Welcome to PHYS50004.01

Introduction to

Algorithms for Data Science and Physics

Lecturer: Dr. Bao-Jun Cai

IMP@Fudan, Spring/2026

Topics to be covered of this course

A. perturbative theories, numerical calculus, statistics and probability

B. optimization algorithms (1st and 2nd), conjugate-gradient, Levenberg-Marquardt, bias-variance decomposition, annealing and statistical mechanics, Metropolis sampling (Monte Carlo)

C. Kalman filter: data-driven optimization, Bayesian inference (state/parameter estimate), 3D reconstruction from 2D images (large-scale matrix computing)

D. sparsity/clustering, SVD/PCA (MNIST), randomized algorithms, massive datasets

E. FFT, differencing schemes for ODE, random walk simulation, convolution and deconvolution

F. quantum linear algebra, Grover's search, path-integral simulation, algorithmic interference

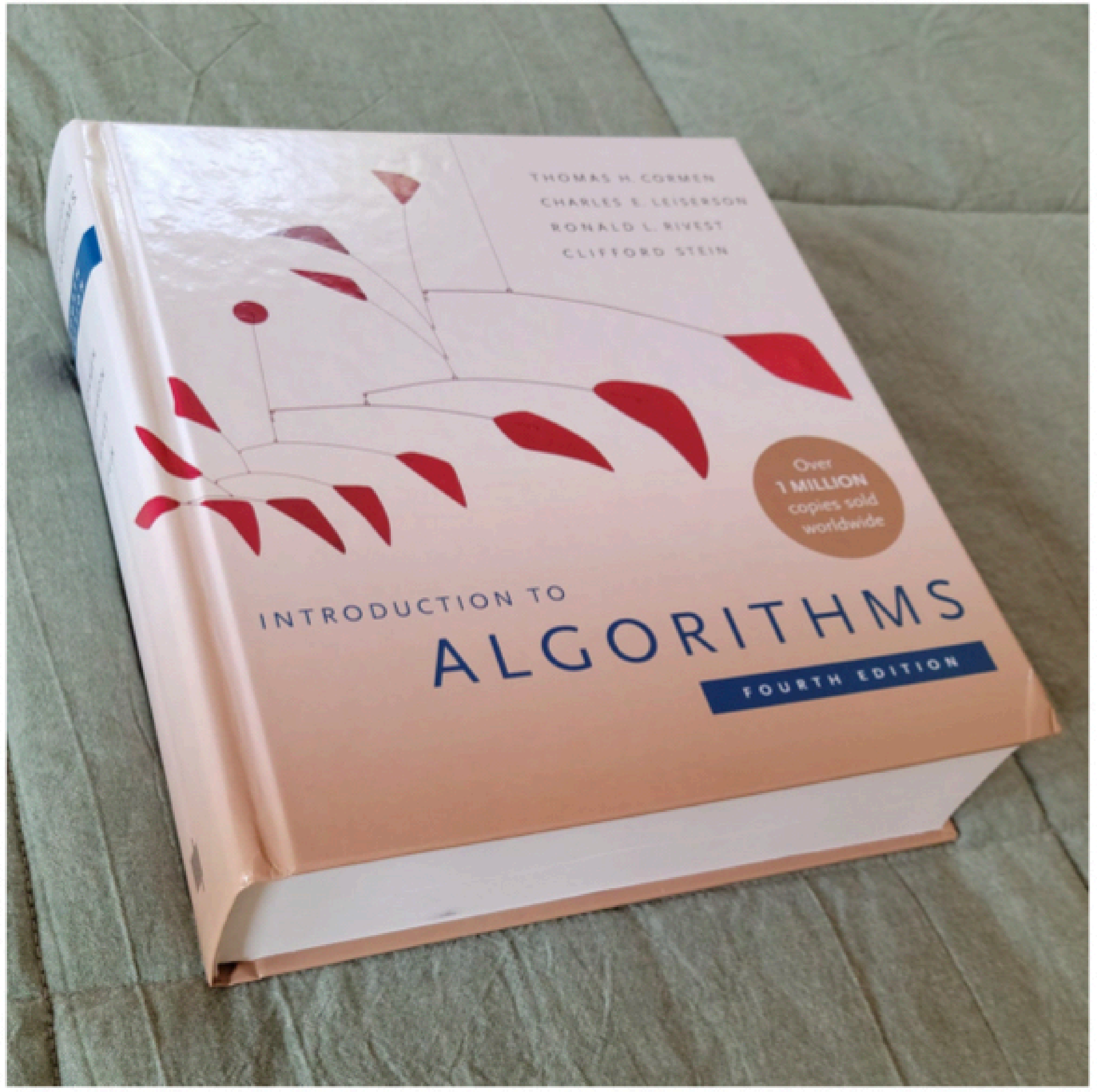
G. EFT/RG for deep neural network (if time permits)

topics not to be introduced: streaming algorithms, graphical models (interval scheduling, shortest path, maximum-flow and minimum-cut, random graph, phase transition), quantum FFT, …

References: no required textbooks

A few useful books for further readings:

- Mackay: *Information theory, Inference and Learning Algorithms*, 2003
- Kleinberg&Tardos: *Algorithm Design*, 2005
- Bishop: *Pattern Recognition and Machine Learning*, 2006
- Mitzenmacher&Upfal: *Probability and Computing*, 2017
- Vershynin: *High-dimensional Probability*, 2018
- Blum&Hopcroft&Kannan: *Foundations of Data Science*, 2020
- Skiena: *The Algorithm Design Manual*, 2020
- Bottcher&Herrmann: *Computational Statistical Physics*, 2021
- CLRS: *Introduction to Algorithms*, 2022
- Prince: *Understanding Deep Learning*, 2023



If you want to learn from them, do exercises!

Assignments/homework

- Theoretical exercises

  derivation, estimate, proof, …

- Programming exercises

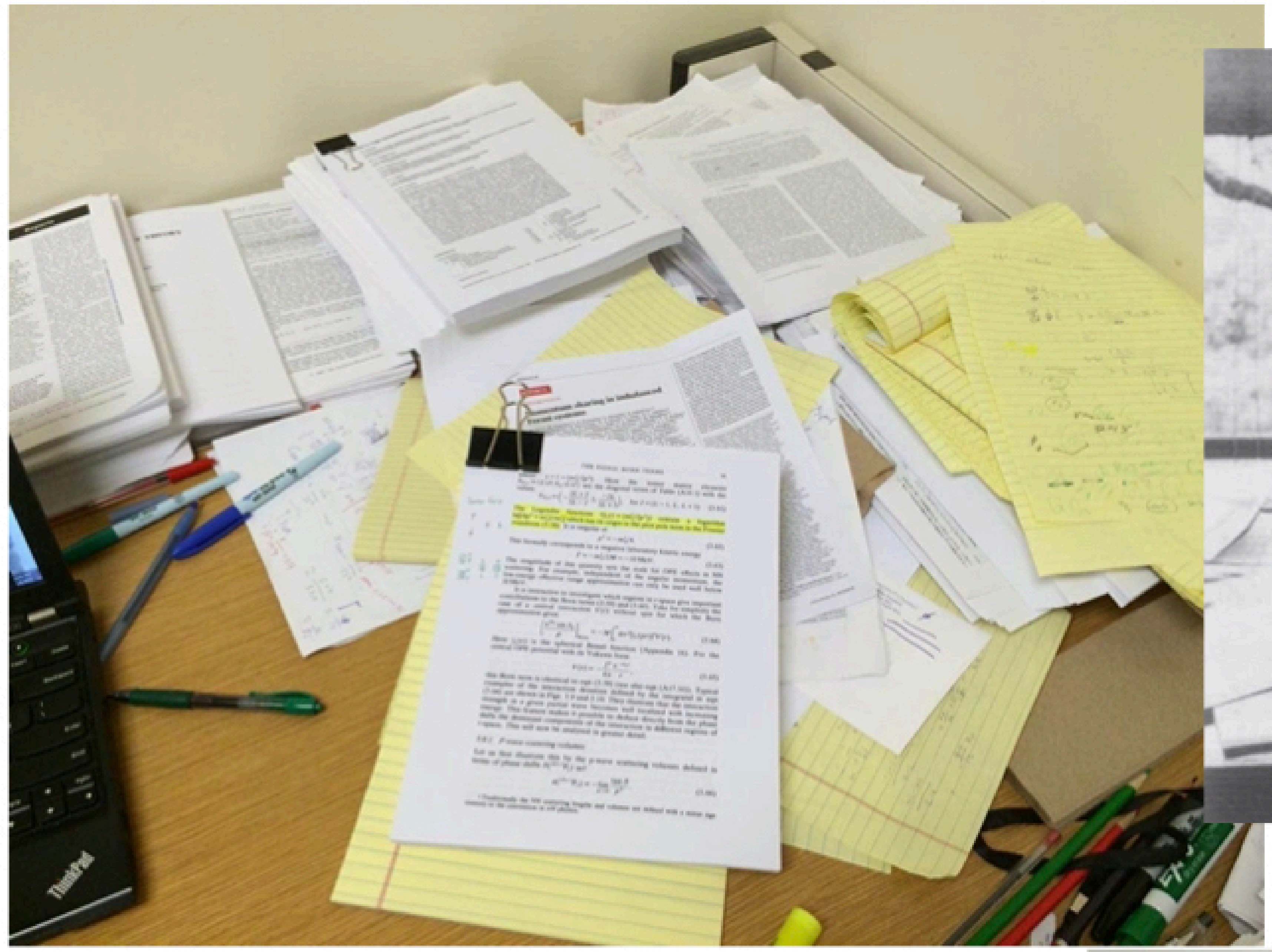
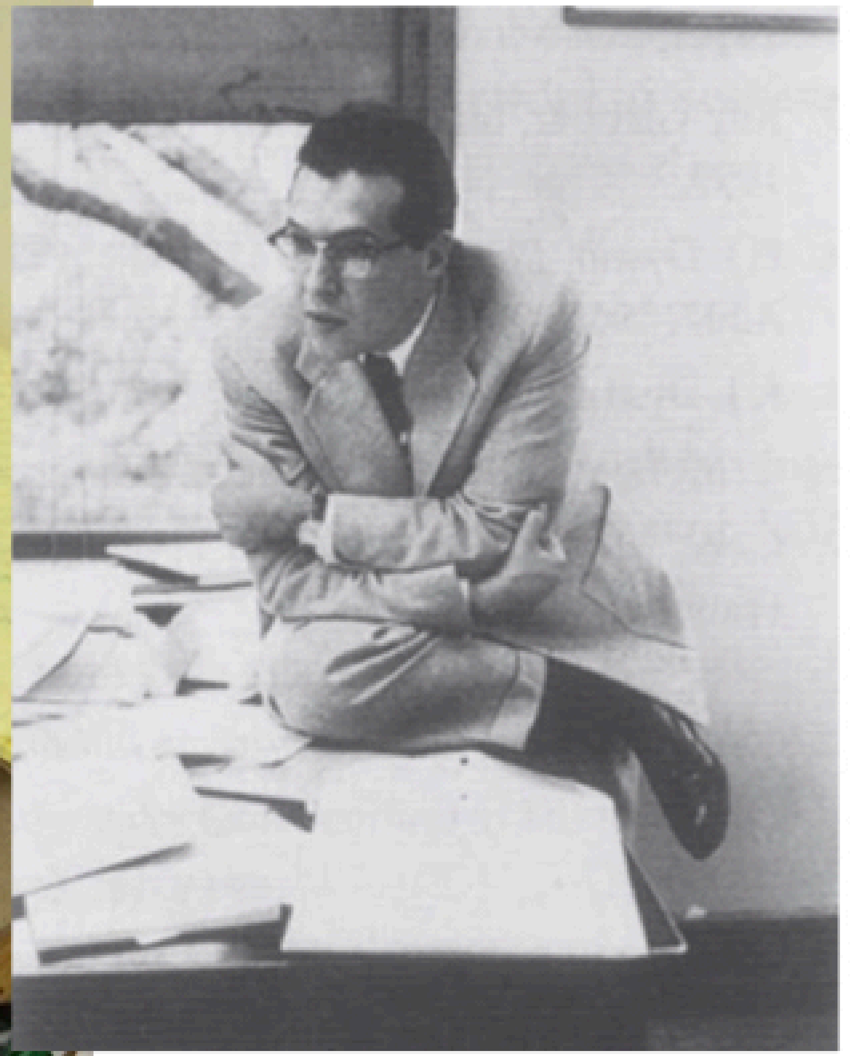
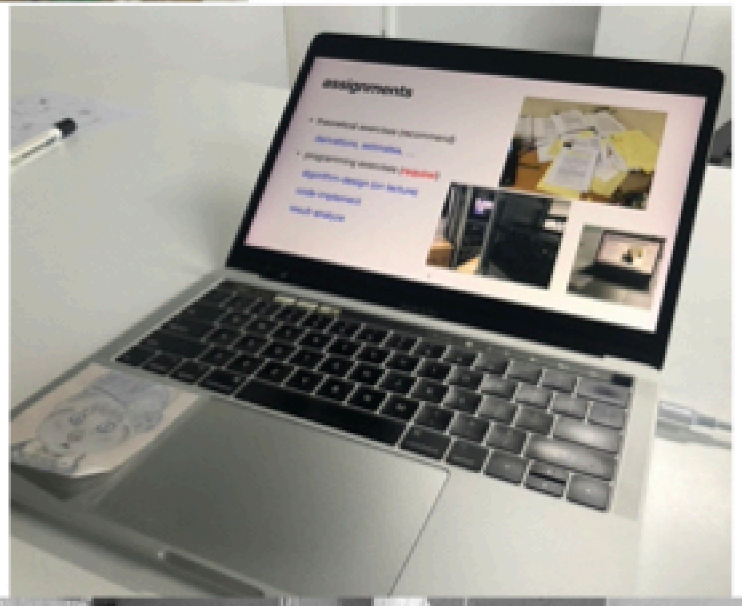
  algorithm-designing

  code-implementing

  result-analyzing

  suggestion: Python/Julia

- Send to bjcai@fudan.edu.cn before the next lecture

Grading policy and office hour

- Homework: 30%, exclude module F

  per week (or per lecture)

  3-5 problems, 1-2 computational/programming

- Quiz: 30%, exclude module F

  conceptual understanding, quick estimate

  around 10-20 minutes for each

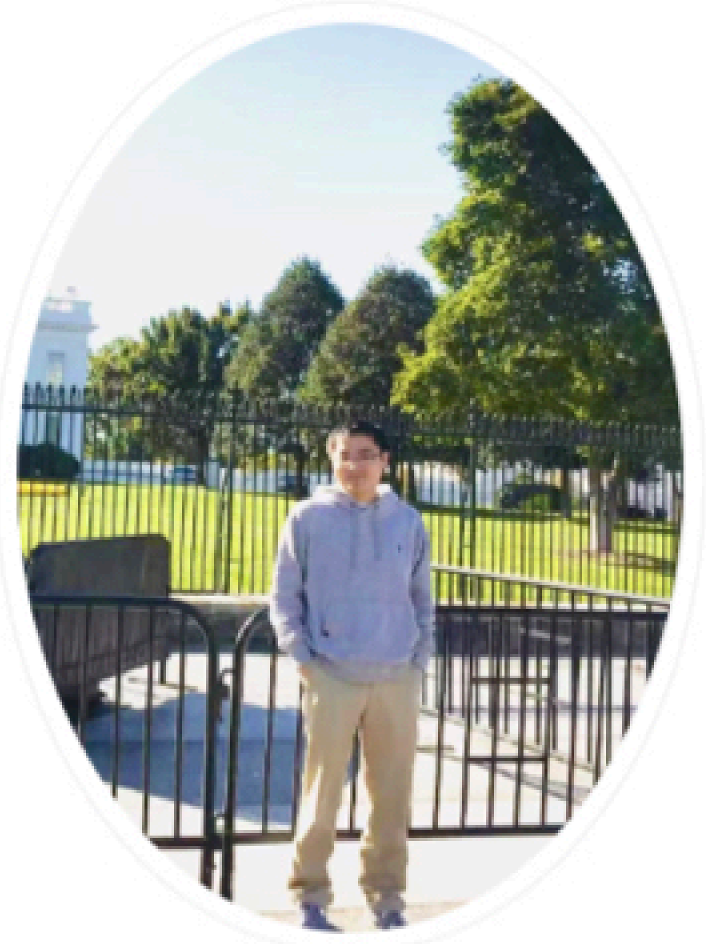
  irregular, total 8-12 times

- Final Exam: 40%, exclude module F

Jiangwan Campus
inter-discipline building 1
B5004
bjcai@fudan.edu.cn

Course website https://bjcai-phys.github.io/algo2026/

Bao-Jun Cai's Website Research Publications Talks Algorithms-2026 Lecture Notes Journal Club CV



Bao-Jun Cai

Dense Matter Equation of State, Short Range Correlations, Physics of Neutron Stars, Statistical Learning Algorithms

Shanghai

Fudan University

Introduction to Algorithms for Data Science and Physics

Institute of Modern Physics, Fudan University, Spring/2026

Syllabus

Course De
algorithms
and moder
principles
basic prob
essential i
with the ra
of contem
statistical

Lecture Notes

There will be no designated textbook for the course, lecture notes will be provided progressively as the course develops. Several reference books that offer broad coverage of relevant topics may be helpful for those seeking a deeper understanding and are provided in the syllabus. You do not need to read everything.

- Lecture 1: Order of Magnitude, Estimate and Divide-and-Conquer

Homework

- Assignment 1: submit before or on 3/10/2026

Lecture 1

Order of Magnitude, Estimate and Divide-and-Conquer

Bao-Jun Cai, 3/4/2026

Introduction to Algorithms for Data Science and Physics IMP@Fudan, 2026

Topics of this lecture:

- order of magnitude $W = \omega \times 10^{\delta}$
- perturbative calculations $f \approx f_0 + f_1 + f_2 + \cdots$
- period of a simple pendulum $T = 2\pi\sqrt{\ell/g} \cdot (1 + \cdots)$
- algebraic equation: $x^n(t) = \Omega + tx(t)/\Lambda$
- divide-and-conquer, computational complexity
- examples: Fibonacci numbers, matrix multiplication

Estimate of a quantity

guess, order of magnitude

$$W = \omega \times 10^{\delta}$$

sign of $W \leftrightarrow$ positive or negative?

value of $\delta \leftrightarrow$ rough estimate

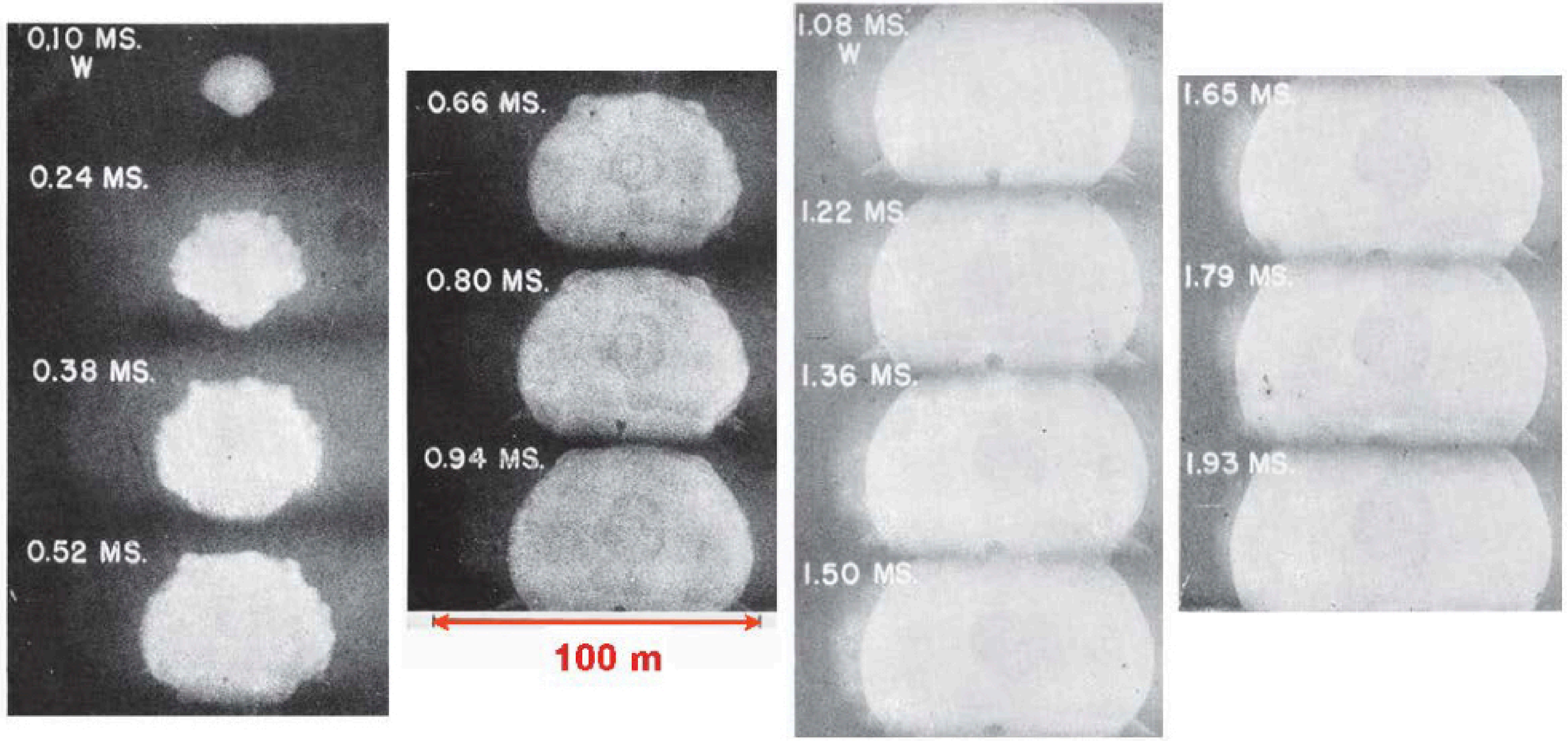
value of $\omega \leftrightarrow$ complete theory

quantity dimensions $\rightarrow$ dimensionless

The formation of a blast wave by a very intense explosion
I. Theoretical discussion

By Sir Geoffrey Taylor, F.R.S.

(*Received 6 October* 1949)



$$R \sim t^{\mu}, \mu = ?$$

Example: fate of false vacuum

$$\Gamma/V \approx Ae^{-B/\hbar}[1 + \mathcal{O}(\hbar)]$$

PHYSICAL REVIEW D VOLUME 15, NUMBER 10 15 MAY 1977

Fate of the false vacuum: Semiclassical theory*

Sidney Coleman
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 24 January 1977)

It is possible for a classical field theory to have two homogeneous stable equilibrium states with different energy densities. In the quantum version of the theory, the state of higher energy density becomes unstable through barrier penetration; it is a false vacuum. This is the first of two papers developing the qualitative and quantitative semiclassical theory of the decay of such a false vacuum for theories of a single scalar field with nonderivative interactions. In the limit of vanishing energy density between the two ground states, it is possible to obtain explicit expressions for the relevant quantities to leading order in h; in the more general case, the problem can be reduced to solving a single nonlinear ordinary differential equation.

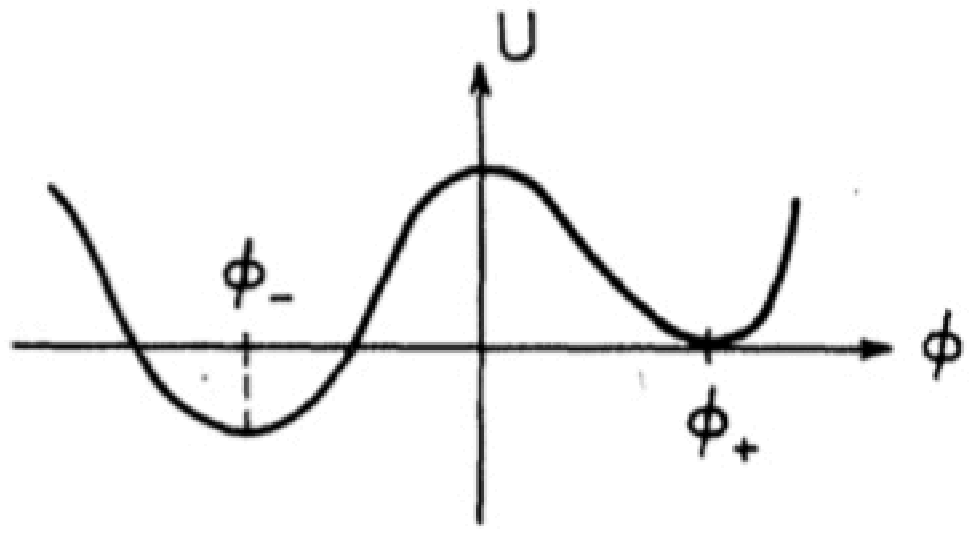


FIG. 1. The nonderivative part of the Lagrangian, $U(\phi)$, for a theory with a false vacuum.

PHYSICAL REVIEW D VOLUME 16, NUMBER 6 15 SEPTEMBER 1977

Fate of the false vacuum. II. First quantum corrections*

Curtis G. Callan, Jr.
Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08540

Sidney Coleman
Lyman Laboratory, Department of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 15 June 1977)

It is possible for a classical field theory to have two homogeneous stable equilibrium states with different energy densities. In the quantum version of the theory, the state of higher energy density becomes unstable through barrier penetration. In the first paper in this series, it was argued that the relevant quantity to study was a decay probability per unit time per unit volume, $\Gamma/V = Ae^{-B/\hbar}[1 + O(\hbar)]$, and the theory of the coefficient B was given. This paper gives the theory of the coefficient A.

$$B = S_E = \int d\tau d\vec{x} \left[\frac{1}{2}\left(\frac{\partial\phi}{\partial\tau}\right)^2 + \frac{1}{2}(\vec{\nabla}\phi)^2 + U(\phi) \right]$$

$$\Gamma/V = \frac{B^2}{4\pi^2\hbar^2} e^{-B/\hbar} \left| \frac{\det'[-\partial^2 + U''(\phi)]}{\det[-\partial^2 + U''(\phi_+)]} \right|^{-1/2}$$

Energy released from atomic bomb: dimensional analysis

energy $E \sim \mathrm{J} \sim \mathrm{kg} \cdot \mathrm{m}^2/\mathrm{s}^2$

gas density $\rho \sim \mathrm{kg/m}^3$

radius $R \sim \mathrm{m}$

time $t \sim \mathrm{s}$

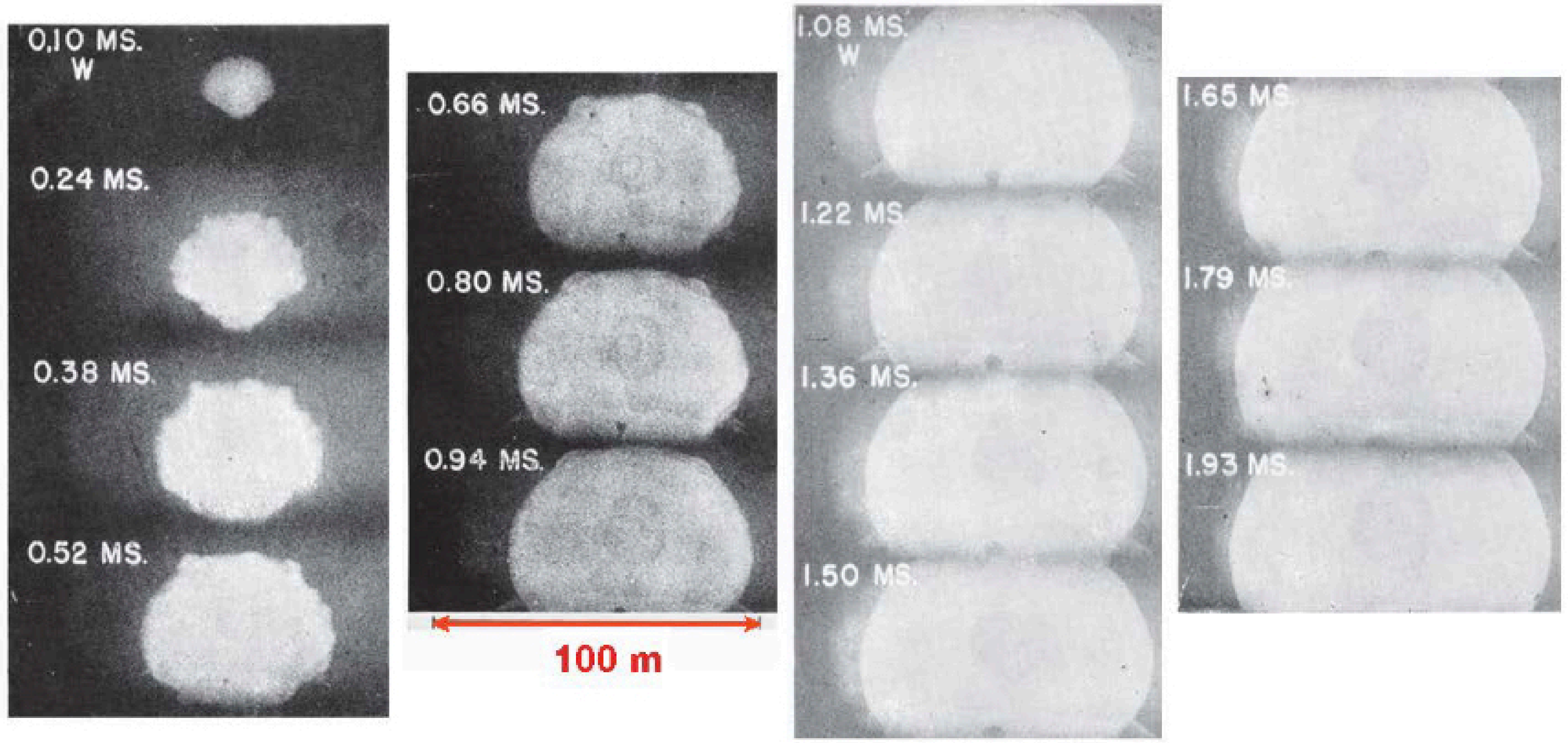
specific heat $\gamma \sim$ dimensionless

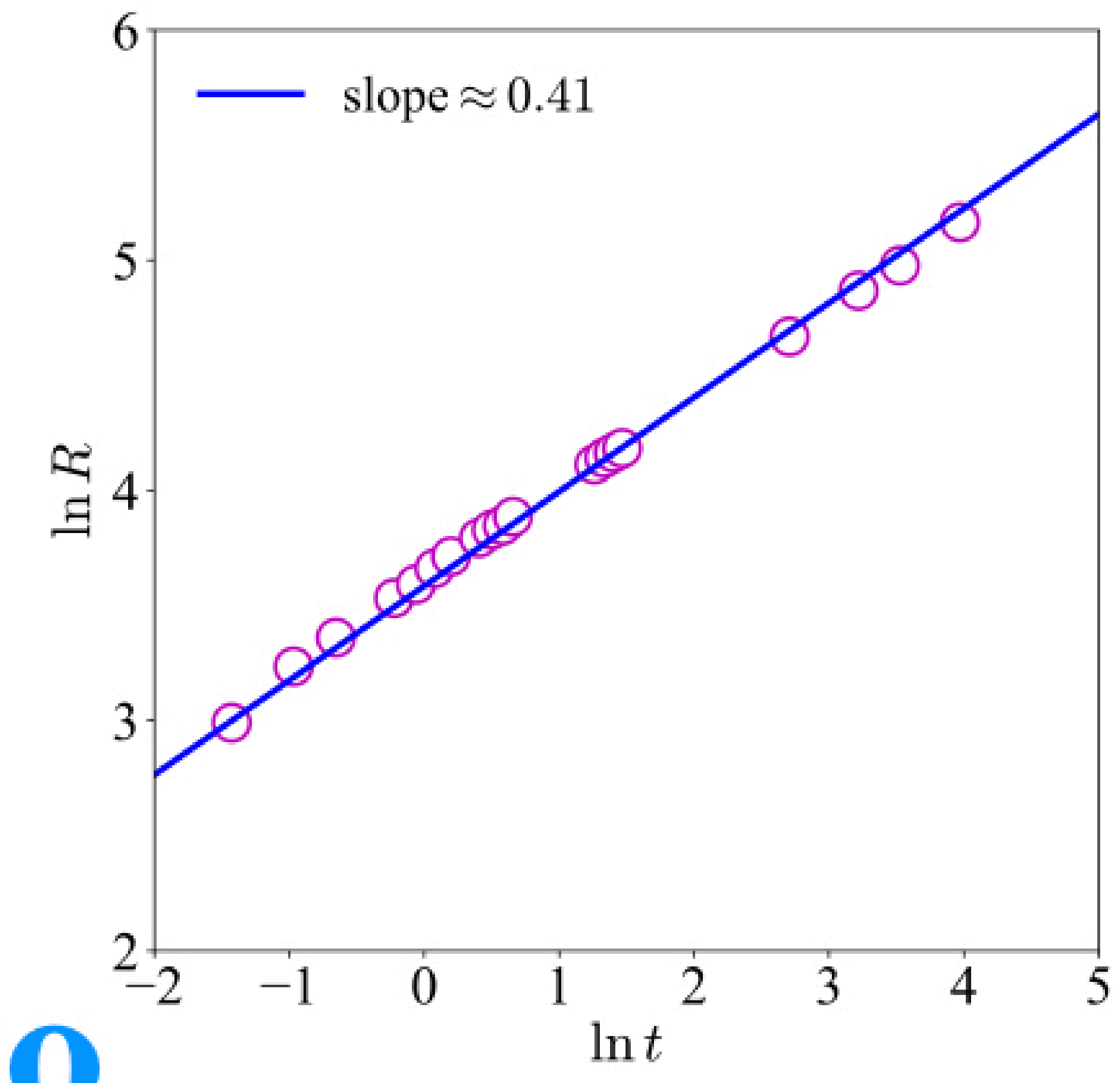
dimensionless $\Pi = E^a \rho^b R^c t^d \sim \mathrm{kg}^{a+b} \cdot \mathrm{m}^{2a-3b+c} \cdot \mathrm{s}^{-2a+d}$

$$\rightarrow b = -a,\ c = -5a,\ d = 2a \rightarrow \Pi \sim \left(\frac{Et^2}{\rho R^5}\right)^a \rightarrow \frac{Et^2}{\rho R^5} \sim \text{dimensionless}$$

$$R \sim t^\mu,\ \mu = ?$$

$$R = \left(\frac{E}{\rho}\right)^{1/5} t^{2/5}\Phi(\gamma) \rightarrow \ln R \sim \frac{2}{5}\ln t \rightarrow \mu \approx 0.4$$





$$E = \overbrace{\Phi^{-5}(\gamma)}^{\Psi(\gamma)\sim\mathcal{O}(1)} \frac{\rho R^5}{t^2} \sim \frac{\rho R^5}{t^2}$$

$R^5/t^2 \sim 10^{13\sim14}\mathrm{m}^5/\mathrm{s}^2$

$\rho \sim 1\mathrm{kg/m}^3 \rightarrow E \sim 10^{13\sim14}\mathrm{J}$

Perturbative calculations: $f/f_0 \approx 1 + \cdots$

warm-up (middle school) example: $ax^2 + bx + c = 0$

$$x^* = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{b}{2a}\left(\sqrt{1-k} - 1\right),\ k = \frac{4ac}{b^2} \ll 1$$

$$\approx -\frac{c}{b}\left(1 + \frac{k}{4}\right) = -\frac{c}{b}\left(1 + \frac{ac}{b^2}\right) \to bx + c = 0$$

$$0.0001x^2 + x - 1 = 0$$

$$\hat{x}_{\text{app}} \approx 1$$

$$0.0001\hat{x}_{\text{app}}^2 = 0.0001$$

$$x_{\text{exact}} \approx 0.9999$$

$$\sqrt{1-k} \approx 1 - \frac{1}{2}k - \frac{1}{8}k^2 \to \sqrt{1-k} - 1 \approx -\frac{1}{2}k\left(1 + \frac{k}{4}\right)$$

Taylor's expansion: $f(x) \approx f(x_0) + f'(x_0)\delta x + \frac{1}{2}f''(x_0)\delta x^2 + \frac{1}{6}f'''(x_0)\delta x^3 + \cdots \leftrightarrow \boxed{\sum_{i=0}^{\infty} \frac{1}{i!} f^{(i)}(x_0)\delta x^i}$

Period of a simple pendulum: perturbations

$$T = 2\pi\sqrt{\frac{\ell}{g}} \times \left(1 + a\chi_{\max}^2 + \cdots\right)$$

Why starts from 2nd order?

How to estimate the coefficient a?

$$b\chi_{\max}^4 \to b = ?$$

equation of motion: $K + U +$ const.

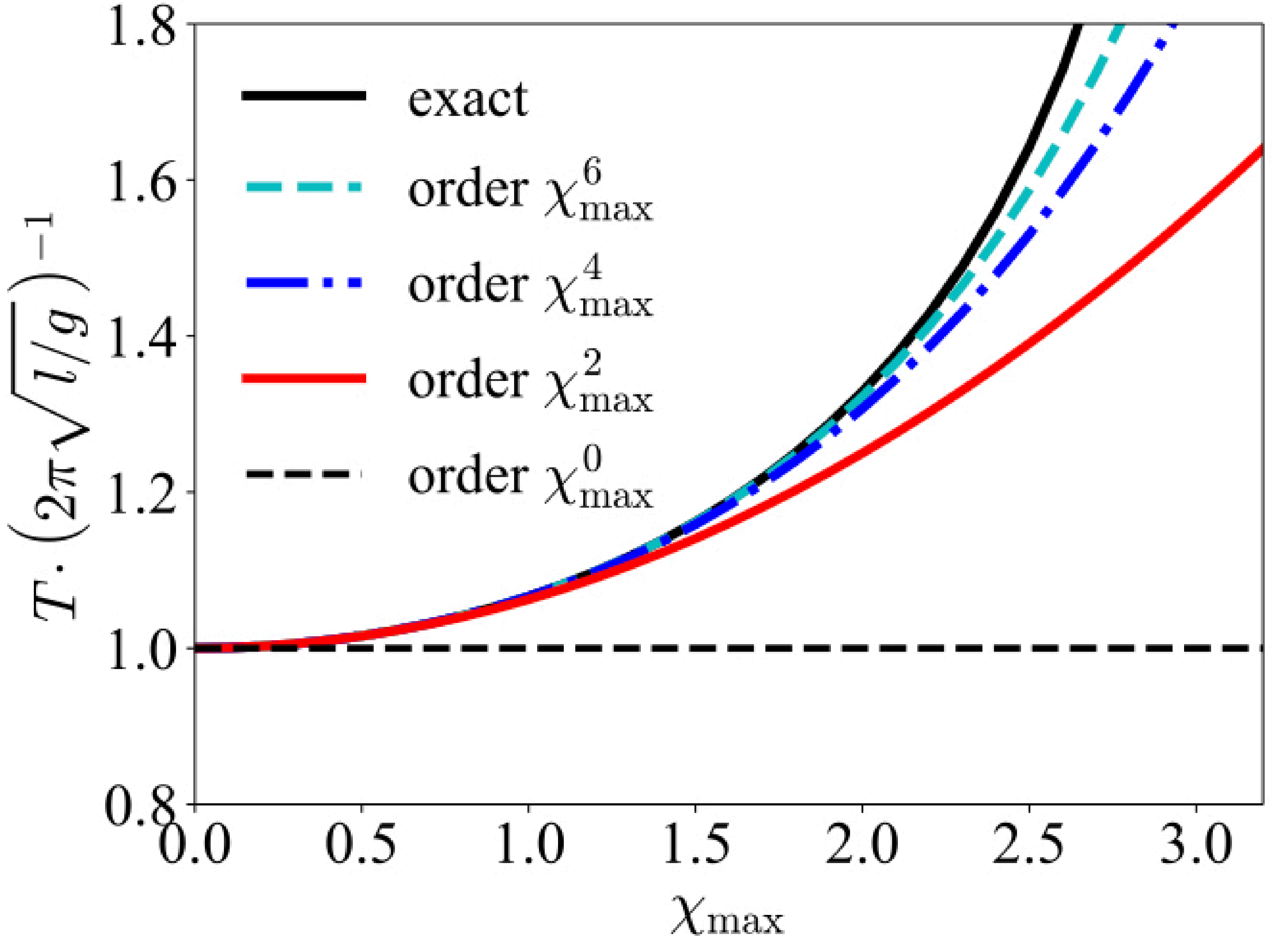
$$K = 2^{-1}mv^2(t),\ U = -mg\ell\cos\chi,\ v(t) = dx/dt,\ x = \chi\ell$$

$$2^{-1}mv^2(t) - mg\ell\cos\chi = -mg\ell\chi_{\max}$$

elliptic integration

$$T = 4\sqrt{\frac{\ell}{g}} \times \int_0^{\pi/2} \frac{d\vartheta}{\sqrt{1 - \sin^2\frac{\chi_{\max}}{2}\sin^2\vartheta}}$$

first-order: $T \approx 4\sqrt{\frac{\ell}{g}}\int_0^{\pi/2} d\vartheta\left(1 + \frac{1}{8}\chi_{\max}^2\sin^2\vartheta\right) = 2\pi\sqrt{\frac{\ell}{g}} \times \left(1 + \frac{1}{16}\chi_{\max}^2\right)$



How to solve if there is no formula?

t: control parameter, e.g., time

$$x^n(t) = \Omega + tx(t)/\Lambda,\ x(t) \in \mathrm{R}^+,\ n \in \mathrm{N}^+,\ n \geq 5,\ t \geq 0,\ \text{e.g., } x^6(t) = 1 + tx(t)$$

small-t limit: perturbative solution

$|\delta(t)| \ll 1 \leftrightarrow t \lesssim t_{\max} \approx sn\Lambda\Omega^{1-1/n},\ s \ll 1$

$$t = 0 \rightarrow x_0 = x(0) = \Omega^{1/n}$$

$$t \approx 0^+ \rightarrow x(t) \approx x_0[1 + \delta(t)] \rightarrow \delta(t) = \Omega^{1/n-1}t/n\Lambda$$

large-t limit: asymptotic solution

small-t: $x(t) \approx x_0[1 + \delta(t) + \alpha\delta^2(t) + \cdots]$

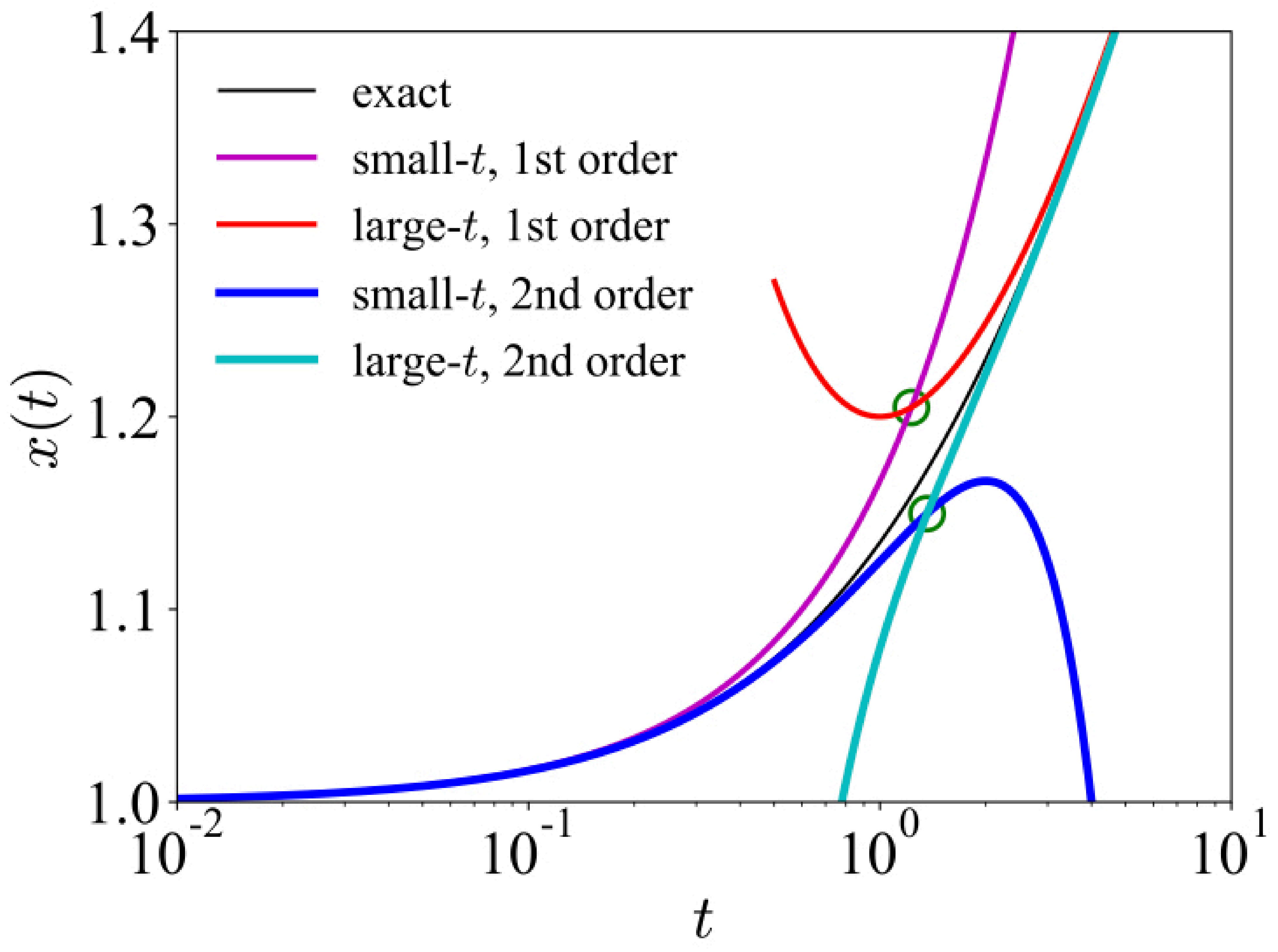
large-t: $x(t) \approx x_\infty[1 + \phi(t) + \beta\phi^2(t) + \cdots]$

$$t = \infty \rightarrow x_\infty = x(\infty) = (t/\Lambda)^{1/(n-1)}$$

$$t < \infty \rightarrow x(t) = x_\infty[1 + \phi(t)] \rightarrow \phi(t) = \frac{\Omega}{n-1}\left(\frac{\Lambda}{t}\right)^{\frac{n}{n-1}}$$

$|\phi(t)| \ll 1 \leftrightarrow t \gg t_{\mathrm{asp}} \equiv \Lambda \exp\left(-\frac{n-1}{n}\ln\frac{n-1}{\Omega}\right)$

Exact solution: numerical method



small-*t*:

$$x(t) \approx \Omega^{1/n}\left(1 + \frac{1}{n\Lambda}\Omega^{1/n-1}t - \frac{n-3}{2n^2\Lambda^2}\Omega^{2/n-2}t^2\right)$$

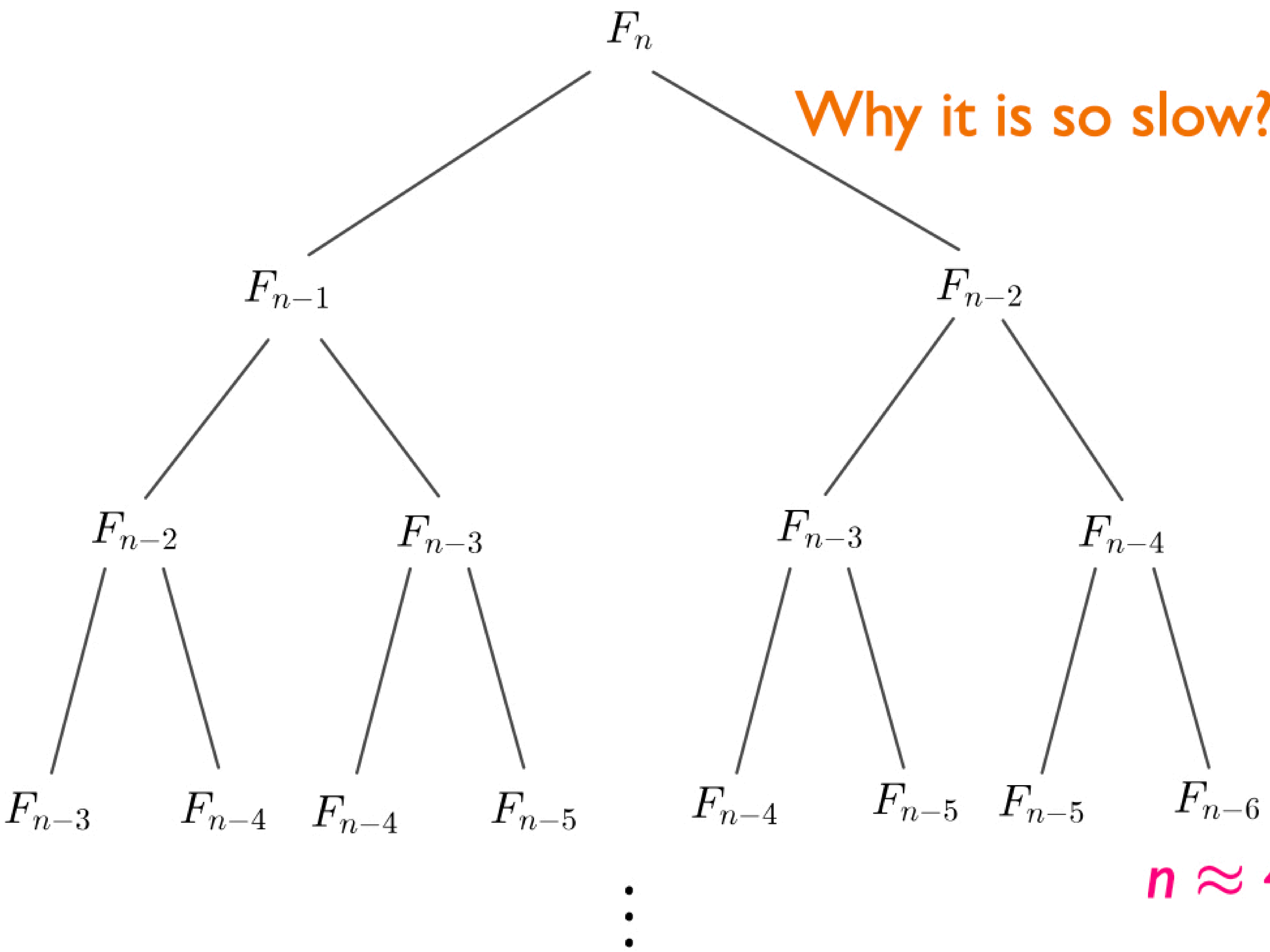
What is the effective range for *t* at this order?

large-*t*:

$$x(t) \approx \left(\frac{t}{\Lambda}\right)^{\frac{1}{n-1}}\left[1 + \frac{\Omega}{n-1}\left(\frac{\Lambda}{t}\right)^{\frac{n}{n-1}} - \frac{n\Omega^2}{2(n-1)^2}\left(\frac{\Lambda}{t}\right)^{\frac{2n}{n-1}}\right]$$

iteration: $x^{(j+1)}(t) = \left(\Omega + \frac{tx^{(j)}(t)}{\Lambda}\right)^{1/n}$, starting from some initial $x^{(0)}(t)$ for a given t

Algorithm and algorithm design: first example



Why it is so slow?

Fibonacci number

$$F_n = F_{n-1} + F_{n-2},\ F_0 = 0, F_1 = 1$$

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

$$\approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n \approx 2^{\log_2(1/\sqrt{5})+0.694n}$$

$n \approx 46 \rightarrow 0.694n \approx 32 \rightarrow$ exponential growth

recursion: $T(n) = T(n-1) + T(n-2) + \text{const.} \rightarrow T(n) \geq F_n,\ T(n) = \text{time/operations}$

Right algorithm makes all the difference

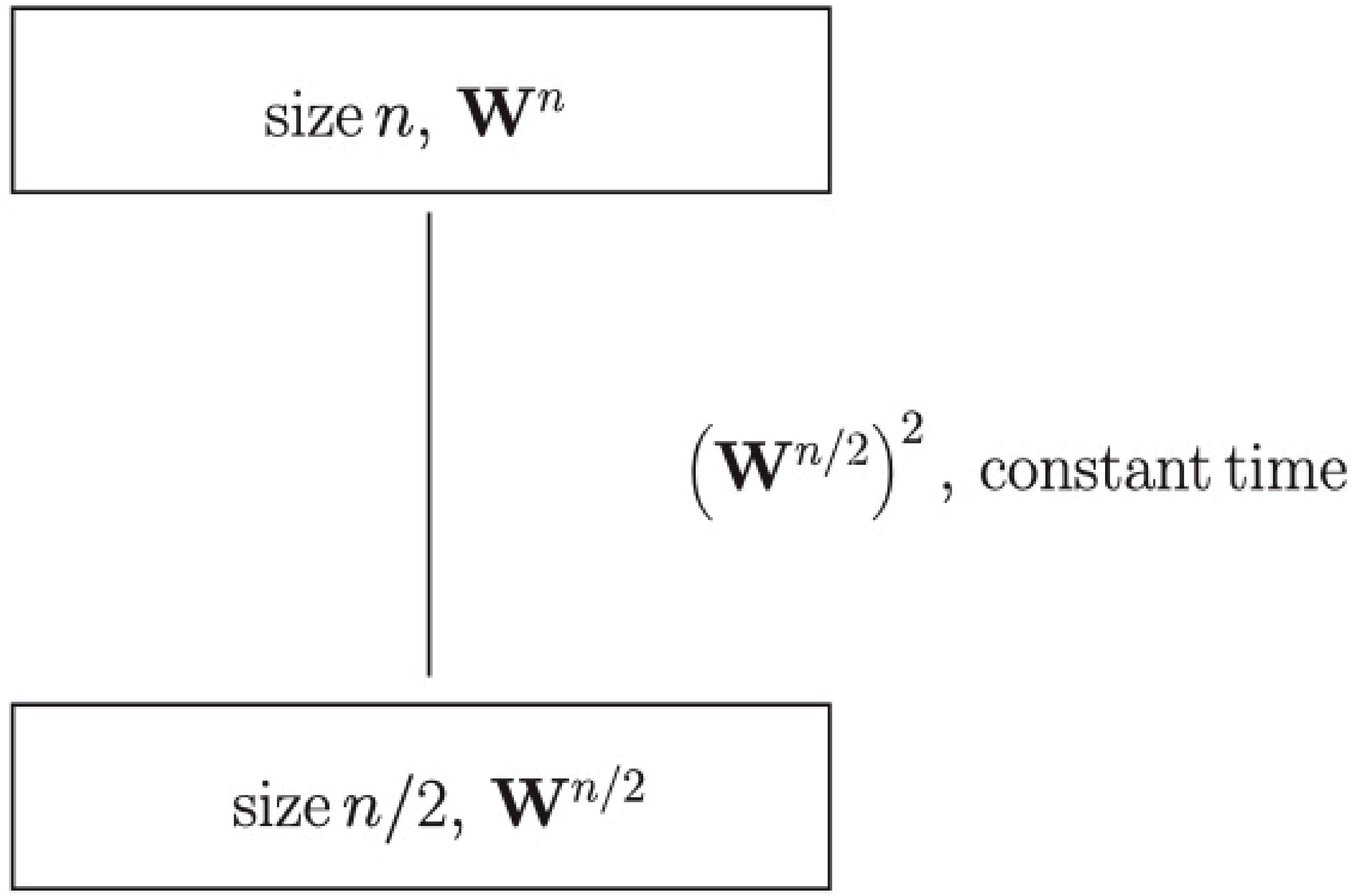
What is the implication?

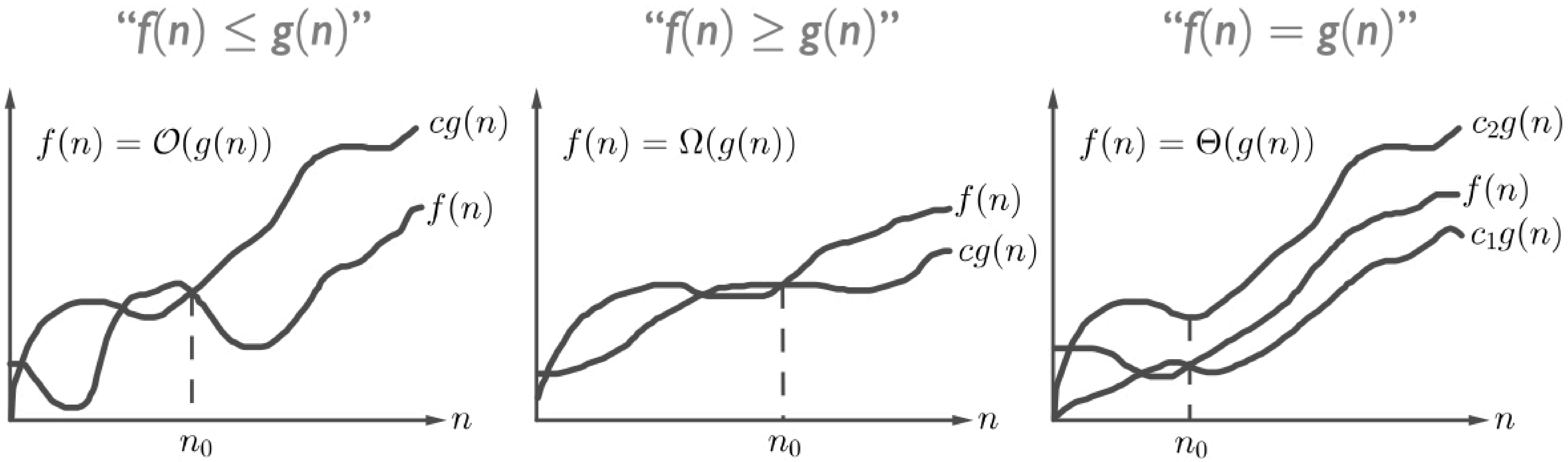
$$\begin{pmatrix} F_n \\ F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_n \end{pmatrix} = \cdots = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} F_0 \\ F_1 \end{pmatrix} = \mathbf{W}^n \begin{pmatrix} F_0 \\ F_1 \end{pmatrix}$$

$\mathcal{O}, \Omega, \Theta$

$$\mathbf{W}^n = (\mathbf{W}^{n/2})^2 = \cdots, \ n = 2^m \rightarrow T(n) = T(n/2) + \mathcal{O}(1)$$

once one knows $\mathbf{W}^{n/2}$ it is straightforward to obtain $\mathbf{W}^n$ in constant time (matrix multiplication)

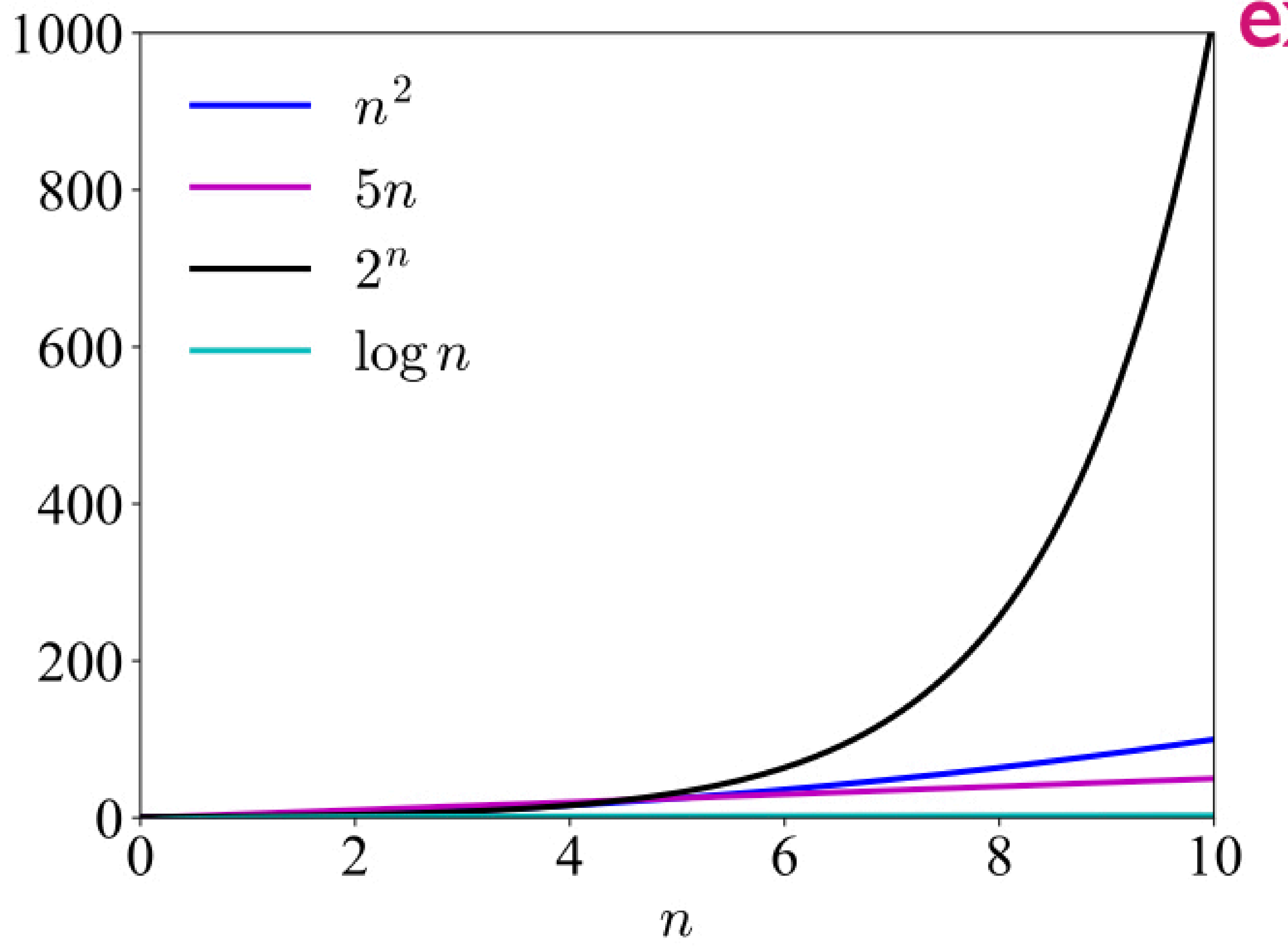




How to solve the recursion T(n)=T(n/2)+O(1)?

Master theorem (divide-and-conquer):

$$T(n) = aT(n/b) + \mathcal{O}(n^d) \rightarrow T(n) = \begin{cases} \mathcal{O}(n^d), & d > \log_b a \\ \mathcal{O}(n^d \log n), & d = \log_b a \\ \mathcal{O}(n^{\log_b a}), & d < \log_b a \end{cases}$$



exponential grows superfast

Fibonacci number:

$$T(n) = T(n/2) + \mathcal{O}(1)$$

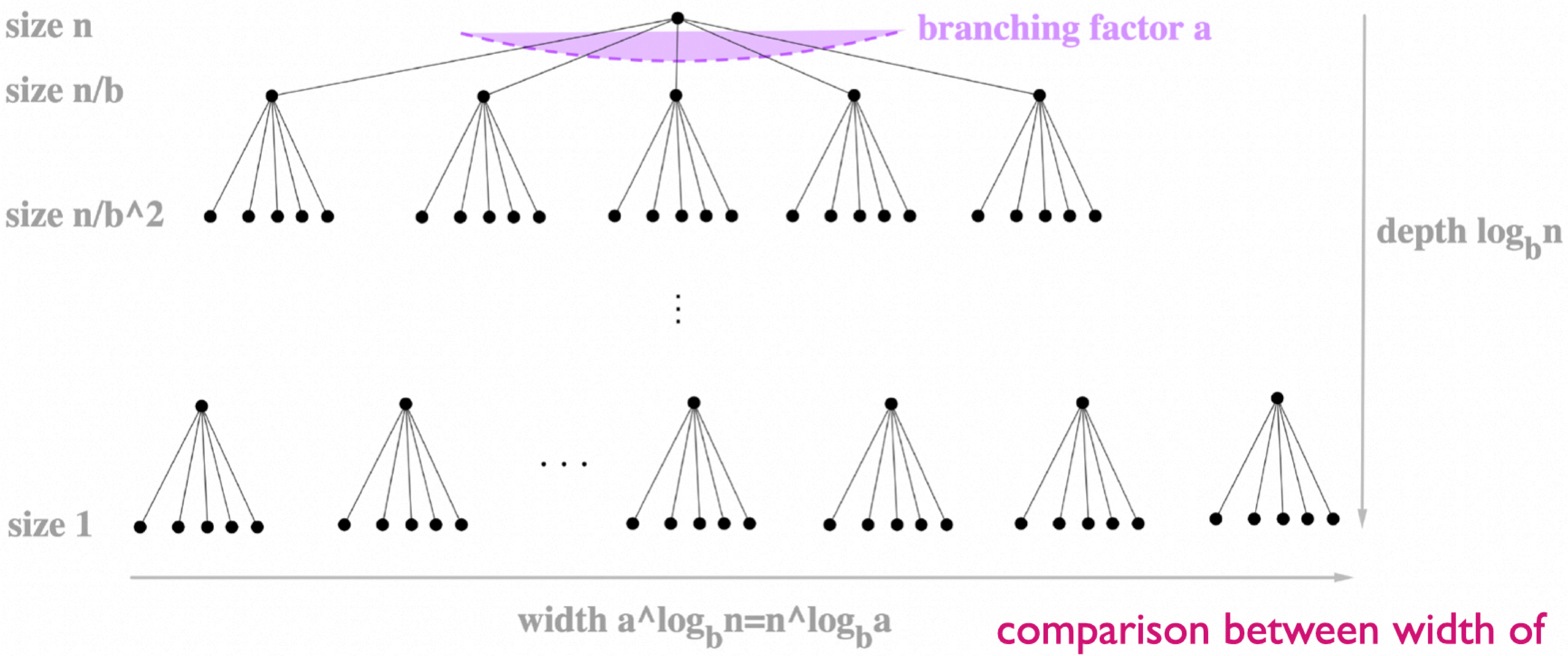
$$a = 1, b = 2, d = 0 \rightarrow d = \log_b a$$

$$\rightarrow T(n) = \mathcal{O}(n^d \log n) = \mathcal{O}(\log n)$$

Intuition behind divide-and-conquer

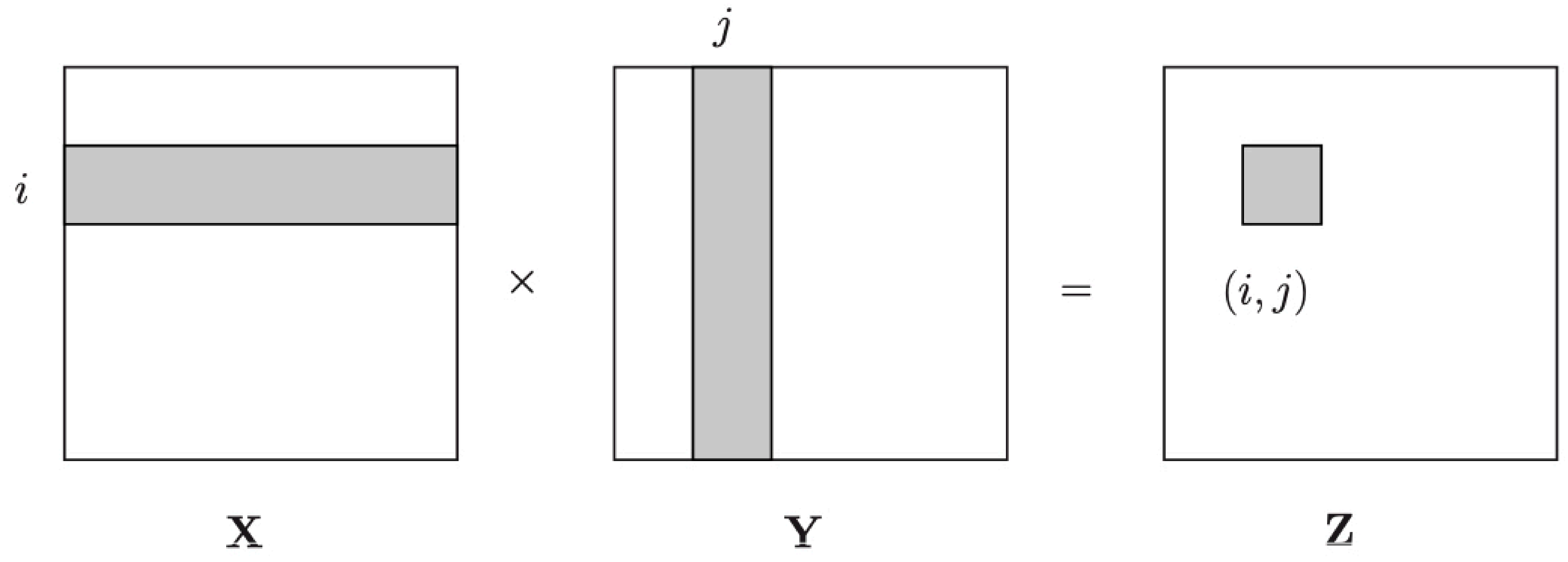
tackle a size-n problem by recursively solving a sub-problems of size n/b and combining them in polynomial time

Master Theorem Construction $T(n) = aT(n/b) + \mathcal{O}(n^d)$



comparison between width of the tree and the adding time

Matrix multiplication: naive divide-and-conquer



$$\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathrm{R}^{n \times n}$$

$$Z_{ij} = \sum_{k=1}^{n} X_{ik} X_{kj} \sim \mathcal{O}(n^3)$$

$$\mathbf{X} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}, \; \mathbf{Y} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} \rightarrow \mathbf{XY} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{pmatrix}$$

$$T(n) = \underbrace{8T(n/2)}_{\text{sub-problems}} + \overbrace{\mathcal{O}(n^2)}^{\text{adding time}} \rightarrow T(n) = \mathcal{O}(n^3)$$

$a = 8, b = d = 2 : \mathcal{O}(n^{\log_b a})$

computational complexity?

* sum: $\mathrm{A} + \mathrm{B}, \; \mathrm{A}, \mathrm{B} \in \mathrm{R}^{n \times n}$
* inner product: $\vec{x} \cdot \vec{y}, \; \vec{x}, \vec{y} \in \mathrm{R}^n$
* matrix inversion: $\mathrm{A}^{-1}, \; \mathrm{A} \in \mathrm{R}^{n \times n}$

Strassen's algorithm (clever rearrangement)

$$\mathbf{XY} = \begin{pmatrix} \mathbf{P}_5 + \mathbf{P}_4 - \mathbf{P}_2 + \mathbf{P}_6 & \mathbf{P}_1 + \mathbf{P}_2 \\ \mathbf{P}_3 + \mathbf{P}_4 & \mathbf{P}_1 + \mathbf{P}_5 - \mathbf{P}_3 - \mathbf{P}_7 \end{pmatrix}$$

$$\mathbf{P}_1 = \mathbf{A}(\mathbf{F} - \mathbf{H})$$
$$\mathbf{P}_2 = (\mathbf{A} + \mathbf{B})\mathbf{H}$$
$$\mathbf{P}_3 = (\mathbf{C} + \mathbf{D})\mathbf{E}$$
$$\mathbf{P}_4 = \mathbf{D}(\mathbf{G} - \mathbf{E})$$
$$\mathbf{P}_5 = (\mathbf{A} + \mathbf{D})(\mathbf{E} + \mathbf{H})$$
$$\mathbf{P}_6 = (\mathbf{B} - \mathbf{D})(\mathbf{G} + \mathbf{H})$$
$$\mathbf{P}_7 = (\mathbf{A} - \mathbf{C})(\mathbf{E} + \mathbf{F})$$

$$\mathrm{X} = \begin{pmatrix} \mathrm{A} & \mathrm{B} \\ \mathrm{C} & \mathrm{D} \end{pmatrix}, \ \mathrm{Y} = \begin{pmatrix} \mathrm{E} & \mathrm{F} \\ \mathrm{G} & \mathrm{H} \end{pmatrix}$$

$a = 7, b = d = 2 : \mathcal{O}(n^{\log_b a})$

$$T(n) = \underbrace{7T(n/2)}_{\text{sub-problems}} + \overbrace{\mathcal{O}(n^2)}^{\text{adding time}} \to T(n) = \mathcal{O}(n^{\log_2 7}) \approx \mathcal{O}(n^{2.81})$$

Coppersmith-Winograd algorithm:

$$\sim \mathcal{O}(n^{2.37})$$

example:

$n = 10^5$ (typical neural networks)

acceleration: $n^{3-2.37} \approx 10^{3.15} \approx 1400$

Even better?

PRL **103,** 150502 (2009) PHYSICAL REVIEW LETTERS week ending 9 OCTOBER 2009

Quantum Algorithm for Linear Systems of Equations

Aram W. Harrow,[1] Avinatan Hassidim,[2] and Seth Lloyd[3]

[1]*Department of Mathematics, University of Bristol, Bristol, BS8 1TW, United Kingdom*
[2]*Research Laboratory for Electronics, MIT, Cambridge, Massachusetts 02139, USA*
[3]*Research Laboratory for Electronics and Department of Mechanical Engineering, MIT, Cambridge, Massachusetts 02139, USA*

(Received 5 July 2009; published 7 October 2009)

Solving linear systems of equations is a common problem that arises both on its own and as a subroutine in more complex problems: given a matrix A and a vector $\vec{b}$, find a vector $\vec{x}$ such that $A\vec{x} = \vec{b}$. We consider the case where one does not need to know the solution $\vec{x}$ itself, but rather an approximation of the expectation value of some operator associated with $\vec{x}$, e.g., $\vec{x}^\dagger M \vec{x}$ for some matrix M. In this case, when A is sparse, $N \times N$ and has condition number κ, the fastest known classical algorithms can find $\vec{x}$ and estimate $\vec{x}^\dagger M \vec{x}$ in time scaling roughly as $N\sqrt{\kappa}$. Here, we exhibit a quantum algorithm for estimating $\vec{x}^\dagger M \vec{x}$ whose runtime is a polynomial of $\log(N)$ and κ. Indeed, for small values of κ [i.e., poly $\log(N)$], we prove (using some common complexity-theoretic assumptions) that any classical algorithm for this problem generically requires exponentially more time than our quantum algorithm.

DOI: 10.1103/PhysRevLett.103.150502 PACS numbers: 03.67.Ac, 02.10.Ud, 89.70.Eg

HHL algorithm: $T(n) \sim \mathcal{O}(\text{poly}(\log n, \kappa, \epsilon^{-1}))$

ϵ: precision (loss of accuracy→speedup)

best classical solvers:

$\mathbf{Ax} = \mathbf{b} \sim \mathcal{O}(n\sqrt{\kappa})$, sparse $\mathbf{A} \in \mathbb{R}^{n \times n}$

condition number: $\kappa = |\lambda|_{\max}/|\lambda|_{\min}$

matrix-based algorithms lie at the heart of modern algorithm design