NSF CyberTraining Winter Workshop

Introduction to Proper Orthogonal Decomposition

concepts, formulation and applications

Part I: Physics Simulations and POD Fundamentals

Ming-Cheng Cheng

Department of Electrical & Computer Engineering
Clarkson University, Potsdam, NY 13699

Supported by NSF OAC-2118079

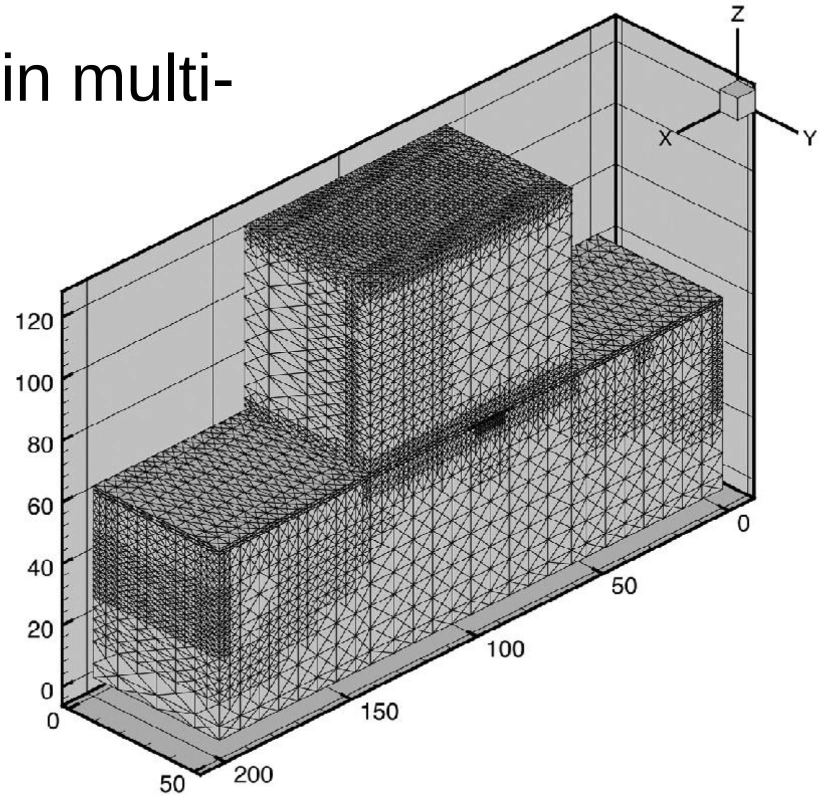
Applications of POD Simulation Methodology

Physics Simulation of Domain Structures with Spatial Details

Numerical solution with detailed spatial solution in multi-dimensional structures:

Extremely time consuming. Why?

- Because of a large number of grids (nodes, or degrees of freedom) needed to solve the physical quantity numerically
- Usually, several 100 thousands or millions of nodes for a 3D problem
- Several 100 thousands or millions of coupled differential equations are needed



Mesh of a MOSFET

https://www.researchgate.net/figure/Mesh-of-a-MOSFET-with-a-polysilicon-gate_fig1_267688735

Applications of POD Simulation Methodology

Physics Simulation of Domain Structures with Spatial Details

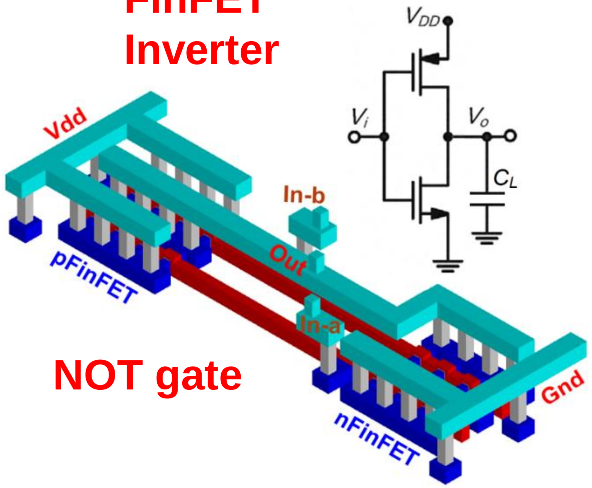
A few examples in different areas of research and industrial applications:

- **Heat Transfer Problems:** Heat Transfer Equation
- **Nanostructures and Materials:** Schrödinger Equation - Quantum Wave Equation (Quantum Eigenvalue problem), relevant to DFT simulation
- **Electromagnetics and Photonics:** Eigenvalue Problem or dynamic Wave Propagation
- **Charge Carrier Transport in Semiconductor Devices:** Electron/hole transport equations in (6D) phase space
- **Phonon Transport in nanostructures/nanodevices:** Phonon Boltzman Transport Equation (6D) phase space
- **Fluid dynamics:** Navier Stokes Equations, conservation of energy, momentum, and mass

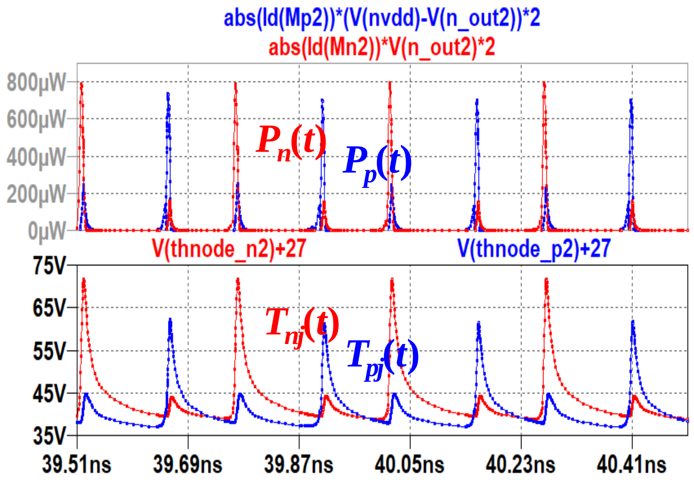
Heat Transfer problems, for example, for semiconductor technology: semiconductor devices, integrated circuits, etc.

$$\rho C_h \frac{\partial T(\vec{r},t)}{\partial t} = \nabla \cdot k \nabla T(\vec{r},t) + P_d(\vec{r},t)$$

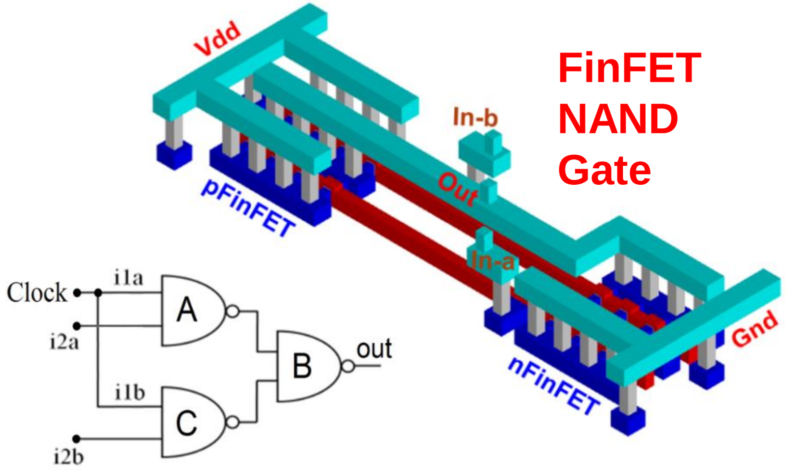
FinFET Inverter

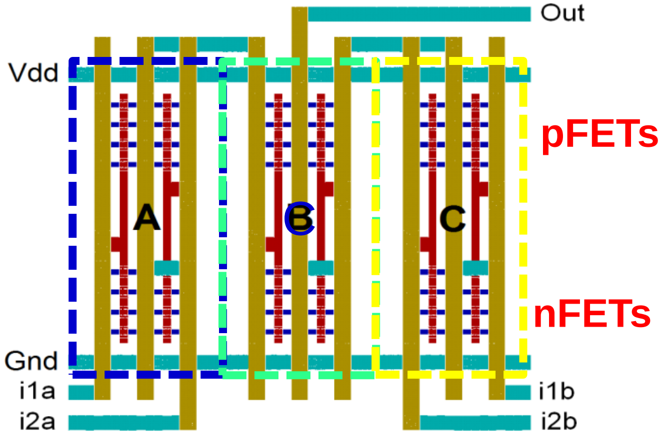


NOT gate



FinFET NAND Gate

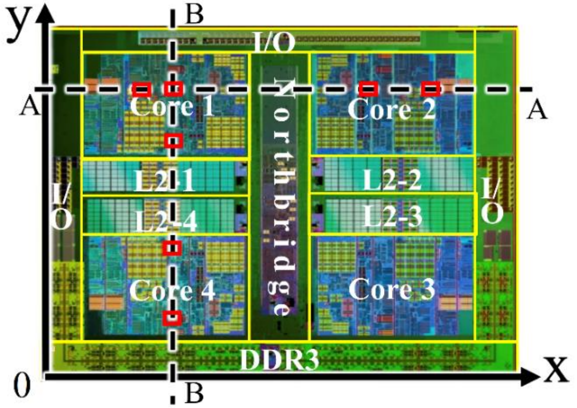




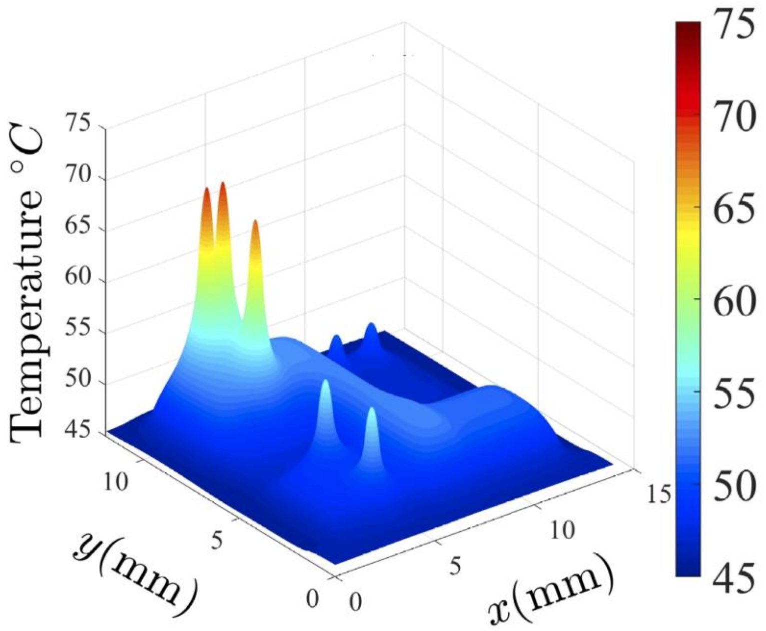
Layout of the IC structure

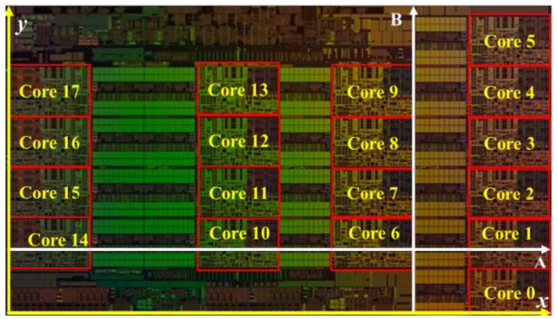
Heat Transfer problems: CPUs, GPUs, etc.

$$\rho C_h \frac{\partial T(\vec{r},t)}{\partial t} = \nabla \cdot k \nabla T(\vec{r},t) + P_d(\vec{r},t)$$

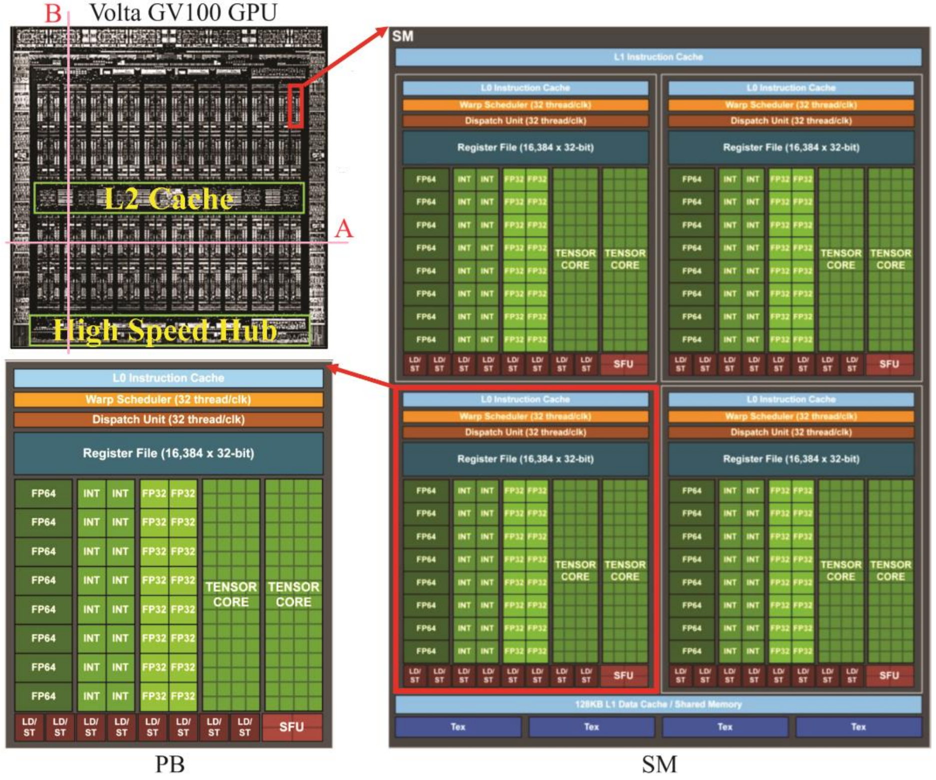


Quad-core AMD ATHLON II X4 610e





Intel Xeon 18 core E5-2699v3 CPU

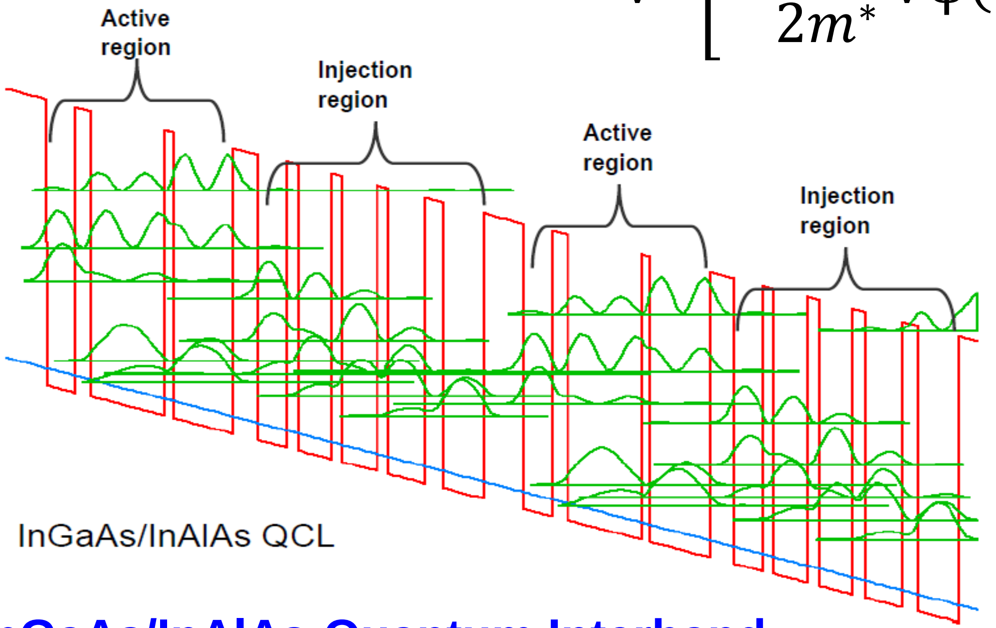


Tesla Volta GV100 GPU with 13,440 cores, including FP32, FP64, INT32 and Tensor Cores.

Quantum eigenvalue problem, Schrödinger equation

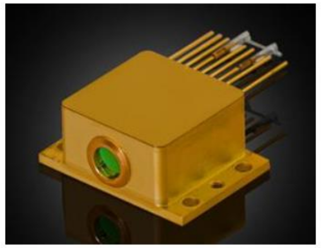
nanostructure, superlattice materials, density functional theory (DFT)

$$\nabla \cdot \left[-\frac{\hbar^2}{2m^*} \nabla\psi(\vec{r}) \right] + U(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}),$$



InGaAs/InAlAs Quantum Interband Cascade Laser (IEEE JQE Vol. 40, p. 1663, 2004)

Quantum Cascade Laser Power Supply

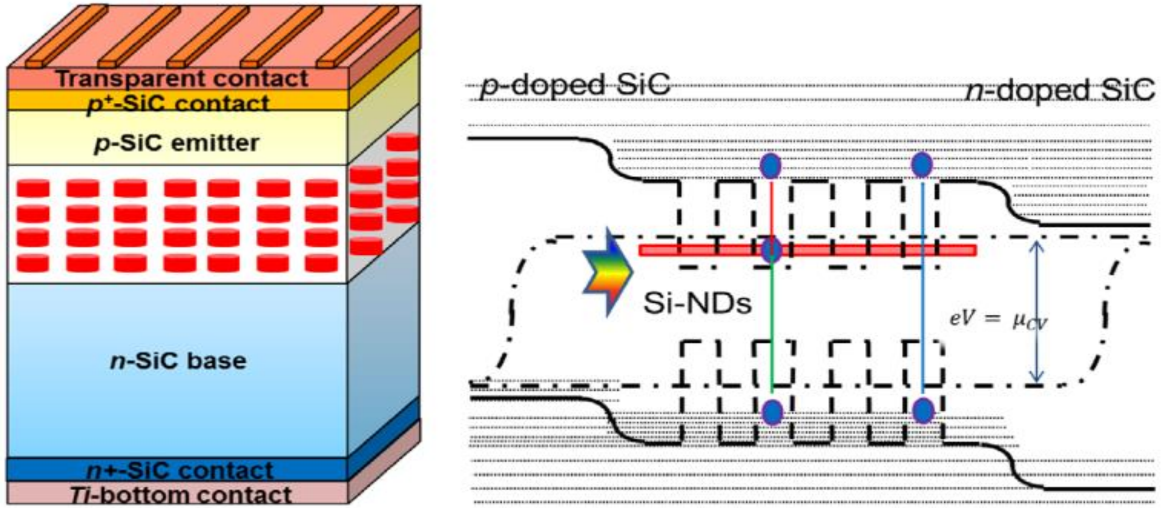


Quantum Cascade Laser (QCL) Systems

Stock #15-957

$5,599.00

Qty 1+	Volume Pricing
$5,599.00	Request Quote



Si quantum-dot intermediate band solar cell

Nanotechnology, 24, 265401, 2013

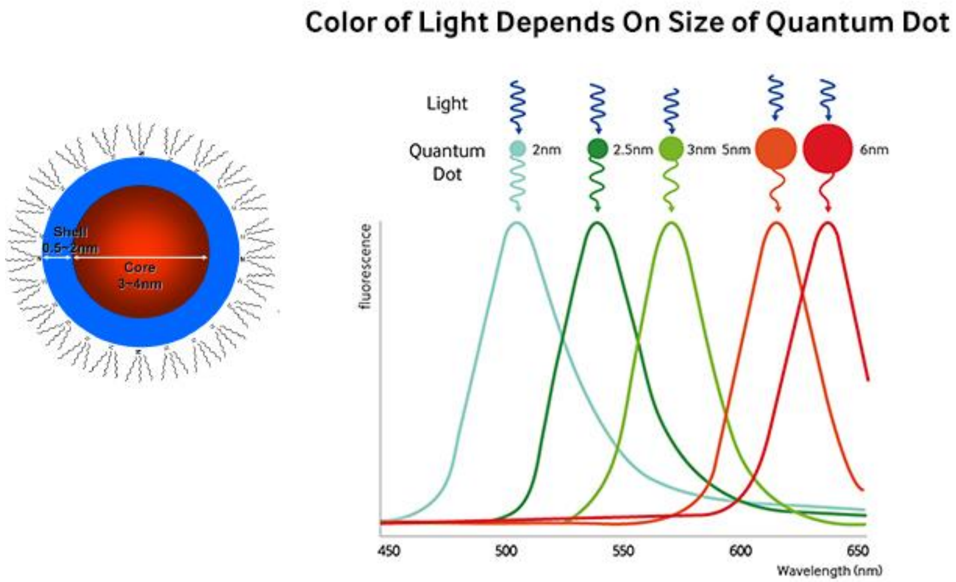
Quantum eigenvalue problem, Schrödinger equation

nanostructure, superlattice materials, density functional theory (DFT)

$$\nabla \cdot \left[-\frac{\hbar^2}{2m^*} \nabla \psi(\vec{r}) \right] + U(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}),$$

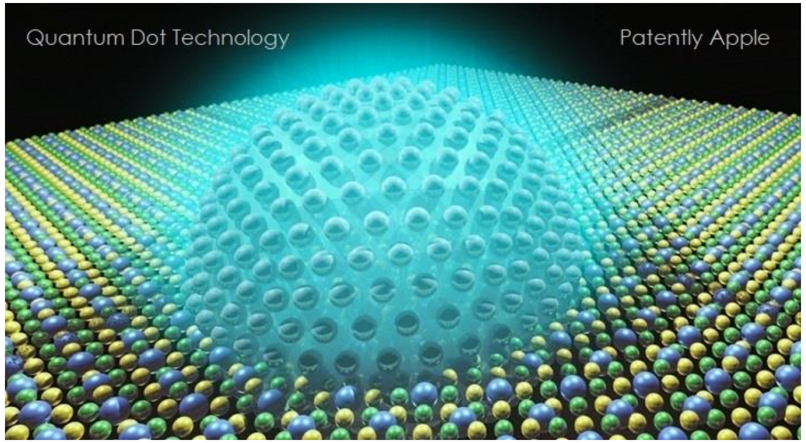
Quantum Dot LEDs for TV and Cell phone Displays

Samsung: Quantum Dot LEDs



https://news.samsung.com/global/why-are-quantum-dot-displays-so-good

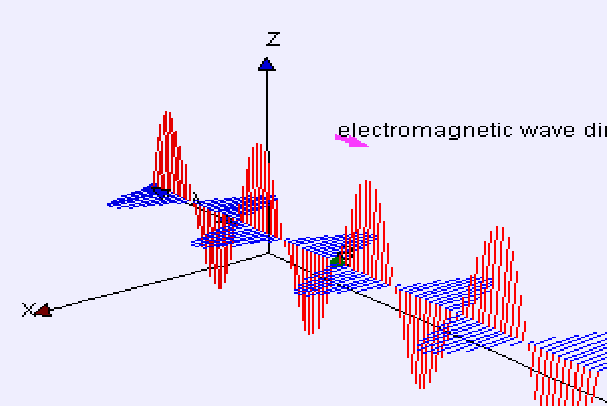
Apple: Hybrid Quantum Dot LED and OLED Displays



https://www.patentlyapple.com/2019/10/apple-won-42-patents-today-covering-hybrid-quantum-dot-displays-guis-supporting-3d-ar-models-a-health-study-more.html

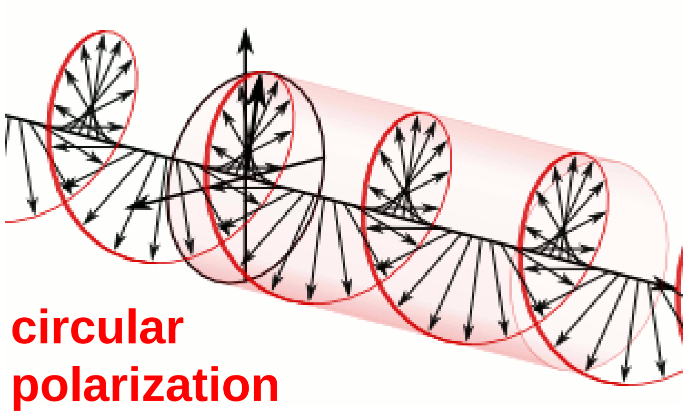
Electromagnetics or photonics structures: Dynamic wave equation, wave equation in frequency domain, EM eigenvalue problem)

- **Dynamic EM Wave equation:** $\mu\epsilon \frac{\partial^2 \vec{E}(\vec{r},t)}{\partial t^2} + \mu\sigma \frac{\partial \vec{E}(\vec{r},t)}{\partial t} - \nabla^2 \vec{E}(\vec{r},t) = -\mu \frac{\partial \vec{j}(\vec{r},t)}{\partial t}$



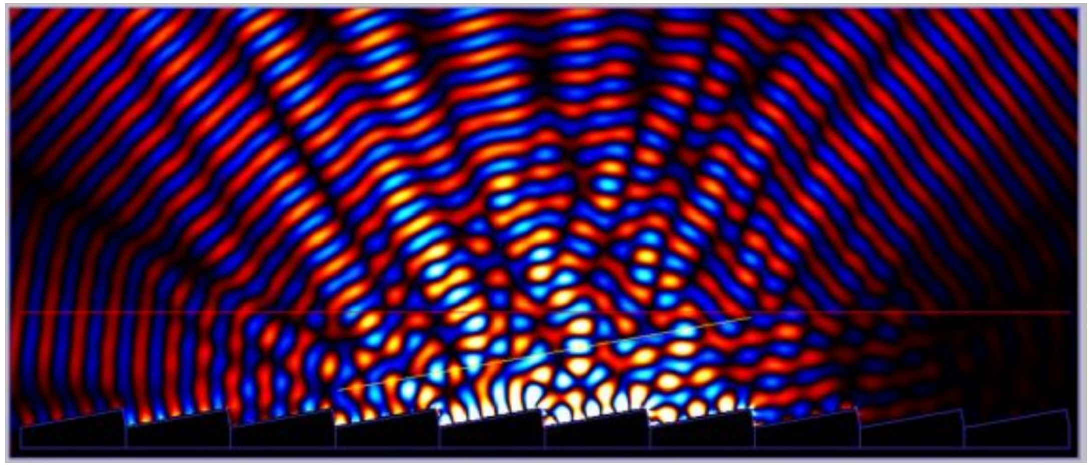
Linear polarization

https://en.wikipedia.org/wiki/Electromagnetic_radiation



circular polarization

https://en.wikipedia.org/wiki/Electromagnetism



Diffraction of a Gaussian beam on a grating structure simulated using the FDTD (finite-difference time-domain) method.

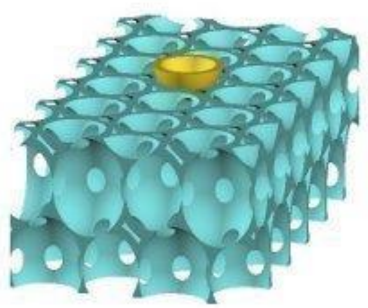
(https://www.photond.com/products/omnisim/omnisim_applications_06.htm)

Electromagnetic Eigenvalue Problems

Electromagnetic Band Gap (EBG) Structures

$$\nabla^2 \vec{E}(\vec{r}) = -\left(\frac{\omega}{c}\right)^2 \epsilon_r \vec{E}(\vec{r})$$

$$\text{Bloch Function: } \vec{E}(\vec{r}) = \vec{u}(\vec{r}) e^{i\vec{k}\cdot\vec{r}}$$

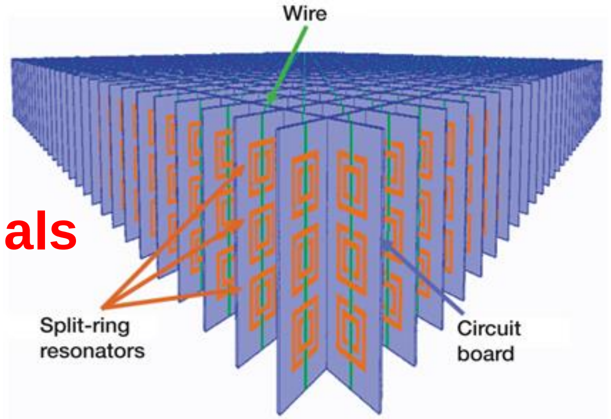


Photonic Crystals

Applications:

- Color selections
- Sensors
- Optical filters
- Photonic Crystal fibers
- Optical computing

Metamaterials



https://www.tech-faq.com/what-are-photonic-crystals.html
https://en.wikipedia.org/wiki/Photonic_crystal

https://en.wikipedia.org/wiki/Negative_index_metamaterials

Charge carrier transport in semiconductor devices (TCAD):

Boltzmann transport equation (BTE), hydrodynamic, energy transport model, drift-diffusion model

10μm   1μm   0.1μm   0.01μm

Semiclassical BTE

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f + \frac{q\mathbf{E}}{\hbar}\cdot\nabla_k f = \left(\frac{\partial f}{\partial t}\right)_C$$

Probability density function in 6D space: $f(\vec{r}, \vec{k}, t)$

Quantum Model

Hydrodynamic model

$$\partial n / \partial t + \nabla\cdot n\bar{\mathbf{v}} = G_n - R_n$$

$$\partial n\bar{\mathbf{p}} / \partial t + \nabla(n\bar{\mathbf{p}}\cdot\bar{\mathbf{v}}) + \nabla(nk_B T_e) = qn\mathbf{E} - n\bar{\mathbf{p}} / \tau_m$$

$$\partial n\bar{\varepsilon} / \partial t + \nabla\cdot n\mathbf{S} = qn\mathbf{E}\cdot\bar{\mathbf{v}} - n(\bar{\varepsilon} - \varepsilon_o) / \tau_\varepsilon ,$$

Energy transport model

$$\partial n / \partial t - \nabla\cdot\mathbf{J}_n / q = G_n - R_n$$

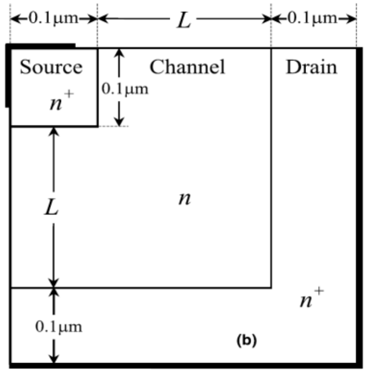
$$\mathbf{J}_n = qn\mu_n\mathbf{E} + \mu_n n\nabla k_B T_e + \mu_n k_B T_e \nabla n$$

$$\partial n\bar{\varepsilon} / \partial t + \nabla\cdot n\mathbf{S} = qn\mathbf{E}\cdot\bar{\mathbf{v}} - \partial\, n(\bar{\varepsilon} - \varepsilon_o) / \tau_\varepsilon$$

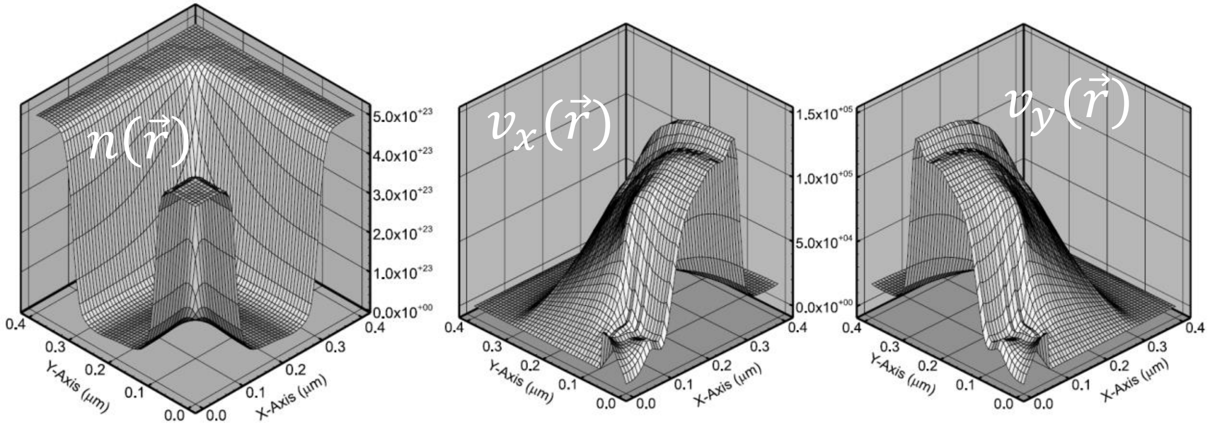
Drift-diffusion model

$$\partial n / \partial t + \nabla\cdot(n\bar{\mathbf{v}}) = G_n - R_n$$

$$\mathbf{J}_n = -qn\bar{\mathbf{v}} = qn\mu_n\mathbf{E} + qD_n\nabla n$$



Hydrodynamic Model



M. Shen, T. Zhou, MC Cheng, R. Fithen, *Computer Methods Appl. Mech & Eng*, Vol. 190, 2875-2891, February 16, 2001.

How to develop a very efficient model (compact model) for a structure that requires the solution with spatial details?

Reduce the **numerical degrees of freedom (DoF)** by several orders of magnitude!

How? Projection or Mathematical Transformation

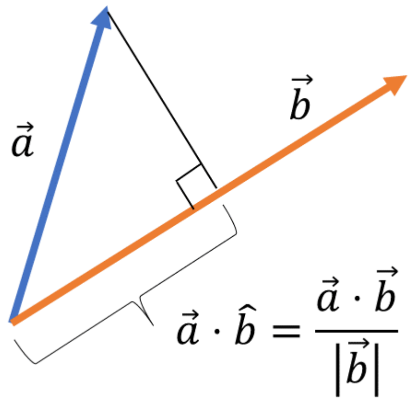
Inner Product of vectors:

$$\vec{a} \cdot \vec{b} = |a||b|cos\theta = \sum_{i=1}^{M} a_i b_i$$

$= a_x b_x + a_y b_y + a_z b_z$ (in real space)

L2 Norm: $|b| = \sqrt{\sum_{i=1}^{M} b_i^2}$

Projection of $\vec{a}$ onto $\vec{b}$:



Inner Product of functions:

$$f \cdot g = \int_a^b f(x)g(x)dx$$

L2 Norm: $|g| = \sqrt{\int_a^b g(x)^2 dx}$

Projection of f onto g : $\frac{f \cdot g}{|g|}$

basis vector

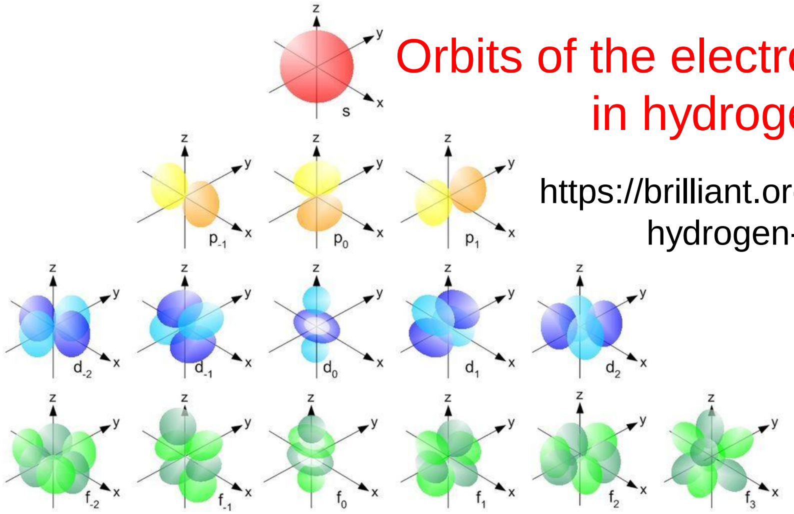
POD: Maximizing the average of (Projection)^2 $\left\langle \left(\frac{f \cdot g}{|g|} \right)^2 \right\rangle$

- A popular projection, **Fourier series** (frequency domain projection)

$$f(t)=\sum_{n=1}^{M}a_n e^{jn\omega t}=\sum_{n=1}^{M}a_n\eta_n(t),\quad \text{or in space}\quad f(x)=\sum_{n=1}^{M}a_n e^{jnkx}=\sum_{n=1}^{M}a_n\eta_n(x)$$

where the basis functions (or modes) $e^{jn\omega t}$ or e^{jnkx} are effective only for **periodic functions**; $\omega=\frac{2\pi}{T}$ and $k=2\pi/\lambda$

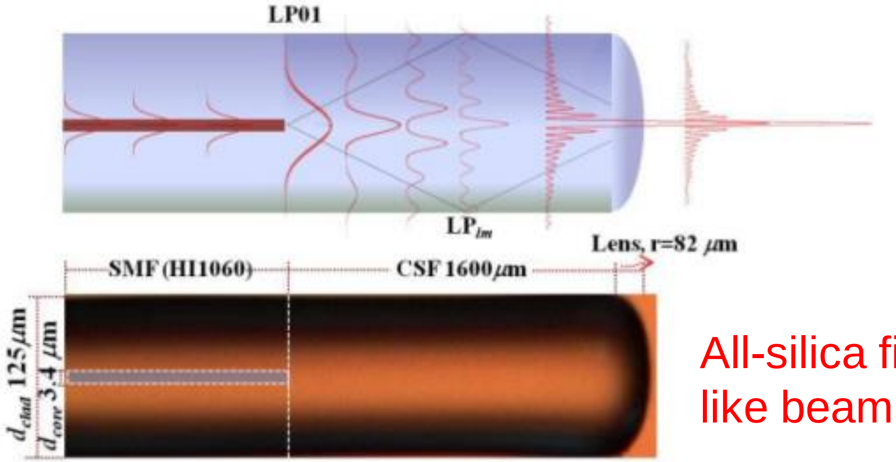
- Spherical Harmonic Expansion: **Legendre Polynomials**



Orbits of the electron in hydrogen

https://brilliant.org/wiki/hydrogen-atom/

- Cylindrical harmonic expansion: **Bessel Functions**



All-silica fiber Bessel-like beam generator

https://opg.optica.org/oe/fulltext.cfm?uri=oe-18-24-25299&id=208252

- Instead of assuming the functions for the modes, a learning algorithm can be used to extract the modes.

  The most commonly used projection-based learning methods:

  - **PCA:** Principal component analysis
  - **SVD:** Singular value decomposition

- The machine learning methods only provide statistical models to minimize the statistical variation in the prediction

  \- Offer no guideline on physical principles in the governing equation for the spatial variation and/or dynamic evolution for the physical quantity.

  - Only work well for prediction of one-to-one correspondence relation
  - Not accurate for dynamic (initial value) problems
  - Poor performance for sudden change in physical quantity
  - Work well for interpolation but poorly for extrapolation

Effective projections of both the data and the governing equation onto the POD space

$$Q(\vec{r},t) = \sum_{j=1}^{M} a_j(t)\,\eta_j(\vec{r})$$

- POD finds a mode $\eta(\vec{r})$ that maximizes its mean square inner product (projection) with the solution data

$$\left\langle \left[\frac{Q\cdot\eta}{|\eta|}\right]^2 \right\rangle = \left\langle \left(\int_{\Omega} Q(\vec{r},t)\eta(\vec{r})d\Omega\right)^2 \right\rangle \Big/ \int_{\Omega} \eta(\vec{r})^2 d\Omega$$

$\langle\ \rangle$ indicate the average of the data collected over numerical observations (or snapshots in time or over the samples) **accounting for the parametric variations.**

- This maximization process leads to an eigenvalue problem of 2-point correlation of data

$$\int_{\Omega'} \langle Q(\vec{r},t) \otimes Q(\vec{r}',t)\rangle \eta(\vec{r}')d\Omega' = \lambda\eta(\vec{r}),$$

which offers a minimum least square error with a smallest number of modes (DoF) **if the training is properly done (i.e., if the data quality is sufficient)**.

Continuous eigenvalue problem

$$\int_{\Omega'} \langle Q(\vec{r},t) \otimes Q(\vec{r}',t) \rangle \eta(\vec{r}') d\Omega' = \lambda \eta(\vec{r}),$$

Q data sets are collected from detailed numerical simulation (DNS)

Discretization

$$\langle Q(\vec{r},t) \otimes Q(\vec{r}',t) \rangle = \frac{1}{N_s} \sum_{j=1}^{N_s} \vec{Q}(\vec{r},t_j) \otimes \vec{Q}(\vec{r}',t_j) = \frac{1}{N_s} \sum_{j=1}^{N_s} \vec{Q}(t_j)\, \vec{Q}^T(t_j)$$

for each sample: $$\vec{Q} \otimes \vec{Q} = \vec{Q}\, \vec{Q}^T = \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_{N_r} \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 & \dots Q_{N_r} \end{bmatrix} = \begin{bmatrix} Q_1 Q_1 & \cdots & Q_1 Q_j & \cdots & Q_1 Q_{N_r} \\ \vdots & \ddots & \vdots & ⋰ & \vdots \\ Q_i Q_1 & \cdots & Q_i Q_j & \cdots & Q_i Q_{N_r} \\ \vdots & ⋰ & \vdots & \ddots & \vdots \\ Q_{N_r} Q_1 & \cdots & Q_{N_r} Q_j & \cdots & Q_{N_r} Q_{N_r} \end{bmatrix},$$

A matrix dimension of N_r x N_r may be too large to manage in 3D problems for a dense matrix.
The method of snapshots will be introduced later to minimize the computational effort

Example, heat conduction in semiconductor Chips

$$Q(\vec{r},t) = T(\vec{r},t) = \sum_{j=1}^{M} a_j(t)\,\eta_j(\vec{r})\,; \quad a_j(t) \text{ needs to be determined}$$

Heat conduction equation: $$\frac{\partial \rho C T(\vec{r},t)}{\partial t} - \nabla \cdot k\nabla T(\vec{r},t) = P_d(\vec{r},t)$$

- **Galerkin projection** (transformation) of the heat conduction equation onto the the *i*th POD mode, $\eta_i(\vec{r})$

$$\int_\Omega \left(\frac{\partial \rho C T(\vec{r},t)}{\partial t} - \nabla \cdot k\nabla T(\vec{r},t) = P_d(\vec{r},t) \right) \eta_i(\vec{r})\, d\Omega$$

Using the following identities:

$$\nabla \cdot (\eta_i k \nabla T) = \nabla \eta_i \cdot k\nabla T + \eta_i \nabla \cdot k\nabla T$$

$$\text{Gauss's Law: } \int_\Omega \nabla \cdot (\eta_i k \nabla T)\, d\Omega = \int_S \eta_i k \nabla T \cdot d\vec{S}$$

vol. surface

The projection leads to the weak form for the heat condition equation

$$\int_\Omega \left(\eta_i \frac{\partial \rho C T}{\partial t} + \nabla \eta_i \cdot k\nabla T \right) d\Omega = \int_\Omega \eta_i P_d\, d\Omega + \int_S \eta_i k \nabla T \cdot d\vec{S}$$

Example, heat conduction in semiconductor Chips

- **Galerkin projection** (transformation) onto the the *i*th POD mode, $\eta_i(\vec{r})$

$$\int_\Omega \left(\eta_i \frac{\partial \rho C T}{\partial t} + \nabla \eta_i \cdot k \nabla T \right) d\Omega = \int_\Omega \eta_i P_d \, d\Omega + \int_S \eta_i k \nabla T \cdot d\vec{S}$$

- Using $T(\vec{r}, t) = \sum_{j=1}^{M} a_j(t) \eta_j(\vec{r})$ ➔ *M* mode POD model, a set of *M*-dimensional ODEs

$$\sum_{j=1}^{M} c_{i,j} \frac{da_j}{dt} + \sum_{j=1}^{M} g_{i,j} a_j = P_{pod,i},$$

$$c_{i,j} = \int_\Omega \rho C \eta_i \eta_j \, d\Omega, \quad g_{i,j} = \int_\Omega k \nabla \eta_i \cdot \nabla \eta_j \, d\Omega, \quad P_{pod,i} = \int_\Omega \eta_i P_d(\vec{x}, t) \, d\Omega - \int_\Gamma \eta_i (-k \nabla T) \cdot d\vec{S}$$

Procedure for Constructing the POD Simulation Model

1. Data Collection from direct numerical simulation (DNS) ➔ $Q(\vec{r}, t)$

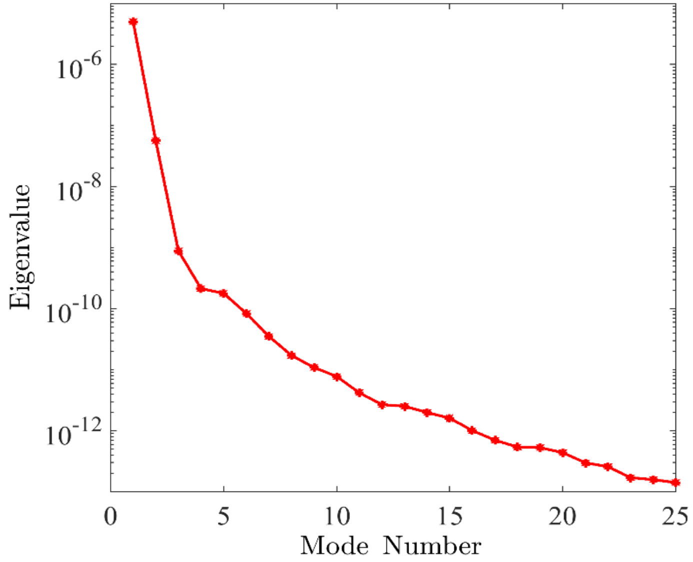
2. Solving the 2-point correlation eigenvalue problem for λ_j & $\eta_j(\vec{r})$.

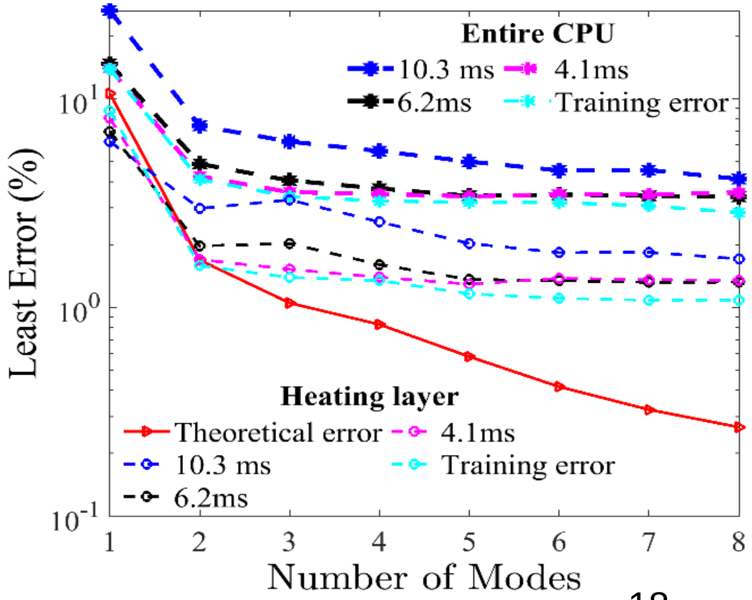
   Observe the Eigenvalue spectrum to determine the number of modes, *M,* where λ_j represents the mean squared information captured by η_j Or estimate the least square error based on

$$Err_{LS,M} = \sqrt{\sum_{i=M+1}^{N_s} \lambda_i \Big/ \sum_{i=1}^{N_s} \lambda_i}$$

3. Project the governing equation onto the POD Space (accounting for physical principles) ➔ a set of *M* ODEs for a_j

4. Evaluate the model parameters (coefficients of the ODEs





Implementation of POD in physics simulation

- Solve the ODEs to obtain a_j
- Post processing: the solution $Q(\vec{r}, t) = \sum_{j=1}^{M} a_j(t)\, \eta_j(\vec{r})$