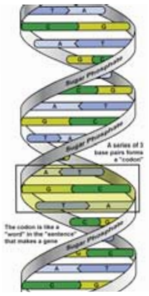
ANALYSIS OF DNA STRUCTURE AS A 2D WALK BY COMPLEX WAVELET TRANSFORM

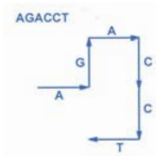
E.B.Postnikov*, A. Loskutov, S.A.Larionov, E. Ryadchenko, R.V. Kiseliov*
*Theoretical Physics Deptartment, Kursk State University, Kursk, Russia
Physics Faculty, Moscow State University, Moscow, Russia
http://chaos.phys.msu.ru/postnicov/



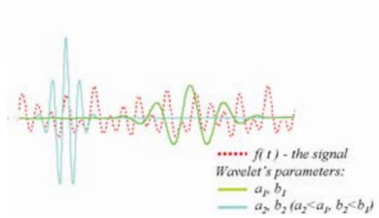
A - Adenine
T - Thymine
C - Cytosine
G - Guanine

Genome: ...AGACCT...

Representation of Genome as 2D walk (M. Gates, 1985)

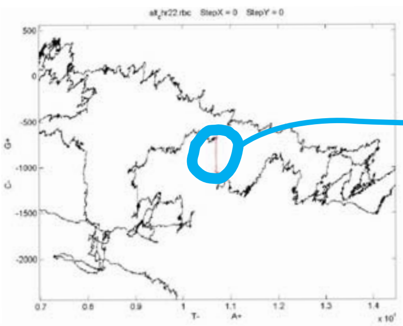


Wavelet transform with the Morlet wavelet



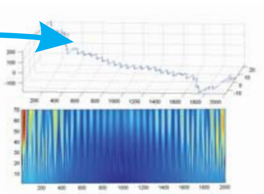
$$w(a,b)=\frac{1}{a}\int_{-\infty}^{+\infty} f(t)e^{i\omega_0\frac{t-b}{a}}e^{-\frac{1}{2}\left(\frac{t-b}{a}\right)^2}dt$$

Genome as 2D walk on a complex plane *A: +1, T: -1, C: +I, G: -i*

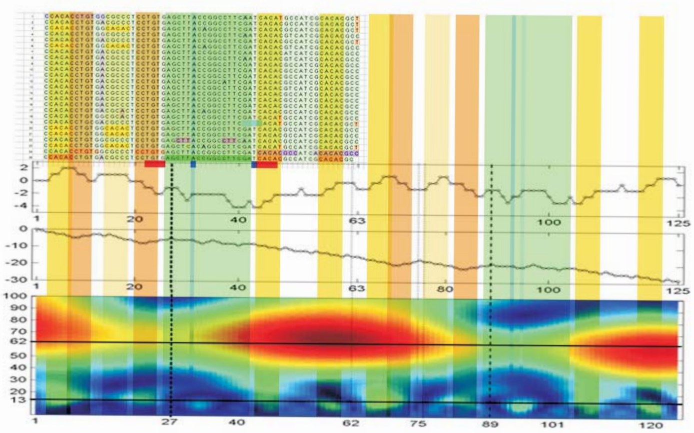


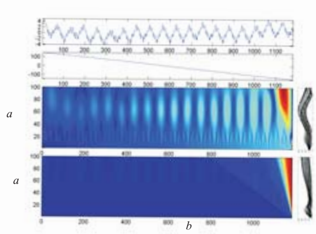
Complex signal DNA

$$f(n)=\sum_{j=1}^{n}\theta(n)=\sum_{j=1}^{n}\theta_{AT}(n)+i\theta_{CG}(n)$$



...And modulus of its wavelet transform

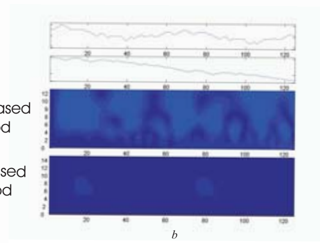




A-T
Re(f(t))
C-G
Im(f(b))

PDE-based method

FFT-based method



Advantages of PDE-based method over FFT-based algorithm for the evaluation of the wavelet transform :

* robust algorithms for the numerical solving of the diffusion-type PDE;
* there are almost no restrictions on the small scale-steps
* boundary effects can be eliminated by the choice of suitable boundary conditions for finite samples (it eliminates errors of periodization in FFT algorithms)
* PDE-based algorithm is faster then FFT-based for large samples

$O(N_a N_b)$ vs. $O(N_a N_b \log N_b)$

Algorithm based on PDE

This CWT admits the following equation (M. Haase, 2000):

$$\left(a\frac{\partial^2}{\partial b^2}-\frac{\partial}{\partial a}-i\omega_0\frac{\partial}{\partial b}\right)w(a,b)=0$$

Idea: *to find corresponding initial value and to represent integral transform as a solution of PDE*

(E. B. Postnikov, Evaluation of a Continuous Wavelet Transform by Solving the Cauchy Problem for a System of Partial Differential Equations // Computational Mathematics and Mathematical Physics, 2006, Vol. 46, No. 1, pp. 73–78.)

$$w(a,b)=\int_{-\infty}^{+\infty} f(t)\frac{e^{-\frac{1}{2}\left(\frac{t-b}{a}+i\omega_0\right)^2}}{\sqrt{2\pi a}}dt \qquad \int_{-\infty}^{+\infty}|\psi(\xi)|d\xi=e^{-\frac{\omega_0^2}{2}}$$

$$\lim_{a\to 0}\int_{-\infty}^{+\infty}\frac{e^{-\frac{1}{2}\left(\frac{t-b}{a}+i\omega_0\right)^2}}{\sqrt{2\pi a}}dt=\delta(t-b)$$

The Cauchy problem for real and imaginary parts

w(a,b)=u(a,b)+iv(a,b)

u(0,b)=Re(f(b),
v(0,b)=Im(0,b).)

$$\begin{cases}\dfrac{\partial u}{\partial a}=a\dfrac{\partial^2 u}{\partial b^2}+\omega_0\dfrac{\partial v}{\partial b}\\[2mm] \dfrac{\partial v}{\partial a}=a\dfrac{\partial^2 v}{\partial b^2}-\omega_0\dfrac{\partial u}{\partial b}\end{cases}$$

Numerical implementation

$$\frac{\beta_{j-1}}{6}\frac{\partial w_{j-1}}{\partial a}+\frac{\beta_{j-1}+\beta_j}{3}\frac{\partial w_j}{\partial a}+\frac{\beta_j}{6}\frac{\partial w_{j+1}}{\partial a}=\left(\frac{a}{\beta_{j-1}}+\frac{i\omega_0}{2}\right)w_{j-1}-\left(\frac{a}{\beta_{j-1}}+\frac{a}{\beta_j}\right)w_j+\left(\frac{a}{\beta_j}-\frac{i\omega_0}{2}\right)w_{j+1}.$$

$$\mathbf{M}(\{\beta_j\})\frac{\partial \mathbf{w}(a)}{\partial a}=\mathbf{F}(a,\{\beta_j\})\mathbf{w}(a), \qquad \beta_j=b_{j+1}-b_j$$

Step-by-step solution

$$\left(\mathbf{M}-\frac{1}{2}\mathbf{F}_{n+1}\Delta a\right)\mathbf{w}_{n+1}=\left(\mathbf{M}+\frac{1}{2}\mathbf{F}_n\Delta a\right)\mathbf{w}_n,$$

$$\Delta\left(\mathbf{M}-\frac{1}{2}\mathbf{F}_{n+1}\Delta a\right)=-\frac{1}{2}(\Delta a)^2\mathbf{A} \qquad \Delta\left(\mathbf{M}+\frac{1}{2}\mathbf{F}_{n+1}\Delta a\right)=\frac{1}{2}(\Delta a)^2\mathbf{A}$$

$$\mathbf{A}=\mathrm{tridiag}\left(\frac{1}{\beta_{j-1}},-\frac{1}{\beta_{j-1}}-\frac{1}{\beta_j},\frac{1}{\beta_j}\right). \qquad \mathbf{M}=\mathrm{tridiag}\left(\frac{\beta_{j-1}}{6},\frac{\beta_{j-1}+\beta_j}{3},\frac{\beta_j}{6}\right)$$

Number of operations: $O(N_a N_b)$

For the grid $N_a \times N_b$:
N_a times to solve the matrix system by the Thomas method ($O(N_b)$ operations)

See some additional mathematical details:

Postnikov E.B., *Evaluation of a Continuous Wavelet Transform by Solving the Cauchy Problem for a System of Partial Differential Equations //* Computational Mathematics and Mathematical Physics. 2006. V. 46. No. 1. P. 73-78.