

АКАДЕМИЯ НАУК СССР  
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ  
ИССЛЕДОВАНИЙ

В.И.КРЮКОВ, Г.Н.БОРИСЮК,  
Р.М.БОРИСЮК, А.Б.КИРИЛЛОВ, Е.И.КОВАЛЕНКО

# МЕТАСТАБИЛЬНЫЕ И НЕУСТОЙЧИВЫЕ СОСТОЯНИЯ В МОЗГЕ

ПУЩИНО · 1986

**АКАДЕМИЯ НАУК СССР  
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ  
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР**

**В.И.КРЮКОВ, Г.Н.БОРИСЮК,  
Р.М.БОРИСЮК, А.Б.КИРИЛЛОВ, Е.И.КОВАЛЕНКО**

**МЕТАСТАБИЛЬНЫЕ  
И НЕУСТОЙЧИВЫЕ  
СОСТОЯНИЯ В МОЗГЕ**

**ПУЩИНО · 1986**

УДК 519.21:57

**В.И.Крюков, Г.Н.Борисюк, Р.М.Борисюк, А.Б.Кириллов, Е.И.Коваленко.** Метастабильные и неустойчивые состояния в мозге. Пущино, ОНТИ НЦБИ АН СССР, 1986, 114 с.

Данная монография посвящена проблеме анализа сенсорной информации в мозге и является результатом много летнего коллективного труда сотрудников лаборатории статистических методов НИВЦ АН СССР.

Полученные результаты дают новую точку зрения на ряд классических проблем электрофизиологии и вместе с тем, по-видимому, закрывают несколько взаимосвязанных ее проблем, например проблему внутреннего торможения.

В работе представлены теоретические и экспериментальные доказательства существования в мозге долгоживущих микроочагов возбуждения, являющихся нейронным субстратом доминантного очага А.А.Ухтомского.

Отв. редактор проф. *A.М.Молчанов*  
Отв. за выпуск *B.И.Крюков*

Рецензенты канд. физ.-мат. наук *C.А.Пирогов*  
и д-р мед. наук профессор *Л.Г.Охнянская*

## ПРЕДИСЛОВИЕ

Одной из наиболее актуальных проблем современной физики является проблема метастабильности — экспериментально наблюдаемого, но теоретически не объясненного, существования долгоживущих неустойчивых состояний. Аналогичное по виду явление имеется в нейрофизиологии — незатухающий длительное время очаг возбуждения, являющийся предполагаемой нейронной основой доминанты А.А.Ухтомского. Однако, никто до сих пор не придавал серьезного значения указанной аналогии. Отчасти это, по-видимому, связано с тем, что идеи статистической физики пока мало известны нейрофизиологам. Более существенная причина состоит в том, что математическая физика пока еще не дала исчерпывающего объяснения явлению метастабильности. Поэтому особое значение приобретают вычислительные (имитационные) эксперименты и приближенные теоретические построения, позволяющие исследовать явления метастабильности в различных ситуациях.

Естественно, что коль скоро речь идет об изучении явлений, связанных с проявлениями метастабильности в функционировании нервной системы, то предпочтение должно быть отдано имитационным моделям и приближенным схемам, концептуально наиболее близким к предполагаемой структуре центральной нервной системы, и позволяющим объяснить большое число феноменов метастабильного поведения в мозге. Работа «Метастабильные и неустойчивые состояния в мозге» направлена на построение именно таких, реалистических, имитационных моделей и приближенных теорий относительно устойчивых (метастабильных) очагов возбуждения в центральной нервной системе. В этой работе математически строго сформулирована упрощенная, но достаточно близкая к реальности модель сети из неформальных (физических) нейронов и доказано существование фазового перехода 1-го рода в этой модели. Чтобы исследовать явления метастабильности в нейронных сетях, ав-

торы широко пользуются имитационным моделированием. С этой целью они формулируют еще один вариант своей основной модели фрагмента нейронной сети, специально приспособленный для вычислительного эксперимента. В этом варианте модели авторы обнаруживают два фазовых перехода и явление хорошо выраженной метастабильности («эффект сохранения пятна») в области пространства параметров, заключенной между двумя фазовыми переходами. Небольшая модификация основной имитационной модели (введение тормозного нейрона) приводит к системе, способной генерировать колебания суммарного уровня активности с периодом, значительно превышающим время релаксации отдельного элемента. Численный эксперимент обнаруживает в этой системе четко проявляющуюся периодичность по времени как суммарного уровня активности сети, так и импульсных последовательностей возбуждающих и тормозного нейронов. Стабильность колебаний может быть повышена введением в сеть малой пространственной неоднородности, играющей роль ведущего центра автоколебаний. Сами колебания носят выраженный релаксационный характер, т.е. каждый период колебаний состоит из относительно медленной фазы нарастания активности и быстрой фазы падения активности до нуля.

Все эти замечательные свойства, обнаруженные авторами в основной имитационной модели нейронной сети, применяются ими затем к построению моделей отдельных, функционально отличающихся, участков центральной нервной системы. Таким образом, авторам удалось с единой точки зрения объяснить целый ряд эффектов, наблюдавшихся ранее в динамике электрической активности мозга, существенно развить реверберационную теорию памяти и дать «микроскопическую» интерпретацию доминанты Ухтомского. Все эти результаты представляют собой единственный в мире пример математической теории вероятностных нейронных сетей, обнаруживающих при моделировании как совершенно неожиданные с чисто математической точки зрения свойства (эффект метастабильности в области параметров между двумя фазовыми переходами), так и позволяющих воспроизвести характерные особенности реальной деятельности мозга, вплоть до характерной формы электроэнцефалограмм при выработке и при угашении условного рефлекса. Столь широкий охват темы и глубина полученных результатов делают эту работу достойной внимания специалистов в различных областях математики, кибернетики и физиологии.

С.А.Пирогов

## **1. ВВЕДЕНИЕ. КРАТКИЙ ОБЗОР БОЛЕЕ РАННИХ МОДЕЛЕЙ**

Задача, которая решается в этой работе, тесно связана с основным принципом функционирования мозга. Однако чтобы не обсуждать пока вопрос о том, в чем этот принцип состоит или существует ли он вообще, мы поясним основную идею на нескольких частных случаях, представляющих, однако, большой самостоятельный интерес. В электрофизиологии нервной системы хорошо известны следующие три проблемы:

1. Проблема генератора пачечной ритмической активности.
2. Проблема кратковременной памяти.
3. Проблема местного и распространяющегося возбуждения.

Несмотря на кажущуюся разнородность, они, на наш взгляд, содержат общую особенность, состоящую в том, что у нейронной сети довольно неожиданно может появиться некоторое свойство кооперативности, а именно инерционность, которого не было у отдельных нейронов сети. Поясним это, раскрыв кратко содержание каждой из проблем в терминах межнейронного взаимодействия.

1. Многие нейронные структуры генерируют в определенных условиях пачечную ритмическую активность при поступлении на их входы неструктурированного случайного потока импульсов. При этом период повторения пачек может меняться в широких пределах и его величина во многих случаях на порядок больше постоянной времени мембранны одиночного апериодического нейрона, а вариабельность (неустойчивость) периода на порядок меньше вариабельности межспайковых интервалов одиночных апериодических нейронов. Спрашивается, каков механизм генерации подобной активности и, в частности, как могут возникать довольно медленные и регулярные колебания на основе быстрых, хаотических элементов?

2. Современная нейрофизиология рассматривает память человека и животных как сложный динамический процесс, разделенный на кратковременную и долговременную память. Первая связана с первичной электрофизиологической реакцией на внешние стимулы, вторая — с последующими структурно-функциональными изменениями, например с модификацией си-

напсов или каких-то других субструктур. Под кратковременной памятью обычно понимается способность нейронной сети удерживать или периодически (с периодом в несколько сотен миллисекунд) повторять пространственные конфигурации совозбужденных нейронов в течение времени, достаточно большого, чтобы вызвать структурные или функциональные изменения в нейронах, активных в этих конфигурациях. Спрашивается, как возникает столь большая задержка и большой период повторения в замкнутых цепочках взаимосвязанных нейронов и почему входные воздействия и сигналы не разрушают рециркуляцию импульсов в таких цепочках?

3. До недавнего времени импульсное возбуждение, распространяющееся по аксонам нейронов, считалось единственной формой возбуждения, исключающей существование локального, стационарного или медленно меняющегося возбуждения, что нашло свое отражение в известном принципе «все или ничего». Однако уже давно известно, что кроме бегущей волны существует медленное, почти стационарное, градуальное возбуждение, причем эта вторая форма часто оказывается более существенной, чем распространяющаяся импульсная активность. Спрашивается, какова взаимосвязь между этими видами возбуждения и, в частности, возможен ли переход из одной формы в другую как коллективный эффект, внезапно возникающий при изменении силы взаимодействия между нейронами.

Во всех трех проблемах присутствует одна и та же трудность: как согласовать упорядоченность и большое замедление кинетики на уровне системы нейронов и быстроизменяющуюся стохастическую активность на уровне одиночных нейронов. Естественно возникает вопрос, не существует ли похожей ситуации в физике? Оказывается, подобных физических явлений с резко различающимися микро- и макрохарактеристиками очень много. Так, явление метастабильности состоит в том, что некоторое специально приготовленное термодинамически неустойчивое состояние (например, переохлажденный пар или перегретая жидкость) имеет время релаксации (перехода) в устойчивое состояние, превышающее на много порядков время релаксации элементарных процессов. Все, что осталось сделать — это установить «изоморфизм» физической и физиологической модели и учесть малосущественную специфику каждой из трех проблем.

К сожалению, динамическое поведение нейрона сильно отличается от поведения атомов и молекул физических систем и естественно возникает законное сомнение в применимости указанной аналогии к мозгу. Необходимо не только объяснить «макроповедение» нейронной сети, но и его согласие с фактами, полученными в поведенческих опытах при регистрации электрической активности мозга с помощью макро- и микроэлектродной техники. Это накладывает жесткие ограничения на допустимый класс объяснений и моделей. Положение осложняется тем, что математическая физика пока еще не дала удовлет-

ворительной «микроскопической» теории явления метастабильности. Теория нейронных сетей, таким образом, столкнулась с серьезными трудностями и вынуждена искать свои собственные пути их преодоления. Приведем краткий обзор предыдущих попыток.

Узловым вопросом теории нейронных сетей является анализ роли взаимодействия между нейронами: связь активности одиночного нейрона и активности большой совокупности нейронов, связь микро- и макроописаний, возможность новых коллективных эффектов и, в частности, нейронной памяти.

Одна из первых попыток решения этих вопросов содержится в работах Э.А.Лямина (1973, 1975). В них дан критический разбор двух наиболее популярных моделей нейронной памяти и отмечен разрыв, существовавший между структурно-функциональным и нейрофизиологическим уровнями. Для его частичного устранения была предложена двухпопуляционная модель, средняя активность которой подчиняется примерно тем же закономерностям, что и активность отдельных нейронов, и описывается системой уравнений Ходжкина-Хаксли. С ее помощью удалось объяснить некоторые важные режимы работы гиппокампальной системы, тесно связанной с процессами памяти, и вплотную подойти к объяснению ее тета-ритмической активности. При этом был высказан ряд весьма важных для дальнейшего идей: о необходимости иметь иерархию времен релаксации различных подсистем, о решающей роли S-образной нелинейности и необходимости ее обоснования в терминах более низкого структурно-функционального уровня. Для их реализации и развития нам пришлось временно спуститься на уровень одиночного нейрона, чтобы затем снова вернуться к первоначальным узловым вопросам.

В серии работ, посвященных моделированию активности одиночных нейронов [Крюков, 1973а, 1973б, 1974, 1976; Кгуков, 1974, 1976], предложен ряд вероятностных моделей, характеризующихся учетом рефрактерности в такой форме, которая может описывать импульсную активность как независимых, так и взаимодействующих нейронов. Показано, что распределение межимпульсных интервалов удовлетворяет интегральному уравнению Вольтерра, пригодному для численных расчетов. Кроме того, в некоторых частных случаях получен ряд формул для таких распределений или их преобразований Лапласа. Попутно приведен обзор многих более ранних результатов. Основной метод, специально разработанный для этой задачи, состоит в распространении известного в теории вероятностей тождества Вальда на более широкий класс случайных процессов. Метод имеет также некоторое самостоятельное значение [Казанович, Крюков, 1976].

Полученные результаты позволили продолжить изучение нейронных сетей теоретико-вероятностными методами. В другой

серии работ [Крюков, 1975, 1977; Кгуков, 1978а; 1978б; Кириллов, Крюков, 1982] предложено описывать нейронную активность взаимодействующими марковскими процессами. Исходя из физиологически приемлемых допущений, выведено обратное уравнение Колмогорова, причем характеристический оператор процесса выражен через распределение межимпульсных интервалов одиночных нейронов. Тем самым установлена связь между двумя ранее независимыми областями, моделирующими активность нейронных сетей и соответственно одиночных нейронов. Затем были найдены довольно естественные достаточные условия для того, чтобы стационарная активность сети описывалась гибсовским распределением. Это, в свою очередь, позволило связать нашу модель с широким классом процессов, уже хорошо изученных в статистической физике, в частности в теории фазовых переходов.

Первой математической моделью нейронной сети с локальным взаимодействием, в которой возникал кооперативный эффект, соответствующий физическому фазовому переходу, была бесконечная сеть из формальных нейронов в дискретном времени, предложенная Ставской и Пятецким-Шапиро (1968). Условия отсутствия фазовых переходов для аналогичной модели в непрерывном времени привел Добрушин (1971а, 1971б). Он же построил пример, основанный на кинетической модели Изинга, в котором фазовый переход действительно существует, и высказал гипотезу: для конечной сети локальное взаимодействие компонент приведет к очень длительному переходу в стационарное состояние, если параметры и начальные условия выбраны должным образом. Работа содержала по существу первую строгую постановку задачи об описании долгоживущих состояний в довольно широком классе марковских процессов. Существенным ограничением здесь, однако, оказалось условие обратимости процесса во времени, а также применимость лишь к сетям из формальных нейронов. Потребностями физиологического моделирования была мотивирована работа Четаева (1979), в которой для конечной неприводимой обратимой марковской цели строится понятие фазы как метастабильного состояния.

Little (1974) высказал гипотезу о существовании в мозге долгоживущих состояний (*persistent states*) по аналогии с магнитными доменами, возникающими в модели Изинга при температуре ниже критической. Нейроны у него снова формальные, время дискретное и марковская цепь обратимая, что несколько снижает доверие к гипотезе. Убедительности гипотезы мешает трудность физиологической интерпретации этих состояний, определенных здесь весьма специальным образом — как собственные векторы матрицы переходных вероятностей, соответствующие наименьшему вырожденному собственному значению. Недавно Roney, Shaw (1980) заявили, что эта трудность ими устранена, и тем самым вроде бы они доказали первоначаль-

ную гипотезу Little. Однако такое заключение будет, по-видимому, преждевременным, поскольку их решение задачи на собственную функцию получено в линейном, «высокотемпературном» приближении и в его применимости не убеждены сами авторы [Roney, Shaw, 1980]. Кроме того, из подобного определения долгоживущих состояний, как отметил Schulman (1980), следует, что вероятности некоторых из этих состояний неизбежно должны быть отрицательными, что снова поднимает проблему интерпретации, на этот раз уже не физиологическую, а чисто математическую. Таким образом, необходимы альтернативные подходы.

В настоящей работе, в разделе 2 предложена сравнительно несложная, но достаточно близкая к физиологии модель сети из неформальных нейронов и доказано существование в ней фазового перехода [Круиков, 1984]. Модель учитывает импульсную природу нейронной активности, экспоненциальное затухание мембранныго потенциала и рефрактерность. Предполагается, что без учета этих фактов говорить о существовании фазовых переходов в нервной системе крайне рискованно. Модель представляет из себя марковскую систему, полученную из большого числа  $N$  первоначально независимых компонент, объединенных в систему с помощью локальных связей типа физического парного потенциала. Из микротеории систем такого типа известно [Cappuccio et al., 1974], что существование фазового перехода при  $N \rightarrow \infty$  влечет при некоторых дополнительных условиях существование метастабильных состояний для конечных  $N$ . Тем самым существование фазового перехода в базовой модели открывает путь для разработки адекватной динамической теории нейронной памяти как критического явления, существующего вблизи точки фазового перехода.

Довольно неожиданным оказалось, что небольшая модификация теории среднего поля, изложенная в разделе 3, качественно объясняет большое количество очень трудно интерпретируемых фактов из самых различных отделов ЦНС, в том числе все основные свойства доминанты А.А.Ухтомского (см. раздел 8 и заключение). Тем не менее необходима дальнейшая разработка теории в направлении более точного описания динамики микросостояний сети. Трудности построения микротеории этих состояний еще пока значительные и обусловлены в основном указанной выше спецификой нейронных сетей. Поэтому неоцененную помощь здесь может оказать имитационное моделирование.

В последние годы метод имитационного моделирования (метод Монте-Карло) достиг высокого уровня развития в физике и превратился в своеобразный и очень мощный инструмент исследования трудных проблем, особенно связанных с кооперативными, критическими явлениями. Наш подход к имитационному моделированию, изложенный в разделе 4, отличается от физического в двух аспектах:

- 1) динамика одиночного элемента (нейрона) определяется не

энергетическими соображениями (модельным гамильтонианом), а кинетикой формирования мембранныго потенциала и его пересечением критического порогового уровня срабатывания;

2) принцип обратимости динамики процесса во времени или иначе принцип детального баланса (неприменимый вследствие экспоненциального затухания мембранныго потенциала во времени, импульсной природы связи между элементами, наличия относительной и абсолютной рефрактерности) заменен естественной динамикой синаптических взаимодействий.

Основной вопрос математического моделирования нейронных сетей состоит в том, при каких условиях возможны сравнительно долгоживущие состояния в сети из малонерционных нейронов. Это, по существу, вопрос об условиях существования нейронной памяти.

Кроме этого основного вопроса интересны следующие более специальные вопросы, имеющие, однако, фундаментальное значение для ЦНС:

1. Возможны ли долгоживущие и метастабильные состояния в сети, имеющей только локальные возбуждающие связи, или для этого принципиально необходимы еще и тормозные связи?

2. Существует ли критический режим активности в сетях с локальными случайными связями? Какова при этом роль синаптических шумов?

3. Существуют ли и каковы проявления нейронных аналогов таких физических критических явлений, как спонтанная намагниченность, критическое замедление, переход беспорядок—порядок?

4. Помогут ли критические явления в решении узловых вопросов теории нейронной памяти и, в частности, в какой мере необходима гипотеза синаптической модификации и гипотеза реверберации импульсов по замкнутым путям?

5. Как совместить высокую надежность нервной системы с критическим и, по-видимому, неустойчивым режимом долгоживущих и метастабильных состояний?

Эти вопросы обсуждаются в заключительном разделе работы. Некоторые из них еще остаются открытыми. Однако несомненно, что сочетание имитационного моделирования и аналитических моделей в нейрофизиологии в ближайшем будущем станет стандартным методом научных исследований, ценность которого будет сравнима с экспериментальными методами на животных. Наиболее перспективным инструментом моделирования ЦНС остается теория многокомпонентных стохастических систем с локальным взаимодействием.

## **2. БАЗОВАЯ МОДЕЛЬ И ДАЛЕКИЙ ПОРЯДОК В НЕЙ**

Предложенная ниже модель является специальным случаем стохастического процесса, полученного наложением своеобразного взаимодействия многих изначально независимых марковских подсистем. В результате взаимодействия подсистема утрачивает марковское свойство, однако система как целое остается марковской. Такой процесс может иметь довольно необычное асимптотическое поведение и мы полагаем, что это существенно для понимания функционирования центральной нервной системы. Мы доказываем, в частности, что в такой системе существует фазовый переход. Это позволяет нам предложить новую интерпретацию кратковременной памяти как метастабильного состояния, существующего вблизи точки фазового перехода.

Наша основная задача состоит в следующем. Предположим, что электрическую активность одиночного независимого нейрона можно описать довольно простой, но все еще достаточно физиологичной моделью (скажем, с тремя—пятью параметрами) нейрона в непрерывном времени [Кгуков, 1976].

Введем локальные межнейронные связи так, что пресинаптический импульс, приходящий к данному нейрону, вызовет заметное изменение мембранныго потенциала, если только он пришел от соседних нейронов. Вопрос состоит в следующем: можно ли выбрать силу синаптических связей для данного набора параметров одиночного нейрона таким образом, чтобы появилась кратковременная память? Трудность заключается в том, что время удержания данной конфигурации должно превышать на 2–3 порядка время релаксации одиночного независимого нейрона. Время релаксации определяется главным образом постоянной времени затухания мембранныго потенциала и составляет несколько единиц или десятков миллисекунд. Поэтому положительное решение может существовать только благодаря кооперативному эффекту большой группы взаимосвязанных нейронов, поскольку мы не можем увеличить постоянную времени введением, например, медленных метаболических процессов.

Модель, описанная в работе Little (1974), наиболее близка к нашей и интерпретирует кратковременную память как ана-

лог далекого порядка в магнитной системе, описываемой кинетической моделью Изинга в дискретном времени. Существует несколько других попыток описать некоторые отделы нервной системы в терминах фазовых переходов [Крюков, 1975]. Все они основаны на формальных нейронах, которые однако не учитывают следующие факторы:

- импульсный характер нейронной активности;
- экспоненциальное затухание мембранныго потенциала в отсутствие афферентных импульсов;
- относительную рефрактерность.

Эти факторы затрудняют появление фазового перехода и, несомненно, нарушают так называемый принцип детального баланса, который является обычно неотъемлемым атрибутом кинетической модели Изинга. Таким образом, мы думаем, что проблема описания кратковременной памяти остается нерешенной.

В этом разделе мы доказываем возможность появления эффектов типа фазового перехода в нервной системе. При этом модель явным образом учитывает три указанные выше фактора. На этой основе мы выдвигаем гипотезу об условиях существования нового феномена — метастабильности. По этой гипотезе существуют точно определенные начальные состояния, которые имеют необычно большое время жизни (по сравнению с постоянной времени мембранны), если параметры модели выбраны предписанным образом. Гипотеза была проверена с помощью имитационной модели. Результаты проверки изложены в разделе 4, применение к моделированию конкретных нейронных структур дается в разделах 5—8.

## 2.1. Базовая нейронная модель

Предположим, что электрическая активность одиночного нейрона как подсистемы описывается некоторой моделью первого пересечения криволинейной границы [Кгуков, 1978b]. В соответствии с этой моделью нейрон срабатывает и генерирует нервный импульс, когда его мембранный потенциал (обычно дробовой шум или процесс Орнштейна-Уленбека) пересекает порог (убывающая непрерывная кривая), после чего восстанавливаются начальные условия и процесс продолжается до следующего пересечения и т.д. Таким образом, нейрон генерирует последовательность событий, времена появления которых образуют процесс восстановления. Обратное время возвращения этого процесса  $s_i^t$  (возраст последнего импульса к моменту  $t$ ) мы выбираем в качестве состояния  $i$ -й подсистемы. Предположим далее, что большая система таких нейронов, независимых друг от друга, размещена в узлах двумерной квадратной решетки. Введем, наконец, взаимные зависимости, соединив каждый нейрон с его ближайшими соседями, положительными связями силы  $a_{ij}$ , и с постоянной времени  $a_i^{-1}$ . Последнее означает

чает, что мембранный потенциал  $i$ -го нейрона в результате взаимодействия с  $j$ -м нейроном изменяется скачком по направлению к порогу на величину  $a_{ij}$  в точности в момент срабатывания  $j$ -го нейрона и затем влияние этого скачка затухает во времени экспоненциально со скоростью  $a_i$ .

Основной вопрос заключается в том, можно ли выбрать матрицу  $(a_{ij})$  таким образом, чтобы появилось необычное асимптотическое поведение, а именно очень медленная по сравнению с  $a_i$  сходимость процесса  $\{s_i^t\}$  к равновесному состоянию, образуя некоторое подобие памяти о начальном состоянии системы. Пытаясь ответить на этот вопрос, мы вывели (см. приложение 10.0) из указанных выше предположений следующее обратное уравнение Колмогорова:

$$\frac{\partial u_t(s)}{\partial t} = \sum_{i=1}^N c_i(s) [u_t(^0 s_i) - u_t(s)] + \frac{\partial u_t(s)}{\partial s_i}, \quad (2.1)$$

где  $s$  — вектор значений обратных времен возвращения и  $^0 s$  — тот же самый вектор, но в котором  $i$ -я компонента заменена нулем;  $c_i(s)$  — зависящая от возраста интенсивность отказов  $i$ -го процесса восстановления, модифицированного взаимозависимостью, — «известная» функция (см. далее соотношение (2.14)).

## 2.2. Достаточные условия фазового перехода

Пусть  $\mu(s)$  — инвариантная плотность вероятности процесса, описываемого уравнением (2.1), то есть  $\mu(s)$  — стационарное решение уравнения, сопряженного к (2.1):

$$\frac{\partial \mu_t}{\partial t} = \sum_{i=1}^N \left\{ \left[ \int_0^\infty c_i(x_s) \mu_t(x_s) dx \right] \delta(s_i) - c_i(s) \mu_t(s) - \frac{\partial \mu_t}{\partial s_i} \right\}, \quad (2.2)$$

где  $x_s$  — вектор  $s$ , в котором  $i$ -я компонента заменена на  $x$ . Следуя Круков (1978), назовем марковский процесс  $\{s_i^t\}$  псевдообратимым, если для всех  $i$

$$\frac{\partial \mu_i(s_i/x_s)}{\partial s_i} + c_i(s_i/x_s) \mu_i(s_i/x_s) - v_i(x_s) \delta(s_i) = 0, \quad (2.3)$$

где  $\mu_i(s_i/x_s) = \mu(s)/\mu(x_s)$  — условная инвариантная плотность возрастов  $i$ -го нейрона при условии, что известны состояния  $x$  всех его соседей,  $\delta(x)$  — дельта-функция. Из (2.2) ясно, что (2.3) просто обозначало бы компонентную стационарность, если бы все нейроны были взаимно независимы. Но в случае взаимной зависимости оно обозначает несколько большее: сохранение этой стационарности в случае, когда состояния соседей данного нейрона «заморожены» (фиксированы). Более подробно вероятностный смысл псевдообратимости объяснен в приложении 10.2. В частности, там показано, что условие (2.3) за-

меняет «принцип детального баланса» и подобно ему обуславливает доказываемый ниже факт, что  $\mu(s)$  — гиббсовская плотность, хотя и с другими условиями согласованности компонент.

**Предложение 2.1** [Крюков, 1978]. Пусть каждую функцию  $c_i$  можно представить в виде:

$$c_i(s) = \tilde{c}_i(s_i) + \gamma_i(s_i)L_i(s), \quad (2.4)$$

где  $\gamma_i$  и  $\tilde{c}_i$  зависят только от состояния  $i$ -го нейрона, в то время как  $L_i$  зависит только от состояний его соседей. Пусть  $\gamma_i$  — интегрируемая на  $(0, \infty)$  и

$$L_i(s) = \sum_j a_{ij} \int_{s_j}^{\infty} \gamma_i(x) dx / \int_0^{\infty} \gamma_i(x) dx, \quad (2.4a)$$

$$a_{ij} = a_{ji}. \quad (2.4b)$$

Тогда инвариантная плотность вероятностей псевдообратимого марковского процесса  $\{s_i\}$  имеет гиббсовскую форму

$$\mu(s) = Z^{-1} \prod_i \tilde{\mu}_i(s_i) \exp \left\{ - \sum_i b_i K_i(s_i) + \sum_{i < j} b_{ij} K_i(s_i) K_j(s_j) \right\}, \quad (2.5)$$

$$\text{где } b_i = \sum_j a_{ij}, \quad b_{ij} = a_{ij}/\beta_i, \quad \beta_i = \int_0^{\infty} \gamma_i(x) dx,$$

$$K_i(s_i) = \int_0^{s_i} \gamma_i(x) dx, \quad \tilde{\mu}_i(s_i) = \exp \left\{ - \int_0^{s_i} \tilde{c}_i(x) dx \right\}.$$

На бросок доказательства. Единственное решение уравнения (2.3) имеет вид:

$$\mu_i(s_i/s) = v_i \exp \left\{ - \int_0^{s_i} c_i(x) dx \right\}, \quad \text{где } i = 1, 2, \dots, N.$$

После подстановки (2.4) и (2.4a) получим выражение для одноточечной условной плотности вероятности стационарного марковского поля, глобальную плотность вероятности которого можно вычислить с помощью теоремы Аверинцева и леммы Бесага, приведенных в приложении 10.1. Подробное доказательство для несколько более общей модели нейронной сети приведено в работе [Крюков, 1978b].

**Теорема 2.1.** Пусть верны (2.4) и (2.4a). Предположим, что каждая функция  $\gamma_i(x)$  для некоторого постоянного  $h_i$  удовлетворяет уравнению:

$$\frac{d\gamma_i(x)}{dx} + \tilde{c}_i(x)\gamma_i(x) - \left( \frac{1}{2} \sum_j a_{ij} - \beta_i h_i \right) \gamma_i^2(x) = 0. \quad (2.6)$$

Тогда псевдообратимый марковский процесс  $\{Y_i^t\}$ , определяемый согласно выражению

$$Y_i^t + 1/2 = \int_{s_i^t}^{\infty} \gamma_i(x) dx / \int_0^{\infty} \gamma_i(x) dx, \quad (2.7)$$

имеет инвариантную плотность вероятности в гиббсовской форме, совпадающей с моделью Изинга с непрерывным спином:

$$p(Y) = Z^{-1} \exp \left\{ -\beta \left[ \sum_i h_i Y_i - \sum_{i < j} a_{ij} Y_i Y_j \right] \right\} v_1(dY_1) \dots v_N(dY_N) \quad (2.8)$$

с обратной температурой  $\beta = \int_0^{\infty} \gamma_i(x) dx$  (если все  $\gamma_i$  одинаковы), парным потенциалом  $a_{ij} = a_{ji}$ , внешним магнитным полем  $h_i$  и равномерным на  $(-1/2, 1/2)$  распределением состояний одиночного спина

$$v_i(dY_i) = \begin{cases} dY_i, & -1/2 < Y_i < 1/2 \\ 0, & \text{в противном случае.} \end{cases}$$

Если дополнительно предположить, что для всех  $i$  и  $j$

$$h_i = 0, \quad a_{ij} = a_{|i-j|} > 0, \quad (2.9)$$

тогда система  $\{Y_i^t\}$  для достаточно больших  $\beta$  имеет далекий порядок в термодинамическом пределе  $N \rightarrow \infty$ , то есть

$$\lim_{|i-j| \rightarrow \infty} \int_{R^{Z^2}} Y_i Y_j \mu(Y) dY \geqq A > 0. \quad (2.10)$$

**Доказательство.** Плотность (2.8) следует из (2.5) после замены переменных по формуле (2.7):

$$p(Y) = \mu(s) \prod_i \left| \frac{ds_i}{dY_i} \right| = Z^{-1} \exp \left\{ \sum_i \left[ - \int_0^{s_i} \tilde{c}_i(x) dx - \frac{1}{2} \sum_j a_{ij} Y_i \beta_i - \ln \gamma_i(s_i) \right] + \sum_{i < j} a_{ij} Y_i Y_j \beta_i \right\} v_1(dY_1) \dots v_N(dY_N).$$

Здесь выражение в квадратных скобках равно  $\beta_i h_i Y_i$  тогда и только тогда, когда выполняется (2.6). Доказательство (2.10) в наших условиях следует из того факта, что оно справедливо для изинговского ферромагнетика с непрерывным спином, точная формулировка которого приведена в приложении 10.3.

### 2.3. Задача о первом достижении

В этом разделе мы обсуждаем возможность согласовать условия теоремы 2.1, наложенные на функции  $c_i(s)$ , с базовой нейронной моделью.

Пусть  $x_i^t$  — мембранный потенциал  $i$ -го нейрона,  $n_i^t$  — его шумовая компонента,  $y_i^t$  — его порог в момент  $t$ . Тогда согласно базовой нейронной модели

$$x_i^t = \sum_j a_{ij} e^{-a_j s_j^t} + n_i^t + x_i^0 \quad (2.11)$$

$$y_i^t = y_i^\infty + (y_i^0 - y_i^\infty) e^{-a_i s_i^t} \quad (2.12)$$

$$(y_i^0, y_i^\infty > 0, y_i^0 > y_i^\infty).$$

Нейрон генерирует импульс в некоторый момент  $\tau$  после предыдущего импульса тогда и только тогда, когда  $\tau$  — момент первого пересечения процесса  $n_i^t$  и «эффективного» порога  $r_i^t$ , где

$$r_i^t = y_i^t - \sum_j a_{ij} e^{-a_j s_j^t}. \quad (2.13)$$

Следовательно, фиксируя состояния  $s_j^t$  всех нейронов, являющихся соседями  $i$ -го нейрона в каждый момент  $t$  (то есть фиксируя кривую  $r_i^t$ ), можно в принципе вычислить условную плотность вероятностей межимпульсных интервалов  $f_i(t|s)$  точно так же, как в случае независимых нейронов с криволинейным порогом [Круиков, 1976]. Тем самым можно получить необходимую информацию относительно функции  $c_i(s)$ , так как

$$c_i(s) \equiv c_i(s_i/s) \equiv f_i(s_i/s) / \left( 1 - \int_0^{s_i} f_i(x/s) dx \right). \quad (2.14)$$

Очевидно,  $c_i(s) = \tilde{c}_i(s_i)$  для независимых нейронов. Точные формулы для  $c_i$  и  $\tilde{c}_i$  неизвестны. Простейшая догадка приводит к (2.4). Но может оказаться, что (2.4) не обязательно верно в общем случае и его нужно понимать только как аппроксимацию, приводящую к более простым условиям далекого порядка. В частности, можно считать, что (2.4) дает линейную часть тейлоровского разложения функции  $c_i$  в ряд по степени  $L_i$ , если

$$L_i(s) = \sum_j a_{ij} e^{-a_j s_j},$$

так как согласно (2.13)  $L_i(s)$  пропорционально «возмущению» порога вследствие взаимодействия нейронов. Тогда согласно (2.4а)

$$\gamma_i(t) = k_1 e^{-a_i t} \quad (k_1 > 0).$$

Но эта функция может быть решением уравнения (2.6) тогда и только тогда, когда

$$\tilde{c}_i(t) = a_i + k_2 e^{-a_i t} \quad (k_2 > 0), \quad (2.15)$$

в чем можно убедиться, разрешая уравнение (2.6) относительно  $\tilde{c}_i$  при данной  $\gamma_i$ . С другой стороны,  $\tilde{c}_i$  определяется исключительно моделью одиночного независимого нейрона. Следовательно, чтобы теорема 2.1 согласовалась с базовой нейронной моделью, выражение (2.15) должно быть решением задачи о первом пересечении для независимых нейронов. Что это возможно, по крайней мере, асимптотически, видно из следующего предложения.

**Предложение 2.2.** Пусть шум  $n_i^t$  описывается стохастическим процессом Орнштейна-Уленбека с инфинитезимальным сном  $a_i$  и дисперсией  $\sigma_i^2$ , а порог  $y_i^t$  имеет вид (2.12). Тогда при  $t \rightarrow \infty$

$$\tilde{c}_i(t) = a_1 + \frac{a_1}{a_2}(a_2 - a_1) \frac{\beta_2}{\beta_1} e^{-(a_2 - a_1)t} + 0(e^{-(a_2 - a_1 + \epsilon)t}), \\ \epsilon > 0, \quad (2.16)$$

где  $-a_1$  и  $-a_2$  — наименьшие нули функции параболического цилиндра; и  $\beta_1$  и  $\beta_2$  — вычеты функции  $\Phi(\lambda)$  в полюсах  $-a_1$  и  $-a_2$ , где (при  $\operatorname{Re} \lambda > 0$ )

$$\Phi(\lambda) = \frac{D_{-\lambda/a_1}((x_i^0 - y_i^0 + y_i^\infty)/\sigma_i)}{D_{-\lambda/a_1}(-y_i^\infty/\sigma_i)} \exp \left\{ \frac{(x_i^0 - y_i^0 + y_i^\infty)^2 - (y_i^\infty)^2}{4\sigma_i^2} \right\} \quad (2.17)$$

**Доказательство.** Формулу (2.17) получил Кгуков (1976) как преобразование Лапласа плотности вероятности межимпульсных интервалов  $f_i(t)$  для независимых нейронов. Используя обратное преобразование и соотношение

$$\tilde{c}_i(t) = -\frac{d}{dt} \ln \left[ 1 - \int_0^t f_i(x) dx \right],$$

получаем (2.16).

#### 2.4. Существование метастабильных состояний

Математически фазовый переход — явление бесконечномерное и требующее отсутствия магнитного поля в случае изинговского ферромагнетика. Первое требование можно оправдать огромным числом ( $10^{11}$ ) нейронов в мозге. Второе, как мы видели в разделе 2.3, не так уж легко реализуется в базовой нейронной модели и тем более в реальной нервной системе. Поэтому вряд ли далекий порядок в модели Изинга является достаточно близким аналогом кратковременной нейронной памяти. Существует другое явление вблизи точки фазового перехода, которое, по-видимому, лучше имитирует кратковременную память, а именно метастабильность. Это явление можно наблюдать при малых магнитных полях и оно типично для конечных систем. Точное определение дано Саросаккиа et al., (1974). Мы приведем лишь перечень его основных свойств:

- в состоянии метастабильности присутствует только одна фаза;
- системе с таким начальным состоянием, вероятнее всего, потребуется много времени, чтобы выйти из него;
- когда система вышла из него, маловероятно, что она вернется в него.

Следующие факты доказаны Саросаккиа et al., (1974). (Мы не имеем возможности привести полные результаты, поскольку их формулирование потребовало бы многих новых понятий и обозначений).

Рассмотрим изинговскую  $1/2$ -спиновую систему со взаимодействием ближайших соседей с силой  $J > 0$  и магнитным полем  $h < 0$  при обратной температуре  $\beta$ , заключенную в ящик  $\Lambda$  с граничным условием  $+$ , временная эволюция которой удовлетворяет условию детального баланса. Тогда начальное состояние, определенное как равновесная гиббсовская мера на конфигурациях, которые имеют площадь любого кластера меньше некоторого числа  $l^2$  и нулевую вероятность для любых других конфигураций, является метастабильным, если  $\beta > 2 \ln 6 / J$  и  $l < 4(\beta J - 2 \ln 6) / |\beta| h$ . Имеются строгие, не зависящие от  $\Lambda$  нижние и верхние границы для времени релаксации, которые показывают, что последнее уменьшается линейно по объему  $|\Lambda|$ , но тем не менее может быть достаточно большим при соответствующем выборе параметров. Далее, для

$$J > |h| \geq AJe^{-\beta J}, \quad A = \text{const} \quad (2.18)$$

приведена следующая оценка размеров, соответствующая максимальным временам релаксации:

$$\left| l - \frac{2J}{|h|} \right| \leq \sqrt{\frac{2(J - |h|)}{|h|}} \left( 1 + O\left(\frac{1}{\beta}\right) \right). \quad (2.19)$$

Предварительный анализ доказательств работы Carosaccia et al. (1974) показывает, что подобные результаты, вероятно, останутся справедливыми и для кинетической модели Изинга с непрерывным спином. Однако детальное доказательство все еще отсутствует. Поэтому мы сформулируем предположение, которое, по-видимому, не противоречит нашим результатам имитационного моделирования, изложенным в разделе 4.

**Предположение.** В условиях теоремы 2.1 положим  $\gamma_i = \gamma$  для всех  $i$  и  $a_{ij} = J > 0$ , если  $|i - j| = 1$ , и  $a_{ij} = 0$  в противном случае. Тогда существует такое  $\Theta > 0$ , что марковский процесс, определенный преобразованием

$$\eta_i^\ell(\Theta) = \text{th}\left(B \int_0^{s_i^\ell} \gamma_i(x) dx\right), \quad B > 0$$

на квадратной решетке, содержащей  $N$  нейронов, с периодическими граничными условиями для достаточно больших  $\beta$  и малых  $l$ , имеет метастабильное состояние с тем же самым распределением, как и у кинетической модели изинговского ферромагнетика со спином  $1/2$  с парным потенциалом  $J$ , внешним полем  $h$ , определенным по (2.6), и  $\beta = \int_0^\infty \gamma_i(x) dx$ . Если выполнено (2.18), тогда справедлива оптимальная оценка (2.19).

### **3. ДИФФУЗИОННАЯ АППРОКСИМАЦИЯ**

#### **3.1. Введение**

В этом разделе продолжено изучение кооперативных эффектов в модели нейронной сети, предложенной в разделе 2. Основной результат состоит, грубо говоря, в том, что состояния неустойчивого равновесия динамической системы, описывающей сеть, становятся под действием внутренних шумов при некоторых значениях параметров модели относительно долгоживущими.

Получено качественное согласие с результатами имитационного моделирования и предложена новая интерпретация «трудных» для обычного понимания данных по электрофизиологии коры и мозжечка. Результаты подтверждают гипотезу о существовании в мозге инерционных микроочагов возбуждения как нейронного субстрата кратковременной памяти, внимания, поведенческих реакций животных.

В предыдущем разделе предложена модель нейронной сети, предположительно обладающая необычным кооперативным свойством: некоторые начальные состояния ее активности имеют сравнительно большое время жизни, если параметры подобраны оптимальным образом. Эти состояния, по-видимому, аналогичны метастабильным состояниям статистической физики, но их время жизни не столь велико, не так сильно зависит от размеров системы. Понимание этих вопросов пока недостаточно ясное, поскольку существующая в физике микроскопическая теория метастабильности, изложенная, например, в работе Сарасакка *et al.* (1974), предполагает существование детального баланса, что не может быть выполнено в нейронной сети в силу причин, указанных в начале предыдущего раздела.

Поэтому мы вынуждены обратиться к более простой диффузионной аппроксимации типа теории среднего поля в ферромагнетизме. Несмотря на свою простоту, она оказалась весьма успешной в объяснении качественного поведения сложных физических систем, а также в качестве исходного пункта для последующих уточнений и усовершенствований. В данном разделе она развита с учетом потребностей и возможностей физиологического эксперимента и поэтому не содержит, например, традиционного для физики обсуждения критических ин-

дексов. Вместо этого более пристальное внимание уделено динамическим эффектам и оценке времени жизни неустойчивых состояний.

Диффузионная модель нейронной сети ранее рассматривалась в работе Venzi (1976). В отличие от нее, мы не постулируем лежащую в основе модели динамическую систему, а выводим ее из первых принципов ЦНС. В результате механизм образования метастабильности, динамического гистерезиса и кратковременной памяти у нас оказывается другой. Этот механизм вместе с результатами других работ этой серии позволяет предложить новую модель организации памяти и внимания животных, структура которой намечена в работе Крюкова (1982).

### 3.2. Вывод динамической системы

Постулаты базовой модели были даны ранее, и здесь мы приведем лишь результирующее обратное уравнение Колмогорова и некоторые дополнительные предположения, которые понадобятся в дальнейшем. Это уравнение имеет вид

$$\frac{\partial u_i(s)}{\partial t} = \sum_{i=1}^N c_i(s)[u_i(^0s) - u_i(s)] + \frac{\partial u_i(s)}{\partial s_i}, \quad (3.1)$$

где  $u_i(s)$  — достаточно гладкая функция, описывающая эволюцию некоторого функционала марковского процесса как функцию начального состояния  $s$ , где  $s = \{s_i\}$  — вектор обратных времен возвращения системы взаимодействующих точечных процессов;  $^0s$  — тот же вектор, но с  $i$ -й компонентой, замененной нулем;  $c_i(s)$  — функция условной интенсивности срабатывания  $i$ -го нейрона, определение и некоторые свойства которой приведены в разделе 2.3.

Теперь в дополнение к постулатам базовой модели мы будем считать что, во-первых, все параметры не зависят от номера нейрона; во-вторых, каждый нейрон имеет период абсолютной рефрактерности  $t_{\text{абс}} = 1/a$ , и, наконец, что каждый нейрон сети связан с каждым нейроном, причем

$$a_{ij} = \frac{a}{N}, \quad i, j = 1, \dots, N; \quad a = \text{const.} \quad (3.2)$$

Тогда эволюция суммарной активности  $\xi'$

$$\xi' = \frac{1}{N} \sum_{j=1}^N e^{-as_j} \quad (3.3)$$

описывается следующим обратным уравнением Колмогорова:

$$\frac{\partial U_i}{\partial t} = -[a\xi - S(a\xi)] \frac{\partial U_i}{\partial \xi} + \frac{D(a\xi)}{N} \frac{\partial^2 U_i}{\partial \xi^2}, \quad (3.4)$$

Его можно получить из (3.1), если взять  $u_t(s) = U_t(\xi)$ , где  $U_t(\xi)$  — достаточно гладкая функция обоих аргументов; разложить выражение в прямоугольных скобках уравнения (3.1) в ряд Тейлора около точки  $\xi$  и учесть, что при  $N \rightarrow \infty$

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (1 - e^{-as_i}) c_i(s) &\xrightarrow{P} S(a\xi), \\ \frac{1}{2N} \sum_{i=1}^N (1 - e^{-as_i})^2 c_i(s) &\xrightarrow{P} D(a\xi), \end{aligned} \quad (3.5)$$

понимая сходимость в слабом смысле, то есть как сходимость суммы одинаковых функций от идентично распределенных величин  $\{s_i\}$ ,  $i = 1, 2, \dots, N$  при условии, что фиксировано значение суммы (3.3).

Уравнение (3.4) эквивалентно уравнению динамической системы с шумами:

$$\frac{d\xi_t}{dt} = -a\xi_t + S(a\xi_t) + \sqrt{\frac{D(a\xi_t)}{N}} \frac{dW(t)}{dt},$$

где  $W(t)$  — винеровский процесс. Нужные нам временные свойства процесса  $\xi_t$  будут рассмотрены в конце этого раздела. Они в значительной степени определяются детерминистической частью системы. Однако удобно и полезно вначале рассмотреть систему кинетических уравнений для среднего  $E(t) = M\xi_t$  и дисперсии  $V(t) = M\{(\xi_t - E(t))^2\}$ . Она имеет вид:

$$\frac{dE}{dt} = -aE + S(aE) + \frac{a^2 S''}{2} (aE) V \quad (3.6)$$

$$\frac{dV}{dt} = -2V(a - aS'(aE)), \quad (3.7)$$

где  $S'$ ,  $S''$  — производные по  $E$ . Эта система получается следующим образом. Запишем уравнение, сопряженное к (3.4).

$$\frac{\partial \mu_t}{\partial t} = \frac{\partial}{\partial \xi} \{(a\xi - S(a\xi))\mu_t\} + \frac{1}{N} \frac{\partial^2}{\partial \xi^2} \{D(a\xi)\mu_t\}, \quad (3.8)$$

где  $\mu_t$  — плотность вероятностей процесса  $\xi_t$  в момент  $t$ . Умножим обе части на  $\xi$  и проинтегрируем по  $\xi$  от 0 до  $\infty$ . Слева получим производную  $E(t)$ . Предполагая, что  $\mu_t(\xi)$  стремится к нулю на бесконечности быстрее, чем  $1/\xi^2$ , и производя в правой части (3.8) интегрирование по частям, получим

$$\frac{dE}{dt} = -aE + MS(aE). \quad (3.9)$$

Аналогичные выкладки с умножением обеих частей (3.8) на  $(\xi - E)^2$  показывают, что

$$\frac{dV}{dt} = -2aV + 2M\{(\xi_i - E)S(a\xi_i)\} + MD(a\xi_i)/N. \quad (3.10)$$

Разлагая функции  $S(a\xi_i)$  в ряд Тейлора в окрестности точки  $E$ , беря математические ожидания, указанные в (3.9) и (3.10), и пренебрегая старшими моментами, получим (3.6) и (3.7). Для дальнейшего анализа нам понадобятся некоторые свойства функции  $S(x)$ .

### 3.3. Обоснование сигмоидной нелинейности (задача Э.А.Лямина)

**Предложение 3.1.** Уравнение

$$ax - S(ax) = 0, \quad (3.11)$$

в зависимости от значения параметров модели имеет одно, два или три решения, причем в последнем случае два устойчивых ( $x_1, x_2$ ) разделены неустойчивым решением  $x_0$ , то есть

$$a - aS'(ax_0) < 0 \quad (3.12)$$

$$a - aS'(ax_{1,2}) > 0 \quad (3.13)$$

Эскиз доказательства. Согласно вероятностному смыслу функции  $c_i(s)$ , приведенному в разделе 2 (формулы (2.11)–(2.14)), она зависит не от каждой компоненты  $s_i$  в отдельности, а только от  $s_i$  и от следующей суммы

$$\sum_{j=1}^N a_{ij} e^{-as_j} = \frac{1}{N} \sum_{j=1}^N a e^{-as_j} = a\xi.$$

Поэтому вместо (3.5) по закону больших чисел имеем

$$S(a\xi) = \int_0^\infty p(s_i)(1 - e^{-as_i})c_i(s)ds_i,$$

где  $p(s_i)$  — плотность вероятностей возрастов  $s_i$  в стационарном состоянии, причем для достаточно больших  $N$

$$c_i(s) = \tilde{c}_i(s_i) + O\left(\frac{1}{N}\right),$$

$$r_i^t = y_i^\infty + (y_i^0 - y_i^\infty)e^{-as_i} - a\xi^t.$$

Это означает, что учет взаимодействия между нейронами сводится к уменьшению величины эффективного порога на величину  $a\xi$ . Поэтому используя предложение 2.2, получим

$$S(a\xi) = k a_1 + O(1/N) + \text{дополнительные слагаемые}, \quad (3.14)$$

где  $k = \int p(s_i)(1 - e^{-as_i})ds_i$ ,  $a_1 \equiv a_1(x)$  — наименьший ноль параболического цилиндра  $D_{\lambda/a}(x)$ , вычисленный при

$$x = (y_i^\infty - a\xi)/\sigma.$$

Если использовать численные значения  $\alpha_i$ , приведенные в работе Beekman (1975), и построить график зависимости  $\alpha_i$  от  $x$ , то можно убедиться, что с хорошей точностью  $\alpha_i(x) \cong \alpha e^{-x}$ . Это означает, что уравнение (3.11) имеет самое большое два решения, если дополнительные слагаемые в (3.14) не вносят существенных поправок. Третье решение возникает, если учесть, что абсолютная рефрактерность, исключенная выше из рассмотрения, ограничивает сверху функцию  $\tilde{c}_i(s_i)$  величиной  $1/t_{abc}$ . Фактически при этом вместо  $\tilde{c}_i$  следует взять

$$\tilde{c}_i(s_i)_{abc} = 1/(t_{abc} + 1/\tilde{c}_i(s_i)).$$

Если для этой функции повторить приведенные рассуждения, то получим

$$S(x) = ka / \left( 1 + \exp \left\{ -\frac{x - y_i^\infty}{\sigma} \right\} \right) + \text{дополнительные слагаемые.}$$

Теперь уже первое слагаемое функции  $S(x)$  удовлетворяет условиям (3.11)–(3.13) и можно показать, что дополнительные слагаемые не изменят этого заключения.

### 3.4. Динамические эффекты вблизи точки неустойчивого равновесия

$S$ -образный характер кривой  $S(x)$  означает, что детерминированная часть нашей динамической системы, имеющая вид

$$\frac{dx}{dt} = -ax + S(ax),$$

обладает *гистерезисом*, причем критические значения  $a_c$  и  $x_c$ , показанные на рис. 3.1, определяются из уравнения

$$a = a_{c1,2} S'(a_{c1,2} x_{c1,2}) = 0.$$

Рассмотрим роль флуктуаций на основе системы уравнений (3.6) и (3.7), в которой фиксированы все параметры, кроме  $a$ . Основные результаты качественного анализа этой системы следующие (подробное исследование системы приведено в приложении 10.4):

1. При фиксированном  $a \in (a_{c1}, a_{c2})$  в системе одновременно существует два устойчивых стационарных состояния  $(E_1, 0)$  и  $(E_2, 0)$  и одно неустойчивое  $(E_0, 0)$ . Если начальное состояние системы находится вблизи точки  $(E_1, 0)$  или  $(E_2, 0)$ , то система довольно быстро приходит в равновесное состояние (фрагменты  $a$ , в рис. 3.2А, Б, В). Если же начальное состояние системы находится вблизи точки  $(E_0, 0)$ , то наблюдается аномальное увеличение дисперсии  $V(t)$  от сколь угодно малых значений до макроскопически значимых при относительно небольших изменениях  $E(t)$  (фрагмент с рис. 3.2А, Б, В). Подобный

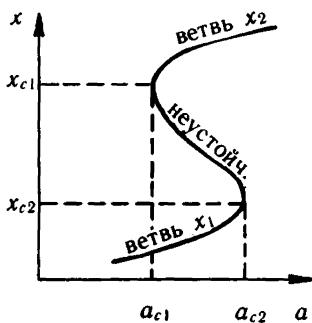


Рис. 3.1. Критические значения  $a$  и  $x$

деле 4. В разделе 8 он будет использован для объяснения необычного постепенного удлинения (до 850 мс) вызванных ответов ЭЭГ, регистрируемых в процессе выработки условных рефлексов

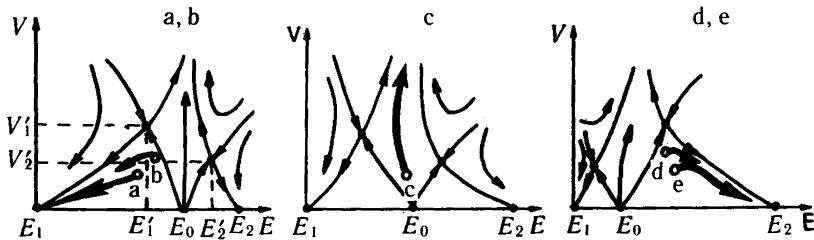


Рис. 3.2А. Возникновение критического режима при изменении параметра связи  $a$  ( $a_{c1} < a < a_{c2}$ ). Начальный уровень активности сети фиксирован

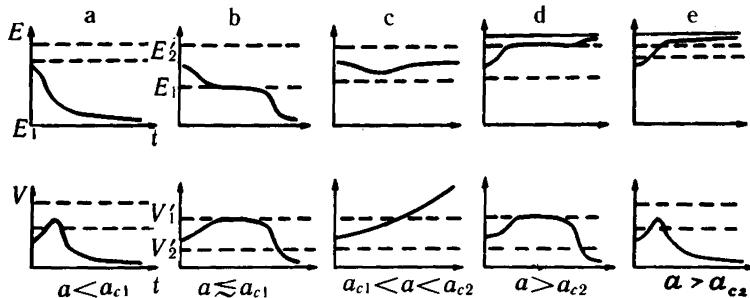


Рис. 3.2Б. Изменение динамики активности  $E(t)$  и дисперсии  $V(t)$  при различных значениях параметра связи  $a$

рост дисперсии можно сопоставить с аналогичным эффектом в имитационной модели в разделе 4, а также с эффектом маскировки шумами вызванных ответов ЭЭГ [Мнухина, 1977].

2. При постепенном изменении силы связи  $a \uparrow a_{c1}$  или при  $a \downarrow a_{c2}$  время выхода на одно из устойчивых стационарных состояний неограниченно растет как  $(a - a_c)^{-1}$ . Этот эффект при  $a_{c1} \approx a_{c2}$  напоминает критическое замедление кинетики в ферромагнетиках и его можно рассматривать как макроскопический аналог эффекта сохранения пятна, приведенного в раз-

Рис. 3.2В. Схематическое изображение среднего уровня активности при фиксированных параметрах, но различных начальных состояниях, примерно соответствующих изображающей точке, показанной на рис. 3.2А

[Мнухина, 1977]. Этим же эффектом легко объясняется необычное удлинение затухания авто- и кросскорреляционных функций активности основных нейронов мозжечка (до 200—500 мс) при стимуляции [Ebner, Bloedel, 1981а, 1981б]. Более подробно об этом будет изложено в разделе 8.

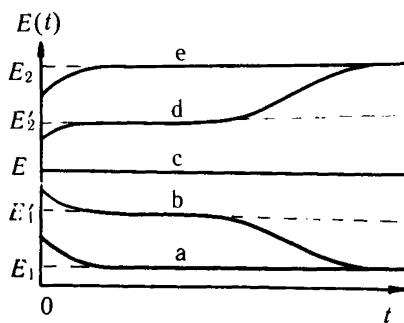
3. При фиксированном начальном состоянии ( $E(0)$ ,  $V(0)$ ) и постепенном монотонном увеличении  $a$  время выхода вначале возрастает (рис. 3.2а,в), а затем (при  $a > a_{c2}$ ) падает (рис. 3.2д,е), причем отклонение  $E(t)$  от  $E(0)$  меняет свой знак на противоположный: сначала при малых  $a$  система приходит в стационарное состояние ( $E_1$ , 0), а потом по мере увеличения параметра  $a$  — в состояние ( $E_2$ , 0). При этом смена знака отклонения с минуса на плюс происходит при  $a = a_{c1}$ , а с плюса на минус — при  $a_{c2}$  ( $a_{c1} < a_{c2}$ ). Поразительно похожее поведение вызванных потенциалов ЭЭГ наблюдала Мнухина (1977) при выработке условных рефлексов, включая удлинение, смену полярности, укорочение и обратный порядок этой последовательности при постепенном угашении рефлекса.

Отметим, что аналогичные эффекты будут наблюдаться при изменении любого другого параметра модели при фиксированных остальных параметрах. Значит, кратковременная память в нашем понимании не обязательно связана со структурными изменениями синаптического аппарата.

### 3.5. Оценки времени жизни метастабильных и неустойчивых состояний

Пусть  $x$  — начальное значение диффузационного процесса, описываемого уравнением (3.4). Обозначим  $\tau_x$  — время первого достижения стационарного состояния  $x_2$  и  $t_x$  — время первого достижения какого-либо из стационарных состояний  $x_1$  или  $x_2$ , исходя из состояния  $x$ ,  $x_1 < x < x_2$ . Пусть  $T_x = M\tau_x$  и  $t_x = M t_x$ . На рис. 3.3 приведен качественный вид этих функций в связи с кривой  $S(x)$  и потенциалом  $\phi(x)$ .

**Предложение 3.2.** В окрестности неустойчивой точки  $x_0$  существует такое начальное состояние  $x_*$ , что  $t_{x_*} = \max_x t_x$  (см. рис. 3.3). Точка  $x_*$  может быть найдена как единственное



решение уравнения, получаемого приравниванием правых частей следующих выражений:

$$t_x = \int_{x_1}^x e^{N\varphi(y)} \int_{y_*}^y \frac{N}{D(az)} e^{-N\varphi(z)} dz dy \quad (x_1 < x \leq x_*) \quad (3.15)$$

$$t_x = \int_x^{x_2} e^{N\varphi(y)} \int_{x_*}^y \frac{N}{D(az)} e^{-N\varphi(z)} dz dy \quad (x_* \leq x < x_2), \quad (3.16)$$

где  $\varphi(x) = \int_0^x \frac{ay - S(ay)}{D(ay)} dy$ , причем

$$\frac{(x_* - x_1)N}{ax_{c1} - S(ax_{c1})} > t_{x_*} > \frac{(x_0 - x_1)}{ax_{c1} - S(ax_{c1})} \left( \begin{array}{l} x_0 \neq x_1 \\ \varphi(x_1) > \varphi(x_2) \end{array} \right). \quad (3.17)$$

Для достаточно больших  $N$

$$T_{x_1} \gg t_{x_*} \quad (3.18)$$

за исключением точек  $x_1$  вблизи  $x_{c1}$ , где

$$T_{x_1} \sim 2t_{x_*} \quad (3.19)$$

при этом [Agarwal, Shenoy, 1981]

$$T_{x_1} \sim \frac{\sqrt{2}N0.83}{D(ax_{c1})} (a_{c1} - a)^{1/2} \left( \frac{6D(ax_{c1})}{\varphi'''(x_{c1})} \right)^{1/3} \left( \frac{\partial^2 a}{\partial x_{c1}^2} \right)^{-1/2} \quad (3.20)$$

Доказательство. Согласно Феллеру (1958) функция  $t_x$  является единственным решением уравнения

$$A \frac{\partial t_x}{\partial x} + B \frac{\partial^2 t_x}{\partial x^2} = 1 \quad (3.21)$$

$$(A = -ax + S(ax), \quad B = D(ax)/N),$$

для которого  $t_{x_1} = t_{x_2} = 0$ . Отсюда, обозначая  $t'_x = \frac{\partial t_x}{\partial x}$ , получим

$$t'_x = -e^{N\varphi(x)} \int_{x_*}^x \frac{N}{D(az)} e^{-N\varphi(z)} dz.$$

Отсюда получаем (3.15) и (3.16). Применяя метод перевала к внутреннему интегралу (3.15) и оценивая внешний интеграл (3.15) по теореме о среднем, получим (3.17) при  $x_* > x_0$ . Аналогичным образом, находя  $T_x$  как решение уравнения (3.21) с граничными условиями  $\frac{\partial}{\partial x} T_{x=0} = T_{x=x_2} = 0$ , получим:

$$T_x = \int_x^{x_2} e^{N\varphi(y)} \int_0^y \frac{N}{D(az)} e^{-N\varphi(z)} dz dy, \quad (0 < x < x_2).$$

Откуда двукратным применением метода перевала находим

$$T_{x_1} \sim \frac{2\pi N}{D(ax_1) \sqrt{\varphi''(x_1)\varphi''(x_0)}} e^{N[\varphi(x_0) - \varphi(x_1)]}. \quad (3.22)$$

Отсюда для больших  $N$  очевидно неравенство (3.18) ввиду (3.17). При  $\varphi''(x_1)=0$  формула (3.22) заменяется формулой (3.20), полученной Agarwal, Shenoy (1981).

Формула (3.19) следует из того, что при  $\varphi''(x_1)=0$  снос процесса на интервале  $(x_1, x_*)$  практически отсутствует.

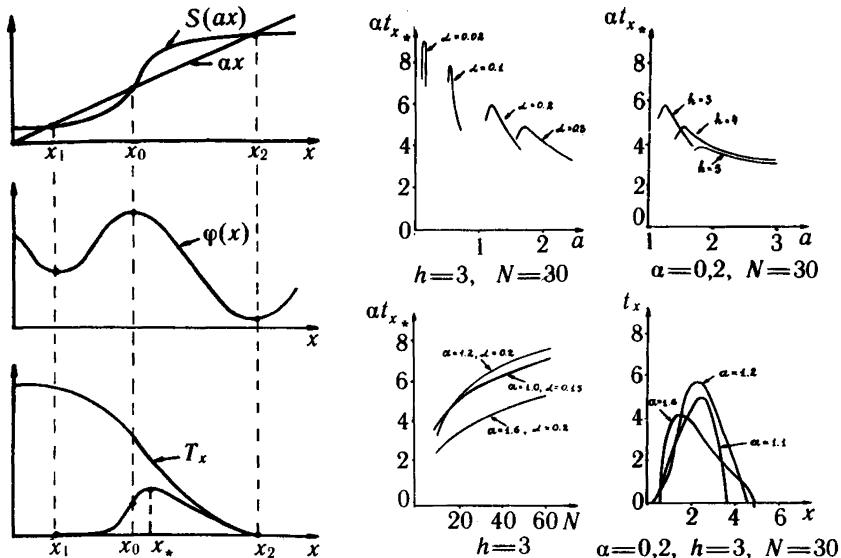


Рис. 3.3. Связь между различными функциями в Предложении 3.2

Рис. 3.4. Решение уравнения (3.21) при  $A = -ax + (1 + e^{-ax+h})^{-1}$ ;  $B = 1/N$

Детальное численное исследование поведения  $t_x$  при различных значениях параметров приведено на рис. 3.4. Качественно оно согласуется с результатами имитационного моделирования, приведенного далее в разделе 4, хотя количественно  $at_{x_*}$  значительно меньше наблюдаемого в экспериментах главным образом ввиду допущения (3.2) о малости взаимодействия.

Важнейшим фактом, подтверждающим возможность существования в мозге метастабильных и неустойчивых состояний, является экспериментальное наблюдение гистерезисного поведения активности клеток Пуркинье мозжечка, а также спонтанные переходы в нем между двумя устойчивыми состояниями активности со средними временами жизни  $T_{x_1}$  порядка нескольких минут [Colin et al., 1980]. Подробный анализ и интерпретация этих фактов в терминах нашей модели будут приведены в разделе 7.

## 4. ИМИТАЦИОННАЯ МОДЕЛЬ ОСНОВНОГО СУБМОДУЛЯ

В этом разделе основная модель нейронной сети аналогична той, которая описана в разделе 2, но имеет некоторые отличия, частично обусловленные реализацией на ЭВМ, а частично связанные с уточнением структуры связей (по типу модели Изинга). Так, время в модели дискретно, множество внутренних состояний нейрона непрерывно, а выходные состояния сети принадлежат  $\{0, 1\}^N$ , где  $N$  — число нейронов в сети.

### 4.1. Машинный вариант базовой модели

1. Порог нейрона учитывает абсолютную и относительную рефрактерность и имеет вид, изображенный на рис. 4.1, причем везде скорость экспоненциального затухания  $\alpha = 0,3/\Delta t$ , а время абсолютной рефрактерности  $t_{\text{абс}} = \Delta t$ , где  $\Delta t$  — шаг дискретности по времени.

2. Каждый нейрон двусторонне связан с четырьмя соседними элементами, расположенными в узлах плоской решетки. Все веса связей одинаковы и равны амплитуде скачка  $a$ , на которую увеличивается мембранный потенциал данного нейрона при поступлении на него возбуждающего импульса от соседей. Границные условия периодические.

3. В отсутствие входных импульсов мембранный потенциал каждого нейрона затухает экспоненциально до потенциала покоя со скоростью  $\alpha = 0,3/\Delta t$ .

4. После каждого пересечения мембранным потенциалом порога нейрон посылает возбуждающий импульс на соседние нейроны, его мембранный потенциал становится равным потенциальному покоя. Порог после возбуждения принимает значение  $+\infty$ , последующее изменение порога показано на рис. 4.1 (при этом ноль на оси времени соответствует моменту возбуждения нейрона).

5. К описанному в п. 3 мембранныму потенциальному добавлен стационарный гауссовский шум с нулевым средним и дисперсией  $\sigma^2$ .

Модель имеет три основных параметра  $a$ ,  $\sigma$ ,  $r$ . Остальные параметры фиксированы, причем в большей части экспериментов  $r_\infty$  тоже фиксируется и приведенные значения  $a$  и  $\sigma$  нормированы на  $r_\infty$ .

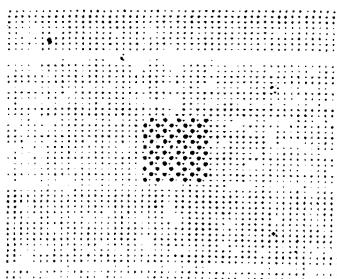
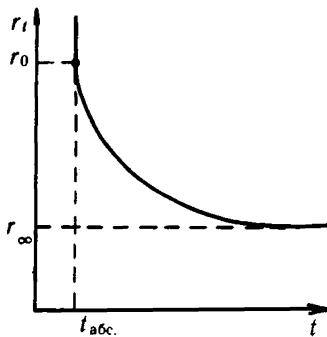


Рис. 4.1. Кривая порога

Рис. 4.2. Начальная конфигурация: • — состояние  $x_i^t = 0$ , \* — состояние  $x_i^t = 1$

Состоянием сети называется решетчатая конфигурация из нулей и единиц, соответствующих возбужденному ( $x_i^t = 1$ ) или невозбужденному ( $x_i^t = 0$ ) нейрону.

Уровнем активности сети будем называть величину

$$\bar{x}^t = \frac{1}{N} \sum_{i=1}^N x_i^t,$$

где  $N$  — число нейронов в сети. В наших экспериментах рассматривались сети, в которых было от 400 до 2500 модельных нейронов.

#### 4.2. Динамика уровня активности: выход на стационарный уровень

Исследование нашей модели мы начнем с анализа достаточно простой характеристики работы сети — уровня активности  $\bar{x}^t$ . В этом разделе мы рассмотрим реализации  $\bar{x}^t$  при различных значениях параметров модели  $a$  и  $\sigma$ , но при одной и той же начальной конфигурации (пятно из единиц на фоне нулей с  $\bar{x}^0 = 0,024$ , см. рис. 4.2). Задавая разные начальные значения датчика псевдослучайных чисел, мы получили разные реализации  $\bar{x}^t$  для фиксированных значений параметров сети и начальной конфигурации.

Все реализации  $\bar{x}^t$  ведут себя следующим образом: некоторое время после начала имеется тренд уровня активности, затем тренд исчезает и  $\bar{x}^t$  колеблется вокруг некоторого стационарного значения, зависящего от  $a$  и  $\sigma$  (рис. 4.3). Зависимость стационарного значения уровня активности от параметров модели показана на рис. 4.4.

Оказывается, время  $t$  выхода на стационарный уровень также существенно зависит от значений параметров сети (рис. 4.5).

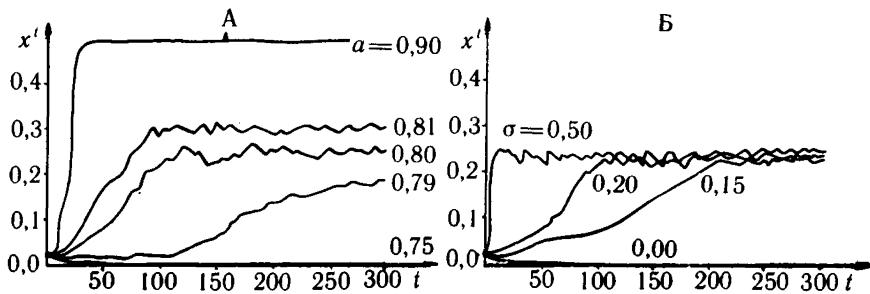


Рис. 4.3. Реализации  $\bar{x}'$ : А — при различных значениях  $a$  ( $\sigma=0,2$ ), Б — при различных  $\sigma$  ( $a=0,80$ ), (в обоих случаях  $r_0=30$ ,  $r_\infty=10$ )

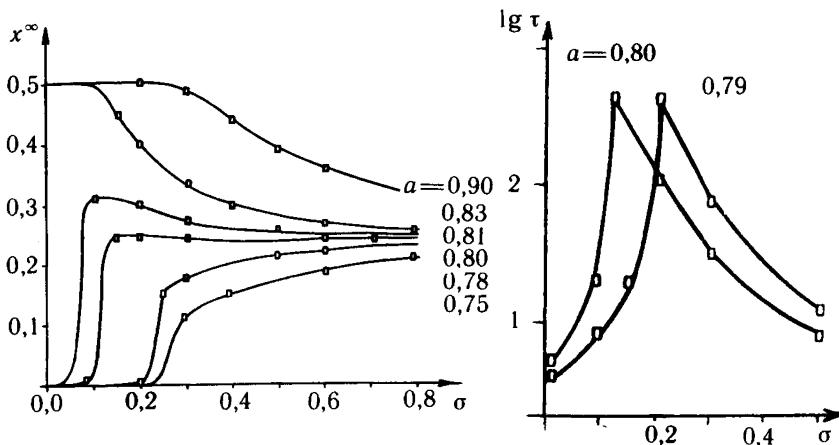


Рис. 4.4. Зависимость стационарного значения уровня активности от  $\sigma$  при разных значениях  $a$  (параметры порога как на рис. 4.3)

Рис. 4.5. Время выхода на стационарный уровень активности сети при изменении  $\sigma$  ( $a$  — фиксировано)

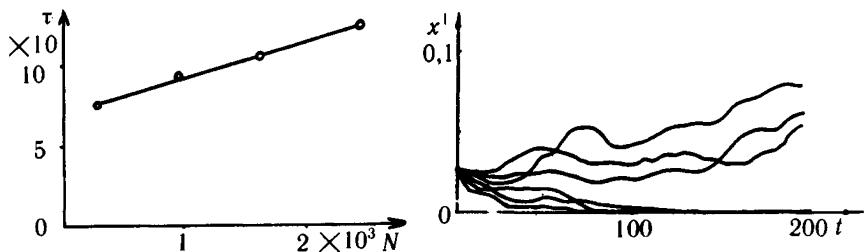


Рис. 4.6. Время выхода на стационарный уровень в зависимости от размеров сети ( $a=0,80$ ;  $\sigma=0,2$ )

Рис. 4.7. Начальные участки реализаций  $\bar{x}'$  при критических параметрах ( $a=0,79$ ;  $\sigma=0,2$ ) (параметры порога как на рис. 4.3)

При  $a=0,79$ ;  $\sigma=0,2$  наблюдается резкое увеличение  $t$ . Эти значения параметров  $a$  и  $\sigma$  будем называть критическими. Время выхода на стационарный уровень в критической ситуации возрастает с увеличением размеров сети (рис. 4.6). Для сети с критическими параметрами характерен еще один эффект. Разные реализации  $x^t$  сходятся не к одному (как в некритическом случае), а к двум различным стационарным значениям (рис. 4.7).

Еще одним весьма важным показателем динамики сети является дисперсия уровня активности. Она вычислялась по 10 реализациям по формуле:

$$D' = \frac{1}{10} \sum_{i=1}^{10} \left( \bar{x}_{(i)}^t - \frac{1}{10} \sum_{j=1}^{10} \bar{x}_{(j)}^t \right)^2, \quad 2 \cdot 10^{-2}$$

здесь  $i, j$  — номера реализаций.

Для  $a$  и  $\sigma$  вдали от критических значений  $D'$  вначале возрастает, затем по мере приближения  $\bar{x}^t$  к стационарному уровню  $D'$  убы-

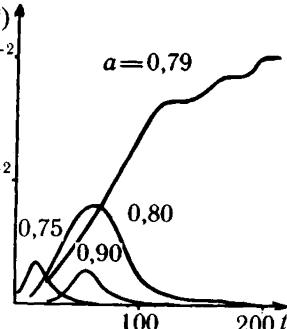


Рис. 4.8. Дисперсия уровня активности ( $\sigma=0,2$ )

вает; на стационарном участке дисперсия мала. Другая картина в критической ситуации. Здесь дисперсия возрастает, пока не будет достигнут стационарный уровень, и затем остается почти неизменной (рис. 4.8). Большая дисперсия на стационарном участке объясняется неединственностью стационарного значения уровня активности в критическом случае.

### 4.3. Эффект сохранения пятна

Рассмотрим теперь, как изменяется во времени состояние сети, если начальное состояние — пятно из единиц на фоне нулей (см. рис. 4.2). При определенных параметрах пятно расплывается (рис. 4.9А), и в стационарном режиме активность сети (в некотором диапазоне значений  $\sigma$ ) представляет собой кластерную конфигурацию, изменяющуюся во времени (пример таких состояний сети на рис. 4.11). В другом случае пятно размывается и исчезает, так что устанавливается конфигурация, состоящая в основном из нулей (рис. 4.9В). Как правило, размывание или расплывание пятна происходит довольно быстро ( $t \approx 50$  шагов времени, рис. 4.9А, В), однако сеть с критическими значениями параметров способна длительное время находиться в состоянии, близком к первоначальному (рис. 4.9Б), так сказать, сохранять пятно. Время сохранения пятна существенно зависит от первоначального размера пятна (рис. 4.10). Пятно меньше определенного размера быстро размывается.

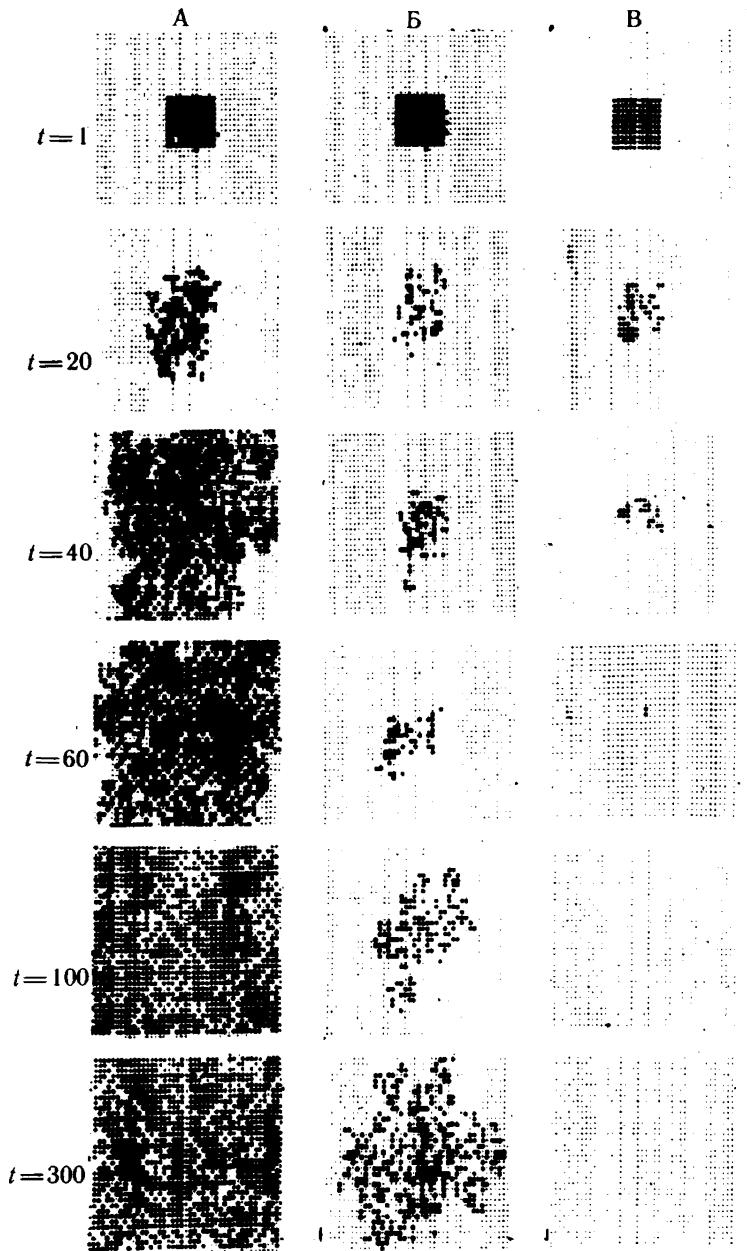


Рис. 4.9. Эффект сохранения пятна: А —  $a=0,90$ ; Б —  $a=0,79$ ; В —  $a=0,75$   
 $(\sigma=0,2; r_0=30, r_\infty=10)$

#### 4.4. Линии критического режима

В ходе эксперимента наблюдалась стационарная активность четырех типов. Оказалось, что каждому из них соответствует своя область на плоскости параметров системы (рис. 4.12). Для области I характерны упорядоченные стационарные состояния сети, при которых всякий элемент возбужден каждый второй такт времени ( $x^\infty = 0,5$ ). Если параметры сети принадлежат области II, то конечные стационарные состояния представляют собой конфигурации, состоящие целиком из

Рис. 4.10. Время сохранения пятна в зависимости от его первоначальных размеров и параметров сети ( $\sigma = 0,2$ )

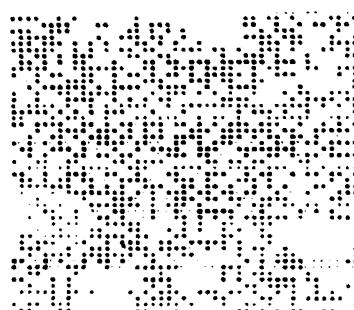
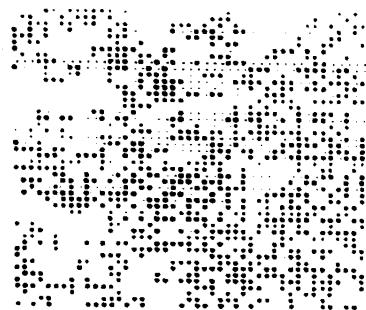
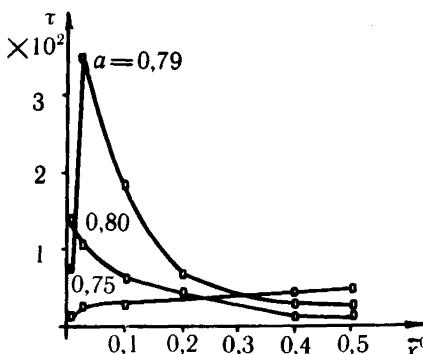
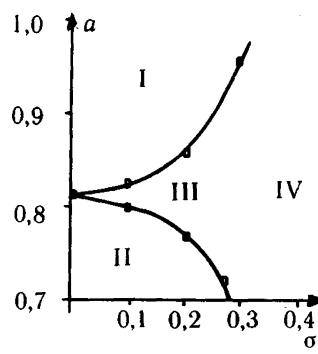


Рис. 4.11. Примеры кластерных конфигураций

нулей ( $x^\infty = 0$ ). Область III соответствует случаям кластерных конфигураций ( $0 < x^\infty < 0,5$ ), изменяющихся во времени (пример таких состояний сети на рис. 4.11). В области IV, соответствующей большим шумам, устанавливаются конфигурации из случайно разбросанных нулей и единиц.

Рис. 4.12. Линии критического режима ( $r_\infty = 10$ ,  $r_0 = 30$ )



#### 4.5. Периодическая модуляция порога

В этом разделе мы немного усложним модель — введем периодическое внешнее воздействие. Для этого к порогу каждого элемента сети добавим величину  $\Delta r^t$ , которая является периодической функцией времени  $\Delta r^t = A_\Theta \sin 2\lambda_\Theta t$ . Нас интересовал вопрос, как обеспечить в условиях такого внешнего воздействия длительное сохранение начальной конфигурации (эффект сохранения пятна). На рис. 4.13 показана зависимость врем-

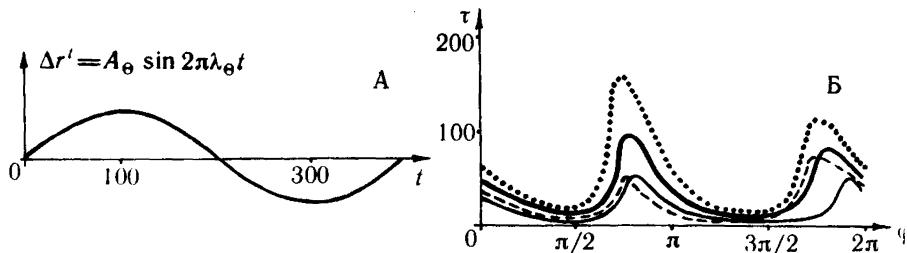


Рис. 4.13. Эффект периодической модуляции порога. А — зависимость приращения порога во времени; Б — зависимость времени сохранения пятна  $\tau$  от фазы  $\varphi_0$  подачи пятна на сеть: (—) —  $A_\Theta = 3,0, T_\Theta = 400$ ; (—) —  $A_\Theta = 1,5, T_\Theta = 400$ ; (---) —  $A_\Theta = 1,5, T_\Theta = 200$ ; (\*\*\*\*) —  $A_\Theta = 0,8, T_\Theta = 400$  ( $N = 2500$ ,  $r_\infty = 15, r_0 = 35, a = 0,56, \sigma = 0,1$ )

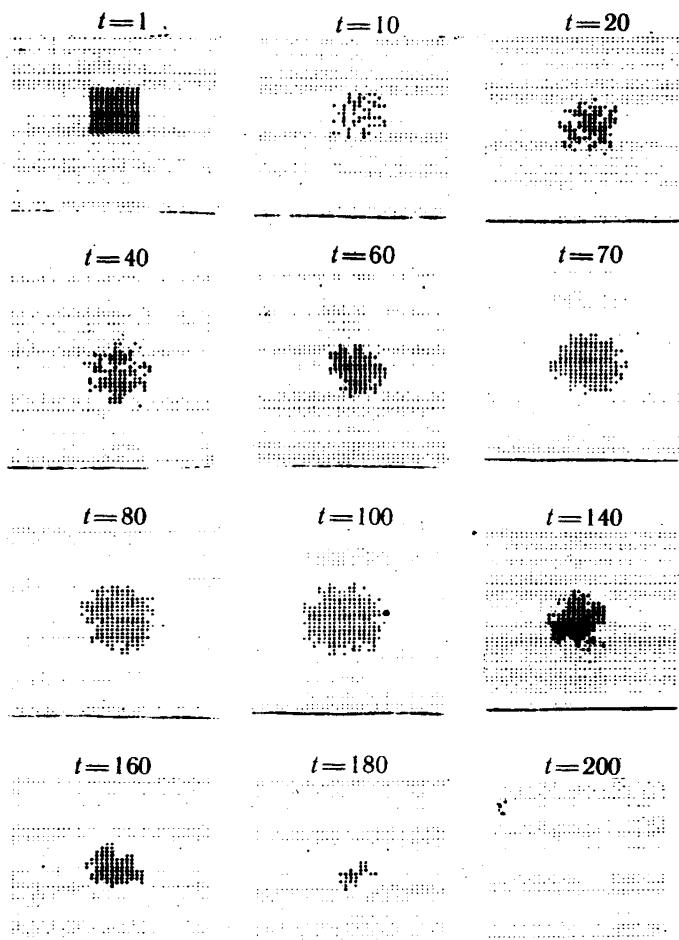
мени сохранения пятна от параметров периодического входа ( $A_\Theta, \lambda_\Theta, \varphi_0$ ), где  $\varphi_0$  — фаза подачи пятна на сеть. Оказалось, что для длительного удержания пятна необходимо выполнение следующих условий:

1. Параметры  $a$  и  $\sigma$  близки к критическим.
2. Период модуляции  $T_\Theta = 1/\lambda_\Theta$  близок к времени выхода на стационарный уровень в критическом случае.
3. Начальная фаза  $\varphi_0$  должна быть из определенного диапазона.

#### 4.6. Сеть с синаптической модификацией

Мы провели серию экспериментов, в которых веса связей между отдельными элементами могли изменяться в процессе работы модели. А именно, всякий раз, когда два соседних элемента возбуждались в один и тот же или в два последовательных момента времени, сила связи между ними возрастала на малую величину  $\Delta a \sim 0.001 \div 0.02$ .

На сеть с параметрами, близкими к критическим, подавалась начальная конфигурация в виде пятна из единиц на фоне нулей. Так как некоторое время сохраняется конфигурация, близкая к начальной, происходит постепенное увеличение связей



*Рис. 4.14. Сохранение пятна при модификации и нарастающем пороге ( $\sigma=0,2$ ;  $a=0,80$ ;  $\Delta a=0,002$ ;  $1/\lambda_a=400$ )*

между элементами пятна. В результате время сохранения пятна значительно возрастает, и состояние сети меньше отличается от начального, чем в случае сети без модификации связей (рис. 4.14).

Восстановление или «проявление» ранее записанного пятна происходит, во-первых, при установке параметров, близких к критическим (например, можно лишь слегка увеличить порог, сохранив остальные параметры теми же, что были при записи), и, во-вторых, при возбуждении части записанной конфигурации; последнее можно заменить подачей случайной конфигурации с равномерным распределением нулей и единиц.

## 4.7. Сеть с возвратным торможением

В этом пункте мы рассмотрим модификацию основной имитационной модели нейронной сети, способную генерировать колебания среднего уровня активности большого периода (по сравнению со временем релаксации отдельного элемента). В основу модели положен тот факт, что нейронная сеть с локальными связями, рассмотренная выше, в области критических значений параметров имеет большое время выхода суммарной активности на стационарный уровень. Модифицированная модель отличается от первоначальной тем, что вводится один тормозной элемент, который осуществляет функции возвратного торможения. На рис. 4.15 представлена схема соединения элементов модели. Сеть состоит из 400 возбуждающих нейронов, расположенных в узлах плоской решетки. Каждый из них соединен с соседями положительными связями одинаковой силы  $a_{ee} > 0$ . Кроме того, вводится один тормозной нейрон, который получает от каждого возбуждающего элемента положительную связь  $a_{ei} > 0$  и, в свою очередь, посылает на все элементы отрицательные связи одинаковой силы  $a_{ie} < 0$ .

Основными переменными параметрами модели являются значения порогов элементов и сила связи тормозного элемента с возбуждающими элементами (модельный аналог величины ТПСП).

При исследовании такой модельной сети были получены следующие результаты.

1. При достаточно высоких значениях порогов возбуждающих элементов  $r_e^e$  тормозный нейрон не срабатывает и колебания не возникают (рис. 4.16А;  $r_e^e = 13$ ). Подробное исследование такого режима нерегулярной одиночной и пачечной активности приведено по существу в предыдущих пунктах.

2. При достаточно большой интенсивности входного потока или (эквивалентно) при малых значениях порога суммарная активность нейронной сети периодически изменяется со временем (см. рис. 4.16А,Б), а импульсная активность отдельных нейронов приобретает пачечный характер с заметной регулярностью (рис. 4.16В).

При этом пачки отдельных элементов могут быть сдвинуты по фазе относительно друг друга, но хорошо коррелируют с колебаниями среднего уровня. Тормозной нейрон генерирует более высокорегулярные последовательности одиночных импульсов (рис. 4.16В).

3. При специальном подборе параметров модели средний период колебаний суммарной активности

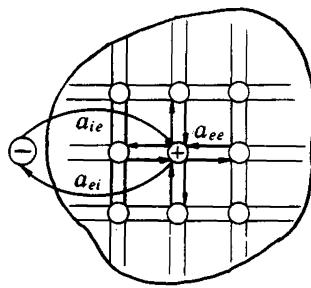


Рис. 4.15. Схема соединения элементов в сети

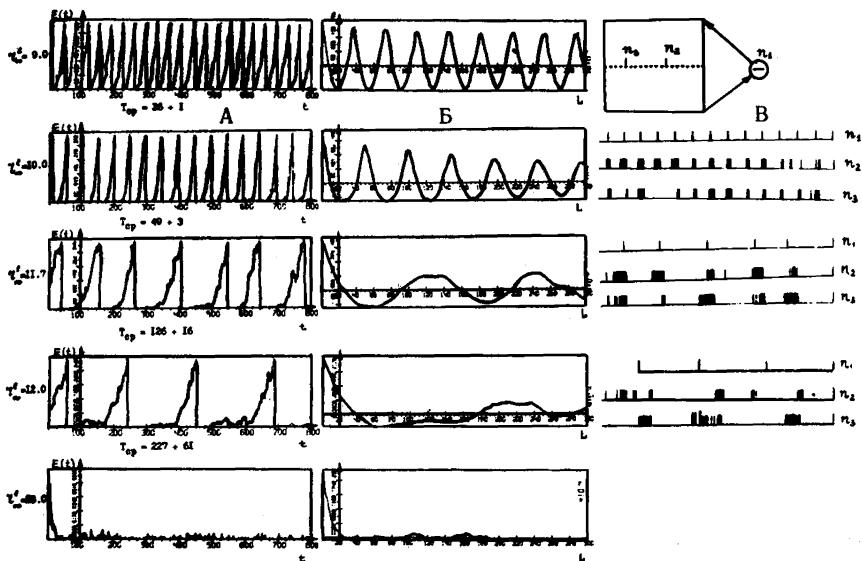
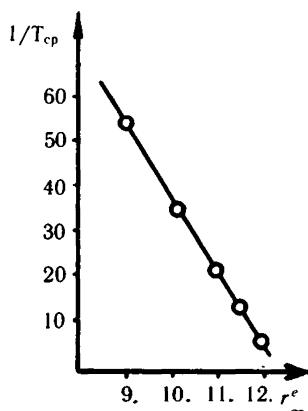


Рис. 4.16. Динамическое поведение модели при различных значениях порогов возбуждающих элементов: А — средний уровень активности сети; Б — автокорреляционные функции среднего уровня активности; В — реализации импульсных последовательностей тормозного ( $n_1$ ) и возбуждающих ( $n_2$ ,  $n_3$ ) элементов ( $N=400$ ;  $a_{ee}=8,6$ ;  $\sigma=3,0$ ;  $\tau^e=\tau_{PSP}^e=3,0$ ;  $a_{ei}=0,08$ ;  $a_{ie}=20,0$ ;  $\tau^i=\tau_{PSP}^i=2,0$ ;  $\tau_{TPSP}=10,0$ ;  $r_\infty^i=25,0$ )

Рис. 4.17. График зависимости средней частоты колебаний суммарной активности от значений порогов возбуждающих элементов

может на порядок превосходить средний межимпульсный интервал одиночного нейрона. Широкая регулировка частоты достигается изменением  $r_\infty^e$  (рис. 4.17), что эквивалентно изменению интенсивности входного потока. Из рисунка видно, что в рассмотренном диапазоне частот зависимость близка к линейной. Для отдельных элементов рост интенсивности входного потока сопровождается увеличением частоты пачек и уменьшением их длины. При сильном снижении  $r_\infty^e$  колебания срываются. Минимальный период определяется постоянной времени затухания тормозного воздействия.



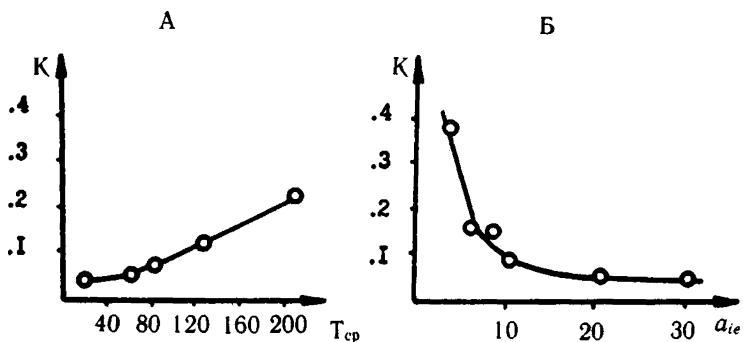


Рис. 4.18. График зависимости коэффициента вариации периода колебаний суммарной активности от среднего периода колебаний (А) и величины модельного ТПСП (Б)

4. Коэффициент вариации периода колебаний суммарной активности (отношение среднеквадратичного отклонения к среднему периоду) может достигать нескольких процентов, однако возрастает на низкочастотном конце диапазона (рис. 4.18А). Стабильность колебаний может быть повышенена введением специальной пространственной неоднородности сети, например, путем исключения тормозного воздействия на некоторую фиксированную часть нейронов.

## 5. ПРИМЕНЕНИЕ К ГИППОКАМПУ: МОДЕЛЬ ПРИВЫКАНИЯ

Самое замечательное и самое загадочное явление в электрофизиологии гиппокампа — привыканье. Оно состоит в том, что нейронные реакции, возникающие в ответ на широкий класс стимулов, имеют необычные свойства:

1. Длительность реакций может значительно превосходить максимальную длительность постсинаптических потенциалов и достигает 10—15 с.

2. При повторении одних и тех же стимулов длительность реакций постепенно укорачивается вплоть до полного угашения.

3. Угашенные реакции полностью восстанавливаются при изменении любых параметров стимула на весьма малую величину, близкую к дифференциальному порогу чувствительности.

В настоящее время существует три основных модели привыкания. Модель Groves, Thomson (1970) основана на теории моносинаптической депрессии (истощение медиатора). Она является вариантом более ранних моделей привыкания беспозвоночных и неспособна объяснить третье свойство при внезапном уменьшении интенсивности стимула. Модель нервного стимула Соколова (1960) свободна от этого недостатка, так как постулирует запоминание детальной информации о свойствах стимула для сравнения с новыми стимулами, но не дает конкретных нейронных механизмов. В концепции лимбической системы Виноградовой (1975, с. 266—268) уточняется, что подобное сравнение стимулов и генерация угасающих реакций осуществляется в гиппокампе, но при этом возникает ряд трудностей, описанных ниже.

Заметим, что ни одна из моделей не объясняет свойства 1, считая его, по-видимому, очевидным. Предлагаемое ниже обсуждение основывается на принципиальной, ведущей роли этого свойства по отношению к остальным. В то же время наше объяснение наиболее близко к третьей модели, которую мы опишем несколько подробнее.

Основная концепция Виноградовой состоит в том, что поле CA<sub>3</sub> гиппокампа — это нейронный компаратор двух сигналов, выделяющий не конкретные информационные свойства сигналов, а относительное качество новизны. Новизна оценивается

как состояние, при котором сигнал, поступающий по одному (ретикуло-септальному) входу, не находит своего аналога в другом (кортикальном) входе. В случае такого рассогласования возникают реакции нейронов поля СА<sub>3</sub>. Совпадение сигналов в обоих входах приводит к восстановлению устойчивого (фонового) состояния системы. Во время действия нового сигнала большинство нейронов гиппокампа переходит в тормозное состояние, меньшинство — активируется, что может приводить к снижению суммарного сигнала, выходящего на ретикулярную формацию. Поскольку в опытах подтверждено, что гиппокамп, действительно, может оказывать тоническое тормозное влияние на активирующую ретикулярную формацию, то снижение его суммарного сигнала должно обеспечивать повышение активности, восходящей от ретикулярной формации, что приводит к общему повышению рабочего уровня мозга. По мере повторения раздражителя и его частичного узнавания реакции гиппокампальных нейронов будут блокироваться все раньше, и это будет соответствовать известному феномену постепенного укорочения общей активационной реакции до полного прекращения. На этом этапе поддержание высокого уровня активации мозга биологически нецелесообразно, поскольку анализ раздражителя и фиксация соответствующего следа завершены.

Однако многие детали в этой модели остаются не вполне ясными. В частности, неясно каков конкретный механизм градуального укорочения тонических реакций. Большие осложнения также вызывает следующий факт, существенно дополняющий три описанных выше свойства привыкания.

4. Сохранность кортикального входа к полю СА<sub>3</sub> необходима для угашения реакций, и в то же время потенциация синапсов зубчатой фасции (структуры, стоящей на входе гиппокампа) приводит к блокаде тонических ответов пирамидных нейронов поля СА<sub>3</sub> на новые сенсорные раздражители, хотя хорошо известно, что эти синапсы возбудительные [Виноградова, 1975, с. 265—269].

Необходимо учесть еще один существенный факт, который тоже затрудняет объяснение привыкания.

5. Почти все нейроны поля СА<sub>3</sub> реагируют на стимулы любой модальности [Виноградова, 1975, с. 236].

Этот факт делает маловероятным существование специфической памяти о стимулах в гиппокампе, а без этого неясно, как возникает высокая чувствительность восстановления угасших реакций к малым изменениям параметров стимулов.

Правильное понимание привыкания, по-видимому, невозможно без учета влияния тех структур, с которыми гиппокамп тесно связан, особенно с септумом, ретикулярной формацией и неокортексом. Однако, чтобы подчеркнуть основную идею, мы временно упростим реальность, ограничившись процессами, происходящими в основном поле СА<sub>3</sub> гиппокампа. Необ-

**ходимые уточнения и развитие модели будут приведены позже.**

Итак, предположим, что сеть, описанная в разделе 4, имитирует ту часть пирамидных нейронов поля СА<sub>3</sub>, которая содержится в одном сегменте гиппокампа, причем кортикальный вход (через клетки зубчатой фасции) формирует пятно активных нейронов, а ретикулосепタルный вход (через баскетные клетки) модулирует пороги нейронов. Тогда из приведенных результатов вытекает следующая интерпретация привыкания (порядок обсуждаемых вопросов сохранен прежним).

1. *Длительные тонические реакции пирамидных нейронов поля СА<sub>3</sub>* — естественное следствие достаточно сильных (критических) связей между нейронами (см. рис. 4.3; 4.5; 4.6). О том, что межнейронные связи поля СА<sub>3</sub> достаточно сильные, косвенно свидетельствует высокий уровень стационарной фоновой активности (есть и прямые данные): а ее нерегулярный залповый характер тоже косвенно свидетельствует о существовании критического режима. Чтобы возник длительный переходный тонический процесс, необходимо вначале генерализованно погасить фоновую активность, установив с помощью септального входа необходимую «тишину». После этого сеть готова воспринять слабый кортикальный сигнал и очень долго его удерживать даже после того, как сам этот сигнал исчезнет с входа СА<sub>3</sub>. Установленное сигналом «пятно» постепенно «расплывается» и сеть возвращается в исходное состояние высокой фоновой активности. С приходом следующего стимула реакция повторяется без изменений.

2. Предположим теперь, что один из параметров модели может постепенно меняться под действием входных сигналов, например в результате длительной потенциации синапсов. Тогда пятно в соответствии с графиками, показанными на рис. 4.5; 4.10, расплывается тем быстрее, чем дальше уйдут параметры от критических значений. Длительные тонические реакции будут постепенно укорачиваться до фазных. *Градуальность привыкания — непосредственное проявление постепенного нарушения критического режима под действием входных сигналов.*

3. Для возобновления угасших тонических реакций необходимо каким-то образом восстановить критический режим в сети. Однако быстро изменить параметры сети нельзя, поэтому остается варьировать только начальные условия пятна. Можно, например, сформировать новое пятно в другом участке сети, где соответствующие синапсы оказались еще непотенцированными. Однако в нашей модели восстановление реакций (а также и их угашение) можно получить другим, несколько неожиданным способом — путем сравнительно небольших вариаций запаздывания кортикального сигнала по отношению к септальному.

Действительно, учтем, что септальный выход при действии нового раздражителя обычно имеет пачечный тета-ритмический

характер, позволяющий, кроме первоначального внезапного торможения пирамид СА<sub>3</sub>, осуществить необходимую нам периодическую модуляцию их порогов. Тогда в соответствии с рис. 4.13 можно сформировать долгоживущее пятно только в том случае, если кортикальный сигнал тоже имеет тета-ритмическую компоненту и ее фаза примерно на 110—120° отстает от фазы септального тета-ритма. При неизменных параметрах пятно, пульсируя с частотой тета-ритма, будет «сохраняться» до тех пор, пока не будет нарушена синхронность и синфазность между двумя входными сигналами. Следовательно, необходимо предположить, во-первых, что «кортикальный тета-ритм» и «септальный тета-ритм» имеют общий источник возбуждения (вероятно, ретикулярная формация, активность которой, как известно, модулируется гиппокампом) и, во-вторых, в головном мозге должна существовать фазовая следящая система, поддерживающая автоматически нужную синхронность и синфазность в течение длительного времени, пока стимул привлекает к себе внимание животного. Постепенное нарушение режима слежения (вследствие выхода сигнала из диапазона слежения или ухода параметров) обеспечит постепенное угашение реакций. Новый стимул, захватывая «внимание» следящей системы и обеспечивая необходимый критический сдвиг фаз, вызовет возобновление угасших реакций. Тем самым мы признаем, что наша модель поля СА<sub>3</sub> неспособна полностью объяснить явление привыкания, имеющее, несомненно, системный характер. Зато наша модель позволяет и даже вынуждает высказать гипотезу о существовании в мозге функциональной системы регистрации информации и внимания, основанной на фазовой автоподстройке частоты тета-ритма, причем поле СА<sub>3</sub> выполняет в ней ключевую функцию фазового дискриминатора, выделяющего сигнал ошибки рассогласования. Эта функция по существу та же самая, что в модели Виноградовой: поле СА<sub>3</sub> выделяет «новизну», пропорциональную разности фаз, а если она очень велика (один из сигналов может отсутствовать), сигнализирует о необходимости «поиска» сигнала.

Более детальное описание такой следящей системы выходит за рамки этого раздела. Отметим лишь, что структура и «нейронное обеспечение» такой фазовой следящей системы в принципе известны [Крюков, 1982] и хорошо согласуются со схемой основных взаимодействий между структурами лимбической системы, описанной в книге Виноградовой (1975).

4. Из сказанного очевидно, что сохранность кортикального входа необходима для привыкания и восстановления реакций, причем потенциация синапсов зубчатой фасции может приводить к постепенному нарушению критического сдвига фаз, укорочению реакций и, наконец, к полной блокаде тонических реакций поля СА<sub>3</sub>.

5. Приведенное объяснение привыкания не требует (хотя и не исключает) запоминания детальной информации о стимуле

в самом гиппокампе. Нам достаточно лишь того, чтобы разные новые стимулы обеспечивали нужную синхронность и синфазность кортикального сигнала по отношению к септальному. Это требование легко выполняется с помощью указанной следящей системы, если регистрация специфической информации о стимуле осуществляется вне гиппокампа, причем время распространения нервных импульсов до места хранения и обратно равно (или кратно) периоду тета-ритма. Поскольку привыкание носит универсальный характер, то есть не зависит от модальности стимула, то для поля СА<sub>3</sub> как фазового компаратора специфичность сигналов не имеет значения, и поэтому она «теряется» при смешивании сигналов в зубчатой фасции. Однако, если необходимо, она может быть восстановлена вновь с помощью синхронной селекции на выходе наподобие технических систем связи с временным разделением компонент. Не исключено, что функцию такого селектора выполняет поле СА<sub>1</sub>, где, как известно, регистрируется значительная пачечная тета-ритмическая активность.

Резюмируя, можно сказать, что поле СА<sub>3</sub>, по-видимому, не является единственным источником угасающих реакций, но, несомненно, является одним из главных звеньев системы привыкания. Обнаруженный в модели критический режим и его фазочастотная зависимость позволяет уточнить функцию поля СА<sub>3</sub> как фазового компаратора, выделяющего сигнал ошибки рассогласования двух входных сигналов. Одновременно с помощью эффекта сохранения пятна оно выполняет другую важную функцию — фиксацию сигнала ошибки в течение значительной части периода тета-ритма. Иначе сигнал ошибки не подействует на последующие инерционные звенья системы регулирования частоты генератора тета-ритма.

В заключение отметим, что приведенное объяснение привыкания оказывается неразрывно связано с основным принципом работы мозга — принципом доминанты и усвоения ритма А.А.Ухтомского. Эта связь, здесь только намеченная, будет постепенно проясняться при моделировании других отделов и звеньев функциональной системы и, в частности, при моделировании септума, ретикуло-септогиппокампальной системы, неокортекса и, наконец, всей системы в целом.

## **6. ПРИМЕНЕНИЕ К СЕПТУМУ: МОДЕЛЬ ГЛОБАЛЬНОГО ПЕЙСМЕКЕРА**

Многие нейрофизиологические исследования показывают, что в головном мозге человека и животных существует пейсмекер, то есть специальный нейронный генератор низкочастотной ритмической активности, который, по-видимому, задает единый ритм обработки сенсорной информации путем синхронизации различных структур мозга на частоте пейсмекера. Одновременно с помощью этого ритма может осуществляться дискретизация по времени входных сигналов, что увеличивает пропускную способность, помехозащищенность и надежность системы. Частота пейсмекера, по-видимому, автоматически регулируется в зависимости от текущего состояния организма, и разным значениям частоты соответствуют разные конstellации совозбужденных центров, усвоивших этот ритм.

Среди всех ритмических процессов, возникающих в ЦНС, особое место занимает тета-ритм как наиболее вероятный кандидат на глобальный пейсмекерный ритм. Тета-ритм — это регулярные, высокоамплитудные колебания ЭЭГ в диапазоне 3—6 Гц (иногда до 10—12 Гц), наиболее четко регистрируемые в гиппокампе, хотя сам пейсмекер, по-видимому, локализован не в гиппокампе или не только в нем. Самая яркая особенность тета-ритма, привлекающая к себе огромное внимание исследователей, — возможность четкого выделения его в различных поведенческих ситуациях. Именно «поведенческие корреляты» тета-ритма дают серьезные основания гипотезе о его системообразующей функции глобального синхронизатора различных отделов мозга. Поэтому вопрос о происхождении и локализации первичного источника тета-ритма должен помочь в прояснении основных механизмов работы мозга.

В настоящее время по мнению многих нейрофизиологов источник тета-ритма локализован в медиальном ядре септума. Импульсная активность этой структуры обладает следующими замечательными свойствами [Виноградова, 1975, с. 166—167]:

1. Многие клетки этой структуры при определенном уровне входной аfferентации начинают генерировать регулярную пачечную активность. Однако таких клеток становится значительно меньше при достаточно высоком уровне аfferентации.

2. Частота следования пачек коррелирует с частотой тета-ритма ЭЭГ, но пачки отдельных нейронов сдвинуты друг относительно друга так, что каждый нейрон имеет свою жесткую предпочтительную фазу тета-ритма, в которой возникает его пачка.

3. Вариация межпачечных интервалов невелика и уменьшается на высокочастотном конце диапазона.

4. Частота следования пачек изменяется в значительных пределах при изменении уровня афферентации. Имеется участок линейной зависимости, соответствующий рабочему диапазону частот тета-ритма.

Из этих данных следует, что период колебаний предполагаемого пейсмекера должен быть по крайней мере на порядок выше характерной постоянной времени мембранны нейрона, а его стабильность должна быть примерно на порядок выше, чем у обычных нейронов. Это очень жесткое ограничение осложняется тем, что частота генератора должна линейно регулироваться уровнем входной афферентации, причем на краях диапазона колебания должны срываться, превращаясь в хаотическую активность.

Существует два класса объяснений этих фактов: 1) модель типа одиночного нейрона-пейсмекера, который после внезапного ударного возбуждения генерирует низкочастотную пачечную активность стабильной частоты и синхронизирует основную массу нейронов септума, выполняющих функцию усиления мощности колебаний [Виноградова, Золотухина, 1972]; 2) модель сети с возвратным торможением, в которой пачечная активность одиночных нейронов априорно не предполагается, но возникает как сетевой эффект за счет периодического низкочастотного «вытурмаживания» высокочастотной активности основных нейронов с помощью выделенного тормозного интернейрона [Tömböl, Petsche, 1969].

Первая модель, относя все указанные трудности за счет одиночных пейсмекерных нейронов, все же должна еще объяснить, как в ней совмещены свойства 3 и 4, поскольку конфликт между стабильностью частоты и возможностью ее регулирования, как хорошо известно из техники, неизбежно ведет к большим усложнениям схемы генератора. Вторая модель, свободная от этого недостатка, требует, однако, существования в септуме тормозных интернейронов с постоянной времени ТПСП порядка максимального периода тета-ритма ( $\sim 350$  мс). Столь большая инерционность может вызвать сомнение, хотя существование самих интернейронов не противоречит морфологическим данным [Tömböl, Petsche, 1969].

Предлагаемая модель септального генератора является, по существу, усовершенствованием второй модели с целью увеличения максимального периода генерируемых колебаний без предположения о существовании длительных ТПСП. Это достигается установкой сети в критический режим путем изме-

нения параметров модели, что реализуется в ЦНС, например за счет изменения интенсивности входного потока импульсов. Одновременно генератор приобретает ряд важных новых свойств, описанных ниже.

Итак, предположим, что основные клетки септума локально связаны с помощью возбуждающих связей и существует сравнительно небольшое число тормозных интернейронов, получающих возбуждение от большого числа основных нейронов в сети и, в свою очередь, их тормозящие, так что активность локали (области влияния одного интернейрона) описывается моделью, представленной в разделе 4. Предположим также, что локали различных интернейронов работают независимо, но все получают один и тот же управляющий афферентный приток.

Тогда, используя полученные результаты имитационного моделирования, легко объяснить приведенные выше факты 1—4.

1. При низком уровне афферентации или (эквивалентно) при высоких порогах нейронов периодические колебания либо вообще отсутствуют, либо слабо регулярны (см. рис. 4.16Б;  $r^e = 12$ ). Регулярные колебания возникают лишь при умеренном входном уровне (см. рис. 4.16Б;  $8 < r^e < 12$ ), причем частота колебаний линейно растет с ростом уровня (см. рис. 4.17) и при этом растет регулярность колебаний (см. рис. 4.18), что соответствует аналогичному свойству тета-ритма. При дальнейшем увеличении входного уровня можно наблюдать интересное явление: колебания среднего уровня активности и пачечность у большинства нейронов исчезают, хотя небольшое число нейронов сохраняет очень регулярную пачечную активность. Этот режим, очевидно, соответствует исчезанию тета-ритма в гиппокампе.

2. Одной из важных особенностей модели, помимо возможности получения длительных межпачечных интервалов, является ее способность объяснить свойство 2, до сих пор еще не объясненное, насколько нам известно. Предварительно заметим, что эффект замедления динамики среднего уровня активности, описанный в разделе 4 для сети без обратной связи, имеет место и при наличии тормозной обратной связи. Постепенное заполнение локали горячими элементами (после фазы молчания) приводит к тому, что каждый нейрон локали начнет генерацию очередной пачки лишь тогда, когда граница повышенной активности включит его в свою зону. Это и означает, что каждый нейрон будет иметь свою предпочтительную фазу появления пачек по отношению к тета-ритму (независимо от частоты) и пачки разных нейронов будут иметь взаимные сдвиги.

3. Свойство 3 — основное сетевое свойство ячейки с возвратным торможением. Оно возникает за счет того, что суммарный мембранный потенциал тормозного интернейрона, имеющего  $N$  входов, имеет более регулярный характер, чем каждая из его компонент: среднеквадратичное отклонение растет

примерно пропорционально  $\sqrt{N}$ , а среднее — пропорционально  $N$ . Критический режим, внося взаимную зависимость между компонентами, уменьшает указанную регулярность, что и является причиной повышения коэффициента вариации с ростом периода колебаний (см. рис. 4.18).

4. Свойство 4 и, в частности, довольно хорошая линейность регулировки частоты тоже присуще обычной модели с возвратным торможением. В нашей модели нелинейное возрастание периода колебаний вблизи критической точки (см. рис. 4.17) способствует расширению линейности регулировки в сторону более низких частот колебаний. Однако в непосредственной близости от критической точки небольшие изменения параметров приводят к значительным изменениям периода и могут даже полностью сорвать колебания, переводя систему в обычный режим стохастической активности. Мы получим, таким образом, новое, пятое свойство генератора, которого не было у обычной модели с возвратным торможением, а именно:

5. Однаковые по величине изменения параметров модели могут приводить к резко различным изменениям поведения генератора на высокочастотном и низкочастотном концах диапазона. Так, изменение силы связей на 1—2%, совершенно незаметное на высоких частотах, может полностью срывать колебания на низких частотах. Это свойство, по-видимому, соответствует известному факту сильной чувствительности (вплоть до блокады) низкочастотного тета-ритма атропином (антихолинергический агент) и независимости от него высокочастотного (7—12 Гц) тета-ритма [Vanderwolf, Robinson, 1981]. Трудность обычной интерпретации этого факта и была, по-видимому, причиной появления очень спорной гипотезы Вандервольфа о существовании в ЦНС двух типов тета-ритма с различной локализацией соответствующих источников и с различными путями их воздействия на архи- и неокортекс. Наша модель свидетельствует в пользу представления об унитарности источника тета-ритма (см. обсуждение работы [Vanderwolf, Robinson, 1981] и, в частности, критику Vinogradova, Vertes, Shiromani, Fishbein).

6. Следует отметить один существенный недостаток модели: длительность непрерывной генерации тета-ритма иногда не пре-восходит 5—10 периодов ритма. Это возникает из-за того, что неизбежные небольшие различия в частотах различных локалей, постепенно накапливаясь, приводят после синхронизирующего залпа к рассогласованию и десинхронизации, если интервал между стимулами превышает время затухания автокорреляций в одной локали (рис. 4.16Б). Таким образом, наша модель требует внутренней синхронизации локалей генератора через цепи обратной связи с задержкой на период тета-ритма. Для этого морфологически существует несколько возможностей. Замыкание через гиппокамп уже использовано в одной из моделей [McLennan, Miller, 1974, 1976], а замыкание через нео-

кортикс и гиппокамп — в работе Виноградовой (1975). В первом случае нужная нам задержка может быть получена за счет длительного переходного процесса в критическом режиме гиппокампа, а во втором — за счет времени распространения импульсов до «цели» в неокортикс и обратно; в обоих случаях задержка должна быть кратна периоду тета-ритма. И снова, как в разделе 5, мы вынуждены постулировать замкнутую систему, в которой септальный генератор выполняет роль управляемого генератора ритма. Широкий диапазон, хорошая линейность и чувствительность при изменении управляющих параметров, а также стабильность периода колебаний делают его идеальным интегрирующим звеном системы автоматического регулирования.

## **7. ПРИМЕНЕНИЕ К МОЗЖЕЧКУ: МОДЕЛЬ ВЗАИМОДЕЙСТВИЯ ЛИАННЫХ И МШИСТЫХ ВОЛОКОН**

Основная проблема мозжечка согласно Экклсу (1977) состоит в следующем. Два основных входа в мозжечок, обраzuемые лианными и мшистыми волокнами, соответственно резко различаются по морфологическим и функциональным показателям и в то же время несут мозжечку тем не менее почти одну и ту же информацию из церебрального кортекса и с периферии. Эта биологическая конструкция, дублирующая входы на каждую клетку Пуркинье мозжечка, была создана у самых примитивных животных и затем сохранена на всех последующих стадиях эволюции. В ней, по-видимому, заложен глубокий смысл, отражающий какой-то весьма важный принцип ЦНС.

Решению этой проблемы посвящено огромное количество теорий функций мозжечка, среди которых наиболее развитой и фундаментальной является теория обучения Бриндли-Марра [Магг, 1969]. Согласно этой теории активность лианного входа «инструктирует» каждую клетку Пуркинье о том, какие ее синапсы от параллельных волокон должны быть долговременно модифицированы при совпадении во времени активностей обоих входов. Тем самым мозжечок оказывается местом хранения долговременных следов заученных движений. Теория неоднократно подтверждалась многими исследованиями на крысах, кроликах и кошках, но на приматах столкнулась с рядом серьезных трудностей [Miles et al., 1980]. Для разрешения последних пока предложен лишь один выход [Lisberger, 1982]: допустить, что у различных видов животных лианные волокна могут выполнять существенно различающиеся функции.

Мы считаем, что такое решение равносильно отказу от первоначальной проблемы и предлагаем вернуться к анализу основных экспериментальных данных, связанных со взаимодействием активности лианных и мшистых волокон, без учета которых невозможна современная теория функций мозжечка. Из большого количества таких данных мы выберем лишь наиболее трудные для интерпретации в терминах существующих моделей. Перечислим эти данные.

1. Химическое разрушение нижних олив — основного и, возможно, единственного источника активности лианных волокон

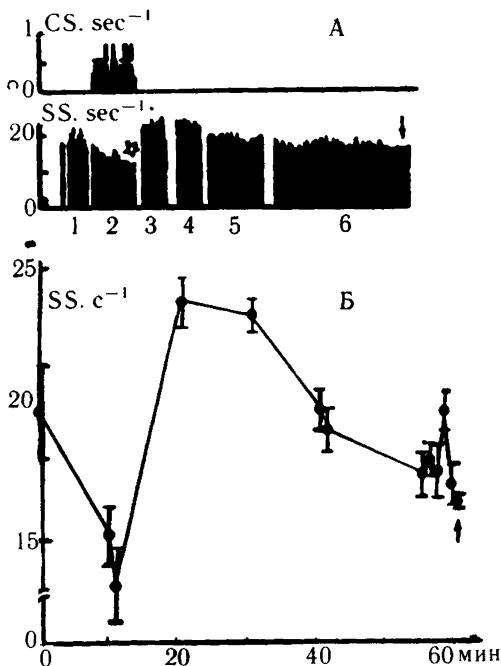


Рис. 7.1. Скачкообразные («фазовые 1-го рода») переходы активности клеток Пуркинье мозжечка после химической инактивации нижних олив [Colin et al., 1980]. А — гистограмма текущей частоты входных (комплексных) и выходных (простых) спайков; Б — график текущего среднего и дисперсии выходной частоты. Ширина бин на 10 мин

почти вдвое увеличивает фоновую частоту простых (не комплексных) спайков клеток Пуркинье, их регулярность, а также приводит к появлению спонтанно возникающих пауз случайной длительности, в течение которых спайковая активность клетки не регистрируется (рис. 7.1). Электрической стимуляцией лиановых волокон можно снизить высокую активность клеток Пуркинье почти до полного подавления. После снижения интенсивности стимуляции высокий уровень активности восстанавливается не сразу, а как бы в две стадии: вначале активность нарастает до 75% от максимальной в течение 9 с, а затем она подрастает еще в течение 10 мин, причем снова ее рост сопровождается увеличением регулярности импульсных последовательностей. График перехода из состояния с высокой частотой спайков в состояние с низкой частотой и обратно в зависимости от интенсивности стимуляции (см. рис. 7.2) имеет вид петли гистерезиса [Colin et al., 1980].

2. Автокорреляционная функция простых спайков у 50% обследованных клеток Пуркинье мозжечка ненаркотизированных десеребрированных кошек имеет в фоне необычно большое время затухания, превышающее 50 мс. Электрическая импульсная стимуляция нижних олив на частоте 2—5 Гц приводит к продлению затухания автокорреляционной функции, причем, как показано на рис. 7.3, автокорреляционная функция с малым временем затухания (до 10 мс) может трансформироваться в автокорреляционную функцию с большим временем затухания (больше 50 мс). Но скорость срабатывания клеток Пуркинье после стимуляции не только не возрастает, но даже несколько снижается [Ebner, Bloedel, 1981а].

3. Кросскорреляционная функция простых спайков в фоне

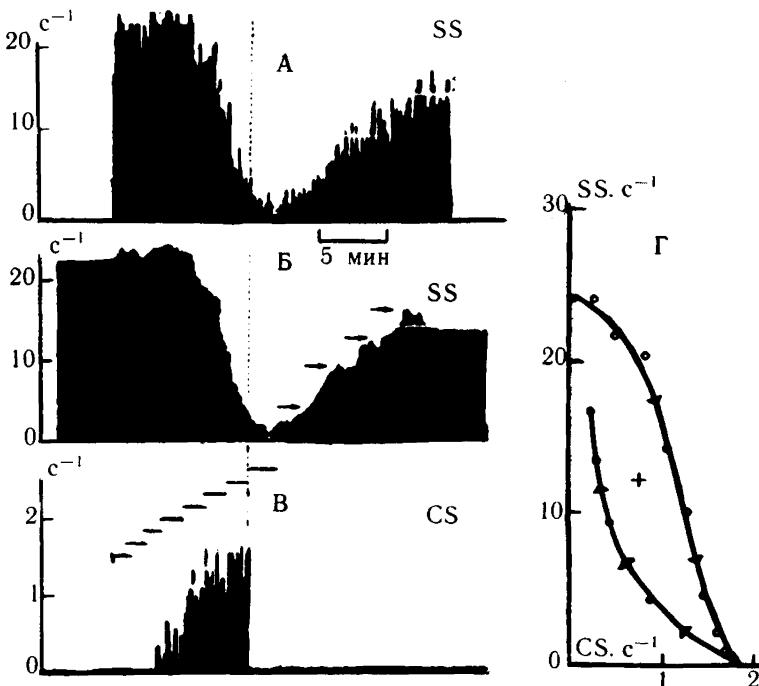


Рис. 7.2. Неравновесные («гистерезисные») переходы активности клеток Пуркинье мозжечка при стимуляции через лианное волокно после химической инактивации нижних олив [Colin et al., 1980]. А — гистограмма текущей частоты выходных (простых) спайков; Б — то же после фильтрации; В — гистограмма текущей частоты входных (комплексных) спайков; Г — гистерезисная кривая зависимости выходной частоты от входной. Ширина бина гистограммы 9,6 с

у 25% изученных пар клеток Пуркинье, отстоящих друг от друга на расстоянии до 750 мкм, имеет симметричную относительно нуля форму кросскорреляционной функции с необычно большим временем затухания (100—500 мс) как положительных, так и отрицательных кросскорреляций. Быстро затухающие положительные коротколатентные (с максимумом слева или справа от нуля) кросскорреляции встречаются только у тех клеток Пуркинье, которые были удалены друг от друга меньше чем на 100 мкм или были записаны на один и тот же микроэлектрод. Естественная периферическая стимуляция с частотой 5—15 Гц часто удлиняла долгозатухающие положительные кросскорреляции, особенно на частоте 10 Гц, причем для многих клеток такие изменения были независимы от изменения средней скорости срабатывания [Ebner, Bloedel, 1981b].

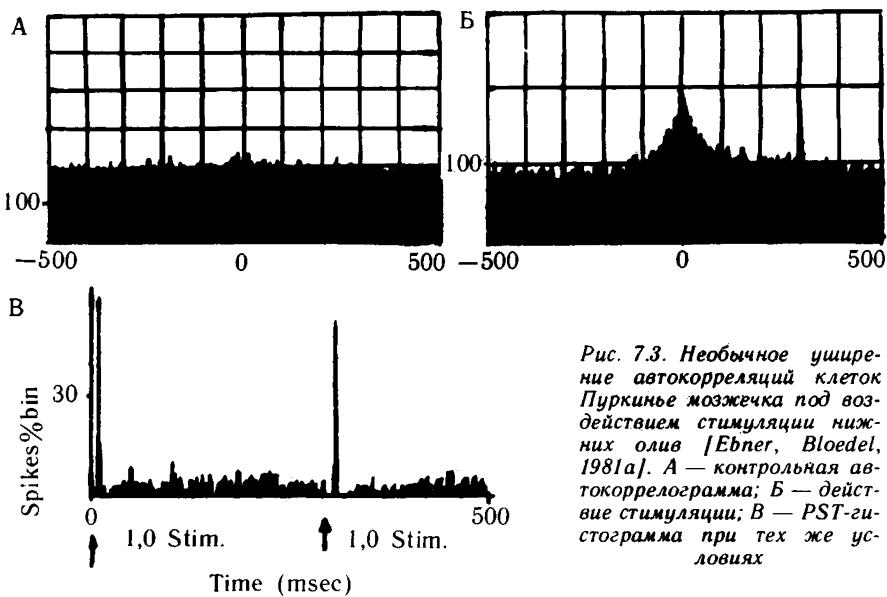


Рис. 7.3. Необычное уширение автокорреляций клеток Пуркинье мозжечка под воздействием стимуляции низких олив [Ebner, Bloedel, 1981a]. А — контрольная автокоррелограмма; Б — действие стимуляции; В — PST-гистограмма при тех же условиях

4. Восприимчивость многих клеток Пуркинье к сигналам мшистых волокон, определяемая как отношение частоты разрядов во время реакции к частоте разрядов в фоне, сильно зависит от времени, прошедшего после появления последнего сигнала лианного волокна, причем она выше у тех клеток Пуркинье, которые имеют автокорреляционную функцию с большим временем затухания. Естественная периферическая стимуляция, запускаемая с задержкой через 20—300 мс после появления каждого комплексного спайка, может увеличивать восприимчивость к сигналам мшистых волокон независимо от того, состоит ли реакция в увеличении или уменьшении скорости разрядов. График восприимчивости четырех клеток в зависимости от величины задержки приведен на рис. 7.4. Для многих клеток отклик наблюдался только, если задержка стимула была небольшой (20—70 мс). Наряду с такими кратковременными изменениями восприимчивости иногда после стимуляции с короткими задержками наблюдались модификации, которые сохранялись и тогда, когда величина задержки увеличивалась до 300 мс [Ebner, Bloedel, 1981c]. Нейронный механизм, объясняющий эти факты, не известен. Среди возможных кандидатов на механизм для данных п. 2—4 Ebner and Bloedel (1981a, 1981b) упоминают:

- долговременные изменения мембранный проводимости после дендритной деполяризации;

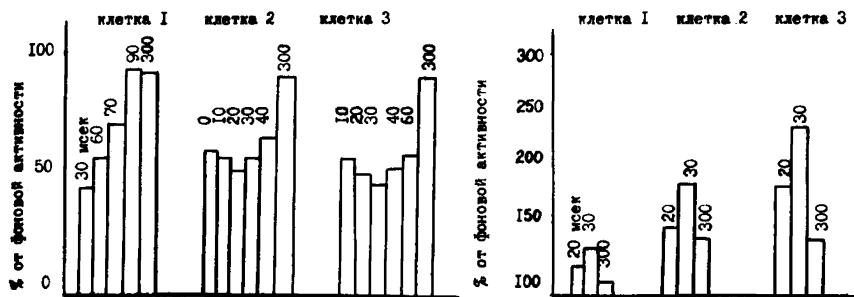


Рис. 7.4. Диаграмма изменений восприимчивости клеток Пуркинье в зависимости от времени, прошедшего после последнего сигнала лианного волокна (указано над столбцами) [Ebner, Bloedel, 1981c]

- участие интернейронов с длительными ТПСП;
- синхронизация лиановых входов;
- влияние возвратных коллатералей.

Объяснение длительного последействия за счет накопления возбуждения во всех этих случаях отвергается, поскольку оно приводило бы к увеличению частоты срабатывания, в то время как чаще наблюдается понижение частоты.

Предлагаемое ниже объяснение связано с импульсной, а не с мембранны-дendритной активностью и охватывает сразу все перечисленные факты. Его физическая природа близка к коллатеральному эффекту Марра (1971), но в отличие от него не обязательно связана с долговременной модификацией синапсов. Скорее речь идет о кратковременной памяти, подробно описанной нами в применении к гиппокампу [Коваленко с соавт., 1982].

Итак, предположим, что описанная в разделе 4 модель имитирует активность совокупности клеток Пуркинье, связанных возбуждающими связями и принадлежащих локали данного тормозного интернейрона. Чтобы воспользоваться полученными там результатами, сделаем два дополнительных предположения. Будем считать, во-первых, что переустановка клетки Пуркинье на низкий уровень активности происходит в момент прихода сигнала по лианному волокну и осуществляется с помощью возвратного торможения либо с помощью механизма ребаунда. Во-вторых, входы от лиановых волокон на ближайшие клетки Пуркинье могут быть активированы одновременно, причем зона синхронности увеличивается в режиме стимуляции. Некоторые физиологические основания для этого предположения изложены в работе Bell, Kawasaki (1972). Переходим к объяснению указанных выше данных в порядке их перечисления.

1. В разделе 3 было показано, что если рассматривать суммарную активность нашей сети в диффузационном приближении,

то потенциал  $\phi(x)$  оказывается бистабильным, то есть имеет два минимума (см. рис. 3.3). Состояние, соответствующее точке  $x_1$ , в которой  $\phi(x)$  достигает локального минимума, называется метастабильным, а состояние, соответствующее абсолютному минимуму в  $x_2$ , — устойчивым. Однако если глубина потенциальных ям примерно одинакова, что возможно лишь при очень малом внешнем поле, оба состояния окажутся метастабильными и тогда возможны спонтанные переходы из одного состояния в другое. Последнее вероятнее всего и реализуется в мозжечке при уменьшении до нуля активности лиановых волокон. Так мы объясняем спонтанно возникающие паузы. Далее, электрическая стимуляция лиановых волокон, эквивалентная включению внешнего поля, приводит с помощью тормозных интернейронов или ребаунда к «размагничиванию», то есть к постепенному подавлению активности клеток Пуркинье и переводу всей системы в состояние  $x_1$ , если она перед этим была в состоянии  $x_2$ . Время до скачкообразного возвращения  $T_{x_1}$  в устойчивое состояние  $x_2$  может быть весьма большим [см. формулу (3.22)], на несколько порядков превышающим постоянную времени ТПСП. Тогда оказывается возможным постепенный, гистерезисный период из  $x_1$  в  $x_2$ . Вообще же гистерезис наблюдается всякий раз, когда один из параметров системы меняется достаточно быстро, так что минимум  $\phi(x_1)$ , в котором сидит система, «двигается» быстрее, чем скорость разложения ( $1/T_{x_1}$ ) этого метастабильного состояния [Agarwal, Shenoy, 1981]. В противном случае переход из  $x_1$  в  $x_2$  будет содержать скачок. Увеличение регулярности спайковой активности при увеличении уровня активности объясняется тем, что на предельно высоких частотах все определяется высокой стабильностью периода абсолютной рефрактерности. Объяснить свойство двухстадийной релаксации количественно мы не в состоянии, поскольку диффузионное приближение его, по-видимому, не описывает, а проверить его на имитационной модели трудно, поскольку требуется очень много машинного времени. Можно лишь заметить, что разложение метастабильного состояния в имитационной модели двумерного изинговского ферромагнетика протекает, действительно, в два этапа. Сначала система довольно быстро релаксирует в промежуточное состояние, а затем довольно долго из него переходит в устойчивое состояние [Stoll et al., 1973]. Качественно это же видно из рис. 3.2 (кривые b и d).

2. Большое время затухания автокорреляционной функции одиночных нейронов возникает при наличии в сети критического режима и при установке специальных начальных условий. В нашем случае эти условия должны быть такими, чтобы сначала активность в лиановых волокнах повысила пороги, обеспечив необходимую «тишину», а затем уже активность параллельных волокон установила начальное пятно или группу

пятен. Поскольку одно лианное волокно воздействует лишь на 10 клеток Пуркинье, а для критического режима необходимо как минимум 25 нейронов [Коваленко с соавт., 1982], то очевидно, что нам необходима синхронизация лиановых волокон, чтобы установить пятно критического размера. Следовательно, если при стимуляции растет синхронность активности лиановых волокон, то время установления стационарного состояния тоже возрастет, и поэтому удлинится время затухания автокорреляционной функции у тех нейронов, которые участвуют в формировании пятна. Наша модель предсказывает, что при дальнейшем повышении интенсивности стимуляции удлинение корреляций прекратится и даже может смениться укорочением.

3. Хорошо известно, что симметричная относительно нуля форма кросскорреляционной функции спайковой активности свидетельствует о наличии общего источника, в то время как сдвиг максимума вправо или влево говорит о более вероятной непосредственной межнейронной связи [Perkel et al., 1967]. В общем случае каждый нейрон в нашей модели испытывает влияние обоих этих источников. Однако если два соседних нейрона принадлежат окрестности одного и того же пятна, то большее значение приобретает непосредственная связь; а если два нейрона удалены и принадлежат окрестности различных пятен, то преобладающим оказывается влияние общего источника, особенно если соответствующие пятна появляются достаточно синхронно во времени. В отсутствие синхронизации, то есть без стимуляции, вероятность встретить симметричную форму кросскорреляции удаленных нейронов будет меньше и равна 0,25 (если вероятность каждого из двух независимых подсобытий принять равной 0,5). Поскольку функция кросскорреляции содержит автокорреляцию как свое слагаемое [Perkel et al., 1967], то согласно предыдущему абзацу стимуляция вызовет тем больше уширение положительной кросскорреляционной функции, чем ближе сеть к критическому режиму. Наибольший эффект стимуляции достигается на «резонансной» частоте  $f \cong 1/t$ , где  $t$  — время сохранения пятна, поскольку при большей частоте эффект сохранения пятна не успевает проявиться вполне; а при малой частоте стимуляции сеть большую часть времени проводит около устойчивой стационарной точки, где время релаксации от малых возмущений невелико. Наконец, независимость указанных эффектов от изменений средней скорости срабатывания — это общее свойство стационарно-связанных процессов.

4. Простейшее объяснение фактов, связанных с восприимчивостью, состоит в предположении, что восприимчивость максимальна в критической точке (это будет вскоре обосновано) и критический режим возникает спустя 20—30 мс после последнего разряда лианового волокна, а затем постепенно нарушается. Еще более простое объяснение могло бы состоять в том,

что все дело в немонотонном изменении порога, достигающего своего минимума спустя 20—30 с после последнего срабатывания и затем постепенно повышающегося до стационарного значения. Однако такое объяснение не проходит, поскольку оно никак не связано с нейронами, имеющими автокорреляции с большим временем затухания, как это имеет место в эксперименте. Поэтому мы продолжим первое объяснение, используя известную аналогию с магнитной системой.

Согласно флуктуационно-диссипационному соотношению статистической физики [Стенли, 1973] магнитная восприимчивость, описывающая реакцию магнитной системы на внешнее воздействие, характеризуется, с одной стороны, как производная от намагниченности по параметру внешнего поля, а с другой стороны, как сумма кросскорреляций всех компонент системы. Отсюда, во-первых, можно дать более точное и более употребительное определение восприимчивости нейронной сети, заменив магнитное поле на суммарную импульсную активность и внешнее магнитное поле на порог нейрона. Это сразу по Теореме 2.1 о далеком порядке дает в критической точке необходимость восприимчивости для бесконечной системы и максимальное значение восприимчивости для конечной системы. Во-вторых, можно корректно определить восприимчивость одиночного нейрона по данному входу как величину кросскорреляции этого входа с выходом нейрона. Для пуассоновского входа оценка этой функции совпадает с PST-гистограммой, и поэтому новое определение восприимчивости согласуется с определением, приведенным в п. 4.

Следовательно, восприимчивость нейрона как функция времени, прошедшего после подачи стимула, изображенная на рис. 7.4 должна быть аналогична правой ветви кросскорреляционной функции вход—выход, причем время ее затухания должно быть того же порядка, что и у автокорреляционной функции этого нейрона. Возможное увеличение восприимчивости после стимуляции следует из приведенных выше объяснений формы авто- и кросскорреляций, за исключением случаев очень большой задержки стимуляции.

В заключение обсудим возможное значение фактов 1—4 для ЦНС. Прежде всего наше единообразное объяснение их свидетельствует о большей вероятности существования в нервной системе коллективных эффектов наподобие фазовых переходов, метастабильности, критического замедления. Но микроскопическим субстратом этих эффектов в физике являются микронеоднородности или «зародыши» нарождающейся новой фазы в окружении более многочисленных компонент старой фазы. Поэтому вся группа данных 1—4 свидетельствует о большой вероятности существования подобных же микронеоднородностей, то есть микроочагов возбуждения в ЦНС. Более того, данные п. 3

и наша интерпретация их позволяют оценить верхнюю границу диаметра этих микроочагов величиной в 100 мкм.

Во-вторых, отдельная диаграмма, приведенная на рис. 7.4, является, выражаясь техническим языком, дискриминационной характеристикой временного (или фазового) компаратора,рабатывающего выходной сигнал, пропорциональный ошибке рассогласования времен прихода сигналов по двум входам (в нашем случае по линнному и мшистому). Следовательно, мозжечок является не столько хранилищем долговременных следов, хотя это пока не исключается (см. ниже п. 5), сколько временным дискриминатором, обеспечивающим другим звеньям двигательной системы сигнал о существовании, величине и знаке ошибки в системе регулирования движением. К этому же выводу, но из совершенно других соображений пришел Lisberger (1982).

Наконец, возникает естественный вопрос: какова возможная структура системы автоматического регулирования, если ее компаратор имеет дискриминационную характеристику, изображенную на рис. 7.4? Ответ для специалиста по системам регулирования почти очевиден: она должна быть системой *фазовой автоподстройки частоты* (ФАПЧ), то есть должна содержать интегратор-генератор управляемой частоты (пейсмекер) и быть, таким образом, астатической системой первого или второго порядка (заметим, что нейронная схема интегратора уже фактически описана нами в конце раздела 4). Но это означает, что регулируемый объект или, точнее, объекты тоже могут быть генераторами почти периодических пачечных колебаний (например, пуль мотонейронов), которые в режиме автоколебательного слежения настраиваются на единый системный ритм, который в силу обратных связей может меняться по ходу выполнения двигательной задачи. Включение или выключение новых групп локальных генераторов определяется их способностью усвоить единый текущий ритм и нужную фазу колебаний, иначе импульсы таких генераторов не попадут в узкий временной интервал дискриминатора и потеряют возможность «подстроить» пейсмекер. Отсюда ясна огромная роль двойного входа в мозжечок как входа в систему выделения ошибки. Короче, описанная выше система регулирования реализует важнейший принцип нервной системы — принцип усвоения единого ритма или, более общим образом, принцип доминанты А.А.Ухтомского. Его универсальность делает неслучайным сохранение одной и той же функции двойного входа в мозжечок на самых разных этапах эволюции. В этом мы видим решение проблемы мозжечка, поставленной Экклсом и описанной в начале этого раздела.

Где же в таком случае хранилище долговременных следов заученных движений? В мозжечке? Но выучить движения в такой системе значит воспроизвести временную последовательность

смены конstellяции локальных генераторов, которая существовала при обучении, что, по-видимому, можно осуществить, задавая какой-то начальный фрагмент движения. При этом очевидно, что информация о прошлых событиях будет храниться не в амплитудах и не в частотах колебаний, и тем более не в весах синапсов, а в фазах колебаний локальных генераторов по отношению к фазе пейсмекера. Да и нужно ли хранить эту фазовую информацию, если для воспроизведения прошлых состояний важны не фазы, а отзывчивость, лабильность генераторов на запросы пейсмекера, что при данной его частоте зависит как от собственной частоты локальных генераторов, так и их задержки прохождения импульсов от пейсмекера к генератору и обратно, то есть от местоположения генераторов внутри организма. Таким образом, мы приходим к такой системе организации движений, которая ничего или почти ничего не должна помнить, кроме постоянства своей морфологии; может ничему или почти ничему не учиться, кроме умения менять свою лабильность, и которая тем не менее имеет довольно большую емкость хранения. Ведь каждый генератор, изменяющий свою частоту от 2 до 4 Гц, при 10%-ной разрешающей способности фазового компаратора имеет 10 различных состояний, что для  $N$  генераторов составит  $10^N$  состояний. Ясно, что стоит проверить высказанные соображения на имитационной модели ФАПЧ.

## **8. ПРИМЕНЕНИЕ К НОВОЙ КОРПУСНОЙ МОДЕЛИ СТАЦИОНАРНОГО И РАСПРОСТРАНЯЮЩЕГОСЯ ВОЗБУЖДЕНИЯ**

### **8.1. «Тихая революция» и связанные с ней проблемы**

Еще до недавнего времени в физиологии господствовал выдвинутый представителями школы Запада принцип «все или ничего». Согласно этому принципу возбуждение в каждом пункте деятельности нервной ткани в каждый данный момент не может градуироваться, углубляться, становиться больше или меньше, но совершенно одинаково и количественно постоянно для каждой ткани. Возбуждение согласно этому принципу — в сущности «импульс», мгновенный стимул.

Первое принципиальное выступление в пользу возможного стационарного возбуждения принадлежит Н.Е.Введенскому, который в 1901 году показал, что кроме бегущей волны возбуждения существует медленное, длительное, градуальное возбуждение и вторая форма возбуждения является общей и более существенной, чем первая. Пропагандируя эти результаты, его ученик А.А.Ухтомский в 20—30-е годы настойчиво протестовал против превращения знаменитого правила «все или ничего» в принцип нервной физиологии, поскольку «адепты этого принципа были склонны представить его чем-то вроде космического закона в стиле теоретической физики, например выводом из квантовой механики» [Ухтомский, 1978]. Но вот в 1956 г. известный физиолог Буллок (см. [Schmitt et al., 1976]) подвел итог предшествующей эволюции взглядов о природе возбуждения и оповестил всех о состоявшейся в этой области «тихой революции». Она состояла в том, что на смену принципу «все или ничего» пришло признание основной, определяющей роли медленных градуальных локальных возбуждений как основы обработки информации в мозге. При этом роль импульсной активности сводится в основном к быстрой сигнализации о пространственно удаленных процессах и рассматривается теперь как новейшее добавление к более эволюционно древним механизмам локального возбуждения, возникшее в связи с увеличением размеров организмов.

В последнее десятилетие генезис локальных потенциалов стали связывать не столько с ионно-мембранными механизмами

одиночных нейронов, сколько с механизмами активности элементарных ансамблей, групп взаимодействующих нейронов. Особое внимание уделяется «локальным цепям» [Schmitt, 1976; Шмитт, 1981], состоящим преимущественно, но не исключительно из мелких нейронов Гольджи II. Этих мелких клеток в головном мозге очень много, у них короткие, иногда кустистые аксоны, которые не входят в крупные тракты белого вещества за пределами данной области. В этой же связи становится весьма популярной гипотеза о модульном строении коры [Маунткасл, 1981], основанная на открытии функционально и морфологически однородных групп нейронов (бочонков, колонок), являющихся, по-видимому, структурными элементами обработки информации в мозге. Вместе с тем осознается, что иерархический принцип работы мозга отчасти теряет свою эвристическую ценность в пользу параллельной, содружественной работы большого числа однотипных модулей [Эделмен, 1981].

Параллельность обработки информации и огромное число «локальных цепей» (в коре 600 млн. мини-колонок) поднимают во всей остроте вопрос о механизмах координации и синхронизации в работе компонент системы и, в частности, вопрос об условиях перехода локального возбуждения к возбуждению, распространяющемуся на большие расстояния, поскольку локальные реакции медленные и распространяются с большим затуханием. С этим же вопросом тесно связана центральная проблема электрофизиологии о связи возбуждения с торможением, о механизмах отключения «ненужных» в данный момент локальных цепей, о способах и путях замыкания и размыкания связей между ними и исполнительными органами.

Итак, можно кратко сформулировать три основные проблемы, возникающие в связи с «тихой революцией».

I. *Проблема перехода локального градуального возбуждения в распространяющееся.* Какой из случаев их взаимодействия имеет место: независимость, полная взаимообусловленность и единство [Bindman, Lippold, 1981] или дуализм по типу «волна—частица» в квантовой механике [Stuard et al., 1979]? Каков конкретный механизм перехода?

II. *Проблема синхронизации.* Каков механизм возникновения низкочастотных ритмов в мозге и какова роль их в координации работы нервной системы?

III. *Проблема торможения.* Существует ли самобытная, структурно закрепленная «субстанциональная» система, обеспечивающая угашение нервных реакций, либо торможение есть модификация, другая форма возбуждения, например стационарное нераспространяющееся возбуждение? Каков механизм координированного торможения?

Все эти проблемы в их взаимосвязи впервые были поставлены и в значительной мере решены школой Введенского-Ухтомского. Начатая ими революция, оказалась на Западе слиш-

ком «тихой» и задержалась там более чем на 60 лет. Чтобы оценить глубину проблем и нетривиальность предложенных решений, познакомимся с краткими основами учения этой школы. Затем, тоже после краткого обзора существующих моделей, сформулируем простую модель стационарного возбуждения на основе «эффекта сохранения пятна» и с ее помощью попытаемся интерпретировать некоторые данные и теоретические положения этой школы, которые раньше было трудно понять в связи с отсутствием ясной физической модели стационарного возбуждения. Наконец, используя нашу модель, обсудим возможный механизм основного принципа работы мозга — принципа доминанты А.А.Ухтомского.

## 8.2. Принцип локального стационарного возбуждения и торможения Н.Е.Введенского

Многочисленные опыты с нервномышечным аппаратом привели Н.Е.Введенского к заключению, что наряду с бегущей волной в возбудимой ткани существует локальное стационарное возбуждение, которое может быть создано как нормальной стимуляцией, так и местным охлаждением, высокой температурой, щипком, кислотой, местными электрическими токами, отравлением кураре и т.д. Состояние такого альтернированного участка, а также его реакции на приходящее от нерва раздражение определяется одним функциональным параметром, называемым физиологической лабильностью. За меру лабильности, согласно Введенскому, принимается то наибольшее число электрических колебаний, которое данный физиологический аппарат может воспроизвести в секунду, оставаясь в точном соответствии с ритмом максимальных раздражений, то есть без какой-либо трансформации ритма.

Закон относительной лабильности Н.Е.Введенского гласит: «По мере того, как частота сильных импульсов, исходящих от нерва (образования наиболее лабильного) приближается к предельной величине (оптимуму частоты), выражающей крайнюю лабильность нервных окончаний (образования наименее лабильного), высота тетануса (см. ниже) увеличивается. Как только этот критический предел перейден, мышечная активность начинает уменьшаться; нервные окончания передают нервные импульсы, придавая им свой собственный, более медленный ритм, как они это делают и со слабыми импульсами различных частот. При дальнейшем учащении раздражения нерва (пессимум частоты), без изменения его силы, нервные окончания впадают в состояние полной невозбудимости. Это физиологический эквивалент отравления кураре. Мышца, став недоступной для импульсов с нерва, полностью расслабляется и работает на восстановление своих сократительных сил» [цит. по Меркулову, 1960].

Степень возбуждения, превышающая максимум, при котором наступает падение лабильности, называется парабиозом, а соответствующий участок ткани — парабиотическим, если произведенное в ней изменение обратимо. Парабиоз (буквально «около—жизнь») граничит непосредственно со смертью, окончением и нормальным возбуждением, отличаясь от последнего лишь степенью подвижности активного состояния. Парабиоз — явление сугубо нелинейное, свидетельствующее о том, что слитное мышечное сокращение или *тетанус* является не простым наложением одиночных мышечных сокращений (и тогда сокращение мышцы было бы пропорционально силе раздражения), а имеет наиболее выгодные силу и частоту раздражения (оптимум частоты и силы раздражения), которые способны вызвать максимальное сокращение мышцы. Вместе с этим существует такая сила и частота раздражения, которые вызывают угнетение и торможение сокращений мышцы (пессимум частоты и силы раздражения).

Таким образом, торможение, согласно основному положению Н.Е.Введенского, есть модификация и продукт возбуждения. Соотношение между интервалом раздражения и интервалом возбуждения определяет, будет ли переход от возбуждения к торможению или наоборот. Несоответствие между ними ведет к затягиванию интервала возбуждения, к снижению лабильности и торможению.

Необычайно смелой мыслью Н.Е.Введенского было положение о справедливости этих закономерностей не только для нервно-мышечного аппарата, но и для центральной нервной системы. Тем самым местное нераспространяющееся возбуждение признавалось не как дефектное состояние и не как специфическое свойство группы агентов, его вызывающих: «парабиоз нерва должен быть признан всеобщей реакцией на самые разнообразные воздействия, реакцией более общей, чем его состояние возбуждения или деятельности в ходячем значении слова» (Введенский, цит. по [Виноградов, 1952, с. 99]).

Перенос парабиоза в ЦНС вызвал, по свидетельству Виноградова (1952, с. 172), недоверие и даже враждебность Шеррингтона и всех учеников Н.Е.Введенского во главе с А.А.Ухтомским. Последний лишь в 1927 г. признал свою ошибку, назвав свою доминанту преддверием парабиоза и использовав явление пессимума для интерпретации сопряженного торможения (см. раздел 8.5). Развивая далее это положение, он в 1928 г. сформулировал принцип усвоения ритма. Согласно этому принципу нервная ткань способна в той или иной мере приспособливать скорость своих электрохимических процессов к ритмике возбуждающих импульсов, повышая (или понижая) свою лабильность; причем возможен как режим *принудительной синхронизации* ритма ткани и ритма раздражителя, так и собственно усвоение ритма, то есть воспроизведение ритма после прекращения раздражения. Согласно этому принципу

усвоение ритма лежит в основе механизма координации и подчинения нервных центров и нейронных групп с различным уровнем лабильности: настраивая свой ритм в резонанс с единым ритмом системы, данная группа нейронов включается в доминирующую констелляцию; сбившиеся с единого ритма нейронные группы, выключающиеся из нее, реализуя, таким образом, «парабиотическое» торможение, сопряженное с данной доминантой. Следующие факты имеют прямое отношение к упомянутым проблемам I—II.

I. Кроме оптимальной частоты и силы раздражения, для возникновения ритмической активности и проводящих спайковых ответов необходимо оптимальное (критическое) значение поляризации и лабильности. (Это так называемое *правило Голикова*). Способность к бездекрементному проведению возбуждения (называемая в дальнейшем *возбудимостью*) исчезает как при значительном повышении лабильности и поляризации, так и при их значительном снижении. Она наибольшая при среднем уровне поляризации и среднем уровне лабильности. Способность же к локальным ответам (называемая в дальнейшем *реактивностью*) сохраняется после утраты возбудимости, причем она наивысшая при низком уровне лабильности, но также снижается как при гиперполяризации, так и при деполяризации. Возбудимость и реактивность функционального состояния возбудимых структур меняются по-разному: параллельные изменения могут сменяться диаметрально противоположными [Голиков, 1970].

II. При оптимальном уровне лабильности и поляризации и повышенной возбудимости и реактивности наблюдается феномен *усвоения ритма*, состоящий в том, что возникающая синхронно с ритмом приходящих раздражений ритмическая активность сохраняется в течение некоторого времени после прекращения подачи раздражения [Голиков, 1970].

Передача возбуждения в коре из центра сигнала в центр эффектора сопровождается синхронностью биопотенциалов в этих центрах, причем распространение возбуждения тем легче, чем большая поверхность вовлечена в синхронную когерентную деятельность и чем выше коэффициент взаимной корреляции «замыкаемых» центров. Для некоторых животных, например кроликов, указанная синхронность и когерентность осуществляется на частоте тета-ритма. Угасательное торможение («размыкание» центров) сопровождается расхождением частот сочетаемых центров, падением синхронности и синфазности медленных биопотенциалов, снижением лабильности в этих центрах [Ливанов, 1975].

III. Реакция нервной ткани на раздражитель обычно имеет три стадии («трехфазная реакция»): при малой силе стимуляции оказывает поляризующее, негэнтропическое действие, затем, снижая лабильность, приводит к деполяризующему воз-

буждающему влиянию и, наконец, при перераздражении — к функциональному парабиозу. Однако в общем случае тип реакции и порядок смены стадий (фаз) определяется исходным функциональным состоянием и спецификой применяемого раздражения. Возможно, например, «парадоксальное» повышение возбудимости при действии гиперполяризующих влияний на предварительно деполяризованные клетки, так что не всякое увеличение поляризации является симптомом торможения [Голиков, 1970].

Особенно четко прослеживается смена фаз и переходы от локального стационарного возбуждения к распространяющемсяся возбуждению и обратно при выработке и угашении условных рефлексов. Один такой сводный пример, составленный нами по данным Мнухиной (1964, 1970, 1977), представлен в табл. 1. Обращает на себя внимание трехфазность многих реакций, а также возобновление похожих событий как при выработке, так и при угашении рефлекса. Например, появление медленных волн и увеличение амплитуды и длительности ВП при угашении кажется парадоксальным и неоднократно обсуждалось в литературе [Мнухина, 1964].

Приведенные факты, подчеркивая взаимосвязь и корреляции различных показателей стационарного и распространяющегося возбуждения, все-таки, на наш взгляд, лишь косвенно свидетельствуют в пользу монистического представления об их единстве, единстве торможения и возбуждения. В частности, весьма вероятно, что торможение не обязательно является результатом действия особой тормозной субстанции и особых тормозных структур, а является моментом, стадией, модификацией и следствием текущего состояния возбуждения, например результатом десинхронизации различных структур. Однако временное совпадение пусть даже большого количества показателей все же не доказывает их причинную взаимообусловленность и «генетическое» единство. Поскольку прямое общеизвестное экспериментальное доказательство этого монизма вряд ли возможно, особую роль приобретает поиск достаточно простой математической модели, которая смогла бы единообразно объяснить все или хотя бы большинство из указанных фактов. Рассмотрим некоторые из существующих моделей, имеющих отношение к этой проблеме.

### **8.3. Существующие модели и представления**

Первой моделью, имеющей дело со всеми основными проблемами стационарного возбуждения и «парабиотического» торможения, была модель А.А.Ухтомского и П.И.Гуляева [Ухтомский, 1978б], до сих пор сохраняющая свое значение как пример глубокого понимания и разностороннего применения теории нелинейных колебаний к анализу физиологических процессов.

**Таблица 1. Динамика изменений электрографических показателей при выработке мигательного условного рефлекса на звук у кролика и собаки (по данным Мнухиной, 1964, 1970, 1977)**

Отведение	Стадии выработки условного рефлекса			
	генерализация	концентрация	упрочнение	угашение
<b>1. Импульсная активность</b>				
а) распространение	-	+	+	+/-
б) пачечность	-	-/+	+	+/-
в) длительность спайка	-/+	+	+/-	-/+/-
<b>2. ЭЭГ моторной коры</b>				
а) амплитуда волн	-/+	+	+/-	-/+/-
б) частота волн	+/-	-	-/+	+/-/+
в) синхронизация	-	+	+	+/-
<b>3. ЭКоГ слуховой коры</b>				
а) амплитуда ВП	-/+	+	+/-	-/+/-
б) длительность ВП	-/+	+	+/-	-/+/-
в) МПП	-	-	+	+/-
<b>4. ЭМГ (двигательный рефлекс)</b>	-	-/+	+	-
<b>5. Типичные потенциалы</b>	рис. 8.1 а, б	рис. 8.1 с	рис. 8.1 д, е	-

Обозначения: ВП — вызванный потенциал; МПП — медленный поверхностно-поступательный потенциал; (+) — велик(а) или присутствует; (-) — мал(а) или отсутствует; (-/+,-+/-/-) — переходные состояния.

Однако модель имеет лишь сосредоточенные параметры и поэтому мало говорит о природе переменной лабильности живой ткани и об условиях перехода локального возбуждения в распространяющуюся активность.

Сравнительно недавно появилась работа Ливанова с соавторами (1974), которую можно рассматривать как пространственно-распределенный вариант предыдущей модели. Она была построена для проверки очень интересной гипотезы Ливанова, состоящей в том, что для облегчения проведения возбуждения в нейронной сети должно существовать строго определенное соотношение между частотами и фазами срабатывания элементов сети и непрерывными локальными потенциалами, периодически меняющими возбудимость (пороги) в сети. В работе Фролова и Шульгиной (1977) формальные нейроны предыдущей модели были заменены на неформальные и, кроме того, было учтено возможное влияние возвратного торможения. Оказалось, что как бы ни менялась возбудимость нейронов, проведение возбуждения в такой сети сильно затруднено. Однако если тормозную популяцию каким-либо образом отключить (в модели это делалось с помощью катодической депрессии), то наилучшие условия проведения

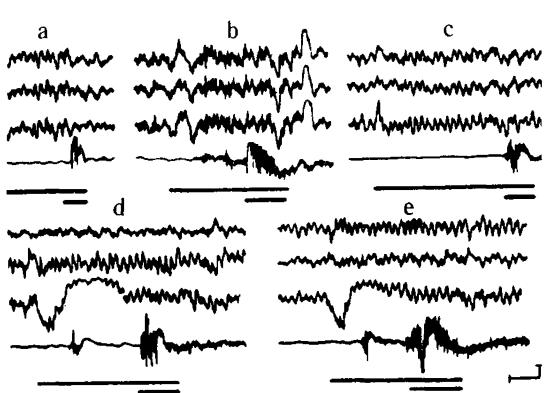


Рис. 8.1. [Мнухина, 1977]  
Динамика изменения ЭЭГ и вызванных потенциалов в процессе выработки и упрочнения двигательного условного рефлекса. а — маскирование ВП в первую фазу выработки условного рефлекса (4-е сочетание); б — генерализация возбуждения, возникновение импульсной активности, увеличение длительности ВП при первом появлении рефлекса (9-е сочетание); с — синхронизация ритма электроактивности в обеих зонах коры, концентрация процесса возбуждения (21-е сочетание); д, е — упрочнение рефлекса сопровождается выявлением в ответ на условный раздражитель медленного позитивного потенциала (41—42-е сочетания). Сверху вниз: две верхних записи — ЭЭГ двигательной зоны коры, третья — ЭЭГ слуховой зоны; биотоки орбиткулярных мышц; отметка условного звукового раздражителя. Калибровка: 1 с, 100 мкВ

Четвертый абзац. Легко заметить удивительную аналогию между динамикой изменения каждого 3-й сверху записи этого рисунка и модельной динамикой активности, изображенной на рис. 3.2. В обоих случаях вначале появляются потенциалы небольшой длительности (случай а), которые к некоторому критическому событию первого проявления рефлекса на рис. 8.1 и первого приближения к сепаратрисе на рис. 3.2 значительно удлиняются, сопровождаемые в обоих случаях возрастанием дисперсии активности (случай б). Затем на обоих рисунках происходит смена полярности длительного отклонения от стационарного состояния (случай д) и, наконец, (случай е) постепенное уменьшение амплитуды и длительности отклонения от стационарного уровня.

возникают при некотором среднем уровне входного возбуждения.

В работе Rotterdam et al. (1982) для сети из неформальных нейронов с возвратным торможением показано расчетами, что при некоторых значениях параметров в сети может возникать пространственно-ритмическая активность на частоте альфа-ритма. Однако при этом авторам неясно, что является определяющим фактором: пространственное ли распределение связей или возвратное торможение? Отмечается также, что для получения более низких частот колебаний в диапазоне тэта-ритма в этой модели потребуется ввести возвратное возбуждение вместо возвратного торможения. Выводы работы не бесспорны, так как они основаны на линейной теории. Это же замечание относится к ряду других линейных моделей корковой ритмической и распространяющейся активности [Жадин, 1982; Freeman, 1983].

Наиболее тяжелое положение сложилось в связи с попытками расшифровать механизм условного или внутреннего торможения. Здесь, несмотря на обилие гипотез и теорий, ни одна не является сколько-нибудь убедительной. Вот мнение

двух известных специалистов. «Одной из самых мучительных загадок условного торможения является вопрос: с помощью какого конкретного физиологического механизма торможению подвергается одновременно вся целостная деятельность организма со всеми его многочисленными эффекторными компонентами?...» [Анохин, 1968, с. 530]. «Внутреннее торможение, по-видимому, еще долго будет мучительной проблемой физиологии, так как здесь необычно сложно переплетаются механизмы интегративной деятельности головного мозга с механизмами межклеточных и внутриклеточных процессов» [Гасанов, 1972, с. 8].

Авторы этих высказываний из трудностей видят выход на пути, указанном учением Введенского-Ухтомского. Правда, первый, адаптируя идею пессимального торможения [Анохин, 1968, с. 460], одновременно отмежевывается от принципа доминанты А.А.Ухтомского. Кроме того, предлагая свою «функциональную систему» как нейронный механизм внутреннего торможения, он вводит в физиологию еще более трудную проблему «приятия решения», родственную, на наш взгляд, известной «проблеме» гомункулуса. Гасанов (1972), наоборот, не имея объяснения механизмов, все же более последовательно сводит проблему внутреннего торможения к доминанте А.А.Ухтомского и, в частности, к режиму нераспространяющегося локального возбуждения.

#### **8.4. Модель «сохранения пятна» и распространения возбуждения**

Предварительно напомним общие сведения об архитектуре коры и некоторые важные дополнительные данные, которые мы собираемся учесть при моделировании. Как известно, кора всегда имеет сходное модульное или колончатое строение. Модуль — это вертикальный цилиндр диаметром 500 мкм и высотой 2000 мкм, в котором все клетки (их примерно  $10^3$ — $10^4$ ) находятся под влиянием одного афферентного волокна одной и той же модальности, оканчивающегося в основном в IV слое. Этот афферент может возбуждать короткоаксонные звездчатые клетки (тип Гольджи II), а также базальные дендриты пирамидных клеток (тип Гольджи I). Аксоны звездчатых клеток в свою очередь оканчиваются на апикальных и базальных дендритах пирамидных клеток, которые возбуждаются, таким образом, моно- или полисинаптически. После некоторой обработки такой локальной целью информация передается в другие центры мозга по аксонам пирамидных клеток. Полагают, что одновременно с этим возбуждаются тормозные нейроны наподобие баскетных клеток, которые в свою очередь тормозят окружающие структуры пирамидных клеток. Таким образом, в каждый данный момент активизируется модуль, связанный

с заданной точкой на сенсорной поверхности тела с помощью специфического афферентного волокна, в то время как окружающие цилиндры более или менее выключены.

Эти данные дополним сведениями о горизонтальном строении коры. Прежде всего каждый из шести слоев имеет свой характерный клеточный состав и свои эфферентные адреса. Затем существенно, что для некоторых животных описанная выше колончатая структура либо не показана, либо вообще отсутствует, а вместо нее в III—V слоях обнаружены параллельные подслои со своими афферентными каналами [Rager, 1983]. Необходимо, таким образом, учитывать как вертикальную, так и горизонтальную структуры коры. В частности, мы считаем важными следующие данные, полученные на кроликах [Petsche et al., 1984]:

а) спайки в слоях появляются не одновременно, но между глубинными и поверхностными слоями существует фазовый сдвиг, соответствующий скорости распространения возбуждения в несколько сантиметров в секунду;

б) функция когерентности ЭКоГ (спектр взаимной мощности) имеет наименьшее значение между IV и III слоями, а также между IV и V слоями;

в) «разряды в одной части вертикально ориентированных колонок вызывают ток между другими частями этих элементов преимущественно в вертикальном направлении».

Теперь сформулируем допущения нашей модели. Предположим, что описанная в разделе 4 модель сети имитирует активность сравнительно небольшой группы короткоаксонных клеток типа Гольджи II (GII), принадлежащих окрестности одной выходной клетки типа Гольджи I (GI) и лежащих в данном корковом слое. Назовем эту группу G-гнездом. Будем считать, что выходная GI-клетка одновременно осуществляет возвратное торможение всех клеток GII, принадлежащих данному G-гнезду. (Можно было бы сразу разделить эти функции между разными клетками, но для удобства, сравнивая с предыдущими моделями, мы их временно объединим). Предположим далее, что на сигнальные входы каждого G-гнезда (в основном в IV слое) поступает по классическим путям сенсорная информация; а неспецифические сигналы от ретикулярной системы модулируют одинаковым образом все пороги основных клеток GII. Порог клетки GI установлен таким образом, что она не срабатывает до тех пор, пока не станут активными больше половины нейронов данного гнезда. Учитывая описанную ранее неоднородность в вертикальном направлении, мы будем считать, что связи между GII-клетками смежных слоев значительно слабее внутристойных связей, так что каждая колонка состоит из шести сравнительно независимых эстафетно-подключенных G-гнезд. Наконец, учтем, что выходы колонки (в основном из VI слоя) могут поступать на входы

той же колонки (в основном в IV слой) с помощью строго топически организованных таламо-корковых связей.

Теперь рассмотрим режимы активности, возникающие сначала в одном G-гнезде, а затем во всей колонке. Во-первых, из результатов имитационной модели следует, что в каждом G-гнезде даже в отсутствие входной афферентации за счет только возбуждающих локальных связей может возникать местная неколебательная «стационарная» активность как высокого, так и низкого уровня. Во-вторых, за счет эффекта «сохранения пятна» возможно «удержание» входной специфической афферентации на время нескольких сотен миллисекунд. И, наконец, превышение суммарной местной активности над некоторым уровнем сигнализируется появлением выходной активности нейрона GI, которая по его аксону бездекрементно распространяется вдоль колонки за пределы G-гнезда. Мы получаем, таким образом, ячейку задержки входных сигналов на время, значительно превышающее синаптические задержки, причем время задержки плавно и в широких пределах линейно регулируется уровнем неспецифической активации. При достаточно высоком ее уровне ячейка задержки благодаря возвратному торможению превращается в автогенератор низкочастотной ритмики с периодом, равным величине задержки.

Величина задержки может значительно возрасти за счет последовательного соединения G-гнезд в колонке. Подобный эстафетный механизм задержки, но без объяснения природы больших задержек, предложил Батуев (1984) в качестве модели кратковременной памяти лобной коры.

Учитывая, что выходы колонки поступают снова на ее входы, получаем модель *ревербератора*, которая отличается от предыдущих моделей следующими особенностями (см. свойства генератора в разделе 4).

1. Период реверберации плавно и в широких пределах линейно регулируется уровнем неспецифической активации.

2. Стабильность периода ревербератора может быть доведена до процента за счет усреднения флуктуаций при суммации активности нейронов ГII на входах нейрона GI.

3. Нестабильность проводящих путей замыкания реверберации практически не влияет на период ревербератора, так как его основную долю составляет большая и довольно стабильная задержка в слое.

4. Ревербератор возбуждается при превышении входным сигналом некоторого порогового уровня («оптимум» силы) и срывается при его достаточно высоком уровне («пессимум» силы).

5. Ревербератор легко синхронизируется слабыми внешними периодическими воздействиями, особенно на низкочастотном конце в диапазоне тета-ритма за счет высокой чувствительности к изменению параметров в критическом режиме.

6. Колебания уменьшаются по амплитуде и затем срываются, если частота внешних синхронизирующих импульсов превышает некоторую предельную величину, зависящую от параметров модели («пессимум» частоты); колебания срываются на низкочастотном конце диапазона.

7. Модель не нуждается в дополнительной гипотезе о существовании проторения и облегчения, обычно привлекаемой для стабилизации траектории и периода реверберации.

Рассмотрим более детально, как модель объясняет приведенные выше данные ленинградской школы.

I. *Правило Голикова* в модели возникает сразу, как только мы идентифицируем поляризацию и лабильность в терминах модели. За меру поляризации можно принять слаженную суммарную активность G-ячейки (средний уровень ее горения), а за меру текущей лабильности — текущую собственную частоту ревербераций (более точное определение лабильности будет дано в разделе 8.5). Тогда правило Голикова говорит не только о существовании порога поляризации для перехода от местной активности к распространяющемуся возбуждению, но и о коллективной природе задержки возбуждения, приводящей к колебаниям. Наивысшая реактивность связана с локальным критическим режимом, когда задержка максимальная, то есть при наименьшей лабильности ревербератора.

II. Морфологическое сходство множества элементарных ревербераторов, а также близость собственной частоты каждого из них к частоте septального тэта-ритма (у них по существу один и тот же механизм) создает наилучшие условия для массовой синхронизации на этой частоте. При внезапном возбуждении системы, даже в случае независимых ревербераторов, их суммарная активность затухает довольно долго, в течение нескольких периодов колебаний, в соответствии с затуханием автокорреляционной функции, изображенной на рис. 4.16. При отключении синхронизирующего воздействия, колебания также в течение некоторого времени сохраняются, но все-таки возвращаются к собственной частоте, так что явление усвоения ритма модель без дополнительных усложнений объяснить не может. Необходимая для этого модификация модели приведена ниже.

III. Чтобы объяснить факты локального и координированного торможения, необходимо более детально, чем прежде, рассмотреть кинетику возбуждения. Мы переходим к интерпретации данных табл. I. Для этого, вообще говоря, нужна системная модель возникновения условных рефлексов в терминах описанных выше ревербераторов. Однако мы пока ограничимся лишь весьма упрощенной локальной моделью рефлекса путем введения в нашу модель G-гнезда следующего дополнительного постулата, который мы предполагаем исключить в будущей системной модели. Предположим, что сила связи между клетками

GII постепенно возрастает по мере повторения числа сочетаний условного и безусловного стимула. Тогда мы можем воспользоваться результатами качественного анализа суммарной активности и, в частности, построениями, приведенными на рис. 3.2. Этому постулату на рисунке соответствует, грубо говоря, постепенный сдвиг изображающей точки вправо по мере роста числа сочетаний. Вначале изображающая точка находится вблизи нижнего стационарного уровня (рис. 3.2а,в), затем в стадии концентрации переходит в критическую область временной устойчивости (рис. 3.2с) и в конце выработки рефлекса, на стадии его упрочнения, — в область верхнего стационарного уровня (рис. 3.2д,е). Динамика  $E(t)$  в каждом отдельном случае соответствует усредненному отклику одного С-гнезда на условный стимул на данной стадии выработки рефлекса. Когда  $E(t)$  впервые превысит некоторый критический уровень, соответствующий порогу клетки GI, появится распространяющаяся спайковая активность, которая, выражаясь образно, замкнет рефлекторную дугу между центром условного сигнала и центром безусловного стимула; в результате чего может появиться первая двигательная реакция рефлекса. Для этого [Ливанов, 1975] необходимо, чтобы локальные « пятно синхронизации» превратились в глобальное « пятно синхронизации» за счет рекрутования новых ревербераторов, работающих в едином ритме. Отсюда и название — стадия генерализации.

С дальнейшим увеличением числа сочетаний вероятность рефлекса возрастает главным образом за счет роста синхронности и синфазности колебаний. Часть ранее вовлеченных колонок, не связанных непосредственно с условными и безусловными сигналами, отключается, откуда и название — стадия концентрации или специализации. Наша модель объясняет, почему на этой стадии возникают медленные волны ЭЭГ и длительные ВП, которые затем исчезают: время перехода  $t_x$  на этой стадии имеет довольно острый максимум по  $x$  (рис. 3.4), а частота ревербераторов соответственно близка к минимальной как следствие критической метастабильности. Максимальную длительность одиночных спайков на этой стадии тоже можно было бы объяснить критическим замедлением, но для этого пришлось бы перенести наш механизм на уровень ионно-мембранных процессов, что выходит за рамки данной модели. Тем не менее отсюда в принципе ясна связь длительности спайка с лабильностью одиночного нейрона, установленная чисто экспериментально [Мнухина, 1977]. Еще один показатель импульсной активности, а именно пачечность, возникающая на месте одиночных разрядов по мере становления условного рефлекса, легко объясняется в нашей модели как признак критического режима и образования локальных ревербераторов. Естественно, что пачечность, возникнув, сохраняется и на поздней стадии упрочнения рефлекса и исчезает при угашении его.

Менее очевидной оказывается интерпретация динамики изменения ВП и, в частности не отмеченное в таблице появление поздних компонент ВП, изменение их полярности и длительности в процессе выработки рефлекса [Мнухина, 1977] (см. рис. 8.1). Для этого снова обратимся к рис. 3.2, но будем иметь ввиду, что различные G-гнезда могут иметь несколько различающиеся параметры, дающие в результате несколько различающиеся начальные состояния  $E_i(0)$ . Так что регистрируемая макроэлектродом суммарная активность пропорциональна не  $E_i(t)$ , а  $\sum_i p_i E_i(t)$ , где  $\{p_i\}$  — распределение начальных

состояний  $E_i(0)$ . Таким образом, «взвешивая» все ситуации, представленные на рис. 3.2В, мы получаем, что вызванные потенциалы, регистрируемые внутрикортикальными или поверхностными электродами, всегда имеют вид биполярного «двухфазного» сигнала с более или менее выраженнымными ранними и поздними компонентами. Соотношения между компонентами определяются долей нейронов с высоким и низким значением  $E_i(0)$ . На ранней стадии выработки рефлекса (рис. 3.2а) преобладают ранние компоненты, на стадии концентрации или упрочнения — поздние (рис. 3.2б,с), причем преобладающая полярность ВП и его длительность плавно меняются с числом сочетаний.

Таким образом, появление длиннолатентных и продолжительных поздних компонент ВП есть симптом критического режима, признак настройки локальных генераторов на единый ритм системы. Развивая эту идею на системном уровне, мы в будущем предполагаем дать единую схему происхождения поздних компонент и объясним экспериментально обнаруженную их связь с направленным вниманием, кратковременной памятью, обучением и привыканием.

А пока ограничимся интерпретацией связанного с этим более грубого и привычного факта — появления на стадии упрочнения рефлекса *медленного поверхности-положительного потенциала* (МПП) коры [Мнухина, 1964, 1977], а также родственного ему (если не тождественного) *уровня постоянного потенциала* (УПП) коры [Русинов, 1969]. И тот и другой имеют длительность более 1 с, могут изменять полярность при усилении стимуляции, но оба исчезают (а не меняют полярность, как можно было бы ожидать) при угашении условного рефлекса. Всеми этими свойствами обладает в нашей модели интеграл от  $\sum_i p_i E_i(t)$  за период стимуляции.

Закончим обсуждение таблицы замечанием о том, что приведенные в ней данные по динамике ЭЭГ согласуются с обычным представлением ЭЭГ как линейной суммы локальных активностей с тем лишь отличием, что у нас элементарной компонентой — «квантом» ЭЭГ оказывается не одиночный спайк и

не ВПСП или ТПСП, а активность Г-гнезда или ревербератора.

Теперь мы можем обсудить проблему внутреннего торможения. Из приведенной выше интерпретации таблицы следует, что знак реакции на нейронном или интегральном уровне (то есть преобладает ли возбуждение или торможение) зависит не только от знака воздействия (угашение или выработка рефлекса), но и от текущего состояния системы и от предыстории. Например, из рис. 3.2а видно, что небольшие возбуждающие деполяризующие воздействия на систему могут привести к локальным тормозным реакциям, если система находится вблизи нижнего стационарного состояния. И наоборот, небольшие тормозные (гиперполяризующие) воздействия могут приводить к возбуждающим реакциям, если система предварительно сильно деполяризована, то есть находится вблизи верхнего стационарного состояния (см. рис. 3.2д,е).

Выработка условного рефлекса в терминах локального возбуждения — это просто постепенное накопление возбуждения (несмотря на его угашение после каждого предъявления стимула), приводящее в конце концов к переходу системы из нижнего стационарного состояния в верхнее, через промежуточное, временно устойчивое состояние. Угашение рефлекса, наоборот, — обратный переход системы из верхнего стационарного состояния в нижнее, очень похожий по кинетике на прямой переход.

Наличие возвратного торможения немого усложняет эту картину: при переходе изображающей точки на рис. 3.2 через зону неустойчивости возникают колебания, амплитуда которых постепенно уменьшается, а частота возрастает, пока колебания не сорвутся (рис. 8.2). Прекращение подкрепления вызывает обратное движение изображающей точки с повторением описанных событий в обратном порядке. Еще раз отметим, что только благодаря возникновению реверберационных колебаний возможна координация возбуждения и торможения, выработка и угашение условного рефлекса.

Таким образом, можно говорить о трех стадиях функционального внутреннего торможения:

1) неглубокое — за счет расфазировки ревербераторов при сохранении монохромности колебаний;

2) основательное — за счет различия частот (гетерохронизма) ревербераторов;

3) полное — за счет срыва колебаний при сохранении местного нераспространяющегося возбуждения.

Важно отметить, что увеличение глубины функционального торможения не обязательно сопровождается уменьшением деполяризации. Так, если колебания срываются вблизи верхнего стационарного состояния, то и остаточная деполяризация будет

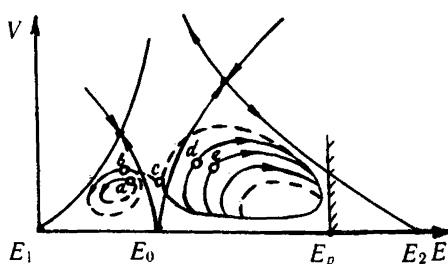


Рис. 8.2. Кинетика уровня возбуждения с учетом возвратного торможения (отражающий барьер в точке  $E_p$ ): пунктиром показано два типа срывов колебаний

высокой; а если вблизи левой сепаратрисы, то низкой. В первом случае, как видно из рис. 8.2, амплитуда и период колебаний перед срывом убывают, а во втором, наоборот, — возрастают. Хотя оба эти случая приводят к системному торможению, но второй выгоднее, поскольку в первом случае уровень деполяризации или остаточного, «скрытого» возбуждения выше. Мы продолжим обсуждение проблемы торможения в следующем разделе.

Crick et al. (1983) на основании машинных экспериментов с нейроноподобными сетями пришел к выводу, что основная функция парадоксальной фазы сна, вопреки существующим гипотезам, — вовсе не поддержание гомеостаза ключевого медиатора, не обработка информации, собранной во время бодрствования, и не консолидация возникающих тогда синаптических связей, а наоборот, ослабление этих связей, «обратное обучение» (unlearning). Этим снимается перевозбуждение, устраняются навязчивые идеи, галлюцинации и другие «паразитные» виды активности, неизбежно сопутствующие сетям с модифицируемыми синапсами. Никакого механизма «обратного обучения» Крик не предложил.

Гипотеза Крика не является новой по существу. Еще И.П.Павлов считал сон и внутреннее торможение одним и тем же процессом. У этой гипотезы примерно та же основная трудность, что и у предыдущих: чем объяснить, что в парадоксальную фазу сна возникает столь сложный комплекс проявлений (быстрое движение глаз, изменение дыхания и сердечно-сосудистой деятельности, подергивание мышц, сновидения, гиппокампальный тета-ритм, кортикальная активация, напоминающая бодрственное состояние организма)?

Полный ответ возможен только в рамках системной модели внимания, однако частичное объяснение вытекает из нашей локальной модели колонки, если учесть, что парадоксальный сон как внутреннее торможение — это фаза возобновления ритмических колебаний ревербератора в процессе угашения (неподкрепления) условных рефлексов, выработанных во время бодрствования. Поэтому «парадоксальное» увеличение корковой активности во время этой фазы не более парадоксально, чем другие описанные выше проявления условного торможения. Можно образно сказать, что сон — это борьба организма с парабиозом.

## **8.5. Интерпретация свойств доминантного очага**

Описанная модель подсказывает, что все проблемы, поднятые в разделе 8.1, тесно взаимосвязаны и решение первой из них дает в значительной мере решение и остальных. Так что имеется возможность рассмотреть физиологический механизм активности коры с единой «парабиотической» точки зрения. Мы проследим за этой возможностью на примере анализа и интерпретации свойств доминанты А.А.Ухтомского — одного из наиболее острых дискуссионных вопросов электрофизиологии. Вначале кратко изложим сам принцип доминанты А.А.Ухтомского. Он состоит в следующем.

В каждый момент времени в нервной системе существует лишь одна активная доминирующая конstellация (созвездие, очаг) совозбужденных локальных нейронных групп или центров, характеризующаяся единым ритмом и единым действием (поведением). Одни и те же отдельные центры или группы нейронов могут входить в состав различных доминирующих конstellаций, причем вхождение в данную конstellацию или выключение из нее определяется способностью этих групп или центров усваивать единый темп и ритм активности. Следы прежних доминант длительно сохраняются в высших отделах нервной системы и при полном или частичном восстановлении первоначальных условий могут вспыхнуть вновь полностью или частично. Доминантный очаг характеризуется следующими свойствами:

- 1) повышенной возбудимостью;
- 2) стойкостью возбуждения;
- 3) способностью суммировать возбуждение;
- 4) инерцией, то есть способностью сохранять это состояние, если первоначальный стимул уже миновал;
- 5) «сопряженным» торможением центров, функционально не совместимых с деятельностью очага.

Последнее свойство сам А.А.Ухтомский считал особенно важным («целая половина принципа доминанты»), но в перечне свойств доминанты, составленном им в 1926 г., оно почему-то отсутствует. Один из учеников А.А.Ухтомского, обративший в печати внимание на эту особенность, считает, что пятое свойство помимо сопряженного торможения, должно включать в себя торможение в самом доминантном очаге [Уфлянд, 1975].

По свидетельству другого ученика, после 1937 г. на основании опытов своих учеников А.А.Ухтомский полагал, что пятым признаком необходимо считать усвоение ритма раздражений, происходящее на фоне понижения лабильности центров доминантного очага [Меркулов, 1959]. Относительно же торможения «самому Ухтомскому было очевидно, что нет ясного и наглядного показа, как и при каких условиях «сопряженное

торможение» обеспечивает протекание одной доминанты, полностью угнетая возникновение других рефлексов» [Меркулов, 1977].

Еще один ученик А.А.Ухтомского, И.А.Аршавский, считает, что в более ранний период А.А.Ухтомский понимал существенно более правильно природу торможения, нежели это имело место позднее, когда он считал, что координированное торможение складывается по типу парабиоза, поскольку последнее означает невероятно расточительнуютрату энергии [Аршавский, 1985].

Сам А.А.Ухтомский писал, что «понять доминанту в терминах парабиоза трудно, так как надо вводить неизвестный фактор, который бы увеличивал лабильность в доминанте и снижал бы ее в тормозимых центрах» [Ухтомский, 1978].

Не менее сложным и противоречивым оказывается четвертое свойство доминанты — ее инерционность. С одной стороны, оно требует, чтобы интервалы возбуждения в доминанте были бы достаточно длительными, продолжающимися даже в отсутствие стимулов, то есть лабильность в доминанте должна быть низкой. С другой стороны, А.А.Ухтомский неоднократно писал, что доминанта связана с повышением лабильности в доминантном очаге. Вопрос о противоречивой природе инерционности наиболее остро сформулирован самим Ухтомским в одной из архивных записей за 1938 год: «Единство события, которым характеризуется возбуждение длительных интервалов, является ли, в самом деле, выражением инерции в очень малолабильном приборе и в силу малой лабильности последнего. В этом случае данный прибор вообще был бы неспособен осуществлять сколько-нибудь быстрые периодики и сигнализацию в себе! Или дело идет об объединениях некоторого другого характера, когда единство «предмета», «события» получается лишь оттого, что его элементы спаиваются между собой исключительно высокими скоростями» [Ухтомский, 1938].

Мы не будем продолжать дальше в этом духе — критику остальных свойств доминанты можно найти в нейрофизиологической литературе [Анохин, 1968; Беленков, 1980; Крыжановский, 1980]. Но и сказанного вполне достаточно, чтобы убедиться, что принцип доминанты до предела обострил противоречия «парабиотического» подхода и разрешение трудностей вряд ли возможно без серьезной математической модели доминанты. Ниже будет показано, что в качестве такой модели «микроочага» доминанты может служить модель G-гнезда и в основе доминанты лежит ее свойство инерционности.

Но сначала мы должны понятие лабильности, основное понятие школы Введенского-Ухтомского, перевести на язык нашей модели. Для этого, на первый взгляд, достаточно приведенное в разделе 8.2 определение лабильности связать с максимальной частотой реверберации и оценить последнюю через

параметры модели. Однако такой подход не учитывает случаев неколебательного возбуждения, а также скоростного вовлечения ревербераторов в большую систему, так что, например, «негугомонный пескарь» окажется лабильнее палкой лежащей «косной» щуки, «опровергающей» такое поспешное определение исходом охоты на пескаря. Поэтому нам придется вначале получше уяснить то понятие лабильности, которое сформировалось у автора доминанты.

«Одна из существенных для нашей школы теорем заключается в положении, что лабильность ткани есть величина изменчивая и притом по ходу реакции, то есть под влиянием приходящих импульсов» [А.А.Ухтомский, 1966, с. 125].

**Аспект 1:** Если  $x$  — вектор состояния ткани в момент  $t$ , то лабильность есть функция  $x$ .

«Можно характеризовать лабильность как степень инерции, понимая ее физиологически в смысле большей или меньшей протяженности активного состояния во времени. Высокая инерция — это малая способность освобождаться от начавшегося возбуждения...» [А.А.Ухтомский, 1954, с. 43].

**Аспект 2:** Лабильность обратно пропорциональна степени инерции. Степень инерции в состоянии  $x$  пропорциональна времени жизни состояния  $x$ .

«Наилучший и наиболее наглядный критерий высокой лабильности заключается в том, насколько данной возбудимой системе доступна срочная мгновенность реакции в смысле ли прихода к возбуждению, в смысле ли возвращения к покоя» [А.А.Ухтомский, 1978, с. 168].

**Аспект 3:** Если  $x$  — состояние системы в начальный момент времени, то лабильность системы обратно пропорциональна минимальному из времен прихода системы из этого состояния в одно из ее стационарных состояний.

На основании всего этого мы предлагаем следующую идентификацию:

$$\text{Модельная лабильность} = (at_x)^{-1},$$

где  $t_x$  — среднее время первого достижения системой любого из ее стационарных состояний, исходя из начального состояния  $x$ ;  $a$  — постоянная. Для оценки модельной лабильности, таким образом, можно воспользоваться формулами и графиками из раздела 3. С их помощью можно проверить качественное согласие модельной лабильности с экспериментальными данными, приведенными в разделе 8.2, принимая за  $x$  уровень деполяризации или средний уровень активности сети.

Модельная лабильность может служить оценкой «качества» системы и пригодна как для локальной, так и для глобальной системы, как колебательной, так и неколебательной. В последнем случае ее можно понимать как обратное время реакции на внезапное одиночное возмущение (аналогичное реакции на скачок в системах регулирования). В случае колебательного

возбуждения локальная модельная лабильность пропорциональна частоте отдельного ревербератора, а глобальная модельная лабильность — частоте синхронных колебаний системы. Если локальная лабильность гетерохронного участка ткани ниже глобальной, то этот участок не в состоянии следовать за общим ритмом и проводить возбуждение. Он обособляется, выпадает из системы, «пессимально тормозится». В этой ситуации модельная лабильность согласуется с определением лабильности по Введенскому (см. раздел 8.2). Но тот же участок не сможет усвоить общий ритм и в том случае, когда его лабильность — частота выше глобальной. Отсюда для нашей модели следует правило Голикова: для возникновения ритмической активности и проводящихся ответов необходимо оптимальное значение лабильности ( $al_x$ )<sup>-1</sup> и поляризации  $x$ .

Применительно к доминанте это означает, что «сопряженному» торможению подвергаются все участки ткани, не имеющие оптимальной лабильности и поляризации. Более того, благодаря зависимости от  $x$  можно исключить из доминирующей конstellации не только те ревербераторы, которые не имеют оптимальной частоты, но и те, которые не имеют нужной фазы колебаний. На этом принципе можно построить модель системы с «обостренным вниманием», у которой «сопряженное» торможение ненужных в данной доминанте ревербераторов мгновенно и очень эффективно управляет самим корковым доминантным очагом, реализуя «произвольное» или словесно направленное внимание, восприятие, воспоминание, узнавание и т.д.

Значит сопряженное торможение, более тесно связанное с высшими функциями, можно моделировать, используя, по существу, те же локальные и системные механизмы, что и внутреннее (условное) торможение, которое обсуждалось в конце раздела 8.4. В обоих случаях в соответствии с учением Введенского-Ухтомского «процессы возбуждения и торможения не противоположные процессы, а родовым образом один и тот же процесс с противоположным конечным эффектом в зависимости от условий его осуществления. Основное и определяющее условие для эффекта дается степенью лабильности действующего эффектора в наличный момент времени» [А.А.Ухтомский, 1966].

Не противоречит ли этому наш постулат о существовании возвратного торможения и тормозных нейронов? Нет, поскольку их роль сводится в основном к локальной переустановке начального состояния ревербератора и/или к организации «тишины», необходимой для развития эффекта сохранения пятна. Кроме того, эту же функцию можно реализовать на основе поствозбудительного торможения без какого-либо участия тормозных интернейронов.

Рассмотрим теперь связь отдельных свойств доминанты с инерционностью. Повышенная возбудимость (первое свойство) качественно объясняется тем, что в системе, находящейся

вблизи точки неустойчивого равновесия (в критическом режиме), даже микроскопически малые входные возмущения могут приводить к макроскопическим заметным изменениям выходной активности. Количественной мерой возбудимости может служить амплитуда колебаний, а в случае отсутствия колебаний — дисперсия локальной активности  $V(t)$ , описываемая уравнением (3.7). Ее поведение во времени и при изменении параметров показано на рис. 3.2Б, нижний ряд. Можно показать, что эта характеристика совпадает с восприимчивостью, которая подробно обсуждалась при моделировании мозжечка. Там было показано, что ее максимум в критическом режиме обусловлен медленным затуханием межнейронных корреляций, то есть прямо связан с инерционностью, критическим замедлением. В зависимости от того, пройден этот максимум или нет, входные воздействия могут как повышать, так и понижать возбудимость и инерционность системы. Это объясняет несколько «парадоксальных» наблюдений: а) повышение возбудимости и инерционности системы при снижении интенсивности стимула [Ухтомский, 1959, с. 202—207]; повышение возбудимости при угашении условного рефлекса [Гасанов, 1972, с. 24 и 116]; появление медленных волн как при выработке условного рефлекса, так и при его угашении [Мнухина, 1964].

Как модельно определить стойкость очага (свойство 2) и какова ее связь с инерционностью? Под модельной стойкостью очага будем понимать время до нарушения синхронности колебаний ревербераторов, отсчитываемое от последнего стимула. За меру локальной стойкости можно принять время затухания автокорреляционной функции активности одиночных независимых ревербераторов (рис. 4.16Б). Благодаря сравнительно высокой стабильности частоты колебаний (см. рис. 4.18), она может превышать период колебаний (меру инерционности) на один два порядка, достигая нескольких секунд. Системную стойкость можно значительно повысить за счет специально организованной фазовой автоподстройки синхронизирующей частоты.

Наконец, способность к суммации возбуждения (свойство 3) неотделима от способности к удержанию возбуждения (свойство 4) и означает существование длительных следовых процессов с большой постоянной времени затухания. Хотя наша модель свойства суммации не объясняет, а лишь постулирует его (в форме постепенного роста силы межнейронных связей по мере повторения стимулов), но это сделано лишь для простоты анализа. Мы его могли бы получить как следствие критического режима, предположив, что модификация межнейронных связей возникает только в условиях, когда «сохраняется» пятно, как это было сделано в имитационной модели (см. рис. 4.14). Необходимая при этом суммарная глубина модификации оказывается очень небольшой (порядка

1—2%), а в системной модели ожидается, что она будет еще меньше.

Этот факт может серьезно изменить представления о нейронной памяти как о глубоких структурных изменениях синаптического аппарата. Вместо них принцип доминанты подсказывает системную организацию памяти на основе синхронизированных ревербераторов. При этом свойство инерционности доминантного очага можно понимать как кратковременную память [Кгурков, 1984], свойство стойкости или «фазности» ревербераторов — как долговременную память [Русинов, 1975], а само доминантное состояние — как процесс организации внимания и перевода следов из кратковременной памяти в долговременную и обратно.

Заканчивая обсуждение свойств доминанты, отметим, что в нашей модели все они тесно связаны, взаимообусловлены и неотделимы друг от друга, потому что отражают различные черты единого физического явления — критической метастабильности.

## 9. ЗАКЛЮЧЕНИЕ. ФАЗОВЫЕ ПЕРЕХОДЫ И МЕТАСТАБИЛЬНОСТЬ — ОСНОВА ФУНКЦИОНИРОВАНИЯ МОЗГА

Идея применить теорию фазовых переходов, неустойчивых и метастабильных состояний в ферромагнетиках для объяснения загадочных функций мозга не нова (см. введение). Интересно отметить, что еще в 30-е годы А.А.Ухтомский в связи с открытием им доминанты как основного принципа работы мозга писал: «На доминанту можно смотреть как на состояние *неустойчивого равновесия* (курсив наш) молекул, способное к катализитическому (детонационному) разложению, в то время как обычные реакции — равновесие *обратимых реакций*» [Ухтомский, 1966, с. 241]. В связи с поиском физического механизма доминанты и, в частности, для объяснения ее основного свойства инерционности и способности длительно удерживать следы прошлых доминант он писал: «...Уж если роль следов во времени заставляет учитывать себя в *магнитах*, в коллоидном *гистерезисе*, в отрезке нерва, то надо думать, что в клетке, да еще в нервной, да еще кортикальной передача следов от момента к моменту должна иметь первенствующую роль. Поэтому мы очень приблизимся к действительному пониманию явлений этого рода, когда физическая химия раскроет природу коллоидного *гистерезиса*» [Ухтомский, 1966, с. 62].

В последние годы концепцию неравновесных фазовых переходов в работе мозга развивает Haken (1983). Он считает, что мозг в известном смысле аналогичен лазеру — световому источнику с зеркальноотражательной обратной связью, способному в зависимости от интенсивности внешней «накачки» излучать световые волны различной структуры: от диффузионного, шумоподобного поля до непрерывного когерентного излучения и регулярных световых импульсов большой амплитуды. Фазовые переходы от одного режима к другому он сопоставляет с переходами мозга к различным функциональным состояниям таким, как сон, бодрствование, эпилепсия. Не настаивая на более детальной аналогии между нейронами ЦНС и атомами или молекулами лазера, Haken, однако, считает, что общая концепция фазовых переходов будет полезна для теории мозга. В частности, он почти дословно повторяет Ухтомского, когда пишет, что «мозг работает вблизи

точек неустойчивости, если сравнивать мозг с некоторой динамической системой». Он полагает также, что «мускульный синергизм» Шеррингтона есть следствие синергизма нервных клеток, управляемых параметром порядка». Правда, на это ему можно возразить словами А.А.Ухтомского, который еще примерно 50 лет назад писал: «Но у Шеррингтона скорее преформированные координации, основанные на *анатомическом антагонизме* — за счет смены или алгебраического суммирования двух противоположно действующих на синапсы химических реагентов, ...а у Введенского нельзя предсказать исход торможения только на основе анатомических данных. Его антагонизм определяется сроками взаимодействующих возбуждений, и его координации складываются по ходу действия, а не преформированы. Его координация есть всегда вновь и вновь настраивающийся, но и относительно легко расстраивающийся процесс, являющийся функцией интервалов возбуждения в отдельных компонентах действующего механизма... У Шеррингтона не было идеи содружества рефлексов, объединяемых не анатомическим соседством, а функциональным сродством» [Ухтомский, 1978, с. 162].

Идея возникновения торможения из возбуждения наподобие физических фазовых переходов от порядка к беспорядку, от геля к золю, от жидкости к пару и т.д., в противоположность торможению за счет структурно-фиксированной тормозной субстанции пронизывает все аспекты учения Введенского-Ухтомского. Это учение до сих пор сталкивается с трудностями и непониманием вследствие того, что слишком сложно, малопонятно и нередко парадоксально лежащее в его основе физическое явление. Значительно легче принять теорию Экклса, ученика Шеррингтона, о самостоятельной, гиперполяризационной природе торможения, возникающего за счет тормозного медиатора на основе взаимоотношений с возбуждением по типу хищник—жертва. С такой теорией нет особых проблем, но и нет путей к пониманию доминанты А.А.Ухтомского.

В данной работе представлены теоретические и экспериментальные доказательства того, что фазовые переходы и связанные с ними критические явления действительно существуют в самых различных мозговых структурах. Теория существенно опирается на недавние результаты по взаимодействующим марковским процессам. Она включает в себя микро-теорию (вывод обратного уравнения Колмогорова из постулатов базовой нейронной модели и доказательства существования далекого порядка в ней), макро-теорию (вывод динамической системы из уравнения Колмогорова, обоснование S-образной нелинейности и оценка времен жизни неустойчивости и метастабильных состояний) и имитационное моделирование («эффект сохранения пятна» и исследование свойств генератора на метастабильных состояниях).

Экспериментальные доказательства существования в мозге равновесных и неравновесных фазовых переходов и временно устойчивых состояний собраны нами из ранее опубликованных данных по четырем нейронным структурам: гиппокампу, септуму, мозжечку и коре. До сих пор интерпретация этих данных в терминах механизма, как правило, отсутствовала, что исключало априорную «настройку» экспериментаторов на проверяемую гипотезу. То обстоятельство, что одна и та же модель и один и тот же эффект оказались пригодными для объяснения большого набора трудных для понимания данных весьма непохожих нейронных структур свидетельствует о том, что мозг работает по единому принципу, а все различия между его структурами имеют количественный, а не качественный характер. Этим апостериорно оправдывается наш интерес прежде всего к качественным и лишь во вторую очередь к грубым количественным характеристикам. Дальнейшее обсуждение результатов мы продолжим в форме ответов на вопросы, поставленные во введении.

1. Теория ферромагнетизма, служащая для нас в качестве основной аналогии для локального описания мозга, построена на основе очень грубого предположения о том, что парный потенциал взаимодействия положительный и равен нулю вне некоторого интервала длины связей. В нашей базовой нейронной модели этому соответствует постулат о том, что все синаптические связи — возбуждающие и короткодействующие. Однако хорошо известно, что в мозге наряду с возбуждающими синапсами повсеместно существуют и тормозные синапсы, за исключением, возможно, коры. Это дало основание некоторым теоретикам мозга (см. например, [Wilson, 1974]) говорить о существенном различии природы кооперативных явлений в физике и нейронных системах, поскольку в физике нет ничего аналогичного тормозным обратным связям. Наши результаты показывают, что для существования таких явлений в нашей базовой модели тормозные связи не являются абсолютно необходимыми. Их роль в некотором смысле вспомогательная: они срывают критический режим по достижении некоторого состояния и переустанавливают сеть на начальное состояние. Эту функцию можно реализовать без изменения локального механизма кооперативности на основе поствозбудительного торможения без какого-либо участия тормозных интернейронов. В этом мы видим теоретическое оправдание веры Н.Е.Введенского в универсальность «пессимального» торможения, понимаемого, правда, в слегка расширенном смысле, указанном в конце раздела 8.4.

2. В какой мере реальные межнейронные связи отражены в базовой нейронной модели, чтобы критический режим не оказался изысканным «артефактом»? Ясно, что связи по характеру дальнодействия отличаются как от модели Изинга, так и от схемы «все на всех». Но поскольку мы доказали, что крити-

ческий режим существует в этих крайних случаях, то он существует и в более реалистическом промежуточном случае. Более того, наши результаты имитационного моделирования говорят, что сила связей может быть случайной величиной с фиксированным средним и довольно большой дисперсией. Последнее имеет особое значение с двух точек зрения: во-первых, известно, что структура связей в мозге представляет парадоксальное сочетание порядка и хаоса; во-вторых, синаптические шумы и другие стохастические факторы играют решающую роль в образовании нейронных ансамблей и пренебречь шумами, значит лишить модель надежности, легкой управляемости и других необычных внезапно появляющихся в критическом режиме свойств.

В самом деле, критический режим существует за счет тонкого равновесия между разрушительным действием шумов и противодействующим влиянием межнейронных связей. Это приводит к тому, что точка неустойчивого равновесия результирующей динамической системы становится временно устойчивой, за счет чего и возникают необходимые в ЦНС свойства инерционности системы и высокой чувствительности к внешним воздействиям. Необычно важная роль флуктуаций в нашей модели подчеркивается тем, что они в решающей степени контролируют параметры S-образной нелинейности динамической системы, благодаря чему дисперсия суммарной активности играет в нашей динамической системе (3.6) и (3.7) роль второй существенной переменной, роль, которая ранее принадлежала отдельной тормозной популяции. В результате пренебречь шумами, значит лишить систему способности задерживать возбуждение на сотни миллисекунд и более. Стохастическая природа активности оказывается не каким-то внешним шумом и источником ненадежной работы мозга, для борьбы с которыми формируются «элементарные нейронные ансамбли», но является существенным аспектом механизма, без которого невозможна инерционность, низкочастотные ритмы, привыканье и т.п. Короче, пренебречь внутренними шумами, значит сделать невозможным «эффект сохранения пятна» и всю основанную на нем интерпретацию экспериментальных данных.

3. Подводя итоги моделирования четырех упомянутых выше нейронных структур, мы приводим таблицу соответствия между фактами ферромагнетизма и данными нейрофизиологии (табл. 2), которая свидетельствует о неслучайности магнитной аналогии для мозга. Обращает на себя внимание прежде всего сходство, иногда тождественность понятий и терминов (например, спонтанная активность, восприимчивость, гистерезис). Затем можно отметить значительное разнообразие весьма трудных для объяснения эффектов, которые кажутся самостоятельными, независимыми друг от друга, а по существу являются различными аспектами одного и того же явления фазового перехода. Наконец, в последней строке таблицы записана гипотеза,

**Таблица 2. Соответствия, установленные с помощью базовой нейронной модели**

Ферромагнетизм	Нейрофизиология
Далекий порядок	Корреляция активности удаленных нейронов
Спонтанная намагниченность	Спонтанная активность
Критическое замедление	Инерционность, «возбуждение длительных интервалов» [Ухтомский, 1978]
Критическое поле	Критический уровень деполяризации [Голиков, 1970]
Критическое уширение автокорреляций	(Критическое) уширение автокорреляций [Ebner, Bloedel, 1981a]
Критическая восприимчивость	Повышенная восприимчивость (responsiveness) [Ebner, Bloedel, 1981c] Повышенная реактивность [Голиков, 1970]
Равновесный фазовый переход I-го рода	Скачкообразные, случайные во времени переходы между возбужденным и тормозным состояниями [Colin et al., 1980]
Неравновесный фазовый переход. Гистерезис	Плавный переход от возбуждения к торможению и обратно при стимуляции. Гистерезис [Colin et al., 1980]. Выработка и угашение условного рефлекса [Минухина, 1977]
Кластеры, домены	Микроочаги возбуждения

которую высказывали прежде очень многие нейрофизиологи и специалисты по моделированию нервной системы. Теперь, на наш взгляд, она является неизбежным следствием всех приведенных нами фактов теоретических и экспериментальных.

Действительно, если микроскопическим субстратом явлений, указанных в левой половине таблицы, являются микронеоднородности и «зародыши» нарождающейся новой фазы в окружении более многочисленных компонент старой фазы, то вся группа данных, приведенная в правой половине табл. 2, несомненно, свидетельствует о существовании подобных же функциональных микронеоднородностей в мозге, которые по аналогии с доминантным очагом возбуждения естественно назвать микроочагами возбуждения. Они отличаются от «элементарных ансамблей» своей относительной подвижностью в пространстве и сравнительно небольшим временем жизни, которое в отсутствие колебаний на 1—3 порядка превышает постоянную времени мембранны. Короче, табл. 2 по совокупности фактов, объясненных с помощью «эффекта сохранения пятна», доказывает существование самого этого эффекта в мозге как нового, ранее неизвестного явления живой природы.

4. Перечислим кратко, какие задачи и проблемы удается решить с помощью этого наблюдения. Существование указанных микроочагов впервые позволило:

— понять механизм инерционности доминанты как критического замедления кинетики микроочагов;

- представить непротиворечивую интерпретацию известных пяти свойств доминантного очага;
- объяснить несколько парадоксов, связанных с динамикой электрической активности мозга при выработке и угашении условных рефлексов;
- устранить основные дефекты реверберационной гипотезы кратковременной памяти.

Отметим специально, что высокая чувствительность критического режима к изменению параметров делает излишней глубокую синаптическую модификацию при заполнении сенсорной информации: достаточно все веса связей микроочага изменить на 1%, чтобы он радикально изменил свое динамическое поведение. Это должно сильно затруднить поиск гипотетического синапса Хебба, на котором основана большая часть современных теорий мозга, требующих глубоких структурных изменений [Haken, 1983].

Существование долгоживущих микроочагов возбуждения позволяет выдвинуть следующее новое предсказание о конкретной форме связи долговременной и кратковременной памяти. Если кратковременную память отождествить с активностью ревербераторов, а долговременную память — с фиксированными локальными микроочагами в корковых слоях, то переход из одной формы в другую должен быть тождественным с переходом от локального стационарного возбуждения к распространяющемуся, причем доминанта А.А.Ухтомского, как констелляция совозбужденных на одной частоте множества колонок — это способ «оживить» «мертвые» долговременные следы, легко и быстро отобрать из них лишь те, которые были «записаны» когда-то в аналогичной ситуации (т.е. при той же частоте и фазе колебаний) и скомбинировать различные модальности в единый многогранный интегральный образ.

5. Существование долгоживущих микроочагов возбуждения закрывает несколько взаимосвязанных проблем электрофизиологии, сформулированных и частично решенных школой Введенского-Ухтомского:

- проблему перехода локального градиального возбуждения в распространяющееся возбуждение;
- проблему синхронизации нейронных структур и усвоение ими ритма раздражителя;
- проблему внутреннего и «сопряженного» торможения.

Однако проблема устойчивости доминанты, поставленная А.А.Ухтомским, остается нерешенной. Теперь с учетом долгоживущих микроочагов возбуждения ее можно было бы сформулировать следующим образом: как совместить высокую надежность и устойчивость нейронной памяти (следы прошлых доминант) с принципиально неустойчивым режимом долгоживущих и метастабильных состояний. Одно из возможных решений [Stuart et al., 1979], предложенное для аналогичной

проблемы, состоит в том, что мозг считается смешанной физической системой, в которой электрохимическая часть взаимодействует с квантовым кооперативным эффектом. В этой же связи интересна работа Цирельсона (1979), в которой рассмотрен квантовый вариант локально-взаимодействующего марковского процесса и показано, что при некоторых ограничениях для него справедлива теорема о неэргодичности. Автор, однако, замечает, что о ее применимости к процессам молекулярной биологии он судить не берется.

Другое возможное решение проблемы устойчивости доминанты состоит в том, что в мозге, кроме долгоживущих микроочагов возбуждения, должна существовать фазовая следящая система, синхронизирующая эти микроочаги и вызывающая в них критический режим лишь на периоды времени, необходимые для фиксации или считывания информации из долговременной памяти. Здесь решение проблемы устойчивости можно уподобить равновесию движущегося велосипедиста: исходно неустойчивое состояние при определенных условиях оборачивается источником большого преимущества. Одна такая система, предложенная на основе фазовой зависимости эффекта сохранения пятна от периодического возмущения порогов нейронов (см. раздел 4.5), способна служить моделью внимания и кратковременной памяти и имеет достаточно обоснованное «нейронное обеспечение» [Крюков, 1982]. Есть основания считать, что на ее основе, используя результаты этой работы, удастся построить системную модель доминанты А.А.Ухтомского.

## 10. ПРИЛОЖЕНИЕ

### 10.0. Общая нейронная модель и ее уравнение Колмогорова

Предположим, что справедливы следующие допущения (рис. 10.1).

1°. Каждый нейрон сети возбуждается и генерирует импульс, когда его мембранный потенциал  $x_i^t$  достигает пороговой функции  $y_i^t > 0$ , которая в отсутствие импульсов затухает экспоненциально с коэффициентом затухания  $a_i$  до некоторой постоянной величины.

2°. На вход каждого нейрона поступает последовательность импульсов от других нейронов. Каждый импульс  $j$ -го нейрона скачком меняет мембранный потенциал  $i$ -го нейрона на величину  $a_{ij}$ , где  $\|a_{ij}\|$  — заданная квадратная  $N \times N$  матрица и  $N$  — общее количество нейронов сети.

3°. В отсутствие входных импульсов мембранный потенциал каждого нейрона затухает экспоненциально с показателем  $a_i$  до равновесного потенциала.

4°. После каждого возбуждения порог нейрона и его мембранный потенциал скачком принимают значения  $y_i^0$  и  $a_{ii}$ , соответственно ( $y_i^0 > a_{ii}$ ), после чего процесс накопления мембранных потенциала повторяется до следующего возбуждения. Значение  $y_i^0$  может быть случайной величиной; в таком случае задано стохастическое ядро  $a_i(z, y)$ , обозначающее плотность вероятности того, что  $y_i^0 = y$ , если скачок был из точки  $z$ .

5°. Входы  $i$ -го нейрона двух сортов: «соседние» и «шумовые». У первых  $|a_{ij}| \geq \Delta$ , где  $\Delta$  — некоторая постоянная; зато вторые — более многочисленные, так что сформированная ими доля мембранных потенциала, которую мы обозначим  $n_i^t$ , образует скачкообразный марковский процесс, не зависящий от  $n_j^t$ ,  $j \neq i$ .

Обозначим  $s_i^t$  — случайное время, прошедшее от момента последнего возбуждения  $i$ -го нейрона до момента  $t$ . Это — «возраст»  $i$ -го импульса или обратное время ожидания. Пусть  $N_i = \{j : |a_{ij}| \geq \Delta, s_j^t \leq s_i^t\}$ . Тогда в соответствии с принятыми допущениями мембранный потенциал  $x_i^t$  и порог  $y_i^t$   $i$ -го нейрона можно записать в следующем виде:

$$x_i^t = \sum_{j \in N_i} a_{ij} e^{-\theta_i s_i^t} + n_i^t, \quad y_i^t = y_i^0 e^{-a_i s_i^t}.$$

Фактически в  $x_i^t$  вносит вклад не только последний перед  $t$  импульс  $j$ -го нейрона, но, возможно, и предпоследний и т.д.; их можно легко учесть с помощью усложнения обозначений.

Из допущений видно, что однокомпонентный процесс  $(x_i^t, y_i^t, s_i^t)$  не обязан быть марковским. Однако он станет марковским, если зафиксируем состояния  $(s_i^t, y_i^t) = (s_j, y_j)$

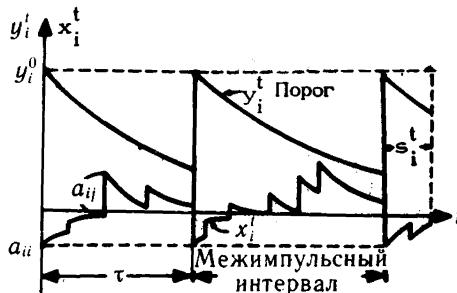
его «соседей»  $j \in N_i \setminus \{i\}$ . Это подсказывает взять в качестве состояний сети область значений процесса  $(s^t, y^t) \in [0, \infty)^N \times R^N$ , траектории которого можно описать следующим образом. В интервале между двумя последовательными возбуждениями  $i$ -го нейрона  $y_i^t$  затухает экспоненциально, а  $s_i^t$  растет с единичной скоростью. В момент возбуждения  $i$ -го нейрона его состояние скачком меняется из  $(s_i^t, y_i^t)$  в  $(0, y_i^0)$ . Отдельные компоненты  $(s_i^t, y_i^t)$  оказываются взаимозависимыми, так как момент появления скачка в  $i$ -й компоненте зависит не только от  $n_i^t$  и  $y_i^t$ , но и от вектора  $s^t$ , а в общем случае и от вектора  $(s^t, y^t)$ . Так как реально начальный порог  $y_i^0$  может быть довольно большим (имитируя так называемую абсолютную рефрактерность), то максимальная частота срабатывания нейрона ограничена. Присутствие шума  $n_i^t$  заставляет разные нейроны срабатывать несинхронно. Учитывая это, полагаем, что выполнено следующее предположение.

6°. Вероятность того, что во временном интервале длиной  $\Delta t$  срабатывает  $i$ -й нейрон ( $i=1, 2, \dots, N$ ), равна  $c_i(s, y)\Delta t + o(\Delta t)$ . Вероятность срабатывания в этом интервале двух и более нейронов равна  $o(\Delta t)$ .

Таким образом, эволюцией активности сети в малый интервал времени управляет система  $\{c_i(s, y)\}$ , которая при каждом  $i$  фактически зависит лишь от состояния  $(s_i, y_i)$  и состояния «соседних» нейронов  $N_i \setminus \{i\}$ . Это позволит нам использовать для анализа больших сетей теорию марковских процессов с локально взаимодействующими компонентами. Наша ближайшая задача выразить характеристический оператор  $L$  марковского процесса в терминах параметров модели. Тогда для любой разумной функции от состояния процесса, например для условного математического ожидания  $E_{s, y} h(s^t, y^t) = u_i(s, y)$ , мы можем записать обратное уравнение Колмогорова

$$\frac{\partial u_i}{\partial t} = L u_i,$$

Рис. 10.1. Схематическое изображение допущений модели



откуда можно в принципе получить любую информацию о динамике и стационарных состояниях сети.

Обозначим  $\tau_i$  — время первого достижения марковским процессом  $n_i^t$ , начинающимся из нуля, кривой  $r_i^t$ ,

$$r_i^t = y_i^t - \sum_{j \in N_i} a_{ij} e^{-a_j s_j}$$

и пусть

$$f_i(s_i|_i s, y) ds_i = P\{\tau_i \leq ds_i |_i s, y\}$$

$$F_i(s_i| \cdot) = \int_0^{s_i} f_i(z| \cdot) dz$$

$${}^u s = (s_1, \dots, s_{i-1}, u, s_{i+1}, \dots, s_N)$$

$${}_i s = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_N)$$

**Предложение 10.1** [Крюков, 1975]. Характеристический оператор марковского процесса  $(s^t, y^t)$  действует на непрерывно дифференцируемую функцию  $u(s, y): [0, \infty)^N \times R^N \rightarrow R^1$  по формуле

$$\begin{aligned} Lu(s, y) = \sum_{i=1}^N \left\{ \int_{R^i} c_i(s, y) a_i(y_i, z_i) [u({}^0_i s, {}^z_i y) - \right. \\ \left. - u(s, y)] dz_i + \frac{\partial u(s, y)}{\partial s_i} - a_i y_i \frac{\partial u(s, y)}{\partial y_i} \right\}, \end{aligned} \quad (10.1)$$

где

$$c_i(s, y) = f_i(s_i|_i s, y) (1 - F(s_i|_i s, y))^{-1}.$$

**Доказательство.** Пусть  $\tau_1$  — момент самого первого после момента  $t=0$  срабатывания какого-то нейрона в сети и пусть  $\tau_\epsilon = \min[\tau_1, \epsilon]$ . Пусть функция  $u(s, y)$  принадлежит области определения характеристического оператора  $L$ , то есть существует

$$\lim_{\epsilon \downarrow 0} \frac{1}{M_{s,y} \tau_\epsilon} [M_{s,y} u(s + \tau_\epsilon, y^{\tau_\epsilon}) - u(s, y)] = Lu(s, y),$$

где  $M_{(\cdot)}$  — условное математическое ожидание при условии  $(\cdot)$ . Для вычисления этого процесса мы вначале заметим, что

$$\begin{aligned} M_{s,y} u(s + \tau_\epsilon, y^{\tau_\epsilon}) = u(s + \epsilon, y e^{-\alpha \epsilon}) P_{s,y} \{\tau_1 > s\} + \\ + M_{s,y} u({}^0_i s, y^{\tau_1}) \chi_{\{\tau_1 < \epsilon\}}, \end{aligned}$$

где  $\chi_D$  — характеристическая функция множества  $D$  и

$$B_0 = P_{s,y} \{\tau_1 > \epsilon\} = \prod_{i=1}^N \frac{1 - F_i(s_i + \epsilon | \cdot)}{1 - F_i(s_i | \cdot)}.$$

Здесь числитель дроби равен вероятности того, что возраст

импульса  $i$ -го нейрона больше, чем  $s_i + \varepsilon$ ; а вся дробь равна условной вероятности того, что  $i$ -й нейрон не срабатывает в интервале  $(0, \varepsilon)$ , если  $(s^i, y^i) = (s, y)$  при  $t=0$ . Произведение таких вероятностей по всем  $i$  появляется благодаря условной независимости соответствующих событий в силу того, что траектории процесса  $(s^i, y^i) = (s+t, y \exp(-as))$  ведут себя детерминистически в интервале  $0 \leq t \leq \tau_1$ . Поэтому все  $x_i^i = n_i^i + \sum a_{ij} e^{-\beta_j(s_i + t)}$

взаимно независимы и, следовательно, независимы все времена первого пересечения кривой  $y^i$  процессом  $x_i^i$  или кривой  $r^i = y^i - \sum a_{ij} e^{-\beta_j s_i}$  процессом  $n_i^i$ . Короче, никакого взаимодействия в сети не возникает до тех пор, пока не возбудится хотя бы один нейрон. Похожие рассуждения могут убедить

в том, что  $P_{s,y}\{\tau_1 \leq \varepsilon\} = \sum_{i=1}^N B_i$ ,

$$B_i = \frac{F_i(s_i + \varepsilon | \cdot) - F_i(s_i | \cdot)}{1 - F_i(s_i | \cdot)} \prod_{i \neq j=1}^N \frac{1 - F_j(s_j + \varepsilon | \cdot)}{1 - F_j(s_j | \cdot)},$$

где  $B_i$  — условная вероятность срабатывания  $i$ -го нейрона в интервале  $(0, \varepsilon)$ , при начальном состоянии  $(s, y)$  и несрабатывания всех других нейронов в том же самом интервале. Такие события при различных  $i$  взаимно независимы и все вместе образуют событие  $\{\tau_1 \leq \varepsilon\}$ . Теперь мы возвращаемся к вычислению предела, предварительно приняв без доказательства, что  $M_{s,y,\tau_\varepsilon}$  асимптотически равно  $\varepsilon$  при малых  $\varepsilon$ . Тогда

$$\begin{aligned} Lu(s, y) &= \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \left\{ u(s + \varepsilon, y e^{-a_i \varepsilon}) B_0 + \right. \\ &\quad \left. + \sum_{i=1}^N \left( \int_{R^i} u(\overset{0}{i}(s+\varepsilon), \overset{z_i}{i}(y e^{-a_i \varepsilon})) B_i a_i(y_i e^{-a_i \varepsilon}, z_i) dz_i \right) - u(s, y) \right\}. \end{aligned}$$

Используя здесь очевидные соотношения

$$B_0 = 1 - \sum_{i=1}^N B_i \quad \text{и} \quad \frac{B_i}{\varepsilon} \rightarrow \frac{f_i(s_i | \cdot)}{1 - F_i(s_i | \cdot)}$$

при  $\varepsilon \downarrow 0$ , мы получим

$$\begin{aligned} Lu(s, y) &= \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \{ u(s + \varepsilon, y e^{-a_i \varepsilon}) - u(s, y) \} + \\ &\quad + \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} \left\{ \sum_{i=1}^N \int_{R^i} (u(\overset{0}{i}(s+\varepsilon), \overset{z_i}{i}(y e^{-a_i \varepsilon})) - u(s + \varepsilon, y e^{-a_i \varepsilon})) \times \right. \\ &\quad \left. \times B_i a_i(y_i e^{-a_i \varepsilon}, z_i) dz_i \right\} = \lim_{\varepsilon \downarrow 0} \frac{1}{\varepsilon} u(s + \varepsilon, y e^{-a_i \varepsilon}) + \\ &\quad + \sum_{i=1}^N \int_{R^i} \{ u(\overset{0}{i}s, \overset{z_i}{i}y) - u(s, y) \} c_i(s, y) a_i(y_i, z_i) dz_i. \end{aligned}$$

Добавляя и вычитая  $u(s, ye^{-a_i s})$  под знаком второго предела, мы при  $\varepsilon \downarrow 0$  получим (10.1). Осталось только доказать, что  $M_{s,y\tau_\varepsilon} \sim \varepsilon$ . Так как

$$M_{s,y\tau_\varepsilon} = \varepsilon P_{s,y}\{\tau_1 > \varepsilon\} + \int_0^\varepsilon t d_t P\{\tau_1 \leq t\} \leq \varepsilon P_{s,y}\{\tau_1 > \varepsilon\} + \varepsilon P_{s,y}\{\tau_1 \leq \varepsilon\},$$

то  $\varepsilon P_{s,y}\{\tau_1 > \varepsilon\} \leq M_{s,y\tau_\varepsilon} \leq \varepsilon$ . Откуда и получаем требуемое, поскольку  $P_{s,y}\{\tau_1 > \varepsilon\} \rightarrow 1$  при  $\varepsilon \downarrow 0$ . Доказательство закончено.

В работе Крюкова (1977) подробно рассмотрены три наиболее важных частных случая оператора (10.1). Здесь мы укажем лишь, что для базовой нейронной модели, кратко описанной в разделе 2.1, плотность  $a_i(z_i, y_i) = \delta(y_i - r_i)$  для всех  $z_i < r_i = \text{const}$ , где  $\delta(x)$  — дельта-функция. Тогда входящие в (10.1) переменные  $s$  и  $y$  оказываются функционально связанными уравнением  $y_i^t = r_i e^{-a_i s_i^t}$  и оператор (10.1) приобретает простой вид, записанный в правой части уравнения (2.1).

## 10.1. Основные определения и общие сведения о гиббсовских случайных полях. Применимость к нейрофизиологии

Одним из фундаментальных понятий статистической физики является канонический ансамбль, описываемый распределением Гиббса:

$$P(x) = Z^{-1} e^{H(x)}, \quad (10.2)$$

где  $H(x)$  — скалярная функция (энергия, потенциал) и  $Z$  — константа. Оно связывает статистическую механику и термодинамику, но применимо к фазовым пространствам любой природы, способно описывать критические явления типа фазовых переходов и, самое главное, имеет структуру функции распределенных независимых компонент:  $H(x)$  равна сумме потенциалов всех взаимодействующих частей системы. Естественно, интересно выяснить, пригодно ли оно для анализа нейронной активности. Это приводит к более общему вопросу: насколько широк класс стохастических систем, описываемых этим распределением. Достаточно полный ответ был получен Аверинцевым (1970, 1972): гиббсовские ансамбли практически совпадают с марковскими случайными полями.

Понятие марковского случайного поля ввел Добрушин (1968а) как обобщение марковской цепи на случай, когда параметр времени становится многомерным и понимается, например, как пространственная координата компоненты. Основная задача теории случайных полей — установить связь локальных и глобальных способов описания поля, обеспечивающих существование, единственность, перемешиваемость и другие свойства поля. Теория тесно связана с новыми методами многомерного статистического анализа [Besag, 1974], но известна прежде

всего своими приложениями к задачам статистической физики.

Суть дела в том, что обычное гиббсовское распределение с попарным потенциалом и конечным радиусом взаимодействия можно рассматривать как частный случай меры, описывающей марковское поле. Отсюда следуют свойства его разнообразных обобщений на случай произвольных фазовых пространств, зависимости от граничных условий, многочастичного потенциала взаимодействия, бесконечного количества компонент, возможной неэргодичности, отсутствия трансляционной инвариантности, нестационарности во времени (процессы) и т.д. На этом пути был получен ряд интересных результатов: о существовании термодинамических пределов, об условиях отсутствия и наличия фазовых переходов, доказан принцип максимума бесконечных систем, об условиях, при которых стационарное распределение модели Изинга становится гиббсовским [Добрушин, 1968б, в; 1969, 1971а, б]. С другой стороны, статистическая физика дает много примеров нетривиального поведения конкретных марковских полей и подсказывает постановку новых задач. Для их обсуждения введем необходимые определения.

*Случайным полем* называется система случайных величин  $x = (x_1, \dots, x_n)$  с заданными условными вероятностями  $P(x_i|\text{все остальные})$ , удовлетворяющими естественным условиям согласованности. Компонента  $j (\neq i)$  называется *соседом* компоненты  $i$  тогда и только тогда, когда  $P(x_i|\cdot)$  зависит от  $x_j$ . Случайное поле, для которого  $P(x_i|\text{все остальные } x_j) = P(x_i|\text{все соседи } x_i)$ , называется *марковским случайным полем*. Множество компонент, которое состоит из одной компоненты, либо все компоненты которого соседи, назовем *цепочкой* («клика»). *Гиббсовским ансамблем* называется случайное поле, совместная вероятность состояний компонент которого задается формулой (10.2), где

$$H(x) - H(0) = \sum_{1 \leq i \leq n} x_i a_i(x_i) + \sum_{1 \leq i < j \leq n} x_i x_j a_{ij}(x_i, x_j) + \\ + \sum_{1 \leq i < j < k \leq n} x_i x_j x_k a_{ijk}(x_i, x_j, x_k) + \dots + x_1 \dots x_n a_{1 \dots n}(x_1, \dots, x_n) \quad (10.3)$$

**Теорема Аверинцева.** Для того, чтобы случайное поле было марковским с условными вероятностями

$$P(x_i|\cdot) > 0$$

необходимо и достаточно, чтобы оно было гиббсовским ансамблем. Соответствие между наборами  $\{P(x_i|\cdot)\}$  и  $\{a_{ij\dots k}\}$  взаимно однозначное, если

$$a_{ij\dots k}(x_i, x_j, \dots, x_k) \neq 0 \Leftrightarrow i, j, \dots, k — \text{цепочка}. \quad (10.4)$$

**Доказательство.** Мы приведем наиболее простое доказательство, предложенное Бесагом [Besag, 1974], чтобы заодно ввести следующее соотношение, которое потребуется нам в дальнейшем:

$$\frac{P(x_i|x)}{P(0|x)} = e^{H(x) - H(^0x)}, \quad (10.5)$$

где  $x = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$   
 $^0x = (x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n)$ .

Имеем

$$\begin{aligned} P(x) &= P(x_n|x_1, \dots, x_{n-1})P(x_1, \dots, x_{n-1}) = \\ &= \frac{P(x_n|x_1, \dots, x_{n-1})}{P(y_n|x_1, \dots, x_{n-1})} P(x_1, \dots, x_{n-1}, y_n). \end{aligned}$$

Аналогично имеем

$$P(x_1, \dots, x_{n-1}, y_n) = \frac{P(x_{n-1}|x_1, \dots, x_{n-2}, y_n)}{P(y_{n-1}|x_1, \dots, x_{n-2}, y_n)} P(x_1, \dots, x_{n-2}, y_{n-1}, y_n).$$

Продолжая дальше такие же преобразования, получим

$$\frac{P(x)}{P(y)} = \prod_{i=1}^n \frac{P(x_i|x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_n)}{P(y_i|x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_n)}.$$

Отсюда, обозначая

$$\begin{aligned} ^i x &= (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n), \\ \frac{P(x)P(y)}{P(y)P(^i x)} &= \frac{P(x_i|x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_n)}{P(y_i|x_1, \dots, x_{i-1}, y_{i+1}, \dots, y_n)}. \end{aligned}$$

Отсюда согласно (10.2) следует (10.5), так как значения  $y_i$  могут быть любыми, в том числе  $y_i = 0$ ,  $y_{i+k} = x_{i+k}$ ,  $k = 1, 2, \dots, n$ . Дальнейший ход рассуждения такой: из (10.2), (10.3) и  $a_{ij \dots k} \neq 0$  следует, что правая и левая части (10.5) зависят от  $j, \dots, k$ . Следовательно,  $i, j, \dots, k$  — соседи, то есть входят в марковскую цепочку. Наоборот, если  $x$  — марковское поле, но компоненты  $i$  и  $l$  — не соседи, то, следовательно, левая, а значит и правая части (10.5) не зависят от  $l$ . Значит (10.3) дает  $a_{ii} = 0$ , так как все его слагаемые за исключением  $x_i x_i a_{ii}$  обращаются в ноль, если положить  $x_k = 0$  для всех  $k$ , кроме  $i$  и  $l$ . Завершается доказательство аналогичными рассуждениями для остальных  $a_{ij \dots k}$ .

Теорема Аверинцева гласит, что распределение Гиббса является удобной универсальной («канонической») формой распределения для марковских случайных полей и указывает, что причина этого — локальный характер взаимодействия элементов: чем сильнее выражены локальные свойства поля, тем меньшим количеством типов слагаемых можно ограничиться в выражении (10.3). Но для установления однозначной связи чисто вероятностного объекта с физическими потенциалами потребовалось обобщить понятие гиббсовских ансамблей сразу в двух направлениях: а) включить в него иерархию нелинейных

взаимодействий и б) ввести граничные условия. Поясним это.

а). Формулы (10.2), (10.3) и (10.4) показывают, что элементарная «рабочая ячейка» марковского поля — не отдельная компонента и не пара компонент, а в общем случае цепочка: достаточно (и необходимо) хотя бы одной компоненте не быть соседом остальных в цепочке, чтобы результирующий потенциал цепочки обратился в ноль. Следовательно, марковское поле описывается парными потенциалами  $\{x_i x_j a_{ij}(x_i, x_j)\}$  тогда и только тогда, когда граф, соединяющий всех соседей, не содержит треугольников. В общем случае допустимо большое разнообразие типов взаимодействия: от полной независимости (например, «сеть—газ») до полностью взаимозависимой системы из одной цепочки (например, сеть типа «все-на всех»). Таким образом, в случае положительных условных вероятностей, цепочка совпадает с областью определения (то есть является носителем) многочастичного потенциала взаимодействия.

б). По формуле условной вероятности из теоремы Аверинцева следует, что

$$P(x_i, x_j, \dots, x_l / \text{все остальные}) = P(x_i, x_j, \dots, x_l / \text{соседи } i, j, \dots, l).$$

То есть любое подполе марковского поля будет снова марковским и, следовательно, гиббсовским при условии (10.4). (Это, между прочим, — обобщение основного свойства гиббсовского ансамбля со слабым взаимодействием: его малая компонента будет тоже гиббсовской с той же температурой). Теперь зависимость от соседей удобно интерпретировать как граничное условие, а само распределение Гиббса — как условную вероятность. Такой подход, обобщенный на случай бесконечного количества компонент, был успешно применен к физическим системам в неограниченных «сосудах» [Добрушин, 1968а, б, в]; он применим и к открытым (энергетически или информационно) гиббсовским ансамблям.

Возвращаясь к случаю попарных взаимодействий компонент отметим один специальный результат, вытекающий из теоремы Аверинцева.

**Лемма [Besag, 1974].** *Если вероятностная структура системы зависит только от цепочек, содержащих не более чем две компоненты, и условная вероятность, связанная с каждой из них, может быть представлена в виде*

$$P_i(x_i | \cdot) = A_i(x_i) B_i(\cdot) \exp(C_i(x_i) D_i(\cdot)),$$

тогда

$$D_i(x) = b_i + \sum_{j=1}^N b_{ij} [C_j(x_j) - C_j(0)],$$

где  $b_{ij} = b_{ji}$  и  $b_{ii} = 0$ , если компоненты с номерами  $i$  и  $j$  — не соседи.

В теореме Аверинцева утверждается эквивалентность ло-

кальных и глобальных свойств марковского поля. Поэтому система условных вероятностей, задаваемых по формуле

$$\frac{P(x_i|x)}{P(0|x)} = e^{x_i a_i(x_i) + \sum_{i < j} x_i x_j a_{ij}(x_i, x_j) + \dots}, \quad (10.6)$$

которая следует из (10.3) и (10.5), столько же информативна, как и безусловная вероятность (10.2). Но система (10.6) гораздо удобнее для статистического анализа данных, чем (10.2), хотя бы потому, что для нее не нужны громоздкие вычисления постоянной  $Z$  (что было бы необходимо для применения метода максимального правдоподобия). Локальность взаимодействия позволяет ограничиться вначале очень грубой статистической моделью, содержащей, например, лишь одиночные и парные потенциалы, а затем провести, если необходимо, последовательное уточнение. В этом суть нового метода статистического анализа, который разработал Бесаг (1974). Он продемонстрировал также ряд других преимуществ своего метода перед традиционным многомерным анализом. Чтобы применить его метод к нейронной активности, конечно, не обязательно, чтобы потенциалы взаимодействия имели физиологический смысл, однако понимание объекта анализа существенно меняется, если построить модель нейронной сети с гиббсовским распределением наблюдаемых переменных.

При построении такой модели сети желательно наделить ее импульсную активность свойствами марковского поля и в то же время учесть основные свойства реальных нейронов: временную сумму приходящего возбуждения, его экспоненциальное затухание и относительную невозбудимость (рефрактерность) после каждого достижения порогового уровня возбуждения. Но выполняются ли при этом условия теоремы Аверинцева?

Основное условие теоремы  $\{P(x_i|x) > 0\}$  требует, чтобы при фиксированных состояниях соседей поведение нейрона было достаточно случайным. Экспериментальные наблюдения во многих случаях не противоречат этому требованию в силу известных шумовых свойств синаптических контактов и постсинаптических мембран. Однако есть группы нейронов с достаточно регулярной, почти детерминированной импульсной активностью. Всегда есть также ограничения, связанные с абсолютной и относительной рефрактерностью.

Но самая главная трудность в том, что нейронная активность по своей природе — процесс динамический, развивающийся во времени под действием внешних сигналов и внутренних шумов. Его стационарное распределение, если и существует, то не обязательно с локальным взаимодействием. Правда, имеется класс так называемых обратимых марковских процессов и цепей, довольно естественный для моделей ста-

тистической физики, который порождает стационарные марковские поля [Добрушин, 1971б; Ставская, 1973; Preston, 1973]. Однако, на наш взгляд, нет ни логических, ни интуитивных оснований для его применения к сетям из неформальных нейронов. Тем не менее дело обстоит не так уж плохо. С помощью условия, приведенного в следующем разделе этого приложения, можно показать, что класс марковских процессов, имеющий распределение Гиббса в качестве инвариантной меры, значительно шире, чем считалось до сих пор.

## 10.2. Обратимые и псевдообратимые марковские процессы

Марковский процесс  $\xi_t, 0 \leq t < \infty$  со значениями в некотором измеримом пространстве  $X$  называется *обратимым*, если при любых  $t, t, 0 \leq \tau < t < \infty$  и любых измеримых  $E \subset X, D \subset X$  вероятность

$$P(\xi_\tau \in E, \xi_t \in D) = P(\xi_t \in E, \xi_\tau \in D), \quad (10.7)$$

то есть если статистические свойства процесса не меняются при изменении направления хода времени. Распределение вероятностей  $P(\xi_t \in C)$  будет при этом стационарным. Чтобы написать это условие в терминах переходного оператора, рассмотрим скачкообразный марковский процесс на  $X = \mathbb{Z}^N$ , вероятности состояний которого эволюционируют в соответствии с уравнением

$$\frac{dP_x(t)}{dt} = \sum_{y \in X} [P_y(t)\lambda_{yx} - P_x(t)\lambda_{xy}], \quad (10.8)$$

где  $\lambda_{xy}\Delta t$  — вероятность перехода скачком из состояния  $x$  в состояние  $y$  в интервале  $(t, t + \Delta t)$ . Тогда по определению обратимости (10.7)

$$P_y(t)\lambda_{yx} - P_x(t)\lambda_{xy} = 0.$$

Заметим, что это же уравнение возникает, если приравнять нулю правую часть (10.8) и потребовать дополнительно, чтобы при этом в ноль обращалось каждое слагаемое под знаком суммы. Отсюда ясно, что условие обратимости — более сильное ограничение на процесс, чем условие существования стационарного распределения: в отличие от последнего оно требует не общего баланса притока и оттока вероятностей, а «детального баланса», то есть равновесия для каждой пары состояний.

Оказывается, что для многокомпонентных систем со взаимодействием имеет смысл условие, промежуточное между обратимостью и стационарностью. Мы его назвали *псевдообратимостью*. Прежде чем дать формальное определение, поясним его интуитивный смысл на простом примере. Для этого предположим, что марковский процесс, описываемый уравнением

(10.8), в каждый данный момент может изменить свое состояние лишь в какой-то одной компоненте, так что из состояния  $x \in \mathbb{Z}^N$  он может перейти за короткое время лишь в состояние  $y = {}^b_i x$ , то есть

$$\lambda_{xy} = c_i(x) a_i(x_i, b),$$

где  $c_i(x)$  — интенсивность скачков из состояния  $x$ ,  $a_i(x_i, y_i)$  — вероятность перехода  $i$ -й компоненты из состояния  $x_i$  в состояние  $y_i$ . Тогда уравнение (10.8) и связанные с ним три упомянутых условия можно записать следующим образом

$$\frac{dP_x(t)}{dt} = \sum_{i=1}^N \sum_{y \in \mathbb{Z}} [P_{{}^b_i x} c_i({}^b_i x) a_i(b, x_i) - P_x c_i(x) a_i(x_i, b)]$$

= 0 — обратимость

= 0 — псевдообратимость

= 0 — стационарность.

Приравнивая нуль выражение под знаком двойной суммы, затем только внутреннюю сумму и, наконец, всю двойную сумму, получим соответственно условие обратимости, псевдообратимости и стационарности. Очевидно, что два последних условия совпадают для однокомпонентной системы. Интуитивно псевдообратимость означает стационарность по каждой компоненте при условии, что состояния ее соседей фиксированы. Чтобы распространить это условие на процессы, возникающие в нейронных сетях, придадим ему следующую форму. Марковский процесс с характеристическим оператором

$$Lu^t(x) = \sum_{i=1}^N L_i u^t(x), \quad u^t(x) \in C_{t,x}^1,$$

называется псевдообратимым, если существует такая плотность вероятностей  $\mu$ , что математическое ожидание от  $L_i u^t$  по отношению к  $\mu$  равно нулю. В операторной форме это запишем так:

$$\mu L_i u^t = 0, \tag{10.9}$$

откуда в силу произвольности функции  $u^t(x)$  видно, что явный вид условия псевдообратимости совпадает с уравнением для инвариантной плотности  $i$ -й компоненты при фиксированных состояниях ее соседей. В частности, для базовой нейронной модели, описанной в разделе 2.1, условие (10.9) сводится к уравнению (2.2).

### 10.3. Определения и общие сведения о модели Изинга с непрерывным пространством состояний

Модель Изинга с непрерывным спином есть тройка  $(L, H, v)$ , где

- 1)  $L$  — множество мест, счетное множество, обычно  $L = \mathbb{Z}^n$ .

С каждым местом  $i \in \mathbb{Z}$  ассоциирован спин  $\sigma_i \in R$  и произведение  $R^L = \prod_{i \in L} R$  называется конфигурационным пространством.

2)  $H$  — гамильтониан, формальный полином по  $\sigma_i$ . Если  $H$  имеет неположительные коэффициенты, то говорят, что соответствующая модель Изинга описывает ферромагнетик. В последнем случае пишут

$$H = - \sum_{K \in \mathcal{F}_c(L)} J_K \sigma_K, \quad J_K \geq 0, \quad \sigma_K = \prod_{i \in K} \sigma_i,$$

где  $\mathcal{F}_c(L)$  — конечное подмножество в  $L$ .

3)  $v$  — мера одиночного спина, четная борелевская вероятностная мера на  $R$ , затухающая достаточно быстро на бесконечности.

Говорят, что модель Изинга трансляционно инвариантна, если

$$J_K = J_{K+1}, \quad \forall K \in \mathcal{F}_c(\mathbb{Z}^n), \quad \forall i \in \mathbb{Z}^n.$$

Гиббсовская мера  $\mu$  конечного ферромагнетика Изинга ( $\Lambda$ ,  $H$ ,  $v$ ) с обратной температурой  $\beta \in [0, \infty)$  — это вероятностная мера на конфигурациях  $R^\Lambda$ , определяемая формулой

$$\mu(E) = Z^{-1} \int_E \exp[-\beta H(\sigma)] \prod_{i \in \Lambda} dv(\sigma_i), \quad E \subset R^\Lambda, \quad (10.10)$$

где  $Z$  — нормирующий множитель.

Физически  $\Lambda$  можно интерпретировать как положение атомов в кристалле или положение нейронов в нейронной структуре, а  $\sigma_i$  — как состояние отдельной  $i$ -й компоненты (спин атома или возраст импульса нейрона), так что точка  $\sigma$  в конфигурационном пространстве  $R^\Lambda$  соответствует состоянию системы, а  $H(\sigma)$  — «энергия» этого состояния, причем  $\mu(E)$  — вероятность нахождения системы в некотором подмножестве  $E \subset R^\Lambda$ , связанная с обменом энергии с термостатом или с другими частями нейронной системы.

В физике  $v$  — температуронезависимый вес, определяемый свойствами атома, причем  $v$  — симметричная функция  $\sigma_i$ . В нейронной системе  $v$  зависит от температуры, то есть зависит как от внутренних шумов нейрона, так и от шумовой компоненты, обусловленной действием окружающих нейронных структур. Кроме того, здесь  $v$  несимметрична, хотя обычно можно подобрать замену переменных по  $\sigma_i$ , чтобы  $v$  сделалась симметричной. Дальнейшее изложение справедливо для симметричных  $v$ .

В бесконечной системе  $\{L, H, v\}$  нельзя воспользоваться формулой (10.10), чтобы определить математическое ожидание

$$\langle \sigma_K \rangle = \int_{R^\Lambda} \sigma_K e^{-\beta H} \prod_{i \in \Lambda} dv,$$

так как  $H$  — формальный полином по бесконечному множеству переменных, не имеющий смысла как функция. Поэтому возникает следующая предельная конструкция. Пусть  $\Lambda \subset L$  —

конечное подмножество. Положим все спины вне  $\Lambda$  нулями и пусть

$$H_\Lambda = - \sum J_k \sigma_k, \quad K \in \mathcal{F}_c(\Lambda),$$

то есть  $H_\Lambda$  есть сужение  $H$  на  $\Lambda$ . Если  $A \in \mathcal{F}_c(\Lambda)$ , то используя (10.10), определим приближенные моменты:

$$\langle \sigma_A \rangle_\Lambda = \langle \sigma_A; H_\Lambda, v, \beta \rangle.$$

Но для конечной системы с симметричной мерой одиночного спина, ввиду того что  $J_k \geq 0$ , справедливы неравенства Гриффитса

$$\langle \sigma_A \rangle \geq 0, \quad A \in \mathcal{F}_c(\Lambda)$$

$$\beta^{-1} \frac{\partial}{\partial J_B} \langle \sigma_A \rangle = \langle \sigma_A \sigma_B \rangle - \langle \sigma_A \rangle \langle \sigma_B \rangle \geq 0. \quad (10.11)$$

А поэтому  $\langle \sigma_A \rangle$  образуют возрастающую последовательность, пронумерованную индексами конечного подмножества  $\Lambda \subset L$ , и в том случае, когда она ограничена сверху, имеет смысл предел

$$\langle \sigma_A; H, v, \beta \rangle = \lim_{\Lambda \rightarrow \infty} \langle \sigma_A, H_\Lambda, v, \beta \rangle.$$

Рассмотрим ферромагнит  $(\mathbb{Z}^n, H, v)$  с трансляционно инвариантным парным потенциалом и внешним полем  $h$ :

$$H = - \sum_{i,j \in \mathbb{Z}^n} J_{(i-j)} \sigma_i \sigma_j - h \sum_{i \in \mathbb{Z}^n} \sigma_i, \quad J_k = 0 \text{ для } k > k_0 \geq 1.$$

В силу (10.11) математическое ожидание  $\langle \sigma_A \rangle$  — монотонно возрастающая функция поля  $h$  и вообще всех коэффициентов связи. Когда  $h$  исчезает, гамильтониан  $H$  инвариантен под действием одновременного обращения всех спинов  $\sigma \rightarrow -\sigma$ , так что нечетные моменты  $\left\langle \prod_{a=1}^{2n+1} \sigma_{ia} \right\rangle = 0$ . Модель  $(\mathbb{Z}^n, H, v)$  называется *спонтанно-намагничивающейся* при нулевом магнитном поле, если  $\langle \sigma_i \rangle$  — разрывная по  $h$  при  $h=0$

$$\lim_{h \downarrow 0} \langle \sigma_i \rangle = m_s > 0.$$

Такое поведение невозможно для конечных моделей, но для бесконечных систем с  $n \geq 2$  — это факт (см. ниже) при достаточно низких температурах. Тесно связано с этим явление *далекого порядка*. Модель  $(\mathbb{Z}^n, H, v)$  имеет далекий порядок при нулевом магнитном поле, если существует  $C > 0$ , так что

$$\langle \sigma_i \sigma_j \rangle|_{h=0} \geq C, \quad \forall i, j \in \mathbb{Z}^n.$$

**Теорема 10.1** [Beijeren, Sylvester, 1978].

Пусть  $(\mathbb{Z}^n, H, v)$  — изинговский ферромагнит в размерности

$n \geq 2$  с непрерывным спином и с парным взаимодействием между ближайшими соседними спинами, где

$$H = - \sum_{k \in \mathbb{Z}^n} \sum_{a=1}^n J_a \sigma_k \sigma_{k+1_a}, \quad J_a > 0, \quad 1_a = (0, \dots, 0, 1, \overbrace{0, \dots, 0}^{n-a}),$$

и четная односиновая мера  $\nu$  не есть  $\delta$ -функция. Если температура  $\beta^{-1}$  достаточно мала, то модель имеет далекий порядок.

**Теорема 10.2** [Веүеген, Sylvester, 1978].

Пусть  $(\mathbb{Z}^n, H, \nu)$  — изинговский ферромагнит с непрерывным спином в размерности  $n \geq 2$  и трансляционно-инвариантным парным гамильтонианом

$$H = - \sum_{l, k \in \mathbb{Z}^n} J_{(k-l)} \sigma_k \sigma_l - h \sum \sigma_k, \quad J_k, h \geq 0,$$

причем  $\nu$  — четная и не дельта-функция. Если температура  $\beta^{-1}$  достаточно низка, то модель спонтанно намагничивается при  $h=0$ .

#### 10.4. Качественное исследование уравнений динамики среднего уровня и дисперсии активности нейронной сети

Предположим, что динамика активности однородной нейронной сети описывается двумя макрохарактеристиками:

- 1)  $E(t)$  — средний уровень активности сети в момент  $t$ ;
- 2)  $V(t)$  — дисперсия среднего уровня активности сети в момент  $t$ .

Как было показано в разделе 3, система дифференциальных уравнений для этих макрохарактеристик имеет вид:

$$\begin{cases} \frac{dE}{dt} = -E + S(aE) + \frac{S''(aE)}{2} a^2 V \\ \frac{dV}{dt} = -2V(1 - aS'(aE)), \end{cases} \quad (10.12)$$

где  $a$  — параметр, характеризующий величину связи между элементами сети;  $S(x)$  — функция, называемая иногда откликом нейронной сети на поступившее воздействие  $x$ . Как показано в разделе 3, графиком этой функции является логистическая кривая. Для дальнейшего исследования нам будет удобно взять функцию  $S(x)$  следующего вида:

$$S(x) = \frac{1}{1 + \exp[-b(x - \Theta)]} - \frac{1}{1 + \exp[b\Theta]}.$$

Функция  $S(x)$  зависит от двух параметров  $b$  и  $\Theta$ , величины которых не влияют на изменение качественной картины поведения сети. Поэтому мы будем считать значения этих параметров фиксированными и изучать перестройки динамического поведения при изменении лишь одного параметра  $a$ .

Заметим, что функция  $S(x)$  выбрана таким образом, что точка с координатами  $E=0, V=0$  является стационарным решением системы при всех значениях параметров. В силу это-

го мы будем рассматривать и отрицательные значения переменной  $E$ , считая, что они соответствуют средней активности сети, которая ниже фиксированного стационарного уровня.

Перечислим свойства функции  $S(x)$ , используемые при дальнейшем исследовании:

- a)  $S(0) = 0$ ;
- б)  $S(\infty) = \frac{\exp[b\Theta]}{1 + \exp[b\Theta]}$ ;
- в)  $S'(x) = \frac{b \exp[-bx + b\Theta]}{(1 + \exp[-bx + b\Theta])^2}$ ;
- г)  $S''(x) = \frac{b^2 \exp[-bx + b\Theta] (\exp[-bx + b\Theta] - 1)}{(1 + \exp[-bx + b\Theta])^3}$ ;
- д)  $\max_x S'(ax) = \frac{b}{4}$ , при  $x = \frac{\Theta}{a}$ .

Найдем *стационарные решения* системы (10.12). Для этого рассмотрим систему нелинейных уравнений:

$$\begin{cases} -E + S(aE) + \frac{S''(aE)}{2} a^2 V = 0 \\ 2V(1 - aS'(aE)) = 0. \end{cases}$$

Эта система эквивалентна двум системам уравнений, дающим две серии решений:

$$\begin{cases} V = 0 \\ S(aE) = 0. \end{cases} \quad (10.13)$$

$$\begin{cases} aS'(aE) = 1 \\ V = 2 \frac{E - S(aE)}{a^2 S''(aE)}. \end{cases} \quad (10.14)$$

При анализе этих систем удобна геометрическая интерпретация. Решениями второго уравнения системы (10.13) являются точки пересечения графиков функций  $f(E) = S(aE)$  и  $f(E) = E$ . Решениями первого уравнения системы (10.14) являются такие значения  $E$ , при которых касательные к графику функции  $f(E) = S(aE)$  параллельны прямой  $f(E) = E$ .

Для определения устойчивости стационарных решений рассмотрим матрицу линеаризации системы (10.12).

$$A = \begin{pmatrix} -1 + aS'(aE) + \frac{S'''(aE)}{2} a^3 V & \frac{S''(aE)a^2}{2} \\ 2Va^2S''(aE) & 2(-1 + aS'(aE)) \end{pmatrix}$$

Для стационарных решений, определяемых системой (10.13), то есть решений вида  $(E_0, 0)$ , матрица имеет вид:

$$A = \begin{pmatrix} -1 + aS'(aE_0) & \frac{S''(aE_0)a^2}{2} \\ 0 & 2(-1 + aS'(aE_0)) \end{pmatrix}.$$

Отсюда находим собственные значения

$$\lambda_1 = -1 + aS'(aE_0), \quad \lambda_2 = 2\lambda_1.$$

Для стационарных решений, определяемых системой (10.14), матрица  $A$  имеет вид:

$$A = \begin{pmatrix} \frac{aS'''(aE)(E-S(aE))}{S''(aE)} & \frac{S''(aE)a^2}{2} \\ 4(E-S(aE)) & 0 \end{pmatrix}.$$

Определитель этой матрицы  $\Delta = -2(E-S(aE))a^2S''(aE)$  позволяет судить об устойчивости стационарного решения (произведение собственных значений равно определителю  $\Delta$ ).

Будем увеличивать значения параметра  $a$ , начиная с  $a=0$ , и перечислим все качественные перестройки в структуре фазовых портретов, которые при этом происходят. В соответствии с этими перестройками весь диапазон изменения параметра  $a$  разбивается на *четыре участка* ( $b$  и  $\Theta$  фиксированы). Для каждого из этих случаев построим соответствующий фазовый портрет (рис. 10.2).

**Случай 1.** В этом случае существует единственное стационарное состояние (устойчивое)  $E=0, V=0$ , к которому эволюционируют все траектории системы (рис. 10.2а). При  $a=a_1$  происходит бифуркация и из бесконечности рождаются два стационарных состояния:  $(E_\text{l}, V_\text{l})$  и  $(E_\text{n}, V_\text{n})$ , причем  $V_\text{l} \xrightarrow[a \uparrow a_1]{\longrightarrow} +\infty$ ,

а  $V_\text{n} \xrightarrow[a \uparrow a_1]{\longrightarrow} -\infty$ . (В момент бифуркации касательная к графику  $f(E)=S(aE)$  становится параллельной прямой  $f(E)=E$ ).

**Случай 2.** В этом случае в полуплоскости  $V>0$  существует неустойчивое (седловое) стационарное решение (рис. 10.2в). При  $a=a_2$  происходит сложная бифуркация, в результате которой рождаются три стационарных состояния: на оси  $V=0$  — устойчивое и неустойчивое состояния и в полуплоскости  $V>0$  — седловой стационар. (В момент бифуркации происходит касание графиков функций  $f(E)=S(aE)$  и  $f(E)=E$ ).

**Случай 3.** В этом случае в полуплоскости  $V \geq 0$  существует пять стационарных состояний (рис. 10.2с): два устойчивых и одно неустойчивое на оси  $V=0$  и два седловых при  $V>0$ . При  $a=a_3$  происходит бифуркация слияния седла  $(E_\text{l}, V_\text{l})$  с устойчивым узлом  $(E=0, V=0)$  и неустойчивым узлом на оси  $E$ , при увеличении  $a$  сложная точка опять распадается на три стационарных решения. (В момент бифуркации происходит касание графиков функций  $f(E)=S(aE)$  и  $f(E)=E$  в точке  $E=0$ ).

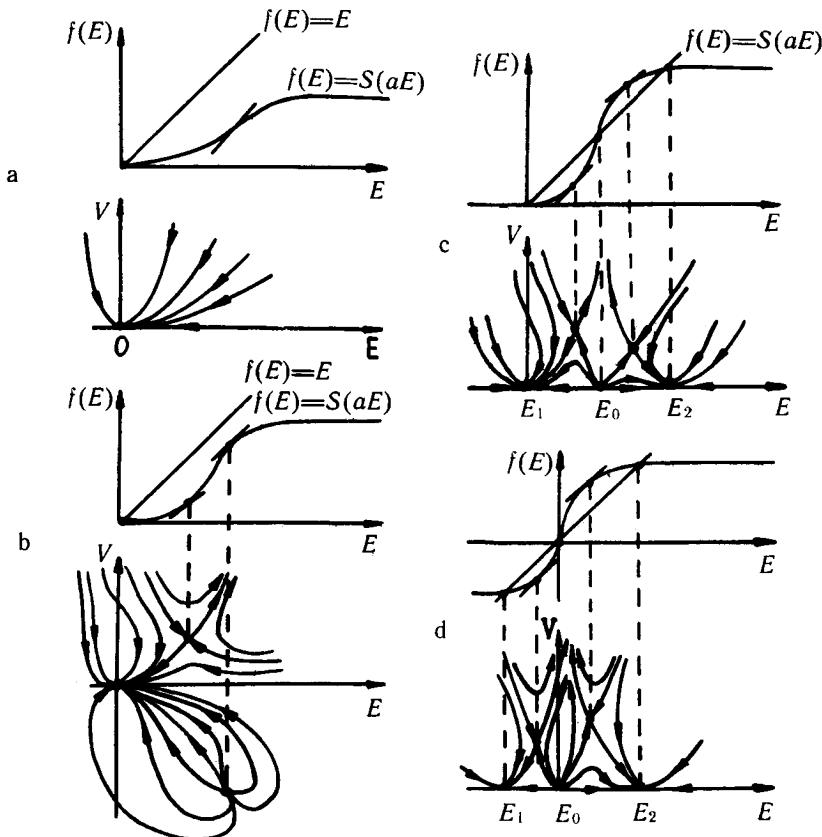


Рис. 10.2. Взаимное расположение функций и соответствующие фазовые портреты при различных значениях параметра взаимодействия а: а) случай 1; б) случай 2; в) случай 3; г) случай 4. Объяснения в тексте

**Случай 4.** В этом случае опять существуют 5 стационарных решений и фазовый портрет аналогичен случаю 3, но теперь одно устойчивое и одно седловое стационарные состояния переместились в квадрант  $E < 0, V \geq 0$  (рис. 10.2д).

Полученные результаты исследования системы применим теперь для описания возникновения критического режима в нейронной сети.

Предположим, что наша система всегда начинает эволюционировать из одних и тех же начальных условий ( $E(0), V(0)$ ). Проследим изменение финальных состояний траектории, выходящей из одной точки, при увеличении параметра связи  $a$  (рис. 3.2). При малых значениях параметра  $a$  траектория будет стремиться к состоянию фоновой активности ( $E=0, V=0$ ) (см. рис. 3.2а, б). При увеличении параметра  $a$  неустойчивый

стационар и оба седла будут сдвигаться влево и при  $a \approx a_{kp}^1$  сепаратриса седла  $(E_l, V_l)$  пройдет вблизи зафиксированной нами начальной точки  $(E_0, V_0)$ . При этой величине связи система будет длительное время эволюционировать к своему стационарному состоянию, поскольку траектория системы пройдет вблизи седла и надолго задержится в его окрестности (см. рис. 3.2б). При  $a_{kp}^1 < a < a_{kp}^2$  траектория, идущая из начальной точки, вообще не попадает в стационар  $(E=0, V=0)$  (рис. 3.2с), так как попадает в область между сепаратрисами седел и переменная будет неограниченно возрастать, а  $E$  будет принимать некоторое промежуточное значение. При  $a > a_{kp}^2$  траектория, идущая из начальной точки, будет стремиться к стационарному состоянию  $(E_2, 0)$ , соответствующему высокому уровню активности.

На рис. 3.2 приведены графики зависимости  $E(t)$  и  $V(t)$  при различных значениях параметра.

Описанная ситуация, по нашему мнению, объясняет возникновение критического режима, а диапазон изменения параметра  $a_{kp}^1 < a < a_{kp}^2$  является критическим для данного начального уровня активности сети  $(E(0), V(0))$  и фиксированных  $b$  и  $\Theta$ . Переход в критический режим связан с увеличением времени приближения к стационарному состоянию. В имитационной модели это соответствует длительному сохранению конфигурации начально активных элементов, что в свою очередь позволяет моделировать кратковременную память. Начальное «пятно» активных элементов как бы раздумывает — погаснуть или разгореться. И в зависимости от результата этих длительных «раздумий» оно либо затухает до фоновой активности, либо разгорается до высокого уровня активности. При этом средний, по множеству реализаций, уровень активности остается промежуточным, а средняя дисперсия растет. Важно подчеркнуть, что критический режим, как в случае имитационного моделирования, так и в системе (10.12), сохраняется для целого интервала изменения параметра взаимодействия  $a \in (a_{kp}^1, a_{kp}^2)$ .

Приведенные на рис. 3.2 фазовые портреты описывали случай, когда начальная активность  $E(0)$  выше фоновой ( $E(0) > 0$ ). Аналогичная картина будет наблюдаться, если взять начальную активность сети ниже фоновой ( $E(0) < 0$ ).

В заключение приведем асимптотическую формулу больших времен релаксации при приближении к критическому режиму. Предположим, что значение параметра  $a$  близко к  $a_{kp}^1$  ( $a < a_{kp}^1$ ). Тогда траектория, идущая из точки  $(E(0), V(0))$ , будет проходить вблизи сепаратрисы седла и, следовательно, время  $\tau$ , которое она будет проводить в окрестности седла, будет увеличиваться по мере приближения  $a$  к  $a_{kp}^1$ . Известно, что асимптотическая формула для  $\tau$  имеет вид:

$$\tau(\delta) \approx \frac{1}{\mu} \ln \frac{1}{\delta},$$

где  $\delta$  — расстояние от начальной точки до сепаратрисы (малый параметр);  $\mu$  — положительное собственное значение седла ( $\mu$  может обращаться в ноль при  $a=a_2$  и  $a=a_3$ ).

## ЛИТЕРАТУРА

*Аверинцев М.Б.* Об одном способе описания случайных полей с дискретным аргументом. — Пробл. передачи информ., 1970, т. 6, вып. 2, с. 100—108.

*Аверинцев М.Б.* Описание марковских случайных полей при помощи гибсовских условных вероятностей. — Теория вероятностей и ее применения, 1972, т. 17, с. 21—35.

*Анохин П.К.* Биология и нейрофизиология условного рефлекса. М., Медицина, 1968.

*Аршавский И.А.* Роль А.А.Ухтомского в создании теоретических предпосылок для исследования физиологических механизмов индивидуального и филогенетического развития (в печати), 1985.

*Батуев А.С.* Нейрофизиология коры головного мозга. Л., Изд-во ЛГУ, 1984, с. 186—192.

*Беленков Н.Ю.* Принцип целостности в деятельности мозга. М., Медицина, 1980.

*Борисюк Г.Н., Борисюк Р.М., Кириллов А.Б., Коваленко Е.И., Крюков В.И.* Об оценке синаптических весов нейронных схем при многомерном анализе импульсной активности. — В сб.: Взаимодействующие марковские процессы и их применение в биологии. Пущино, ОНТИ НЦБИ АН СССР, 1979, с. 152—177.

*Борисюк Г.Н., Коваленко Е.И., Крюков В.И.* Идентификация синаптических связей и обработка многомерных импульсных потоков методами марковских полей. — В сб.: Проблемы нейрокибернетики. Ростов-на-Дону, 1980, с. 9—10.

*Борисюк Г.Н., Борисюк Р.М., Виноградова О.С., Кириллов А.Б., Коваленко Е.И., Крюков В.И.* Метастабильное состояние в имитационной модели поля САз гиппокампа. — В сб.: Тез. докл. I Всесоюз. биофизического съезда, т. 3. Пущино, ОНТИ НЦБИ АН СССР, 1982, с. 36.

*Борисюк Р.М., Борисюк Г.Н., Кириллов А.Б., Коваленко Е.И., Крюков В.И.* О возможности статистической оценки синаптических весов нейронных схем при многомерном анализе импульсной активности. — В сб.: Тез. докл. III Всесоюз. конф. по биологической и медицинской кибернетике, т. 2. Москва—Сухуми, 1978, с. 153—155.

*Виноградов М.И.* Учение Н.Е.Введенского об основных нервных процессах. М., 1952.

*Виноградова О.С., Золотухина Л.И.* Сенсорные характеристики нейронов медиального и латерального ядер септум. — Ж. высш. нервн. деят-сти, 1972, т. 22, с. 1260.

*Виноградова О.С.* Гиппокамп и память. М., Наука, 1975, с. 332.  
*Гасанов У.Г.* Внутреннее торможение. М., Наука, 1972.

*Голиков Н.В.* Проблема местного и распространяющегося возбуждения в современной нейрофизиологии. — В сб.: Механизмы местной реакции и распространяющегося возбуждения. Л., Наука, 1970.

*Добрушин Р.Л.* Описание случайного поля при помощи условных вероятностей и условия его регулярности. — Теория вероятностей и ее применения, 1968а, т. 13, с. 201—229.

*Добрушин Р.Л.* Гиббсовские случайные поля для решетчатых систем с попарным взаимодействием. — Функц. анализ и его прил., 1968б, т. 2, вып. 4, с. 31—43.

*Добрушин Р.Л.* Задача единственности гиббсовского случайного поля и проблема фазовых переходов. — Функц. анализ и его прил., 1968в, т. 2, вып. 4, с. 44—57.

*Добрушин Р.Л.* Гиббсовские случайные поля. Общий случай. — Функц. анализ и его прил., 1969, т. 3, вып. 1, с. 27—35.

*Добрушин Р.Л.* Марковские процессы с большим числом локально взаимодействующих компонент — существование предельного процесса и его эргодичность. — Пробл. передачи информ., 1971а, т. 7, вып. 2, с. 70—87.

*Добрушин Р.Л.* Марковские процессы с большим числом локально взаимодействующих компонент — обратимый случай и некоторые обобщения. — Пробл. передачи информ., 1971б, т. 7, вып. 3, с. 57—66.

*Жадин М.Н.* Теория ритмических процессов в коре головного мозга. Препринт. Пущино, ОНТИ НЦБИ АН СССР, 1982.

*Крыжановский Г.Н.* Детерминантные структуры в патологии нервной системы. М., Медицина, 1980.

*Казанович Я.Б., Крюков В.И.* Фундаментальное тождество Вальда в непрерывном времени. — В сб.: Теория массового обслуживания. М., Изд-во МГУ, 1976, т. 2, с. 153—159.

*Кириллов А.Б., Крюков В.И.* Математическая модель кратковременной нейронной памяти. — В сб.: Тез. докл. I Всесоюз. биофизического съезда. Пущино, ОНТИ НЦБИ АН СССР, 1982, т. 3, с. 32—33.

*Коваленко Е.И., Крюков В.И., Борисюк Р.М., Борисюк Г.Н., Кириллов А.Б.* Нейронная память: имитационная модель поля САЗ гиппокампа. — В сб.: Взаимодействующие марковские процессы и их применение к математическому моделированию биологических систем. Пущино, ОНТИ НЦБИ АН СССР, 1982, с. 77—107.

*Крюков В.И.* Модели случайного блуждания для нейронов с рефрактерностью. Пущино, ОНТИ НЦБИ АН СССР, 1973а.

*Крюков В.И.* Вычисление кривой порога нейрона. — В сб.: Тез. докл. V Всесоюз. конф. по нейрокибернетике. Ростов-на-Дону, Изд-во РГУ, 1973б.

*Крюков В.И.* Модели случайного блуждания нейронов с рефрактерностью. — В сб.: Реф. докл. IV Всесоюз. совещ. по проблемам управления. М., 1974.

*Крюков В.И.* Нейронная активность как взаимодействие марковских процессов. — В сб.: Моделирование возбудимых структур. Пущино, ОНТИ НЦБИ АН СССР, 1975, с. 185—214.

*Крюков В.И.* Нейрон как система M/G/1 с переменной скоростью обслуживания. — В сб.: Теория массового обслуживания. М., Изд-во МГУ, 1976, т. 2, с. 99—104.

*Крюков В.И.* Марковские процессы взаимодействия и нейронная активность. — В сб.: Взаимодействующие марковские процессы в биологии. Пущино, ОНТИ НЦБИ АН СССР, 1977, с. 127—145.

*Крюков В.И.* «Нейролокатор» — модель внимания и оперативной памяти гиппокампальной системы. — В сб.: Тез. докл. I Всесоюз. биофизического съезда. Пущино, ОНТИ НЦБИ АН СССР, 1982, т. 3, с. 32.

*Ливанов М.Н., Крылов Ю.В., Острякова М.В., Шульгина Г.И.* О ритмических колебаниях потенциалов как одном из основных механизмов интеграции нейронов. — В сб.: Механизмы объединения нейронов в нервном центре. Л., Наука, 1974.

*Ливанов М.Н.* Нейронные механизмы памяти. — Успехи физиол. наук, 1975, т. 6, с. 66—89.

*Лямин Э.А.* Модели памяти Бриндли и Марра. — В сб.: Клеточные механизмы памяти. Пущино, ОНТИ НЦБИ АН СССР, 1973, с. 120—135.

*Лямин Э.А.* Пейсмекеры и релаксационный механизм активности в нервной системе. — В сб.: Моделирование возбудимых структур. Пущино, ОНТИ НЦБИ АН СССР, 1975, с. 3—27.

*Маунткасл В.* Организующий принцип мозга — элементарный модуль и распределенная система. — В кн.: Дж.Эделмен, В.Маунткасл «Разумный мозг». М., Мир, 1981.

*Меркулов В.Л.* Принцип доминанты и представления А.А.Ухтомского о хронотопе (временно-пространственном комплексе). — Успехи соврем. биол., т. 47, вып. 2, 1959, с. 204—219.

*Меркулов В.Л.* Алексей Алексеевич Ухтомский. Очерки жизни и научной деятельности (1875—1942). М.—Л., Изд-во АН СССР, 1960.

*Меркулов В.Л.* Принцип общего пути и принцип доминанты А.А.Ухтомского. — В сб.: Механизмы нервной деятельности. Л., Изд-во ЛГУ, 1977, с. 77—86.

*Миухина Р.С.* Электроэнцефалографические исследования условно-рефлекторных реакций. Л., Изд-во ЛГУ, 1964.

*Миухина Р.С.* Переход местного возбуждения в распространяющееся как механизм замыкания временной связи. — В сб.: Механизмы местной реакции и распространяющегося возбуждения. Л., Наука, 1970, с. 114—124.

*Миухина Р.С.* О механизме замыкания связи в свете теории Введенского-Ухтомского. — В сб.: Механизмы нервной деятельности. Л., Изд-во ЛГУ, 1977, с. 112—126.

*Русинов В.С.* Доминанта. М., Медицина, 1969, с. 59.

*Русинов В.С.* Проблема стационарного возбуждения и следовые процессы в коре большого мозга. — В сб.: К 100-летию со дня рождения А.А.Ухтомского. Л., Изд-во ЛГУ, 1975, с. 79—92.

*Соколов Е.Н.* Нервная модель стимула и бриентировочный рефлекс. — Вопр. психологии, 1960, № 4, с. 128.

*Ставская О.Н., Пятецкий—Шапиро И.И.* Об однородных сетях из спонтанноактивных элементов. — Пробл. кибернетики, 1968, вып. 20, с. 91—106.

*Ставская О.Н.* Гиббсовские инвариантные меры для цепей Маркова на конечных решетках с локальным взаимодействием. — Мат. сб., 1973, т. 92, с. 402—419.

*Стенли Г.* Фазовые переходы и критические явления. М., Мир, 1973.

*Уфлянд Ю.М.* Пути развития учения А.А.Ухтомского о доминанте. — Успехи физиол. наук, 1975, т. 6, № 3, с. 23—35.

*Ухтомский А.А.* Заметка на полях календаря-справочника. — ААН, ф 749 оп 3 № 33/1—8, 1938.

*Ухтомский А.А.* Собр. соч., т. IV. Л., Изд-во ЛГУ, 1954, с. 43.

*Ухтомский А.А.* Собр. соч., т. I. Л., Изд-во ЛГУ, 1959.

*Ухтомский А.А.* Доминанта. М., Наука, 1966, с. 125.

*Ухтомский А.А.* Избр. тр. Л., Наука, 1978а, с. 168.

*Ухтомский А.А.* (Совместно с Гуляевым П.И.). Параметр физиологической лабильности и нелинейная теория колебаний. Избр. тр. Л., Наука, 1978б, с. 194.

*Феллер В.* Одномерные диффузионные процессы. — Математика, 1958, т. 2, вып. 2, с. 119—146.

*Фролов А.А., Шульгина Г.И.* Моделирование условий проведения возбуждения при различных режимах работы нервной сети. — В сб.: Функциональное значение электрических процессов головного мозга. М., Наука, 1977, с. 190—198.

*Цирельсон Б.С.* О формальном описании квантовых систем близких к системам стохастических автоматов. — В сб.: Взаимодействующие марковские процессы и их применение в биологии. Пущино, ОНТИ НЦБИ АН ССР, 1979, с. 100—138.

*Четаев А.Н.* Долгоживущие состояния конечных систем. — Пробл. передачи информ., 1979, вып. 1, с. 71—98.

*Шмитт Ф.* Введение к кн. Дж.Эдельмена и В.Маунткасла «Разумный мозг». М., Мир, 1981.

*Эдельмен Дж.* Селекция групп и фазовая повторная сигнализация; теория высших функций головного мозга. — В кн.: Дж.Эдельмен, В.Маунткасл «Разумный мозг». М., Мир, 1981.

*Экклс Дж.* Физиология синапсов. М., Мир, 1966.

*Agarwal G.S., Shenoy S.K.* Observability of hysteresis in first order equilibrium and nonequilibrium phase transitions. — Phys. Rev., 1981, v. A23, no. 5, p. 2719-2723.

*Beekman R.* Asymptotic Distributions for the Ornstein-Uhlenbeck process. — J. Appl. Prob., 1975, v. 12, p. 107-114.

*Beijeren H., Sylvester G.S.* Phase transitions for continuousspin Ising ferromagnets. — J. Funct. Anal., 1978, v. 28, p. 145-167.

*Bell C.C., Kawasaki T.* Relations among climbing fiber responses of nearby Purkinje cells. — J. Neurophysiol., 1972, v. 35, p. 155-169.

*Besag J.* Spatial interaction and the statistical analysis of lattice systems. — J. Roy. Statist. Soc. Ser. B, 1974, p. 192-236.

*Bindman L., Lippold O.* The neurophysiology of the cerebral cortex. Edvard Arnold, 1981.

*Capocaccia D., Cassandro M., Olivieri E.* A study of metastability of the Ising model. — Comm. Math. Phys., 1974, v. 39, p. 185-205.

*Colin F., Manil J., Desclin J.C.* The olivocerebellar system. I. Delayed and slow inhibitory effects: an overlooked salient feature of cerebellar climbing fibers. — Brain Res., 1980, v. 187, p. 3-27.

*Ebner T.J., Bloedel J.R.* Temporal patterning in simple spike discharge of Purkinje cells and its relationship to climbing fiber activity. — J. Neurophysiol., 1981a, v. 45, No. 5, p. 933-947.

*Ebner T.J., Bloedel J.R.* Correlation between activity of Purkinje cells and its modification by natural peripheral stimuli. — J. Neurophysiol., 1981b, v. 45, No. 5, p. 948-961.

*Ebner T.J., Bloedel J.R.* Role of climbing fiber afferent input in determining responsiveness of Purkinje cells to mossy fiber input. — J. Neurophysiol., 1981c, v. 45, No. 5, p. 962-971.

*Eccles J.C.* An instruction-selection theory of learning in the cerebellar cortex. — Brain Res., 1977, v. 127, p. 327-352.

*Freeman W.J.* Dynamics of image formation by nerve cell assemblies. — In: Synergetics of the brain. (E.Basar, H.Flohr, H.Haken, A.J.Mandell, eds.). N.Y., Springer, 1983, p. 102-121.

*Groves P.M., Thompson R.F.* Habituation: a dual-process theory. Psychol. Rev., 1970, v. 77, p. 419-450.

*Haken H.* Synopsis and introduction. — In: Synergetics of the brain. (E.Basar, H.Flohr, H.Haken, A.J.Mandell, eds.). N.Y., Springer, 1983, p. 3-25.

- Kovalenko E.I., Borisyuk G.N., Borisyuk R.M., Kirillov A.B., Kryukov V.I.* Short-term memory as a metastable state. II. Simulation model. — In: *Cybernetics and system research*, v. 2. (R.Trappi, ed.). Elsevier, Science Publishers B.V. (North Holland), 1984, p. 266-270.
- Kryukov V.I.* Wald's identity and random walk models for neuron firing. Preprint Pushchino, 1974.
- Kryukov V.I.* Wald's identity and random walk models for neuron firing. — *Adv. Appl. Prob.*, 1976, v. 8, p. 257-277.
- Kryukov V.I.* Markov interaction processes and neuronal activity. — *Adv. Appl. Prob.*, 1978a, v. 10, p. 297.
- Kryukov V.I.* Markov interaction processes and neuronal activity. — *Lect. Notes in Math.*, 1978b, v. 653, p. 122-139.
- Kryukov V.I.* Short-term memory as a metastable state. I. Master equation approach. — In: *Cybernetics and system research*, v. 2. (R.Trappi, ed.). Elsevier, 1984, p. 261-265.
- Lisberger S.G.* Role of the cerebellum during motor learning in the vestibuloocular reflex (VOR). — *Trends in Neurosci.*, 1982, v. 5, No. 12, p. 437-441.
- Little W.A.* The existence of persistent states in the brain. — *Math. Biosci.*, 1974, v. 19, p. 101-120.
- Marr D.* A theory of cerebellar cortex. — *J. Physiol.*, 1969, v. 202, p. 437-470.
- Marr D.* Simple memory: a theory for archicortex. — *Phil. Trans. Roy. Soc.*, 1971, v. B262, p. 23-81.
- McLennan H., Miller J.J.* The hippocampal control of neuronal discharges in septum of the rat. — *J. Physiol. (Lond.)*, 1974, v. 237, p. 607.
- McLennan H., Miller J.J.* Frequency-related inhibitory mechanisms controlling rhythmical activity in the septal area. — *J. Physiol. (Lond.)*, 1976, v. 254, p. 827-841.
- Miles F.A., Braitman D.J., Dow P.M.* Long-term adaptive changes in primate vestibuloocular reflex. IV Electrophysiological observations in flocculus of monkeys. — *J. Neurophysiol.*, 1980, v. 43, p. 1477-1493.
- Perkel D.H., Gerstein G.L., Moor G.P.* Neuronal spike trains and stochastic point processes. I. The single spike train. II. Simultaneous spike trains. — *Biophys. J.*, 1967, v. 7, p. 391-440.
- Petsche H., Pockberger H., Rappelsberger P.* On the search for the sources of the electroencephalogram. — *Neurosci.*, 1984, v. 11, No. 1, p. 1-27.
- Preston C.J.* Generalized Gibbs states and Markov random fields. — *Adv. Appl. Prob.*, 1973, v. 5, p. 242-261.
- Rager G.* Vertical and tangential organization in the neocortex. — In: *Synergetics of the brain*. (E.Basar, H.Flohr, H.Haken, A.J.Mandell, eds.). N.Y., Springer, 1983.
- Roney K.J., Shaw G.L.* Analytic study of assemblies of neurons in memory storage. — *Math. Biosci.*, 1980, v. 51, No. 1-2, p. 25-41.
- Van Rotterdam A., Lopes da Silva F.H., van der Ende J., Viergever M.A., Hermans A.J.* A model of the spatial-temporal characteristics of the alfa-rhythm. — *Bull. Math. Biol.*, 1982, v. 44, p. 283-305.
- Schmitt F.O., Dev P., Smith B.H.* Electronic processing of information by brain cells. — *Science*, 1976, v. 193, p. 114-120.
- Schulman L.S.* Magnetization probabilities and metastability in the Ising model. — *J. Phys. A: Math. Gen.*, 1980, v. 13, p. 237-250.
- Shiromani S., Fishbein W.* Neocortical activation and adaptive be-

havior: holynergic influences. — *Behav. Brain Sci.*, 1981, v. 4, No. 3, p. 459-514.

Stoll E. et al. Monte Carlo investigation of dynamic critical phenomena in two-dimensional kinetic Ising model. — *Phys. Rev.*, 1973, v. B8, No. 7, p. 3266-3289.

Stuart C.M., Takahashi Y., Umezawa H. On the stability and non-local properties of memory. — *J. Theor. Biol.*, 1978, v. 71, p. 605-618.

Stuart C.M., Takahashi Y., Umezawa H. Mixed-system brain dynamics: neural memory as a macroscopic ordered state. — *Found. Phys.*, 1979, v. 9, p. 301-327.

Tömböl T., Petsche H. The histological organization of the pacemaker for the hippocampal theta rhythm in the rabbit. — *Brain Res.*, 1969, v. 12, p. 414.

Vandervolf C.H., Robinson T.E. Reticulo-cortical activity and behavior: a critique of the arousal theory and a new synthesis. — *Behav. Brain Sci.*, 1981, v. 4, No. 3, p. 459-514.

Venzi G. Statistical fluctuation of activity in localized population. — *J. Theoret. Biol.*, 1976, v. 63, p. 259-274.

Wilson H.R. Mathematical models of neural tissue. — In: Cooperative phenomena. (H.Haken, ed.). North-Holland, 1974, p. 247-261.

Crick F. et al. — *Nature*, 1983, No. 304, p. 111-114.

## Содержание

<b>ПРЕДИСЛОВИЕ . . . . .</b>	<b>3</b>
<b>1. ВВЕДЕНИЕ. КРАТКИЙ ОБЗОР БОЛЕЕ РАННИХ МОДЕЛЕЙ</b>	<b>5</b>
<b>2. БАЗОВАЯ МОДЕЛЬ И ДАЛЕКИЙ ПОРЯДОК В НЕЙ . . . . .</b>	<b>11</b>
2.1. Базовая нейронная модель . . . . .	12
2.2. Достаточные условия фазового перехода . . . . .	13
2.3. Задача о первом достижении . . . . .	15
2.4. Существование метастабильных состояний . . . . .	17
<b>3. ДИФФУЗИОННАЯ АППРОКСИМАЦИЯ . . . . .</b>	<b>19</b>
3.1. Введение . . . . .	19
3.2. Вывод динамической системы . . . . .	20
3.3. Обоснование сигмоидной нелинейности (задача Э.А.Ляминя) . . . . .	22
3.4. Динамические эффекты вблизи точки неустойчивого равновесия . . . . .	23
3.5. Оценки времени жизни метастабильных и неустойчивых состояний . . . . .	25
<b>4. ИМИТАЦИОННАЯ МОДЕЛЬ ОСНОВНОГО СУБМОДУЛЯ . . . . .</b>	<b>28</b>
4.1. Машинный вариант базовой модели . . . . .	28
4.2. Динамика уровня активности: выход на стационарный уровень . . . . .	29
4.3. Эффект сохранения пятна . . . . .	31
4.4. Линия критического режима . . . . .	33
4.5. Периодическая модуляция порога . . . . .	34
4.6. Сеть с синаптической модификацией . . . . .	34
4.7. Сеть с возвратным торможением . . . . .	36
<b>5. ПРИМЕНЕНИЕ К ГИППОКАМПУ: МОДЕЛЬ ПРИВЫКАНИЯ . . . . .</b>	<b>39</b>
<b>6. ПРИМЕНЕНИЕ К СЕПТУМУ: МОДЕЛЬ ГЛОБАЛЬНОГО ПЕЙСМЕКЕРА . . . . .</b>	<b>44</b>
<b>7. ПРИМЕНЕНИЕ К МОЗЖЁЧКУ: МОДЕЛЬ ВЗАЙМОДЕЙСТВИЯ ЛИАННЫХ И МШИСТЫХ ВОЛОКОН . . . . .</b>	<b>49</b>
<b>8. ПРИМЕНЕНИЕ К НОВОЙ КОРЕ: МОДЕЛЬ СТАЦИОНАРНОГО И РАСПРОСТРАНЯЮЩЕГОСЯ ВОЗБУЖДЕНИЯ . . . . .</b>	<b>59</b>
8.1. «Тихая революция» и связанные с ней проблемы . . . . .	59
8.2. Принцип локального стационарного возбуждения и торможения Н.Е.Введенского . . . . .	61
8.3. Существующие модели и представления . . . . .	64
8.4. Модель «сохранения пятна» и распространения возбуждения . . . . .	67
8.5. Интерпретация свойств доминантного очага . . . . .	75
<b>9. ЗАКЛЮЧЕНИЕ. ФАЗОВЫЕ ПЕРЕХОДЫ И МЕТАСТАБИЛЬНОСТЬ — ОСНОВА ФУНКЦИОНИРОВАНИЯ МОЗГА . . . . .</b>	<b>81</b>
<b>10. ПРИЛОЖЕНИЕ . . . . .</b>	<b>88</b>
10.0. Общая нейронная модель и ее уравнение Колмогорова . . . . .	88
10.1. Основные определения и общие сведения о гиббсовских случайных полях. Применимость к нейрофизиологии . . . . .	92
10.2. Обратимые и псевдообратимые марковские процессы . . . . .	97
10.3. Определения и общие сведения о модели Изинга с непрерывным пространством состояний . . . . .	98
10.4. Качественное исследование уравнений динамики среднего уровня и дисперсии активности нейронной сети . . . . .	101
<b>ЛИТЕРАТУРА . . . . .</b>	<b>106</b>

## **THE METASTABLE AND UNSTABLE STATES IN THE BRAIN**

### **Summary**

The theoretical and experimental evidence for the existence in the brain of metastable states with lifetimes of about one second and higher is presented and the basic principle of brain operation is discussed. The theory of phase transitions and metastability similar to that of ferromagnets is developed and applied to interpret very important and puzzling data concerning.

- long lasting tonic reactions and habituation in the hippocampus;
- bursting theta-rhythmic spike activity in the septum;
- the hysteresis and slow-decay correlations of the spike activity in the cerebellum;
- paradoxical increase in amplitude and duration of the late components of evoked potentials in the neokortex.

A tentative solution to the various neurophysiological problems such as synchronization, reverberation and system behaviour inhibition is presented and implication for the theory of memory, learning, attention and oscillatory movement organization is discussed. The existence of the phase control information processing system in the CNS is predicted. This system in particular must be astatic of at least the second order functioning in the alpha- and theta-rhythm frequency range and containing the septohippocampal system as a global synchronizer and phase comparator.

## Content

1. INTRODUCTION AND SHORT REVIEW OF PREVIOUS MODELS . . . . .	3
2. BASIC MODEL AND THE LONG-RANGE ORDER THEREIN . . . . .	11
2.1. Basic neuronal model . . . . .	12
2.2. Sufficient condition for the phase transitions . . . . .	13
2.3. The first passage probability problem . . . . .	15
2.4. The existence of metastable states . . . . .	17
3. THE DIFFUSION APPROXIMATION . . . . .	19
3.1. Introductory remarks . . . . .	19
3.2. Derivation of the dynamic system . . . . .	20
3.3. The proof of the sigmoid nonlinearity . . . . .	22
3.4. The dynamic effects near the point of instability . . . . .	23
3.5. Estimates for the lifetimes of metastable and unstable states . . . . .	25
4. THE SIMULATION MODEL OF THE BASIC SUBMODULE . . . . .	28
5. THE APPLICATION OF THE BASIC MODEL TO THE HIPPO-CAMPUS: THE MODEL OF HABITUATION . . . . .	39
6. THE APPLICATION TO THE SEPTUM: THE MODEL OF THE GLOBAL PACEMAKER . . . . .	44
7. THE APPLICATION TO THE CEREBELLUM: THE MODEL OF MOSSY FIBER AND CLIMBING FIBER INTERACTION . . . . .	49
8. THE APPLICATION TO THE NEOCORTEX: THE MODEL OF STATIONARY AND PROPAGATING EXCITATION . . . . .	59
9. CONCLUSION: PHASE TRANSITIONS AND METASTABILITY AS A BASIS OF BRAIN FUNCTIONING . . . . .	81
10. MATHEMATICAL APPENDICES . . . . .	88
10.0. General neuronal model and its Kolmogorov equation . . . . .	88
10.1. General properties of Gibbs random fields . . . . .	92
10.2. The reversible and pseudoreversible Markov processes . . . . .	97
10.3. General properties of continuous state Ising model . . . . .	98
10.4. Qualitative study of mean and variance of the basic neuronal model activity . . . . .	101
BIBLIOGRAPHY . . . . .	106

**Виталий Иванович Крюков, Галина Николаевна Борисюк,  
Роман Матвеевич Борисюк, Александр Борисович Кириллов,  
Евгения Ивановна Коваленко**

**МЕТАСТАБИЛЬНЫЕ И НЕУСТОЙЧИВЫЕ СОСТОЯНИЯ В МОЗГЕ**

Отредактировано и подготовлено к печати в ОНТИ НЦБИ АН СССР  
Редактор *Л.В.Сотникова*. Технический редактор *Т.М.Печенкина*. Художествен-  
ный редактор *В.М.Рудакова*. Корректоры *Л.М.Орлова, Т.К.Крейцер*. Подписано  
в печать 22.12.86 г. Т-25007. Уч.-изд. л. 6,7. Тираж 635 экз. Заказ 70Р.

Формат 60×90/16. Цена 80 к. Бумага офсетная. Изд. № 130.

Набрано на фотонаборном автомате ФА-1000.  
Отпечатано на ротапринте в Отделе научно-технической информации Научного  
центра биологических исследований АН СССР в Пущине

## CONTENT

1. INTRODUCTION AND SHORT REVIEW OF PREVIOUS MODELS
2. BASIC MODEL AND THE LONG-RANGE ORDER THEREIN
  - 2.1. Basic neuronal model
  - 2.2. Sufficient condition for the phase transitions
  - 2.3. The first passage probability problem
  - 2.4. The existence of metastable states
3. THE DIFFUSION APPROXIMATION
  - 3.1. Introductory remarks
  - 3.2. Derivation of the dynamic system
  - 3.3. The proof of the sigmoid nonlinearity
  - 3.4. The dynamic effects near the point of instability
  - 3.5. Estimates for the lifetimes of metastable and unstable states
4. THE SIMULATION MODEL OF THE BASIC SUB-MODULE
5. THE APPLICATION OF THE BASIC MODEL TO THE HIPPOCAMPUS: THE MODEL OF HABITUATION
6. THE APPLICATION TO THE SEPTUM: THE MODEL OF THE GLOBAL PACEMAKER
7. THE APPLICATION TO THE CEREBELLUM: THE MODEL OF MOSSY FIBER AND CLIMBING FIBER INTERACTION
8. THE APPLICATION TO THE NEOCORTEX: THE MODEL OF STATIONARY AND PROPAGATING EXCITATION
9. CONCLUSION: PHASE TRANSITIONS AND METASTABILITY AS A BASIS OF BRAIN FUNCTIONING
10. MATHEMATICAL APPENDICES
  - 10.0. General neuronal model and its Kolmogorov equation
  - 10.1. General properties of Gibbs random fields
  - 10.2. The reversible and pseudoreversible Markov processes
  - 10.3. General properties of continuous state Ising model
  - 10.4. Qualitative study of mean and variance of the basic neuronal model activity