Персистентные диаграммы: анализ данных и приложения в гемодинамике

Я. В. Базайкин[1] А. П. Чупахин[2] А. А. Черевко[2] А. К. Хе[2]

[1]Институт математики им. С. Л. Соболева СО РАН, Новосибирск
[2]Институт гидродинамики им. М. А. Лаврентьева СО РАН, Новосибирск

V конференция «Математические модели и численные методы в биоматематике»
ИВМ РАН, Москва, 29–30 октября 2013 г.

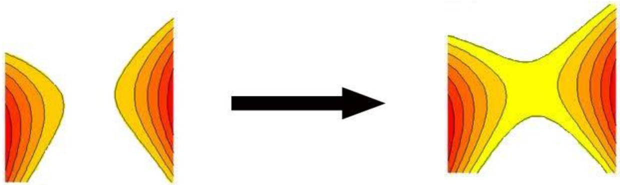
Обработка и анализ цифровых данных

- Как сравнивать числовые данные?

Определения

Введём следующие обозначения:

- X — топологическое пространство
- $f : X \mapsto \mathbb{R}$ — непрерывное отображение
- $X_a = \{p \in X \mid f(p) \leqslant a\}\ \forall a \in \mathbb{R}$
- $C_a = C(X_a)\ \forall a \in \mathbb{R}$ — множество компонент связности $X_a \subset X$
- $f_a^b : C_a \to C_b$ — отображение множеств, индуцированное вложением $X_a \subset X_b\ \forall a < b$.

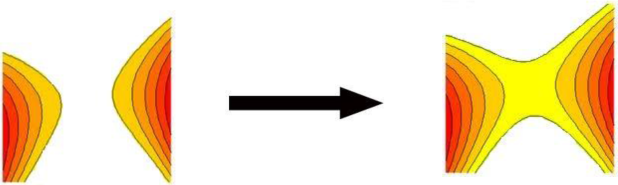


Критические значения

Значение $a \in \mathbb{R}$ будем называть *критическим* для функции a, если для любого достаточно малого $\varepsilon > 0$ отображение $f_{a-\varepsilon}^{a+\varepsilon} : C_{a-\varepsilon} \to C_{a+\varepsilon}$ не является биекцией.

Пусть $a_1, \ldots, a_n$ — все критические значения функции f.

Обозначим $C_i = C_{a_i}$ — множество компонент связности, $i = 1, \ldots, n$

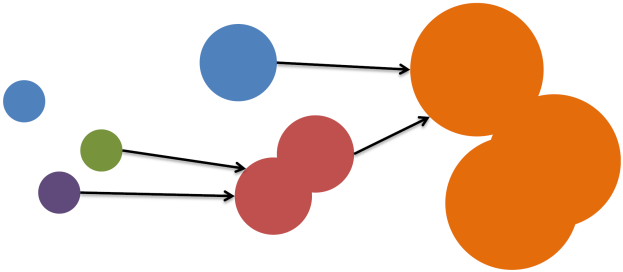


Дерево слияний

Определим граф Γ следующим образом.

1. Множество вершин: $\cup_{i=1}^{n} C_i$.
2. Каждую вершину $c \in C_i$ соединим ребром с вершиной $d = f_i(c) = f_{a_i}^{a_{i+1}}(c)$.

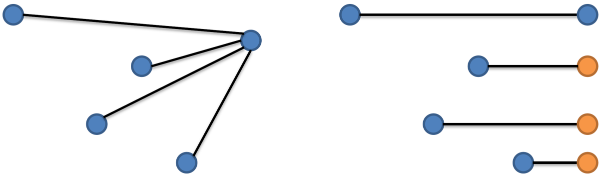
Граф Γ является деревом, которое назовём *деревом слияний*.



Баркод. Персистентная диаграмма

Для $Y \in C_i$ положим $f(Y) = a_i$ и $w(Y) = inf_{p \in Y} f(p)$.
Сконструируем граф Γ' из графа Γ следующим образом.

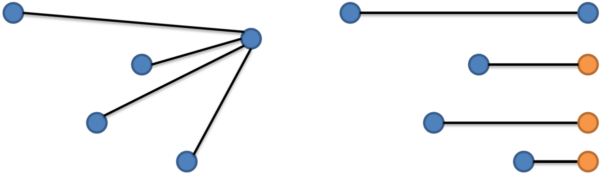
1. Пусть $d \in C_{i+1}$ — вершина Γ и $f_i^{-1}(d) = \{c_1, c_2, \ldots, c_k\} \subset C_i$.
2. Перенумеруем вершины $c_j, j = 1, \ldots, k$ так, что $w(c_1) \geq w(c_2) \geq \ldots \geq w(c_k)$.
3. Отделим каждое ребро $c_j d$, $j = 2, \ldots, k$ от графа Γ, добавив новую вершину d_j, которой это ребро будет заканчиваться.



Определение

Множество B всех таких промежутков $[c, d)$ называется *баркодом* функции f на X.

Персистентной диаграммой функции f называется множество $D(f)$ точек $(c, d) \in \mathbb{R}^2, [c, d) \in B$, объединенное с множеством $\Delta = \{(x, x)|x \in \mathbb{R}\}$ точек диагонали.



Устойчивость

Определим расстояние между двумя множествами D_1 и D_2:

$$d_B(D_1, D_2) = \inf_{\gamma} \sup_{p \in D_1} \|p - \gamma(p)\|_\infty, \quad \gamma : D_1 \to D_2 \text{ — биекция.}$$

Теорема

Пусть X — топологическое пространство, $f, g : X \to \mathbb{R}$. Тогда

$$d_B(D(f), D(g)) \leqslant \|f - g\|_\infty.$$

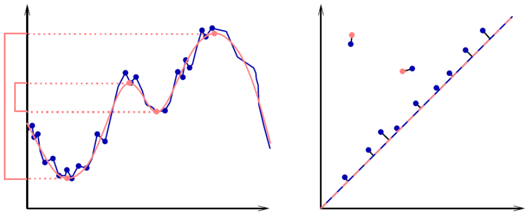
Таким образом, персистентная диаграмма $D(f)$ устойчива по отношению к возмущениям функции f.

[D. Cohen-Steiner, H. Edelsbrunner, J. Harer, 2007]

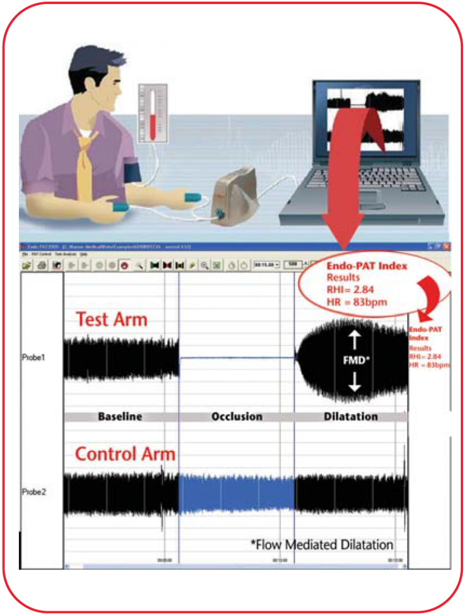
Устойчивость

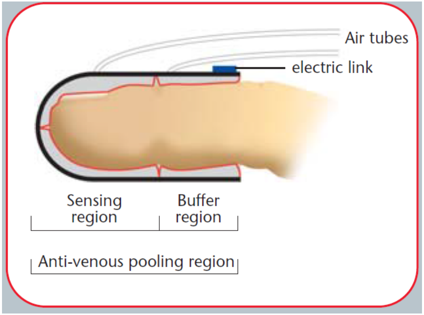
При возмущении функции f с диаграммой $D(f)$:

1. некоторая часть точек сдвигается на малое расстояние;
2. другая часть точек диаграммы лежит недалеко от диагонали и уходит на нее;
3. наконец, ещё одна совокупность точек приходит с диагонали.

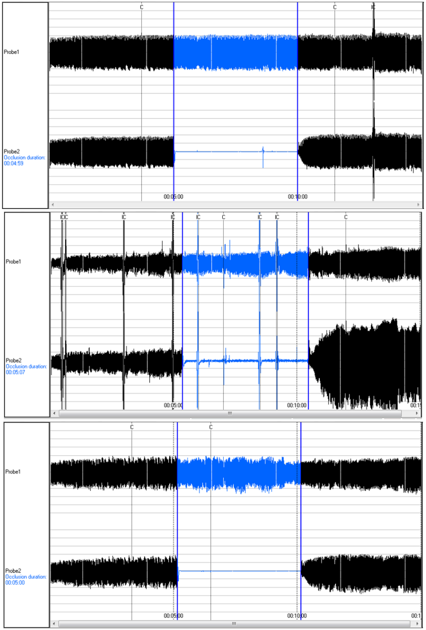


Endo-PAT2000

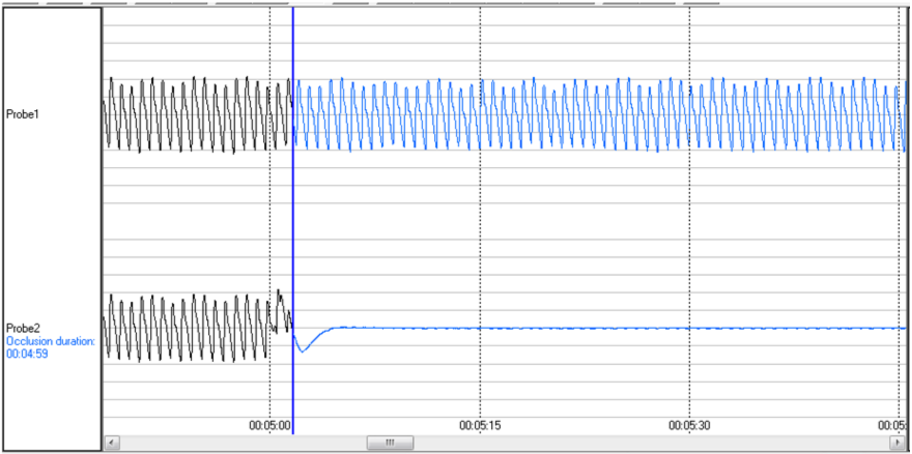




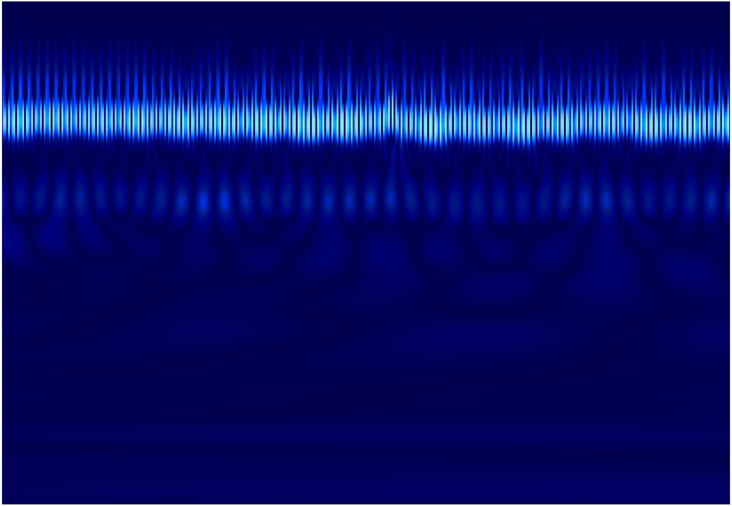
Данные Endo-PAT2000



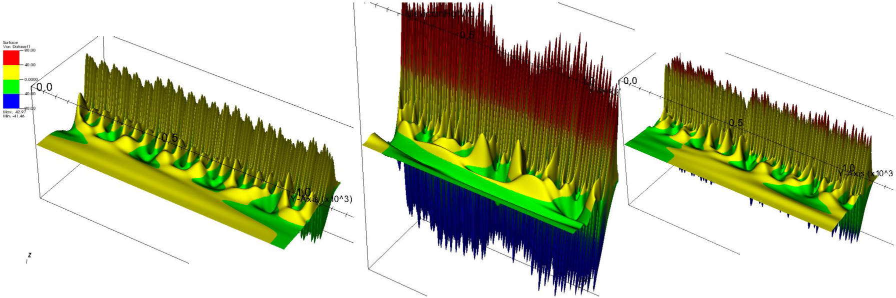
Данные Endo-PAT2000



Вейвлет-скейлограммы

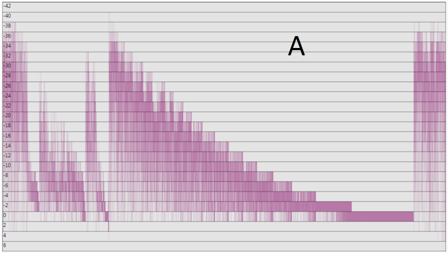


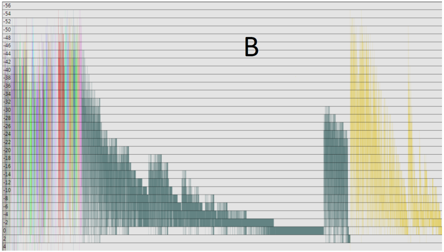
Вейвлет-скейлограммы

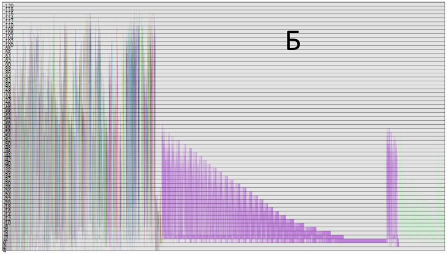


А Б В

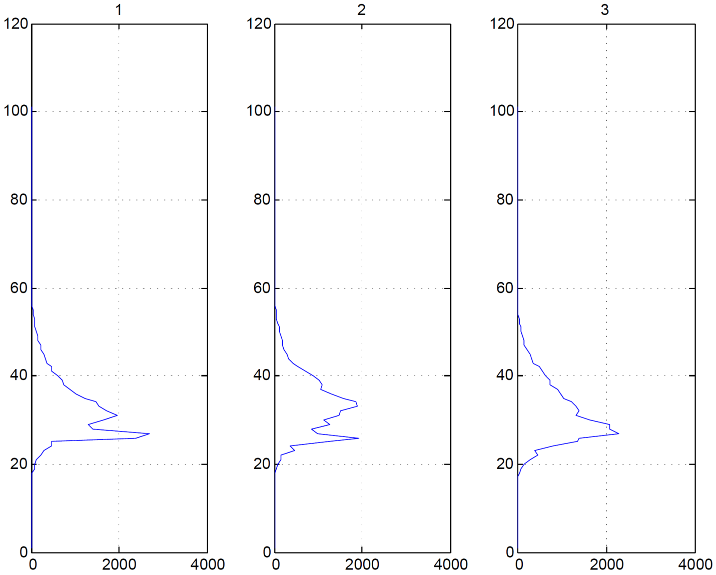
Баркоды



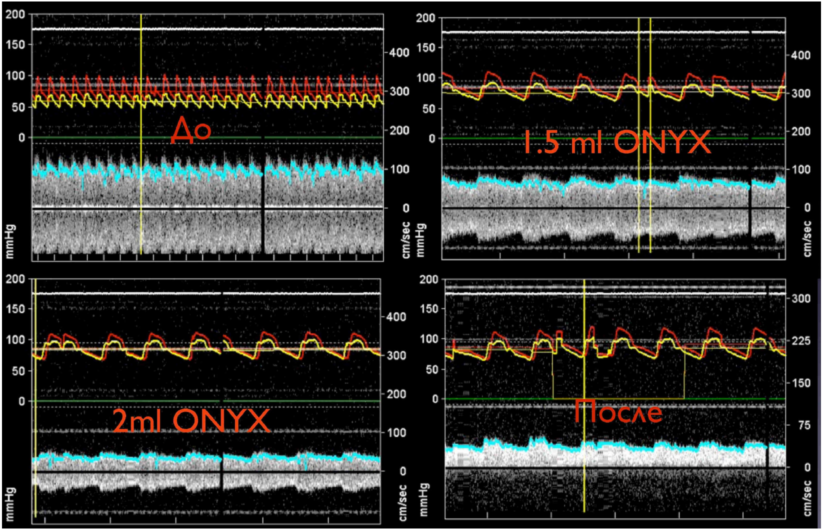




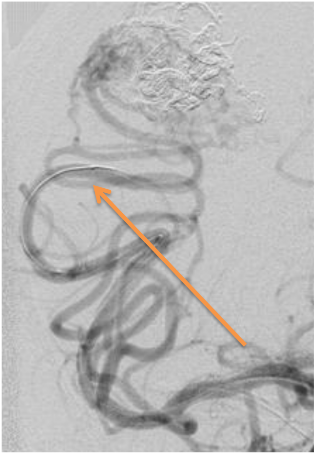
Баркоды. Распределение

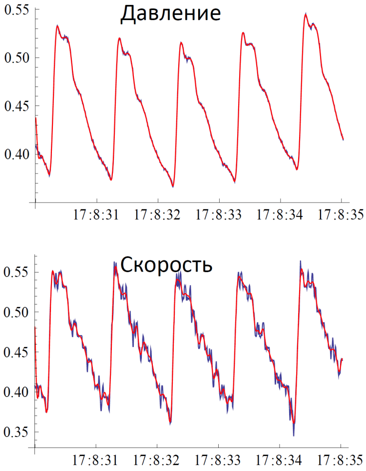


Эндоваскулярные измерения

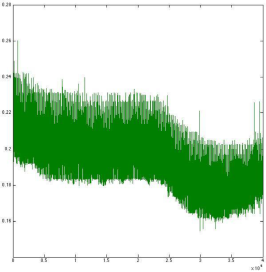


Данные измерений

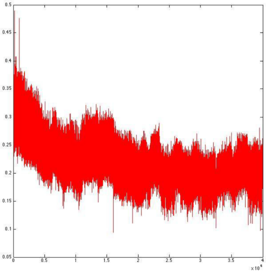




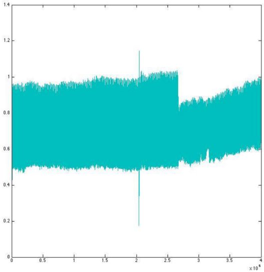
Анализ данных



Давление

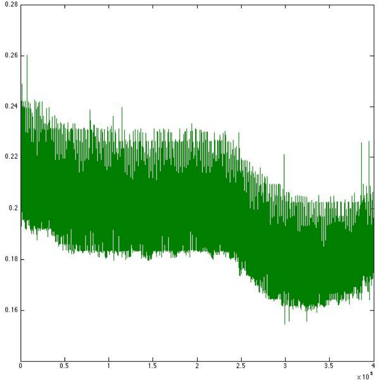


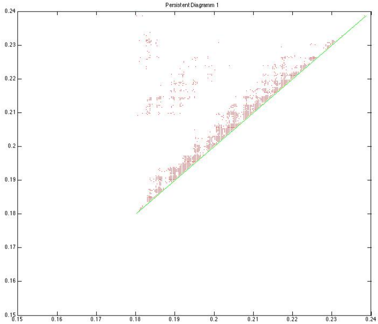
Скорость



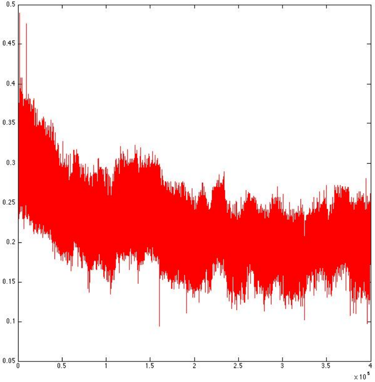
Системное давление

Персистентная диаграмма. Давление





Персистентная диаграмма. Скорость



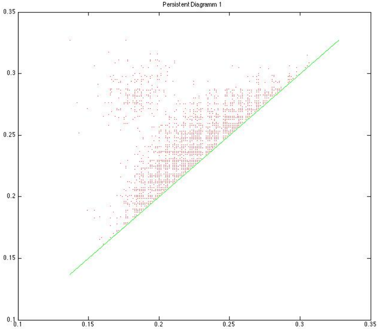
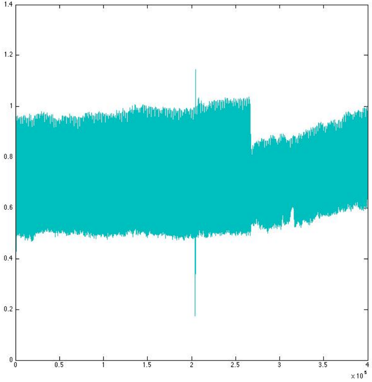
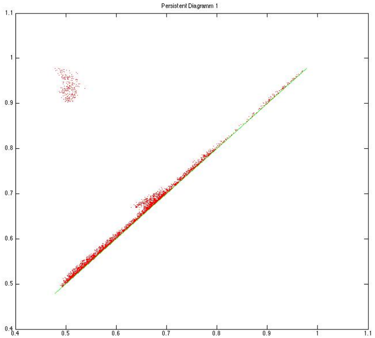


Диаграмма. Системное давление





Спасибо за внимание!