_n(2\pi r)\exp(-i\omega t)$$

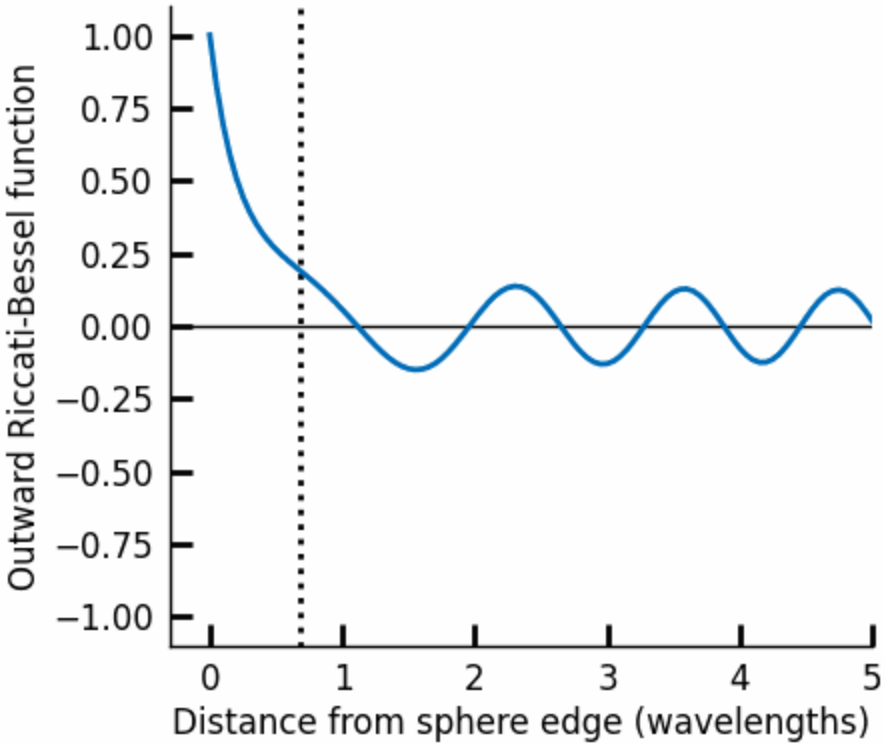
normalized to unit amplitude at the sphere edge

for a sphere of radius 2.9 wavelengths

with $n = 22$

which has an escape radius of

3.58 wavelengths



$r_o = 2.9 \quad n = 22 \quad r_{escn} = 3.58$

Spherical heuristic number

The threshold for tunneling is easy to characterize

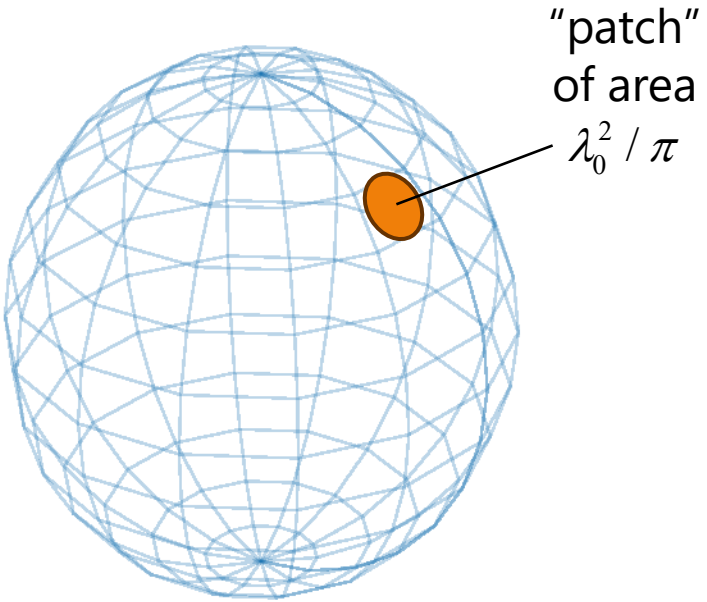
and gives a simple answer for the number of waves that do not need to tunnel

This is well approximated by the spherical heuristic number

$$N_{SH} = (kr_o)^2 \equiv \left(\frac{2\pi r_o}{\lambda_o}\right)^2 \equiv \frac{4\pi r_o^2}{(\lambda_o^2/\pi)} \equiv \frac{A_S}{(\lambda_o^2/\pi)}$$

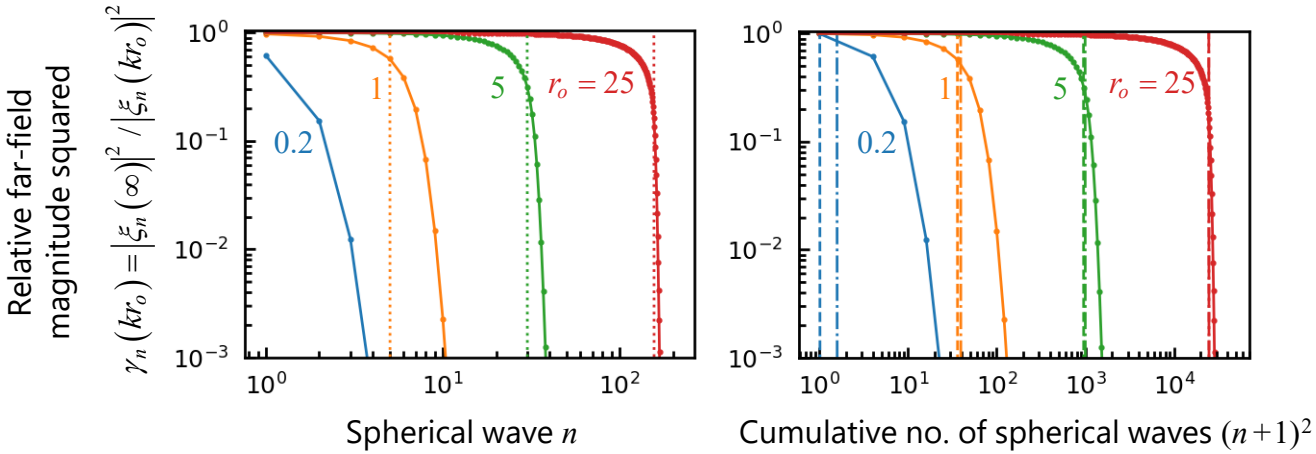
where A_S is the sphere area

so one "propagating" wave for every λ_o^2/π of surface area



D. A. B. Miller, Z. Kuang, O. D. Miller, "Tunneling escape of waves," http://arxiv.org/abs/2311.02744

Relative far-field magnitude squared



dotted lines – n_p
dashed lines – N_{SH}
dot-dashed lines - N_p (the exact counting result)

As the size of the spherical surface increases

the cut-off becomes increasingly relatively abrupt

tending towards the "absolutely abrupt" cut-off of evanescent waves

Note the spherical heuristic number N_{SH} is a good approximation to the total exact number N_p of "propagating" waves even down to ~ 1 wavelength of radius

Defining the diffraction limit

We can now construct a precise definition of

the “diffraction limit”

For a wave interacting with a volume

the wave passes the diffraction limit

if any spherical component of the wave must tunnel to enter or leave the bounding spherical surface enclosing the volume

No n=0 electromagnetic outgoing waves

Electromagnetic waves behave mostly similarly to scalar (e.g., acoustic) waves

but, unlike for acoustic waves, there is no electromagnetic wave that is uniform in angle

Mathematically, there is no $n = 0$ wave in electromagnetism

If the first outgoing electromagnetic waves (so, for $n = 1$)

are not to require tunneling to escape

the bounding spherical volume must be at least

$$r_{esc1} = \lambda_o / \left(\sqrt{2}\,\pi\right) \simeq 0.225\lambda_o$$

in radius or, equivalently, in diameter

$$d = \sqrt{2}\lambda_o / \pi \simeq 0.45\lambda_o$$

(consistent with the well-known Chu limit on antenna Q)

Note: The escape radius for $n = 0$ acoustic waves is, however, zero

so, there is always one acoustic wave that can escape without tunneling

no matter how small the emitter or microphone

Perfect cloaking - An optical “white hole”?

In this “white hole”, incoming light

- appears to be mostly “sucked” into the “white hole” in the middle
  - The phase fronts all “fall” rapidly into the “white hole”
- and then the light is regenerated
  - The phase fronts rapidly re-emerge from the “white hole”

How do we make this optical “white hole”?

- Note: it may be simpler than you think

Perfect cloaking - An optical “white hole”?

So, what does it take to build this cloak?

Absolutely nothing

at least for this wave

If the wave is too complicated

i.e., if it is trying to violate the “diffraction limit”

it can’t even effectively get into the volume

and it “reflects off free space”

This is the “inward wave” version of the tunneling escape

with the wave trying to tunnel to get in

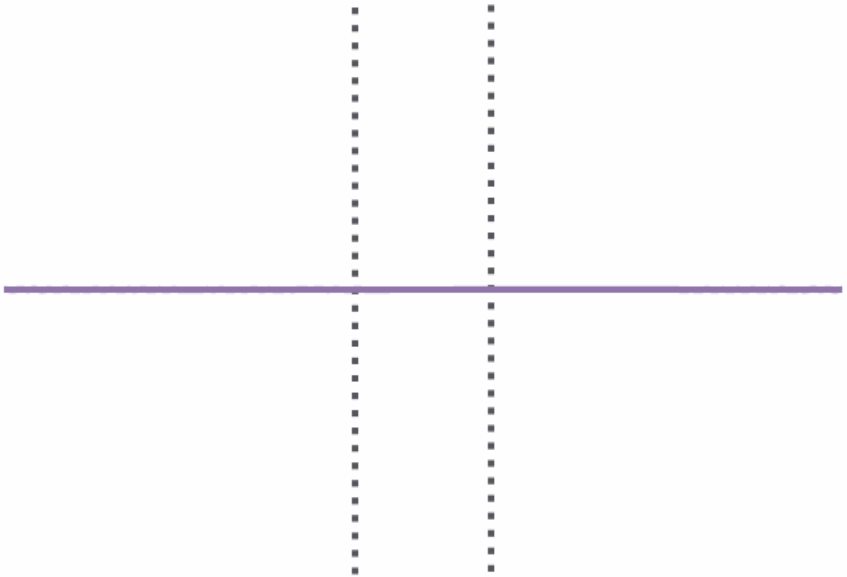
Interestingly, the pulse

looks as if it propagates right through!

see also Z Jacob, L V Alekseyev, and E Narimanov, Opt. Express 14, 8247 (2006)

Perfect cloaking?

Watch the blue dot, which propagates at the usual “phase velocity” of the wave



It appears to move right through the volume at a constant speed

Conclusions

New generations of programmable optics are emerging now

enabling a wide range of things we could not do before

and even setting themselves up

A new modal way of looking at optics

based on singular value decomposition

describes these and other optical devices

and gives us new fundamental and practical results and limits

stanford.io/3NK9w8Z

David Miller

Supported by Air Force Office of Scientific Research
FA9550-17-1-0002 and FA9550-21-1-0312

For a copy of these slides e-mail
dabm@stanford.edu