

Институт возрастной физиологии РАО  
Лаборатория нейрофизиологии когнитивной деятельности

ПОНЯТИЕ  
ВНУТРЕННЕЙ РЕПРЕЗЕНТАЦИИ  
В  
КОГНИТИВНОЙ НЕЙРОНАУКЕ

Андрей Васильевич Курганский

Москва, 2017

# ВНУТРЕННЯЯ РЕПРЕЗЕНТАЦИЯ

Фотоаппарат

полученный снимок является "репрезентацией" пространственного распределения светового потока, попадающего в глаз наблюдателя

Что сохраняет снимок? На черно-белом снимке создается такой двумерный рельеф плотности зерен серебра. Запоминается не световой поток и его параметры.

Сохраняется нечто (рельеф, избирательно рассеивающий падающий свет или двумерный рельеф источников света), что позволяет воссоздать в глазу наблюдателя запечатленное пространственное распределение светового потока.

ЕСЛИ МЫ ГОВОРИМ О РЕПРЕЗЕНТАЦИИ, ВАЖНО ЗНАТЬ:

ЧТО сохраняется – информация, позволяющая восстановить световой поток

КАК сохраняется – в виде рельефа плотности вещества, в виде битов;

в виде кодов; полностью или с потерей информации

Какие ФИЗИЧЕСКИЕ ПРОЦЕССЫ используются для сохранения

(записи и последующего хранения) и воспроизведения (извлечения).

В примере с фотоаппаратом: фотодиоды, разложение бромистого серебра, токи в микросхемах, светодиоды)

# ВНУТРЕННЯЯ РЕПРЕЗЕНТАЦИЯ

В рамках когнитивной психологии исследуется содержание внутренних репрезентаций (ЧТО ХРАНИТСЯ) и формат или способ хранения их Сохраняется (КАК ХРАНИТСЯ).

Эти исследования опираются на психофизические эксперименты.

---

В рамках нейрокогнитивной науки исследуются, какие ФИЗИЧЕСКИЕ (ФИЗИОЛОГИЧЕСКИЕ) ПРОЦЕССЫ в мозге ответственны за формирование внутренней репрезентации.

Эти исследования опираются на физиологические эксперименты и на Вычислительные модели

# Внутренняя репрезентация колебательного движения в рамках теории динамических систем

Математически динамически устойчивое колебательное движение описывается нелинейным дифференциальным уравнением, решением которого является предельный цикл (J.A.S. Kelso and collaborators 1985 - 2009).

$$\ddot{x} + [A(x^2 - r_0^2) + B(\dot{x}^2 - \omega_0^2 r_0^2)] \dot{x} + \omega^2 x = 0 \quad (*)$$

$$\ddot{x} + \omega^2 x \dot{x} = [A(r_0^2 - x^2) + B(\omega_0^2 r_0^2 - \dot{x}^2)] \dot{x}^2$$

$$\frac{d}{dt} \left( \frac{\dot{x}^2 + \omega^2 x^2}{2} \right) = [A(r_0^2 - x^2) + B(\omega_0^2 r_0^2 - \dot{x}^2)] \dot{x}^2$$

скорость изменения полной  
механической энергии  
осциллятора

$$[A(r_0^2 - x^2) + B(\omega_0^2 r_0^2 - \dot{x}^2)] > 0$$

Энергия возрастает

$$[A(r_0^2 - x^2) + B(\omega_0^2 r_0^2 - \dot{x}^2)] < 0$$

Энергия уменьшается

Устойчивое решение соответствует балансу энергии:

$$\int_0^T [A(x^2 - r_0^2) + B(\dot{x}^2 - \omega_0^2 r_0^2)] dt = 0$$

Динамическая система, описываемая уравнением (\*), способна породить колебания требуемой частоты и амплитуды. Сохраненные в памяти «инструкции» по организации такой динамической системы наряду с параметрами (требуемой амплитудой и частотой) составляют внутреннюю репрезентацию колебательного движения.

# Внутренняя репрезентация движения в рамках теории синергий

Мышечные синергии – содружественная работа многих мышц как единого целого – рассматриваются как функциональные единицы движения (Bizzi, Mussa-Ivaldi, 1998). В качестве таких единиц рассматриваются статические и динамические синергии (d'Avella et al., 2006; d'Avella, Lacquaniti, 2013).

Пример динамической синергии:

Сигнал, управляющий движением  $\vec{m}(t) = [m_1(t), m_2(t), \dots, m_M(t)]^T$  представляется как линейная комбинация зависящих от времени «профилей активации мышц»:

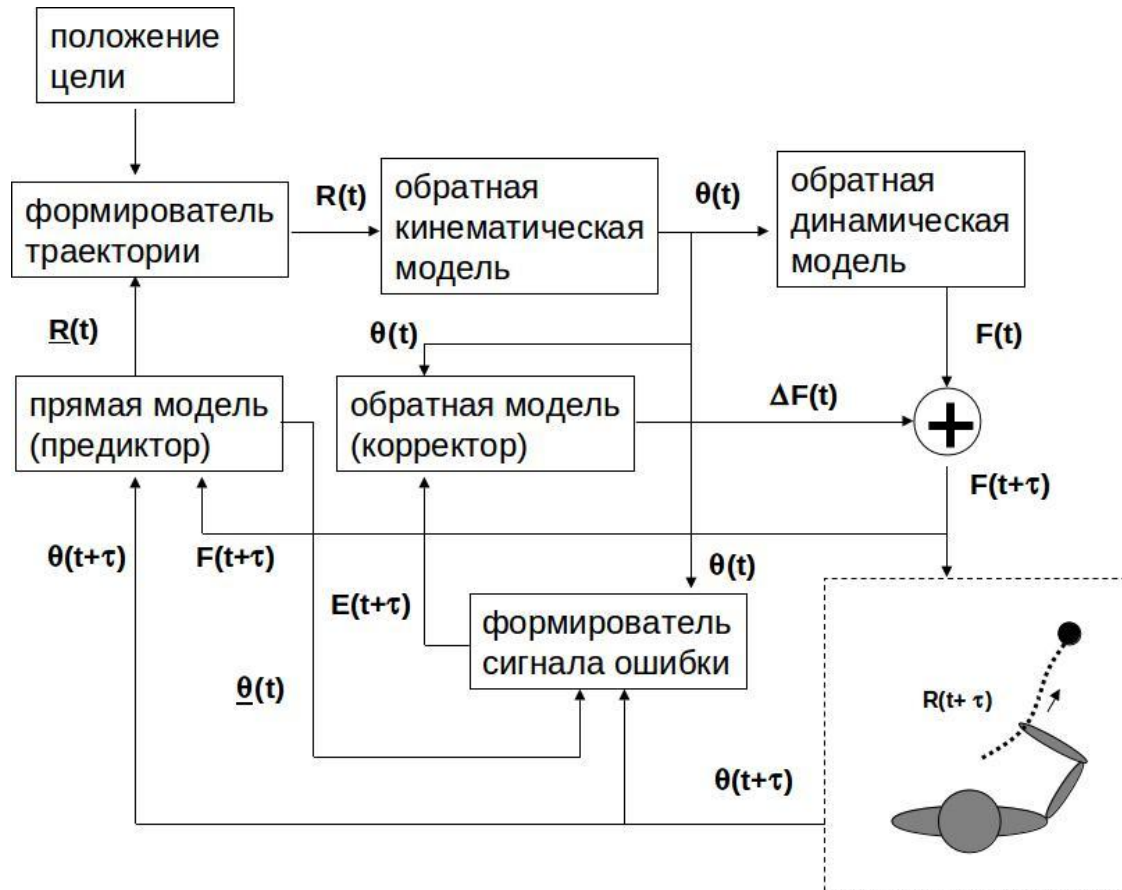
$$\vec{m}(t) = \sum_{n=1}^N c_n \vec{w}_n(t)$$

**Внутренняя репрезентация движения сводится к набору коэффициентов  $c_n$**

Показано, что у человека широкий репертуар движений к пространственной цели может быть выполнен за счет всего четырех – пяти динамических синергий, выбранных из ограниченной малоразмерной "библиотеки" (Muceli et al., 2010; d'Avella, Lacquaniti, 2013).

# Внутренние модели

Движение к пространственной цели в рамках теории внутренних моделей



Прямая модель: зная моторную команду, предсказывает положение

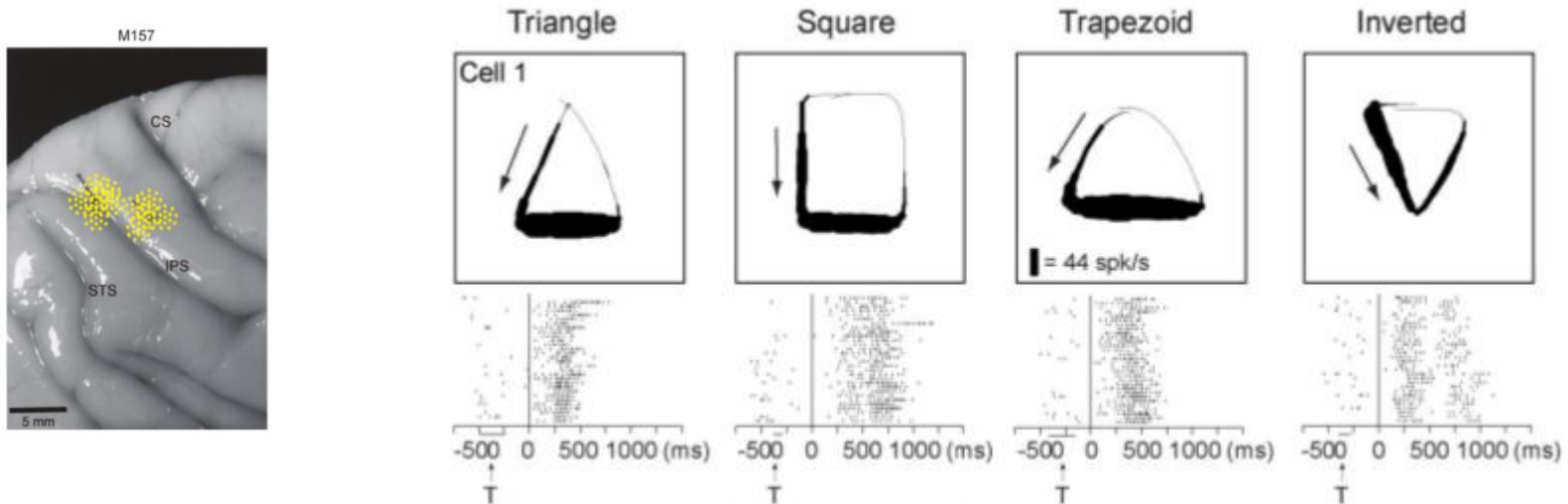
Обратная модель: зная требуемое положение, определяет моторную команду

ФИЗИОЛОГИЧЕСКИЕ МЕТОДЫ  
ИЗУЧЕНИЯ  
ВНУТРЕННЕЙ РЕПРЕЗЕНТАЦИИ

# Анализ нейронной активности

Bruno B. Averbeck, Matthew V. Chafee, David A. Crowe and Apostolos P. Georgopoulos  
Parietal Representation of Hand Velocity in a Copy Task  
J Neurophysiol 93:508-518, 2005. First published 21 July 2004;  
doi: 10.1152/jn.00357.2004

Две макаки рисовали геометрические фигуры на экране, управляя курсором с помощью джойстика. В это время записывалась нейронная активность



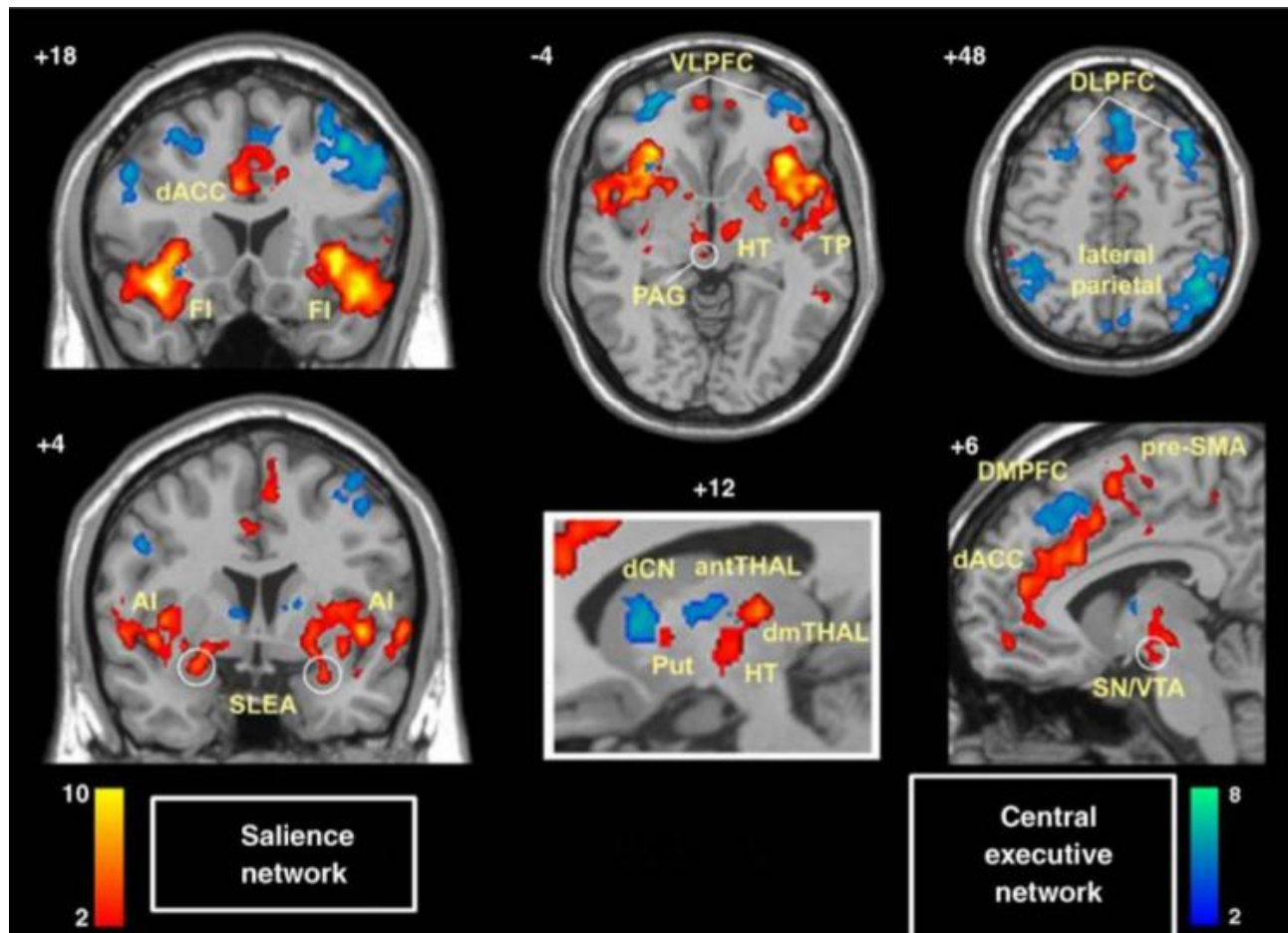
Избирательная реакция клетки на разных этапах рисования геометрических фигур

**Достоинство:** анализ нейронной активности дает наиболее детальное представление о процессах в нейронах.

**Недостатки:** ограничен малыми выборками нейронов.



# фМРТ: Анализ метаболических процессов



(Bressler & Menon, 2010)

**Достоинство:** высокое пространственное разрешение.

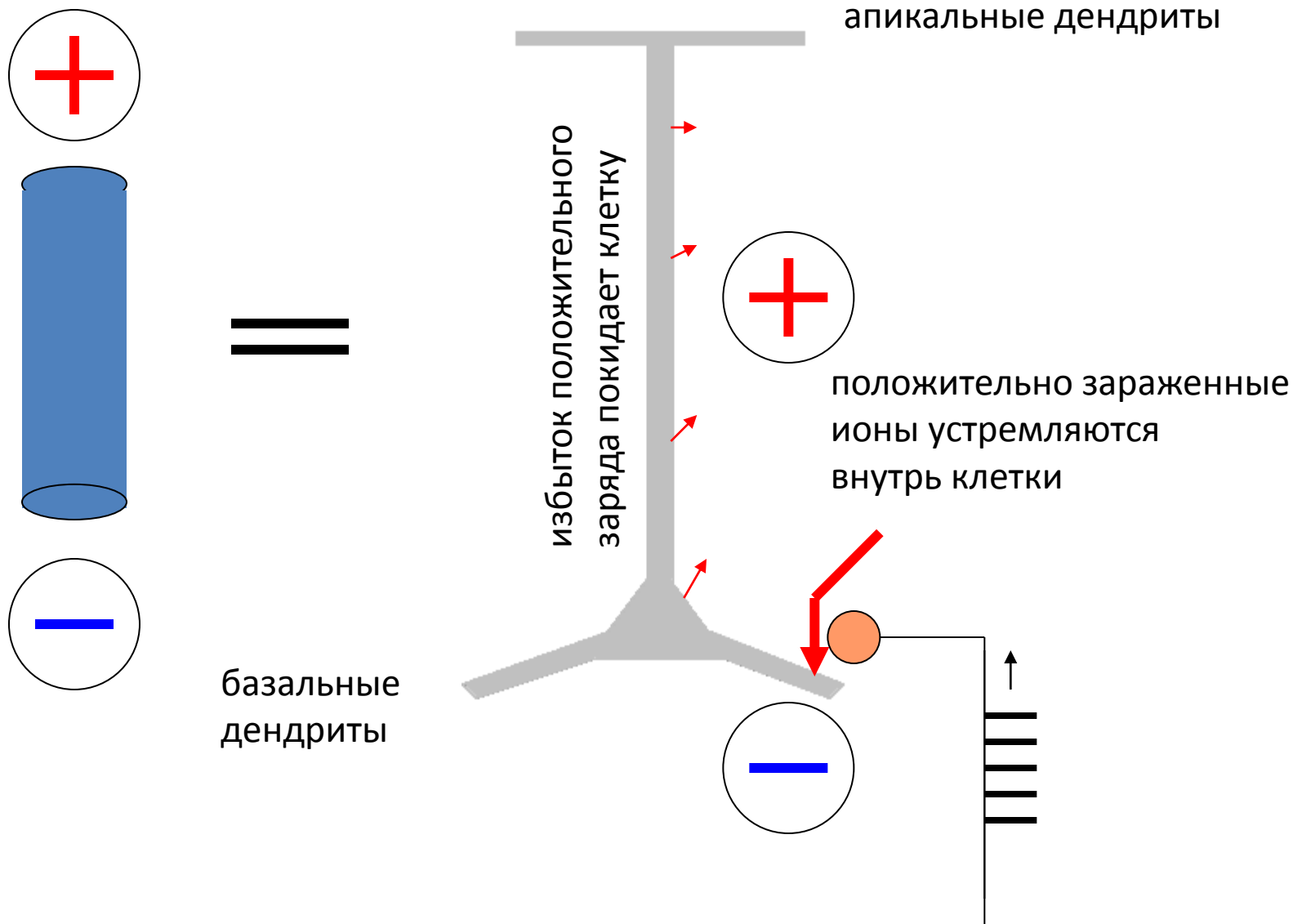
**Недостатки:**

-низкое временное разрешение;

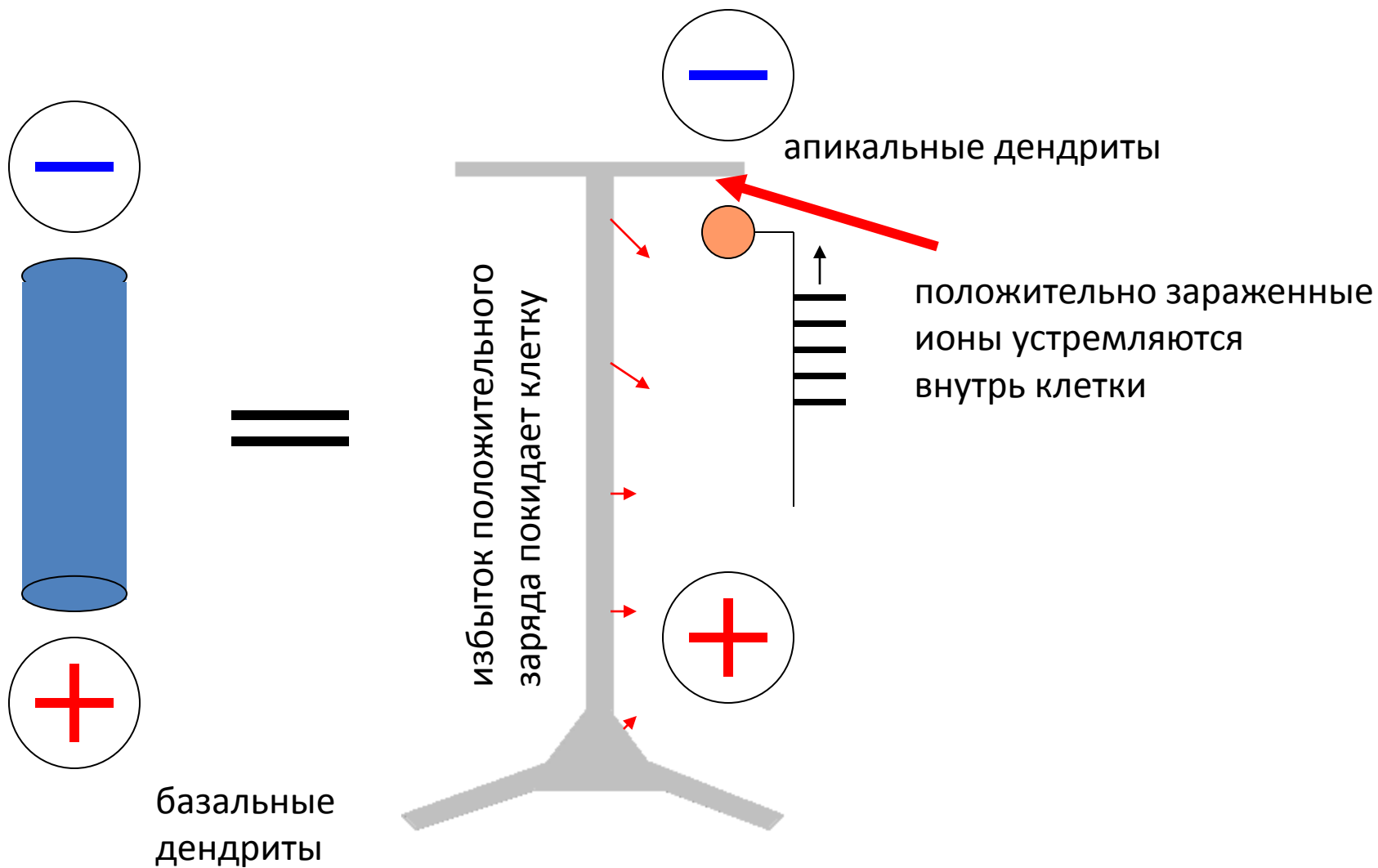
-косвенный характер получаемой информации

# АНАЛИЗ ЭЛЕКТРОЭНЦЕФАЛОГРАММЫ

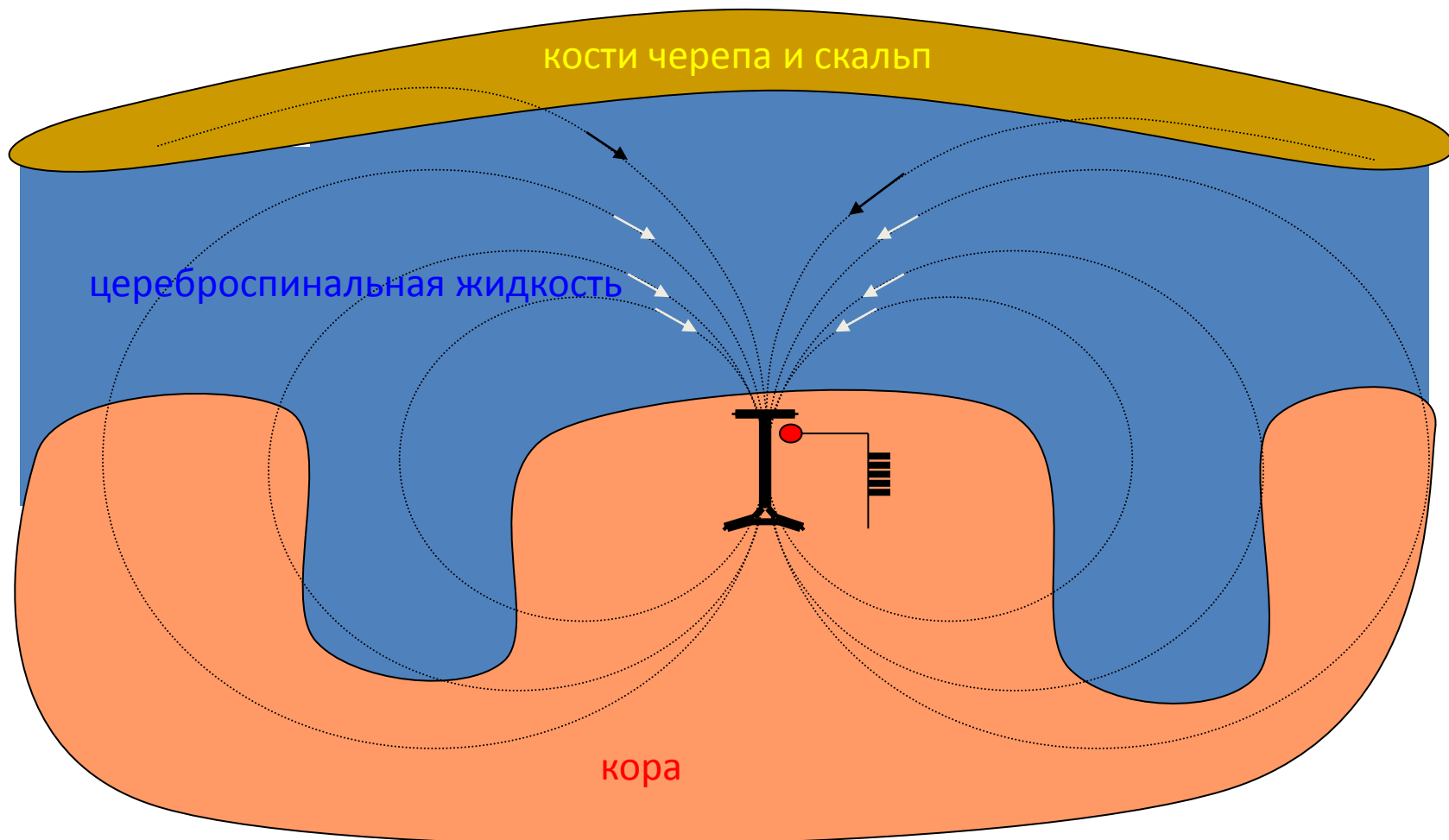
# ПИРАМИДНЫЙ НЕЙРОН КАК ИСТОЧНИК ТОКА



# ПИРАМИДНЫЙ НЕЙРОН КАК ИСТОЧНИК ТОКА



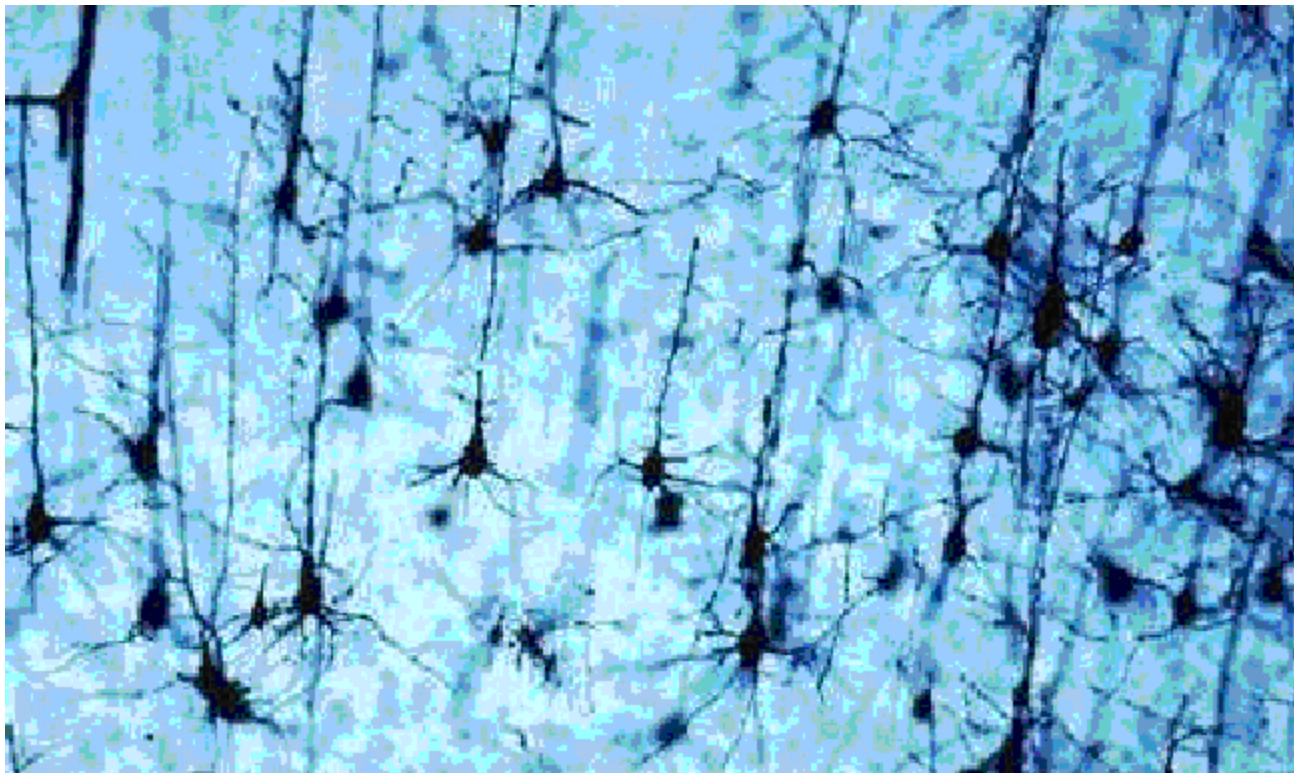
# ВНЕКЛЕТОЧНЫЙ ТОК ОДНОГО НЕЙРОНА



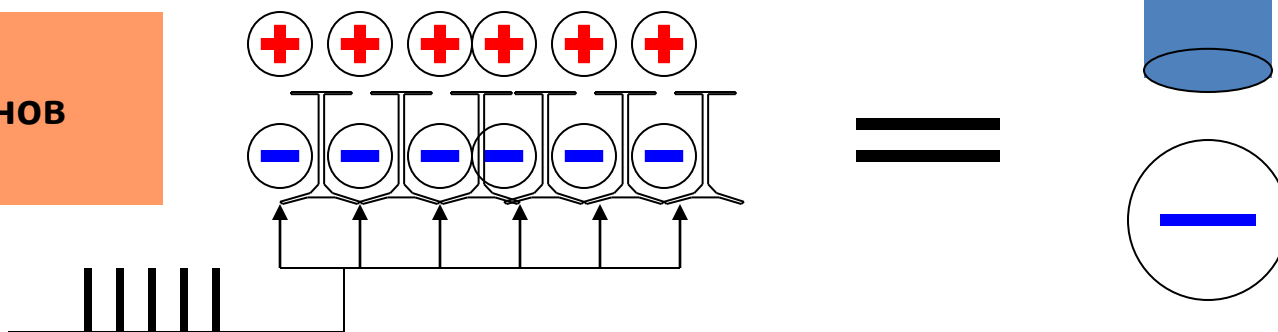
электрический ток, порожденный нейроном, распространяется  
во всей проводящей среде головы

# АПИКАЛЬНЫЕ ДЕНДРИТЫ ПИРАМИДНЫХ НЕЙРОНОВ ПАРАЛЛЕЛЬНЫ

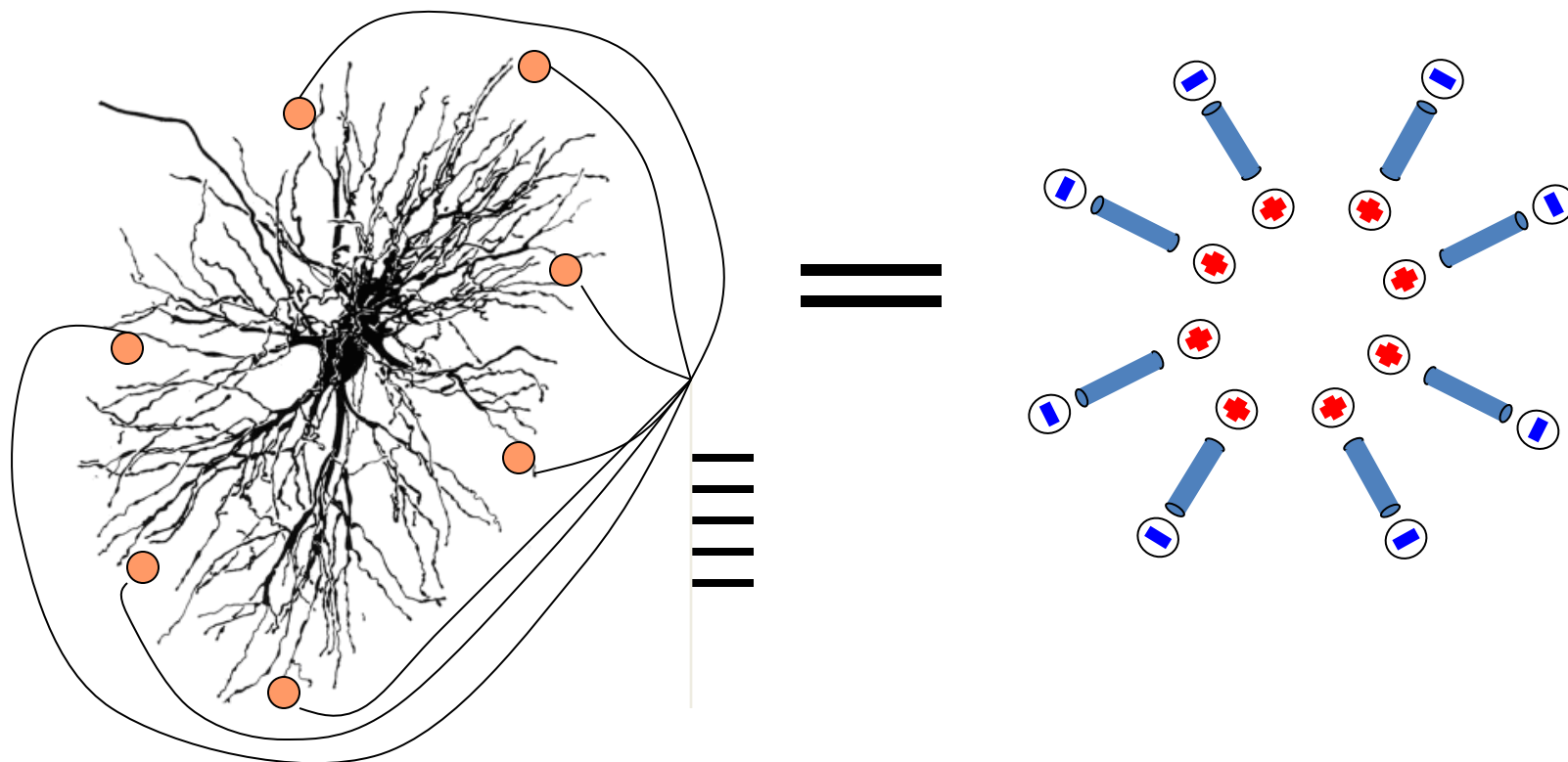
окрашены только пирамидные нейроны



электрические токи пирамидных нейронов суммируются



# НЕЙРОНЫ С КРУГОВОЙ ИЛИ СФЕРИЧЕСКОЙ СИММЕТРИЕЙ ДЕНДРИТНЫХ ДЕРЕВЬЕВ

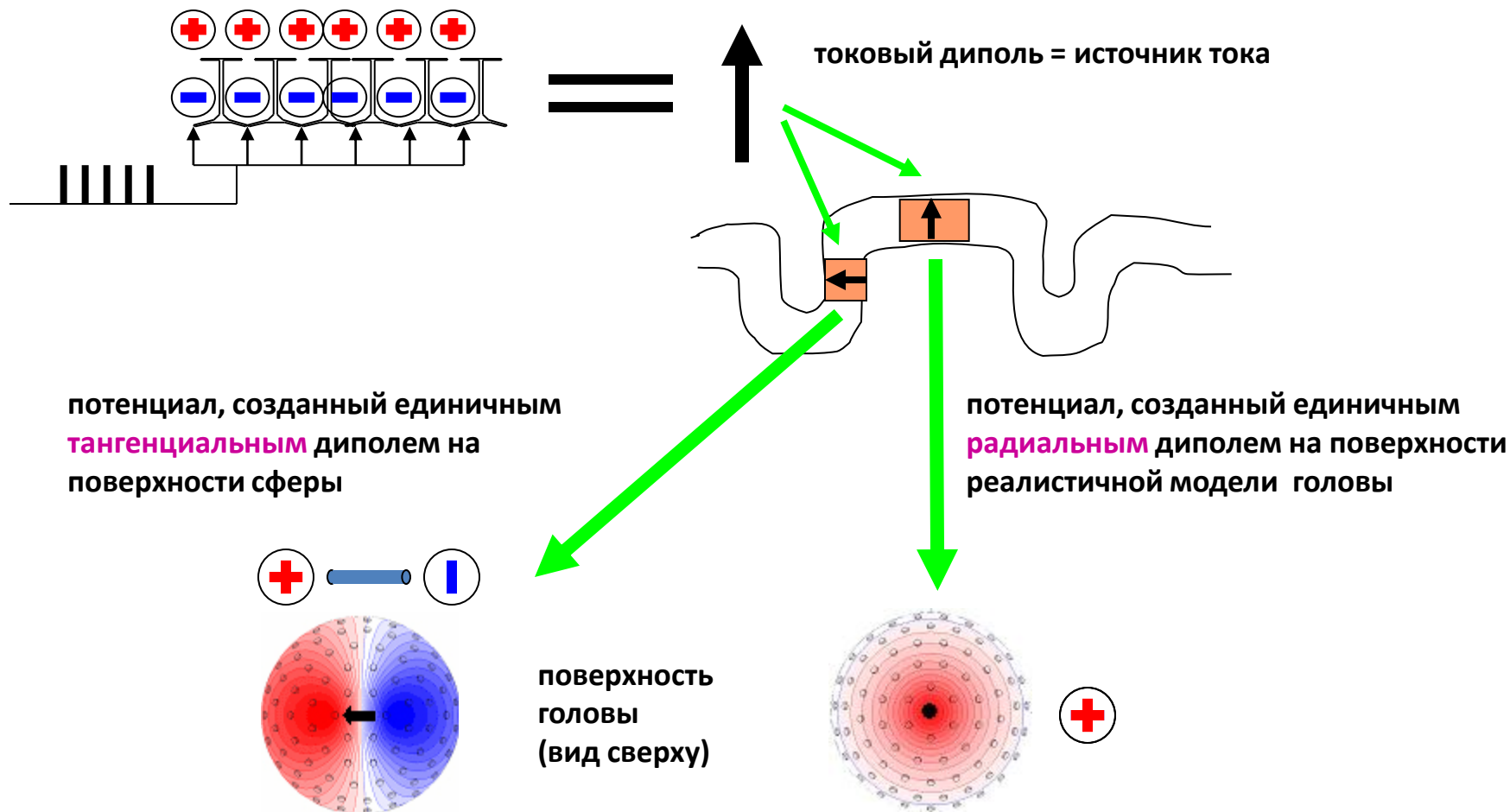


суммарный электрический ток за пределами такой клетки будет очень мал, так как токи взаимно уничтожатся.



**пирамидные нейроны вносят наибольший вклад во внеклеточный электрический ток**

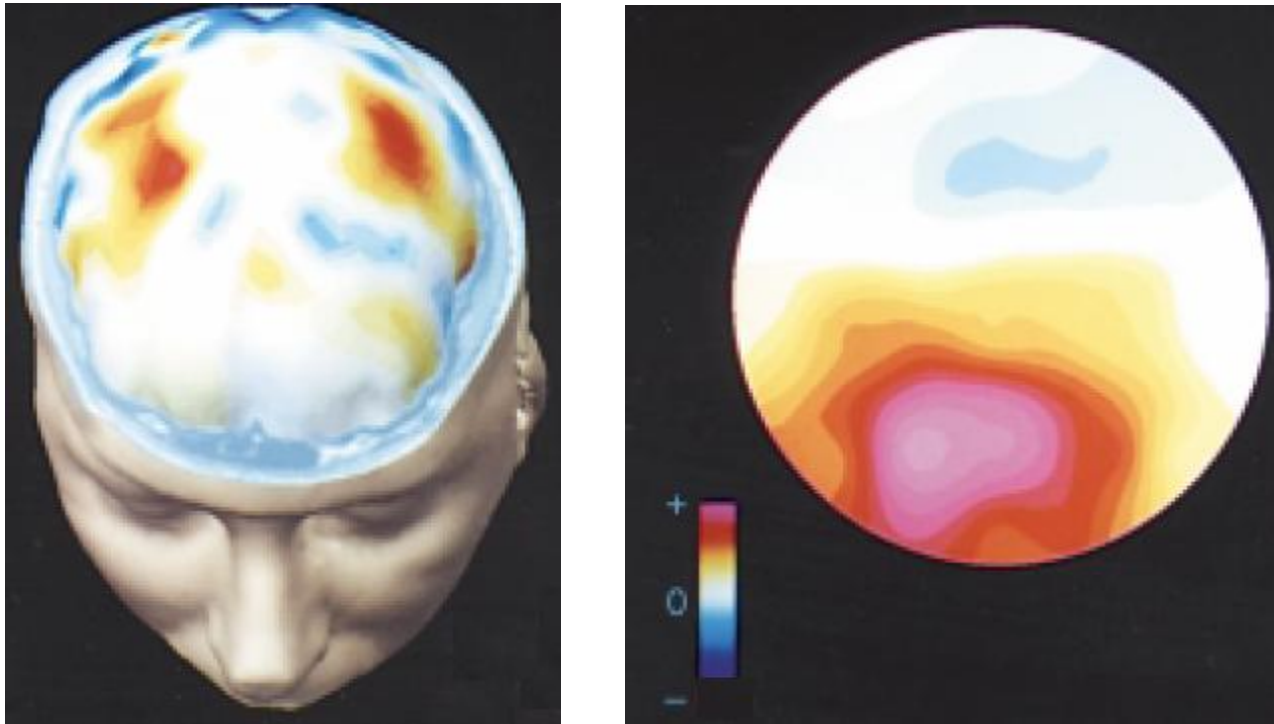
# ЛОКАЛЬНЫЙ ИСТОЧНИК ТОКА ОКАЗЫВАЕТ ГЛОБАЛЬНОЕ ВЛИЯНИЕ



электрический ток, порожденный каждым нейроном и каждой группой синхронно работающих нейронов, распространяется во всем объеме головы и создает неодинаковый потенциал на скальпе



# РАСПРЕДЕЛЕНИЕ СУММАРНОГО ПОТЕНЦИАЛА МНОГИХ ИСТОЧНИКОВ ТОКА

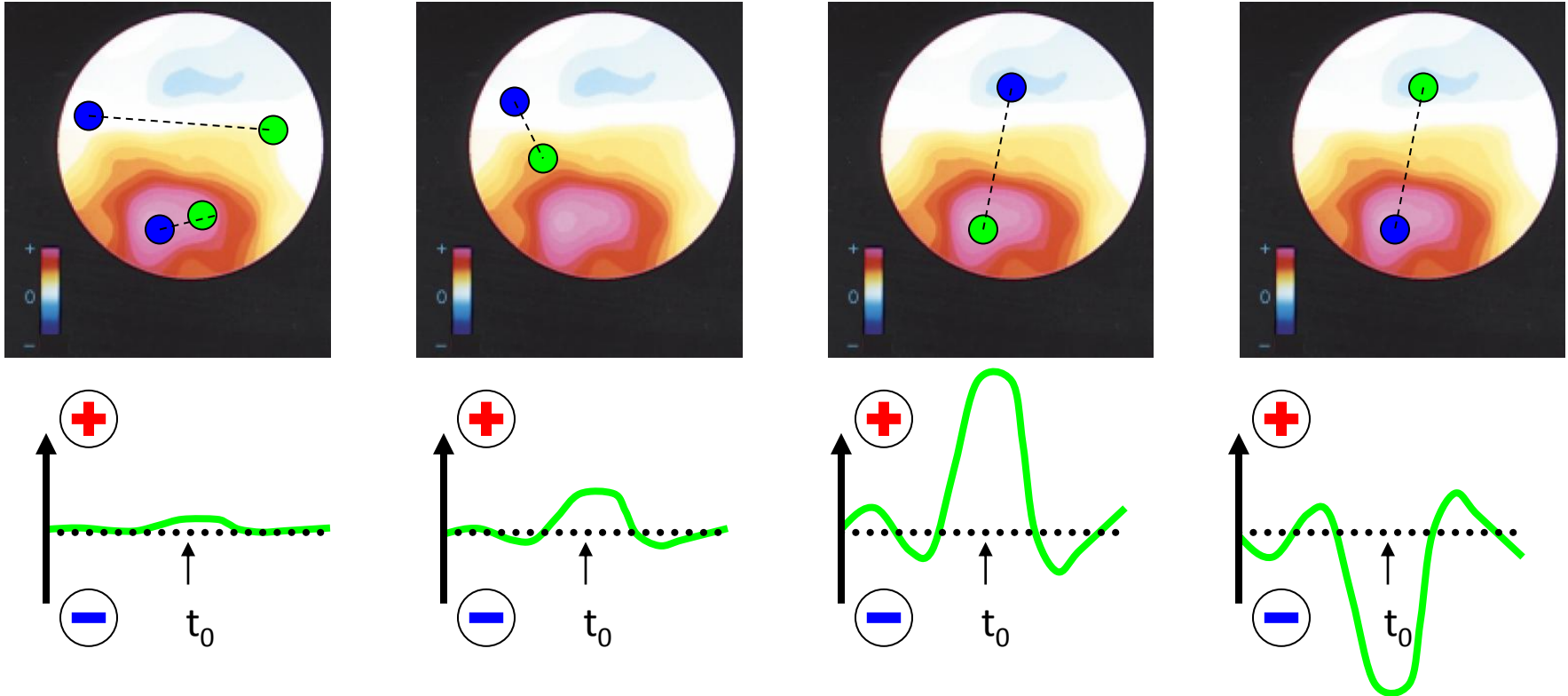


Хотя потенциал в каждой точке скальпа есть сумма вкладов источников электрического тока, распределенных по всему мозгу, наибольший вклад вносят радиально ориентированные источники, расположенные в коре непосредственно под этой точкой.

# РЕГИСТРАЦИЯ ЭЭГ

- референтный электрод
- активный электрод

распределение потенциала по скальпу в момент времени  $t_0$



регистрируемый сигнал есть разность потенциалов активного и референтного электродов

**Достоинства:**

- высокое временное разрешение;
- непосредственно отражает биоэлектрические процессы.

**Недостаток:** низкое пространственное разрешение.

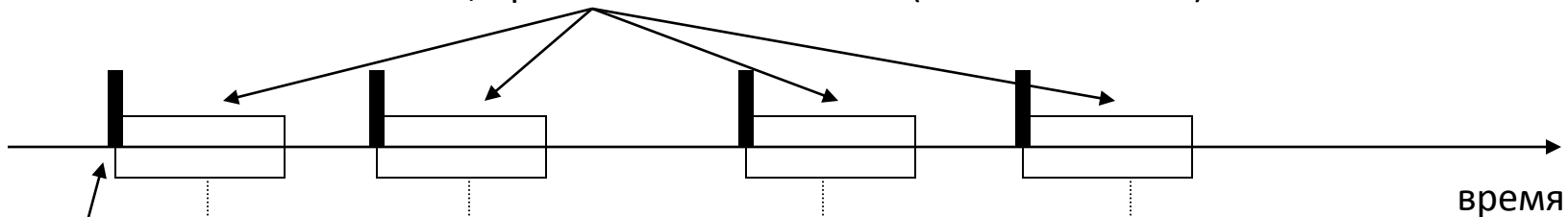
# ИСПОЛЬЗОВАНИЕ ЭЭГ ДЛЯ ИССЛЕДОВАНИЯ МОЗГОВОГО ОБЕСПЕЧЕНИЯ КОГНИТИВНЫХ ПРОЦЕССОВ

Обычно исследуют ЭЭГ-корреляты различных когнитивных операций (кодирование, удержание и извлечение информации из памяти) и пытаются судить о характере мозговой системы, обеспечивающей эти операции.

# СВОЙСТВА ЭЭГ, ПРИВЯЗАННОЙ К СОБЫТИЯМ

Предположение: ЭЭГ = сигнал, вызванный событием + спонтанная активность

Особые свойства ЭЭГ, привязанной к событиям (time-locked EEG)



Синхронизирующее событие (внешнее, напри<sup>м</sup>. сенсорный стимул; или эндогенное, напри<sup>м</sup>. движение)

"+"

Синхронное суммирование – обычный способ оценки ССП

ССП позволяет оценить сигнал, вызванный событием



ССП

# СВОЙСТВА СПОНТАННОЙ ЭЭГ

## КРАТКОВРЕМЕННОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ SHORT-TIME FOURIER TRANSFORM

В отличие от преобразования Фурье

$$F\{x\} = FT\{x\} = X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-i\omega t} dt \quad \text{для непрерывного (не дискретного) времени}$$

кратковременное преобразование Фурье отличается тем, что преобразуется не сам сигнал  $x(t)$ , а сигнал, умноженный на функцию окна (оконную функцию)  $w(\tau)$ :

$$STFT\{x\} = X(\omega, \tau) = \int_{-\infty}^{+\infty} x(t)w(t-\tau)e^{-i\omega t} dt$$

Для того чтобы оконная функция  $w(\tau)$  не искажала мощность сигнала, требуется соблюдение условия нормировки:

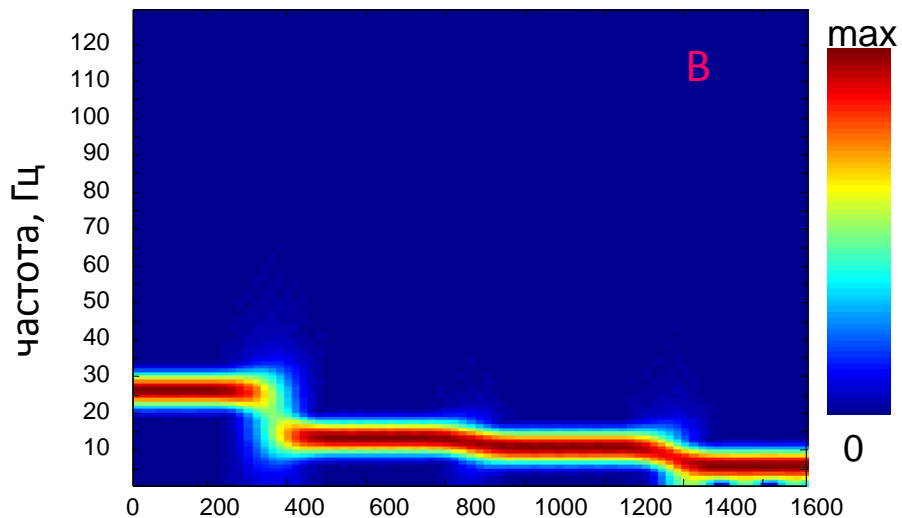
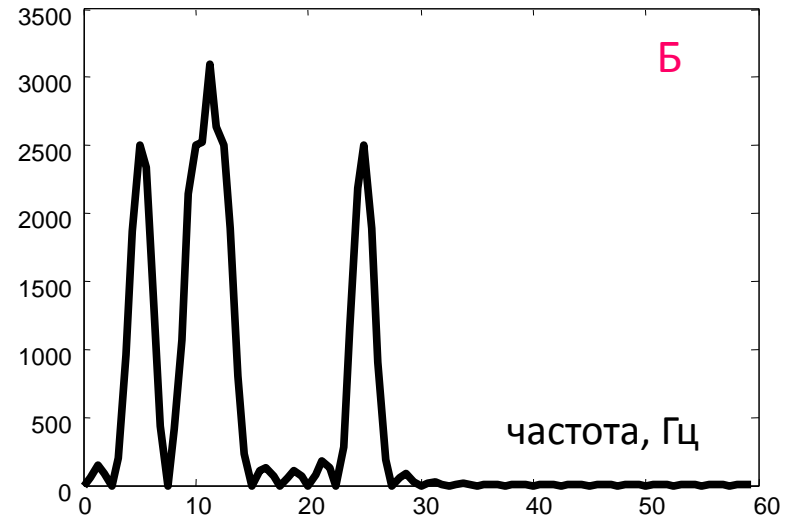
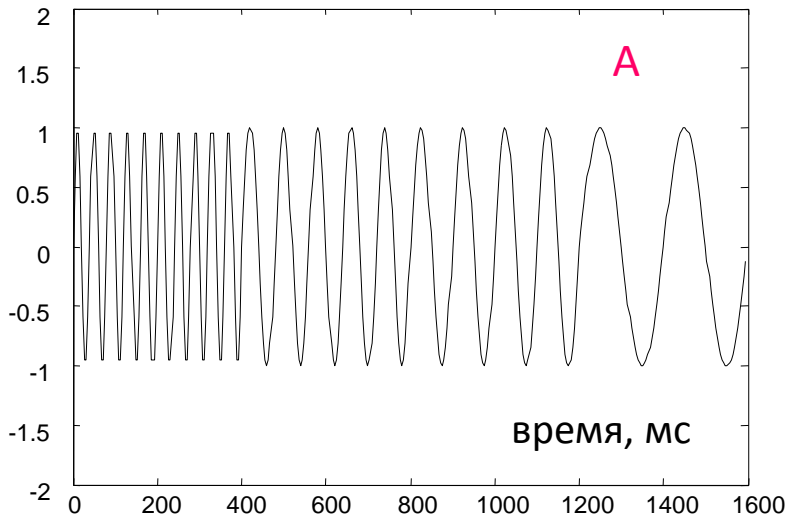
$$\int_{-\infty}^{+\infty} w(\tau) d\tau = \int_{-\infty}^{+\infty} w(t-\tau) d\tau = 1 \quad \forall t$$

Дискретный вариант STFT определяется по аналогии:

$$X(\omega, m) = \sum_{n=-\infty}^{n=+\infty} x(n)w(n-m)e^{-i\omega n}$$

# СПЕКТРОГРАММА

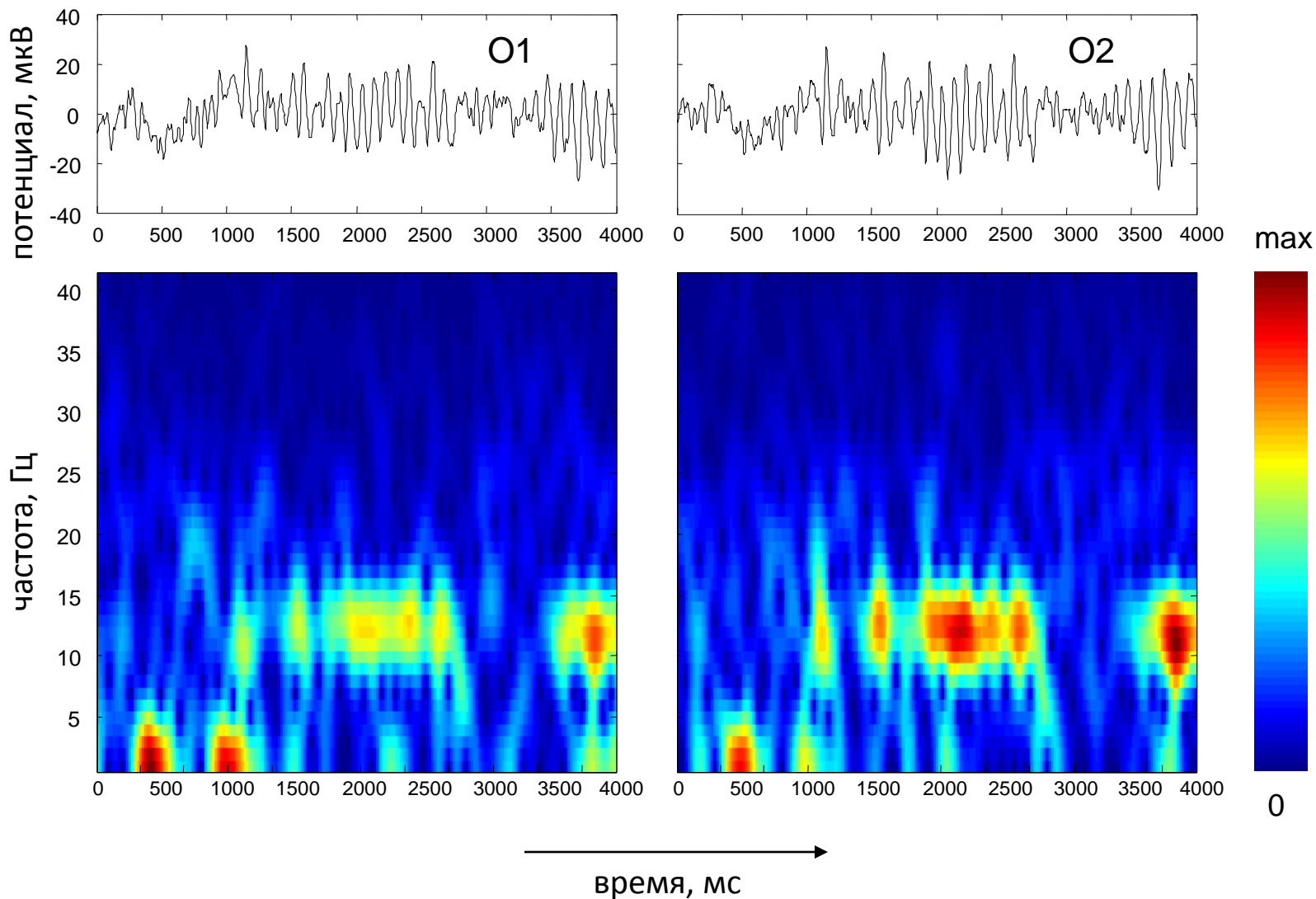
Функция  $|X(\omega, \tau)|^2$  называется спектрограммой (динамическим спектром)



Пример нестационарного процесса в виде 4-х "пристыкованных" друг к другу синусоид (A).  
Спектр мощности этого процесса (B).  
Спектрограмма этого процесса (B).

# ПРИМЕР СПЕКТРОГРАММЫ

Спектрограммы синхронно записанных отрезков ЭЭГ от затылочных электродов O1 и O2. И во временном представлении и в частотно-временном видна сходная динамика.



# ФУНКЦИОНАЛЬНЫЕ ВЗАИМОДЕЙСТВИЯ И КОГНИТИВНЫЕ ФУНКЦИИ

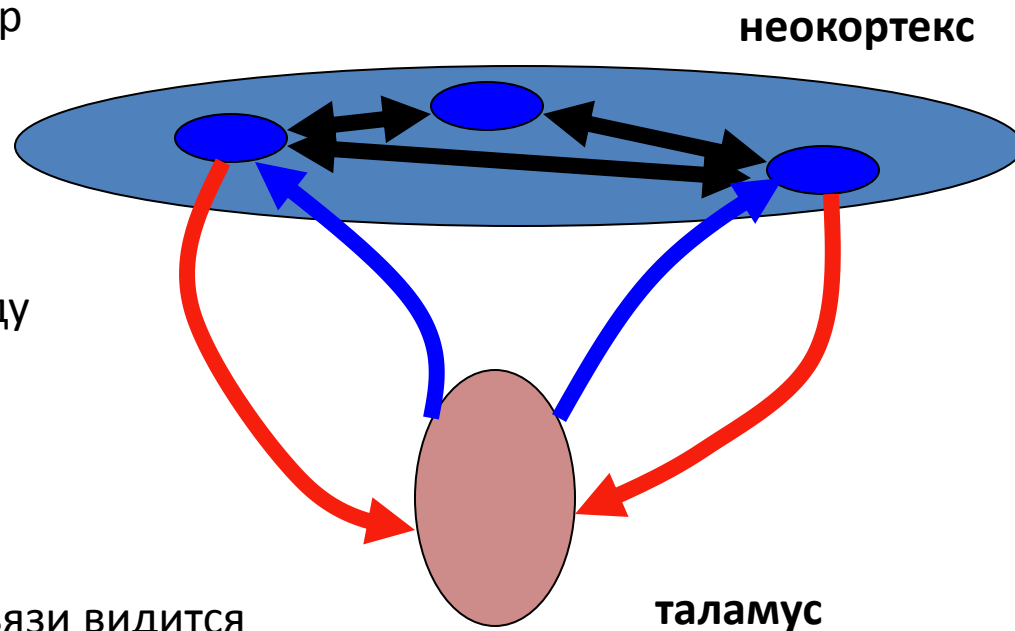
Когнитивные функции реализуются  
посредством ансамблей  
взаимодействующих мозговых структур



Взаимодействие приводит к  
появлению статистической связи между  
активностью участников системы



Исследование такой статистической связи видится  
как основной способ  
оценки «правил» взаимодействия и, следовательно,  
механизмов,  
реализующих когнитивные функции





# VAR-МОДЕЛЬ МНОГОКАНАЛЬНОЙ ЭЭГ

$$x(n) = \sum_{k=1}^{k=P} a(k)x(n-k) + e(n)$$

Линейное описание  $M$  взаимодействующих каналов с помощью AR-модели порядка  $P$

$$x(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \\ \vdots \\ x_M(n) \end{bmatrix} \quad e(n) = \begin{bmatrix} e_1(n) \\ e_2(n) \\ \vdots \\ e_M(n) \end{bmatrix}$$

Здесь сигнал  $x(n)$  и шум  $e(n)$  – это векторы, зависящие от дискретного времени  $n$ .

$$a(k) = \begin{bmatrix} a_{11}(k) & a_{12}(k) & \cdots & a_{1M}(k) \\ a_{21}(k) & a_{22}(k) & \cdots & a_{2M}(k) \\ \vdots & \vdots & \vdots & \vdots \\ a_{M1}(k) & a_{M2}(k) & \cdots & a_{MM}(k) \end{bmatrix}$$

Вот как выглядит матрица коэффициентов авторегрессии для  $k$ -й задержки

# VAR-МОДЕЛЬ В ЧАСТОТНОЙ ОБЛАСТИ

$$x(n) = \sum_{k=1}^{k=P} a(k)x(n-k) + e(n) \quad \xrightarrow{\substack{\text{преобразование} \\ \text{Фурье}}} \quad X(f) = \sum_{k=1}^{k=P} [a(k)e^{-i2\pi kf} X(f)] + E(f)$$

$$X(f) = \left\{ \sum_{k=1}^{k=P} [a(k)e^{-i2\pi kf}] \right\} X(f) + E(f) \quad \longrightarrow \quad X(f) = A(f)X(f) + E(f)$$

$$[I - A(f)]X(f) = E(f) \quad \tilde{A}(f) = [I - A(f)]$$

$$\tilde{A}(f)X(f) = E(f)$$

$$X(f) = [\tilde{A}(f)]^{-1}E(f) \quad H(f) = \tilde{A}^{-1}(f)$$

$$X(f) = H(f)E(f)$$

*Kaminski M., Ding M., Truccolo W.A., Bressler S.I. Evaluating causal relations in neural systems: Granger causality, direct transfer function and statistical assessment of significance. Biol. Cybern. 2001. V.85. P.145-157*

*Baccala L., Sameshima K. Partial directed coherence: a new concept in neural structure determination. Biol. Cybern. 2001. V.84. P.463-474*

# АВТОСПЕКТРЫ, ВЗАИМНЫЕ СПЕКТРЫ, ФУНКЦИЯ КОГЕРЕНТНОСТИ

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1M}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots & \sigma_{2M}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{M1}^2 & \sigma_{M2}^2 & \cdots & \sigma_{MM}^2 \end{bmatrix}$$

Ковариационная матрица остатков AR

$$\sigma_{ij}^2 = C_{ij} = \text{cov}(e_i, e_j)$$

эрмитово сопряжение = транспонирование и  
комплексное сопряжение

$$S(f) = H(f) \cdot \Sigma \cdot H^H(f)$$

Спектральная матрица

Теперь можно вычислить функцию когерентности

$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$$

$$K_{ij}(f) \equiv \Gamma_{ij}(f) = |C_{ij}(f)|^2 = \frac{|S_{ij}(f)|^2}{S_{ii}(f)S_{jj}(f)}$$

# НАПРАВЛЕННАЯ КОГЕРЕНТНОСТЬ

Если ковариационная матрица

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1M}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots & \sigma_{2M}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{M1}^2 & \sigma_{M2}^2 & \cdots & \sigma_{MM}^2 \end{bmatrix} \quad \text{диагональна} \quad \Sigma = \begin{bmatrix} \sigma_{11}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{22}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{MM}^2 \end{bmatrix}$$

то функцию обычной когерентности можно представить следующим образом:

$$C_{ij} = \frac{S_{ij}}{\sqrt{S_{ii}S_{jj}}} = \sum_{k=1}^{k=M} [\overline{H_{ik}} H_{jk} \frac{\sigma_{kk}^2}{\sqrt{S_{ii}}\sqrt{S_{jj}}}] = \sum_{k=1}^{k=M} [\frac{\sigma_{kk} \overline{H_{ik}}}{\sqrt{S_{ii}}} \frac{\sigma_{kk} H_{jk}}{\sqrt{S_{jj}}}] = \sum_{k=1}^{k=M} \overline{\gamma_{ik}} \gamma_{jk}$$

Величину

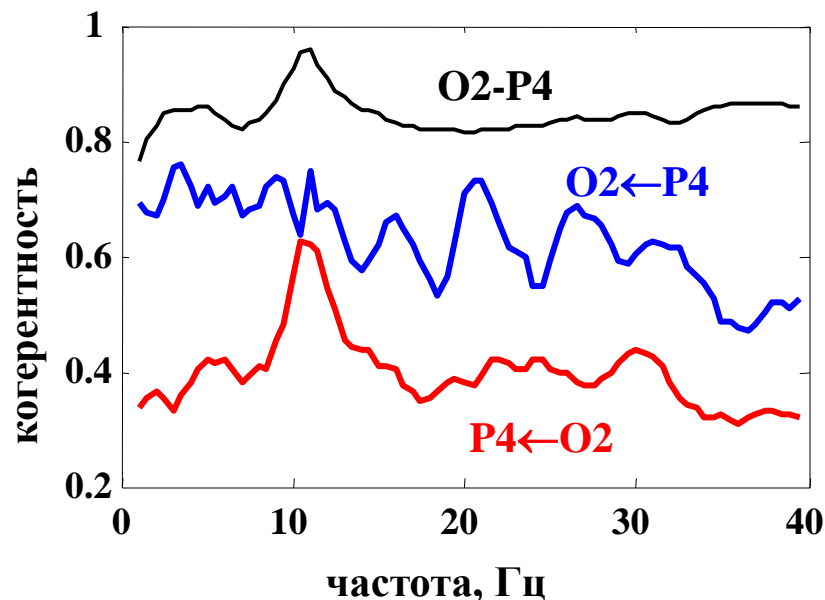
$$\gamma_{ij} = \gamma_{i \leftarrow j} = \frac{\sigma_{jj} H_{ij}}{\sqrt{S_{ii}}}$$

называют направленной когерентностью

$$\gamma_{i \leftarrow j} = \frac{\sigma_{jj} H_{ij}}{\sqrt{S_{ii}}} \neq \gamma_{j \leftarrow i} = \frac{\sigma_{ii} H_{ji}}{\sqrt{S_{jj}}}$$

Знание матрицы  $H$  позволяет вычислить направленную когерентность

*Baccala L., Sameshima K.*  
*Partial directed coherence: a new concept in neural structure determination.*  
*Biol. Cybern. 2001. V.84. P.463-474*



# НЕГАУССОВСКИЕ СИГНАЛЫ

$$x = [x(1), x(2), \dots, x(N)]$$

переменная величина, например  
последовательные отсчеты во времени

математическое ожидание

$$E\{x\} = \langle x \rangle = \bar{x}$$

дисперсия (variance)

$$D\{x\} = E\{(x - E\{x\})^2\}$$

или (в многомерном случае)  
ковариационная матрица

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \dots & \sigma_{1M}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & \dots & \sigma_{2M}^2 \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{M1}^2 & \sigma_{M2}^2 & \dots & \sigma_{MM}^2 \end{bmatrix}$$

статистика 1-го порядка

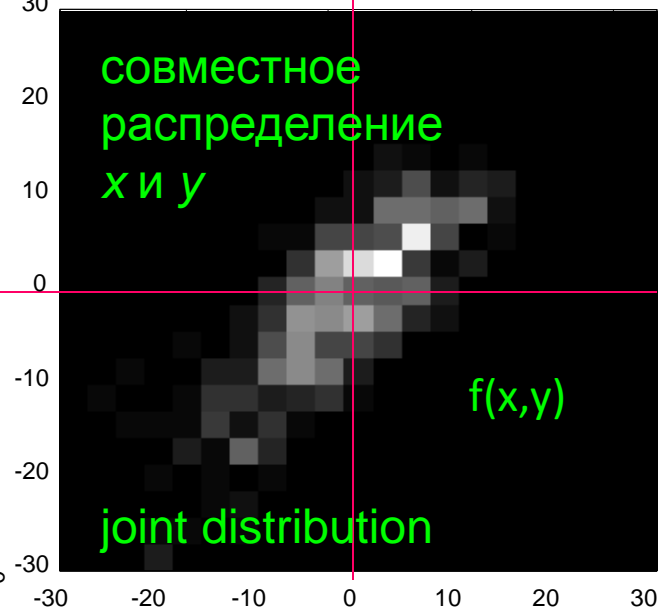
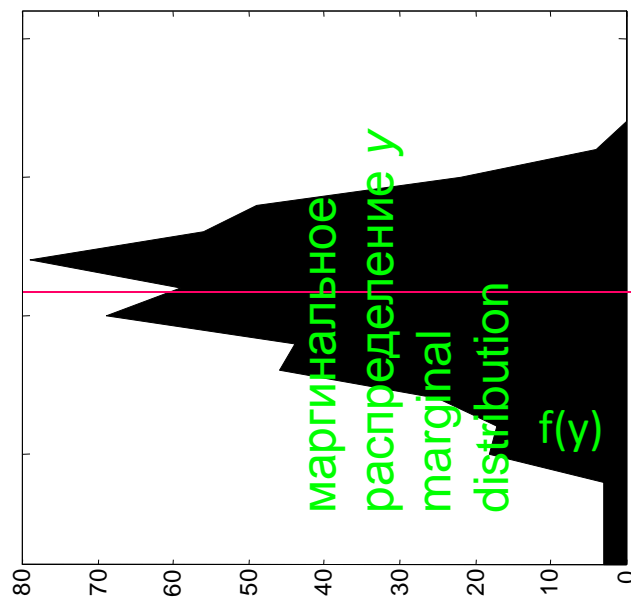
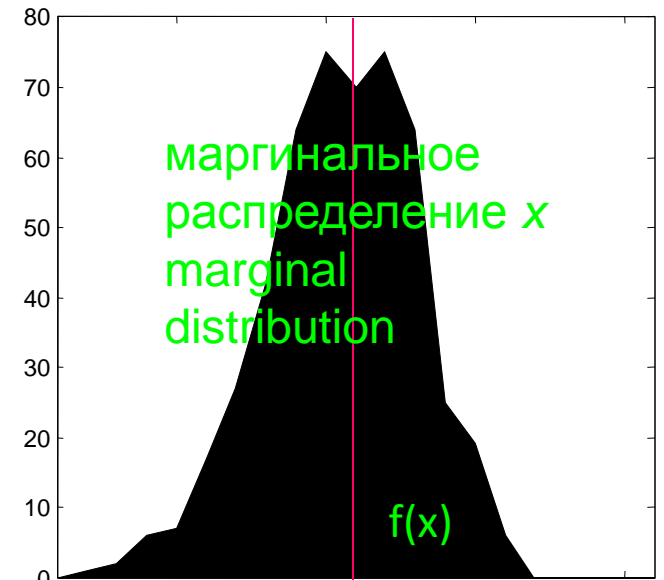
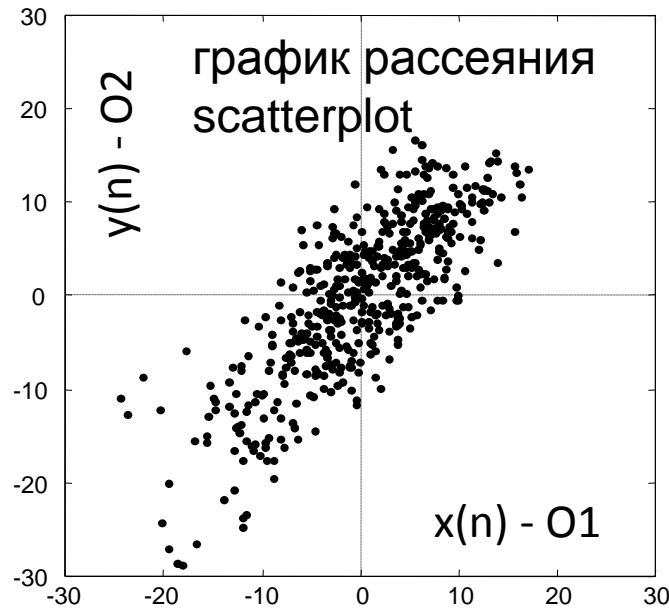
Достаточно для описания  
гауссовского распределения

статистика 2-го порядка

Как быть в случае неауссовского распределения?

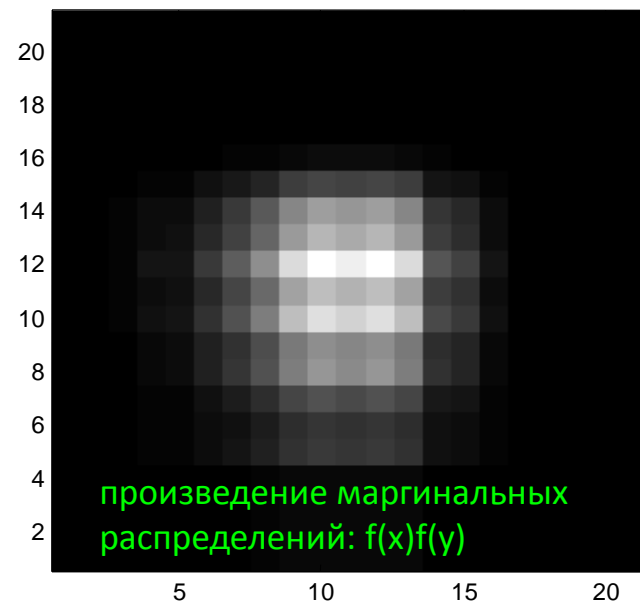
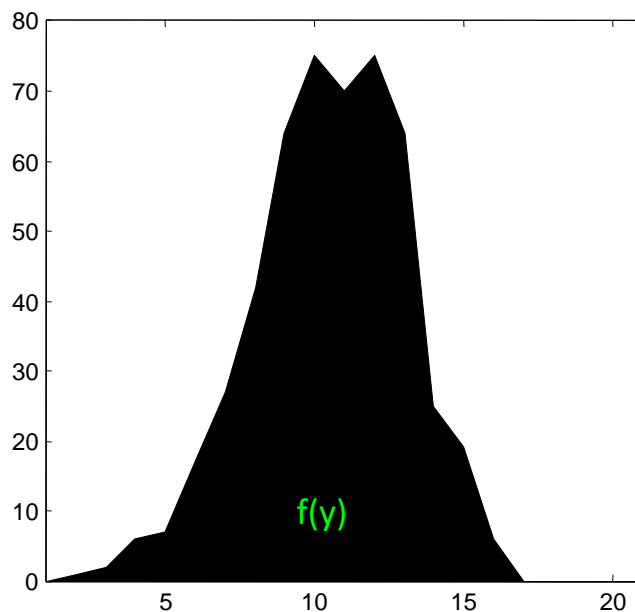
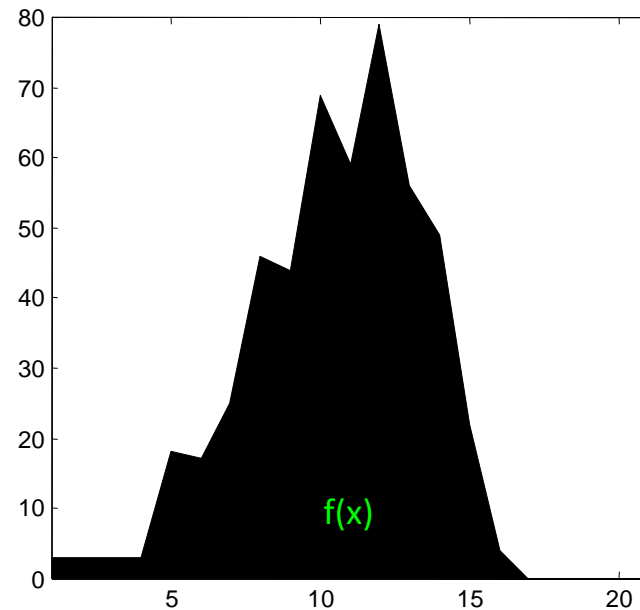
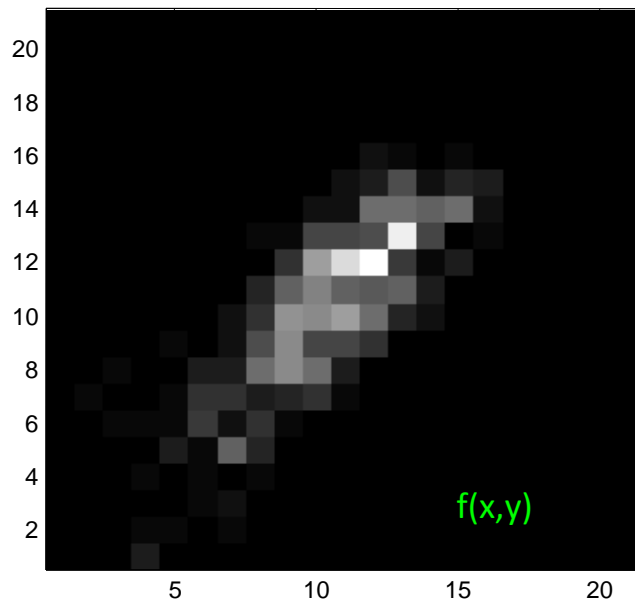
# РАСПРЕДЕЛЕНИЯ

Можно работать  
непосредственно с  
распределениями



# СТАТИСТИЧЕСКАЯ ЗАВИСИМОСТЬ

Основная идея:  
измерить степень  
статистической связи  
по расстоянию  
между совместным  
распределением и  
произведением  
маргинальных  
распределений



Видно, что  $x$  и  $y$   
зависимы

$$f(x, y) \neq f(x)f(y)$$

# СТАТИСТИЧЕСКАЯ ЗАВИСИМОСТЬ: ЭНТРОПИЯ

Насколько информативно событие? Настолько, насколько редко оно встречается, настолько насколько оно неожиданно. Очевидно, неожиданность обратно пропорциональна вероятности события. Вводится мера неожиданности как двоичный логарифм обратной вероятности

$$I = \log_2 \left( \frac{1}{p} \right) = -\log_2(p)$$

Энтропия определяется как математическое ожидание "неожиданности" (информации), связанной с событиями некоторого "источника".

$$H = \sum_{k=1}^K p_k I(p_k) = -\sum_{k=1}^K p_k \log_2(p_k) \quad \text{энтропия Шеннона}$$

Распределение можно характеризовать его энтропией



Можно посчитать энтропию распределения ЭЭГ сигнала

Рекомендуемая литература:

Nonlinear multivariate analysis of neurophysiological signals

Ernesto Pereda <sup>a,\*</sup>, Rodrigo Quian Quiroga <sup>b,1</sup>, Joydeep Bhattacharya <sup>c,2</sup>



# ВЗАИМНАЯ ИНФОРМАЦИЯ

Взаимная информация двух величин  $x$  и  $y$  определяется следующим образом:

$$MI(x, y) = H(x) + H(y) - H(x, y)$$

Взаимная информация -- симметричная величина, т.е. она не изменяется при перестановке аргументов

$$MI(x, y) = MI(y, x)$$

"Расстояние" между двумя распределениями KULLBACK–LEIBLER distance (divergence)  
сокращение: KL-distance

$$D(f(x) \| f(y)) = \sum f(x) \log_2 \left( \frac{f(x)}{f(y)} \right) \geq 0$$

Равно 0 тогда и только тогда, когда распределения идентичны

$$MI(x, y) = D(f(x, y) \| f(x)f(y)) = \sum f(x, y) \log_2 \left( \frac{f(x, y)}{f(x)f(y)} \right)$$

Взаимная информация двух величин  $x$  и  $y$  есть KL-distance между взаимным распределением этих величин и произведением их маргинальных распределений



Можно использовать взаимную информацию для измерения статистической связи между каналами ЭЭГ

**СПАСИБО ЗА ВНИМАНИЕ!**