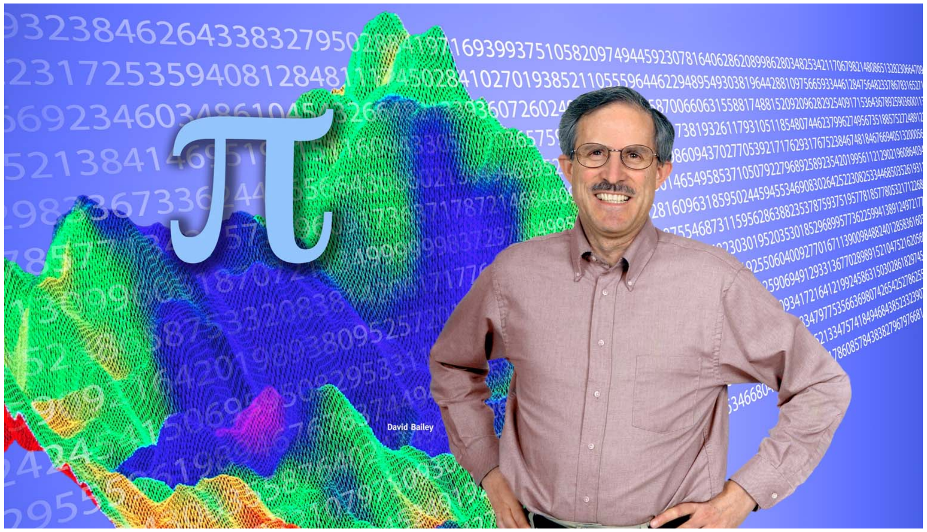
Computing as the Third Mode of Scientific and Mathematical Discovery

David H Bailey

Lawrence Berkeley National Laboratory

http://crd.lbl.gov/~dhbailey



Laplace: Father of Modern Scientific Computing?

Computer simulations can be seen as the modern realization of a centuries-old dream known as the "clockwork universe." This was stated most clearly by Pierre Simon Laplace in 1773:

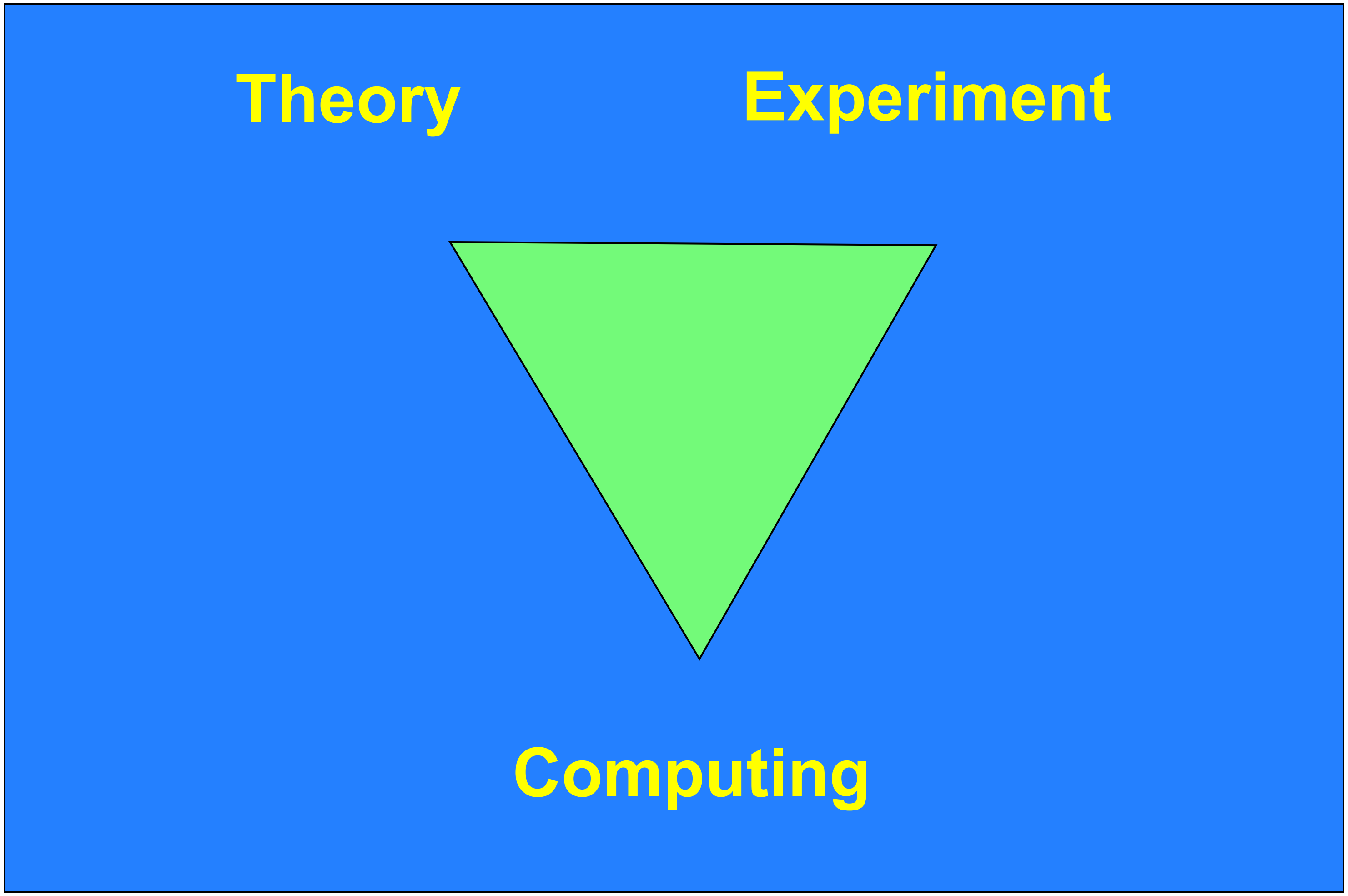
- "An intelligence knowing all the forces acting in nature at a given instant, as well as the momentary positions of all things in the universe, would be able to comprehend in one single formula the motions of the largest bodies as well as of the lightest atoms in the world, provided that its intellect were sufficiently powerful to subject all data to analysis; to it nothing would be uncertain, the future as well as the past would be present to its eyes."

We now know that this dream, taken to its logical extreme, is unrealistic:

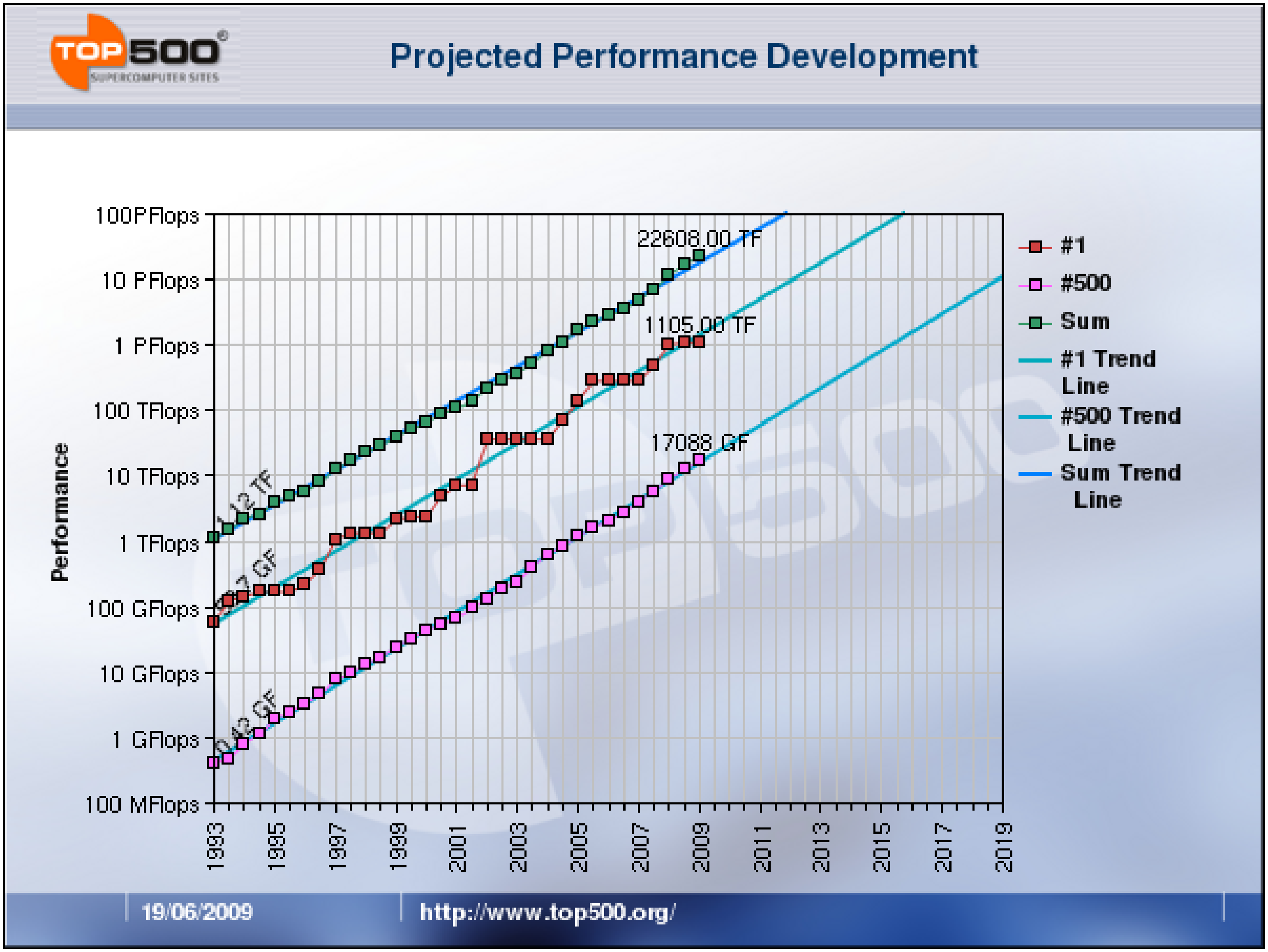
- Chaos theory teaches us that even simple physical systems exhibit chaos: slight changes to the present state are exponentially magnified in future states.
- Quantum theory teaches us that it is fundamentally impossible to know both positions and velocities at any instant to arbitrarily high accuracy.

But many physical systems are amenable to computer simulation, and even chaotic systems can be studied computationally.

Computing:
The Third Mode of Scientific Discovery



Progress of Scientific Supercomputers: Data from the Top500 List



New, Multi-Petaflfops Systems

- Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA: Currently installing “Hopper” system, 1 Pflop/s peak.
- National Center for Supercomputer Applications (NCSA), University of Illinois: Plans to install IBM “Blue Waters” system, 10 Pflop/s peak, in 2012.
- Lawrence Livermore National Laboratory (LLNL), Livermore, CA: Plans to install “Sequoia” system, 20 Pflop/s peak in 2011.
- Japan’s Institute for Physical and Chemical Research (RIKEN), Kobe, Japan: Plans to install Fujitsu system, 10 Pflop/s peak in 2011.

Do We Really Need a Such Powerful Computers?

In 1998, just prior to the time when the first 1 Tflop/s (10^9 flop/sec) computer was installed, a U.S. Congressional Budget Officer asked me, “won’t that finally satisfy the needs of you scientists?”

Needless to say, the answer was a resounding “no.”

In 2010, just after passing the 1 Pflop/s (10^{15} flop/sec) level, the answer is the same:

- For many existing applications, scientists need much finer 3-D grids, or require more sophisticated physical/mathematical models.
- Many new applications require huge amounts of computing and/or storage – chemistry, biology, environmental science, engineering.
- New, larger systems are typically fully utilized within weeks of installation.

Scientists already have sketched applications for 1 Eflop/s (10^{18} flop/sec) systems (probably available in 2020). After that, 1 Zflop/s (10^{21} flop/sec)...

Computing: A Multi-Disciplinary Symphony of 21st Century Science and Engineering

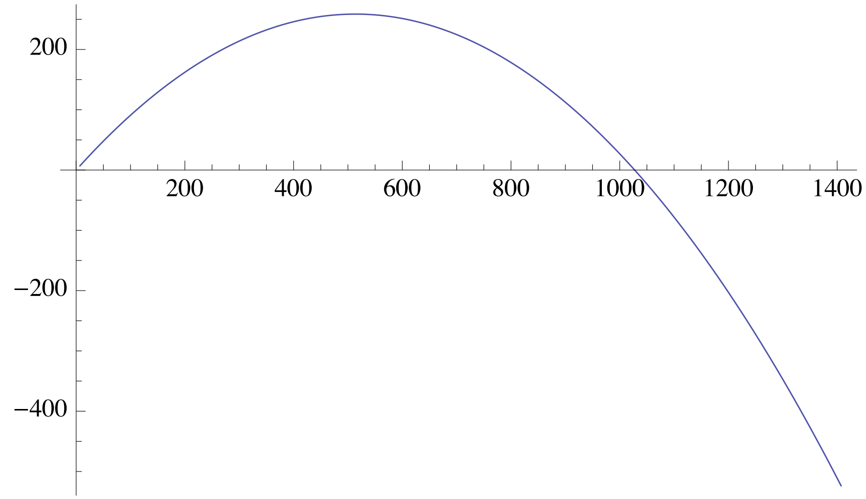
- Basic physical laws.
- Mathematical formulations of these laws.
- Numerical algorithms to solve the mathematics.
- Computational techniques to implement algorithms.
- Grid generation.
- Multidimensional optimization techniques to explore parameter variations.
- Parallel computing methods.
- Scientific visualization.
- Performance monitoring and analysis.
- Computer system software.
- Computer system hardware.

Basic Physical Laws

Many large scientific computations are merely repeated applications of simple physical laws, e.g., Newton's laws, e.g.:

force = mass x acceleration

gravity= G M_1 M_2 / R^2.



Example:

- Suppose initial velocity = 100 m/s at a 45 deg angle (i.e. 70.71 m/s horizontal and 70.71 m/s vertical vel), and is subject to gravity.
- Then after 0.1 sec, horiz vel = 70.71 m/s, vert vel = 70.71 - 0.1 x 9.8 = 69.73 m/s.
- Thus object will have risen 0.1 x 69.73 = 6.973 m, and moved forward 7.071 m.
- Repeated applications yield the above curve. More accurate results can be obtained by using a smaller time interval, e.g. 0.01 sec or 0.001 sec.
- This is just for one object. Large simulations do similar computations for millions or billions of objects, on a large 3-D grid of spatial points.

Mathematical Formulations of Physical Laws

All basic laws of physics can be encapsulated into mathematical equations. For example, Maxwell's equations governing light and electromagnetic radiation can be written as:

$$
\begin{aligned}
\nabla \cdot E &= 4\pi\rho \\
\nabla \times E &= -\frac{1}{c}\frac{\partial B}{\partial t} \\
\nabla \cdot B &= 0 \\
\nabla \times B &= \frac{4\pi}{c}J + \frac{1}{c}\frac{\partial E}{\partial t}
\end{aligned}
$$

Numerical Algorithms

- Advanced numerical algorithms are used to solve the underlying mathematical formulations of physical laws.
- These algorithms dramatically lower the amount of computation normally required – modern scientific computing could not be done without them.
- Examples:
  - Dense linear algebra.
  - Sparse linear algebra.
  - Spectral methods (i.e., fast Fourier transforms).
  - N-body methods.
  - Structured grid methods for partial differential equations.
  - Unstructured grid methods for partial differential equations.
  - Distribute-reduce schemes.
  - Sorting, searching.
  - Optimization, maximization, minimization.
  - Many others.

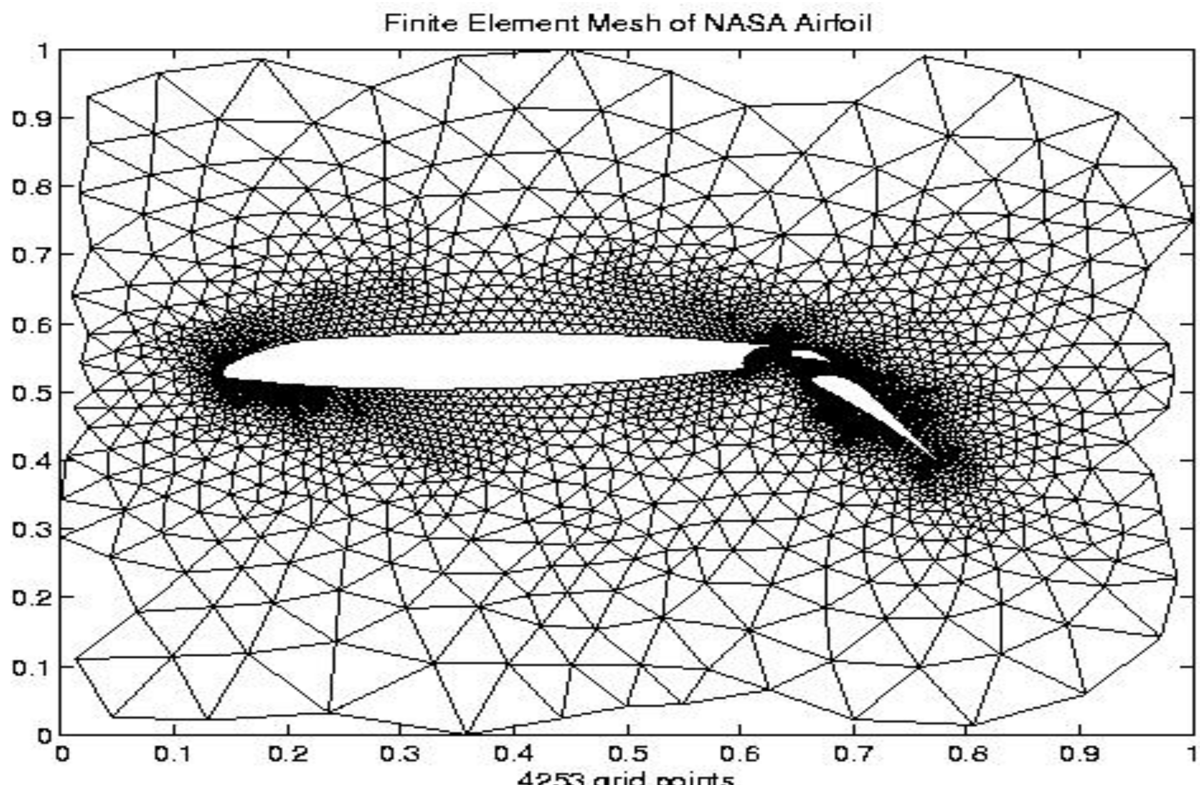
Computational Techniques

Because of the realities of modern computer architectures, efficient implementations of basic algorithms require considerable sophistication:

- Data locality is very important:
  - The “hierarchical” design of processors (Level-1 cache, Level-2 cache, Level-3 cache, main memory, disk memory) means that computations must be carefully structured for good performance.
  - It is often more efficient to recompute a value, rather than store it to memory and later fetch it.
- Changing computer designs means that computer programs that once were “efficient” now must be revised.
  - Floating-point multiply operations once were much more expensive than add/subtract operations, so codes were changed to exploit this. Now there is no difference.
  - Computation (add, subtract, multiply, divide) was once much more expensive than fetching and storing data to memory. Now the opposite is true.

Grid Generation

- In many computational simulations, a key aspect of the computer implementation is the construction of a grid of points or polygons that encompass the physical object under study.
- Techniques for generating efficient and effective grids are an active field of research.



Optimization Techniques

- After a computational simulation program has been successfully developed, say to simulate the operation of a jet engine, then researchers run many instances of this simulation with variations of the input parameters, in an attempt to find an optimal configuration.
- Advanced optimization techniques permit these searches to be done significantly more efficiently than by exhaustive search.
- The resulting discipline of multidimensional optimization now is an essential part of high-performance computing.
- On some large supercomputers, almost all jobs perform multidimensional optimization.

Parallel Computing

- Parallel computing (using multiple processors for a single computation) until recently was the exclusive province of large government laboratories.
- Now, even consumer personal computers have more two, four, eight or more “cores,” and it is essential to take advantage of these cores to fully utilize the power of the system.
- Scientists who do not convert their programs to utilize parallel computers will soon be left behind.
- Constructs must be inserted into computer programs to control:
  - Data layout.
  - Broadcast of data from the control node to other nodes.
  - Synchronization between nodes.
  - Communication between nodes.
  - Collection of data from all nodes back to a single node.
  - Parallel input and output of data to external disk drives.

Amdahl's Law

A simple principle first enunciated by Gene Amdahl in the 1960s places limits on speedup from parallel processing:

- If some fraction P of a computation is amenable to parallelization with speedup S, then the maximum possible speedup of the entire calculation is no more than $1 / ((1-P) + P/S)$.

For instance, if 99.9% of a computation can be effectively run in parallel on 10,000 processors, and the rest must be done serially, then the maximum possible speedup is only 909 (out of 10,000).

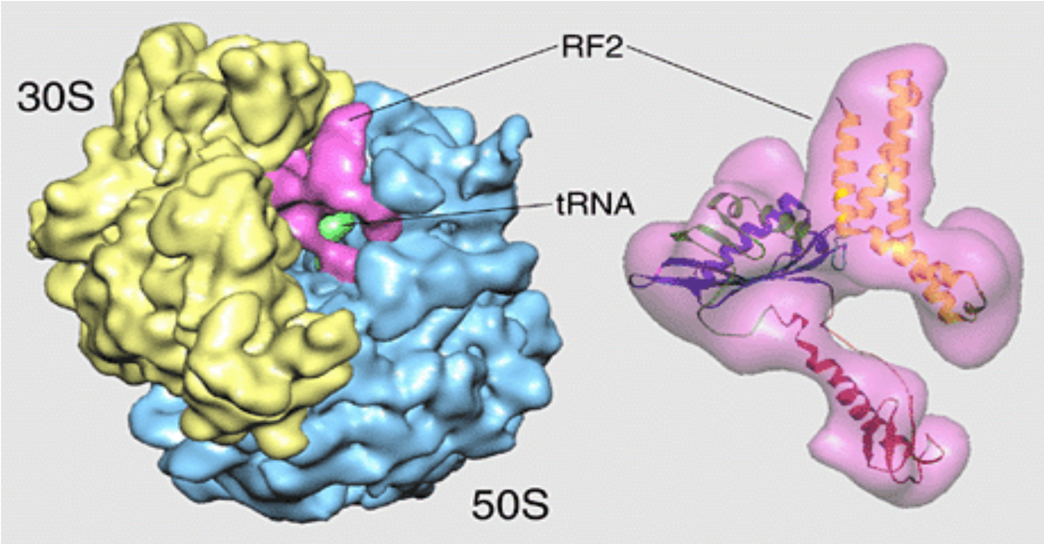
So far scientists working on large-scale computers have been able to keep Amdahl's Law "at bay." But how long will our luck hold out?

- In some arenas, such as climate modeling, computer programs are already pressing the limits of available concurrency.

Present-day leading-edge computations typically must exhibit 10^8-way concurrency at EVERY step of the computation.

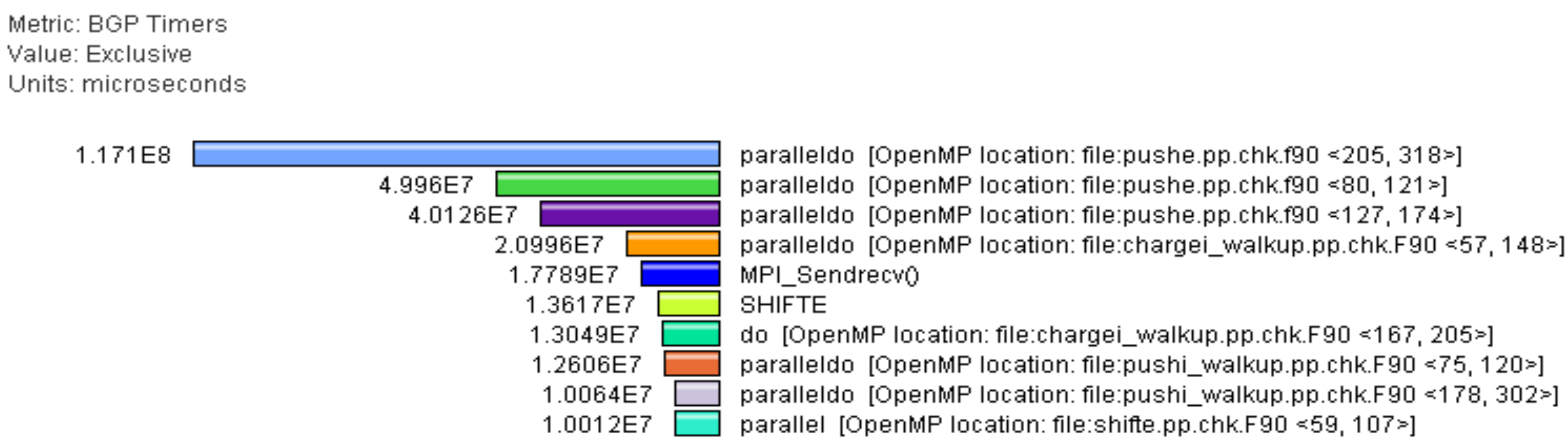
Scientific Visualization

- With the enormous volumes of data now involved in a large scientific simulation, it is no longer possible for a scientist to examine bits of data one by one.
- Instead, sophisticated scientific visualization software must be employed.
- Finding better ways to generate displays, and finding newer approaches to visualization, are active areas of research.



Performance Monitoring and Analysis

- Scientists are often disappointed in the performance (computational speed) of their programs – often only a few percent of the peak theoretical performance is achieved.
- Numerous sophisticated software tools are now available to find bottlenecks and improve performance.
- The development of automatic performance tuning tools for scientific computer programs is an active area of research.



System Software

Underlying all of scientific computing is an enormous body of system software, without which scientific computing would not be possible:

- Operating system – Linux is the most widely used, although BSD Unix and IBM's AIX Unix are also used.
- Compilers – before any program is executed, it must be "compiled", i.e., translated to machine instructions, using sophisticated techniques to obtain the best performance.
- Support for parallel computing.
- Support for large-scale data storage.
- Support for networking.

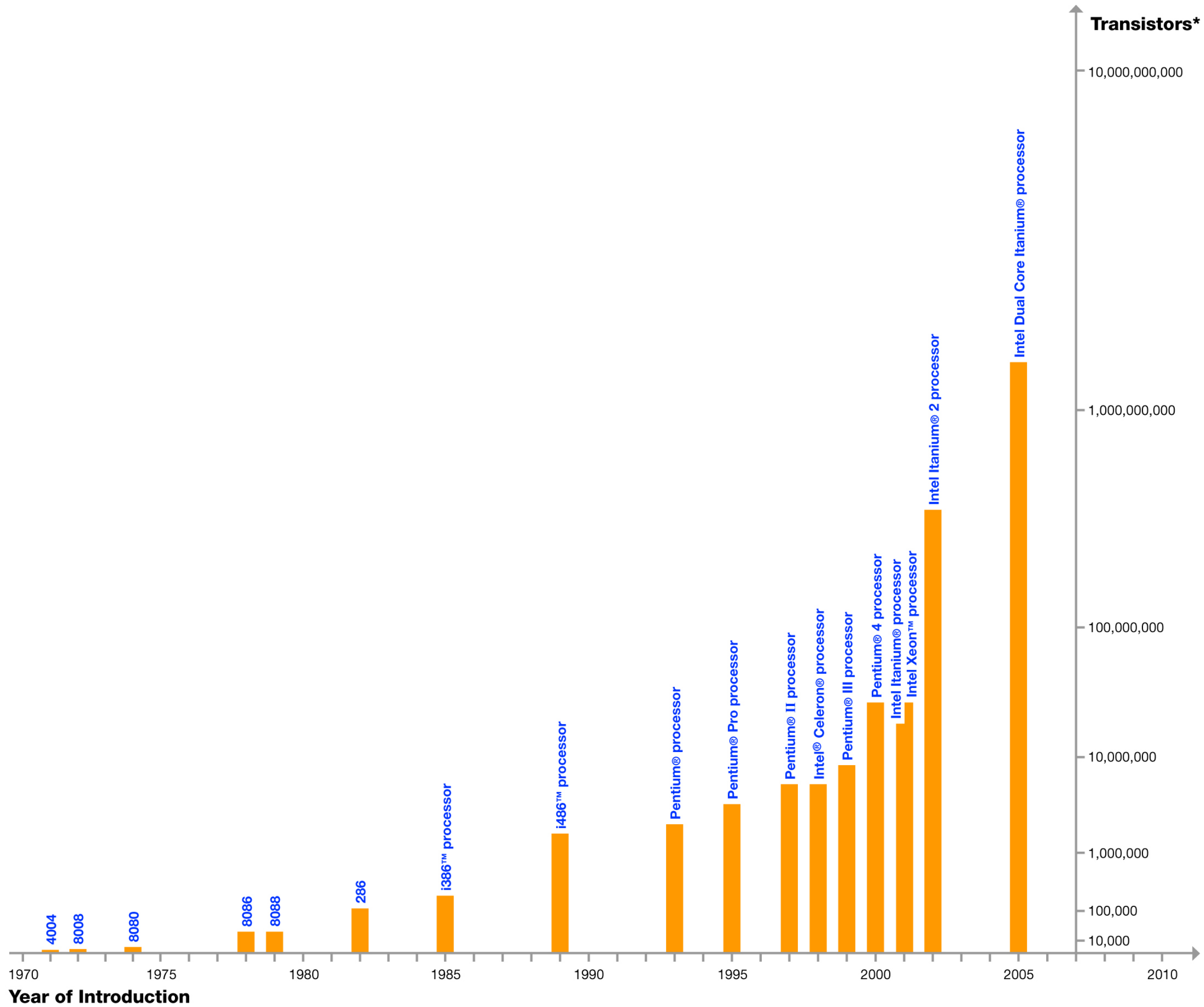
Computer Hardware

A large-scale scientific computer is much more than just a collection of chips – the large-scale system architecture is also important.

- Until about 2000, vector supercomputers were the most common platform for large-scale scientific computing.
- Now most scientific computing is done on large clusters of units, each of which is often an off-the-shelf personal computer system.
- The interprocessor network is very important – without a very high-speed network, many scientific computations would be mired in network congestion.

Moore's Law: 45 Years of Sustained Exponential Growth in Computer Power

From Computer Desktop Encyclopedia
Reproduced with permission.



*Note: Vertical scale of chart not proportional to actual Transistor count.

In 1965 Gordon Moore (co-founder of Intel with Andy Grove and Robert Noyce) predicted that the transistor density of semiconductor chips would double every 18 months or so. After 45 years, no end is yet in sight!

LBNL’s National Energy Research Scientific Computing Center (NERSC)

NERSC serves a large population of users:

~3000 users, ~400 projects, ~500 codes

Allocations managed by the Department of Energy:

- 10% INCITE awards:
  - Open to all of science, not just DOE-funded projects.
  - Large allocations, extra service.
- 70% production (ERCAP) awards:
  - From 10,000 CPU-hours to 5,000,000 CPU-hours.
- 10% each NERSC and DOE reserve for special needs.

NERSC 2009 Configuration

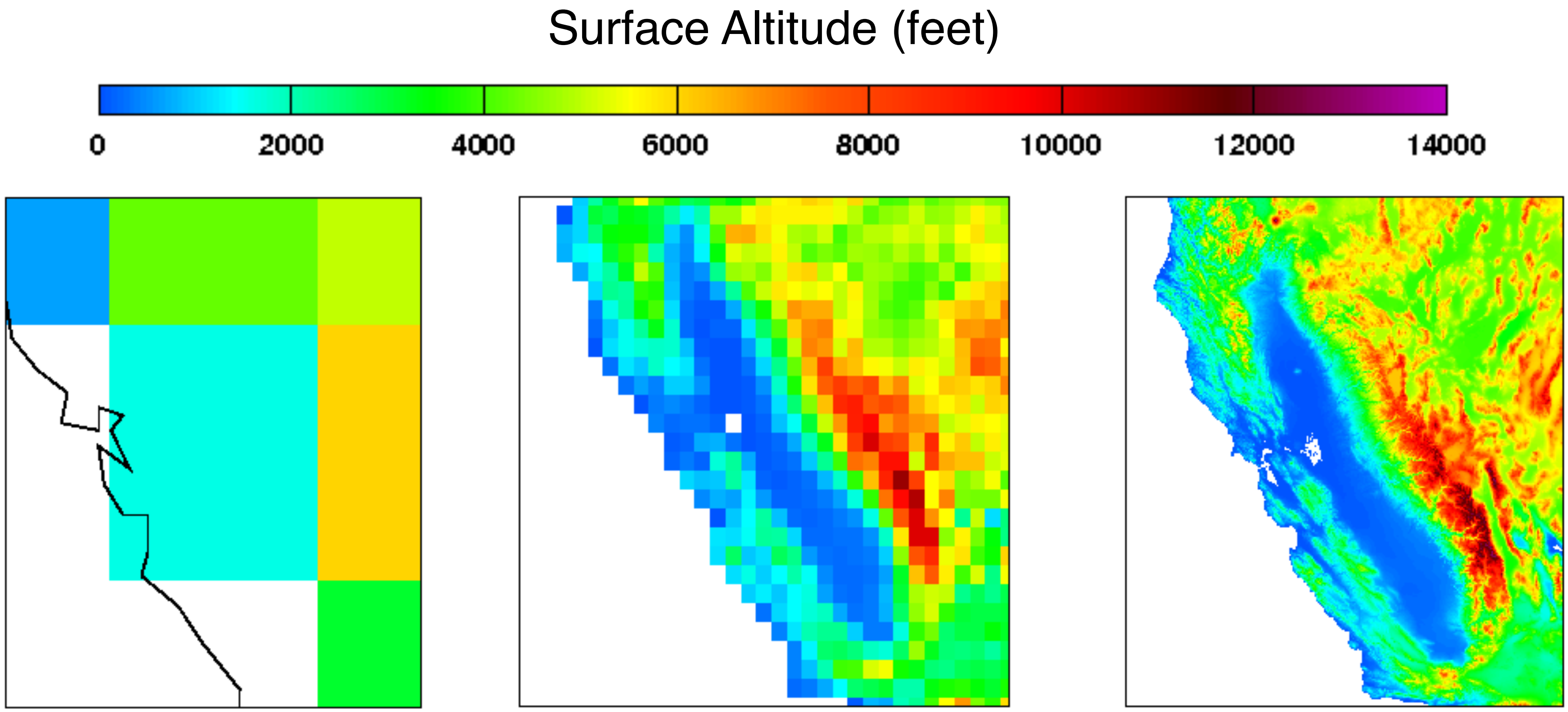
Current flagship system: “Franklin”

- Cray XT4
- 9,740 nodes; 19,480 cores
- 78 Tbyte main memory
- 355 Tflops/s theoretical peak
- ~25 Tflops/s sustained on real scientific work

New upgrade (3Q 2010): “Hopper”

- Cray XT-5, 6,400 nodes, 153,600 cores
- 1.9+ GHz AMD Opteron chips
- 1.17 Pflop/s peak performance (1.17 x 10^{15})
- Expect 100 Tflop/s sustained performance
- 217 TB DDR3 memory total
- Gemini Interconnect
- 2 Pbyte disk, 80 Gbyte/s bandwidth
- Liquid cooled

Global Cloud System Resolving Models



200km
Typical resolution of
IPCC AR4 models

25km
Upper limit of climate
models with cloud
parameterizations

1km
Cloud system resolving
models
are a transformational
change

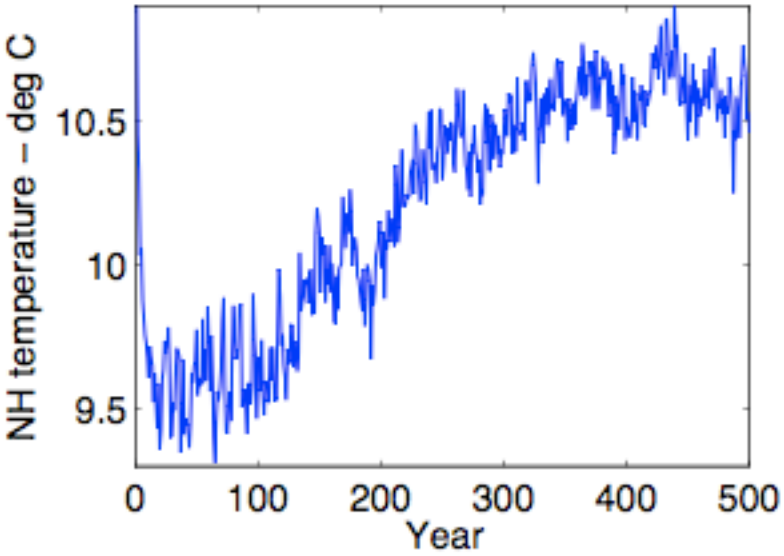
Climate Modeling with CCSM

- Climate Change Simulations with CCSM: Moderate and High Resolution Studies.
- Principal investigator: Warren Washington, NCAR.

- Science Results:
  - 2000-2100 simulation on preserving polar bear habitat by reducing non-CO_2 emissions.
  - Separating human and natural forcings in climate change.
  - Sulfate and carbon sulfate impact isolated.

  Granted 12,000,000 CPU-hours in 2009.

  Typical runs utilize 5,000-6,000 cores.



Material Modeling for Geoscience

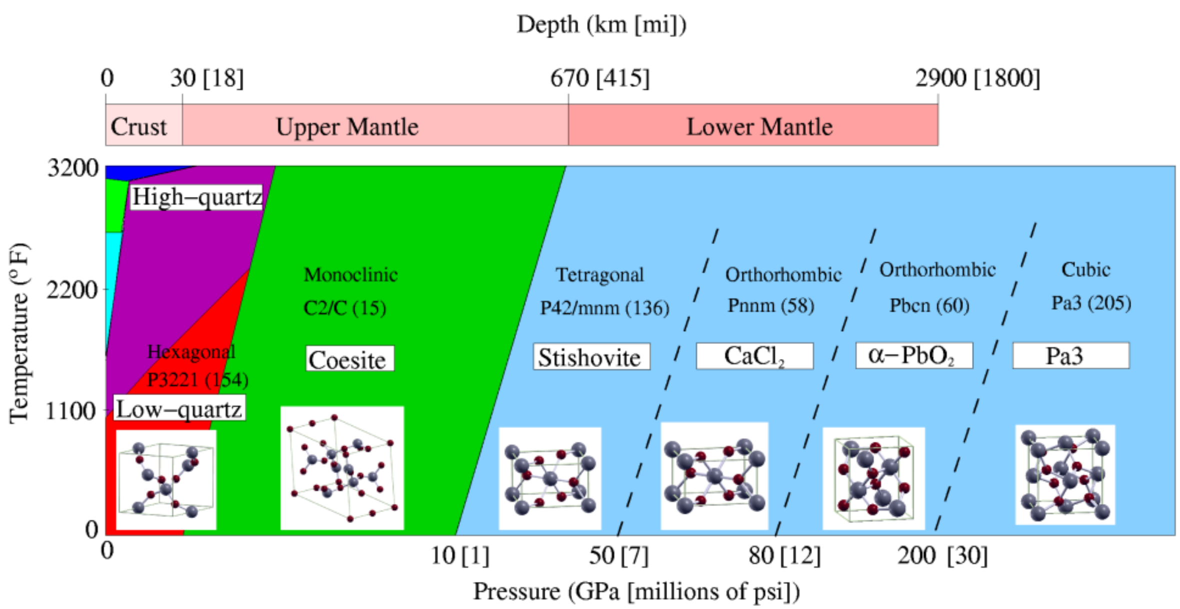
- Calculation: Simulation of seismic waves through silicates, which make up 80% of the Earth's mantle; important for understanding structures in oil well drilling, carbon sequestration, earthquakes, etc.
- PI: John Wilkins, Ohio State University

Science Result:

– Seismic analysis shows jumps in wave velocity due to structural changes in silicates under pressure

First use of Quantum Monte Carlo (QMC) for computing elastic constants.

Typical runs utilize 8,000 cores.

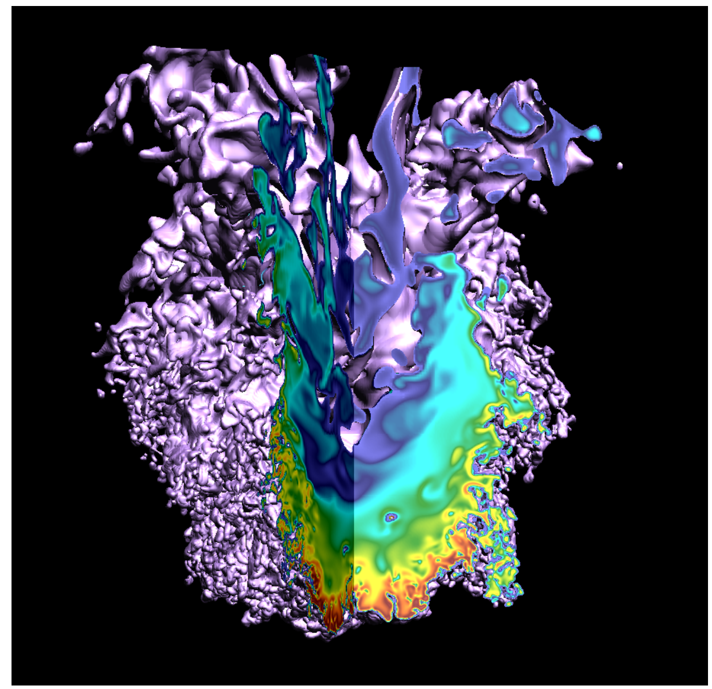


Low-Swirl Combustion Burner Simulation

- Computation: Numerical simulation of a lean premixed hydrogen flame in a laboratory-scale low-swirl burner (LMC code). Uses a low Mach number formulation, adaptive mesh refinement (AMR) and detailed chemistry and transport.
- PI: John Bell, LBNL

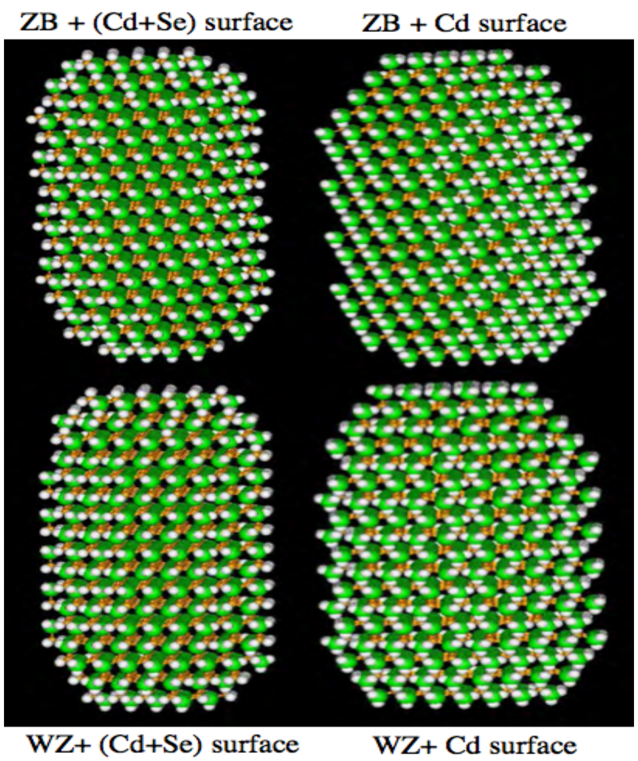
Science Result:

- Simulations capture cellular structure of lean hydrogen flames and provide a quantitative characterization of enhanced local burning structure.
- LMC dramatically reduces time and memory.
- Scales to 4K cores, typically run at 2K
- Used 9,600,000 CPU-hours in 2008; allocated 5,500,000 CPU-hours in 2009.



J B Bell, R K Cheng, M S Day, V E Beckner and M J Lijewski, Journal of Physics: Conference Series 125 (2008) 012027.

New Method for Electronic Structure Calculations



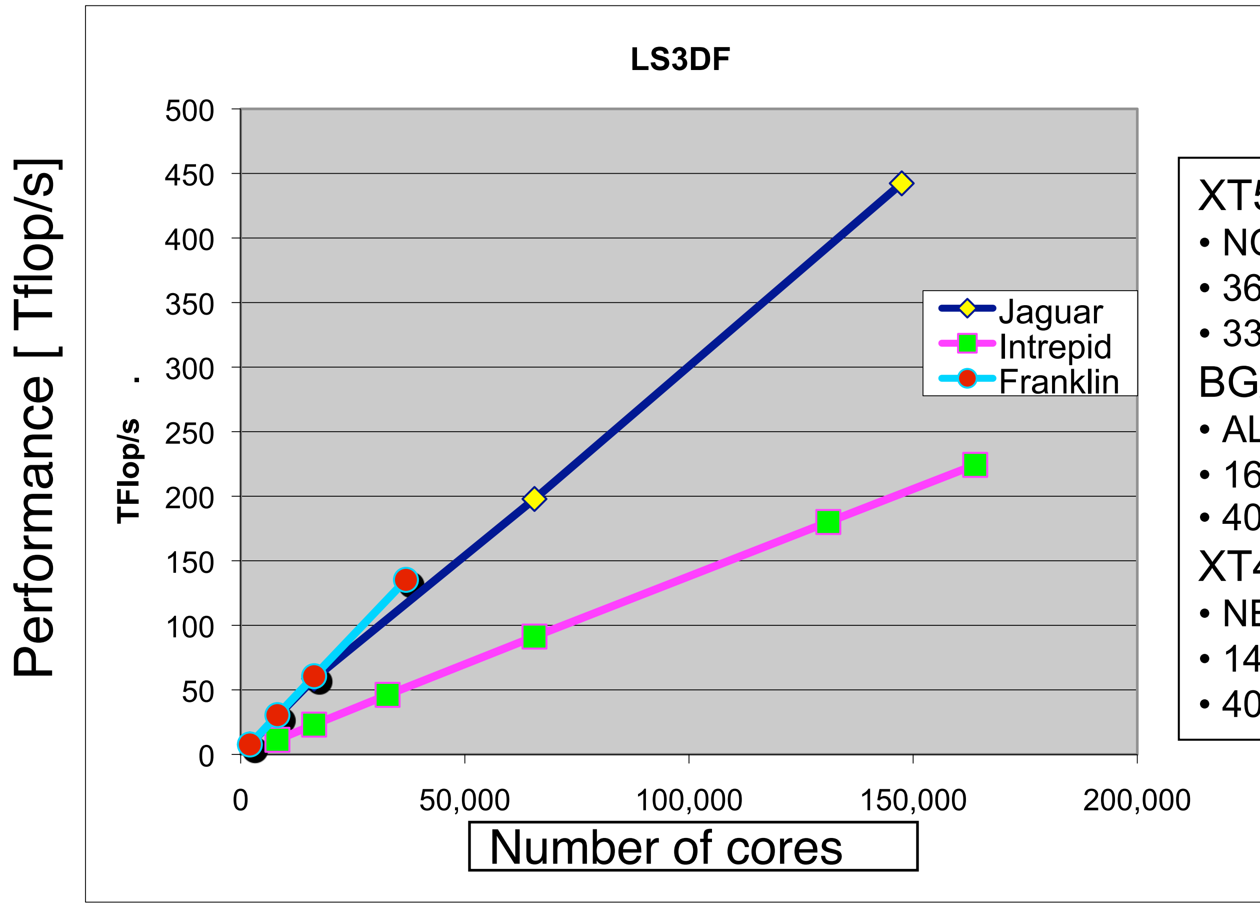
Simulation of CdSe nanorods. The green atoms are Cd, yellow atoms are Se, and white atoms are surface hydrogen. (top) the potential on the central axis; (bottom) the electron state (red) and hole state (green) isosurfaces with an isovalue of 0.0002 e/Bohr^3.

J. Meza, L.-W. Wang, Z. Zhao, LBNL, Nanoscience-Math

- A new linear scaling three dimensional fragment (LS3DF) method for electronic structure calculations now makes possible the simulation of nanostructures with the same accuracy as a direct *ab initio* method.
- The LS3DF method is based on the observation that the total energy of a given system can be broken down into two parts:
  - Long-range electrostatic energy
  - Short-range quantum mechanical energy
- LBNL researchers have used a divide and conquer approach to study the total dipole moments of CdSe quantum dots.

LS3DF Code Wins 2008 ACM Gordon Bell Prize for “Algorithm Innovation”

- Run on Franklin (Cray XT4) at LBNL’s NERSC supercomputer center.
- Subsequent runs on Intrepid (IBM BlueGene/P) at Argonne National Lab and later on Jaguar (Cray XT5) at Oak Ridge National Lab.

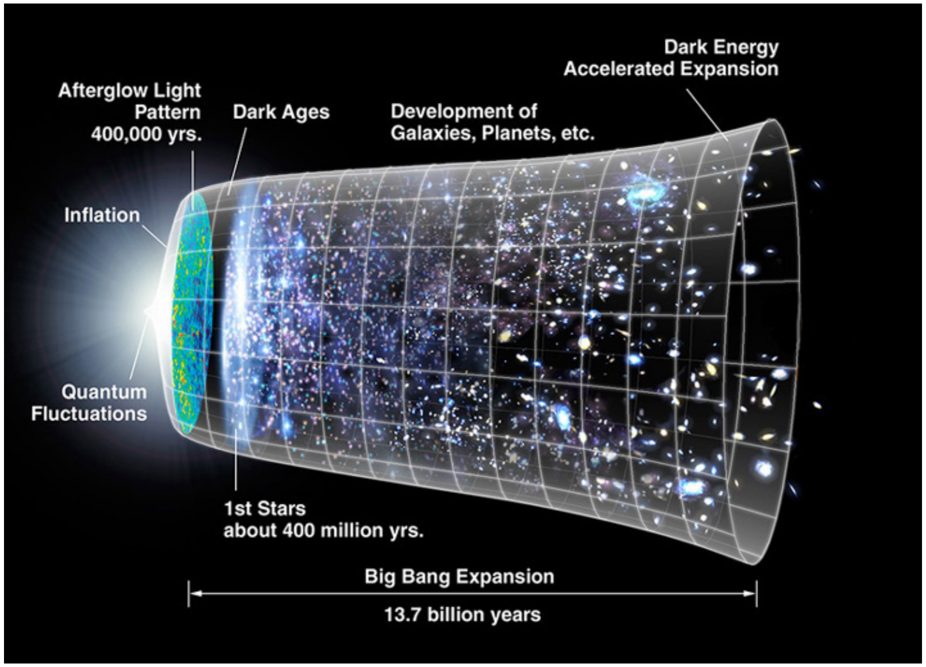


XT5: 442 Tflops
- NCCS Jaguar
- 36,864 cores
- 33% of peak

BG/P: 224 Tflops
- ALCF Intrepid
- 163,840 cores
- 40% of peak

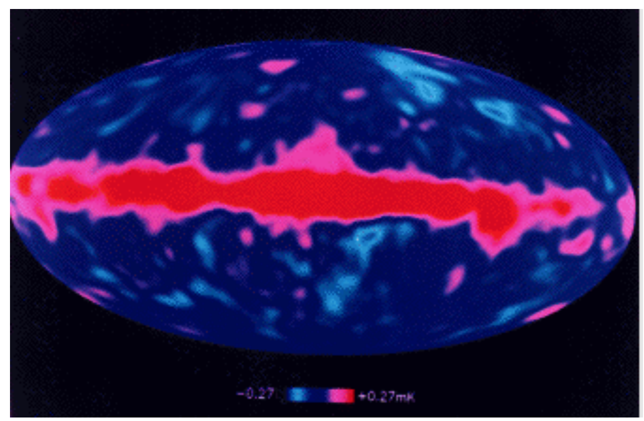
XT4: 139 TFlops
- NERSC Franklin
- 147,146 cores
- 40% of peak

NERSC User George Smoot wins 2006 Nobel Prize in Physics



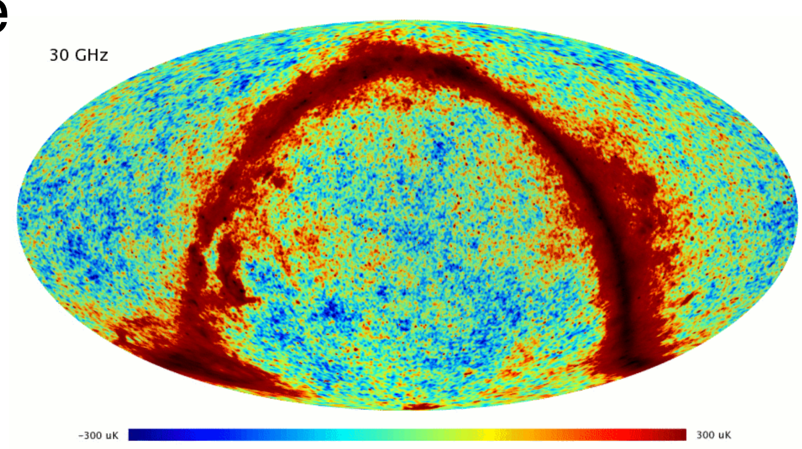
Cosmic Microwave Background Radiation (CMB): an image of the universe at 400,000 years.

COBE Experiment showed anisotropy of CMB



Cosmic Microwave Background Computations

❖ Calculation: Planck full focal plane
  - 1 year simulation of CMB, detector noise and foregrounds.
  - 74 detectors at 9 frequencies.
  - 750 billion observations.
  - 54,000 files, 3 TBbyte data.
- Principal investigator: J. Borrill, LBNL.

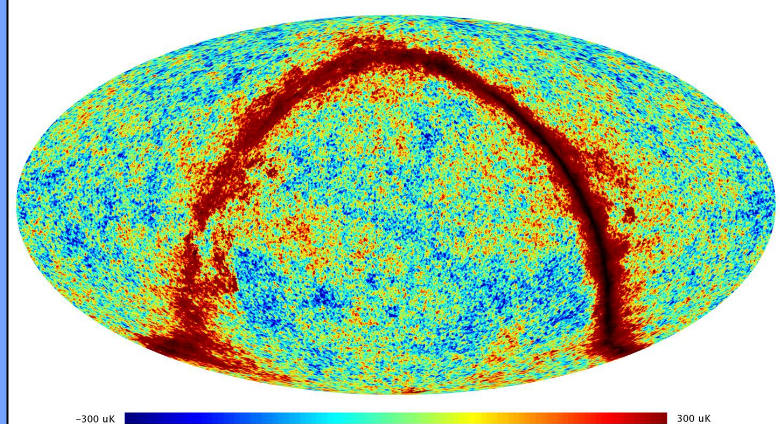


Science Result:

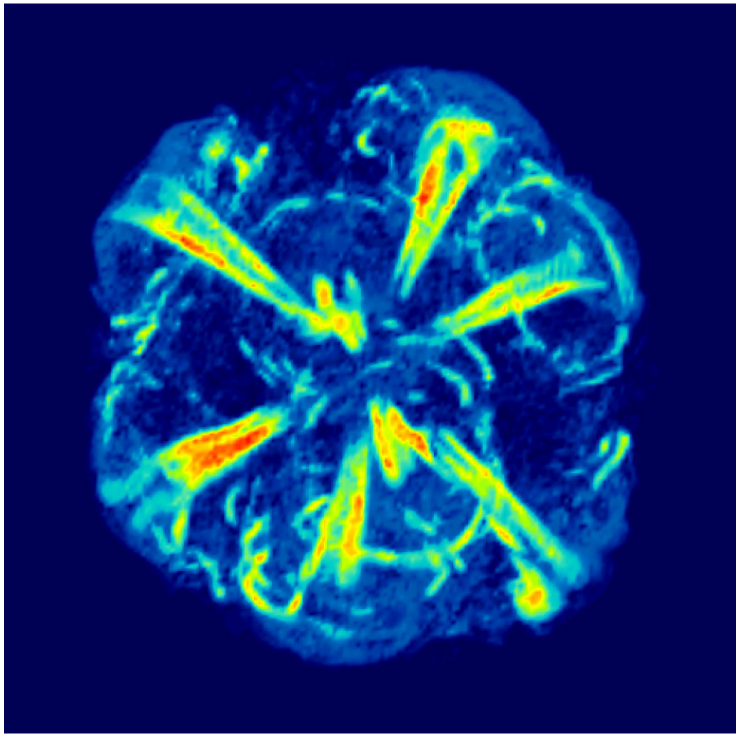
- Unprecedented 9-frequency map with entire simulated Planck data set.

Scaling Result:

- 9-frequency problem ran for < 1 hour on 16,384 processor cores.



Supernovae Simulations

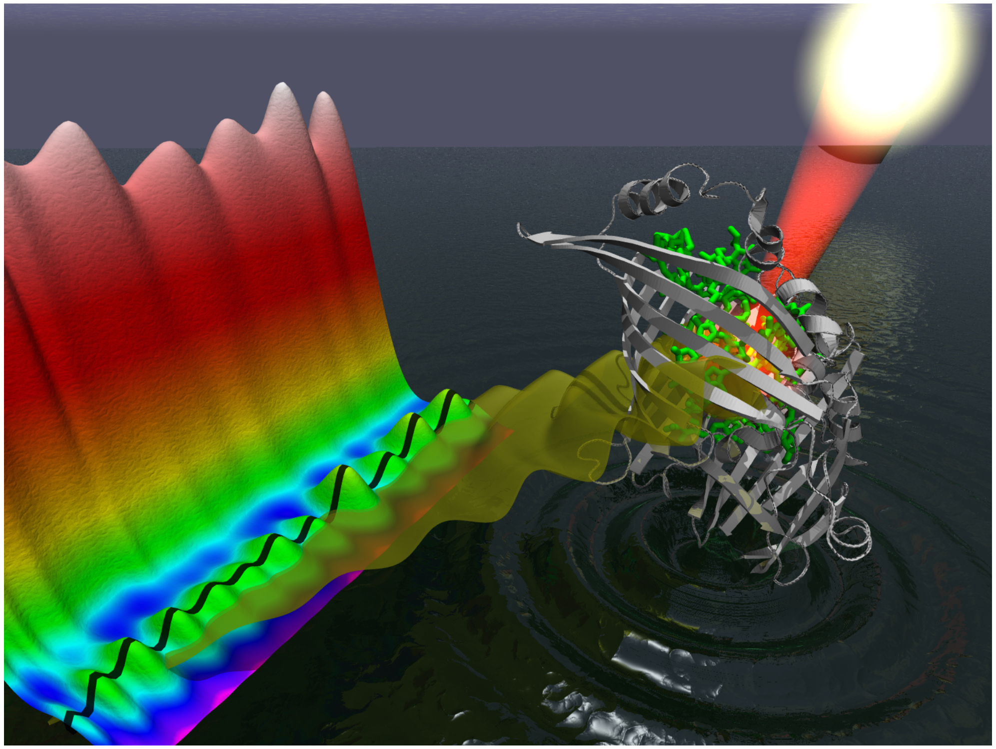


Vortical Structure induced by local heating in a while dwarf. The simulation was performed with the MAESTRO code that is based on the low Mach number analysis developed at LBNL.

A. Almgren, J. Bell, LBNL, Base Math Program

- Researchers at LBNL have developed a new mathematical model based on low Mach number analysis.
- The new method allows astrophysicists to study the ignition process in Type Ia supernovae, which are used as “standard candles” in cosmological studies.
- The low Mach number approach has been validated by comparison with other approaches and promises significant computational savings.

Quantum Effects of Photosynthesis



Sunlight absorbed by bacteriochlorophyll (green) within the FMO protein (gray) generates a wavelike motion of excitation energy.

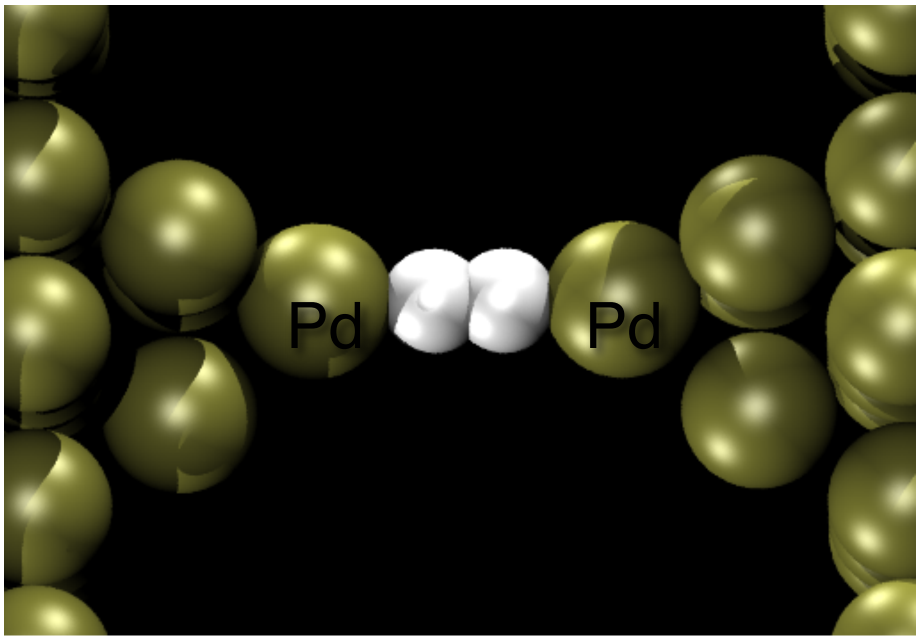
Quantum mechanical properties can be mapped through the use of two-dimensional electronic spectroscopy.

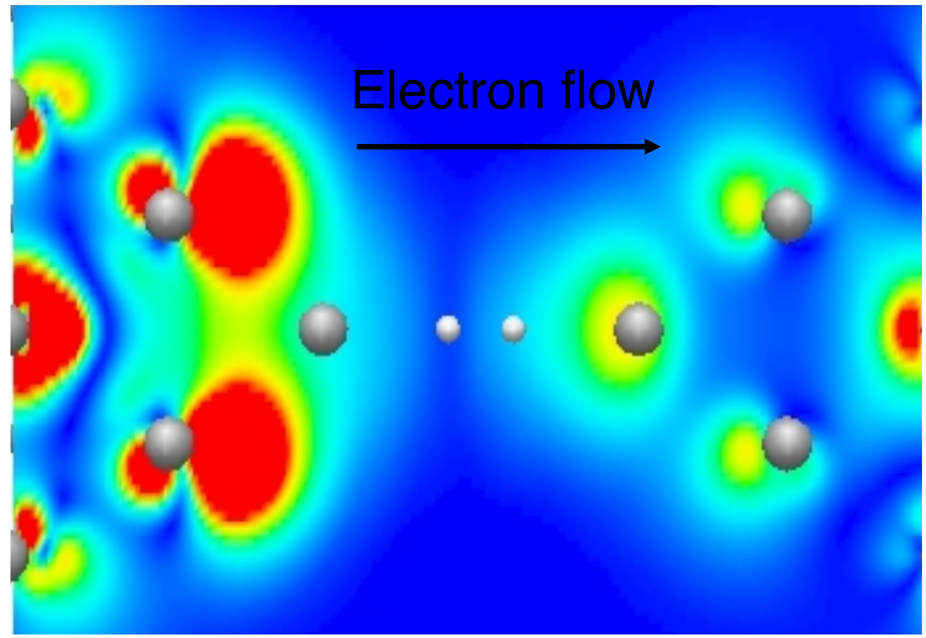
(Image courtesy of G. Engel)

Spectrum modeling and quantum dynamics simulations:

- Oscillations correlate with the quantum coherence in the system.
- Energy transfer pathways shown inside network of photosynthetic pigment-protein complexes.

Single Molecule Electrical Junctions





Shown at left is a single hydrogen molecule (in white) bridging palladium point contacts. At right, a density plot of the dominant transmitting electronic state reveals a significant reflection of charge at the left Pd contact, leading to a high resistance, consistent with recent experiments. (Red is high electronic density in the plot, blue is low.)

Steven Louie,
Marvin Cohen,
UC Berkeley
Jeff Neaton, LBNL
Molecular Foundry

Experimental Math: Discovering New Mathematical Results by Computer

- Compute various mathematical entities (limits, infinite series sums, definite integrals) to high precision, typically 100-1000 digits.
- Use algorithms such as PSLQ to recognize these entities in terms of well-known mathematical constants.
- When results are found experimentally, seek to find formal mathematical proofs of the discovered relations.

Many results have recently been found using this methodology, both in pure mathematics and in mathematical physics.

"If mathematics describes an objective world just like physics, there is no reason why inductive methods should not be applied in mathematics just the same as in physics." – Kurt Godel

Mathematics → Computer science → Scientific computing → Mathematics

The PSLQ Integer Relation Algorithm

Let (x_n) be a given vector of real numbers given to high precision. An integer relation algorithm finds integers (a_n) such that

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0$$

(or within some "epsilon" of zero).

At the present time the "PSLQ" algorithm of mathematician-sculptor Helaman Ferguson is the most widely used integer relation algorithm, although the "LLL" algorithm is also used. PSLQ was named one of ten "algorithms of the century" by *Computing in Science and Engineering*.

1. H. R. P. Ferguson, DHB and S. Arno, "Analysis of PSLQ, an integer relation finding algorithm," *Mathematics of Computation*, vol. 68, no. 225 (Jan 1999), pg. 351-369.
2. DHB and D. J. Broadhurst, "Parallel integer relation detection: Techniques and applications," *Mathematics of Computation*, vol. 70, no. 236 (Oct 2000), pg. 1719-1736.

Ferguson's PSLQ Algorithm Solves Problem Inspired by Ferguson's Sculpture

The complement of the figure-eight knot, when viewed in hyperbolic space, has finite volume

2.02988321281930725004 2…

Recently physicist David Broadhurst found, using PSLQ, that this constant is given by the formula:

$$V = \frac{\sqrt{3}}{9}\sum_{n=0}^{\infty}\frac{(-1)^n}{27^n}\left(\frac{18}{(6n+1)^2}-\frac{18}{(6n+2)^2}\right.$$
$$\left.-\frac{24}{(6n+3)^2}-\frac{6}{(6n+4)^2}+\frac{2}{(6n+5)^2}\right)$$

J. M. Borwein and D. H. Bailey, *Mathematics by Experiment*, A.K. Peters, 2004, pg. 53.

PSLQ Discovery: The BBP Formula for Pi

In 1996, this new formula for π was found using a PSLQ program:

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right)$$

This formula permits one to compute binary (or hexadecimal) digits of π beginning at an arbitrary starting position, using a very simple scheme that can run on any system, using only standard 64-bit or 128-bit arithmetic.

Recently it was proven that no base-*n* formulas of this type exist for π, except $n = 2^m$.

1. DHB, P. B. Borwein and S. Plouffe, "On the rapid computation of various polylogarithmic constants," *Mathematics of Computation*, vol. 66, no. 218 (Apr 1997), pg. 903-913.
2. J. M. Borwein, W. F. Galway and D. Borwein, "Finding and excluding b-ary Machin-type BBP formulae," *Canadian Journal of Mathematics*, vol. 56 (2004), pg 1339-1342.

Some Other New BBP-Type Formulas Discovered Using PSLQ

$$\pi^2 = \frac{1}{8}\sum_{k=0}^{\infty}\frac{1}{64^k}\left(\frac{144}{(6k+1)^2}-\frac{216}{(6k+2)^2}-\frac{72}{(6k+3)^2}-\frac{54}{(6k+4)^2}+\frac{9}{(6k+5)^2}\right)$$

$$\pi^2 = \frac{2}{27}\sum_{k=0}^{\infty}\frac{1}{729^k}\left(\frac{243}{(12k+1)^2}-\frac{405}{(12k+2)^2}-\frac{81}{(12k+4)^2}-\frac{27}{(27k+5)^2}\right.$$
$$\left.-\frac{72}{(12k+6)^2}-\frac{9}{(12k+7)^2}-\frac{9}{(12k+8)^2}-\frac{5}{(12k+10)^2}+\frac{1}{(12k+11)^2}\right)$$

$$\zeta(3) = \frac{1}{1792}\sum_{k=0}^{\infty}\frac{1}{2^{12k}}\left(\frac{6144}{(24k+1)^3}-\frac{43008}{(24k+2)^3}+\frac{24576}{(24k+3)^3}+\frac{30720}{(24k+4)^3}\right.$$
$$-\frac{1536}{(24k+5)^3}+\frac{3072}{(24k+6)^3}+\frac{768}{(24k+7)^3}-\frac{3072}{(24k+9)^3}-\frac{2688}{(24k+10)^3}$$
$$-\frac{192}{(24k+11)^3}-\frac{1536}{(24k+12)^3}-\frac{96}{(24k+13)^3}-\frac{672}{(24k+14)^3}-\frac{384}{(24k+15)^3}$$
$$+\frac{24}{(24k+17)^3}+\frac{48}{(24k+18)^3}-\frac{12}{(24k+19)^3}+\frac{120}{(24k+20)^3}+\frac{48}{(24k+21)^3}$$
$$\left.-\frac{42}{(24k+22)^3}+\frac{3}{(24k+23)^3}\right)$$

where ζ is the Riemann zeta function.

DHB, “A compendium of BBP-type formulas,” 2010, http://crd.lbl.gov/~dhbailey/dhbpapers/bbp-formulas.pdf.

A Connection Between BBP Formulas and Digit Randomness

Let {} denote fractional part. Consider the sequence defined by $x_0 = 0$,

$$x_n = \left\{ 16 x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21} \right\}$$

We showed that π is 16-normal (“random” base-16 digits in a certain sense) if and only if this sequence is equidistributed in the unit interval.

Further, we proved that the following mathematical constant is 2-normal:

$$\begin{aligned} \alpha_{2,3} &= \sum_{n=1}^{\infty} \frac{1}{3^n 2^{3^n}} \\ &= 0.041883680831502985071252898624571682426096\ldots_{10} \\ &= 0.0ab8e38f684bda12f684bf35ba781948b0fcd6e9e0\ldots_{16} \end{aligned}$$

1. D. H. Bailey and R. E. Crandall, "On the Random Character of Fundamental Constant Expansions," *Experimental Mathematics*, vol. 10, no. 2 (Jun 2001), pg. 175-190.
2. D. H. Bailey and R. E. Crandall, “Random Generators and Normal Numbers,” Experimental Mathematics, vol. 11, no. 4 (2002), pg. 527-546.
3. D. H. Bailey, "A Pseudo-Random Number Generator Based on Normal Numbers," manuscript, Dec 2004, http://crd.lbl.gov/~dhbailey/dhbpapers/normal-random.pdf.

Ising Integrals from Mathematical Physics

We recently applied our methods to study three classes of integrals that arise in the Ising theory of mathematical physics – D_n and two others:

$$C_n := \frac{4}{n!}\int_0^\infty \cdots \int_0^\infty \frac{1}{\left(\sum_{j=1}^n (u_j + 1/u_j)\right)^2} \frac{du_1}{u_1} \cdots \frac{du_n}{u_n}$$

$$D_n := \frac{4}{n!}\int_0^\infty \cdots \int_0^\infty \frac{\prod_{i<j}\left(\frac{u_i-u_j}{u_i+u_j}\right)^2}{\left(\sum_{j=1}^n (u_j + 1/u_j)\right)^2} \frac{du_1}{u_1} \cdots \frac{du_n}{u_n}$$

$$E_n = 2\int_0^1 \cdots \int_0^1 \left(\prod_{1\le j<k\le n} \frac{u_k - u_j}{u_k + u_j}\right)^2 dt_2\, dt_3 \cdots dt_n$$

where in the last line $u_k = t_1\, t_2 \ldots t_k$.

DHB, J. M. Borwein and R. E. Crandall, "Integrals of the Ising class," *Journal of Physics A: Mathematical and General*, vol. 39 (2006), pg. 12271-12302.

Limiting Value of C_n: What Is This Number?

The C_n numerical values appear to approach a limit. For instance,

$$C_{1024} = 0.630473503374386796122040192710878904354587078712732 34\ldots$$

What is this limit? We copied the first 50 digits of this numerical value into the online Inverse Symbolic Calculator (ISC):
http://ddrive.cs.dal.ca/~isc or **http://carma-lx1.newcastle.edu.au:8087/**

The result was:

$$\lim_{n\to\infty} C_n \quad = \quad 2e^{-2\gamma}$$

where gamma denotes Euler's constant. Finding this limit led us to the asymptotic expansion and made it clear that the integral representation of C_n is fundamental.

Other Ising Integral Evaluations Found Using PSLQ Calculations

$$
\begin{aligned}
D_2 &= 1/3 \\
D_3 &= 8 + 4\pi^2/3 - 27\,\mathrm{L}_{-3}(2) \\
D_4 &= 4\pi^2/9 - 1/6 - 7\zeta(3)/2 \\
E_2 &= 6 - 8\log 2 \\
E_3 &= 10 - 2\pi^2 - 8\log 2 + 32\log^2 2 \\
E_4 &= 22 - 82\zeta(3) - 24\log 2 + 176\log^2 2 - 256(\log^3 2)/3 \\
&\quad +16\pi^2\log 2 - 22\pi^2/3 \\
E_5 &\stackrel{?}{=} 42 - 1984\,\mathrm{Li}_4(1/2) + 189\pi^4/10 - 74\zeta(3) - 1272\zeta(3)\log 2 \\
&\quad +40\pi^2\log^2 2 - 62\pi^2/3 + 40(\pi^2\log 2)/3 + 88\log^4 2 \\
&\quad +464\log^2 2 - 40\log 2
\end{aligned}
$$

where ζ is the Riemann zeta function and $\mathrm{Li}_n(x)$ is the polylog function. D_2, D_3 and D_4 were originally provided to us by mathematical physicist Craig Tracy, who hoped that our tools could help identify D_5.

Discovering AND Proving New Mathematical Formulas by Computer

For certain types of mathematical formulas, we can discover them using PSLQ, then prove them using the Wilf-Zeilberger algorithm.

Here is one example of a new mathematical result that was both discovered and proven by computer:

$$\sum_{n=0}^{\infty} \zeta(2n+2) x^{2n} = \frac{1 - \pi x \cot(\pi x)}{2x^2} = \sum_{k=1}^{\infty} \frac{1}{k^2 - x^2}$$

$$= 3 \sum_{k=1}^{\infty} \frac{1}{k^2 \binom{2k}{k} (1 - x^2/k^2)} \prod_{m=1}^{k-1} \left(\frac{1 - 4x^2/m^2}{1 - x^2/m^2} \right)$$

DHB, J. M. Borwein and D. M. Bradley, "Experimental Determination of Apery-Like Identities for Zeta(2n+2)," *Experimental Mathematics*, vol. 15 (2006), pg. 281-289.

Summary

- Scientific computing has vastly expanded in sophistication and power over the past 40 years, and is now widely regarded as a third mode of scientific discovery, after theory and experiment.
- Modern scientific computing is a multidisciplinary “symphony” involving scientists and engineers from many fields.
- The power of the leading-edge systems has closely followed Moore’s Law in an exponentially upward path, and no end is yet in sight.
- Computers have recently been applied into mathematical research, discovering many new previously unknown results, including a new formula for pi.

Major challenges:

- In order to utilize future systems efficiently, computer programs must possess and exhibit enormous parallelism (10^{10}-way or more).
- Exotic architectures (e.g., hybrid systems employing game processors) will present difficult programming and software challenges.
- Daunting electric power requirements projected for future systems will require innovation in hardware, software and applications.