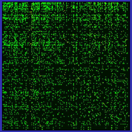
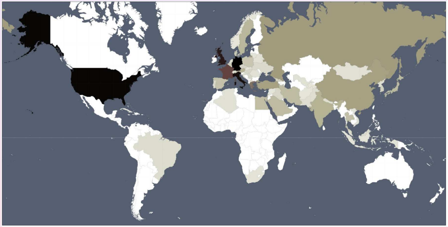
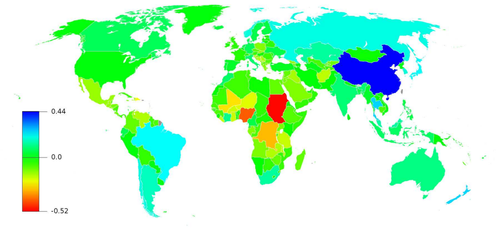
Eigenstates of Google matrix of Wikipedia, world trade and other networks



Dima Shepelyansky (CNRS, Toulouse)
www.quantware.ups-tlse.fr/dima

FET Open NADINE project No 288956 Coordinator





* fractal Weyl law [M6-WP2.3]
* Anderson transition on directed networks [M11-WP2.4]
* 24wiki ranking of historical figures [M13-WP4.3-WP5.2]
* multiproduct world trade network (UN COMTRADE + WTO in short) [M8-WP4.4]

Fractal Weyl law [P1.12, P.15]

invented for open quantum systems, quantum chaotic scattering:
the number of Gamow eigenstates N_γ, that have escape rates γ in a finite bandwidth $0 \leq \gamma \leq \gamma_b$, scales as

$$N_\gamma \propto \hbar^{-\nu} \propto N^\nu, \quad \nu = d/2$$

where d is a fractal dimension of a strange invariant set formed by obits non-escaping in the future and in the past (N is matrix size)

References:
J.Sjostrand, Duke Math. J. **60**, 1 (1990)
M.Zworski, Not. Am. Math. Soc. **46**, 319 (1999)
W.T.Lu, S.Sridhar and M.Zworski, Phys. Rev. Lett. **91**, 154101 (2003)
S.Nonnenmacher and M.Zworski, Commun. Math. Phys. **269**, 311 (2007)

Resonances in quantum chaotic scattering:
three disks, quantum maps with absorption

Perron-Frobenius operators, Ulam method for dynamical maps, Ulam networks, dynamical maps, strange attractors

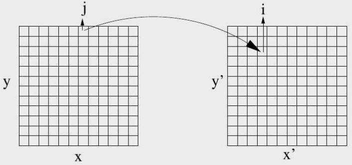
Linux kernel network $d = 1.3$, $N \leq 285509$;
Phys. Rev. up to 2009 $d \approx 1$, $N = 460422$

Ulam networks

Ulam conjecture (method) for discrete approximant of Perron-Frobenius operator of dynamical systems

Discretized phase-space:
Adjacency matrix $\mathbf{A} = P(j \to i)$

$N = N_x \times N_y$ cells.



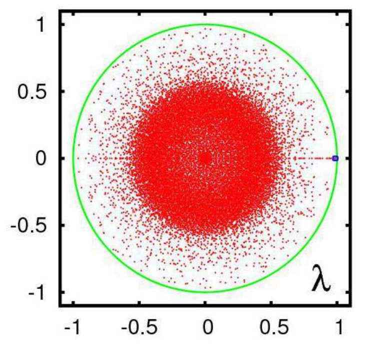
N_c: traj. from cell j
N_i: traj. to call i

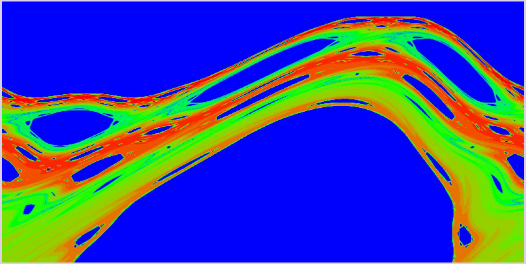
$$\begin{cases} \mathbf{A}_{i,j} = N_i/N_c \\ \sum_i \mathbf{A}_{i,j} = 1 \quad \text{(closed systems)} \end{cases}$$

S.M.Ulam, *A Collection of mathematical problems*, Interscience, **8**, 73 N.Y. (1960)
A rigorous prove for hyperbolic maps:
T.-Y.Li J.Approx. Theory **17**, 177 (1976)
Related works:
Z. Kovacs and T. Tel, Phys. Rev. A 40, 4641 (1989)
M.Blank, G.Keller, and C.Liverani, Nonlinearity **15**, 1905 (2002)
D.Terhesiu and G.Froyland, Nonlinearity **21**, 1953 (2008)

Links to Markov chains: ∞∞∞∞∞∞∞∞∞∞
Contre-example: Hamiltonian systems with invariant curves, e.g. the Chirikov standard map: noise, induced by coarse-graining, destroys the KAM curves and gives homogeneous ergodic eigenvector at $\lambda = 1$

Ulam method for the Chirikov standard map





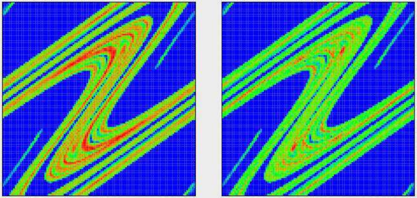
$\bar{y} = y + K \sin x, \ \bar{x} = x + \bar{y} \ \ (mod 2\pi); \ K = 0.971635...$
Left: spectrum $G\psi = \lambda\psi$, $M \times M/2$ cells; $M = 280$, $N_d = 16609$, exact and Arnoldi method for matrix diagonalization; generalized Ulam method of one trajectory.
Right: modulus of eigenstate of $\lambda_2 = 0.99878...$, $M = 1600$, $N_d = 494964$. Here $K = K_G$

Ulam method for dissipative systems

Strange repellers and strange attractors

Scattering

$$\begin{cases} \bar{y} &= y + K\sin(x + y/2) \\ \bar{x} &= x + (y + \bar{y})/2 \quad (\mathrm{mod} 2\pi) \end{cases}$$

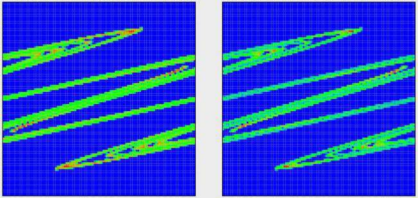


$N = 110 \times 110,\ K = 7, a = 2$

$\lambda_1 = 0.756 \quad \lambda_3 = -0.01 + i0.513$

Dissipation

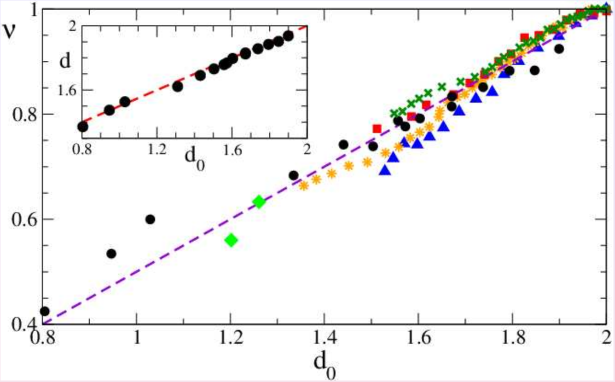
$$\begin{cases} \bar{y} &= \eta y + K\sin x \\ \bar{x} &= x + \bar{y} \quad (\mathrm{mod} 2\pi) \end{cases}$$



$N = 110 \times 110,\ K = 7, \eta = 0.3$

$\lambda_1 = 1 \quad \lambda_3 = -0.258 + i0.445$

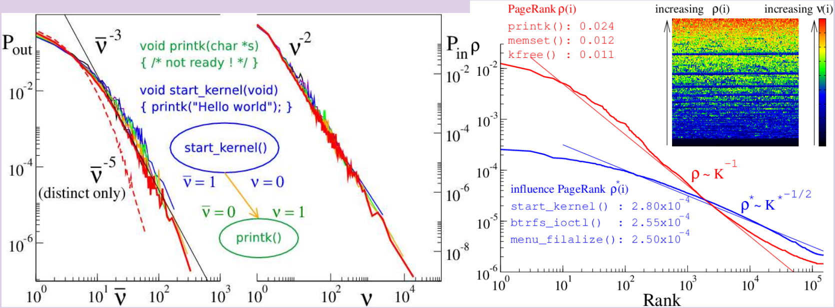
Fractal Weyl law for Ulam networks



Fractal Weyl law for three different models with dimension d_0 of invariant set. The fractal Weyl exponent ν is shown as a function of fractal dimension d_0 of the strange repeller in model 1 and strange attractor in model 2 and Henon map; dashed line shows the theory dependence $\nu = d_0/2$. Inset shows relation between the fractal dimension d of trajectories nonescaping in future and the fractal inv-set dimension d_0 for model 1; dashed line is $d = d_0/2 + 1$.

Linux Kernel Network

Procedure call network for Linux

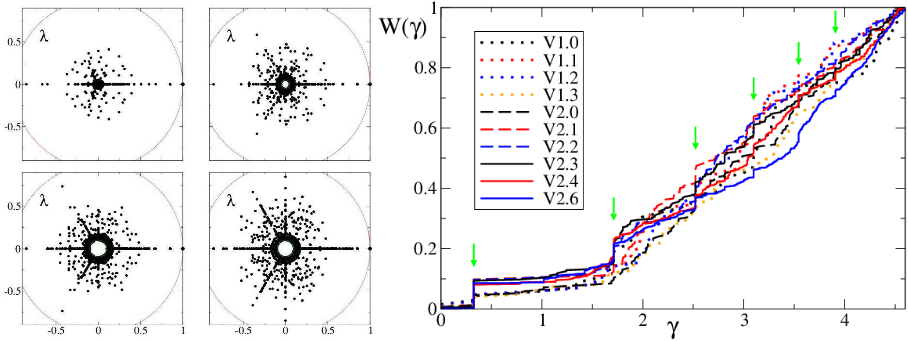


Links distribution (left); PageRank and inverse PageRank (CheiRank) distribution (right) for Linux versions up to 2.6.32 with $N = 285509$ ($\rho \sim 1/j^{\beta}, \beta = 1/(\nu - 1)$).

(Chepelianskii arxiv:1003.5455)

Fractal Weyl law for Linux Network

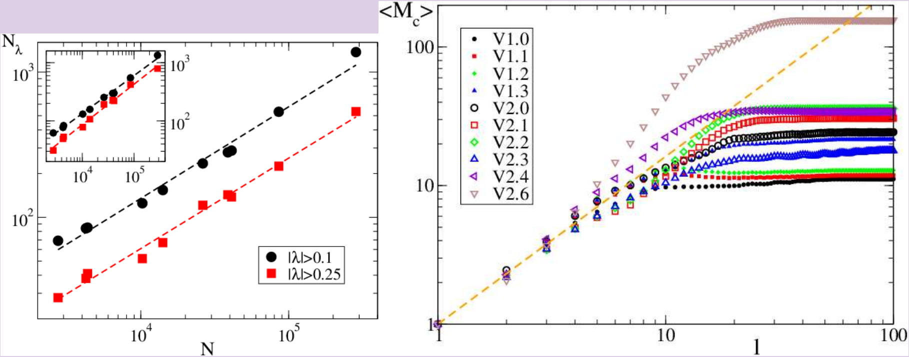
Sjöstrand Duke Math J. 60, 1 (1990), Zworski *et al.* PRL 91, 154101 (2003) $\rightarrow$ quantum chaotic scattering;
Ermann, DS EPJB 75, 299 (2010)$\rightarrow$ Perron-Frobenius operators



Spectrum of Google matrix (left); integrated density of states for relaxation rate $\gamma = -2\ln|\lambda|$ (right) for Linux versions, $\alpha = 0.85$.

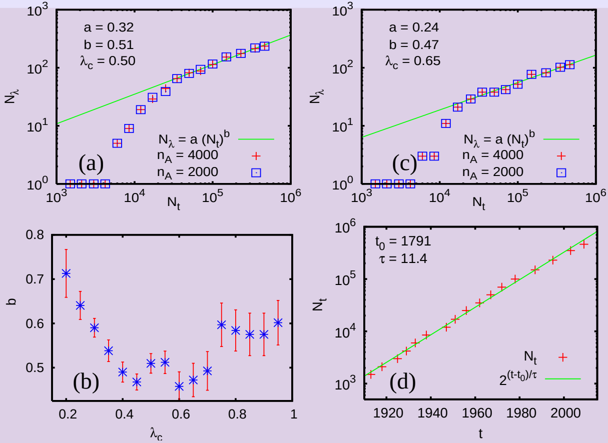
Fractal Weyl law for Linux Network

Number of states $N_\lambda \sim N^\nu, \quad \nu = d/2 \quad (N \sim 1/\hbar^{d/2})$



Number of states N_λ with $|\lambda| > 0.1; 0.25$ vs. N, lines show $N_\lambda \sim N^\nu$ with $\nu \approx 0.65$ (left); average mass $< M_c >$ (number of nodes) as a functon of network distance l, line shows the power law for fractal dimension $< M_c > \sim l^d$ with $d \approx 1.3$ (right).

Fractal Weyl law for Physical Review network



Panel (a) (or (c)): shows the number N_λ of eigenvalues with $\lambda_c \leq \lambda \leq 1$ for $\lambda_c = 0.50$ (or $\lambda_c = 0.65$) versus the network size N_t (up to time t). The green line shows the fractal Weyl law $N_\lambda = a\,(N_t)^b$ with parameters $a = 0.32 \pm 0.08$ ($a = 0.24 \pm 0.11$) and $b = 0.51 \pm 0.02$ ($b = 0.47 \pm 0.04$) obtained from a fit in the range $3 \times 10^4 \leq N_t < 5 \times 10^5$. *Panel (b):* exponent b with error bars obtained from the fit $N_\lambda = a\,(N_t)^b$ in the range $3 \times 10^4 \leq N_t < 5 \times 10^5$ versus cut value λ_c. *Panel (d):* effective network size N_t versus cut time t (in years). The green line shows the exponential fit $2^{(t-t_0)/\tau}$ with $t_0 = 1791 \pm 3$ and $\tau = 11.4 \pm 0.2$. Thus $N = 463348$, $N_\ell = 4684496$, $d \approx 1$, $b = \nu \approx 0.5$.

Anderson transition on directed networks [P1.12, P1.15, P1.20]

Anderson (1958) metal-insulator transition for electron transport in disordered solids (Nobel prize 1977)

$H = \epsilon_n \psi_n + V(\psi_{na+1} + \psi_{n-1}) = E\psi_n \; ; \quad -W/2 < \epsilon_n < W/2$

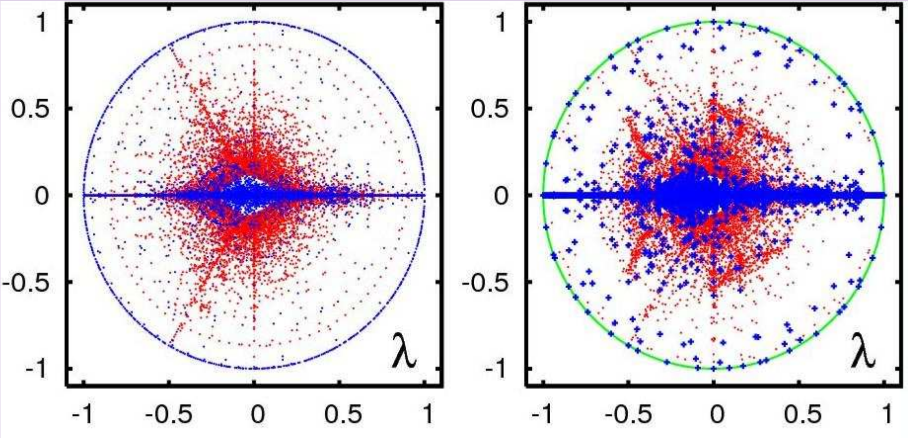
In dimensions $d = 1, 2$ all eigenstates are exponentially localized, insulating phase. At $d = 3$ for $W > 16.5V$ all eigenstates are exponentially localized, for $W < 16.5V$ there are metalic delocalized states, mobility edge, metalic phase

Random Matrix Theory - RMT (Wigner (1955)) for Hermitian and unitary matrices (quantum chaos, many-body quantum systems, quantum computers)

Google matrix, Markov chains, Perron-Frobenium operators:
=> complex spectrum of eigenvalues; new field of research
Can we have the Anderson transition for Google matrix? All the world would go blind if PageRank is delocalized What are good RMT models of Google matrix? Subspaces and core

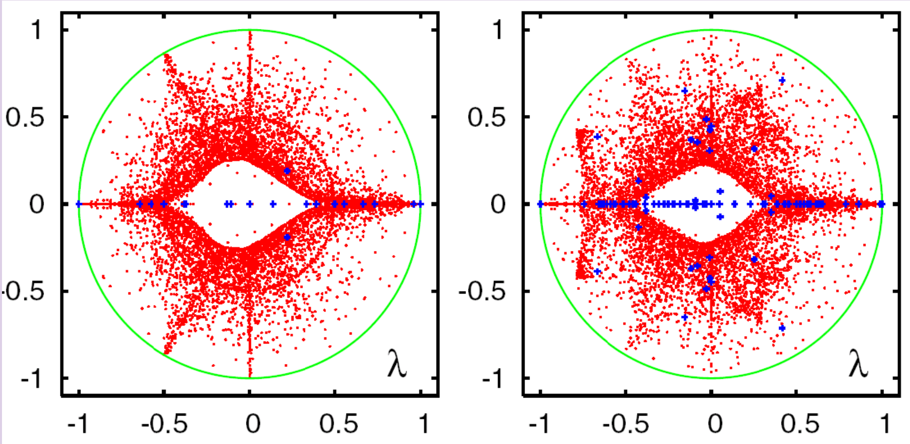
$$\mathbf{S} = \begin{pmatrix} \mathbf{S_{ss}} & \mathbf{S_{sc}} \\ 0 & \mathbf{S_{cc}} \end{pmatrix}$$

Spectrum of UK University networks



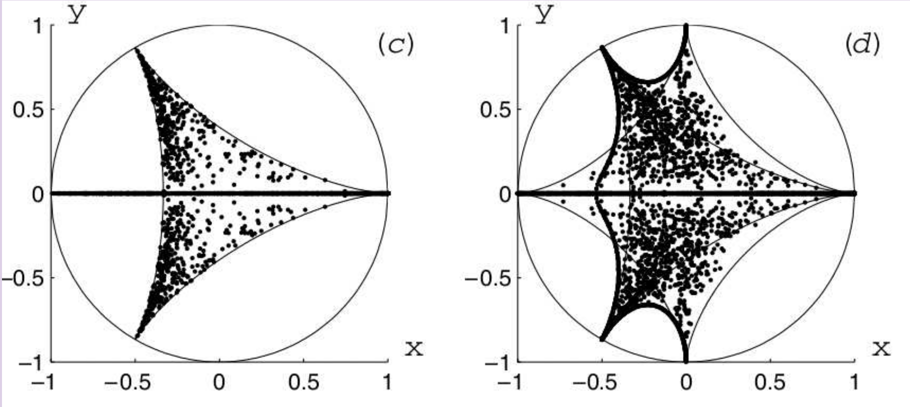
Arnoldi method: Spectrum of Google matrix for Univ. of Cambridge (left) and Oxford (right) in 2006; 2% at $\lambda = 1$ ($N \approx 200000$, $\alpha = 1$).

Spectrum of UK University networks



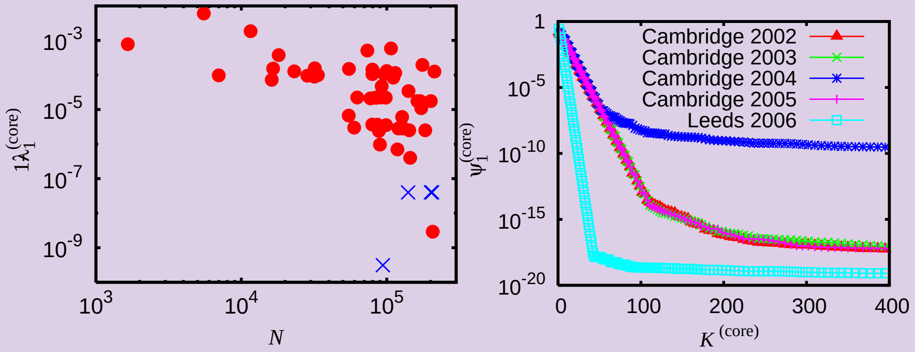
Spectrum of CheiRank Google matrix for Univ. of Cambridge (left) and Oxford (right) in 2006 ($N \approx 200000$, $\alpha = 1$)

Spectrum of random orthostochastic matrices



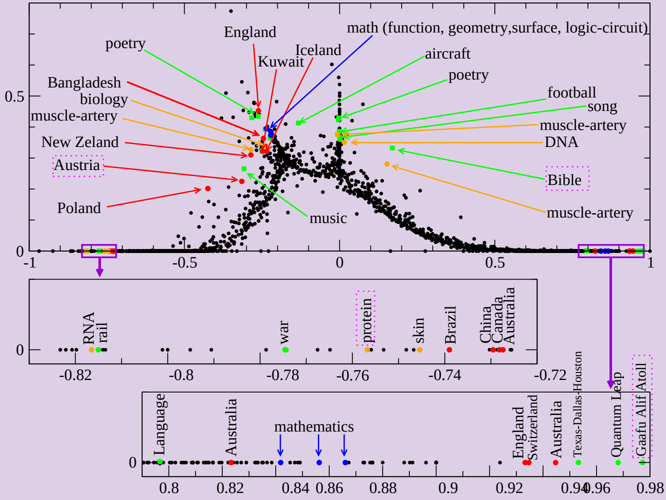
Spectrum $N = 3$ (left), 4 (right) [K.Zyczkowski *et al.* J.Phys. A **36**, 3425 (2003)]

Gap of core space at $\alpha = 1$



(Left) Gap vs *N* for universities Glasgow, Cambridge, Oxford, Edinburgh, UCL, Manchester, Leeds, Bristol and Birkbeck (2002-2006) and Bath,Hull,Keele,Kent,Nottingham, Aberdeen, Sussex, Birmingham, East Anglia, Cardiff, York (2006). Red dots are for gap $> 10^{-9}$ and blue crosses (moved up by 10^9) are for Cambridge 2002, 2003 and 2005 and Leeds 2006 with gap $< 10^{-16}$; point at $2.91 \cdot 10^{-9}$ is Cambridge 2004. (Right) First core eigenstates

Wikipedia spectrum and eigenstates



Spectrum S of EN Wikipedia, Aug 2009, $N = 3282257$. Eigenvalues-communities are labeled by most repeated words following word counting of first 1000 nodes. (Ermann, Frahm, DS 2013)

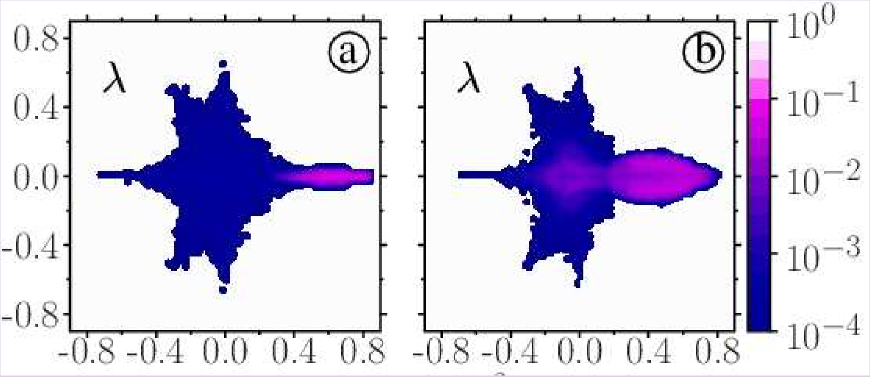
Random Matrix Models of directed networks

We use the usual Anderson model with diagonal disorder terms W_i and transitions V to nearby sites on a lattice in dimension d:

$$W_{\mathbf{i}}\psi_{\mathbf{i}} + V\psi_{\mathbf{i+1}} + V\psi_{\mathbf{i-1}} = \lambda\psi_{\mathbf{i}} \ , \tag{1}$$

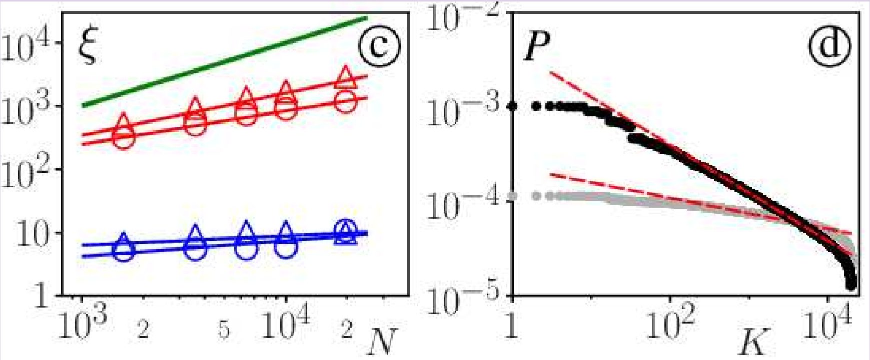
where indexes in bold are vectors in d-dimensional space. On this we construct the matrices S and G for $d = 2, 3$. The matrix S is constructed as follows: each transition matrix element, corresponding to V terms, in the Anderson model in dimension d is replaced by a random number ε_i uniformly distributed in the interval $[0, \varepsilon_{max}/2d]$, the diagonal element $W_{\mathbf{i}}$ is replaced by unity minus the sum of all ε_i over $2d$ nearby sites $(1 - \sum_{i=1}^{2d} \varepsilon_i)$. The asymmetric matrix S constructed in this way belongs to the Google matrix class.
By replacing matrix elements in the model AD2 by blocks B of size 4×4 we obtain the **model AD2Z**. In a similar way for the model AD2S we obtain the **model AD2ZS** with block shortcuts. In this case we restrict our studies only for dimension $d = 2$.

Anderson transition on directed networks



Panels show distribution of IPR values ξ (number of nodes contributing to an eigenstate) on the plane λ or eigenvalues of G matrix for two nodels of directed networks with disorder; color shows the ratio ξ/N, $\alpha = 0.85$.

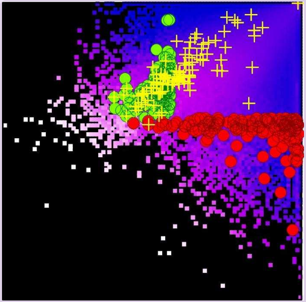
Anderson transition on directed networks

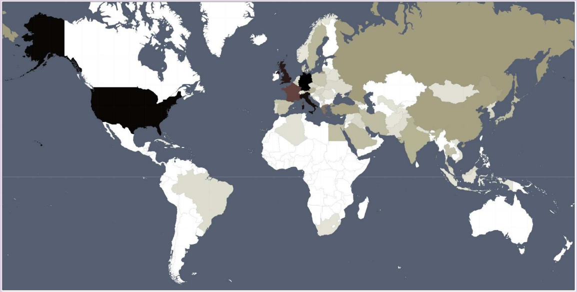


Panel (c): dependence of ξ on *N* for AD2Z with triangles for states with λ located in the delocalized domain $\mathrm{Re}\lambda \in (0.3, 0.85)$ (red triangles, fit gives $\nu = 0.67$) and in the localized domain $\mathrm{Re}\lambda < -0.5$ (blue triangles, $\nu = 0.15$); for AD2ZS at $\delta = 0.25$ with circles for states with λ located in the delocalized domain $\mathrm{Re}\lambda \in (0.2, 0.85)$ (red circles, $\nu = 0.53$) and in the quasi-localized domain $\mathrm{Re}\lambda < -0.5$ (blue circles, $\nu = 0.25$); fits are shown by lines, green line shows $\xi = N$. Panel (d): dependence of PageRank probability *P* on PageRank index *K* for models AD2Z (gray symbols) and AD2ZS at $\delta = 0.25$ (black symbols); the fits for the range $K \in (100, 6000)$ are shown by dashed lines with $\beta = 0.16$ (AD2Z) and $\beta = 0.51$ (AD2ZS) for the parameters of panels (a,b).

Top historical figures of 24 Wikipedia editions [P4.16]

2DRanking of Wikipedia articles; top 100 historical figures;
comparison with historical studies of M.Hart (37 and 43 percent overlap)
35 centures and all countries by birth place; 17 millions wiki-articles

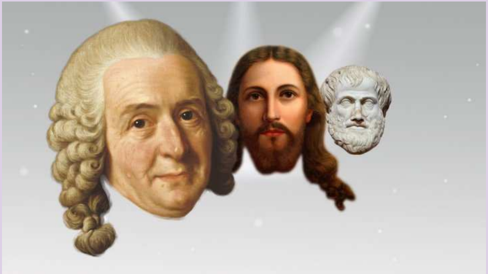




A.Zhirov, O.Zhirov, DLS EPJB (2010); Y.-H.Eom, K.M.Frahm, A.Benczur, DLS EPJB (2013); Y.-H.Eom, DLS PLoS ONE (2013), Y.-H.Eom,P.Aragon, D.Laniado, A.Kaltenbrunner, S.Vigna, DLS arXiv2014 - PLoS ONE (2015)

Top historical figures of 24 Wikipedia editions

Top global PageRank historical figures: Carl Linnaeus, Jesus, Aristotle ...



theguardian

News | Sport | Comment | Culture | Business | Money | Life & style | Jobs

Comment is free

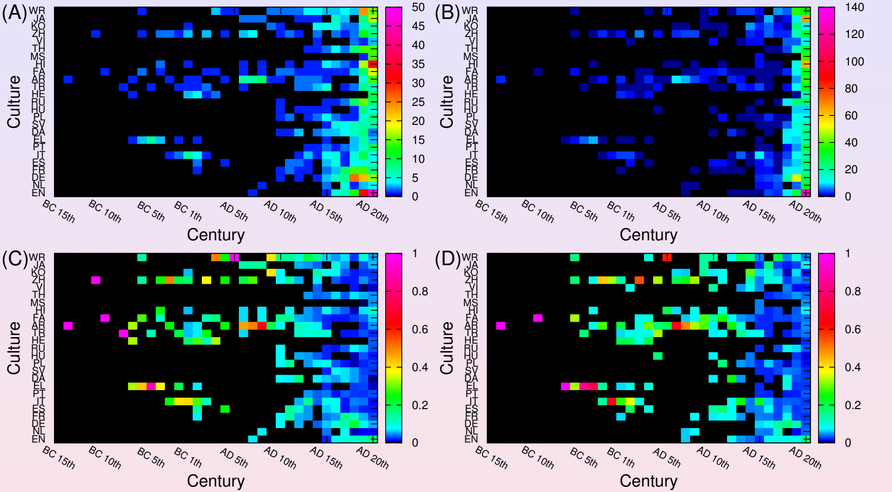
And the winner of Wikipedia's influence list is ... an 18th century botanist. Hear hear

Carl Linnaeus is hardly a household name, but the Swedish doctor who created a global naming system for species deserves this accolade

Patrick Barkham
theguardian.com, Friday 13 June 2014 09.00 BST
Jump to comments (51)

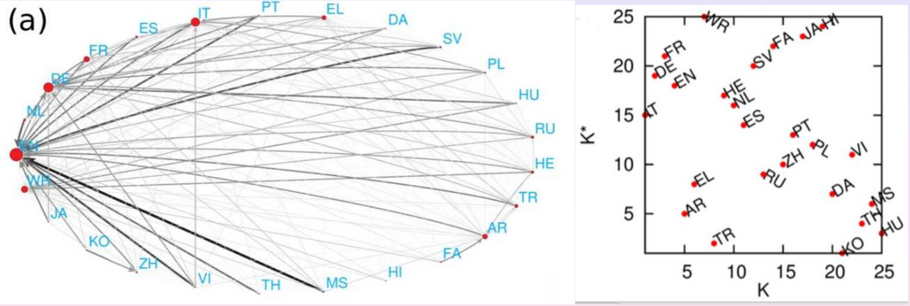
Media highlights: The Guardian, The Independent, The Washington Post, France24, EC CORDIS, Uppsala Universitet: “Carl Linnaeus ranked most influential person of all time” ... (about 20 countries)
Competitors: MIT Pantheon project http://pantheon.media.mit.edu (2014); Stony-Brook NY http://www.whoisbigger.com/ (2014)

Time evolution of 35 centuries



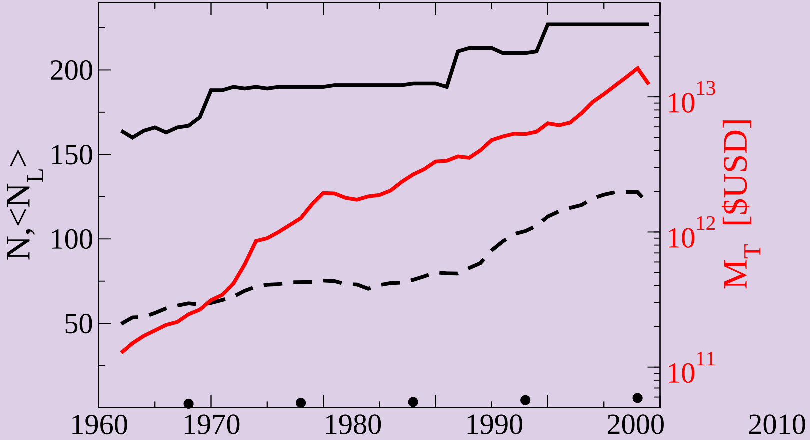
Birth date distribution of historical figures from the global PageRank list (A,C, 1045 persons) and 2DRank list (B,D, 1616 persons). Each historical figure is attributed to her/his own language according to her/his birth place as described in the paper (if the birth place is not among our 24 languages then a person is attributed to the remaining world (WR)). Color in panels (A,B) shows the total number of persons for a given century, while in panels (C,D) color shows a percent for a given century (normalized to unity in each column).

Entanglement of cultures



(Left) Top PageRank historical figures, the link width and darkness are proportional to a number of foreign historical figures quoted in top 100 of a given culture, the link direction goes from a given culture to cultures of quoted foreign historical figures
(Right) PageRank-CheiRank plane fro the network of cultures based on all PageRank historical figures boorn before AG 19th century.

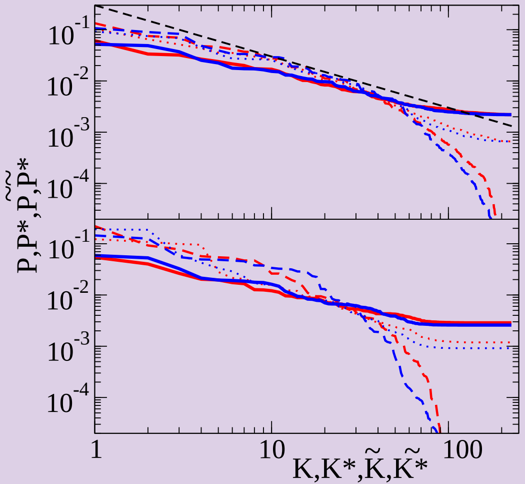
World trade network (WTN) of United Nations COMTRADE 1962-2009



Number of countries (black), links (dashed/points) and mass volume in USD (red)

Leonardo Ermann, DS arxiv:1103.5027 (2011)

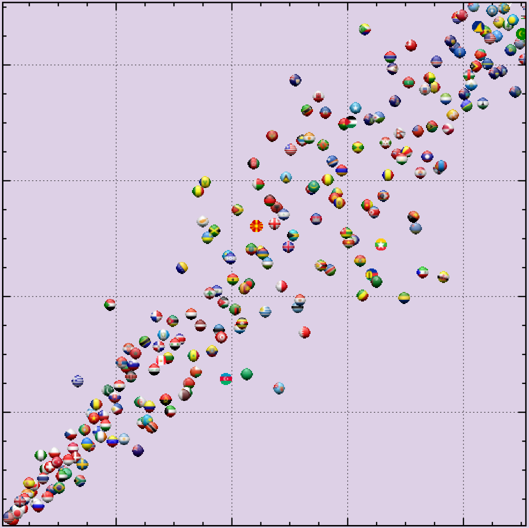
PageRank, CheiRank of World Trade



Year 2008: Probabilities of PageRank $P(K)$ (red), CheiRank $P^*(K^*)$ (blue) for all commodities (top) and crude petroleum (bottom), $\alpha = \mathbf{0.5}$; 0.85 (full/dotted); (dashed curves are for ImportRank, ExportRank); dashed line Zipf law $P \sim 1/K$; 227 countries

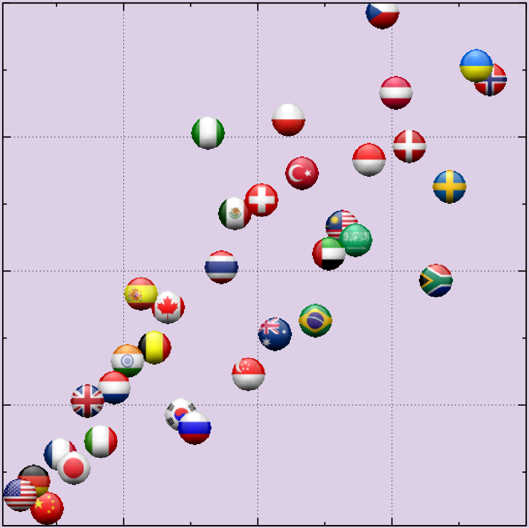
Ranking of World Trade

2008: All commodities



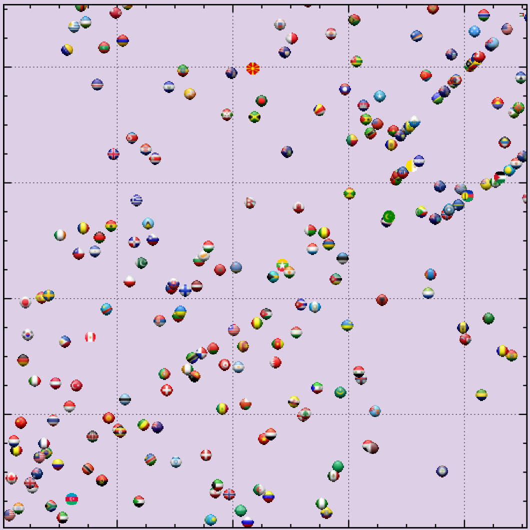
Ranking of World Trade

2008: All commodities



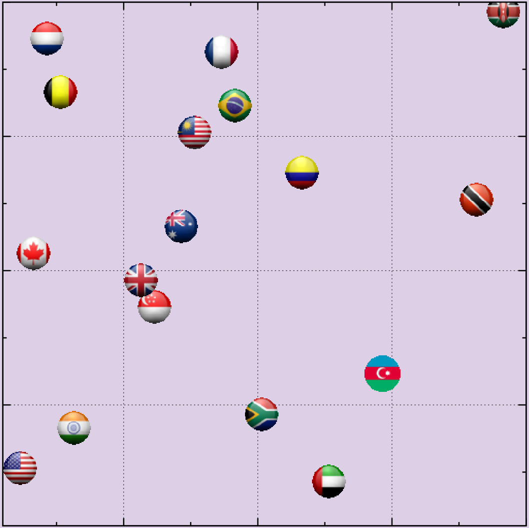
Ranking of World Trade

2008: Crude petroleum



Ranking of World Trade

2008: Crude petroleum



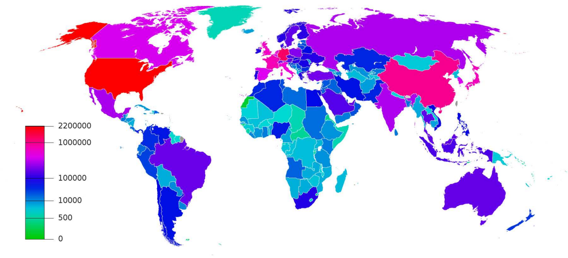
Rank table 2008 (74% of countries of G20)

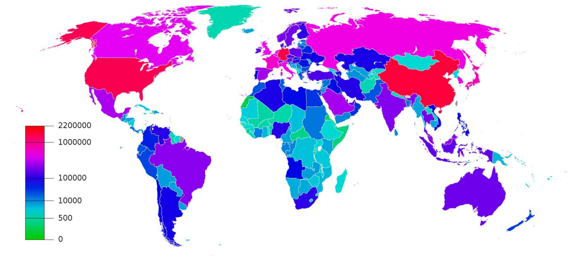
Table 1. Top 20 ranking for *all commodities* – 2008.

Ran	K	K^*	K_2	$\tilde{K}$	$\tilde{K}^*$
1	USA	China	USA	USA	China
2	Germany	USA	China	Germany	Germany
3	China	Germany	Germany	China	USA
4	France	Japan	Japan	France	Japan
5	Japan	France	France	Japan	France
6	UK	Italy	Italy	UK	Netherlands
7	Italy	Russian Fed.	UK	Netherlands	Italy
8	Netherlands	● Rep. of Korea	Netherlands	Italy	Russian Fed.
9	India	UK	India	Belgium	UK
10	Spain	Netherlands	Rep. of Korea	Canada	Belgium
11	Belgium	● Singapore	Belgium	Spain	● Canada
12	Canada	● India	Russian Fed.	Rep. of Korea	● Rep. of Korea
13	Rep. of Korea	Belgium	Canada	Russian Fed.	Mexico
14	Russian Fed.	Australia	Spain	Mexico	Saudi Arabia
15	Nigeria	Brazil	Singapore	Singapore	● Singapore
16	Thailand	● Canada	Thailand	India	Spain
17	Mexico	Spain	Australia	Poland	Malaysia
18	Singapore	South Africa	Brazil	Switzerland	Brazil
19	Switzerland	Thailand	Mexico	Turkey	● India
20	Australia	U. Arab Emir.	U. Arab Emir.	Brazil	Switzerland

Multiproduct WTN + WNEA of WTO [P1.19, P1.20]

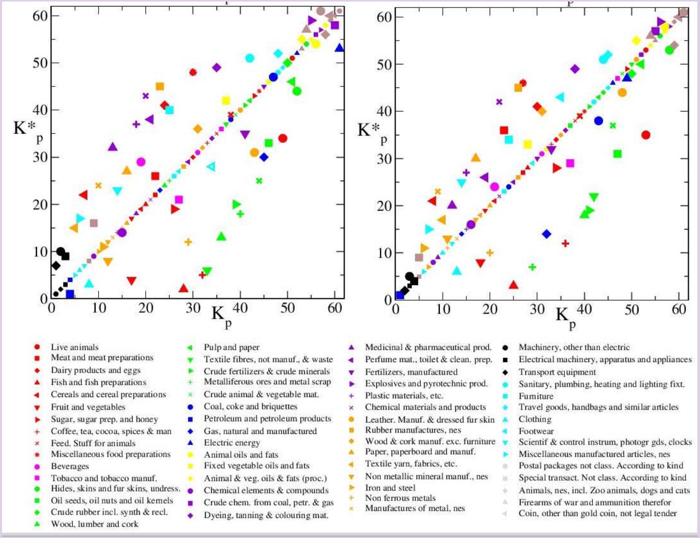
Example: year 2008, $N_c = 227$, $N_p = 61$, $N = 13847$, import (top) - export (bottom) in millions USD





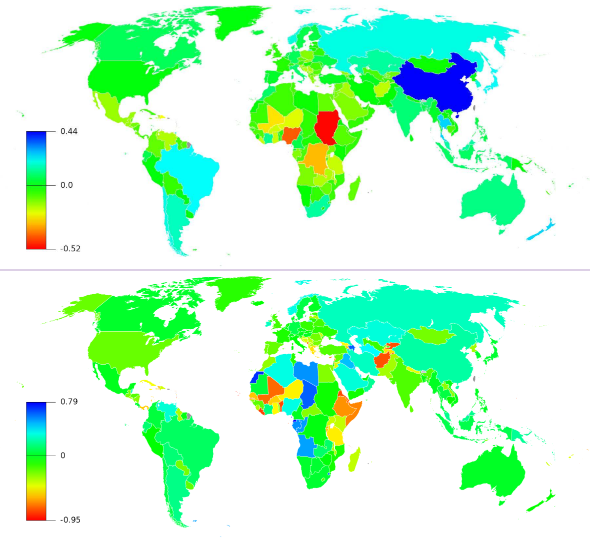
Multiproduct WTN: ranking of products

Democracy in countires, volume fraction in products => personalized vector in *G*. Left: 1993, right: 2008



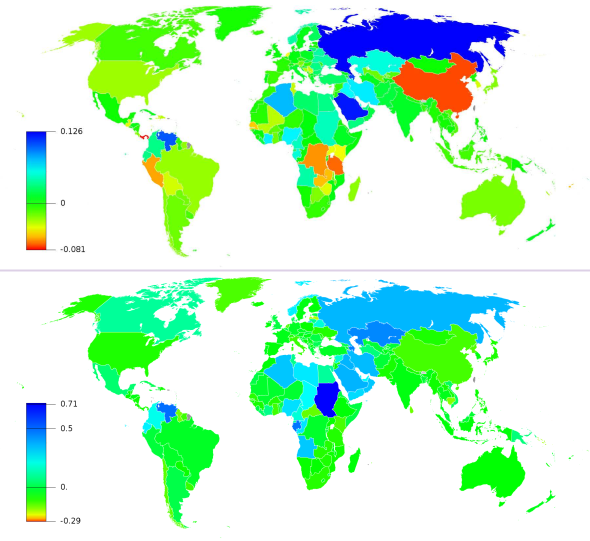
CheiRank-PageRank balance (2008)

$B_c = (P_c{}^* - P_c)/(P_c{}^* + P_c)$ (top - CheiRank-PageRank; bottom -Export-Import volume)



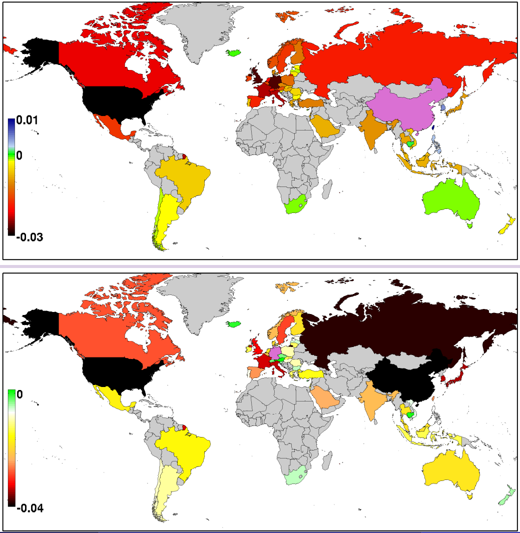
Sensitivity to price of petroleum (2008)

$dB_c/d\delta_p$ (top - CheiRank-PageRank; bottom -Export-Import volume)

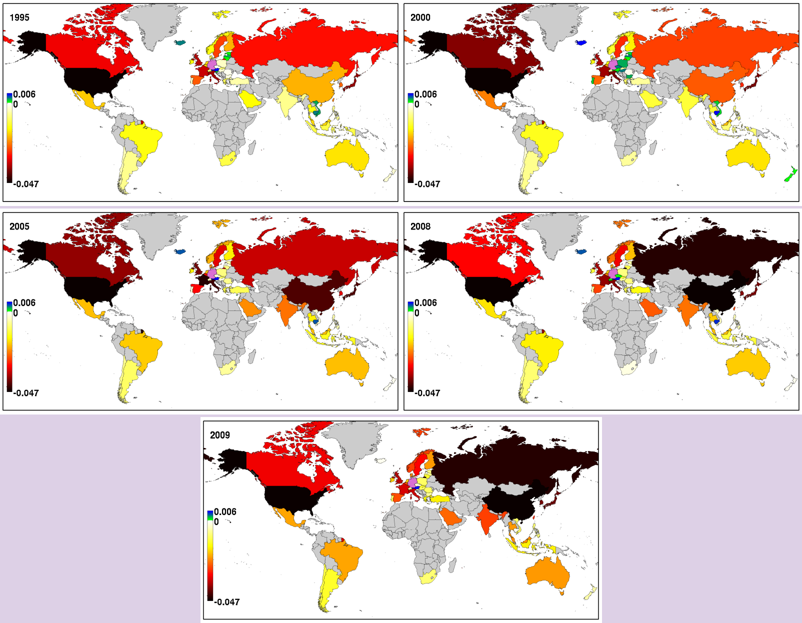


WNEA of OECD-WTO (2008)

World network of economic activities: countries $N_c = 58$, activity sectors $N_s = 37$ (with V.Kandiah and H.Escaith (WTO Geneve)); $dB_c/d\sigma_{c'}$ sensitivity to labor cost of China (top), Germany (bottom)



Future FAPLIDIN: WNEA-WTO-OECD of DEU



Future FAPLIDIN: WNEA EUROZONE