Universality

Terence Tao
University of California Los Angeles
Trinity Mathematical Society Lecture
24 January, 2011

The problem of complexity

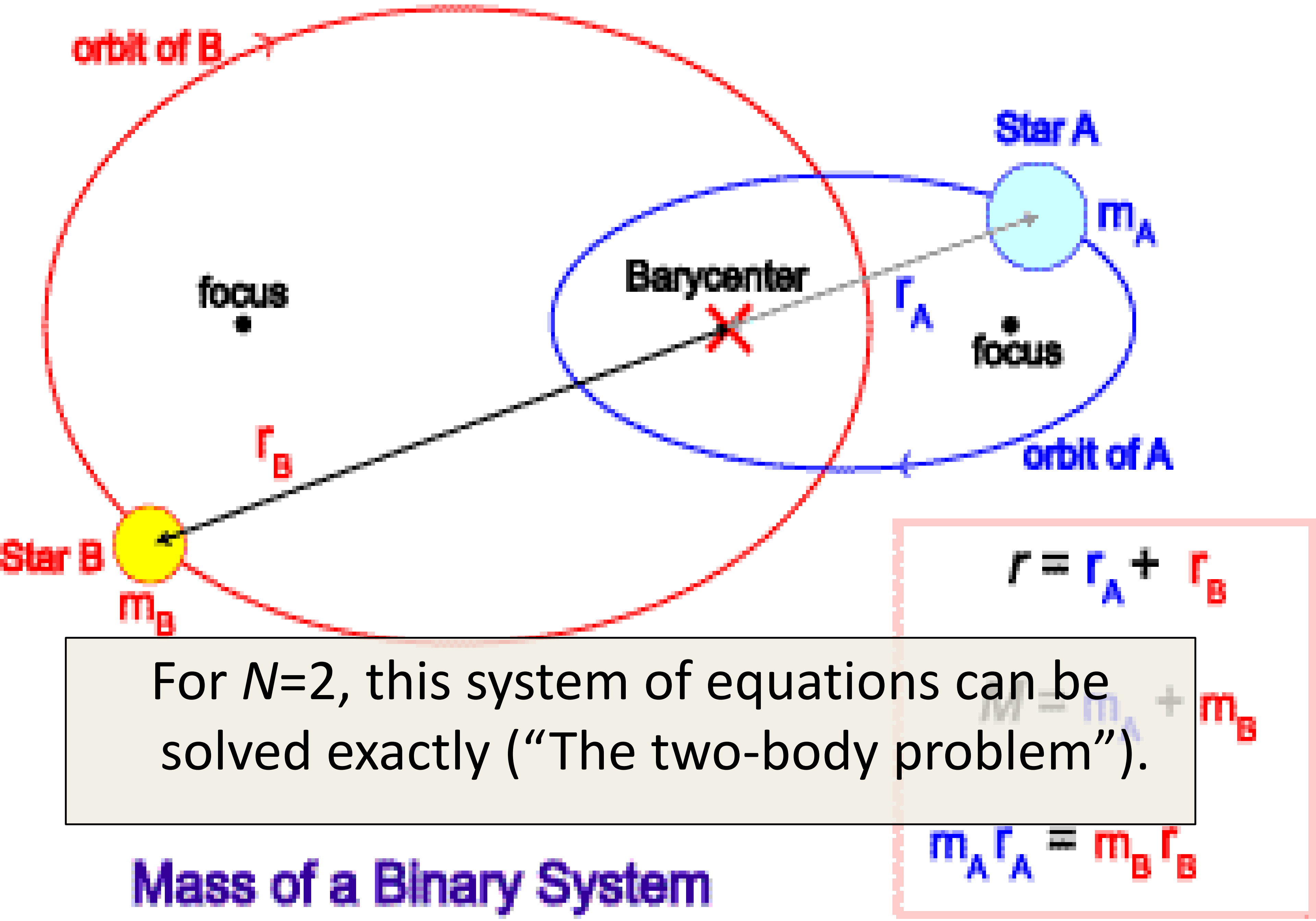
- The laws of physics let us (in principle, at least) predict the behaviour of systems of many interacting objects.
- However, in practice, the equations given by these laws become too complicated to solve exactly when the number *N* of objects becomes large.

Example: Newton's law of gravity

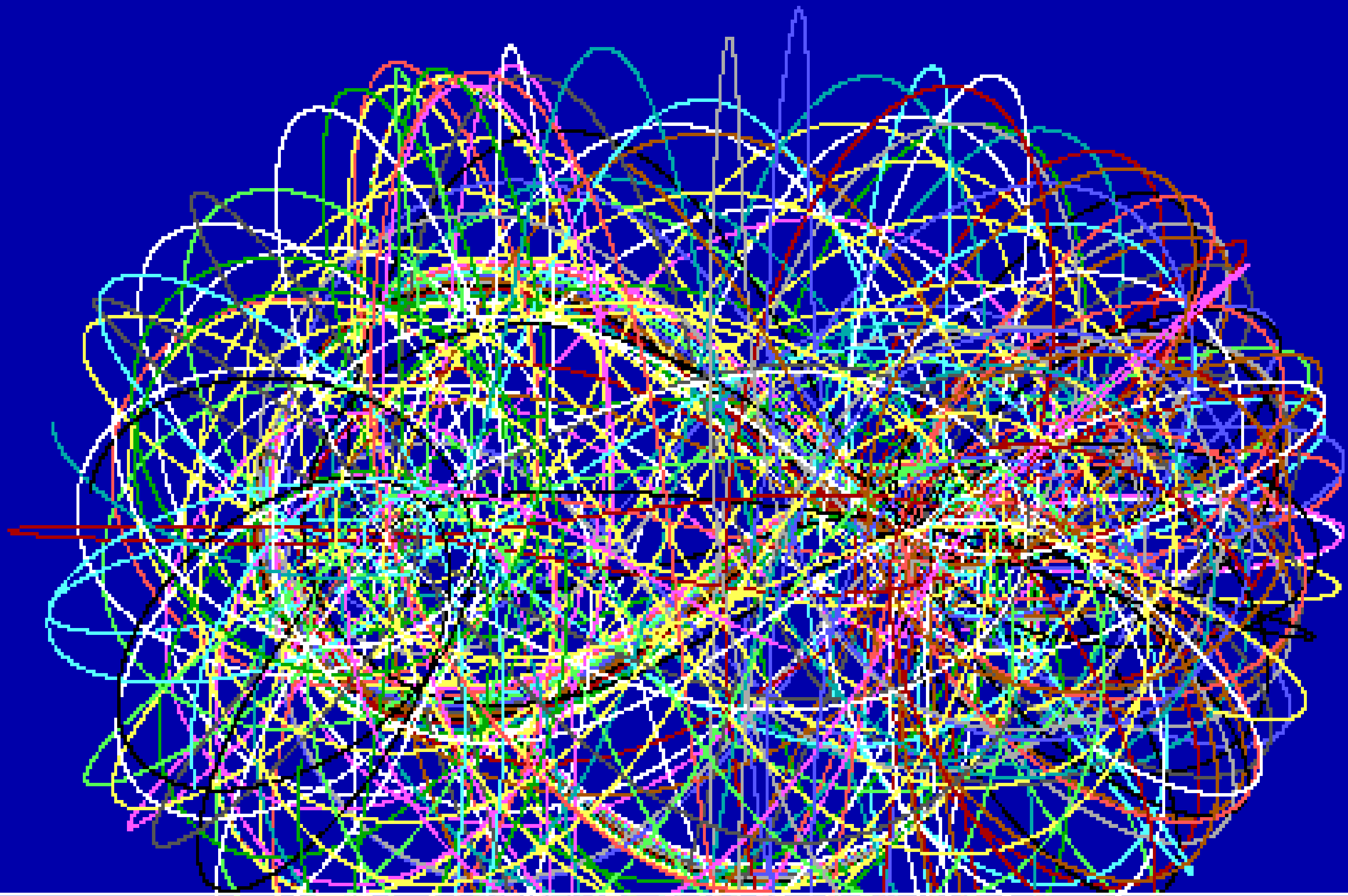
The equations of motion for the positions $x_1(t), \ldots, x_N(t)$ of N particles of masses $m_1, \ldots, m_N$ under Newtonian gravity is given by the system of equations

$$m_i \frac{d^2}{dt^2} x_i(t) = -\sum_{j \neq i} \frac{G m_i m_j}{|x_j(t) - x_i(t)|^3} (x_j(t) - x_i(t))$$

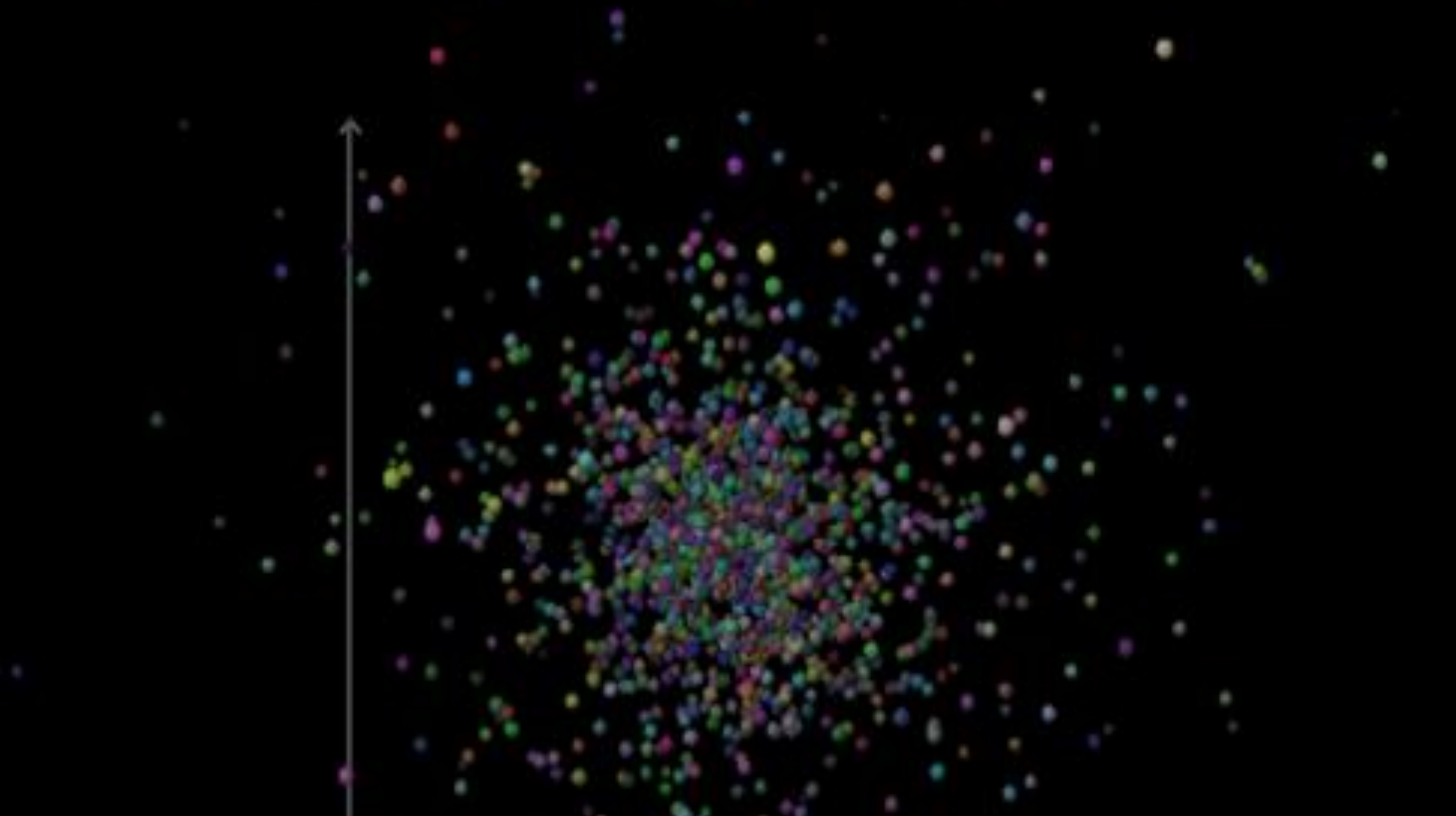
for $i=1,\ldots,N$.



For N=2, this system of equations can be solved exactly ("The two-body problem").

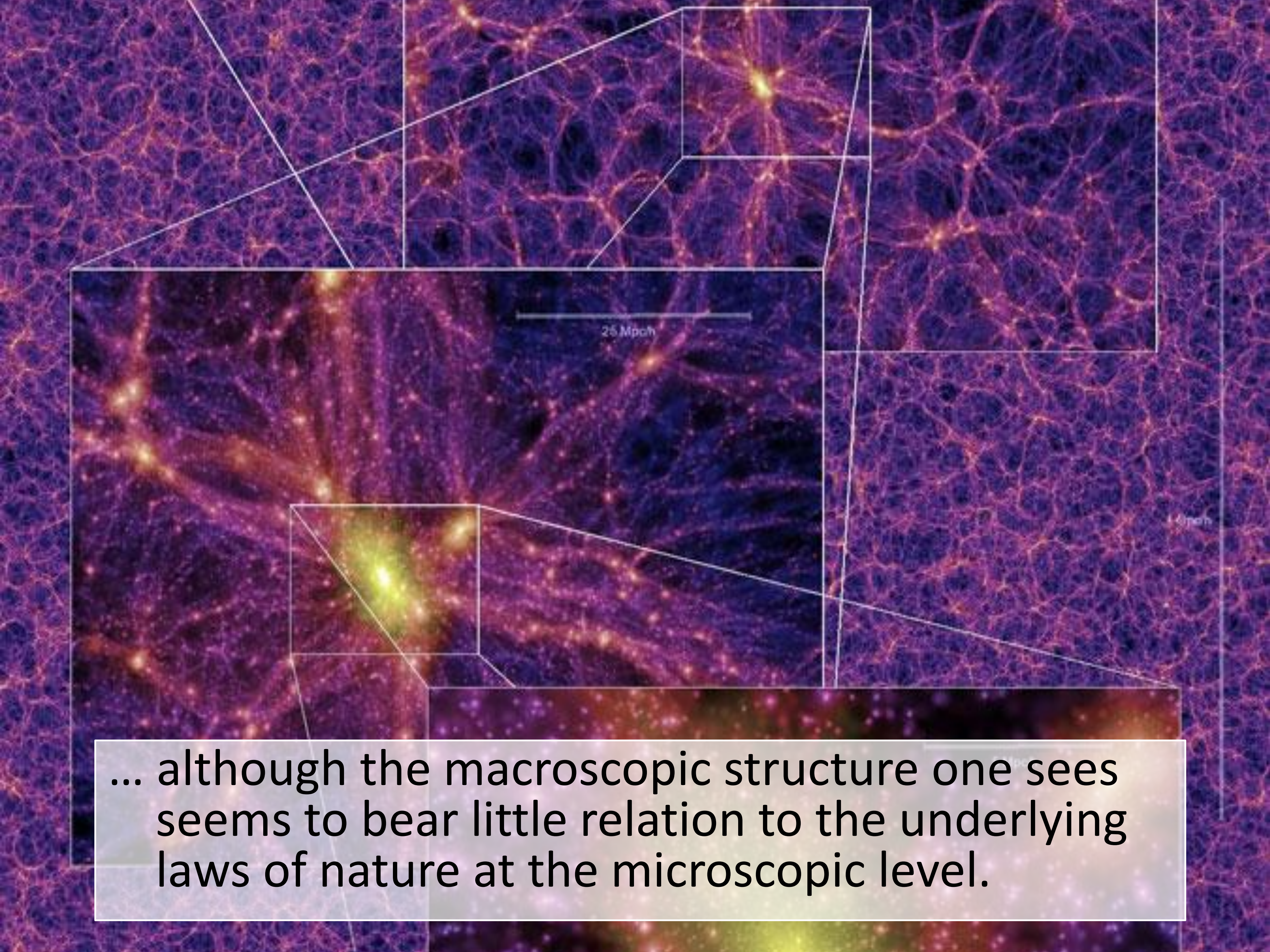


But even for N=3 -"the three-body problem" (the only problem to give Newton severe headaches) - there is no closed-form general solution to the equations.

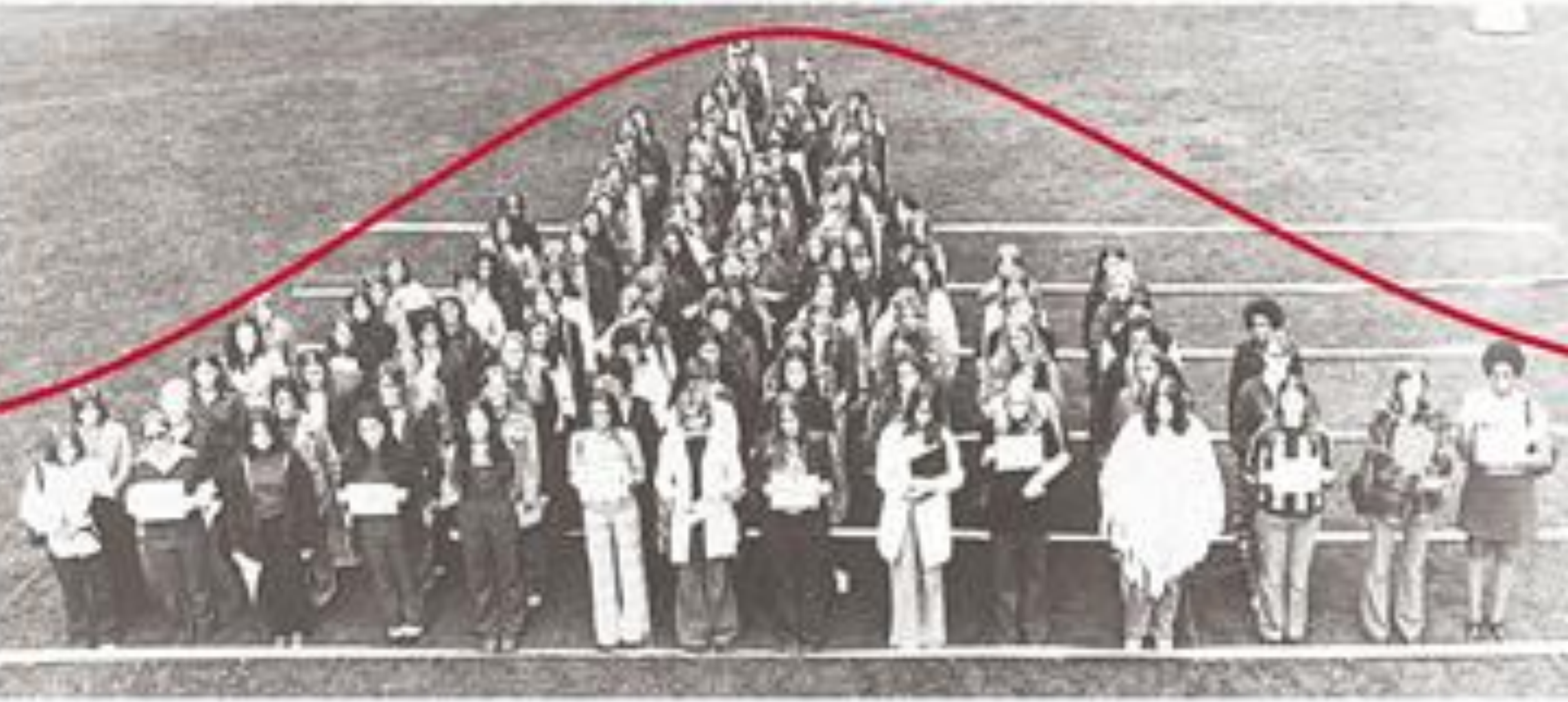


It would seem that the problem only gets worse for larger values of *N*.

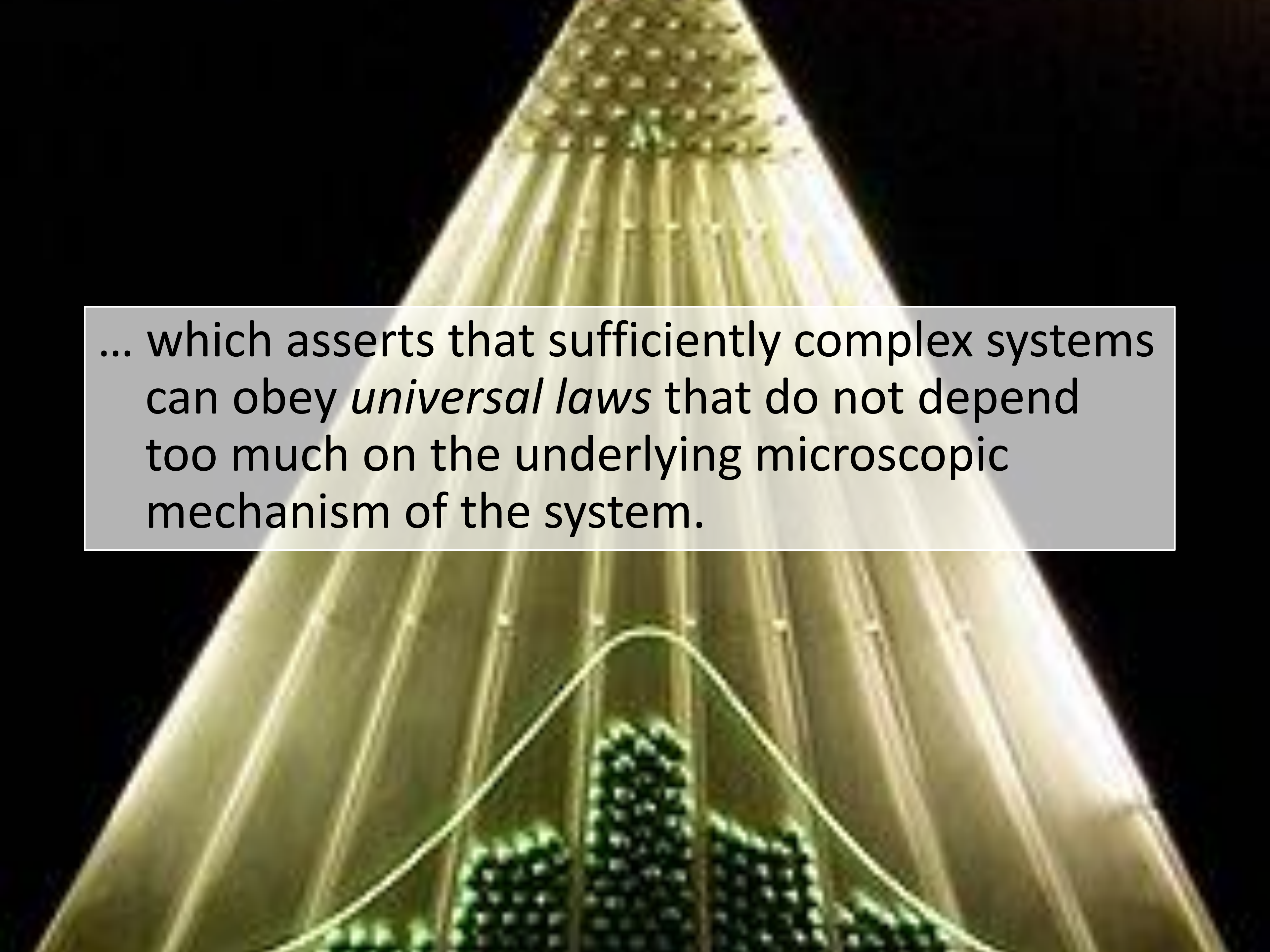
But when *N* becomes extremely large, solutions begin to exhibit regularity and predictability at macroscopic scales…



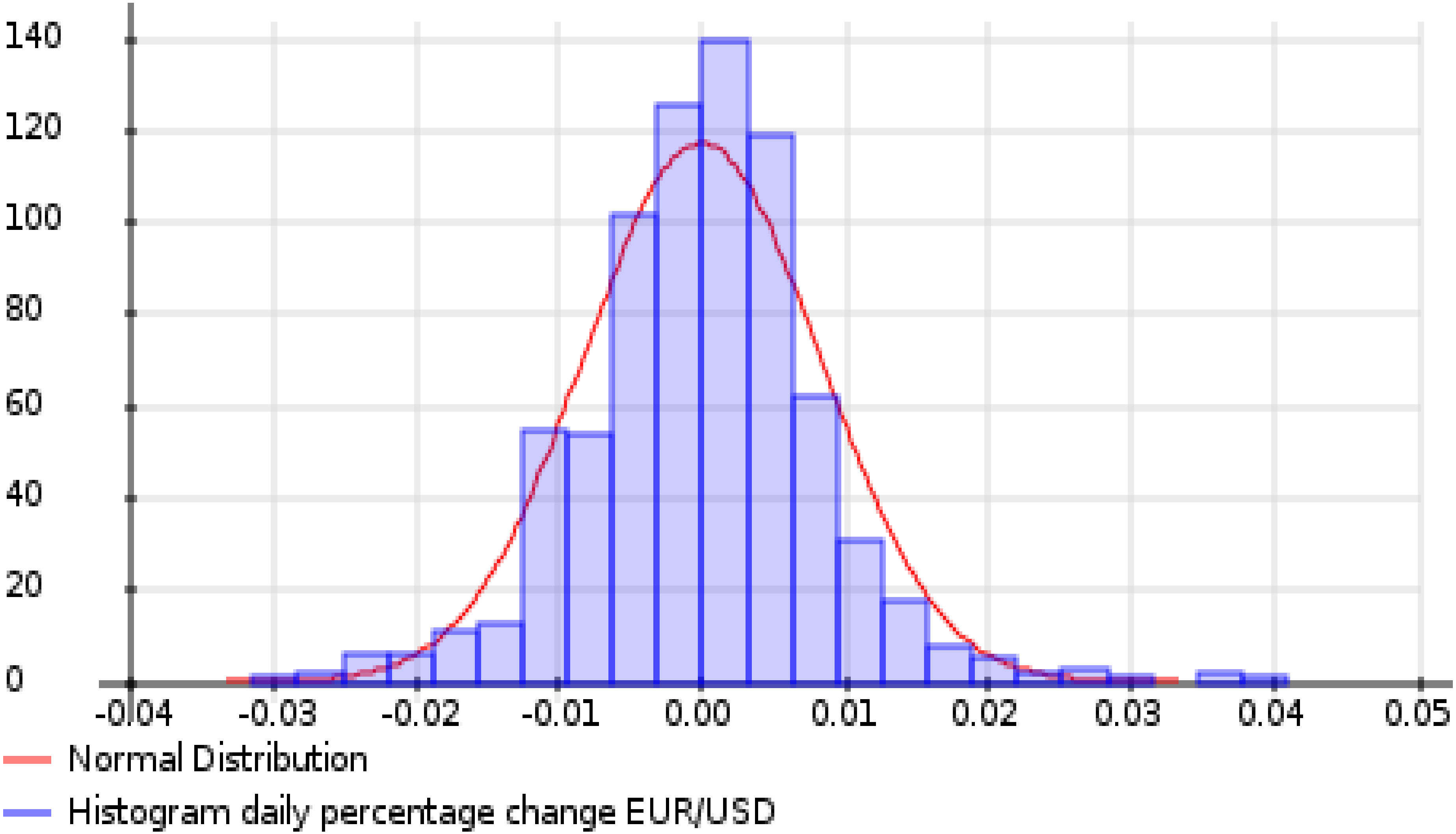
… although the macroscopic structure one sees seems to bear little relation to the underlying laws of nature at the microscopic level.



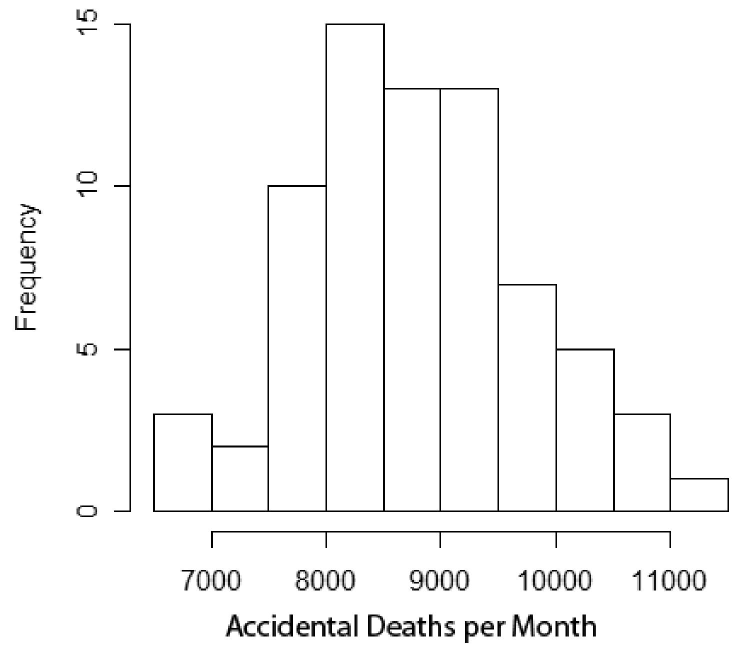
This is a manifestation of a general phenomenon known as *universality*…



… which asserts that sufficiently complex systems can obey *universal laws* that do not depend too much on the underlying microscopic mechanism of the system.



In some cases, we have a plausible mathematical justification for universality; but in other cases, the phenomenon is still only poorly understood.



For instance, it is an empirical fact that many real-life variables are approximately distributed according to a normal distribution (or "bell curve", or "gaussian distribution")

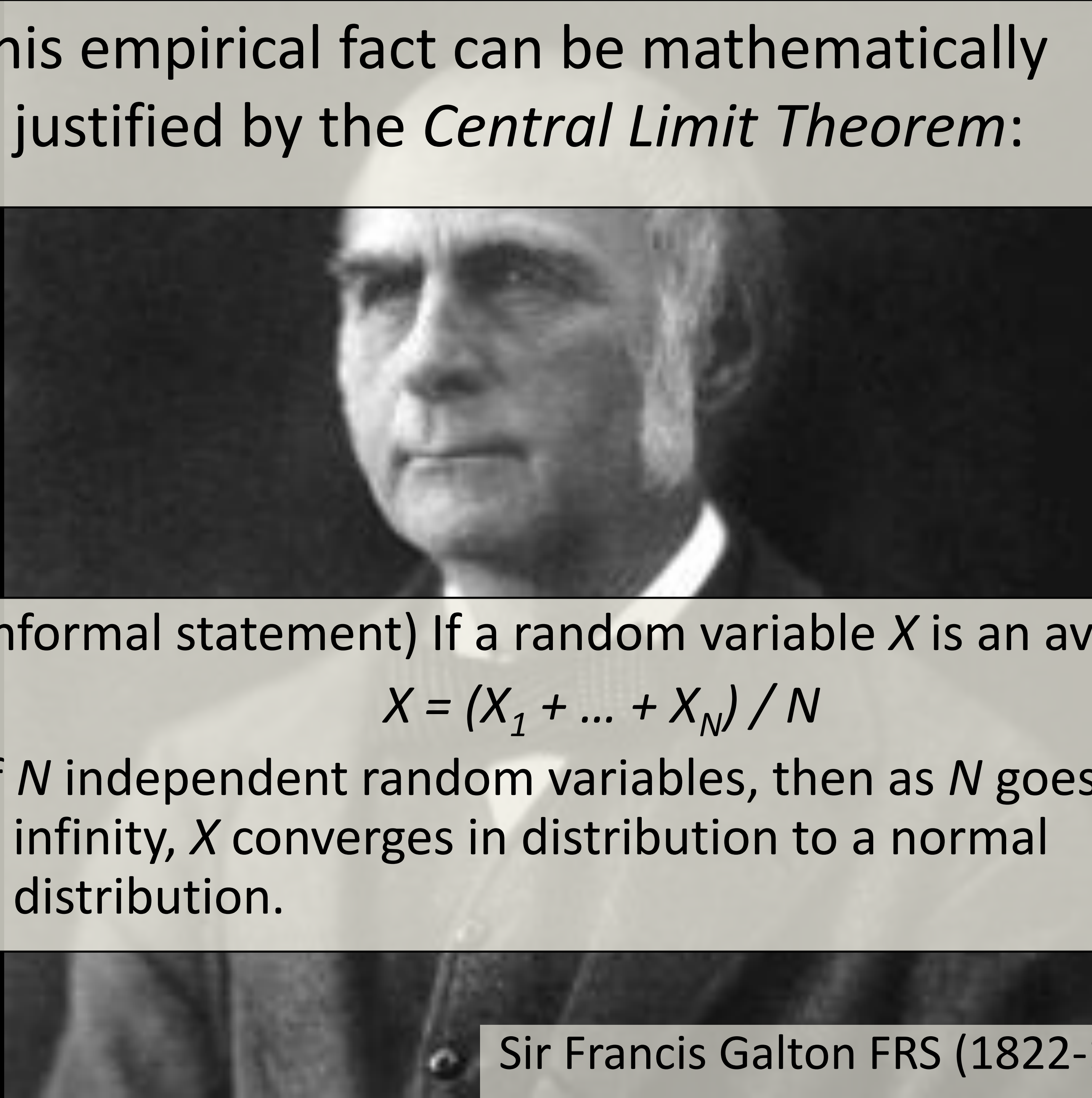
$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \, dx$$

This empirical fact can be mathematically justified by the *Central Limit Theorem*:

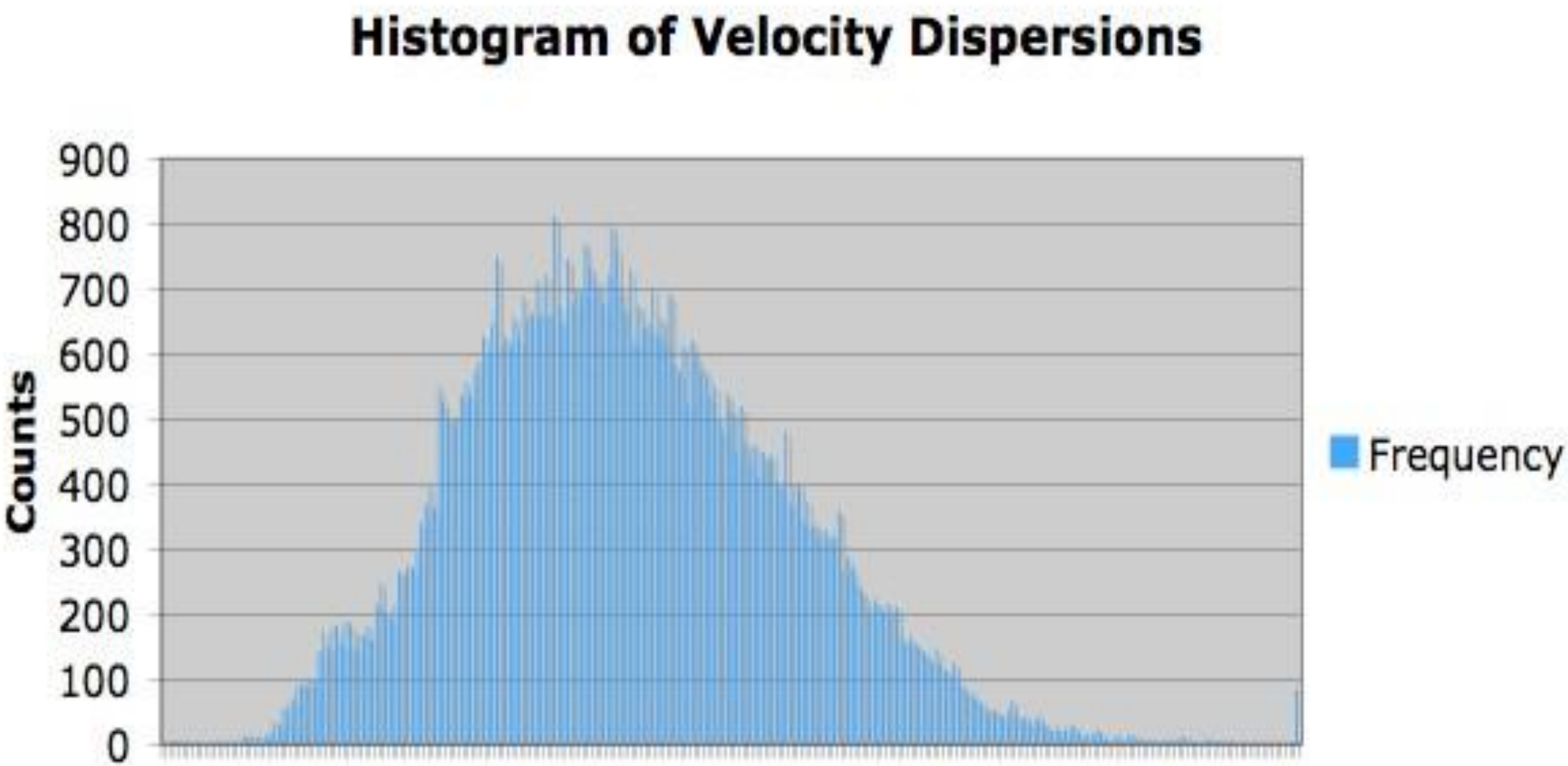
(Informal statement) If a random variable X is an average

$$X = (X_1 + \ldots + X_N) / N$$

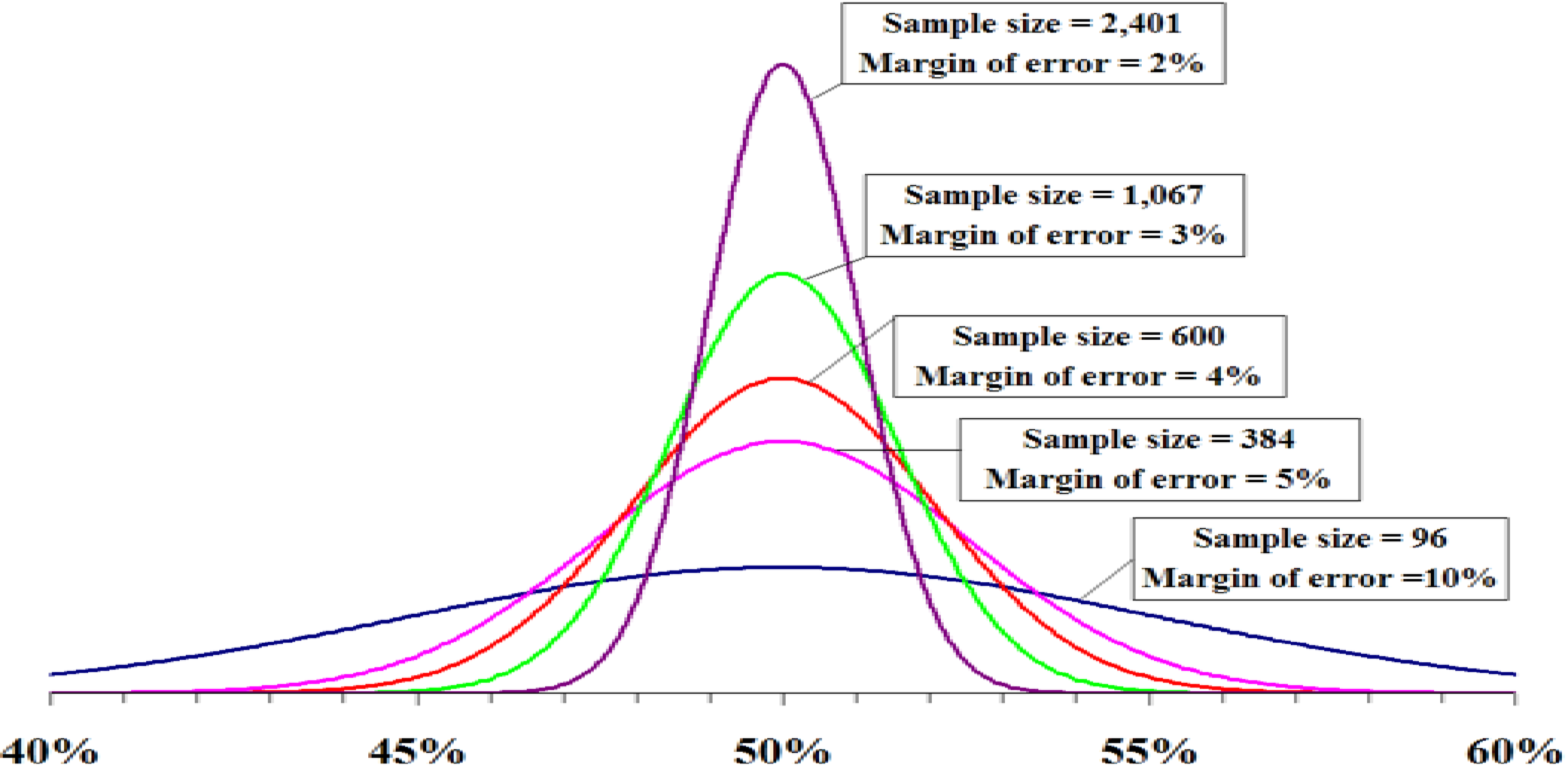
of N independent random variables, then as N goes to infinity, X converges in distribution to a normal distribution.



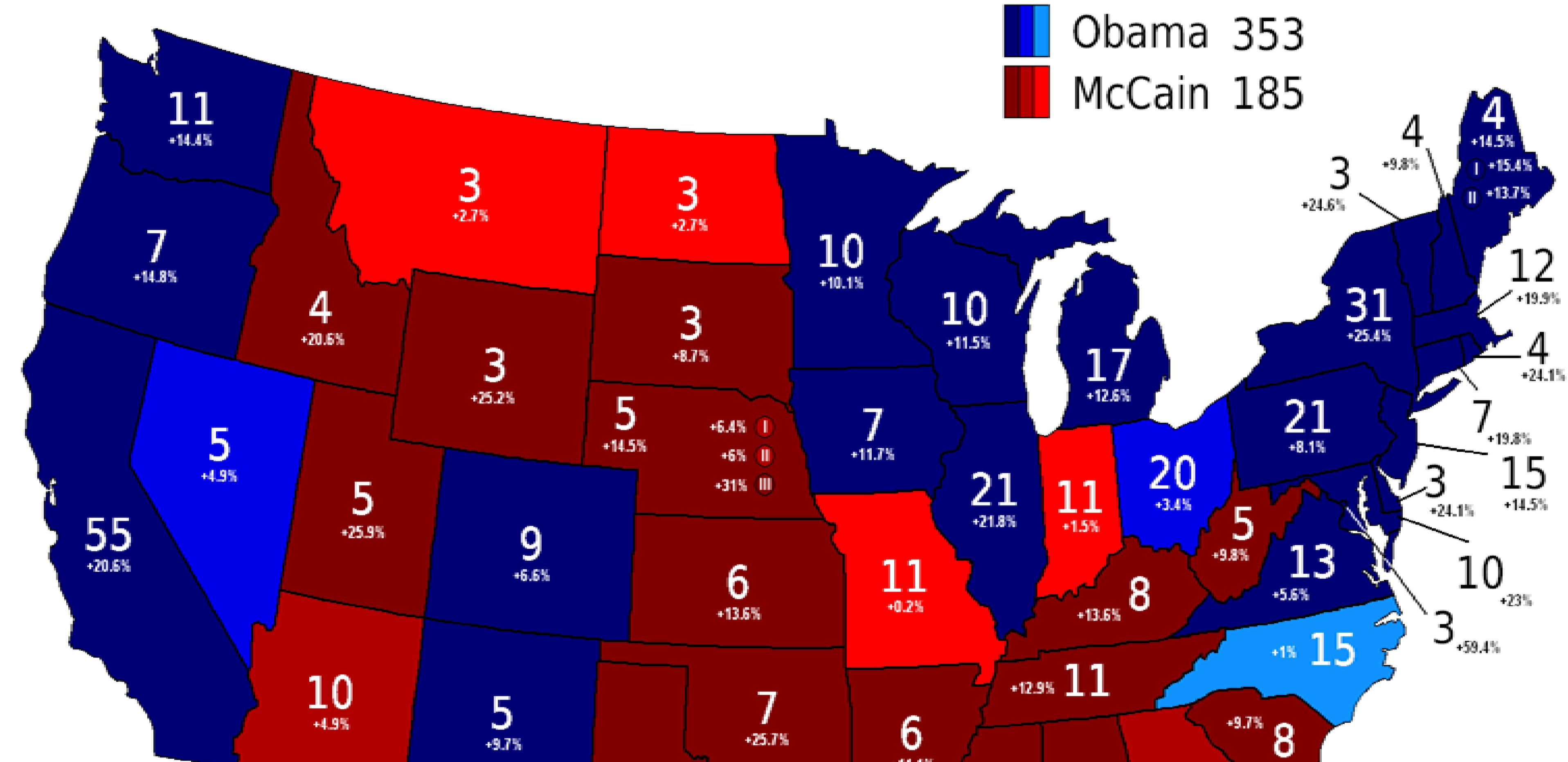
Sir Francis Galton FRS (1822-1911)



The central limit theorem explains the normally distributed behaviour of many empirical variables that can be plausibly modeled as additive averages of many independent factors.



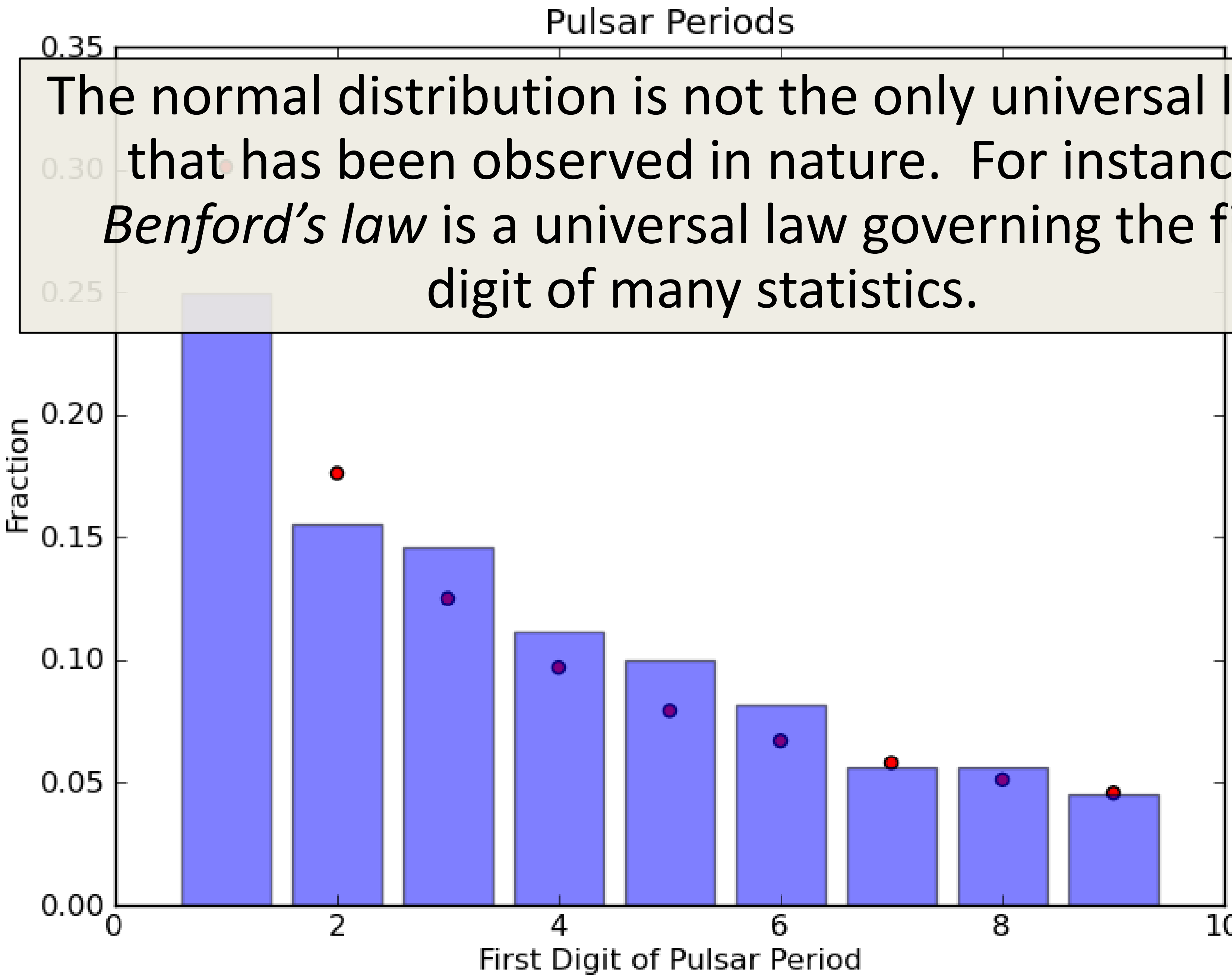
For instance, it can be used to show that the margin of error of a poll of *N* people is asymptotically proportional to $1 / N^{1/2}$ (under ideal conditions) – regardless of the size of the entire voting population.



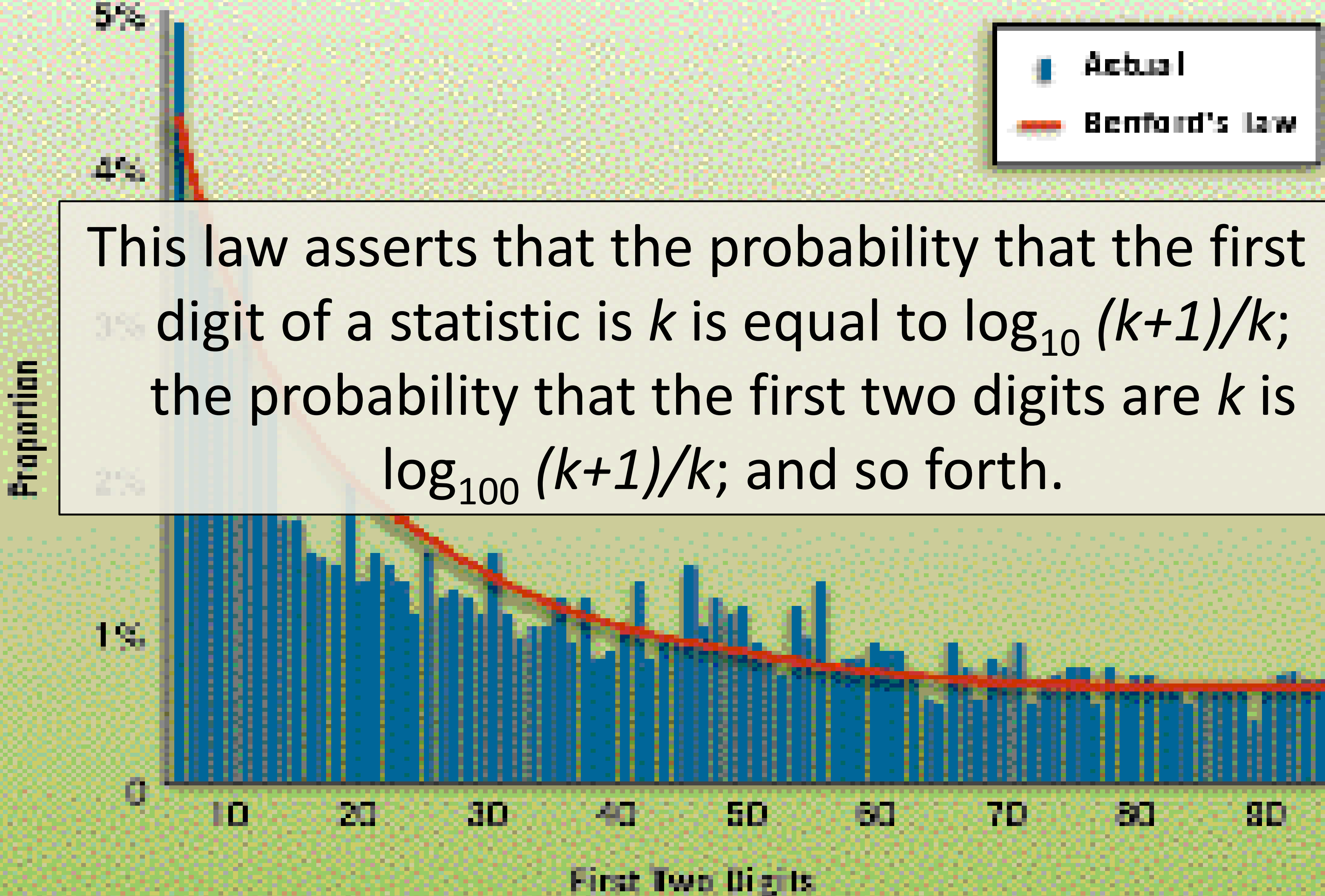
By combining together many polls, carefully weighted by reliability and accuracy, Nate Silver gave electoral predictions for the 2008 US presidential election on the election night…

… that correctly predicted the presidential election in 49 of 50 states (as well as all 35 of 35 Senate races).

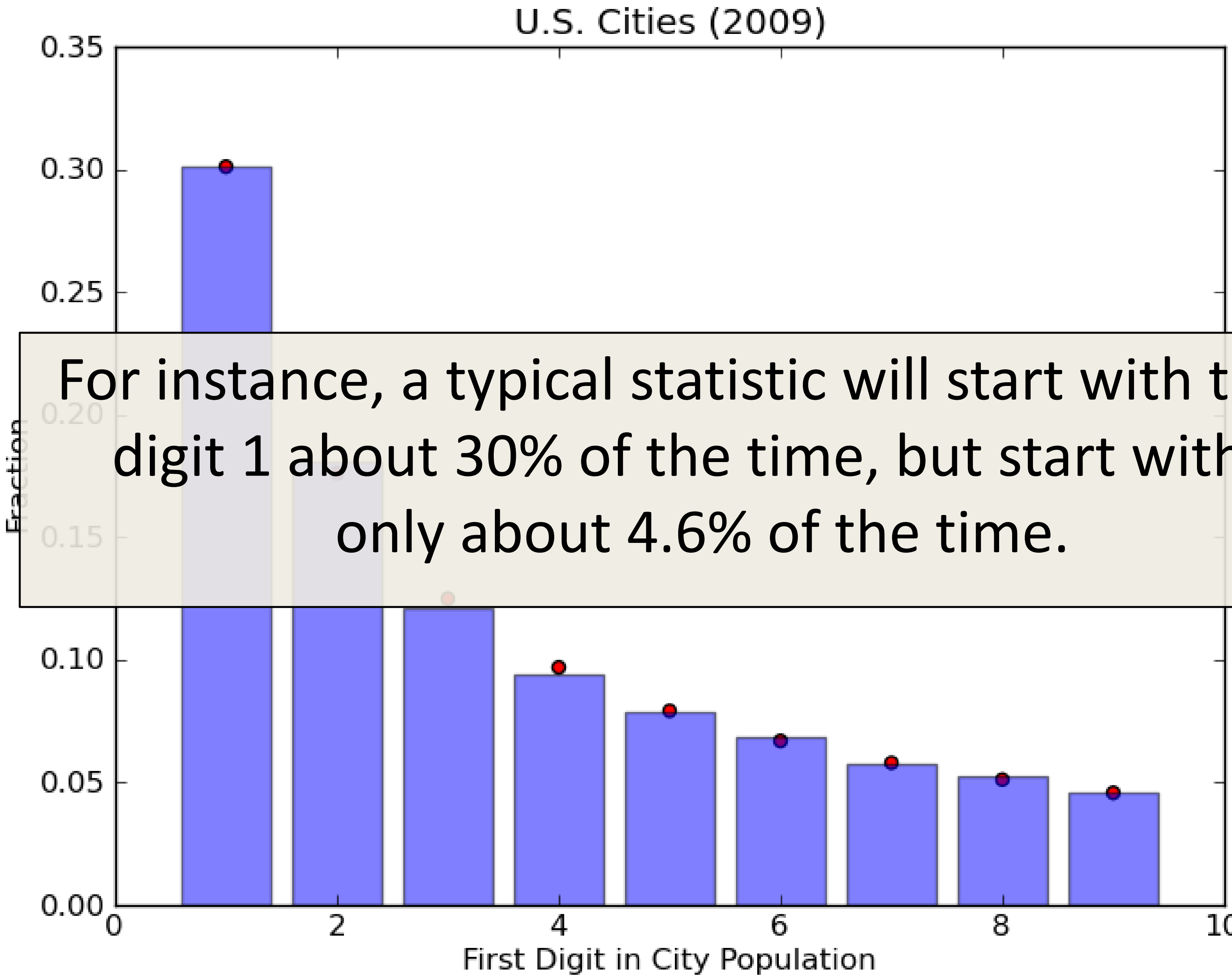
The normal distribution is not the only universal law that has been observed in nature. For instance, *Benford's law* is a universal law governing the first digit of many statistics.



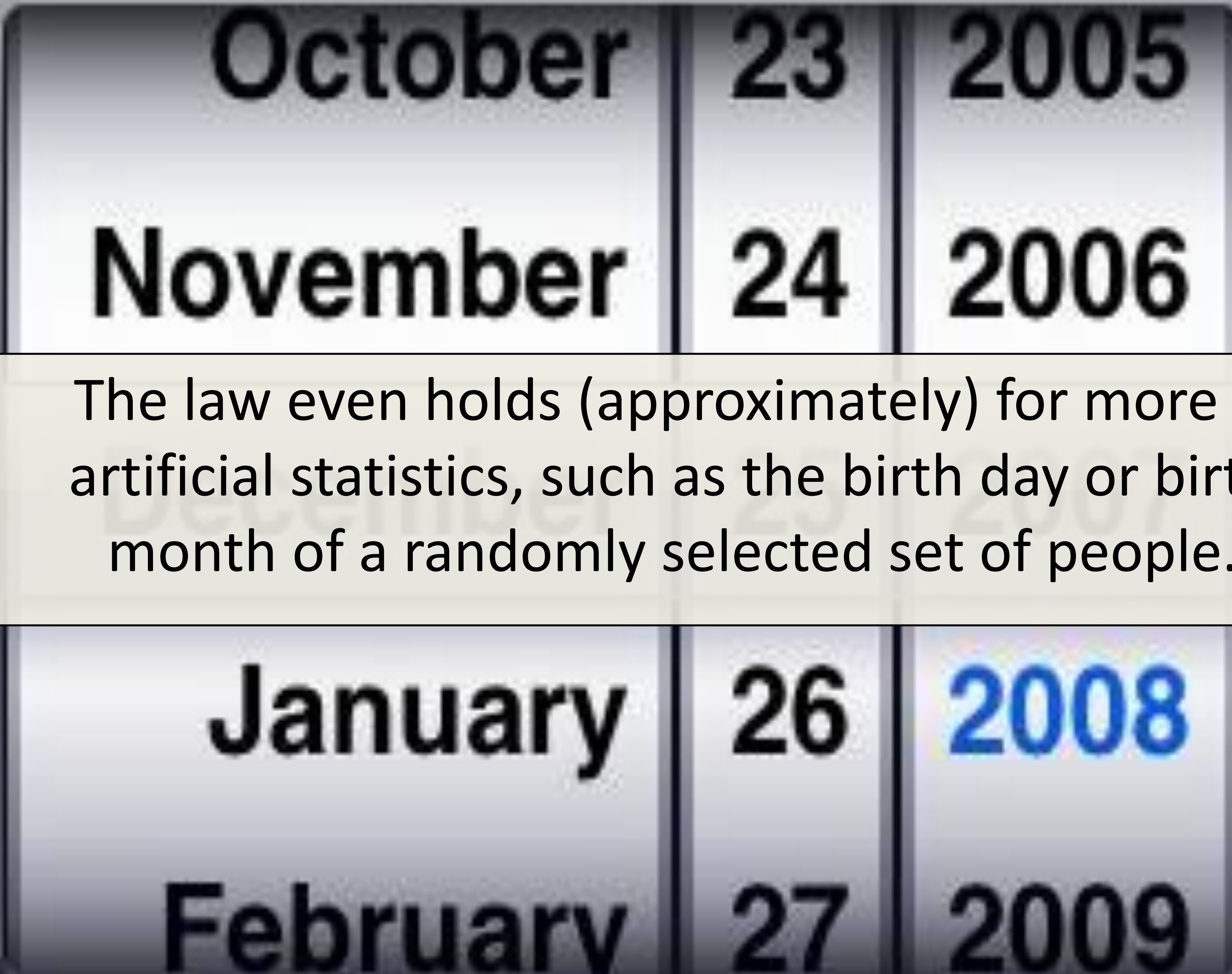
First two digits of a data set of accounts payable data



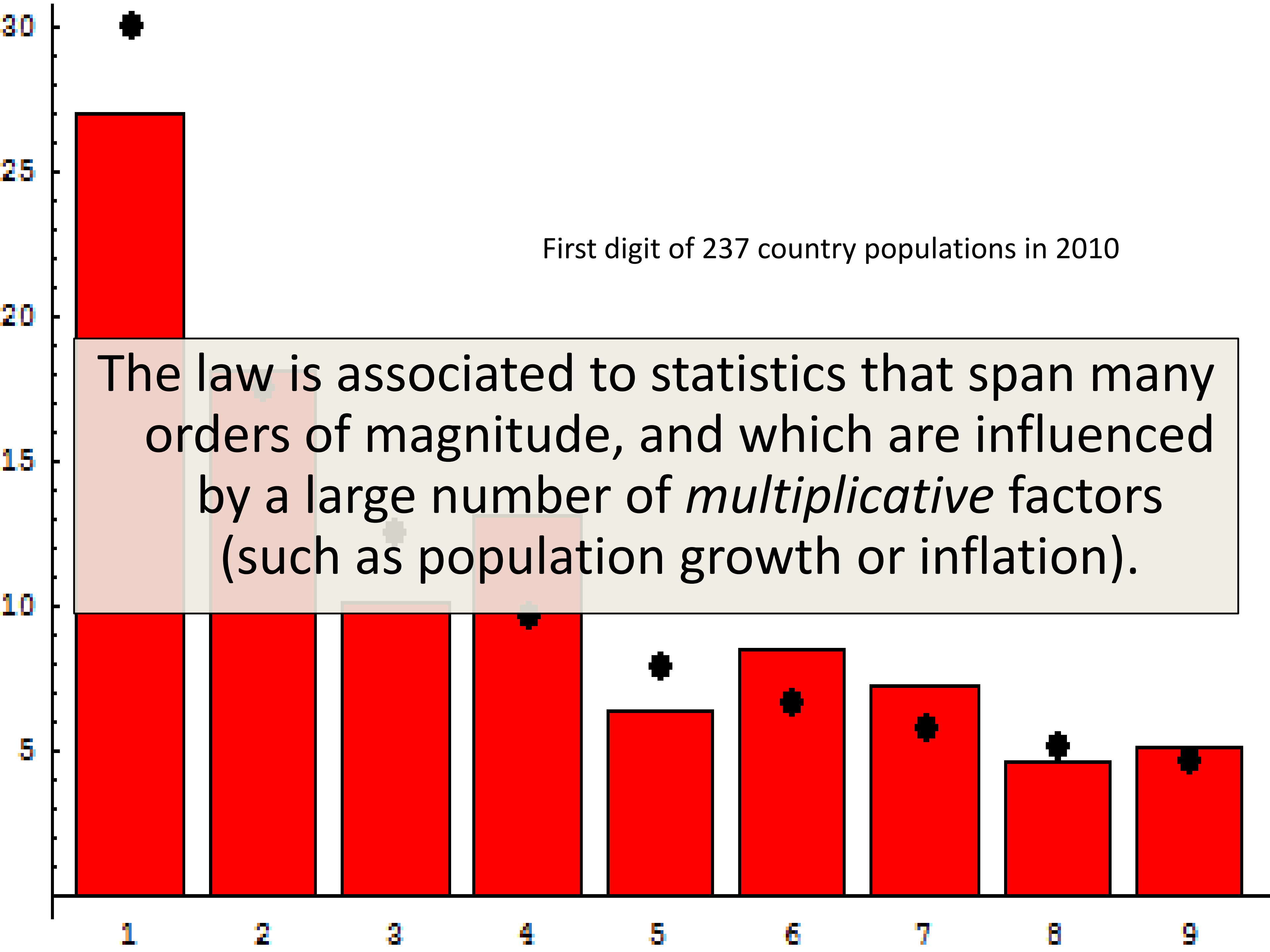
This law asserts that the probability that the first digit of a statistic is *k* is equal to $\log_{10} (k+1)/k$; the probability that the first two digits are *k* is $\log_{100} (k+1)/k$; and so forth.



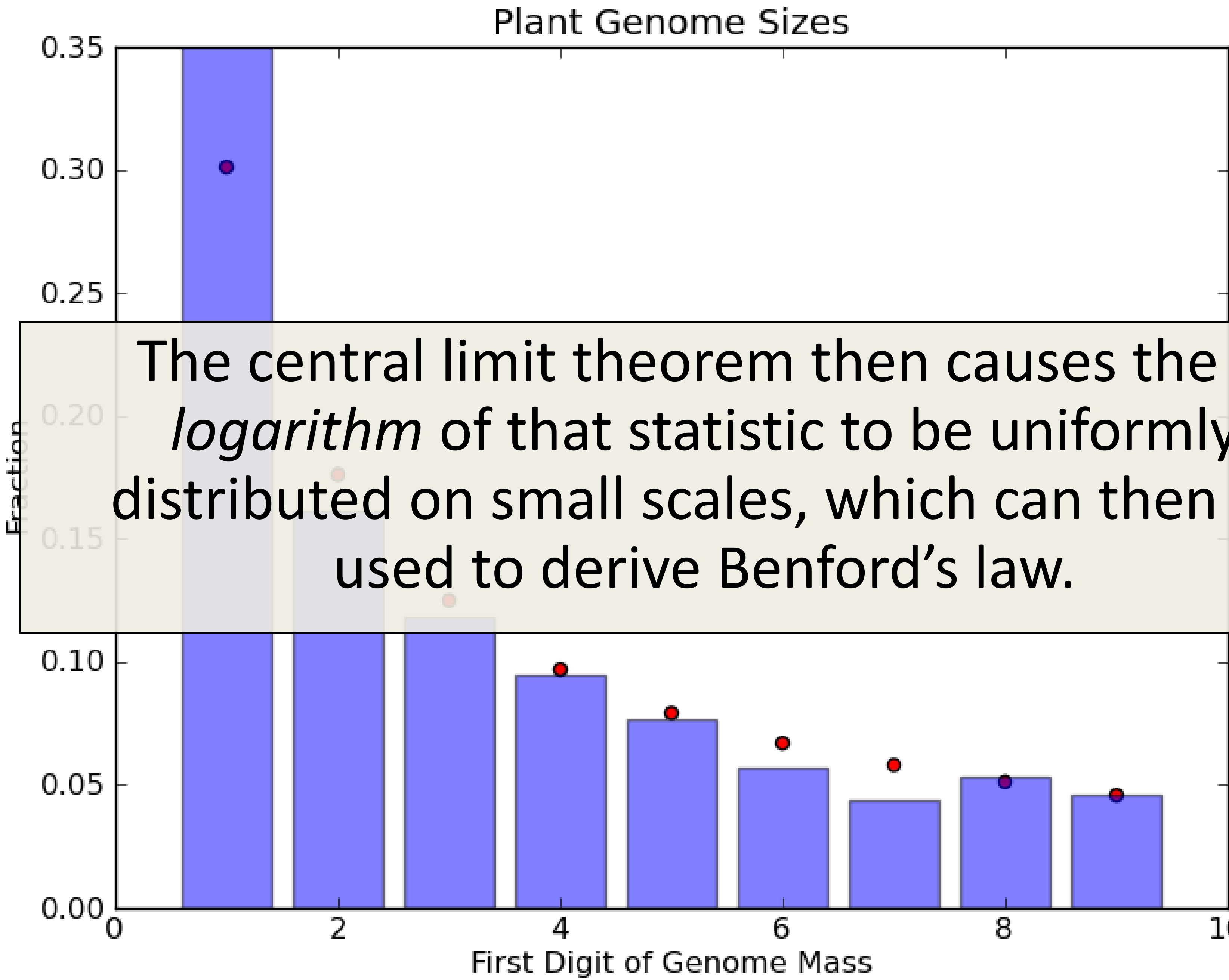
For instance, a typical statistic will start with the digit 1 about 30% of the time, but start with 9 only about 4.6% of the time.



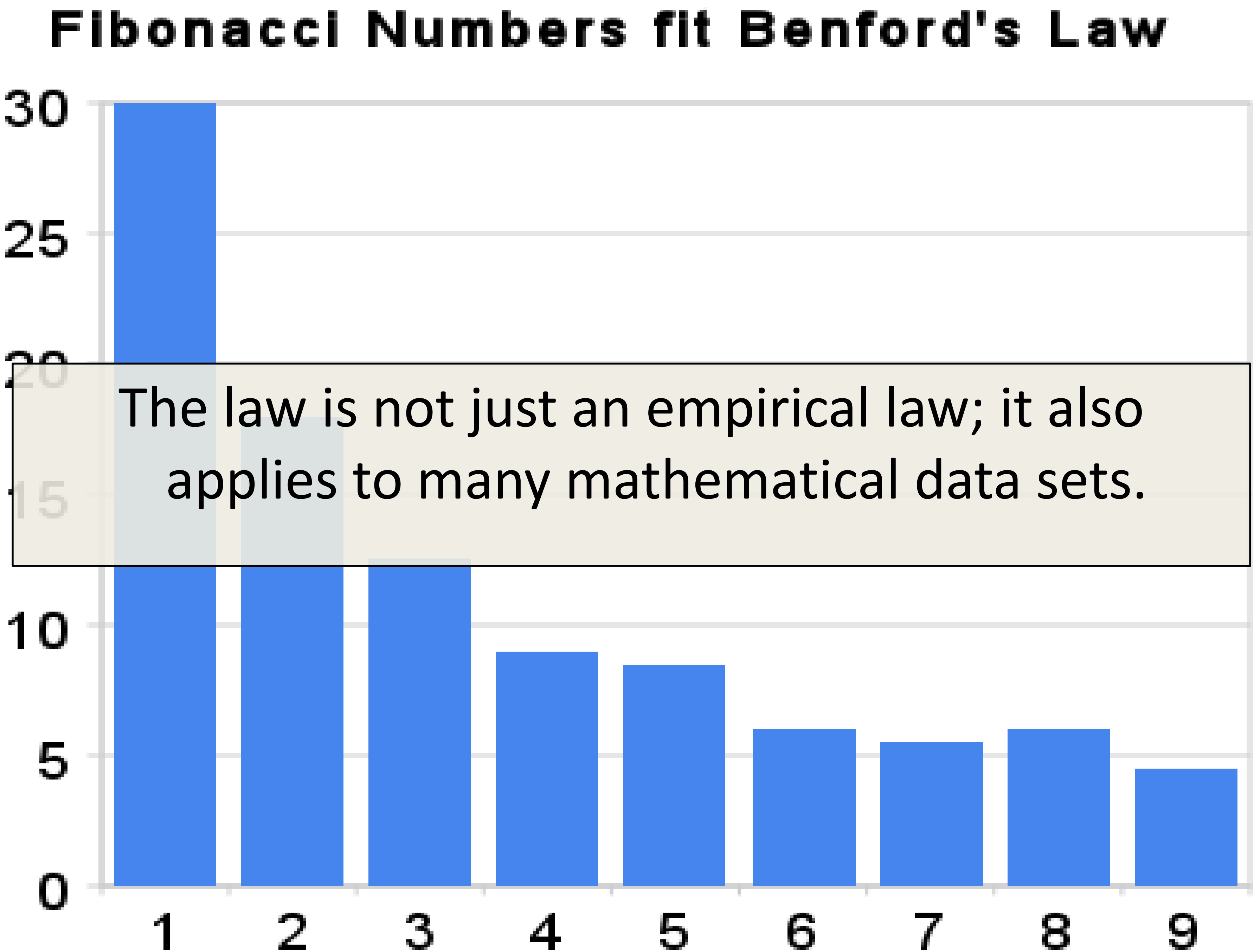
The law even holds (approximately) for more artificial statistics, such as the birth day or birth month of a randomly selected set of people.



The law is associated to statistics that span many orders of magnitude, and which are influenced by a large number of *multiplicative* factors (such as population growth or inflation).

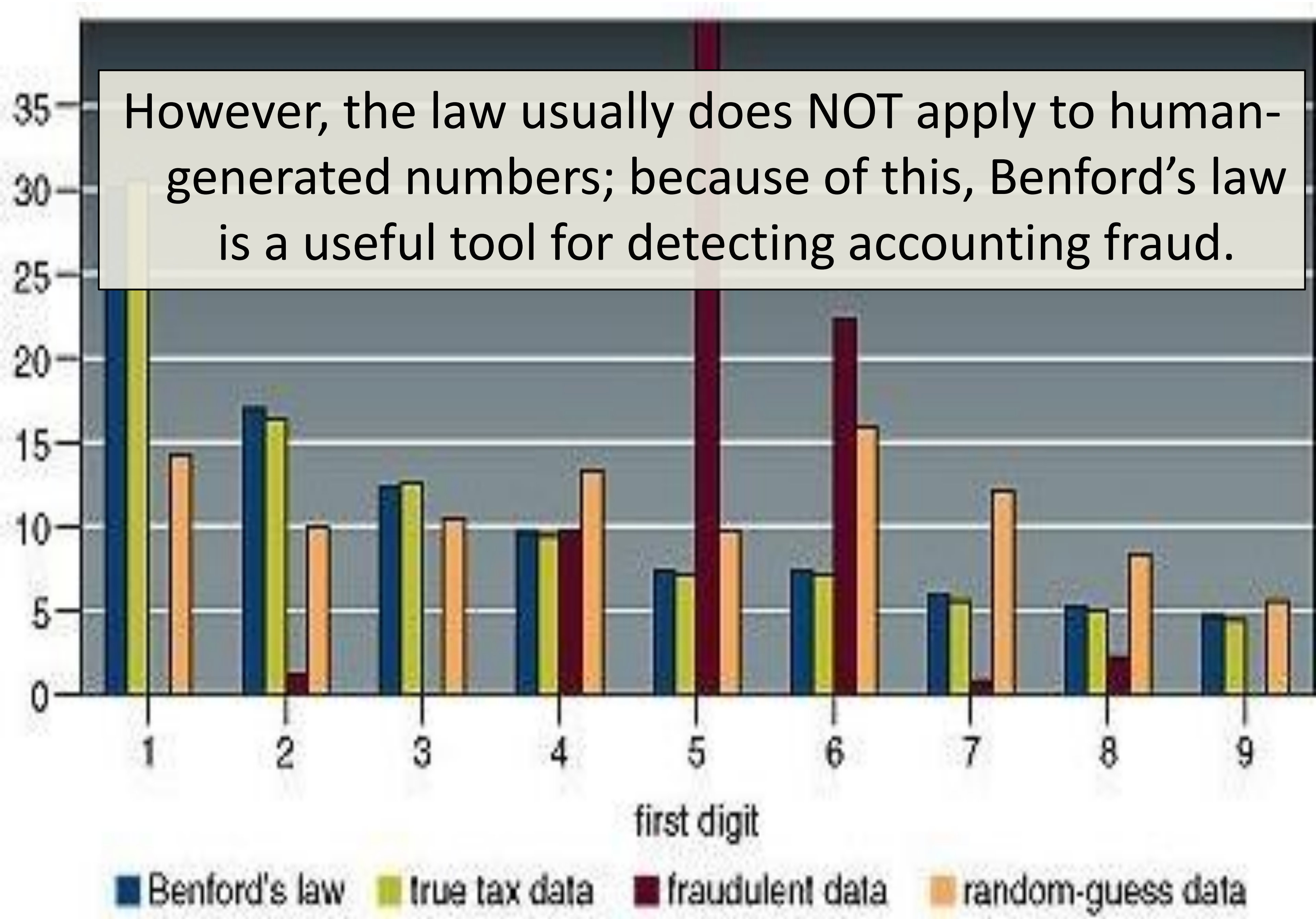


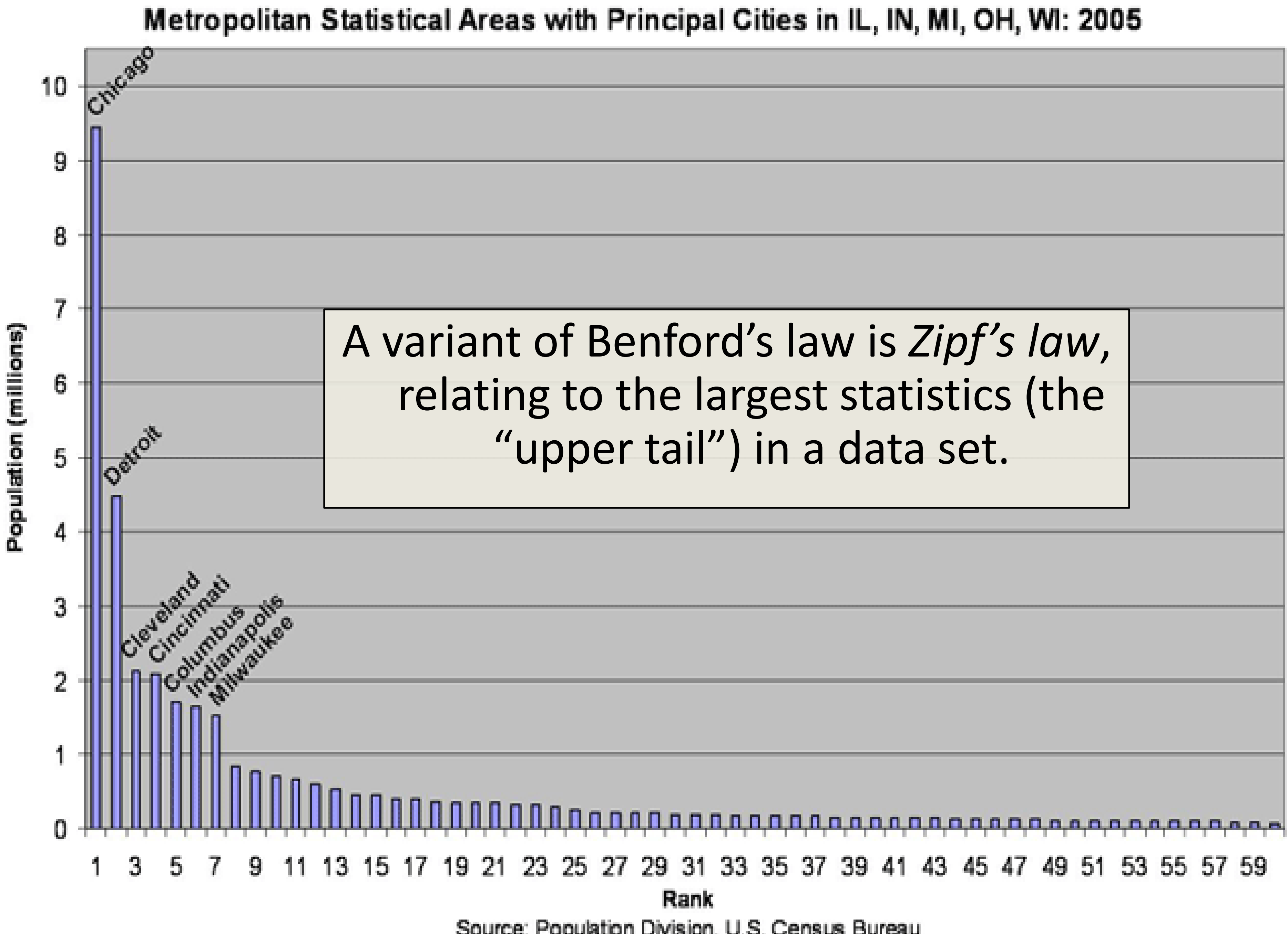
The central limit theorem then causes the *logarithm* of that statistic to be uniformly distributed on small scales, which can then be used to derive Benford's law.



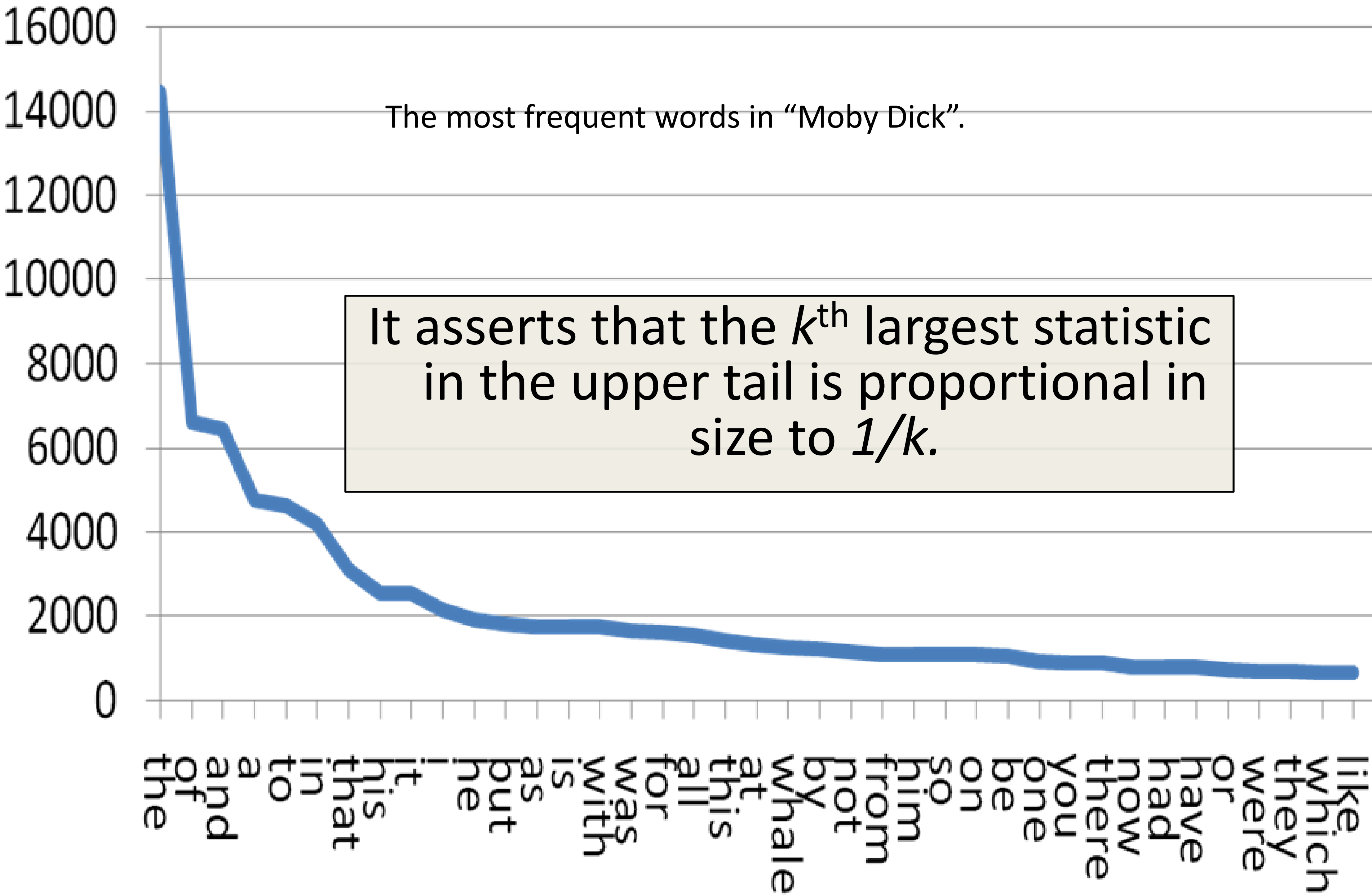
The law is not just an empirical law; it also applies to many mathematical data sets.

However, the law usually does NOT apply to human-generated numbers; because of this, Benford’s law is a useful tool for detecting accounting fraud.

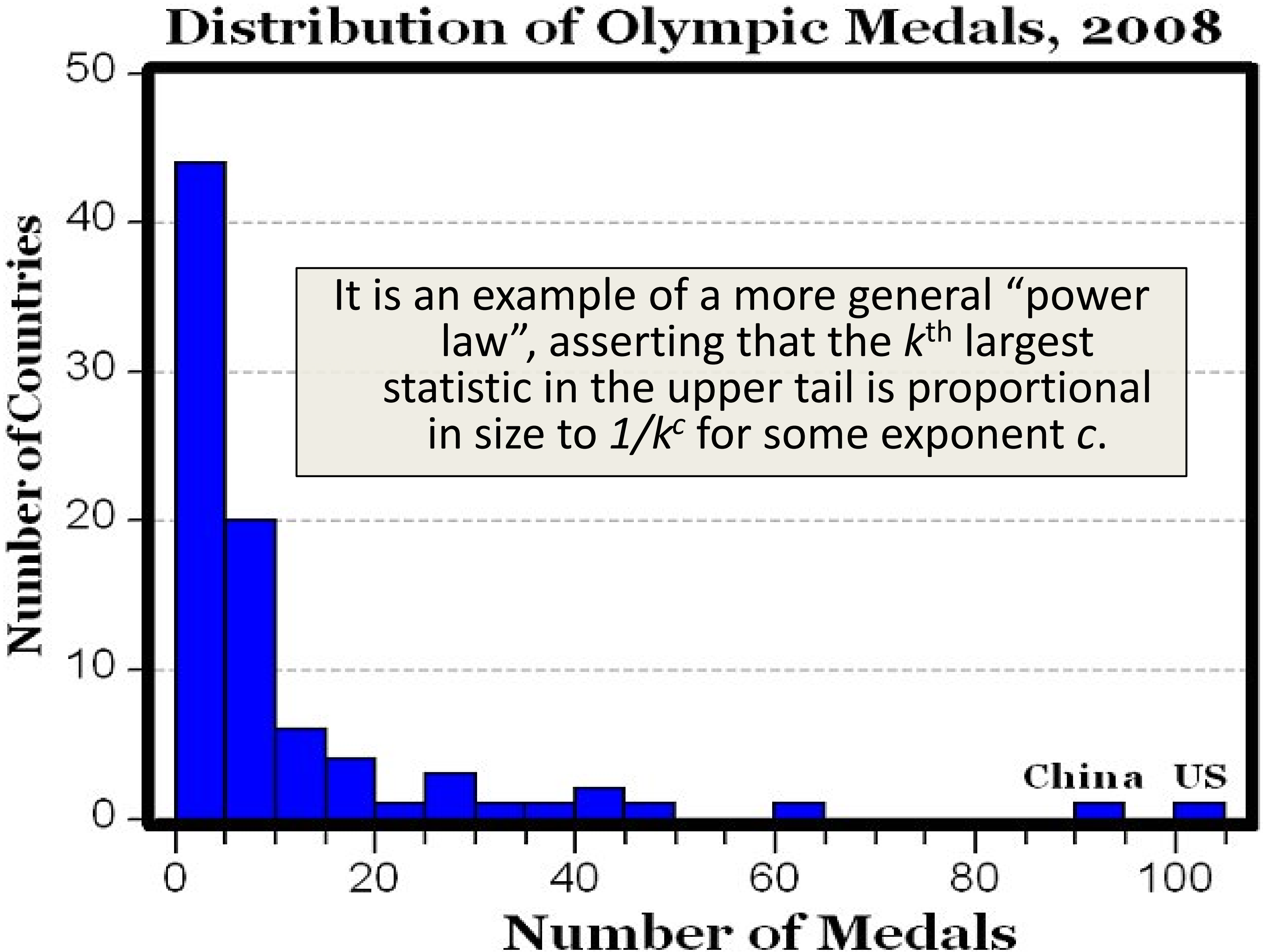




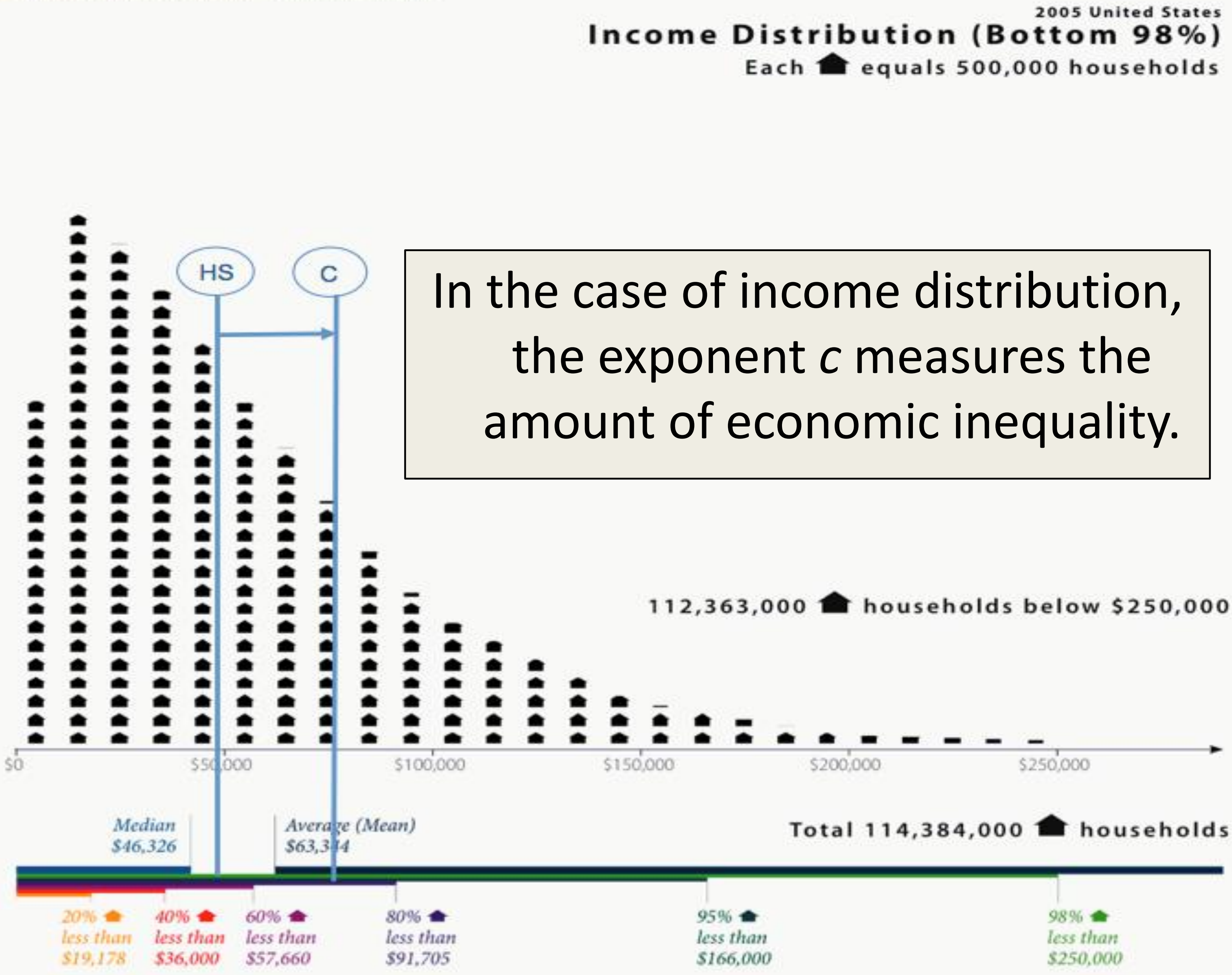
A variant of Benford's law is *Zipf's law*, relating to the largest statistics (the "upper tail") in a data set.



It asserts that the k^{th} largest statistic in the upper tail is proportional in size to $1/k$.

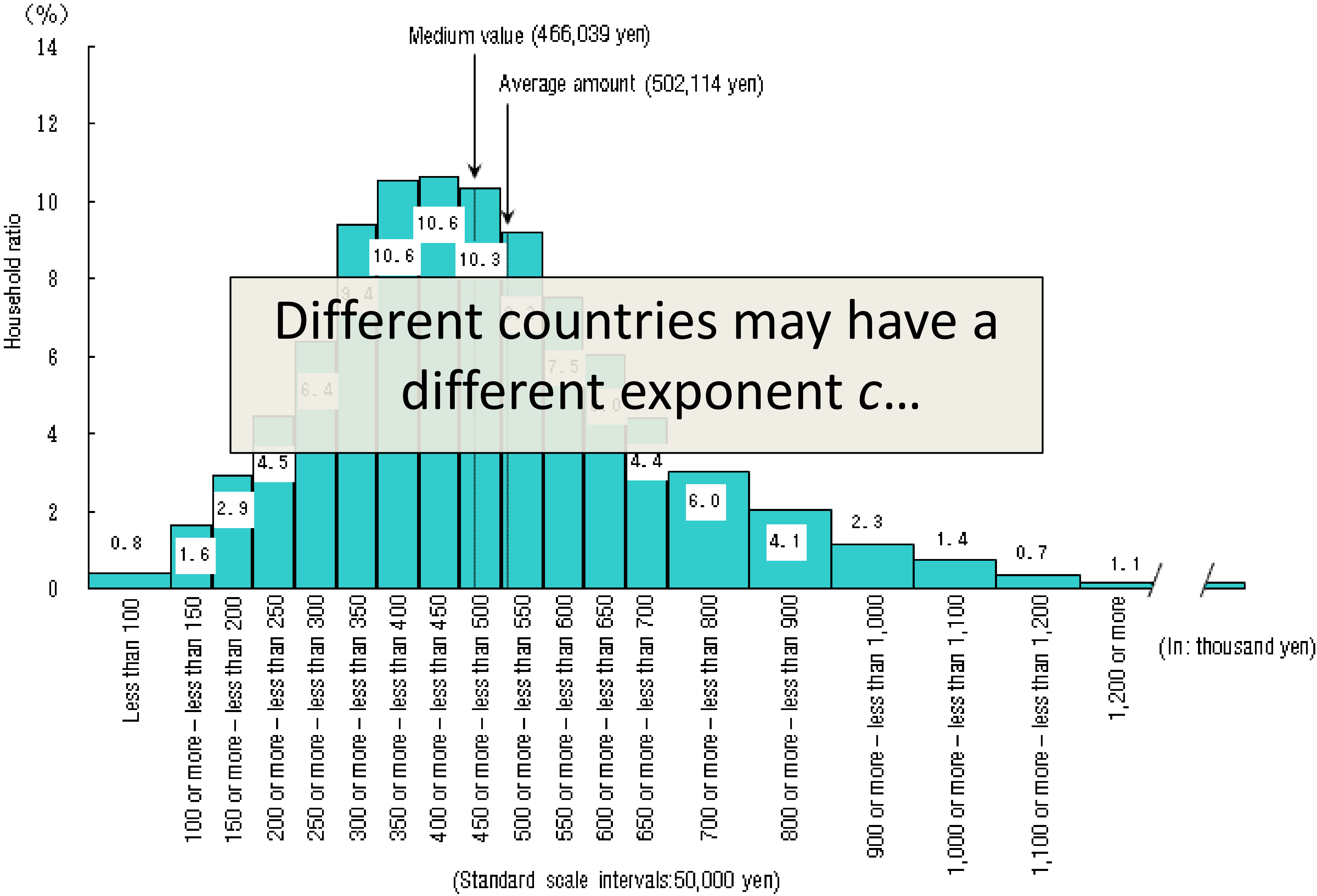


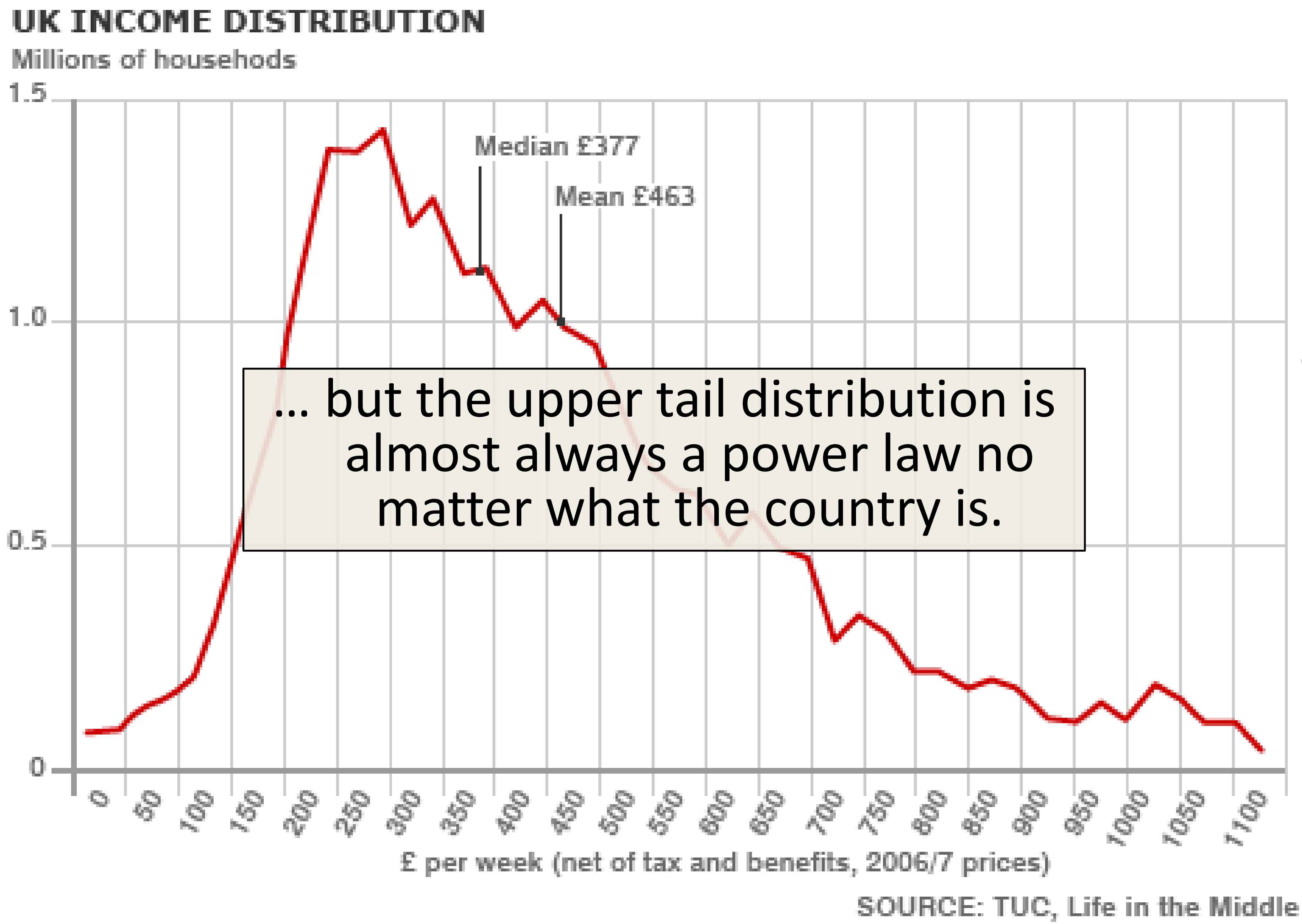
It is an example of a more general “power law”, asserting that the k^{th} largest statistic in the upper tail is proportional in size to $1/k^c$ for some exponent c.



In the case of income distribution, the exponent *c* measures the amount of economic inequality.

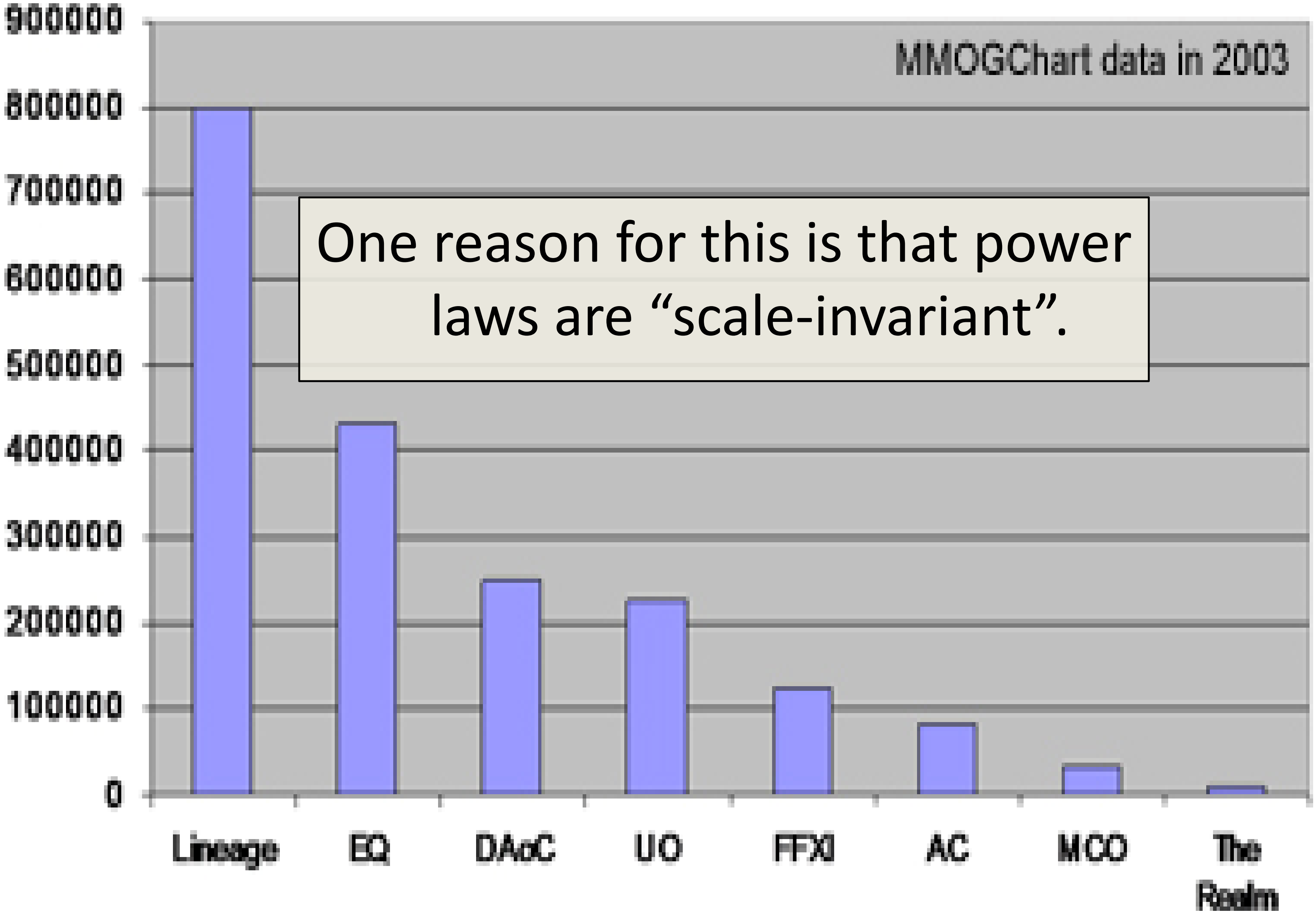
Different countries may have a different exponent *c*…



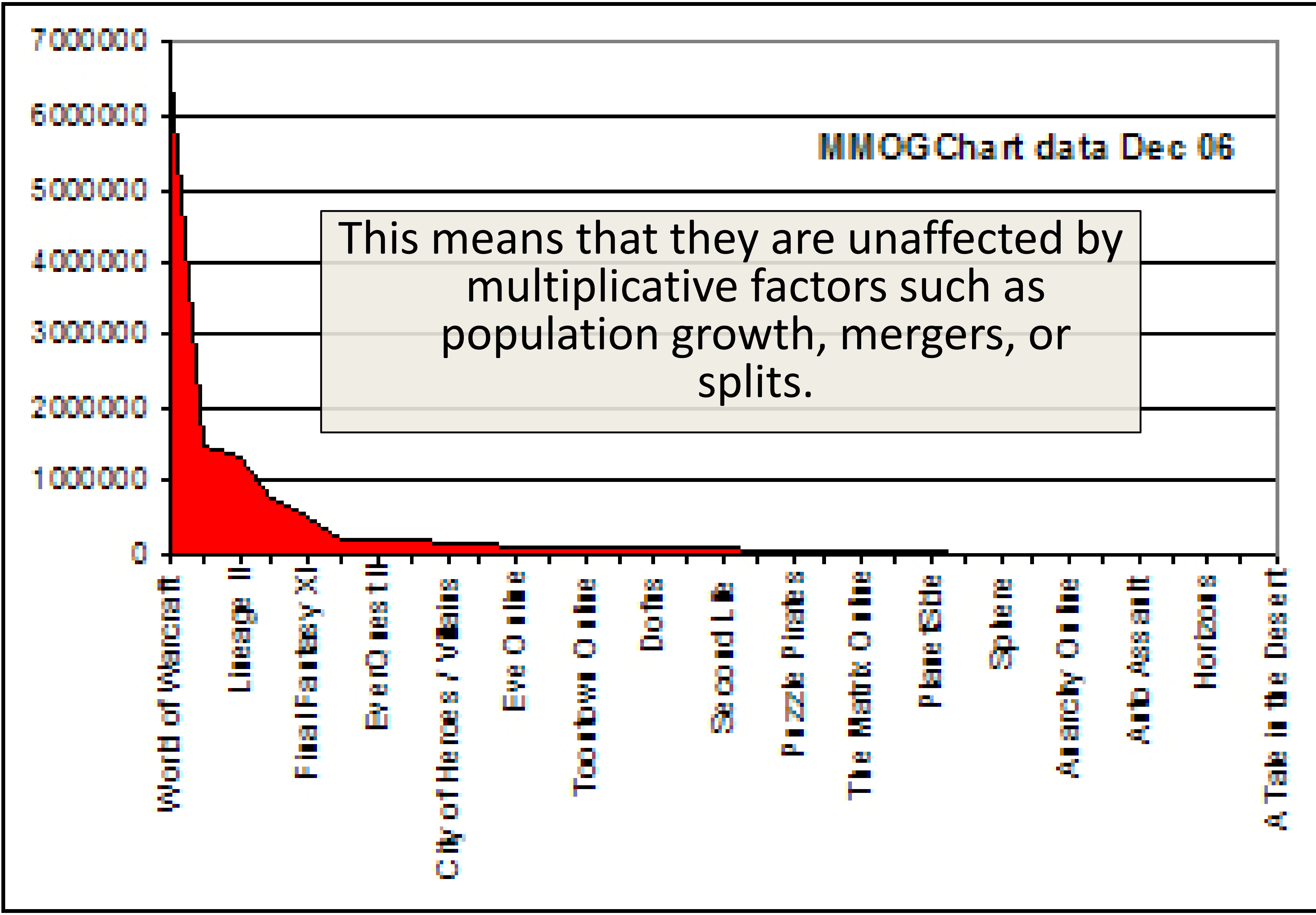


… but the upper tail distribution is almost always a power law no matter what the country is.

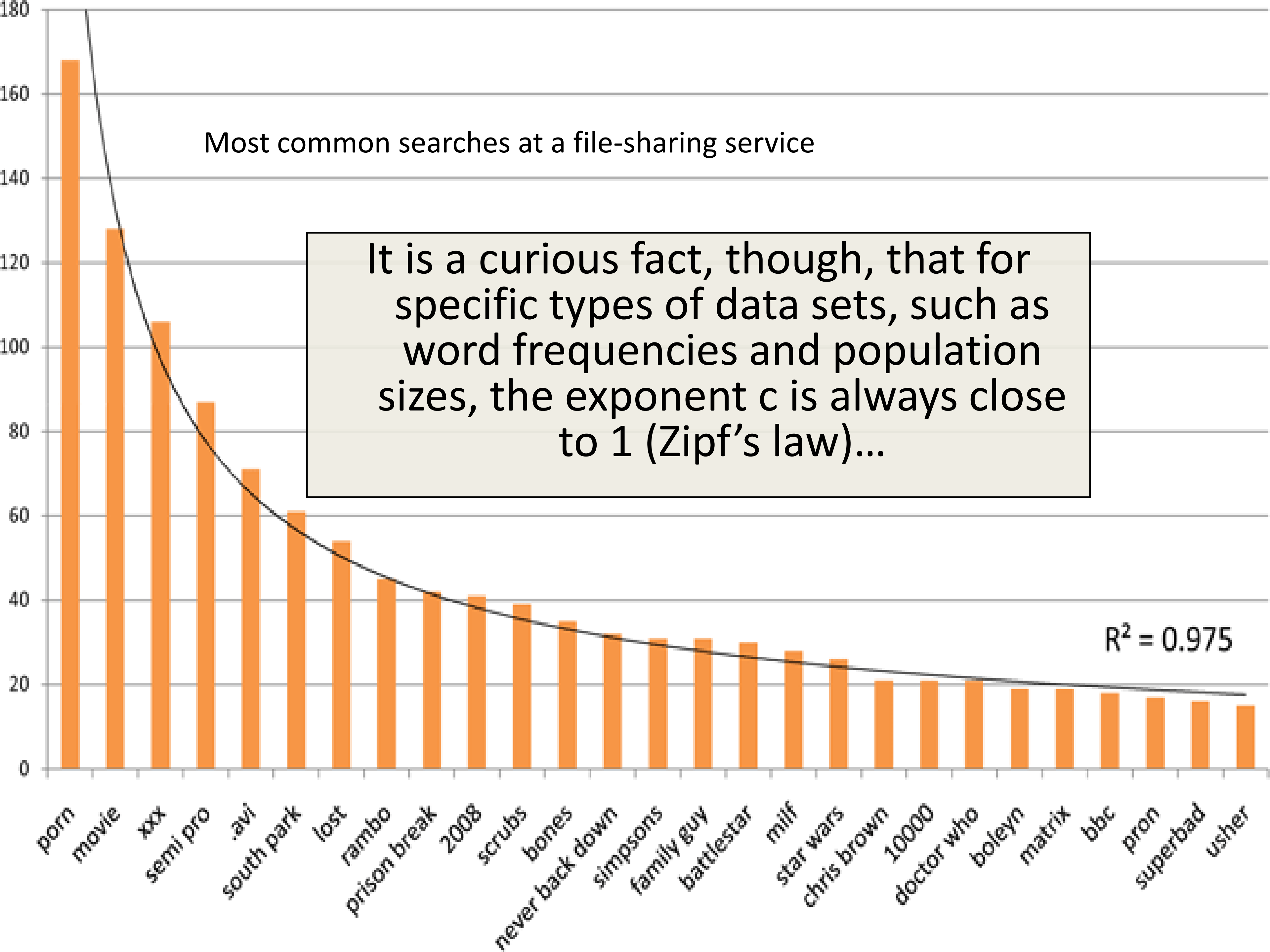
The UK income distribution is quite uneven



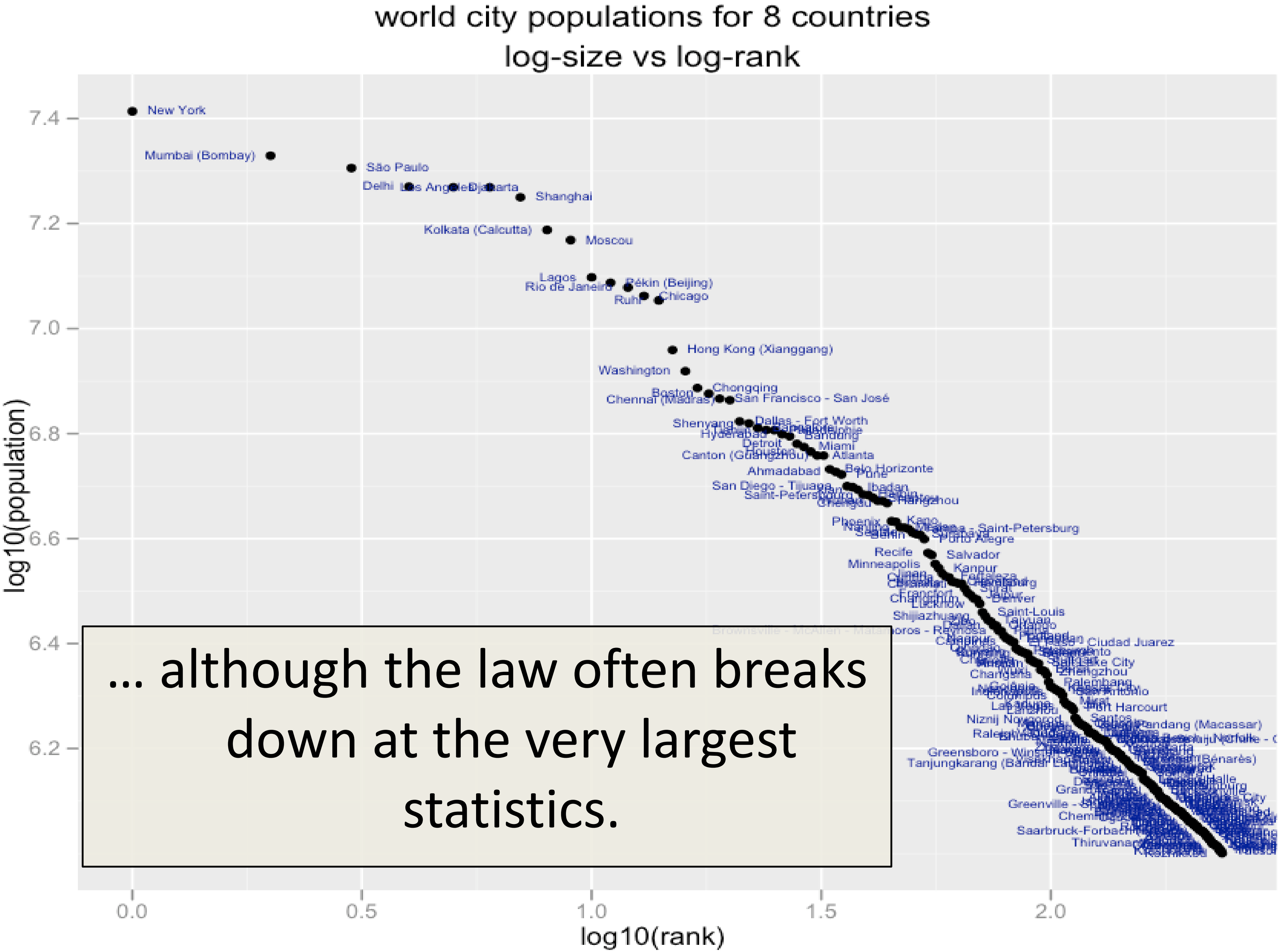
One reason for this is that power laws are "scale-invariant".



This means that they are unaffected by multiplicative factors such as population growth, mergers, or splits.

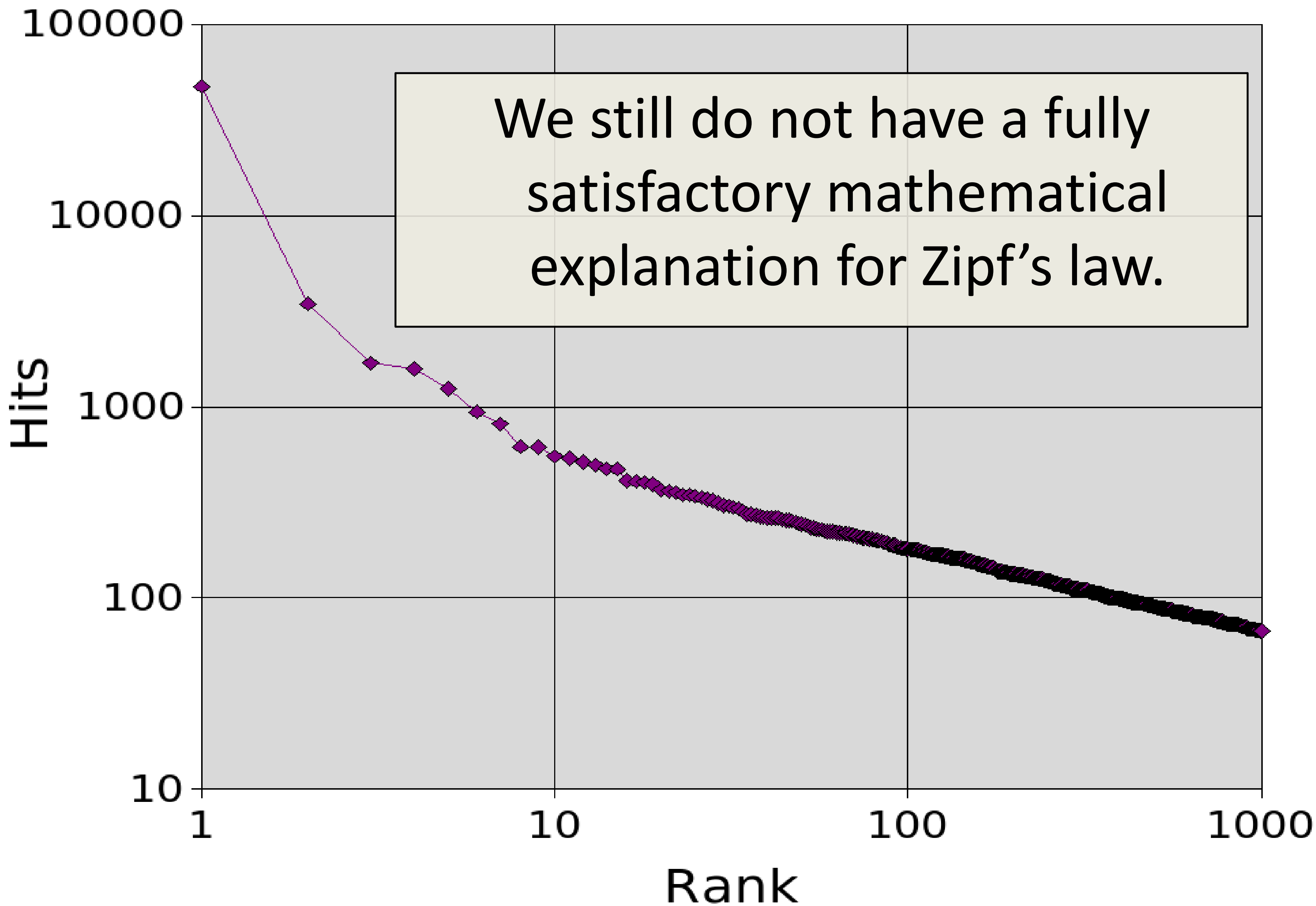


It is a curious fact, though, that for specific types of data sets, such as word frequencies and population sizes, the exponent c is always close to 1 (Zipf's law)...



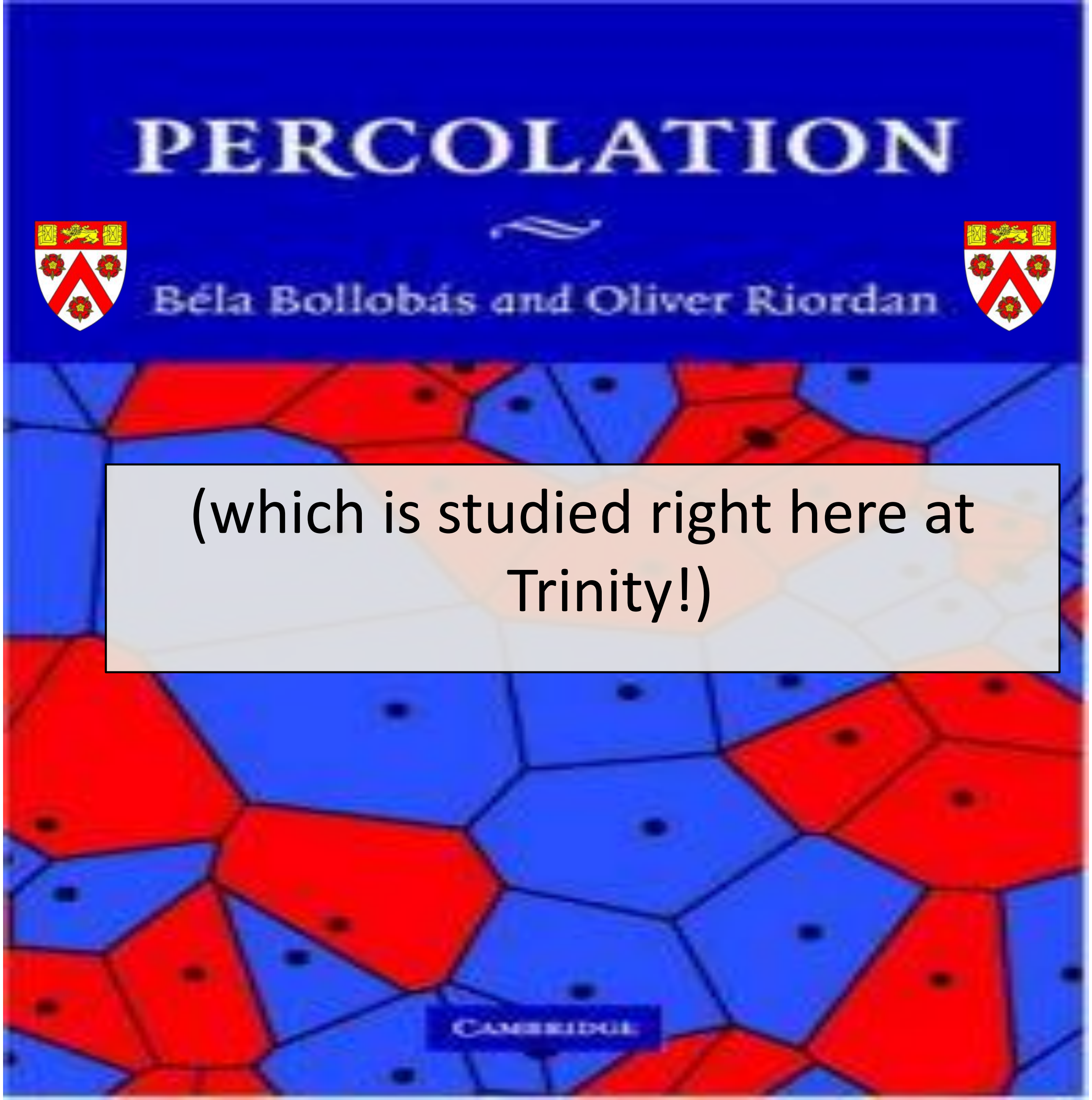
… although the law often breaks down at the very largest statistics.

Wikipedia page hits vs. rank

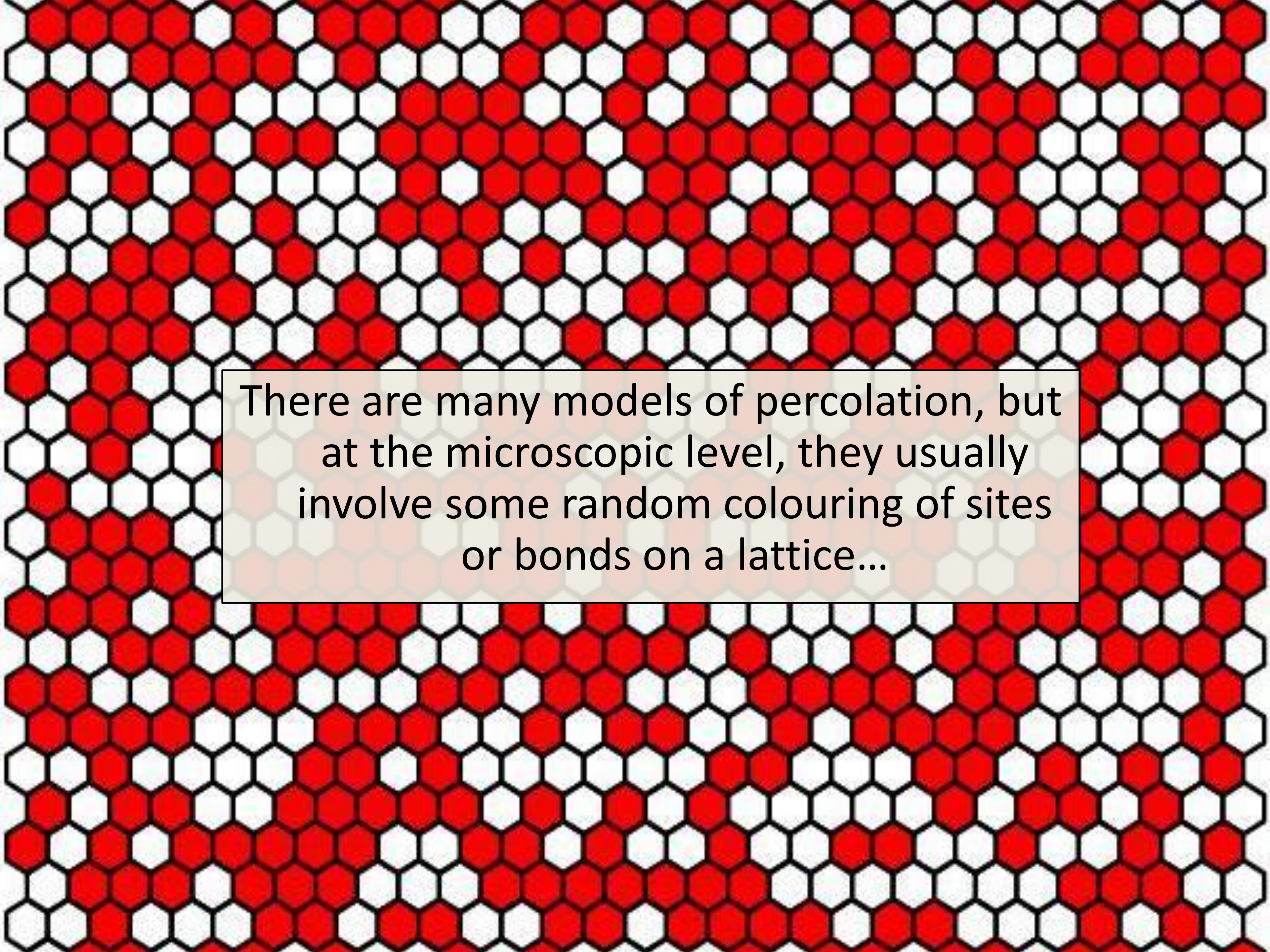


We still do not have a fully satisfactory mathematical explanation for Zipf's law.

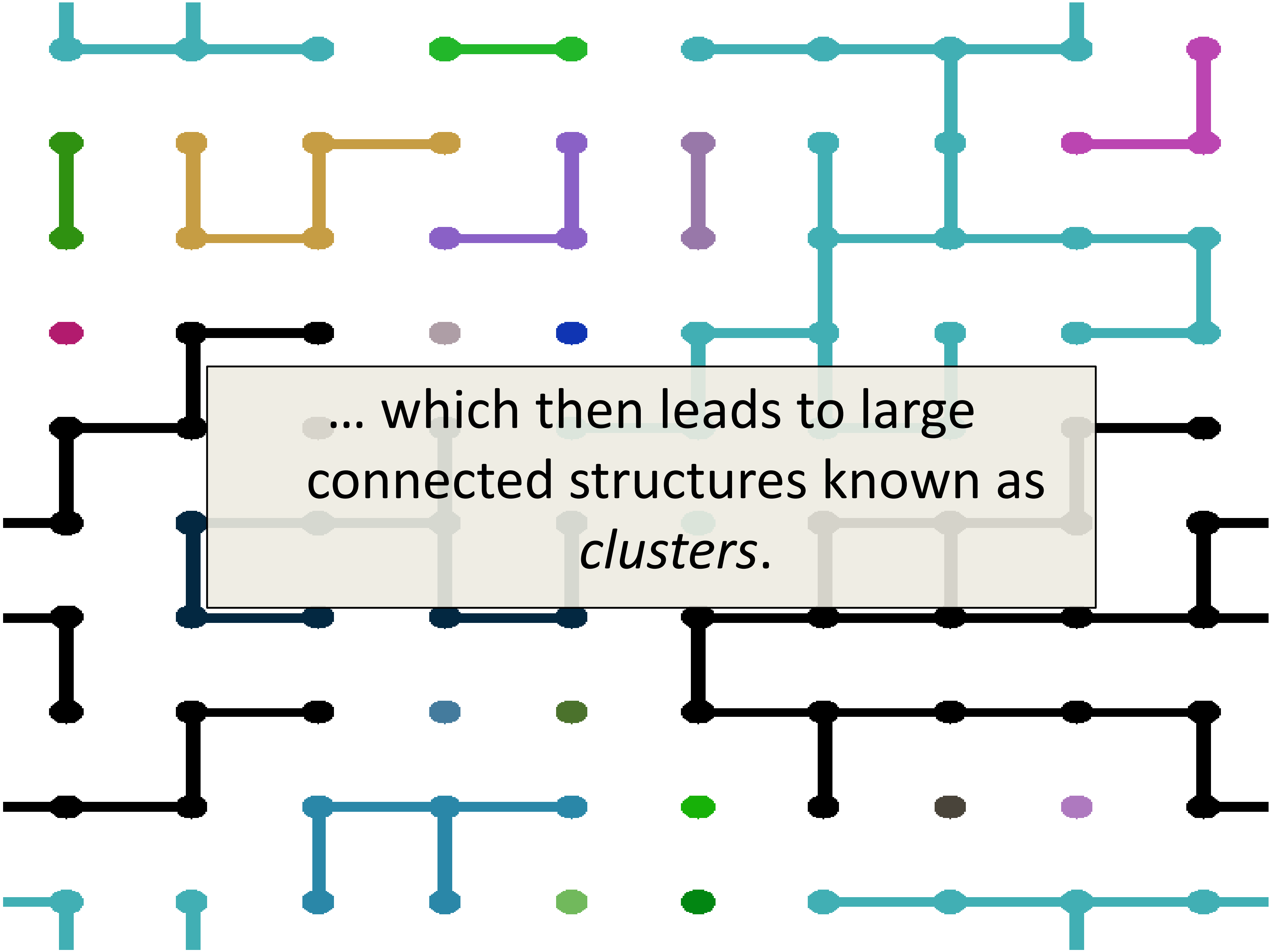
Universality phenomena show up in many areas of mathematics, such as the theory of *percolation*.



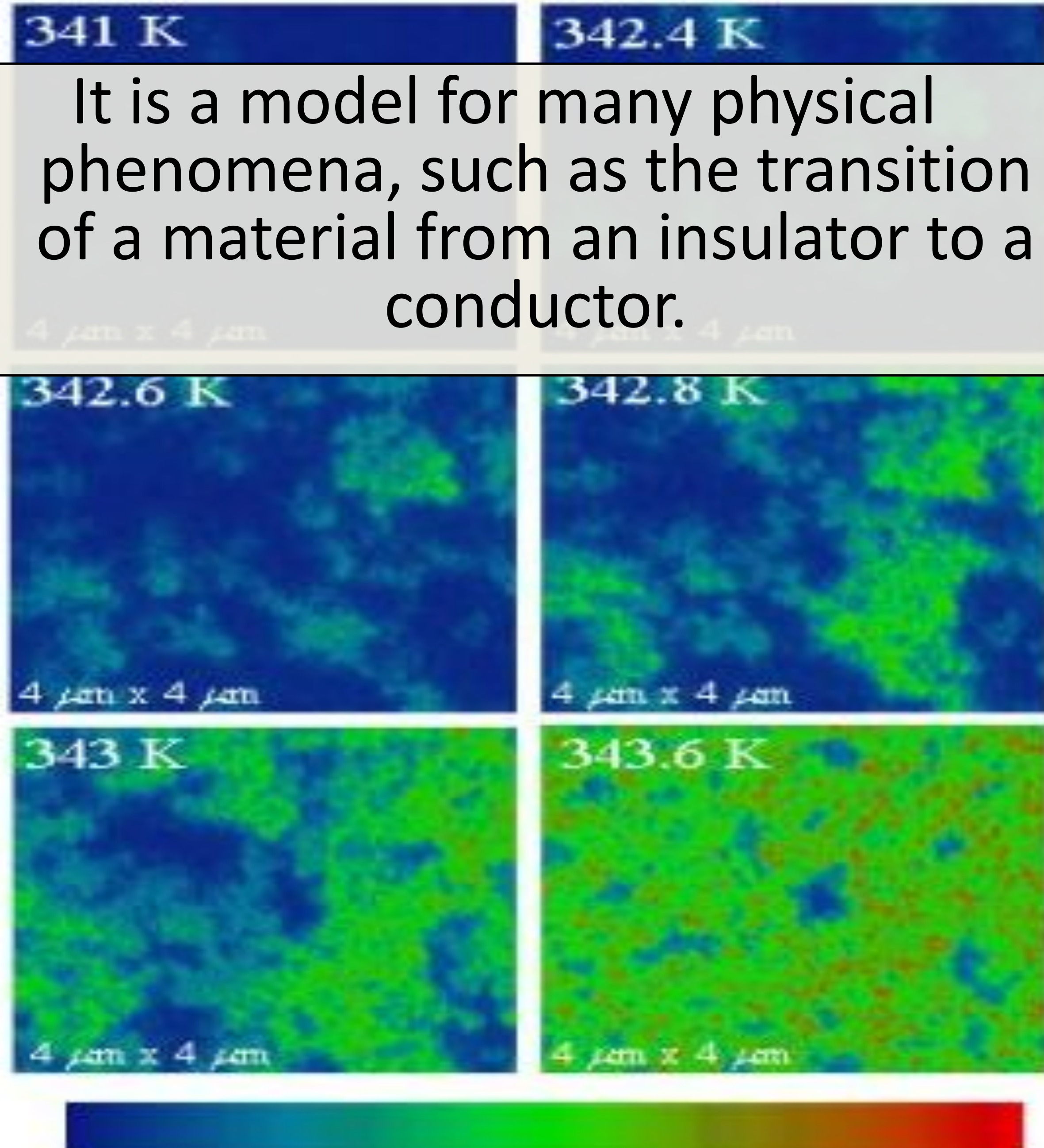
(which is studied right here at Trinity!)



There are many models of percolation, but at the microscopic level, they usually involve some random colouring of sites or bonds on a lattice…

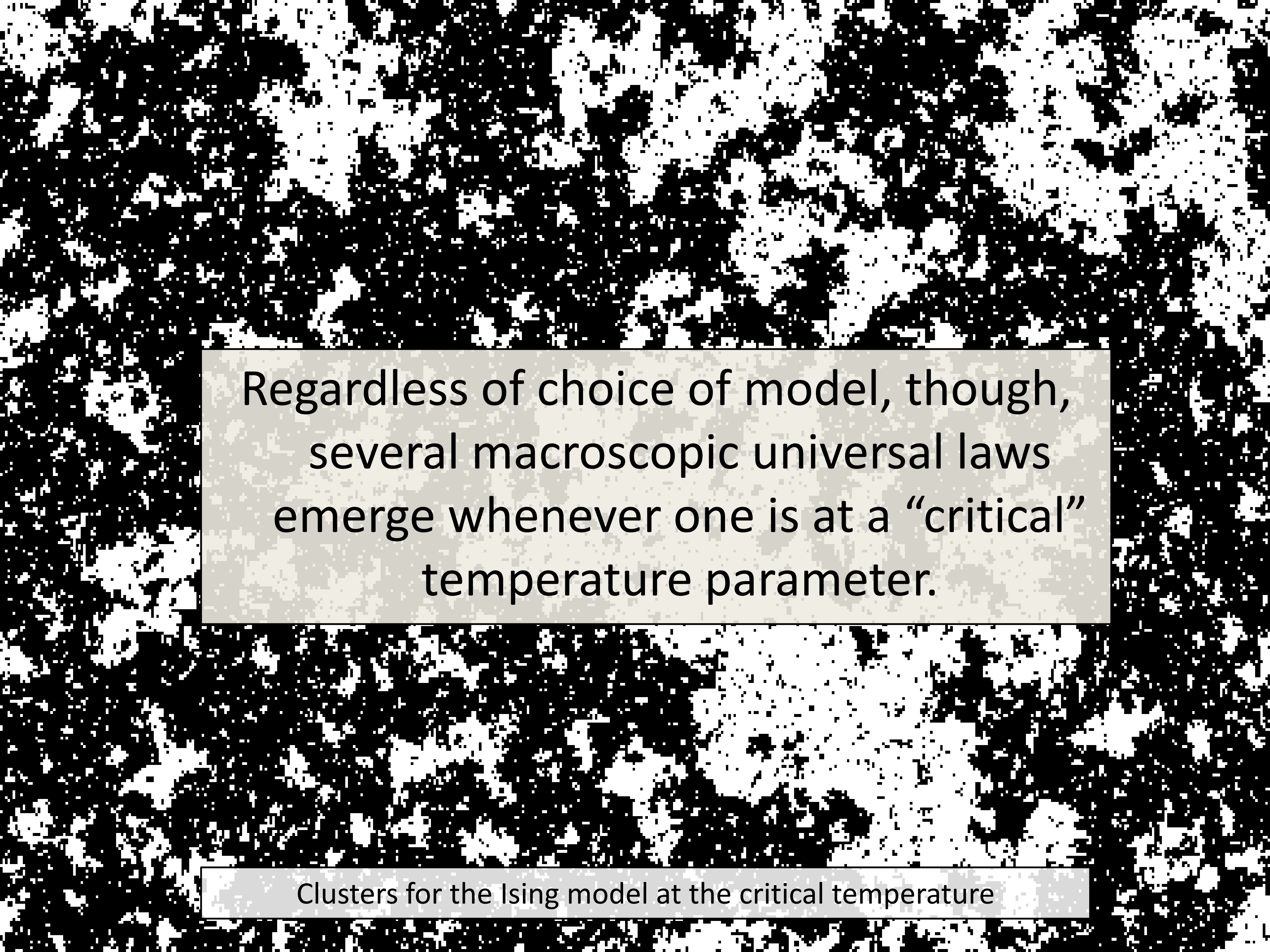

… which then leads to large connected structures known as *clusters*.

It is a model for many physical phenomena, such as the transition of a material from an insulator to a conductor.

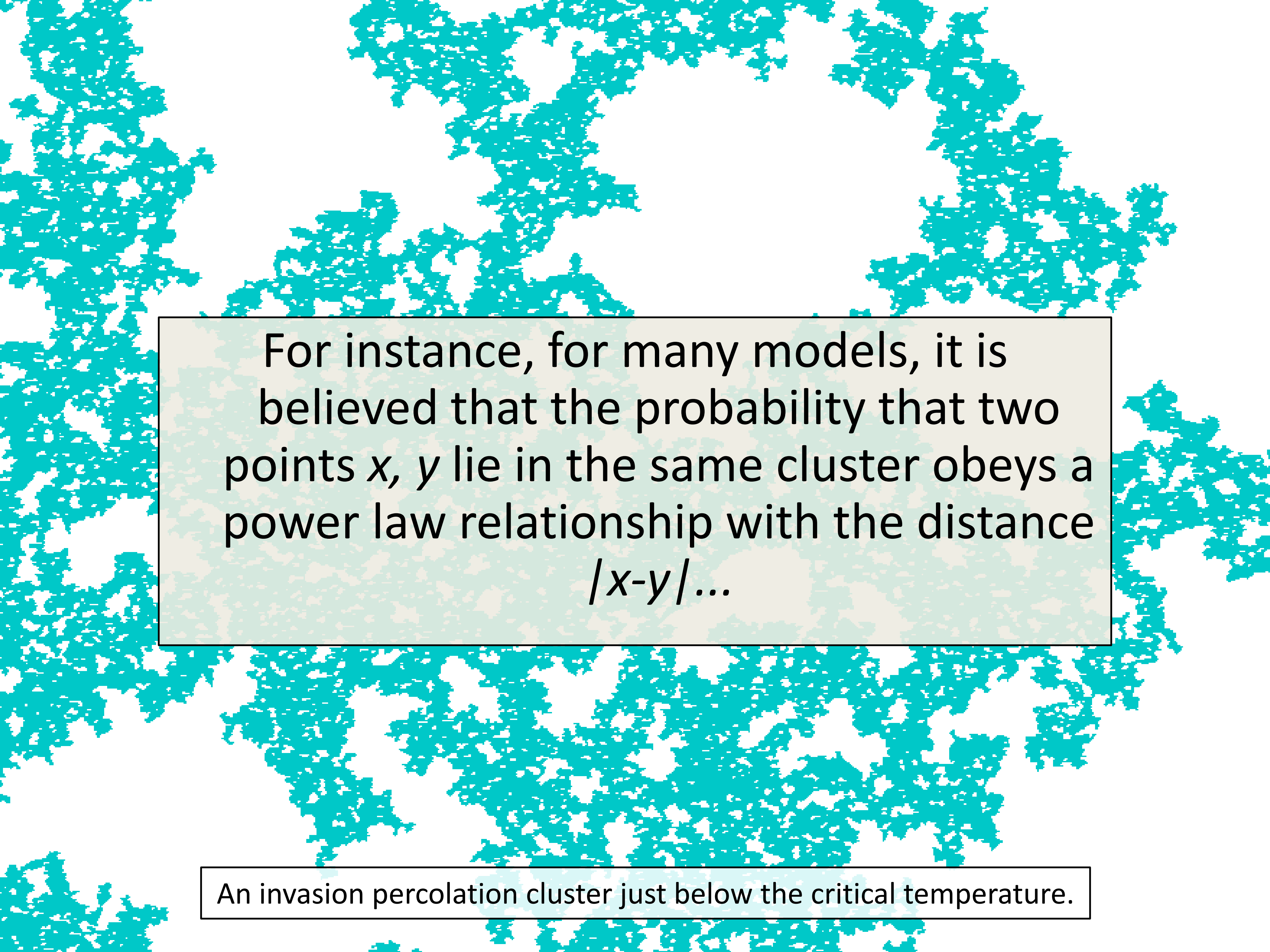


Insulator (green) and conductor (blue) behaviour of vanadium dioxide near the critical temperature

Regardless of choice of model, though, several macroscopic universal laws emerge whenever one is at a “critical” temperature parameter.

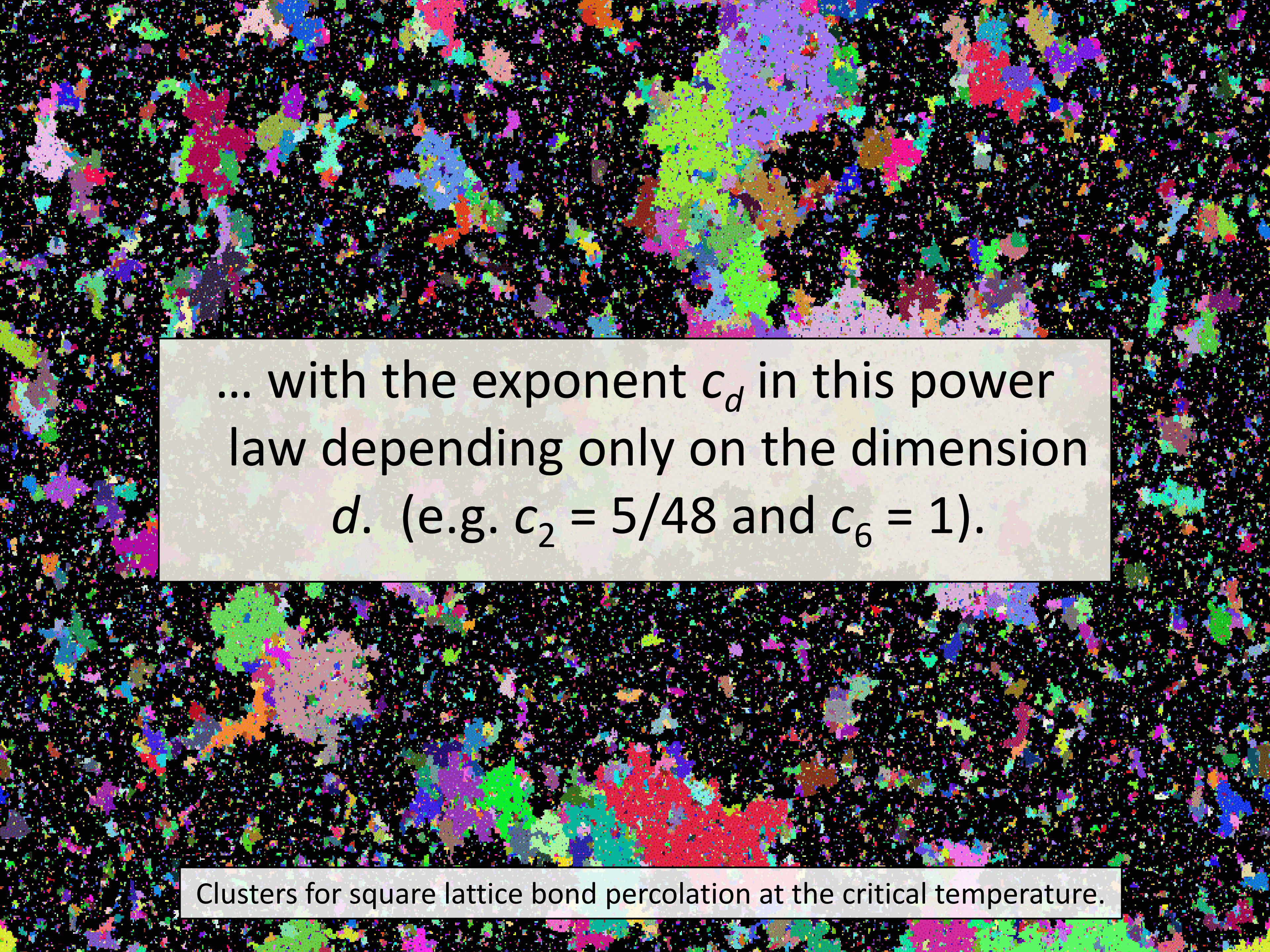


Clusters for the Ising model at the critical temperature



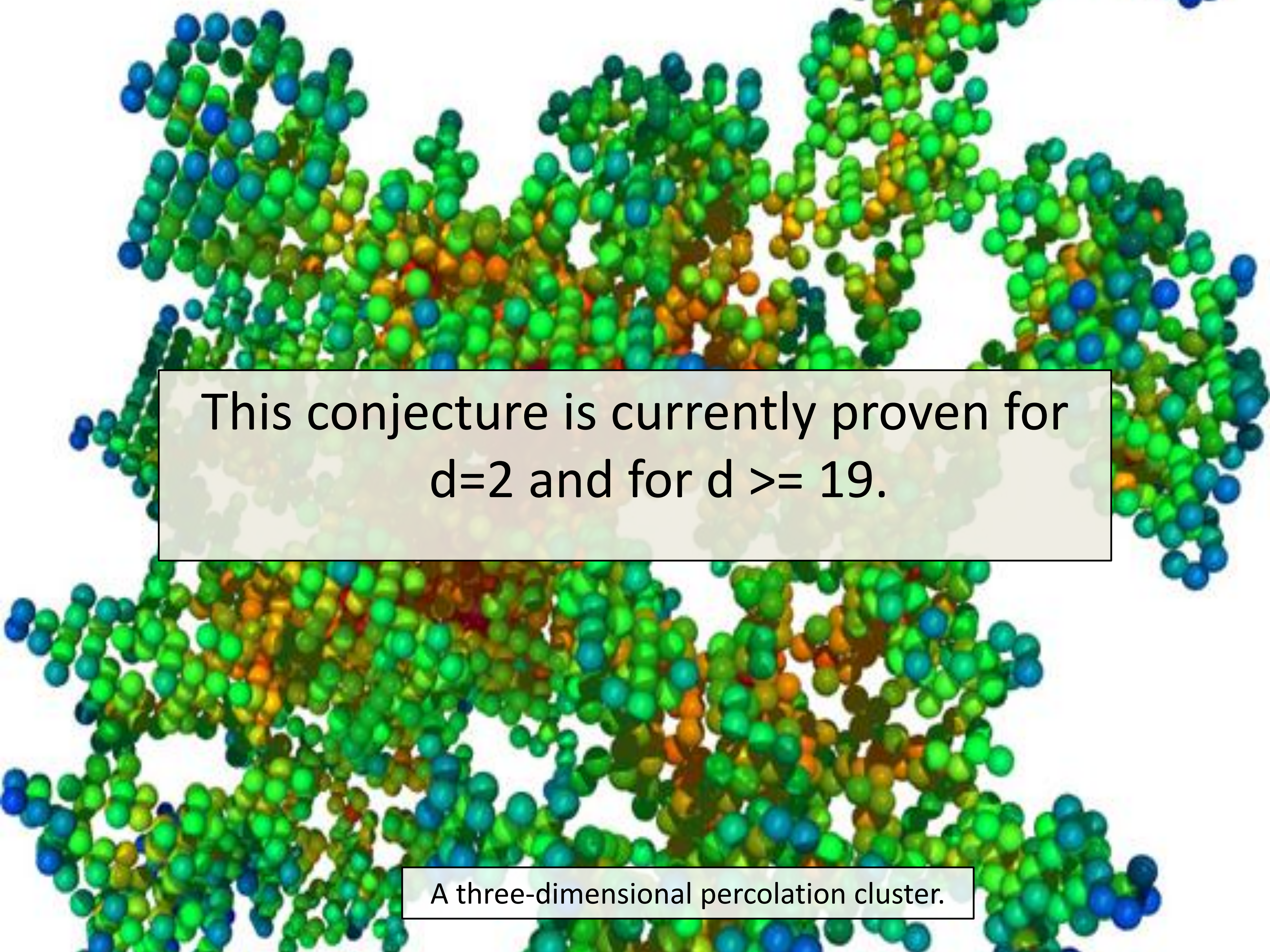
For instance, for many models, it is believed that the probability that two points *x, y* lie in the same cluster obeys a power law relationship with the distance $|x-y|$...

An invasion percolation cluster just below the critical temperature.



… with the exponent c_d in this power law depending only on the dimension d. (e.g. $c_2 = 5/48$ and $c_6 = 1$).

Clusters for square lattice bond percolation at the critical temperature.



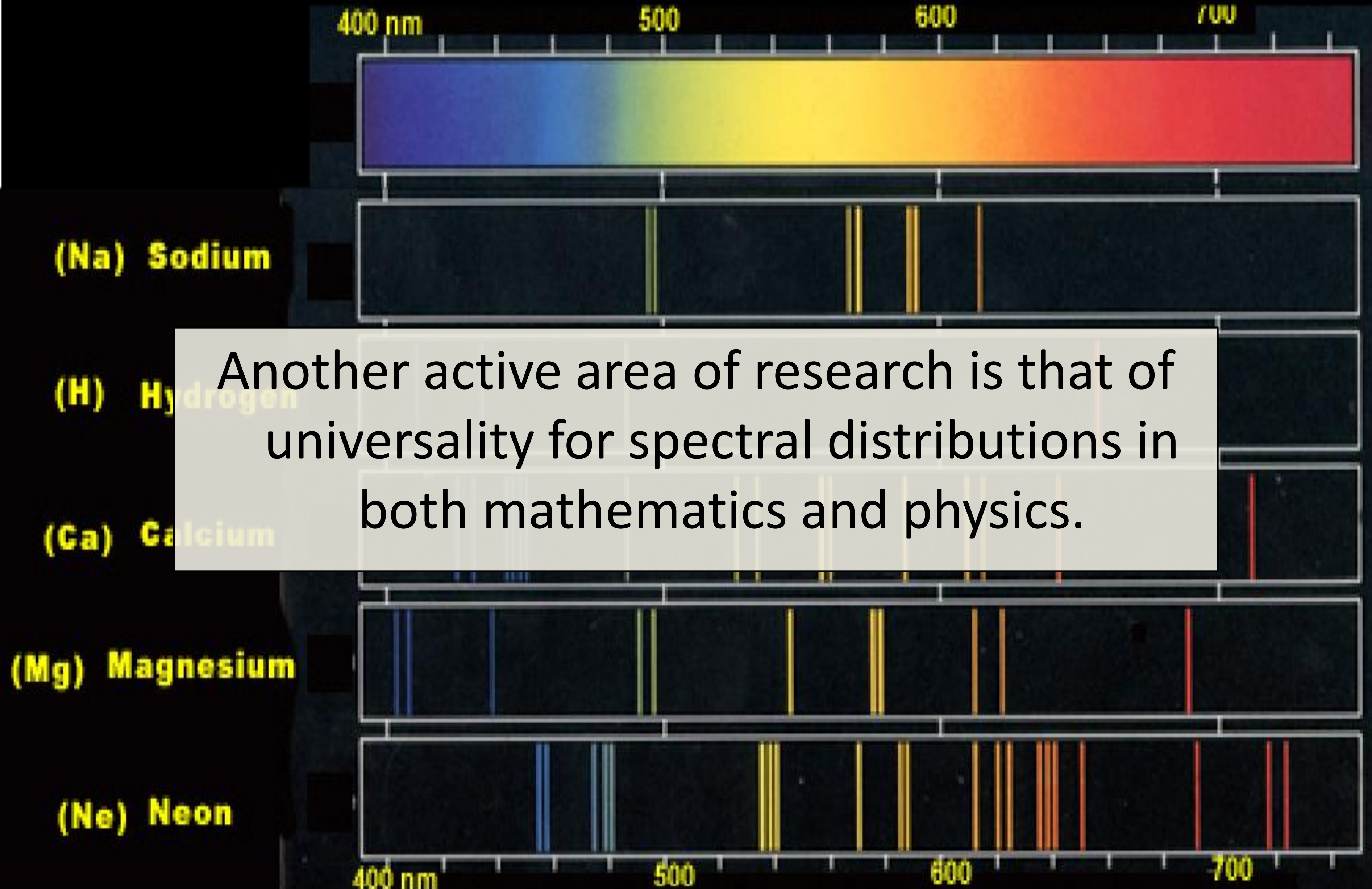
This conjecture is currently proven for d=2 and for d >= 19.

A three-dimensional percolation cluster.

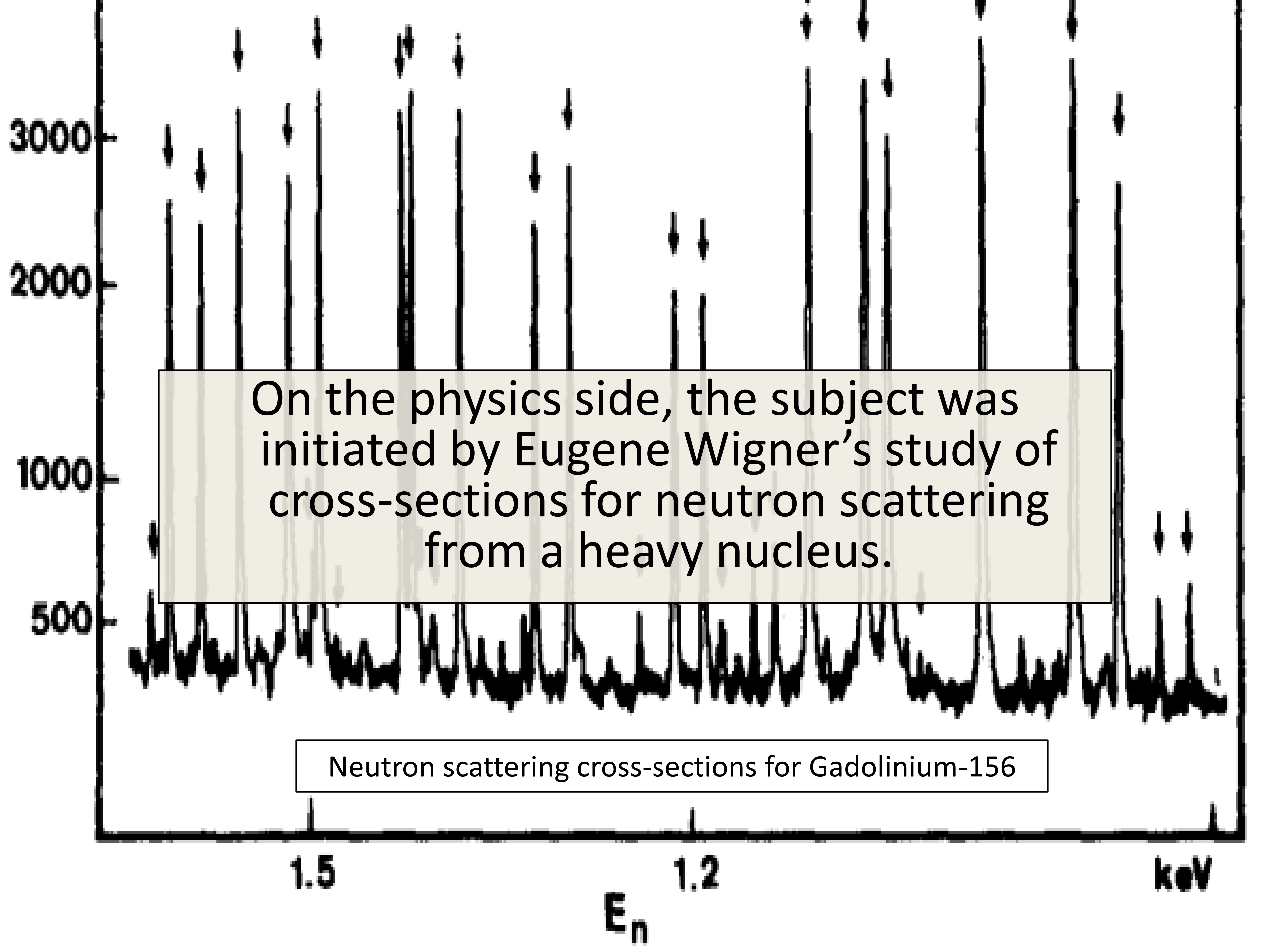
Work in this area formed part of the citation for Stas Smirnov's (third from left) Fields medal in 2010.

International Mathematical Olympiad, Bremen 2009

Another active area of research is that of universality for spectral distributions in both mathematics and physics.

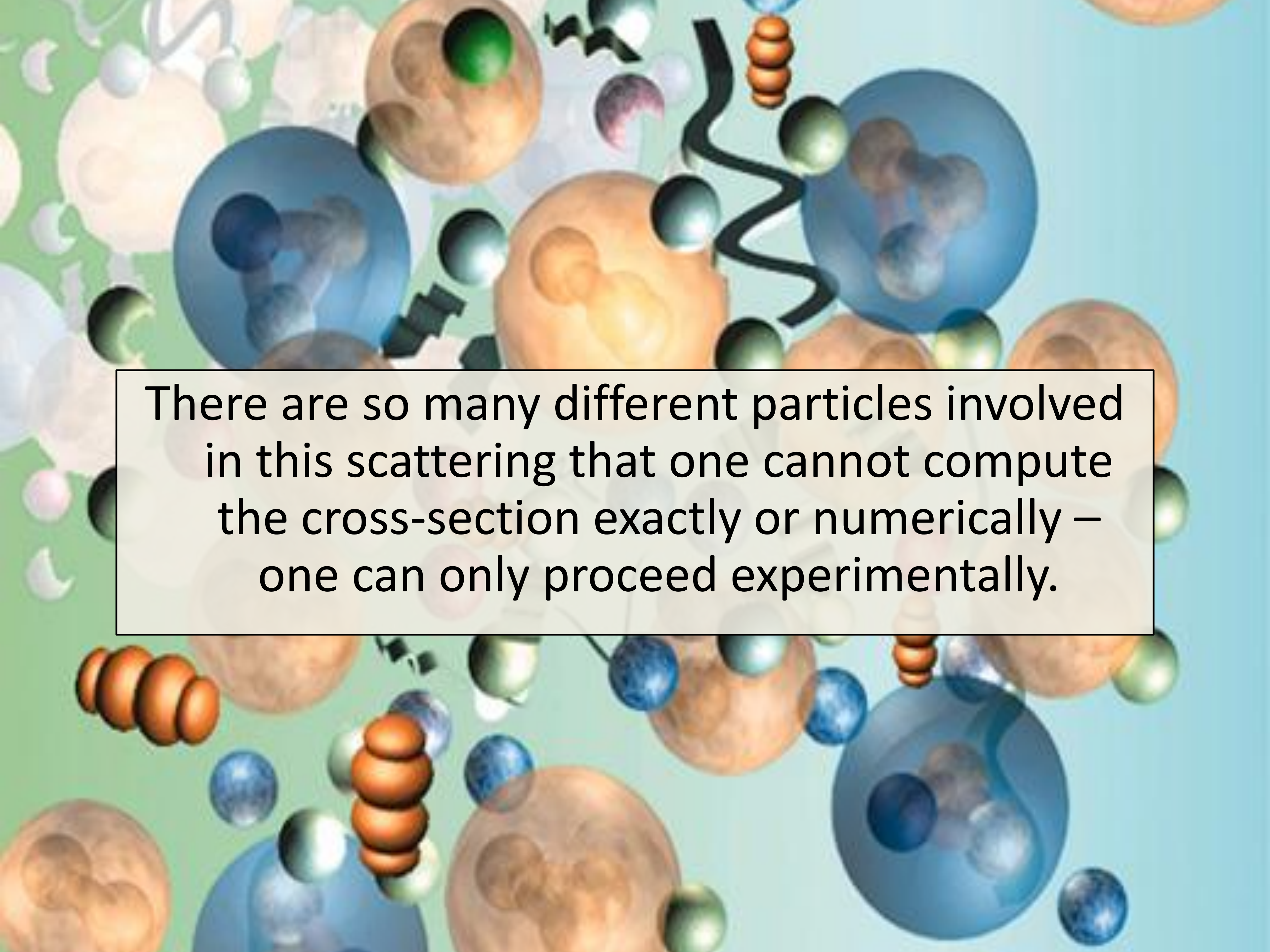


Adapted from Honolulu Community College Science 122 site



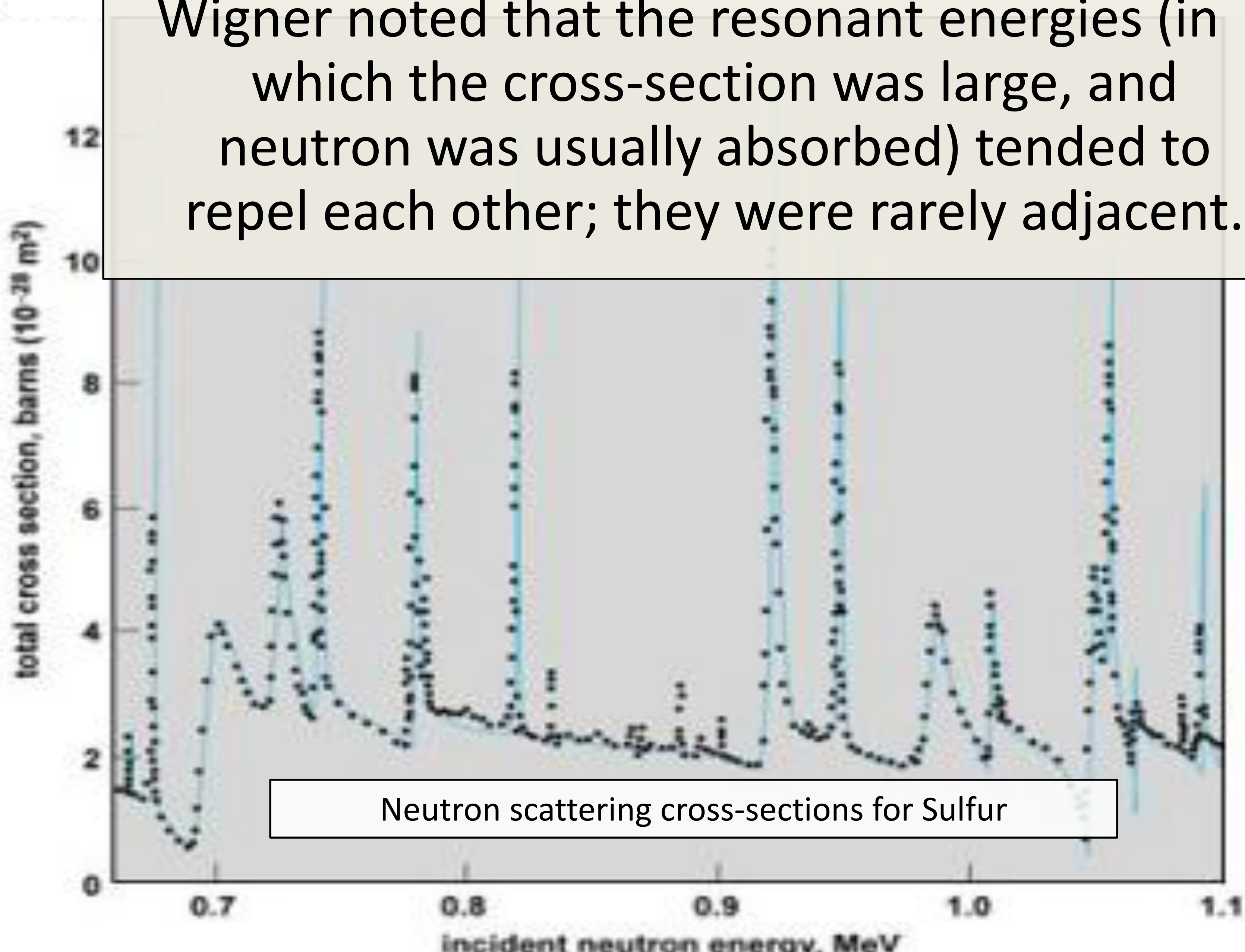
On the physics side, the subject was initiated by Eugene Wigner's study of cross-sections for neutron scattering from a heavy nucleus.

Neutron scattering cross-sections for Gadolinium-156

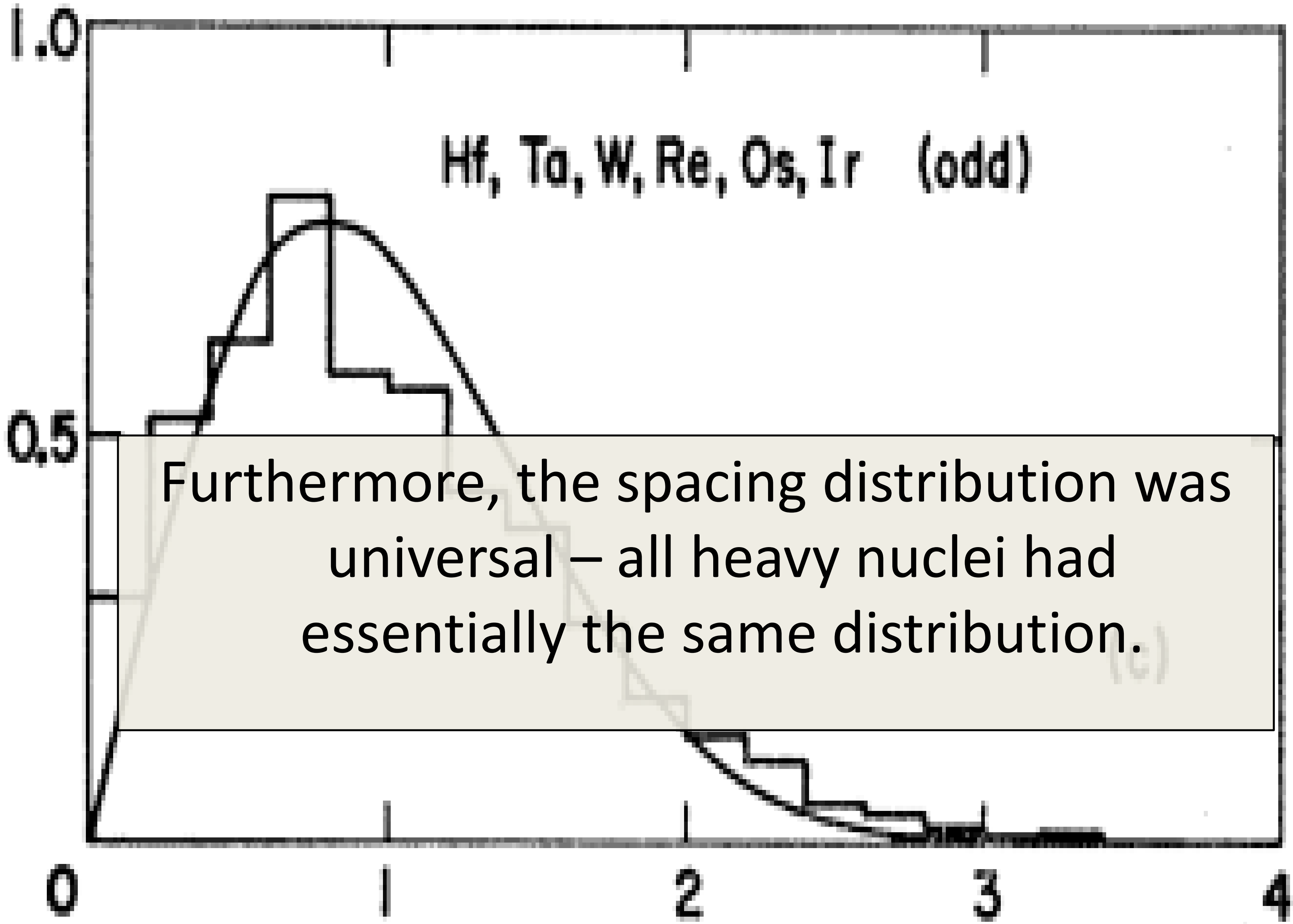


There are so many different particles involved in this scattering that one cannot compute the cross-section exactly or numerically – one can only proceed experimentally.

Wigner noted that the resonant energies (in which the cross-section was large, and neutron was usually absorbed) tended to repel each other; they were rarely adjacent.



Neutron scattering cross-sections for Sulfur



Furthermore, the spacing distribution was universal – all heavy nuclei had essentially the same distribution.

Cross-section spacing distribution for several heavy nuclei

6 2 3 4 7 8 1 2

5 4 9 1 4 2 6 5

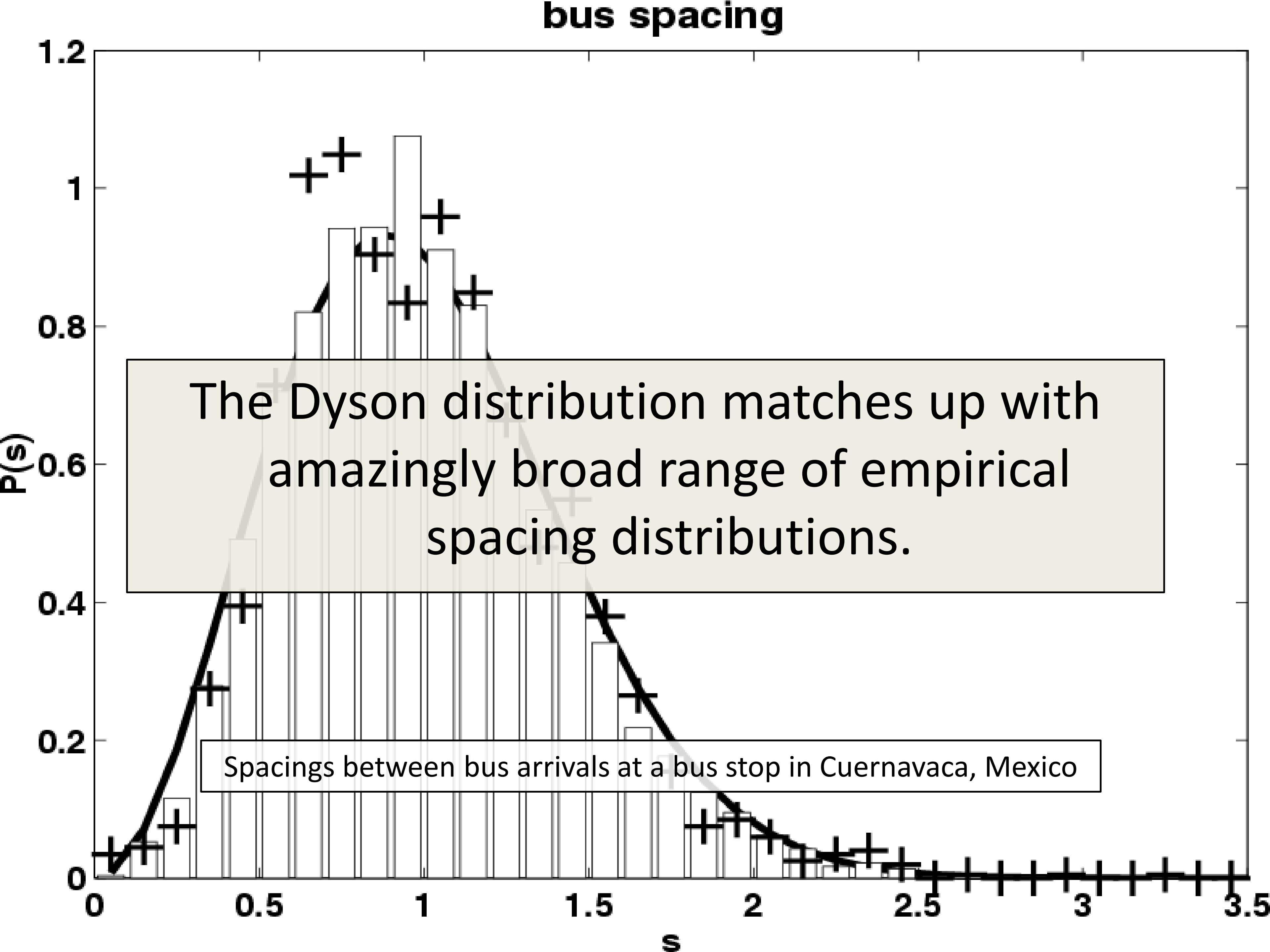
Wigner proposed modeling these spectral lines by the *eigenvalues* of a *random matrix*.

1 0 2 7 4 6 3 9

5 1 9 8 7 0 0 3

$$\rho_k(x_1,\ldots,x_k)=\det\left(\frac{\sin\pi(x_i-x_j)}{\pi(x_i-x_j)}\right)_{1\le i,j\le k}$$

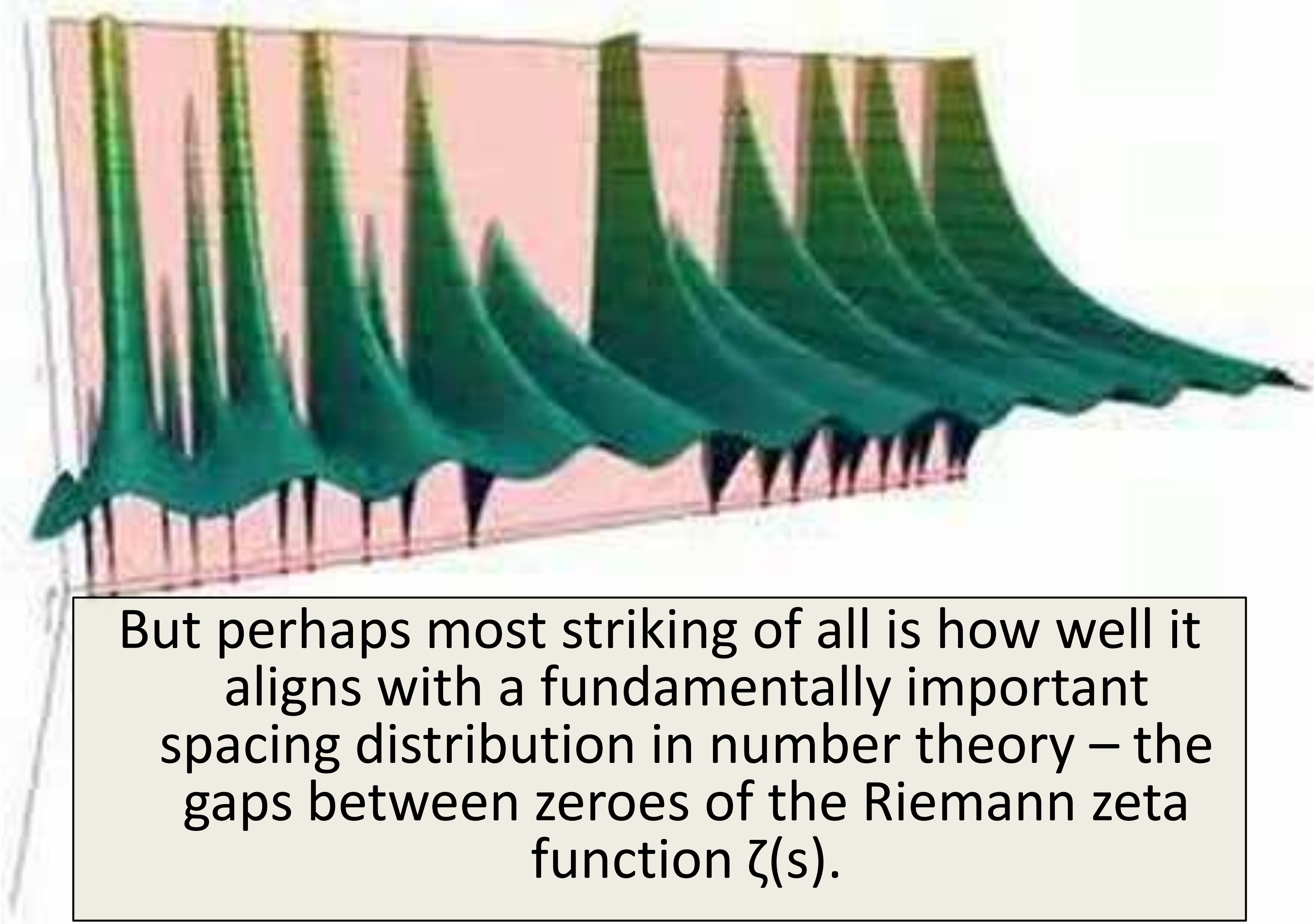
For a special type of random matrix model (the *Gaussian Unitary Ensemble*, or GUE), the asymptotic eigenvalue distribution was computed by Freeman Dyson.



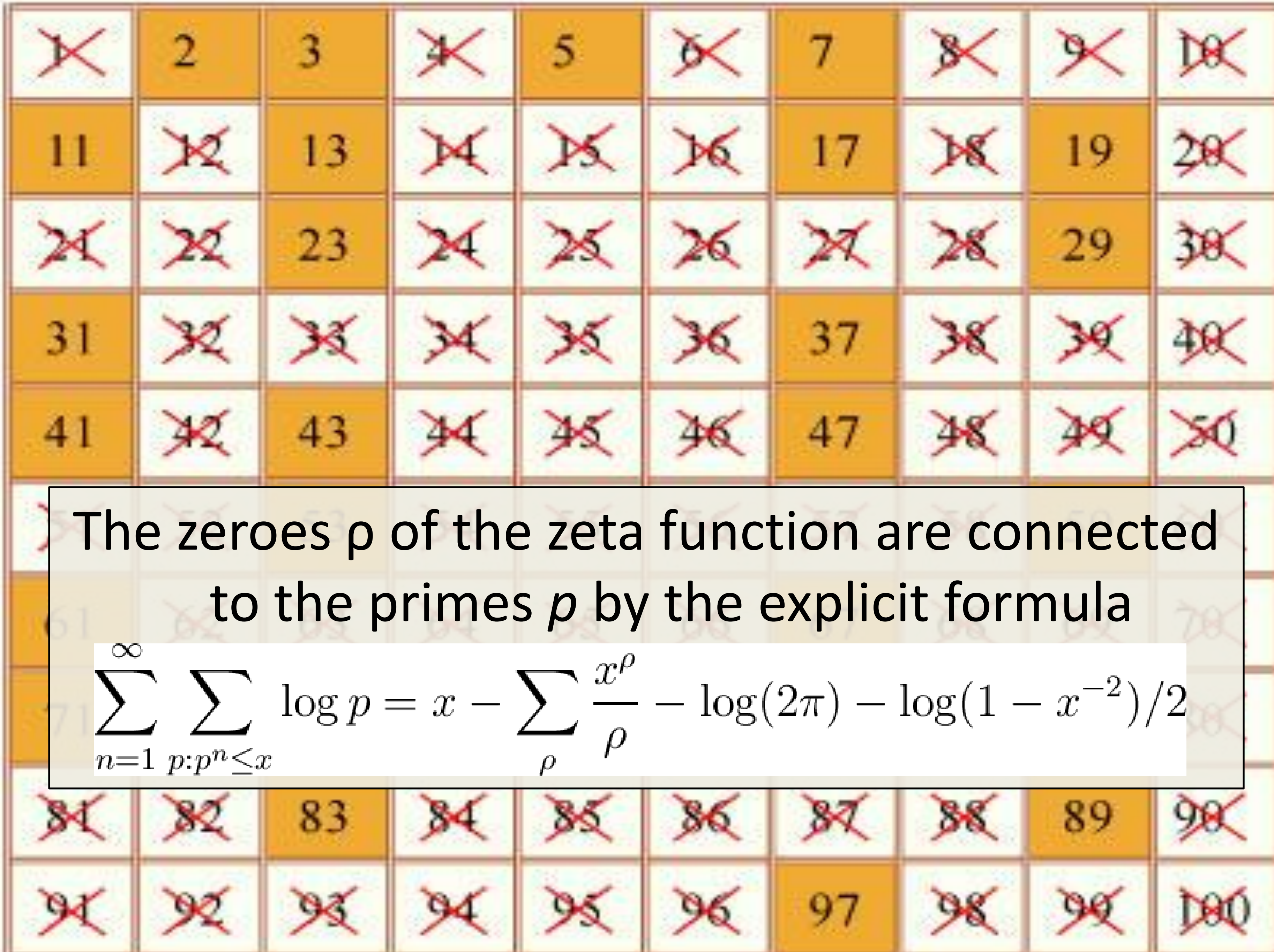
The Dyson distribution matches up with amazingly broad range of empirical spacing distributions.

Spacings between bus arrivals at a bus stop in Cuernavaca, Mexico

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p (1 - \frac{1}{p^s})^{-1}$$



But perhaps most striking of all is how well it aligns with a fundamentally important spacing distribution in number theory – the gaps between zeroes of the Riemann zeta function ζ(s).



The zeroes ρ of the zeta function are connected to the primes *p* by the explicit formula

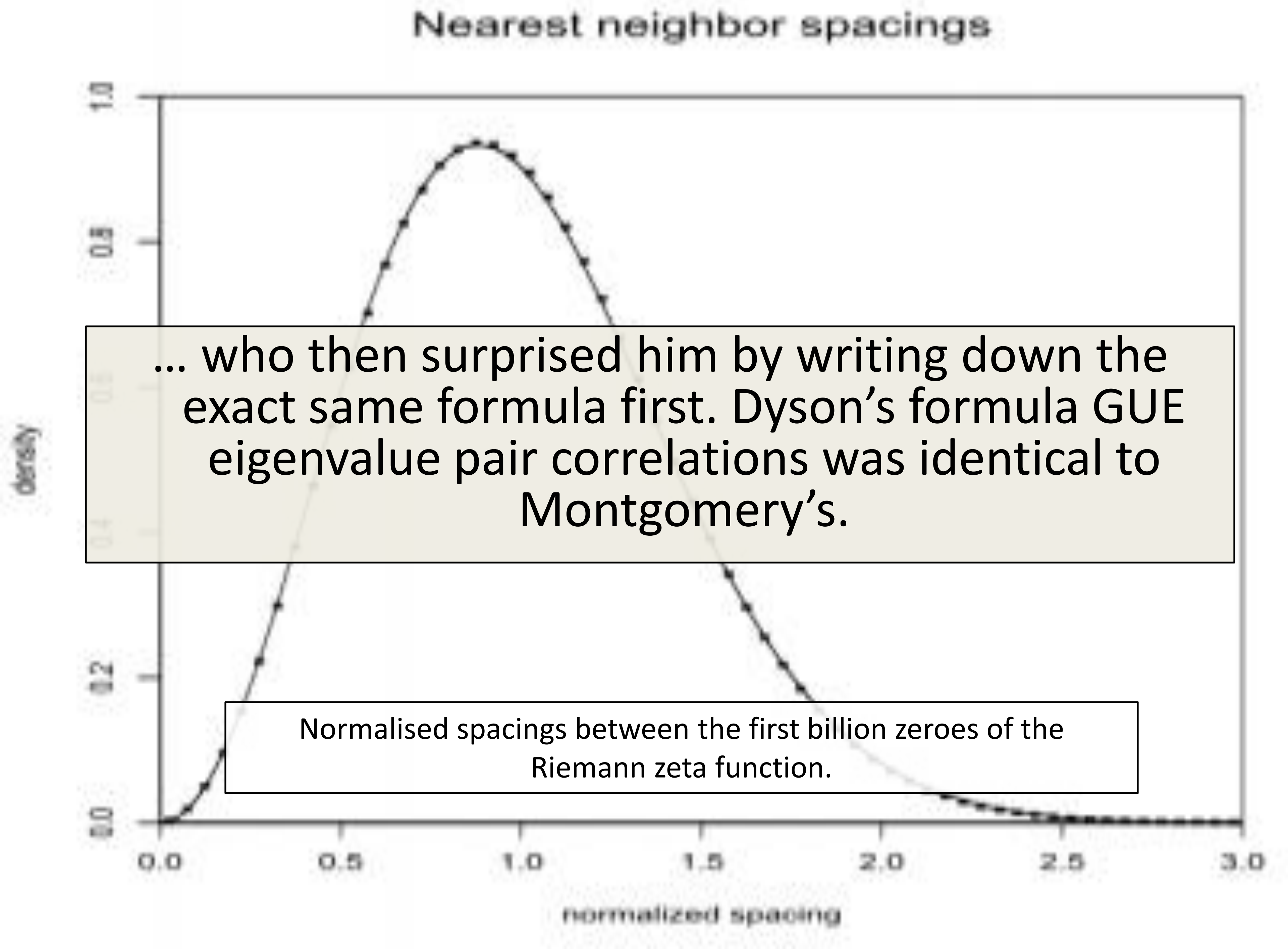
$$\sum_{n=1}^{\infty} \sum_{p:p^n \leq x} \log p = x - \sum_{\rho} \frac{x^\rho}{\rho} - \log(2\pi) - \log(1 - x^{-2})/2$$

It was shown by Hardy and Littlewood that infinitely many of these zeroes ρ lie on the critical line $\{s : \mathrm{Re}(s) = 1/2\}$. One can then study the spacing between these zeroes on this line.

G. H. Hardy FRS (1877-1947)

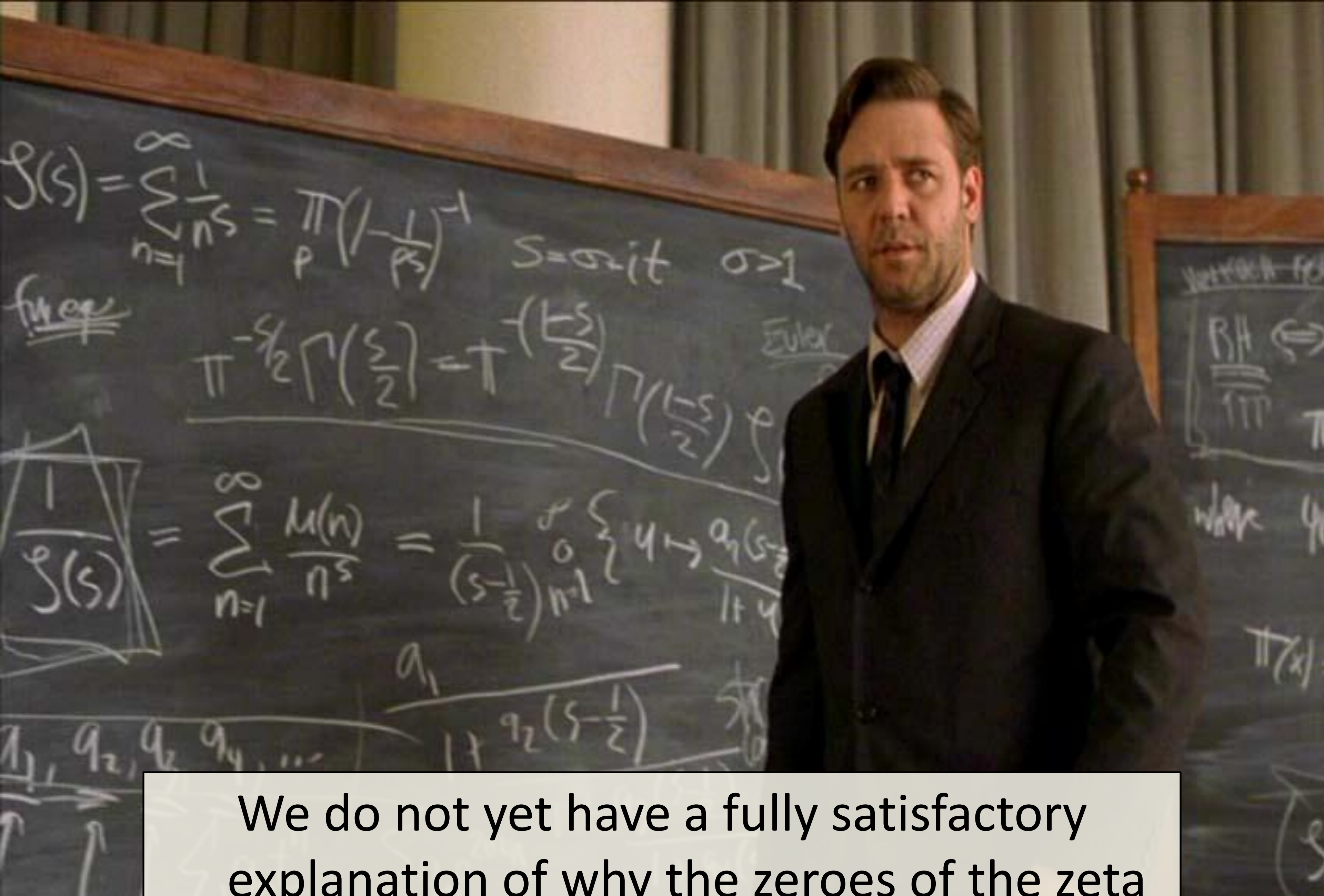
John Edensor Littlewood (1885-1977)

Hugh Montgomery formulated the *pair correlation conjecture* for these zeroes, after obtaining some partial results. He was about to describe this conjecture to Dyson at the Institute for Advanced Study…



… who then surprised him by writing down the exact same formula first. Dyson's formula GUE eigenvalue pair correlations was identical to Montgomery's.

Normalised spacings between the first billion zeroes of the Riemann zeta function.



We do not yet have a fully satisfactory explanation of why the zeroes of the zeta function obey this universal spacing law.

Math. Res. Lett. **17** (2010), *no.* 04, 667–674

BULK UNIVERSALITY FOR WIGNER HERMITIAN MATRICES WITH SUBEXPONENTIAL DECAY

László Erdős, José Ramírez, Benjamin Schlein, Terence Tao, Van Vu, and Horng-Tzer Yau

Abstract. In this paper, we consider the ensemble of $n \times n$ Wigner Hermitian matrices $H = (h_{\ell k})_{1 \leq \ell, k \leq n}$ that generalize the Gaussian unitary ensemble (GUE). The matrix elements $h_{k\ell} = \bar{h}_{\ell k}$ are given by $h_{\ell k} = n^{-1/2}(x_{\ell k} + \sqrt{-1} y_{\ell k})$, where $x_{\ell k}, y_{\ell k}$ for $1 \leq \ell < k \leq n$ are i.i.d. random variables with mean zero and variance 1/2, $y_{\ell\ell} = 0$ and $x_{\ell\ell}$ have mean zero and variance 1. We assume the distribution of $x_{\ell k}, y_{\ell k}$ to

But last year, it was shown by a number of mathematicians (including myself) that this universal law held for a large class of models known as *Wigner random matrix models*.

Let $(p'_n)^{(k)}$ be the k-point correlation function associated to H'. The first step is to show that Theorem 2 is valid for H':

Proposition 4. *The quantity*

$$\frac{1}{2\varepsilon}\int_{u-\varepsilon}^{u+\varepsilon}\int_{\mathbb{R}^k} f(\alpha_1,\dots,\alpha_k)$$
$$\times \frac{1}{[\rho_{sc}(u')]^k}(p'_n)^{(k)}\left(u'+\frac{\alpha_1}{n\rho_{sc}(u')},\dots,u'+\frac{\alpha_k}{n\rho_{sc}(u')}\right)d\alpha_1\dots d\alpha_k du'$$

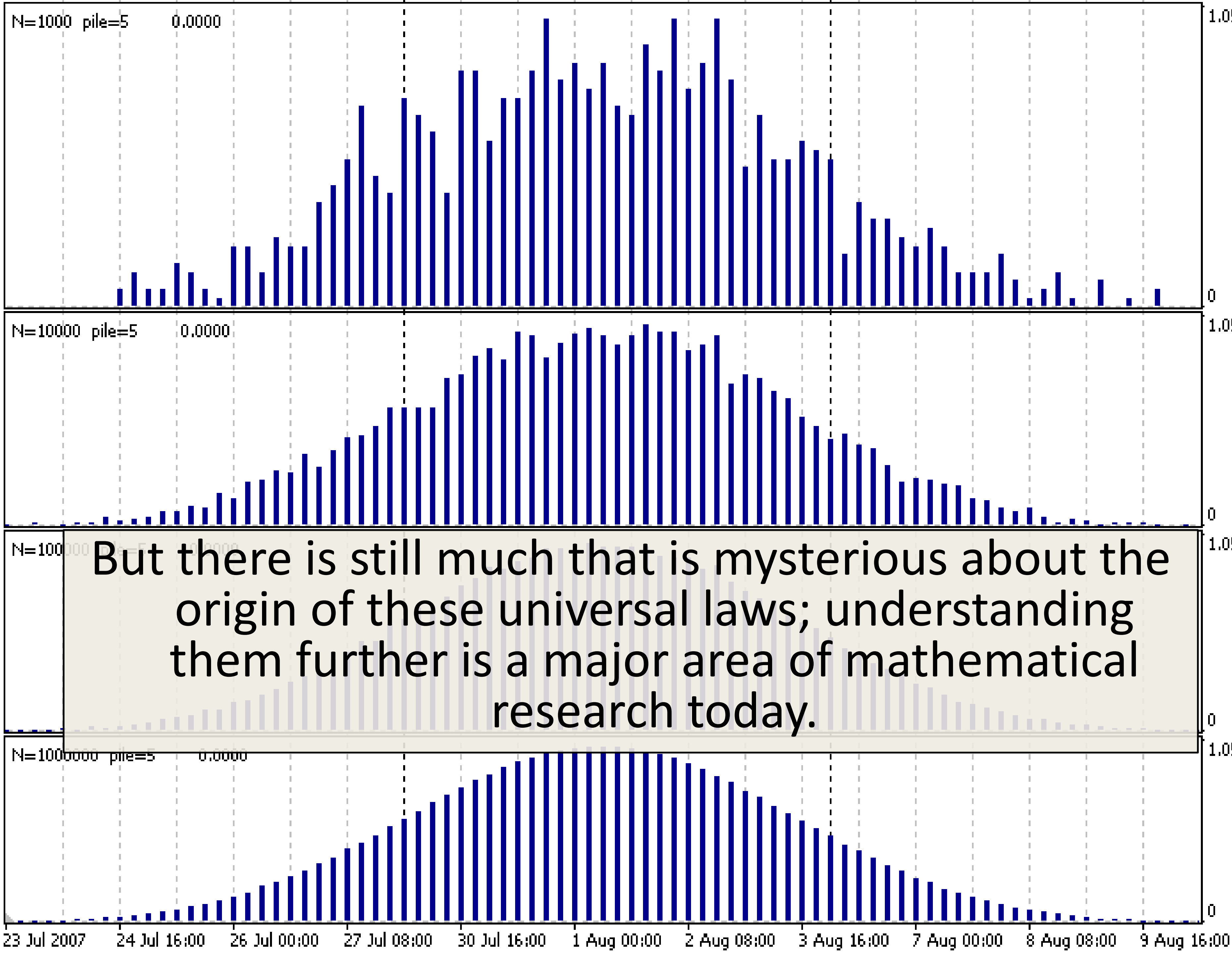
converges as $n\to\infty$ to

$$\int_{\mathbb{R}^k} f(\alpha_1,\dots,\alpha_k)\det(K(\alpha_i,\alpha_j))_{i,j=1}^k\, d\alpha_1\dots d\alpha_k.$$

Proof. By the dominated convergence theorem, it suffices to show that

The proof is too technical to give here, but let me just mention that previous universality results (such as the central limit theorem) are an essential ingredient in the proof.

for every fixed $-2<u<2$. The dominated convergence theorem can be applied since the integral on the l.h.s. of (9) can be bounded by $C_k\,\mathbf{E}\,N_I^k$, where N_I is



But there is still much that is mysterious about the origin of these universal laws; understanding them further is a major area of mathematical research today.