Computer Simulation – setting the scene

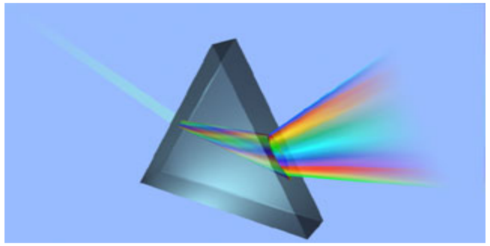
Professor Arthur Trew
Director, EPCC
a.s.trew@ed.ac.uk
+44 131 650 5025

science … the traditional view

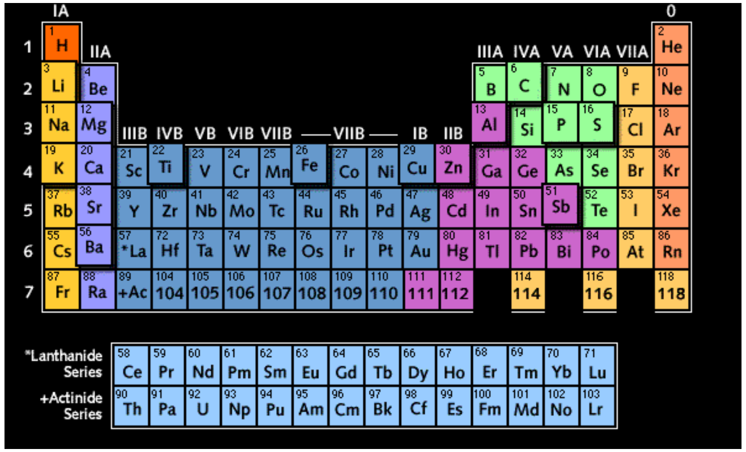
- *Theory*
  - Greece 500 BC

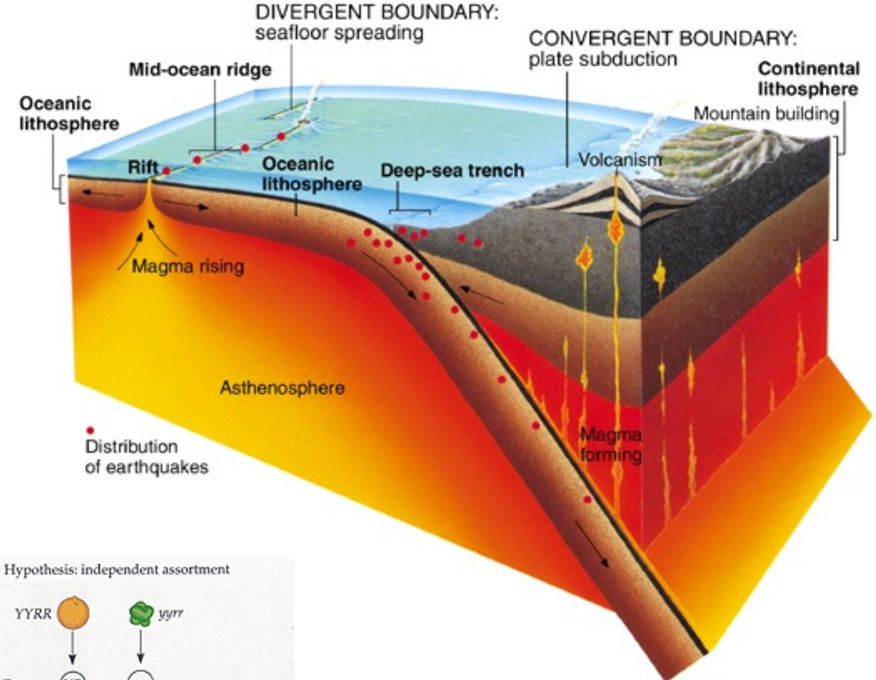
- *$F = ma$*
- *$E = mc^2$*

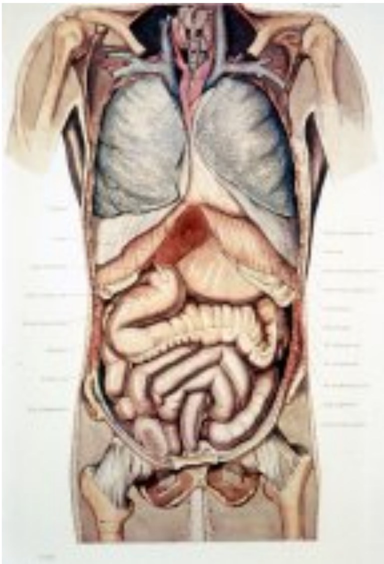
- *Experiment*
  - Europe 1600 AD

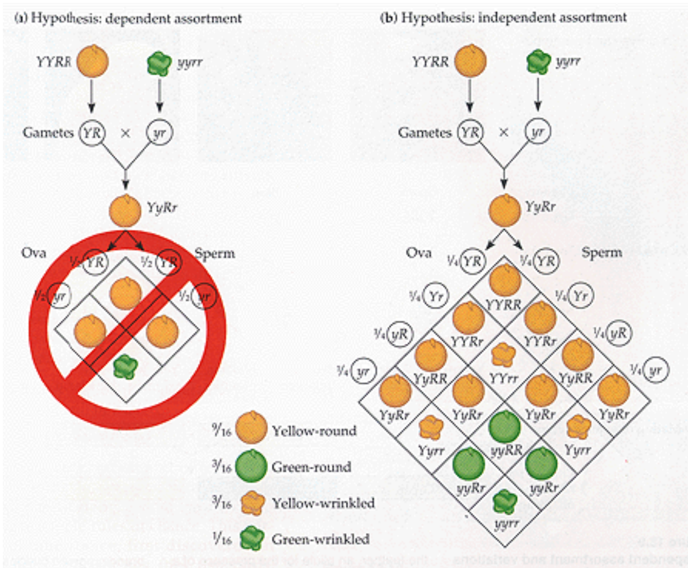


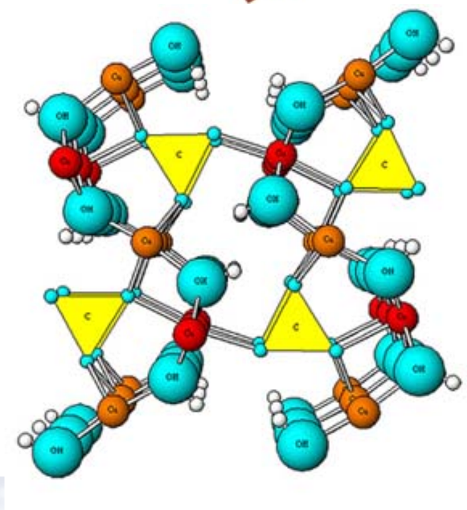
400 years of success



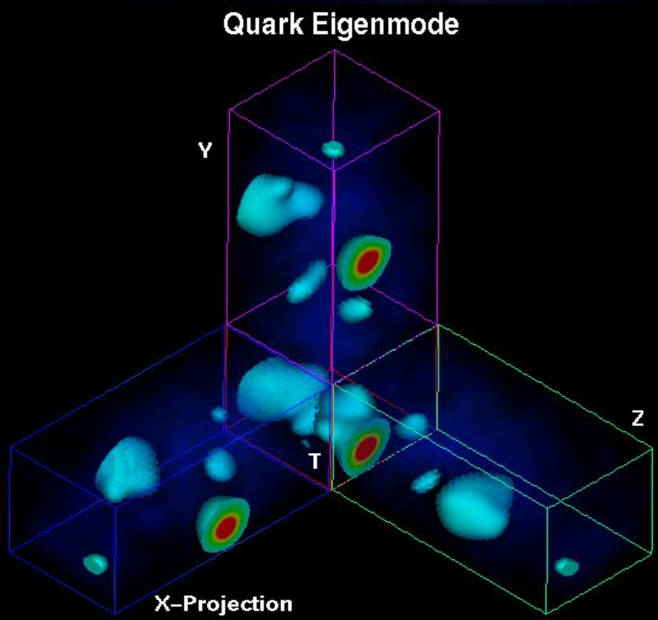









NASA


Quark Eigenmode
Y
Z
T
X–Projection


www.saabnet.com

science … the newcomer

•*Simulation*

•Edinburgh 1980 AD

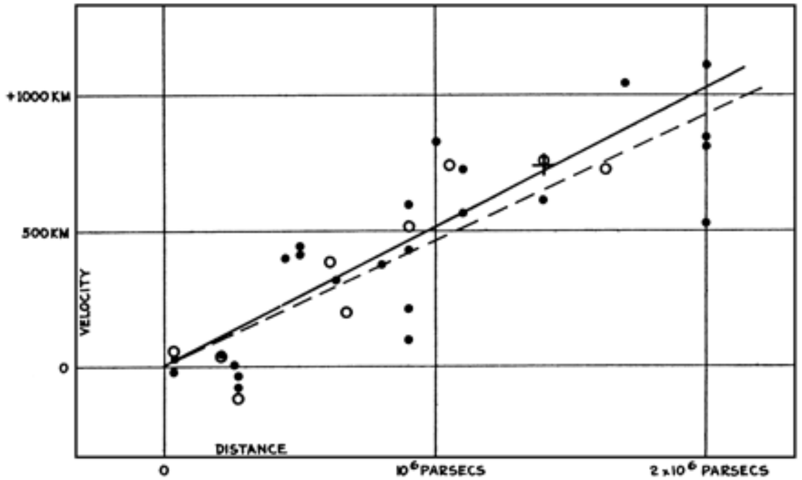
```
prob = exp(-beta*(newenergy - oldenergy))

if (drand(seed) .lt. prob) then

  if (newangle .lt. 0.0    ) newangle = newangle + 2.0*pi
  if (newangle .ge. 2.0*pi) newangle = newangle - 2.0*pi

  angle(ix, iy) = newangle

  dx(ix,iy) = cos(angle(ix,iy))
  dy(ix,iy) = sin(angle(ix,iy))
  col(ix,iy) = cos(0.5*angle(ix,iy))**2

  accept = accept + 1

end if
```

the Big Bang

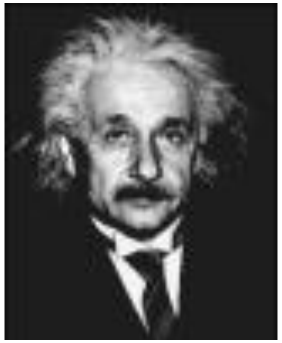
- 200 years ago
  - catalogues of nebulae (clouds) were produced to avoid people mistaking them for comets
- 100 years ago
  - recognised that some nebulae were in our own galaxy and others were whole galaxies

Hubble and the Redshift problem

- **1929**: the Hubble Diagram



- **1915**: Einstein had actually predicted the expansion of the Universe 15 years earlier – but refused to believe his own equations
  - "it was my greatest blunder"



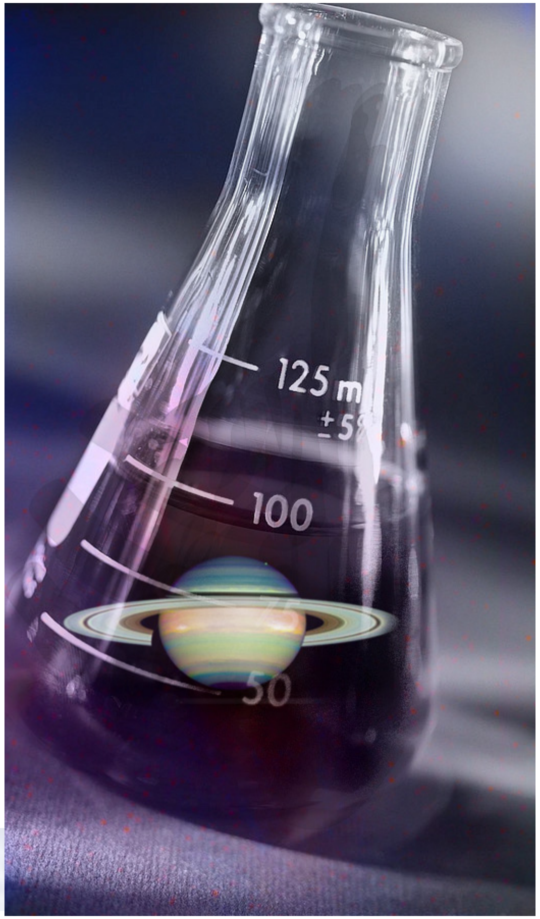
– but this means that we have an equation with which to study the Universe

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}$$

looks simple, but isn't

$$G_{\mu\nu} = 8\pi T_{\mu\nu}$$

- this may look like one equation, but actually it's 10!
- … and they are so complex that Einstein thought they would never be solved
- in fact, they can be solved mathematically – but only in a few few simple cases
- … a job for the computer!
- because, clearly, *we can't experiment on the Universe!*

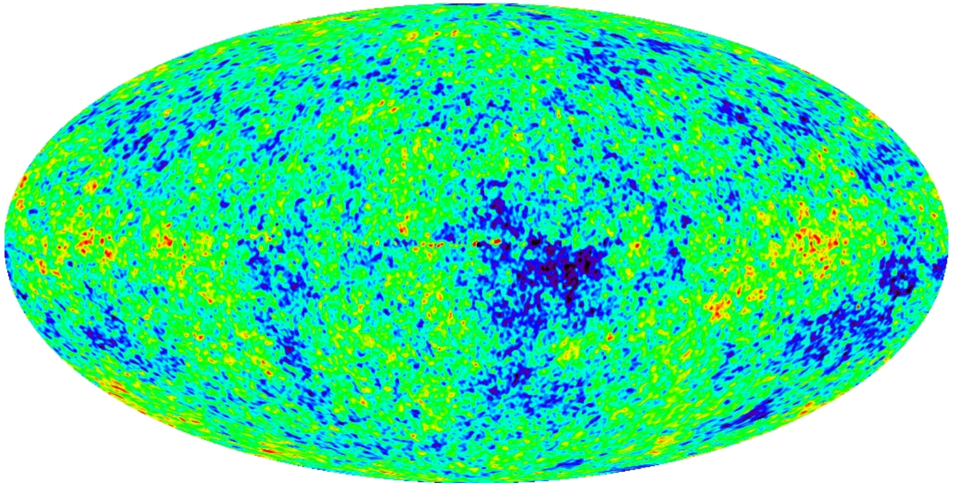


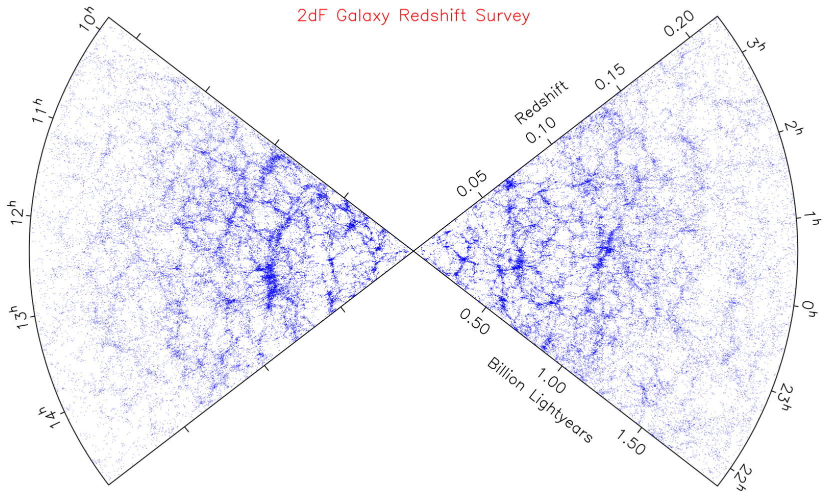
what's the problem?

- actually, there are two (related) problems:
  - 1: what is the Universe made of?
  - 2: and what structure does it have?

- we can only find <10% of the mass
- and theories of the Big Bang say that the missing mass (dark matter) is **not** matter as we know it
- … and now we also have dark energy!!!

structure from (almost) nothing





- can so much structure be created from so smooth a start?
- … well that depends on the how much matter is in the Universe … and what it is made of
- while the acceleration in expansion is due to dark energy
  - is this Einstein's "greatest blunder" in real life (ie $\Lambda \neq 0$)?

- similar approaches can be used, for example, to predict the weather, or model the climate

$$\frac{\partial \rho}{\partial t} + \nabla . \rho \mathbf{u} = 0$$

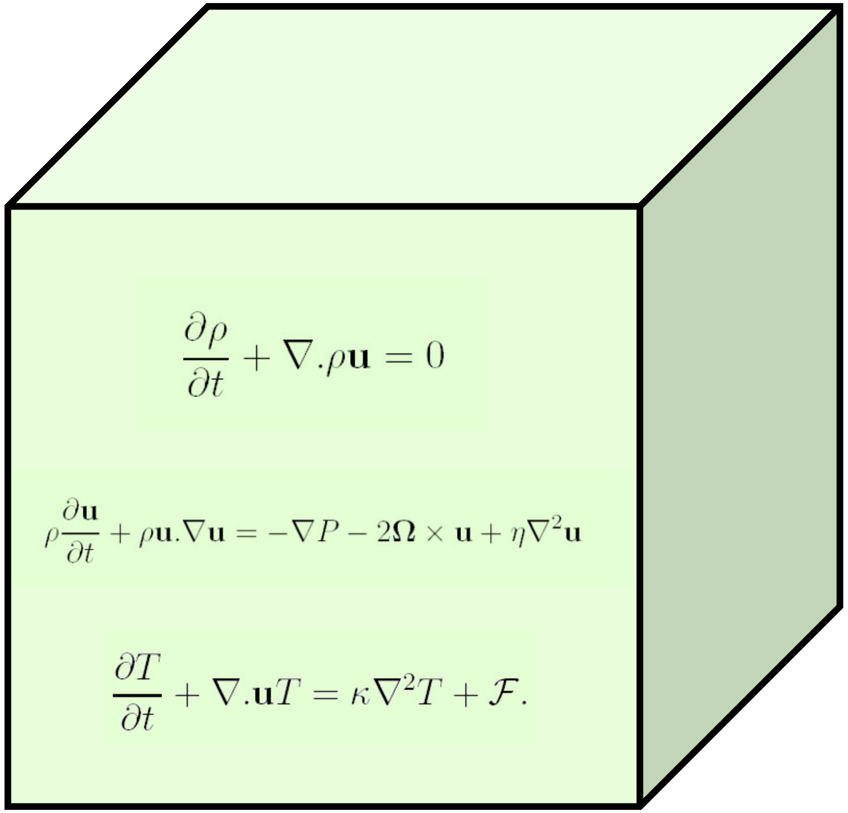
conservation of mass

$$\rho \frac{\partial \mathbf{u}}{\partial t} + \rho \mathbf{u} . \nabla \mathbf{u} = - \nabla P - 2 \mathbf{\Omega} \times \mathbf{u} + \eta \nabla^2 \mathbf{u}$$

Navier Stokes

$$\frac{\partial T}{\partial t} + \nabla . \mathbf{u} T = \kappa \nabla^2 T + \mathcal{F} .$$

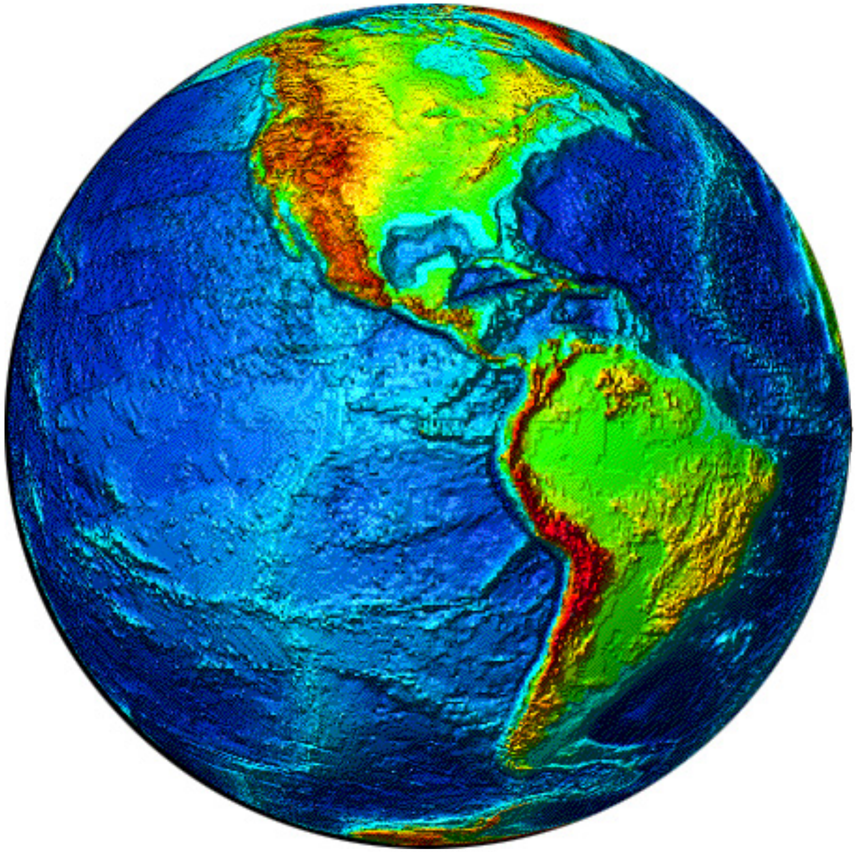
energy transfer



$$\frac{\partial \rho}{\partial t} + \nabla.\rho\mathbf{u} = 0$$

$$\rho\frac{\partial \mathbf{u}}{\partial t} + \rho\mathbf{u}.\nabla\mathbf{u} = -\nabla P - 2\mathbf{\Omega} \times \mathbf{u} + \eta\nabla^2\mathbf{u}$$

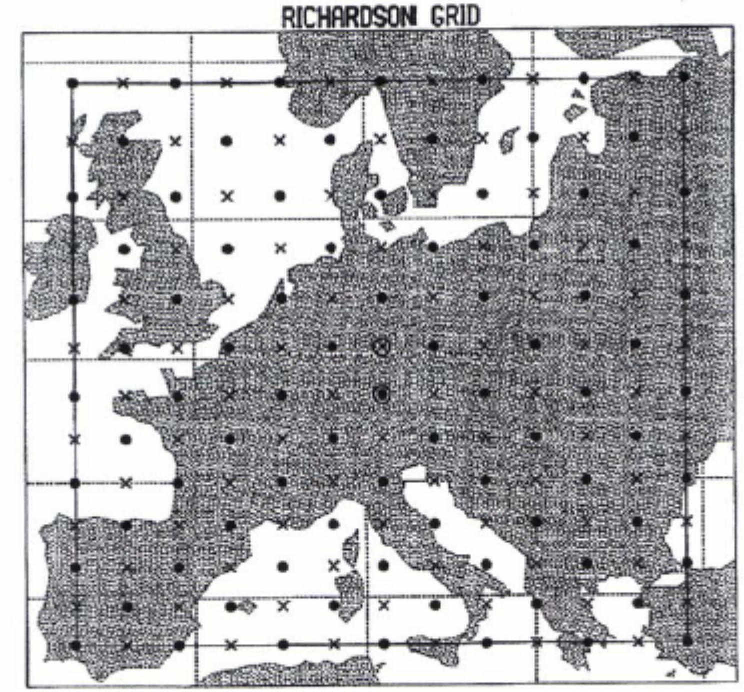
$$\frac{\partial T}{\partial t} + \nabla.\mathbf{u}T = \kappa\nabla^2 T + \mathcal{F}.$$



- possible!
- impossible!

Lewis Fry Richardson

- the first numerical prediction was made by Richardson during WW1
- … he did two timesteps of 3 hours on a 7x7x5 (~250 km) grid
- unfortunately, these took 2 years to perform by hand
- … and he got the answers wrong because of poor input data



but now we have computers

- UK daily forecast: 12 km grid + 70 levels in the atmosphere
  - runs every 6 hours

- Each requires 2,000,000,000,000,000 (2×10^{15}) calculations

- Your PC could do that calculation but it would take 3 months to predict tomorrow's weather!

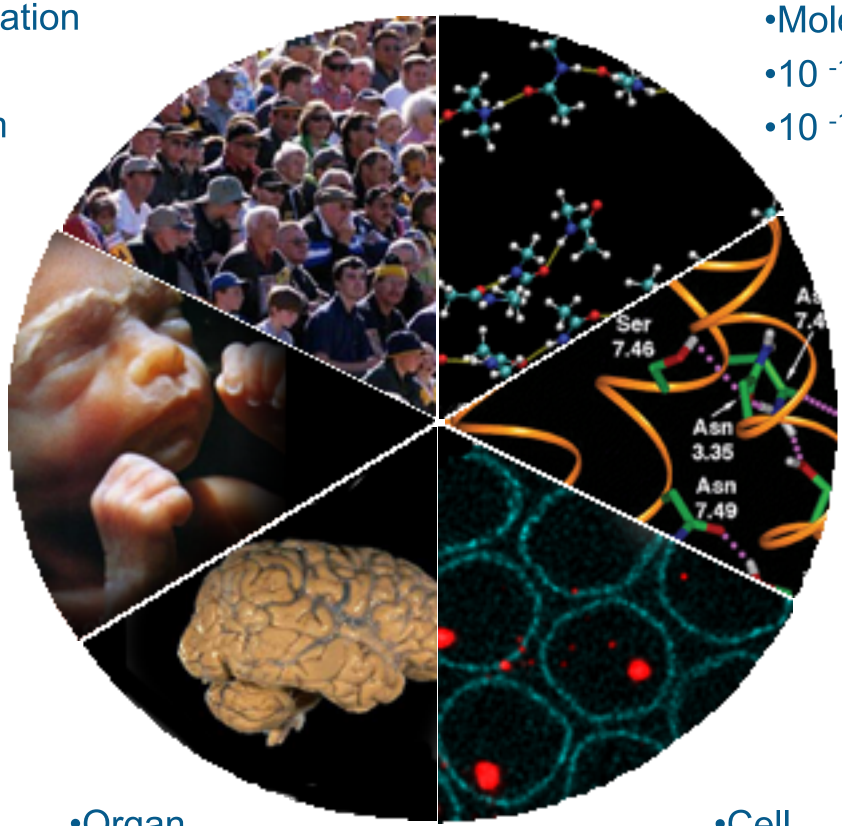
grand challenges

- beyond physics there are even bigger problems

•Darwinian evolution

•Population
•10^{9} s
•10^{5} m

•Molecular dynamics
•10^{-15} s
•10^{-10} m



•Organism
•10^{7} s
•1 m

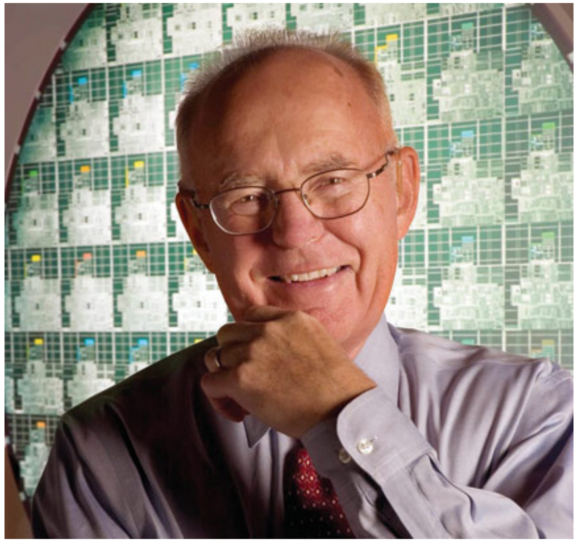
•Molecular
•conformation
•10^{-3} s
•10^{-9} m

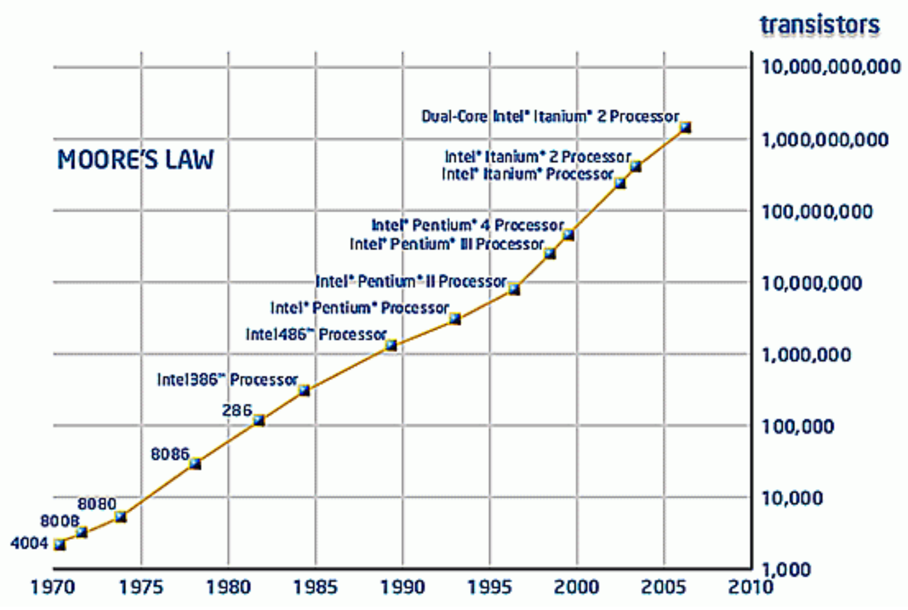
•Organ
•10^{3} s
•10^{-3} m

•Cell
•10^{-3} s
•10^{-7} m

the engine room

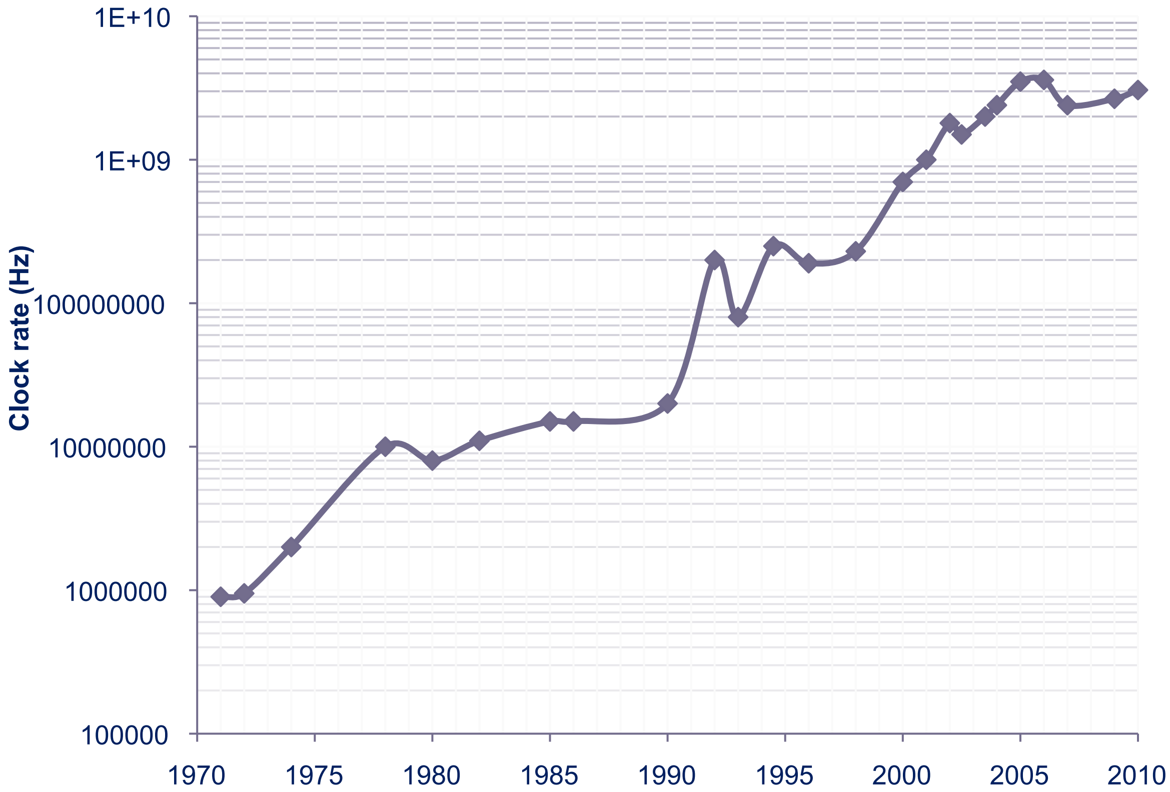
- for the last 40 years we have relied on Moore's Law to deliver faster computers
- … and there is no obvious end in sight, so no problem?
- WRONG!





(not) Moore's Law

- Moore's Law is often confused with its corollary: increasing clock rate … and hence performance



- … and this has peaked

Dennard's Law

- in 1974 Dennard et al established a set of scaling rules for MOSFETs (cf 1)
  - subsequently extended to include CMOS

Parameter	**Scaling Factor**
doping concentration	k
device size	1/k
voltage	1/k
power dissipation	$1/k^2$
power density	1
Switching time	1/k

- but as k rises, carrier mobility degrades
- and leakage currents increase through quantum tunnelling

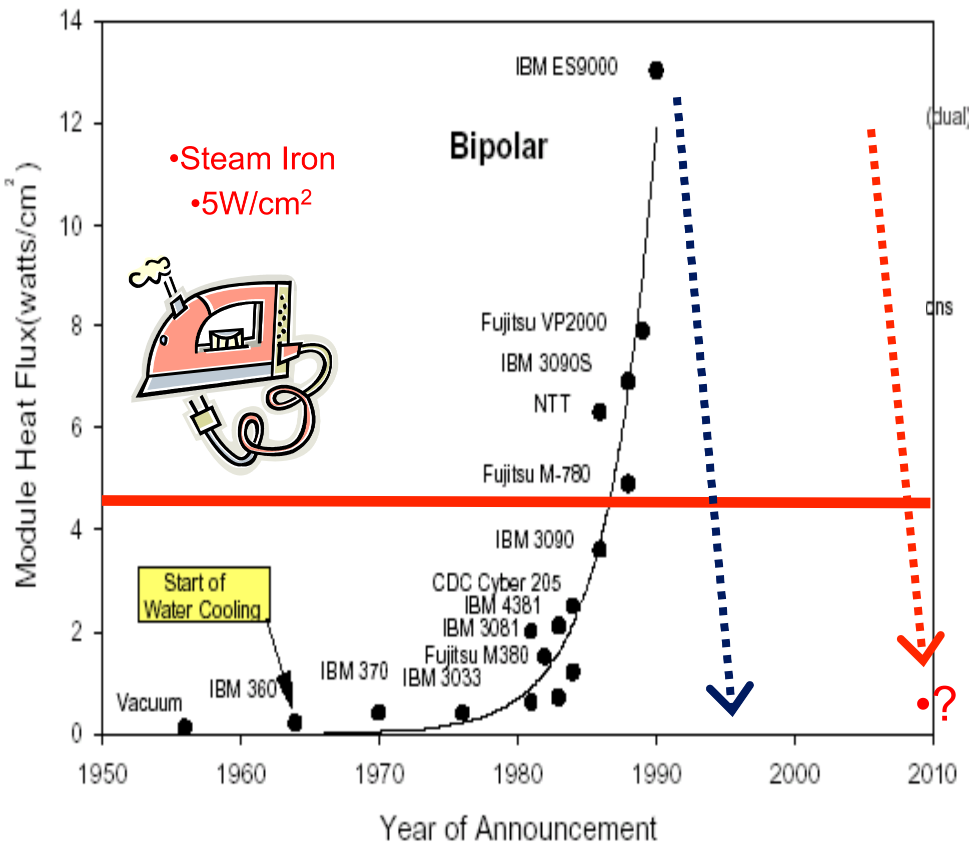
1:http://www.ieee.org/portal/site/sscs/menuitem.f07ee9e3b2a01d06bb9305765bac26c8/index.jsp?&pName=sscs_level1_article&TheCat=6010&path=sscs/07Winter&file=Bohr.xml

quantum effects

- when one atom high "bumps" look significant in a photo of your transistors, both their manufacture and behaviour will be "exciting"[1]

- without a technology change, further shrinkages will produce bizarre behaviours
  - transistors use the same power doing nothing as when they work
  - the new generation of transistors consume dramatically more power than prior experience would have predicted
  - performance gains will be slower than past experience would have indicated

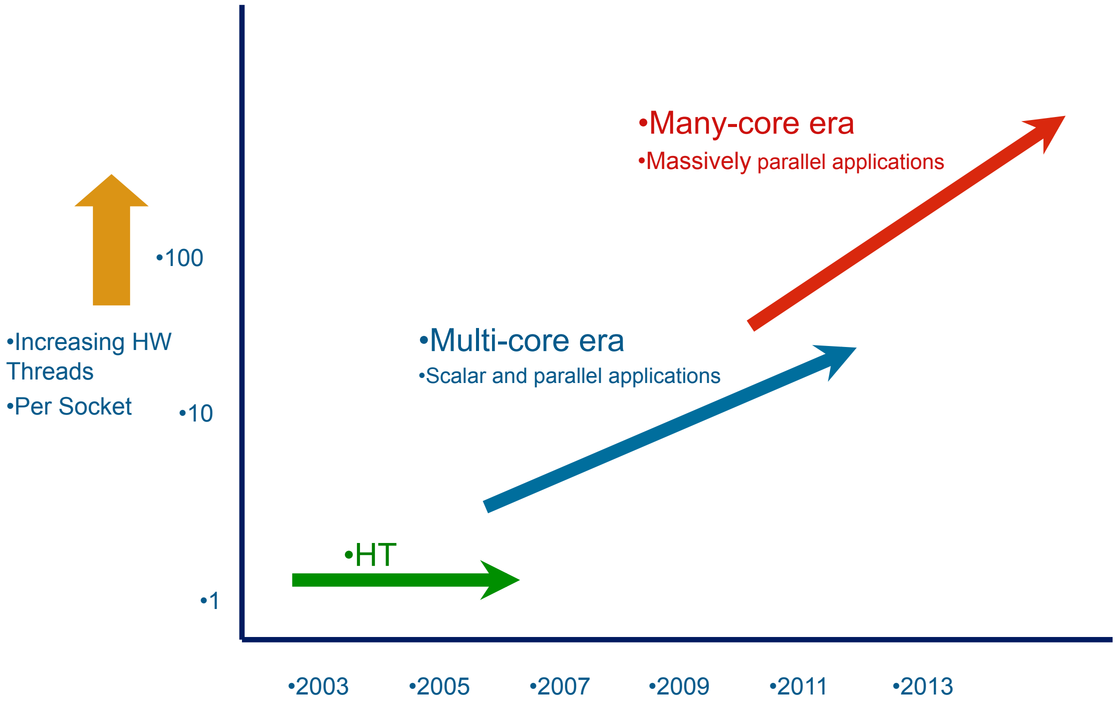
1: Bernie Meyerson, Chief Technologist, IBM



- •high-k gate dielectrics
- •? •power management intelligence

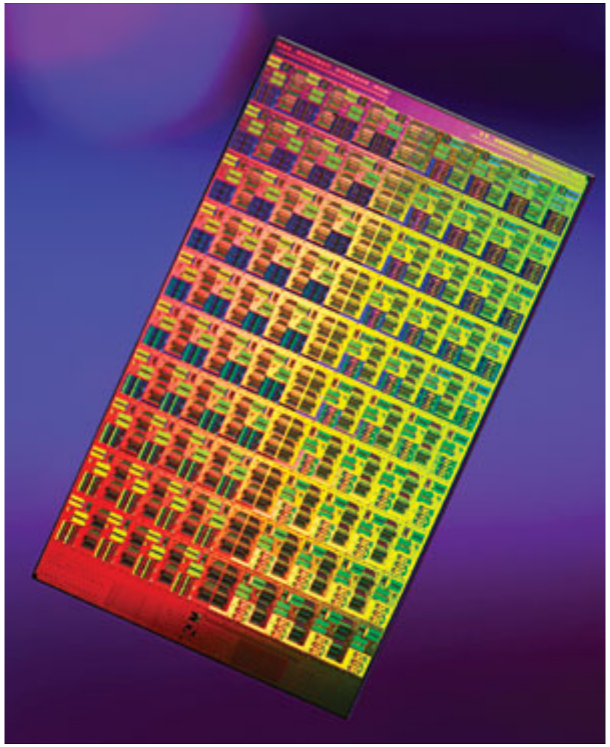
continuing Moore's Law

- Intel Roadmap: Platform 2015



parallelism on the desktop

- functional parallelism easy to implement and effective on dual-core



- … but how do you use an 80-core microprocessor effectively?
- this is **THE** problem for the computer industry today

how can you possibly mean that?

- the microprocessor is ubiquitous:

- and the economics are based on volume sales to the commodity market
- … remove the incentive/need to upgrade the home PC every 2/3 years and those sales will fall
- that will happen if we can’t use the potential of current microprocessors

parallelism

- has been used for HPC applications for 30 years to help sate applications’ need for performance
- it is, however, a second best to increasing single-node performance
  - first, because it is harder to program
  - second, because its benefits don’t necessarily scale

$$\alpha_1 = \alpha_{serial} + \alpha_{parallel}$$

$$\alpha_n = \alpha_{serial} + \frac{\alpha_{parallel}}{n}$$

$$f\,or\,n \rightarrow \infty \quad \frac{\alpha_n}{\alpha_1} = \frac{\alpha_{serial}}{\alpha_{serial} + \alpha_{parallel}}$$

- for α_{serial} = 3%, maximum speedup ~ *30x*
- Amdahl’s Law assumes fixed workload: *strong scaling*

Gustafson's Law

- the problem with Amdahl's Law is that the workload is fixed
- … what if we let this increase with the number of processors?

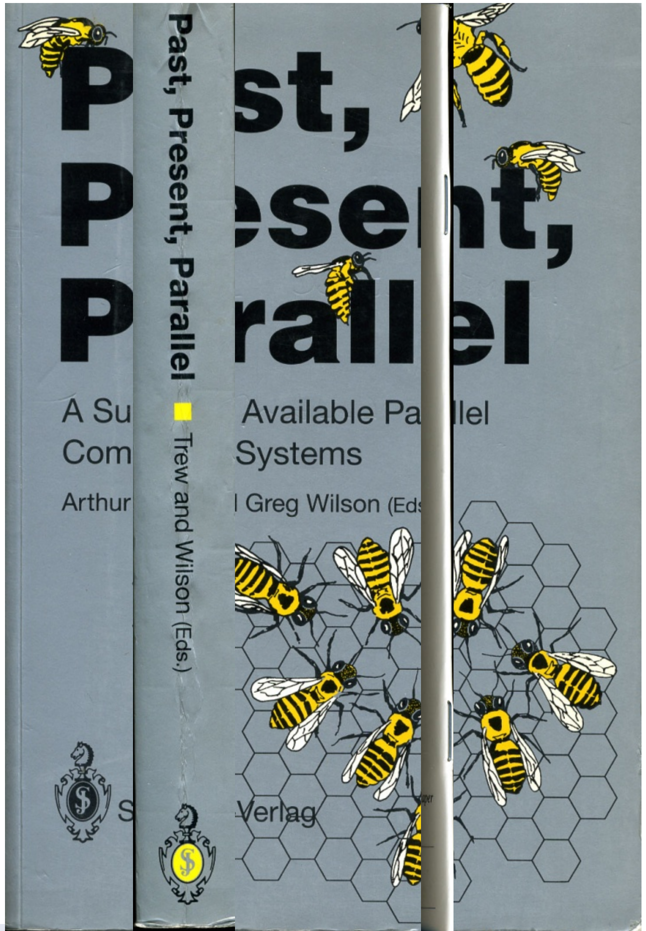
$$\alpha_1 = \alpha_{serial} + 1.\alpha_{parallel}$$

$$\alpha_n = \alpha_{serial} + \frac{n.\alpha_{parallel}}{n} = \alpha_1$$

- hence, speedup α n … this is known as *weak scaling*
- and we have relied on this for the past 30 years

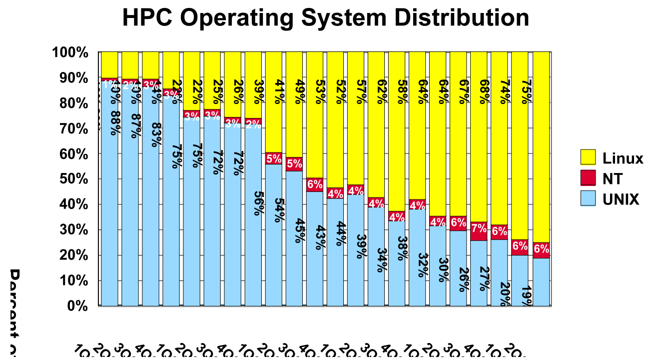
economic pressures

- in the early 90’s parallel HPC was a niche area
  - large computer vendors were not (really) engaged
- medium-sized companies differentiated themselves through technological innovation
  - custom microprocessors
  - custom networks
  - custom O/S
  - custom languages
- lack of standards made portability poor
- … and expensive



parallelism today

- now, after a decade of standardisation, we have
  - –MPI, OpenMP
  - –Fortran90, C++
  - –Linux
  - –PBS, TotalView, ScaLAPACK …



- standard microprocessors (AMD, Intel, Power)
- …only the memory architecture and inter-processor network for companies’ USP
- with clusters even these areas are “commoditised”
  - –so, while clusters have a role in HPC
  - –cluster vendors compete on support, packaging and price

parallelism tomorrow

- are today's programming methodologies fit for the future?
- … that's the key issue for Lecture 2