ВЫСШЕЕ ОБРАЗОВАНИЕ

В.Г. Лещенко, Г.К. Ильич

МЕДИЦИНСКАЯ И БИОЛОГИЧЕСКАЯ ФИЗИКА

УЧЕБНОЕ ПОСОБИЕ



ВЫСШЕЕ ОБРАЗОВАНИЕ
серия основана в 1996 г.

В.Г. ЛЕЩЕНКО
Г.К. ИЛЬИЧ

МЕДИЦИНСКАЯ И БИОЛОГИЧЕСКАЯ ФИЗИКА

Допущено
Министерством образования Республики Беларусь
в качестве учебного пособия для студентов
учреждений высшего образования
по медицинским специальностям

Минск
«Новое знание»

Москва
«ИНФРА-М»

2012

УДК [61:53+577.3](075.8)
ББК 53.6я73
Л54

Рецензенты:
кафедра медицинской и биологической физики Гродненского государственного медицинского университета (зав. кафедрой — кандидат физико-математических наук, доцент *И.М. Бертель*);
зав. кафедрой «Конструирование и производство приборов» Белорусского национального технического университета, доктор технических наук, профессор *М.Г. Киселев*

Лещенко, В.Г.

Л54 Медицинская и биологическая физика : учеб. пособие / В.Г. Лещенко, Г.К. Ильич. — Минск : Новое знание ; М. : ИНФРА-М, 2012. — 552 с. : ил. — (Высшее образование).

ISBN 978-985-475-456-7 (Новое знание)
ISBN 978-5-16-005338-7 (ИНФРА-М)

Рассмотрены физические процессы и явления, сопровождающие и обеспечивающие жизнедеятельность организма, а также физические основы как традиционных, так и современных методов диагностики и лечения. Использован удобный для усвоения порядок изложения материала: сначала физическая сторона изучаемого явления, а затем его медицинские и биологические приложения. В пособие включены также элементы высшей математики, теории вероятностей, основы статистики и корреляционного анализа в соответствии с типовой программой дисциплины.

Для студентов медицинских вузов. Может быть полезно аспирантам и студентам технических вузов, изучающим медицинскую и биологическую физику.

УДК [61:53+577.3](075.8)
ББК 53.6я73

ISBN 978-985-475-456-7 (Новое знание)
ISBN 978-5-16-005338-7 (ИНФРА-М)

© Лещенко В.Г., Ильич Г.К., 2012
© ООО «Новое знание», 2012

ОГЛАВЛЕНИЕ

Предисловие 10

Глава 1. Элементы высшей математики

1.1. Производная функции 12
1.2. Исследование функции с помощью производной. Нахождение экстремумов функции 15
1.3. Дифференциал функции 17
1.4. Частные производные 18
1.5. Частные дифференциалы. Полный дифференциал функции 19
1.6. Первообразная функция и неопределенный интеграл 21
1.7. Определенный интеграл 21
1.8. Дифференциальные уравнения 24
1.9. Этапы решения задач при использовании дифференциальных уравнений 26
1.10. Примеры использования дифференциальных уравнений 27
Контрольные вопросы и задания 30

Глава 2. Основы теории вероятностей

2.1. Классическое и статистическое определение вероятности 31
2.2. Виды случайных событий 33
2.3. Теоремы сложения и умножения вероятностей 34
2.4. Принципы вероятностных подходов к задачам диагностики и прогнозирования заболеваний 37
Контрольные вопросы и задания 40

Глава 3. Случайные величины и их распределение

3.1. Закон распределения дискретной случайной величины 42
3.2. Распределение непрерывной случайной величины. Плотность распределения вероятности 44
3.3. Числовые характеристики распределения случайных величин 46
3.4. Нормальный закон распределения случайной величины 50
Контрольные вопросы и задания 53

Глава 4. Основы математической статистики
4.1. Предмет и задачи математической статистики. Генеральная и выборочная совокупность 54
4.2. Статистическое распределение выборки 56
4.3. Методы описательной статистики 64
4.4. Оценка параметров генеральной совокупности по параметрам выборки 66
4.5. Определение объема выборки при заданной точности интервальной оценки 69
4.6. Понятие нормы для медицинских показателей. 70
4.7. Элементы корреляционного анализа 70
Контрольные вопросы и задания 73

Глава 5. Элементы биомеханики
5.1. Кинематика и динамика движений человека. 74
5.2. Виды деформаций и механические характеристики упругих тел 78
5.3. Механические свойства биотканей 84
Контрольные вопросы и задания 93

Глава 6. Механические колебания и волны
6.1. Гармонические колебания 95
6.2. Затухающие колебания 98
6.3. Вынужденные колебания 100
6.4. Сложение гармонических колебаний 102
6.5. Разложение сложных колебаний в ряд Фурье. 104
6.6. Принципы применения Фурье-анализа для обработки диагностических данных 105
6.7. Механические волны 107
6.8. Эффект Доплера 110
Контрольные вопросы и задания 112

Глава 7. Акустика
7.1. Физические и физиологические характеристики звуковых волн 113
7.2. Область слышимости, порог слышимости и порог болевого ощущения 116
7.3. Закон Вебера–Фехнера. Уровни интенсивности и уровни громкости звука 118
7.4. Отражение и поглощение акустических волн 120
7.5. Ультразвук и его медицинское применение. Инфразвук 122
Контрольные вопросы и задания 130

Глава 8. Свойства жидкостей. Поверхностные явления

8.1. Поверхностная энергия 132
8.2. Силы поверхностного натяжения 134
8.3. Добавочное давление под изогнутой поверхностью жидкости. Формула Лапласа 136
8.4. Явление смачивания 137
8.5. Капиллярные явления 139
8.6. Газовая эмболия 141
8.7. Роль альвеолярного сурфактанта в процессе дыхания 143
8.8. Методы определения коэффициента поверхностного натяжения жидкостей 144
Контрольные вопросы и задания 147

Глава 9. Физические основы гемодинамики

9.1. Основные гидродинамические понятия и законы 149
9.2. Вязкость жидкости. Ньютоновские и неньютоновские жидкости 156
9.3. Течение вязкой жидкости. Формула Пуазейля. 158
9.4. Методы определения вязкости жидкости 161
9.5. Условия перехода ламинарного течения жидкости в турбулентное 165
9.6. Некоторые особенности движения крови по сосудам 166
9.7. Роль эластичности сосудов в системе кровообращения.Пульсовые волны. 170
9.8. Распределение давления и скорости кровотока в сосудистой системе. 174
9.9. Некоторые методы определения давления и скорости крови. 177
9.10. Работа и мощность сердца. 180
Контрольные вопросы и задания 182

Глава 10. Физические свойства и функции биологических мембран

10.1. Строение и физические свойства биологических мембран 184
10.2. Транспорт молекул и ионов через мембраны. 190
10.3. Математическое описание пассивного транспорта. .. 195
10.4. Активный транспорт ионов. 198
Контрольные вопросы и задания 201

Глава 11. Мембранные потенциалы клетки

11.1. Равновесные потенциалы Нернста. 202
11.2. Потенциал покоя клетки. 203
11.3. Возбуждение клетки. Генерация потенциала действия 207
11.4. Подпороговые раздражители и местные потенциалы 216
11.5. Распространение потенциала действия по аксонам. . . 217
Контрольные вопросы и задания. 224

Глава 12. Электрические поля органов и тканей. Методы их регистрации

12.1. Электрическое поле и его характеристики. 226
12.2. Электрический диполь и его поле. 230
12.3. Физические основы электрографии органов и тканей 236
Контрольные вопросы и задания 243

Глава 13. Электропроводимость биологических тканей и жидкостей для постоянного тока

13.1. Постоянный ток в электролитах. 245
13.2. Особенности электропроводимости биологических тканей. 248
13.3. Некоторые лечебные методы, основанные на использовании постоянного тока. 251
Контрольные вопросы и задания. 253

Глава 14. Переменный ток. Электрический импеданс живой ткани. Основы реографии

14.1. Основные характеристики переменного тока. 254
14.2. Различные нагрузки в цепи переменного тока. 255
14.3. Сопротивление живой ткани переменному току. 261
14.4. Физические основы реографии. 264
Контрольные вопросы и задания. 268

Глава 15. Магнитное поле

15.1. Общие свойства и характеристики магнитного поля. . . 269
15.2. Силовое действие магнитного поля. 273
15.3. Магнитный поток, закон электромагнитной индукции. . . 277
15.4. Самоиндукция. Индуктивность контура. Энергия магнитного поля. 279
15.5. Магнитные свойства веществ. 281

15.6. Магнитные свойства биотканей. 283
15.7. Биологическое действие магнитного поля. 285
Контрольные вопросы и задания. 287

Глава 16. Основы электростимуляции тканей и органов

16.1. Параметры импульсных токов. 289
16.2. Физиологическое обоснование параметров импульсных токов при электростимуляции. 292
16.3. Электростимуляция сердечной мышцы. 298
Контрольные вопросы и задания. 299

Глава 17. Использование электромагнитных полей высокой частоты в медицине

17.1. Тепловое действие ЭМП высокой частоты. 301
17.2. Нетепловое действие ЭМП высоких частот. 308
Контрольные вопросы и задания. 311

Глава 18. Получение и регистрация биофизических данных

18.1. Регистрация биопотенциалов. 314
18.2. Регистрация неэлектрических параметров. Датчики. . . . 317
18.3. Усиление электрических сигналов. 324
18.4. Регистрирующие устройства. 332
Контрольные вопросы и задания. 334

Глава 19. Общие свойства и поляризация электромагнитных волн

19.1. Общие свойства и шкала электромагнитных волн. . . 336
19.2. Виды поляризации электромагнитных волн. 339
19.3. Методы получения поляризованного света. 343
19.4. Прохождение света через поляризатор. Закон Малюса. 349
19.5. Оптическая активность. 350
19.6. Методы исследования в поляризованном свете. 351
Контрольные вопросы и задания. 356

Глава 20. Тепловое излучение тел

20.1. Основные характеристики теплового излучения. . . . 358
20.2. Законы теплового излучения 360
20.3. Излучение Солнца. 364

20.4. Инфракрасное излучение и его применение в медицине. 365
20.5. Тепловое излучение тела человека. 366
20.6. Понятие о термографии. 368
Контрольные вопросы и задания. 369

Глава 21. Оптические спектры атомов и молекул

21.1. Поглощение света. 370
21.2. Рассеяние света 374
21.3. Энергетические состояния атомов 377
21.4. Спектр атома водорода. 381
21.5. Спектры сложных атомов. 384
21.6. Спектры молекул. 387
21.7. Люминесценция. 392
Контрольные вопросы и задания. 401

Глава 22. Вынужденное излучение. Лазеры

22.1. Вынужденное излучение и его свойства. 403
22.2. Условия усиления света. Лазеры. 405
22.3. Применение лазерного излучения в медицине. 410
Контрольные вопросы и задания. 411

Глава 23. Элементы фотобиологии

23.1. Фотобиологические процессы. 413
23.2. Типы фотохимических реакций. 416
23.3. Фотосенсибилизированные фотобиологические процессы. 417
23.4. Основы фотодинамической терапии. 419
Контрольные вопросы и задания. 421

Глава 24. Биофизические основы зрения

24.1. Строение глаза. 422
24.2. Оптическая система глаза. 424
24.3. Биофизические основы фоторецепции. 431
Контрольные вопросы и задания. 441

Глава 25. Приборы для исследования микрообъектов

25.1. Линза. 443
25.2. Лупа. 449
25.3. Микроскоп. 450
25.4. Электронный микроскоп. 453
25.5. Сканирующая зондовая микроскопия. 458
Контрольные вопросы и задания. 460

Глава 26. Электронный парамагнитный и ядерный магнитный резонанс

26.1. Свободные радикалы в биологических системах.... 462
26.2. Электронный парамагнитный резонанс 464
26.3. Ядерный магнитный резонанс 470
26.4. Принципы магнитно-резонансной томографии 475
Контрольные вопросы и задания..... 480

Глава 27. Рентгеновское излучение

27.1. Тормозное рентгеновское излучение..... 482
27.2. Характеристическое рентгеновское излучение..... 486
27.3. Первичные эффекты взаимодействия рентгеновского излучения с веществом. 488
27.4. Ослабление рентгеновского излучения в веществе .. 490
27.5. Физические принципы рентгенодиагностики..... 492
Контрольные вопросы и задания..... 495

Глава 28. Радиоактивность

28.1. Основные характеристики ядер..... 497
28.2. Виды радиоактивного распада..... 500
28.3. Простейшие ядерные реакции. Получение радионуклидов..... 503
28.4. Основной закон радиоактивного распада. Активность радионуклидов. 507
28.5. Взаимодействие ионизирующего излучения с веществом. 512
28.6. Принципы радионуклидных методов диагностики... 517
28.7. Физические основы лучевой терапии. 521
Контрольные вопросы и задания..... 524

Глава 29. Дозиметрия ионизирующих излучений

29.1. Дозы излучения..... 525
29.2. Принципы расчета доз внутреннего облучения. 534
29.3. Детекторы ионизирующих излучений..... 538
29.4. Дозиметрические приборы. 542
29.5. Естественный радиационный фон, фоновое облучение человека. 545
Контрольные вопросы и задания..... 548
Литература..... 549

Предисловие

При подготовке медицинских специалистов современного уровня постоянно увеличивается роль базовых естественных наук, что обусловлено внедрением новых технических решений в современную медицину. Это относится, в частности, к физике и связанным с ней смежным техническим дисциплинам (медицинская электроника, материаловедение и др.).

Практически в каждой стратегической и утилитарной задаче современного медицинского образования можно отметить необходимость знания и понимания физических законов и явлений. Например, значительный прогресс в области морфологических наук и в связанных с ними методиках диагностики заболеваний стал возможен и благодаря использованию таких чисто физических методов, как оптическая и электронная микроскопия, атомно-силовая микроскопия, рентгеновская и магнитно-резонансная томография, позитронно-эмиссионная томография и др. Что касается физиологических процессов, протекающих в организме, то еще великий физиолог И.П. Павлов подчеркивал, что понять механизм функционирования живого организма невозможно без знаний законов физики.

Значительная часть современных методов диагностики заболеваний основана на физических методах, а их приборная реализация в виде физических устройств осуществляется средствами медицинской электроники.

В лечебном процессе физические методы и аппаратура также нашли широкое применение. Вся физиотерапия основана на использовании дозированного воздействия на организм токов, полей и излучений. Возможности хирургического лечения заболеваний существенно расширяются благодаря использованию ультразвуковой, лазерной и высокочастотной техники. Палаты реанимации и интенсивной терапии оснащены специализированной физической аппаратурой, которую необходимо грамотно и безопасно эксплуатировать.

Оценка состояния окружающей человека среды и продуктов питания, определение их санитарно-гигиенических характеристик тоже осуществляется физическими методами и приборами наряду с биохимическими и микробиологическими технологиями.

Необходимость грамотного количественного описания медико-биологических процессов и корректной обработки медицинских данных требует освоения будущими врачами определенных элементов математического аппарата, без которых невозможно овладеть современными медицинскими информационными технологиями. Поэтому наряду с физическими разделами в учебник включены элементы высшей математики, теории вероятностей и математической статистики.

Из перечисленного следует, что важнейшим базовым элементом для усвоения современных методов диагностики и лечения заболеваний, контроля окружающей среды и других медицинских задач является комплекс физико-математических знаний, который и рассматривается в курсе медицинской и биологической физики.

Главы 1, 6, 7, 9, 13, 14, 27, 28 и 29 учебника написаны доцентом Г.К. Ильичом, остальные — доцентом В.Г. Лещенко.

Авторы выражают глубокую признательность доценту Н.И. Инсаровой за ценные замечания и существенный личный вклад при подготовке глав «Основы теории вероятностей», «Основы математической статистики», «Элементы биомеханики», а также искренне благодарны доктору технических наук профессору М.Г. Киселеву, кандидату физико-математических наук доценту И.М. Бертелю и кандидату технических наук доценту Н.Т. Минчене за нелегкий труд по рецензированию данного учебника.

Глава 1 Элементы высшей математики

Математические методы долгое время плодотворно используются в медицине для количественного описания и моделирования физиологических процессов, обработки и анализа медицинских данных, создания моделей возникновения, развития и прогнозирования заболеваний, для решения других важнейших задач.

Даже для общего знакомства с принципами и возможностями медицинского применения математических методов необходимо иметь представление соответствующем математическом аппарате, в частности об основополагающих разделах высшей математики — дифференциальном и интегральном исчислении.

1.1. Производная функции

Количественно описать сложные изменяющиеся процессы жизнедеятельности с помощью элементарной математики невозможно, поскольку используемые для этой цели математические величины должны сами обладать способностью к «движению». Высшая математика, в отличие от элементарной, оперирует зависимостями и величинами, подверженными изменениям, происходящим по определенным законам.

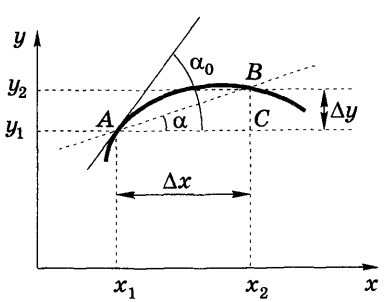


Рис. 1.1. Геометрический смысл производной

Величиной, определяющей темп изменения функциональных зависимостей в высшей математике, является производная функции. Для пояснения этого понятия рассмотрим рис. 1.1, где графически представлена некоторая произвольная функциональная зависимость $y = f(x)$.

Отметим на графике некоторые значения аргумента x_1 и

x_2, разница между которыми есть *приращение аргумента*: $\Delta x = x_2 - x_1$. *Приращение функции* равно: $\Delta y = y_2 - y_1$. Для непрерывных функций, если $\Delta x \to 0$, то и $\Delta y \to 0$. Однако к чему стремится отношение $\frac{\Delta y}{\Delta x}$ при неограниченном убывании Δx заранее сказать нельзя, это зависит от конкретного вида функции $y(x)$ и характеризует темп ее изменения.

Производной функции в данной точке называют предел отношения приращения функции к приращению аргумента при его неограниченном убывании. Производная функции одного аргумента обозначается y' или $\frac{dy}{dx}$. Таким образом,

$$y' \equiv \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \qquad (1.1)$$

Производная функции имеет простой *геометрический смысл*. Из рис. 1.1 видно, что отношение

$$\frac{\Delta y}{\Delta x} = \frac{BC}{AC} = \operatorname{tg}\alpha,$$

где α — угол наклона секущей AB к оси абсцисс.

Если Δx неограниченно убывает (x_2 стремится к x_1), то секущая вырождается в касательную к графику функции в точке A, наклоненную к оси абсцисс под углом α_0:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \operatorname{tg}\alpha_0. \qquad (1.2)$$

Таким образом, геометрический смысл производной состоит в том, что тангенс угла между касательной, проведенной к графику функции в данной точке, и осью абсцисс численно равен значению производной функции в данной точке.

К *физическому смыслу производной* подойдем, рассмотрев механическое движение. Если за время Δt тело проходит путь ΔS, то средняя скорость движения за это время

$$v_{ср} = \frac{\Delta S}{\Delta t}.$$

Но на пути ΔS скорость может иметь различные мгновенные значения ($v_{\text{мгн}}$), которые определяются как предел отношения ΔS к Δt при $\Delta t \to 0$:

$$v_{\text{мгн}} = \lim_{\Delta t \to 0} \frac{\Delta S}{\Delta t} = \frac{dS}{dt}. \qquad (1.3)$$

Следовательно, мгновенная скорость движения в данной точке представляет собой значение в данный момент времени производной от пути по времени.

Итак, производная по времени имеет смысл скорости некоторого процесса. Если рассматривается ускорение (a) механического прямолинейного движения, то мгновенное ускорение представляет собой первую производную от скорости или вторую производную от пути:

$$a_{\text{мгн}} = \frac{dv}{dt} = \frac{d}{dt}\left(\frac{dS}{dt}\right) = \frac{d^2S}{dt^2}. \qquad (1.4)$$

Следовательно, вторая производная пути по времени имеет физический смысл ускорения.

Если некоторая величина y зависит от пространственной координаты x, то производная $\frac{dy}{dx}$ характеризует скорость пространственного изменения величины y. Производная функции по пространственной координате называется ***градиентом*** этой функции.

Поясним смысл градиента. Например, некоторое вещество неравномерно распределено вдоль координаты x, т.е. его концентрация C является некоторой функцией x (рис. 1.2). Тогда скорость изменения концентрации вдоль x определяется производной $\frac{dC}{dx}$ и называется градиентом концентрации.

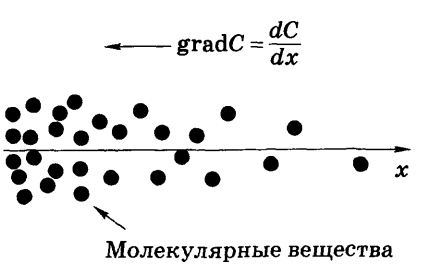


Рис. 1.2. Пояснение понятия «градиент концентрации»

Аналогично существуют градиенты давления, темпе-

ратуры и других величин. В физике и математике градиент $\frac{df}{dx}$ — это вектор, который направлен в сторону возрастания величины f.

Градиенты характеризуют неравномерность распределения величины в пространстве и являются одной из первопричин обменных процессов, происходящих в физических и биологических системах.

Так, наличие градиента концентрации является причиной и количественной мерой диффузии, заставляющей молекулы вещества распределяться более равномерно, перемещаясь в направлении меньшей концентрации, т.е. противоположно вектору градиента концентрации. Аналогично перенос тепла в некотором направлении (теплообмен) возникает при наличии градиента температуры; движение электрически заряженных частиц побуждается градиентом электропотенциала и т.п.

1.2. Исследование функции с помощью производной. Нахождение экстремумов функции

Для исследования функции $y(x)$ широко применяются ее производные $y'(x)$ и $y''(x)$. Их используют для нахождения областей возрастания и убывания функции, для нахождения и исследования ее экстремумов.

В точках экстремумов производная функции $y'(x)$ всегда равна нулю. Эти точки, а также точки разрывов функции делят всю область ее определения на участки, в которых знак производной сохраняется постоянным (рис. 1.3):

- если $y'(x) < 0$, то функция $y(x)$ на этом участке *убывает*: (x_1, x_2), (c, d);
- если $y'(x) > 0$, то функция $y(x)$ на этом участке *возрастает*: (a, x_1), (x_2, x_3) и (x_3, b);
- если $y'(x) = 0$, то функция $y(x)$ на этом участке *постоянна*: (b, c).

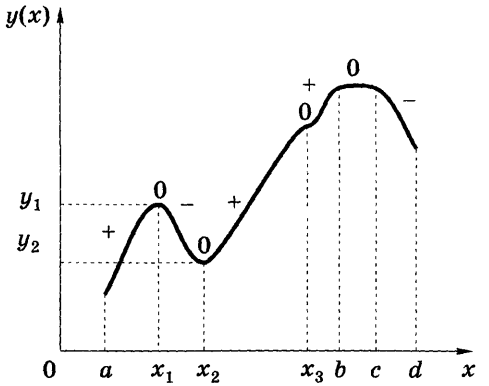


Рис. 1.3. Максимум (x_1, y_1), минимум (x_2, y_2) функции, области ее возрастания (+), убывания (–) и постоянства (bc). Сверху проставлены знаки и нули производной функции

Если $y'(x) = 0$ только в отдельных точках x_1, x_2, x_3, ..., то функция в этих точках может иметь ***экстремум***, т.е. максимум или минимум.

Функция $y = f(x)$ имеет *максимум* в точке x_1, если ее значение в этой точке больше значений функции в любых ближайших соседних точках: $f(x_1) > f(x_1 - \Delta x)$ и $f(x_1) > f(x_1 + \Delta x)$.

Функция $y = f(x)$ имеет *минимум* в точке x_2, если ее значение в этой точке меньше значений функции в любых ближайших соседних точках: $f(x_2) < f(x_2 - \Delta x)$ и $f(x_2) < f(x_2 + \Delta x)$.

Как уже отмечалось, в точках экстремумов производная функции всегда равна нулю: $y'(x) = 0$. Это свойство и используют для нахождения экстремумов функции.

Порядок исследования функции $y = f(x)$ на экстремум:

1. Найти производную функции $y'(x)$.

2. Решить уравнение $y'(x) = 0$. Корни этого уравнения x_1, x_2, x_3, ..., x_n являются точками, подозрительными на наличие экстремума функции.

3. Для установления вида экстремума в этих точках пользуются двумя способами.

1-й способ. Определить знак производной $y'(x)$ слева и справа от подозрительной точки x_i, т.е. в соседних точках $(x_i - \Delta x)$ и $(x_i + \Delta x)$:

➢ если при переходе через подозрительную точку x_i знак производной изменяется с «–» на «+», т.е. $y'(x_i - \Delta x) < 0$, а $y'(x_i + \Delta x) > 0$, то в точке x_i будет *минимум* функции (–, 0, +);

➢ если при переходе через подозрительную точку x_i знак производной изменяется с «+» на «–», т.е. $y'(x_i - \Delta x) > 0$, а $y'(x_i + \Delta x) < 0$, то в точке x_i *максимум* функции (+, 0, –);

➢ если при переходе через подозрительную точку x_i знак производной не изменяется (–, 0, –) или (+, 0, +), то в точке x_i экстремума нет.

2-й способ. Найти вторую производную $y''(x)$ исследуемой функции и определить ее знак в подозрительных точках:

➢ если $y''(x_i) < 0$, то в точке x_i будет *максимум* функции;

➢ если $y''(x_i) > 0$, то в точке x_i будет *минимум* функции;

➢ если $y''(x_i) = 0$, то экстремум функции в точке x_i *не определен* и необходимо воспользоваться 1-м способом

Второй способ обычно менее трудоемкий и поэтому более предпочтительный.

1.3. Дифференциал функции

Дифференциалом аргумента dx называют приращение аргумента Δx, т.е. $dx = \Delta x$.

Дифференциалом функции dy называют произведение производной функции на приращение (или дифференциал) аргумента:

$$dy = y'\Delta x = y'dx. \qquad (1.5)$$

Дифференциал функции dy не равен ее приращению Δy, но представляет собой его главную часть. В этом заключается его аналитический смысл. При малых приращениях аргумента можно считать, что $\Delta y \approx dy$.

Из смысла дифференциала следует его важное практическое применение: дифференциал функции позволяет приближенно определить, насколько изменяется функция при небольших изменениях ее аргумента.

Пример. Имеется куб с длиной ребра $l = 1$ м. На какую величину ΔV изменится объем куба, если длина его ребра увеличится на $\Delta l = 1$ см?

Эту задачу можно решить и методами элементарной математики:

$$\Delta V = (l + \Delta l)^3 - l^3 = 3l^2\Delta l + 3l(\Delta l)^2 + (\Delta l)^3.$$

Однако даже в данном элементарном примере необходимо выполнять довольно длительные вычисления.

Учитывая, что приращение объема куба (функции) при малых изменениях длины его ребра (аргумента) примерно равно дифференциалу объема, получим

$$\Delta V \approx dV = (l^3)'\Delta l = 3l^2\Delta l = 3 \cdot 1 \cdot 0{,}01 = 0{,}03 \text{ м}^3.$$

Полученное значение изменения объема крайне незначительно отличается от рассчитанного по точной формуле.

1.4. Частные производные

Понятие производной было введено для функции одной переменной. Однако часто необходимо количественно описывать процессы, зависящие от нескольких переменных. Например, состояние организма можно рассматривать как некоторую функцию $U(x, y, z, ..., t)$, значение которой зависит от ряда параметров $x, y, z, ..., t$, где x может означать температуру тела; y — систолическое давление; z — содержание гемоглобина в крови и т.д. Как определить производную такой функции?

Для функций многих переменных вводят понятие ***частной производной***, которая характеризует скорость изменения функции только по одной из независимых переменных, в то время как остальные переменные считаются не изменяющимися. Частная производная от функции двух переменных $U(x, y)$ по переменной x:

$$U'_x \equiv \frac{\partial U(x, y)}{\partial x} = \lim_{\Delta x \to 0} \frac{U(x+\Delta x, y) - U(x, y)}{\Delta x}. \qquad (1.6)$$

В числителе данной дроби — разность значений функции U при «наращенном» $(x + \Delta x)$ и прежнем (x) значениях аргу-

мента, а в знаменателе — приращение этого аргумента Δx. При этом аргумент y не меняется.

Аналогичным образом определяется частная производная функции $U(x, y)$ по переменной y:

$$U'_y \equiv \frac{\partial U(x, y)}{\partial y} = \lim_{\Delta y \to 0} \frac{U(x, y + \Delta y) - U(x, y)}{\Delta y}. \qquad (1.7)$$

Нахождение частных производных (дифференцирование функций многих переменных) не представляет особых сложностей.

Пример. Нужно найти частные производные функции $U = x^3 \sin y$.

Вычисляя частную производную от функции U по x, остальные переменные считаем зафиксированными и обращаемся с ними при нахождении производной как с постоянными величинами, т.е. в данном случае обращаемся со множителем $\sin y$ как с постоянной величиной:

$$U'_x \equiv \frac{\partial U}{\partial x} = 3x^2 \sin y .$$

Аналогично поступаем при нахождении частной производной по y, считая теперь множитель x^3 постоянной величиной и определяя производную от $\sin y$:

$$U'_y \equiv \frac{\partial U}{\partial y} = x^3 \cos y .$$

1.5. Частные дифференциалы. Полный дифференциал функции

Рассмотрим функцию нескольких (для простоты — двух) переменных $U(x, y)$.

Частным дифференциалом этой функции по x (обозначение $d_x U$) называют произведение частной производной по x и дифференциала аргумента dx:

$$d_x U = \frac{\partial U}{\partial x} dx.$$

Аналогичным образом определяют частный дифференциал по y:

$$d_y U = \frac{\partial U}{\partial y} dy .$$

Сумма всех частных дифференциалов называется ***полным дифференциалом функции.*** Для функции $U(x, y)$ полный дифференциал

$$dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy . \qquad (1.8)$$

Частные дифференциалы позволяют оценить изменение функции при небольших изменениях одного из ее соответствующих аргументов, а полный дифференциал — изменение значения функции при изменении всех переменных, от которых она зависит.

Пример. Рассмотрим параллелепипед с длинами ребер: $x = 1$ м, $y = 2$ м, $z = 3$ м. Допустим, первые два ребра увеличились на $\Delta x = 0{,}01$ м и $\Delta y = 0{,}02$ м, а третье уменьшилось на $\Delta z = -0{,}01$м. На какую величину ΔV изменился объем параллелепипеда?

Эту задачу можно решить и без применения высшей математики:

$$\Delta V = (x + \Delta x)(y + \Delta y)(z + \Delta z) - xyz .$$

Однако использование полного дифференциала позволяет существенно упростить решение. Поскольку изменения длин ребер по сравнению с их первоначальными значениями невелики, будем считать, что искомое приращение ΔV объема V (функции трех переменных — длин ребер) примерно равно полному дифференциалу dV:

$$\Delta V \approx dV = d(xyz) = \frac{\partial V}{\partial x}\Delta x + \frac{\partial V}{\partial y}\Delta y + \frac{\partial V}{\partial z}\Delta z.$$

Найдя частные производные и подставив в это выражение численные значения, получим

$$dV = yz\Delta x + xz\Delta y + xy\Delta z = 2 \cdot 3 \cdot 0{,}01 +$$

$$+ 1 \cdot 3 \cdot 0{,}02 - 1 \cdot 2 \cdot 0{,}01 = 0{,}10 \text{ м}^3 .$$

1.6. Первообразная функция и неопределенный интеграл

В элементарной математике сложение и вычитание, умножение и деление, возведение в степень и извлечение корня — примеры взаимообратных математических операций. Эти операции, последовательно примененные к одному и тому же числу, самого числа не изменяют:

$$a+b-b=a,\ \frac{ab}{b}=a,\qquad \sqrt[n]{a^n}=a.$$

Взаимообратные операции существуют и в высшей математике. *Дифференцирование* (нахождение производной) позволяет по некоторой заданной функции найти скорость ее изменения. Операцией, обратной дифференцированию, является *интегрирование* — нахождение самой функции (первообразной) по заданной скорости ее изменения.

Функцию $F(x)$ называют ***первообразной*** для функции $f(x)$, если для всех x из области определения функции $F'(x)=f(x)$ или $dF(x)=f(x)dx$.

Иными словами, первообразная функция — это такая функция, производной от которой является заданная. Например, задана функция $f(x)=2x$. Ее первообразной функцией будет $F(x)=x^2$, так как $F'(x)=f(x)=2x$. Однако $F(x)=x^2+C$, где C — произвольная постоянная, также будет первообразной для $f(x)=2x$, поскольку $(x^2+C)'=2x$.

Совокупность всех первообразных функций для заданной функции $f(x)$ называют ***неопределенным интегралом***:

$$\int f(x)dx=F(x)+C.$$

1.7. Определенный интеграл

К понятию определенного интеграла подойдем путем рассмотрения геометрической задачи. Допустим, некоторая функция задана в виде графика, представленного

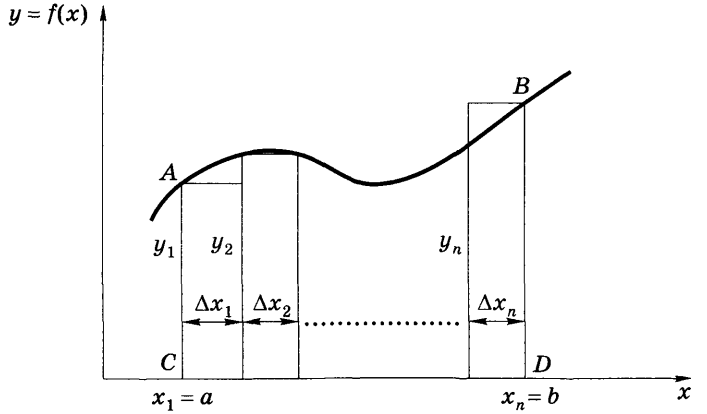


Рис. 1.4. Вычисление площади криволинейной трапеции

на рис. 1.4, и требуется вычислить площадь криволинейной трапеции S_{ABDC}, образованной графиком функции, осью абсцисс и ординатами, восстановленными из точек $x_1 = a$ и $x_n = b$.

Приближенно значение искомой площади можно найти, разбив криволинейную трапецию на отдельные прямоугольники и сложив их площади. Основаниями этих прямоугольников служат малые интервалы Δx_1, Δx_2, Δx_3, ..., Δx_n, а высотами — ординаты y_1, y_2, y_3, ..., y_n. Если основания Δx_i малы, то

$$S_{ABDC} \approx \sum_{i=1}^{n} y_i \Delta x_i. \qquad (1.9)$$

Соотношение (1.9) выполняется тем точнее, чем меньше основания прямоугольников Δx_i. Точное же значение искомой площади можно найти при предельном переходе:

$$S_{ABDC} = \lim_{\Delta x_i \to 0} \sum_{i=1}^{n} y_i \Delta x_i = \int_a^b y dx. \qquad (1.10)$$

Сумма всех произведений $y_i \Delta x_i$, стоящая под знаком предела, называется ***интегральной суммой***, а ее предел — ***опре-***

деленным интегралом от функции $y = f(x)$ на участке $[a, b]$. Значения a и b называют *нижним* и *верхним пределами интегрирования* соответственно.

Из решения рассмотренной геометрической задачи следует:

1. Определенный интеграл имеет геометрический смысл площади фигуры, ограниченной графиком функции, осью абсцисс и ординатами, восстановленными из значений аргумента в пределах интегрирования.

2. Саму операцию интегрирования можно описать как сложение бесконечно большого количества бесконечно малых величин. Действительно, в формуле (1.10) при $\Delta x_i \to 0$ каждое слагаемое $y_i \Delta x_i \to 0$, т.е. является бесконечно малой величиной, каждый отдельный прямоугольник стремится выродиться в линию, но число этих слагаемых стремится к бесконечности.

Определенный интеграл вычисляется с помощью формулы Ньютона — Лейбница:

$$\int_a^b f(x)dx = F(x)\Big|_a^b = F(b) - F(a). \qquad (1.11)$$

Для нахождения определенного интеграла необходимо найти для подынтегральной функции $f(x)$ первообразную $F(x)$ и взять разность значений этой функции на верхнем и нижнем пределах интегрирования.

Пример 1. Возьмем параболу $y = x^2$ и поставим задачу: вычислить площадь S фигуры, образованной графиком параболы, осью x и ординатами, восстановленными из значений $x_1 = -1$ и $x_2 = 2$. На рис. 1.5 эта искомая площадь заштрихована.

Решение данной задачи сводится к нахождению определенного интеграла:

$$S = \int_{-1}^{2} x^2\,dx = \frac{x^3}{3}\Big|_{-1}^{2} = \frac{8}{3} - \left(-\frac{1}{3}\right) = \frac{9}{3} = 3.$$

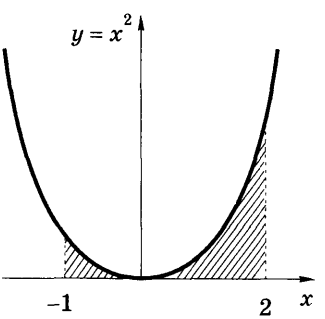


Рис. 1.5. Нахождение площади фигур интегрированием

Пример 2. Скорость движения тела изменяется во времени по закону $v = 3t^2 - 2t$ м/с. Нужно определить, какой путь L пройдет тело за 5 с от начала движения.

Решение сводится к нахождению определенного интеграла

$$L = \int_0^5 \left(3t^2 - 2t\right)dt = \left(t^3 - t^2\right)\Big|_0^5 = 125 - 25 - \left(0 - 0\right) = 100 \text{ м}.$$

1.8. Дифференциальные уравнения

Дифференциальным называют уравнение, связывающее аргумент x, искомую функцию y и ее производные y', y'', ..., y^n различных порядков. В общем виде дифференциальное уравнение можно записать так:

$$F(x, y, y', y'', ..., y^n) = 0.$$

Порядок дифференциального уравнения определяется наивысшим порядком входящей в него производной.

Примером дифференциального уравнения является второй закон Ньютона, определяющий силу F как произведение массы тела m и приобретенного под действием силы ускорения a: $F = ma$.

Учитывая, что ускорение есть первая производная от скорости v, запишем второй закон Ньютона в виде дифференциального уравнения первого порядка:

$$F = m\frac{dv}{dt}. \qquad (1.12)$$

Поскольку ускорение является второй производной от пути S, этот закон можно представить в виде дифференциального уравнения второго порядка:

$$F = m\frac{d^2S}{dt^2}. \qquad (1.13)$$

Если известна зависимость $F(t)$ действующей на тело силы от времени, то, решив уравнение (1.13), найдем закон движения тела, т.е. зависимость пройденного пути от времени $S(t)$.

Решением дифференциального уравнения является функция, которая обращает это уравнение в тождество. Например, нужно решить уравнение

$$y' - x = 0. \tag{1.14}$$

Представим это исходное уравнение в виде

$$\frac{dy}{dx} = x \Rightarrow dy = xdx. \tag{1.15}$$

В уравнении (1.15) выполнено разделение переменных: искомая функция y и ее дифференциал dy вынесены в одну часть уравнения, а аргумент x и его дифференциал dx — в другую.

Теперь произведем интегрирование его левой и правой частей:

$$\int dy = \int xdx \Rightarrow y + C_1 = \frac{x^2}{2} + C_2 . \tag{1.16}$$

При нахождении неопределенных интегралов появляются произвольные постоянные C_1 и C_2. Их следует объединить в одну постоянную C:

$$y = \frac{x^2}{2} + C . \tag{1.17}$$

Различают общее и частное решения дифференциального уравнения.

Формула (1.17) является *общим решением* дифференциального уравнения (1.14), содержащим столько произвольных постоянных, каков порядок дифференциального уравнения.

Легко доказать, что функция (1.17) действительно является решением уравнения (1.14), поскольку ее подстановка в это уравнение обращает последнее в тождество.

При любом конкретном значении постоянной C в функции (1.17) получим *частное решение* дифференциального уравнения (1.14). Частных решений может быть бесконечное множество. Чтобы из этого множества решений выбрать одно, необходимо задать некое дополнительное условие, используя

которое найдем значение постоянной C и нужное частное решение.

Например, зададим начальное условие $y = 1$ при $x = 0$. Подставив эти значения x и y в общее решение (1.17), найдем постоянную C:

$$1 = 0 + C \Rightarrow C = 1.$$

Подставляя данное значение C в общее решение (1.17), получим частное решение, удовлетворяющее заданному условию:

$$y = \frac{x^2}{2} + 1. \qquad (1.18)$$

1.9. Этапы решения задач при использовании дифференциальных уравнений

Дифференциальные уравнения — математический аппарат, позволяющий не только решать чисто математические или физические задачи, но и количественно описывать самые разнообразные процессы (медико-биологические, социальные и др.). Несмотря на разнообразие рассматриваемых явлений, использовать аппарат дифференциальных уравнений для их исследования необходимо в общей логической последовательности.

1. *Составление дифференциального уравнения.* Это наиболее сложный и ответственный этап, требующий учета всех факторов, влияющих на течение исследуемого процесса. С этой целью можно сделать некоторые допущения, определить начальные условия, основываясь на твердо установленных экспериментальных фактах или на логических посылках. Например, практическая польза математических моделей работы сердца (получение новых сведений, позволяющих улучшить диагностику сердечно-сосудистых заболеваний и повысить эффективность их лечения) обусловливается полнотой и корректностью математического учета физиологических данных и клинической практики.

2. *Решение уравнения.* Этот этап может считаться более простым, чем первый, поскольку он предполагает выполне-

ние чисто математических операций. Если получить решение дифференциального уравнения в аналитическом виде невозможно, то это можно сделать расчетным путем с применением современной вычислительной техники.

3. *Оценка и анализ результата.* Решив дифференциальное уравнение (или систему уравнений), необходимо оценить теоретическую и практическую пользу полученных результатов, т.е. выяснить, установлены ли новые закономерности протекания того или иного процесса (например, физиологического), определено ли количественно влияние учтенных факторов (к примеру, на степень развития и характер патологии) и т.п.

Кроме того, следует сопоставить полученные результаты с установленными фактами. Если из математического описания физиологического процесса следуют неожиданные и неизвестные ранее выводы, то это может означать один из двух вариантов:

1) действительно установлено новое явление, которое впоследствии может быть подтверждено экспериментальными исследованиями;

2) полученный результат является следствием того, что на этапе составления дифференциального уравнения не были учтены все необходимые факторы или сделаны слишком грубые допущения.

1.10. Примеры использования дифференциальных уравнений

В соответствии с описанными этапами использования аппарата дифференциальных уравнений применим его для рассмотрения некоторых задач.

Пример 1. Необходимо определить характер движения тела (зависимость пути S от времени t), если сила F на тело не действует.

Дифференциальное уравнение, описывающее поведение тела в этом случае, представляет второй закон Ньютона:

$$m\frac{dv}{dt} = 0 \Rightarrow \frac{dv}{dt} = 0 \Rightarrow v = \text{const}\ .$$

Считая, что масса $m \neq 0$, получим, что ускорение равно нулю, и движение осуществляется с постоянной скоростью v.

Для установления зависимости $S = f(t)$ нужно решить уравнение

$$\frac{dS}{dt} = v \Rightarrow dS = vdt \Rightarrow S = v\int dt. \qquad (1.19)$$

В результате интегрирования уравнения (1.19) получим

$$S = vt + C \Rightarrow S = vt + S_0, \qquad (1.20)$$

где C — произвольная постоянная, которая равна пути, пройденному к начальному моменту времени (определяется из начального условия при $S = S_0$ $t = 0$).

Выражение (1.20) представляет собой уравнение равномерного прямолинейного движения. Таким образом, если на тело не действует сила ($F = 0$), то тело сохраняет состояние покоя [частный случай $v = 0$ в формуле (1.20)] или равномерного прямолинейного движения (первый закон Ньютона).

Рассмотрим теперь микробиологическую задачу.

Пример 2. Необходимо установить закон изменения со временем t численности бактерий n, помещенных в питательную среду.

Для составления дифференциального уравнения, отражающего существование бактерий в этих условиях, необходимо иметь сведения об определенном факте, который следует записать в математической форме. На основании экспериментальных данных и общих соображений таким фактом может служить утверждение, что скорость размножения бактерий (математически $\frac{dn}{dt}$) пропорциональна их числу n в данный момент времени.

Таким образом, необходимое дифференциальное уравнение должно иметь следующий вид:

$$\frac{dn}{dt} = kn, \qquad (1.21)$$

где k — доступный экспериментальному определению коэффициент пропорциональности, зависящий от вида бактерий и параметров среды их обитания.

Дополнительные данные, необходимые для решения задачи, следуют из начального условия $n = n_0$ при $t = 0$, т.е. в начальный момент времени количество бактерий считается известным и равным n_0.

Для решения уравнения (1.21) произведем разделение переменных и последующее интегрирование:

$$\int \frac{dn}{n} = k\int dt \Rightarrow \ln|n| = kt + \ln C\,. \qquad (1.22)$$

Произвольную постоянную в уравнении (1.22) удобно представить в виде $\ln C$. Из начального условия $C = n_0$.

Решив логарифмическое уравнение (1.22) с учетом начального условия, получим искомый закон изменения числа бактерий со временем:

$$n = n_0 e^{kt}. \qquad (1.23)$$

Проанализируем полученный результат, определив, в чем заключается его практическая польза, а также возможные более отдаленные выводы.

Зная коэффициент k и начальное число бактерий n_0, легко определить их число в любой момент времени t.

Прирост бактериальной массы определяется через коэффициент k условиями среды обитания бактерий: чем больше значение k, тем быстрее увеличивается число бактерий (рис. 1.6). Если существуют факторы, препятствующие размножению бактерий (пониженная температура, ионизирующие излучения и др.), то коэффициент k в формулах (1.21)–(1.23) уменьшается и может принять отрицательное значение — в этом случае будет иметь место гибель бактерий.

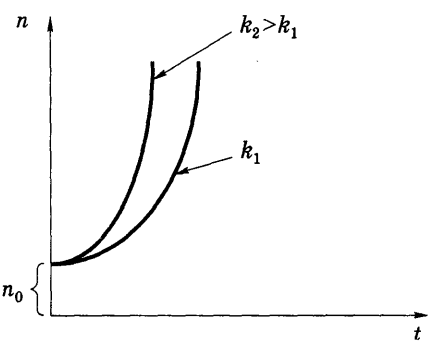


Рис. 1.6. Графики изменения числа бактерий n со временем при различных коэффициентах k

С некоторым риском можно попытаться придать полученному для бактерий результату (1.23) большую общность и сформулировать утверждение: любой биологический вид, находясь в оптимальных для своего существования условиях, экспоненциально увеличивает свою численность со временем. Так, кролики, завезенные в Австралию, где практически нет хищников, которые бы ими питались, увеличили свою численность в соответствии с формулой (1.23) и стали представлять серьезную опасность для сельского хозяйства.

Рассмотренные примеры касались количественного описания с помощью дифференциальных уравнений сравнительно

простых явлений. Естественно, что изучение более сложных задач требует применения более сложного математического аппарата. Однако использование математических методов при анализе процессов жизнедеятельности, изучении внешних воздействий на организм, разработке методов диагностики и лечения совершенно необходимо для понимания сущности этих явлений и несомненно имеет практическое значение.

Контрольные вопросы и задания

1. Дайте определение производной функции, раскройте ее геометрический и физический смысл. Что такое градиент функции, как он направлен?
2. Каков физический смысл второй производной от пути по времени?
3. Что такое экстремумы функции и каковы этапы исследования функций на экстремум?
4. Дайте определение дифференциала функций одной переменной. Проиллюстрируйте на графике функции геометрический смысл дифференциала и сравните его с приращением функции.
5. Дайте определение частных производных. Каков их физический смысл?
6. Что такое частный и полный дифференциал функций? Как используется понятие полного дифференциала для оценки изменения функции многих переменных и в приближенных вычислениях значений функции?
7. Что такое первообразная функция и неопределенный интеграл?
8. Что такое определенный интеграл и каков его геометрический смысл?
9. В чем состоит правило Ньютона—Лейбница для вычисления определенных интегралов?
10. Какое уравнение называют дифференциальным? Чем определяется порядок дифференциального уравнения? Что является его решением?
11. Чем отличаются общее и частное решения дифференциального уравнения? Как из общего решения получить частное?
12. Как проверить, является ли некоторая функция решением заданного дифференциального уравнения?
13. В чем сущность метода разделения переменных, применяемого для решения некоторых простых дифференциальных уравнений?

Глава 2 Основы теории вероятностей

Предмет изучения биологов и медиков — живой организм, зарождение, развитие и существование которого определяется разнообразными, часто случайными, внешними и внутренними условиями. Элементы неопределенности, сложности, многопричинности, присущие случайным явлениям, необходимость установления закономерностей среди случайных явлений — все это потребовало создания вероятностных математических методов для их изучения.

Характерно, что эти закономерности выполняются лишь при массовом проявлении случайных явлений. При этом индивидуальные особенности отдельных событий «взаимно погашаются», а усредненный результат для массы случайных явлений оказывается уже не случайным, а вполне закономерным.

Данное обстоятельство в значительной мере способствовало широкому распространению вероятностных методов исследования в биологии и медицине, где развитие процессов часто зависит от многочисленных, практически не поддающихся учету случайных факторов.

Теория вероятностей — математический аппарат, предназначенный для количественного описания случайных событий, которые при реализации комплекса условий могут произойти или не произойти. Количественной мерой возникновения случайного события является его вероятность.

2.1. Классическое и статистическое определение вероятности

Допустим, некоторое событие A происходит в m случаях из n возможных. Тогда ***классической вероятностью*** $P(A)$ события A называют отношение числа случаев m, благоприятствующих появлению данного события, к общему числу n возможных исходов:

$$P(A) = m/n. \qquad (2.1)$$

Например, экспериментально установлено, что за 8 суток распадается половина атомов одного из радиоактивных изотопов йода. Чему же будет равна вероятность распада одного отдельного атома этого изотопа иода за 8 суток? В данном случае $m = n/2$. Значит, искомая вероятность $P = 0{,}5$.

Далее, пусть в некоторой урне содержится 20 черных и 80 белых шаров, тогда вероятность извлечения наугад из урны черного шара составит $20/100 = 0{,}2$.

Из определения и приведенных примеров следует, что классическая вероятность события принимает значения между нулем и единицей: $0 \le P(A) \le 1$. Если $P(A) = 1$, то событие A достоверно, если же $P(A) = 0$, то оно невозможно.

Рассмотрим теперь статистическое определение вероятности. Проведем простой эксперимент. Допустим, из общего числа N бросков монеты число выпадений герба составило M, решки — K. Если в совокупности случаев N некоторое событие осуществилось M раз, то величину M называют частотой этого события, а отношение M/N — его относительной частотой. При случайном бросании монеты данные табл. 2.1 окажутся самыми разными. Но их всегда будет характеризовать общая особенность: при увеличении числа бросков N относительные частоты этих событий (выпадение герба или решки) все больше приближаются к значению 0,5, которое и определяет их статистическую вероятность.

Таблица 2.1

N	M	K	M/N	K/N
10	4	6	0,40	0,60
100	55	45	0,55	0,45
1000	490	510	0,49	0,51

Таким образом, ***статистической вероятностью*** события называют предел, к которому стремится относительная частота события при неограниченном возрастании общего числа испытаний:

$$P(A) = \lim_{N \to \infty} M/N \ . \qquad (2.2)$$

2.2. Виды случайных событий

Все случайные события можно подразделить на три основных вида.

Несовместными называют события, которые не могут происходить одновременно, поскольку наступление одного из них исключает наступление другого:

- разные оценки, выставляемые по контрольной работе;
- извлечение из урны одного из имеющихся там шаров и др.

Несовместные события A_1, A_2, A_3, ..., A_n образуют полную группу событий, если при испытаниях может происходить только одно из этих событий, и никакое другое.

Независимые события могут происходить одновременно и независимо друг от друга:

- оценки, получаемые по контрольной работе разными студентами;
- оценки, получаемые студентом по разным предметам;
- извлечение по одному шару из разных урн и др.

Зависимыми называют события, если вероятность наступления одного из них зависит от того, произошло или нет другое событие. Поэтому для зависимых событий A и B рассматриваются два вида вероятностей: безусловная и условная.

Безусловной вероятностью $P(B)$ события B называют вероятность его осуществления до наступления события A.

Условной вероятностью события B называют вероятность его осуществления при условии, что событие A уже произошло, что обозначается как $P(B/A)$:

- последовательное извлечение шаров из урны, если они в нее не возвращаются. Так, если в урне 3 черных и 7 белых шаров, то вероятность достать черный шар первым (безусловная вероятность) равна 0,3, а вероятность достать его при второй попытке (условная вероятность) будет уже другой: 2/9, если первым был черный шар или 3/9, если первым достали белый шар;
- вероятность угадать правильный номер при игре в лото изменяется по мере извлечения номеров из игрового барабана и др.

2.3. Теоремы сложения и умножения вероятностей

Вероятность наступления нескольких случайных событий зависит как от вероятностей этих событий, так и от вида событий (несовместные, независимые или зависимые) и определяется соответствующими теоремами.

Вероятность совместного проявления нескольких событий (происходит и одно, и другое событие) определяют теоремы умножения вероятностей, которые несколько различны для независимых и зависимых событий.

Несовместные события одновременно проявляться не могут (по определению этих событий) и для них справедлива теорема сложения вероятностей.

Теорема сложения вероятностей: вероятность наступления одного (безразлично какого) события из нескольких несовместных событий A_1, A_2, ..., A_n равна сумме вероятностей этих событий:

$$P(A_1 \text{ или } A_2, \ldots, \text{ или } A_n) = P(A_1) + P(A_2) + \ldots + P(A_n). \quad (2.3)$$

Пример. При бросании случайным образом игральной кости — кубика, на гранях которого выгравированы цифры от 1 до 6, — классическая вероятность выпадения одной конкретной цифры составит 1/6, а вероятность выпадения любого нечетного числа — 0,5, поскольку при случайном бросании игральной кости вероятность выпадения нечетного числа равна

$$P(1, \text{ или } 3, \text{ или } 5) = P(1) + P(3) + P(5) = 3 \cdot 1/6 = 1/2.$$

Если несовместные события A_1, A_2, ..., A_n образуют полную группу событий, то сумма вероятностей всех этих событий равна единице:

$$P(A_1, \text{ или } A_2, \ldots, \text{ или } A_n) = P(A_1) + P(A_2) + \ldots + P(A_n) = 1. \quad (2.4)$$

В частности, если полную систему событий образуют только два события A и B, то сумма их вероятностей всегда равна единице. Поэтому, зная вероятность одного события, легко найти вероятность второго: $P(A) = 1 - P(B)$.

Теорема умножения для независимых событий: вероятность совместного проявления нескольких независимых событий A_1, A_2 и A_3 равна произведению их вероятностей:

$$P(A_1, A_2, A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3). \quad (2.5)$$

Пример 1. Найдем вероятность того, что при следующих друг за другом двух бросках игральной кости выпадут числа 1 и 5.

Согласно теореме умножения вероятностей находим:

$$P(1\text{ и }5) = P(1)P(5) = \frac{1}{6}\cdot\frac{1}{6} = \frac{1}{36}.$$

Пример 2. В одной урне находятся 2 черных и 8 белых шаров, а в другой — 6 черных и 4 белых. Какова вероятность одновременно извлечь из обеих урн по белому шару?

Вероятность достать белый шар из 1-й урны равна 0,8, а из второй — 0,4, поэтому искомая вероятность равна:

$$P(2\text{ бел. ш.}) = 0{,}8 \cdot 0{,}4 = 0{,}32.$$

Теорема умножения для двух зависимых событий: вероятность совместного появления двух независимых событий равна произведению безусловной вероятности первого события A и условной вероятности второго события B:

$$P(A \text{ и } B) = P(A) \cdot P(B/A). \qquad (2.6)$$

Если же первым происходит событие B, то для него реализуется безусловная вероятность $P(B)$, а для события A реализуется условная вероятность $P(A/B)$. Теперь вероятность их совместного появления равна:

$$P(A \text{ и } B) = P(B) \cdot P(A/B). \qquad (2.6а)$$

Пример. В урне находятся 3 черных шара и 7 белых. Какова вероятность того, что из этой урны последовательно один за другим будут вынуты белые шары?

Вероятность извлечения первого белого шара составит 7/10, после чего в ней останется 6 белых шаров при общем количестве 9. Вероятность извлечения белого шара во второй раз составит 6/9, а искомая вероятность последовательного изъятия двух белых шаров равна:

$$P(2\,\text{бел.ш.}) = \frac{7}{10}\cdot\frac{6}{9} = \frac{42}{90}.$$

Приравнивая формулы (2.6) и (2.6а), можно выразить условную вероятность события B:

$$P(B/A) = \frac{P(B)\cdot P(A/B)}{P(A)}. \qquad (2.7)$$

Важным расширением формулы (2.7) на случай многих зависимых событий является ***формула Байеса.*** Пусть некоторые события B_1, ..., B_k, ..., B_n образуют полную систему несовместных событий и известны их безусловные вероятности $P(B_k)$ (причем $\sum_{k=1}^{n} P(B_k) = 1$), а также условные вероятности $P(A/B_k)$ наступления события A после каждого из них. Тогда вероятность наступления события B_k после осуществления события A

$$P(B_k / A) = \frac{P(B_k) \cdot P(A / B_k)}{\sum_{k}^{n} P(B_k) \cdot P(A / B_k)}. \qquad (2.8)$$

Подчеркнем, что при использовании этой формулы необходимо заранее знать безусловные вероятности $P(B_k)$ несовместных событий B_k и условные вероятности $P(A/B_k)$. Заметим, что знаменатель выражения (2.8) представляет собой полную вероятность наступления события A в сочетании с любым (безразлично каким) событием B_k.

Пример. Установлено, что курящие мужчины в возрасте свыше 40 лет умирают от рака легких в 10 раз чаще, чем некурящие. Полагая, что 60 % мужчин курят, найдите вероятность того, что некоторый мужчина, умерший от рака легких, курил?

События K (курил) и H (не курил) образуют полную группу событий. По условию вероятность того, что мужчина курил $P(K) = 0{,}6$, тогда вероятность, что он не курил $P(H) = 1 - P(K) = 0{,}4$. Вероятность смерти (C) не курящего мужчины от рака легких обозначим $P(C/H)$, а вероятность смерти курящего $P(C/K) = 10 \cdot P(C/H)$. Теперь по формуле Байеса (2.8) рассчитаем вероятность того, что мужчина, умерший от рака легких, курил:

$$P(K/C) = \frac{P(K) \cdot P(C/K)}{P(K) \cdot P(C/K) + P(H) \cdot P(C/H)} =$$

$$= \frac{0{,}6 \cdot 10P(C/H)}{0{,}6 \cdot 10P(C/H) + 0{,}4P(C/H)} = \frac{6}{6{,}4} \approx 0{,}94.$$

Формула Байеса (2.8) широко применяется при решении разных задач, в том числе и при диагностике и прогнозировании заболеваний.

2.4. Принципы вероятностных подходов к задачам диагностики и прогнозирования заболеваний

Согласно классификации Всемирной организации здравоохранения, в перечне возможных диагнозов насчитывается около 10 000 наименований. Для каждого заболевания характерен соответствующий набор симптомов, общее количество которых примерно 100 000. Поэтому в общем случае установление точного диагноза весьма сложная задача: по некоторому набору симптомов врач должен определить у конкретного пациента одно или несколько заболеваний из 10^4 возможных.

Целью введения в диагностику и прогнозирование заболеваний количественных подходов является облегчение решения этих чрезвычайно важных проблем врачебной деятельности и получение более эффективных результатов.

При реализации данных подходов необходимо соблюдать два основных условия: 1) максимальным образом использовать богатейший клинический опыт; 2) эффективно следовать логике клинического мышления врача-диагноста.

Последовательность логических операций при постановке диагноза должна быть отражена и формализована с помощью определенного математического аппарата. Одним из возможных способов реализации такого подхода является использование вероятностных методов.

Составим перечень заболеваний (возможных диагнозов), обозначив их символами M_1, M_2, M_3, ..., M_i, ..., M_n. Например, M_1 — грипп, M_2 — пневмония, M_3 — острое респираторное заболевание и т.д. Полный список включал бы примерно 10^4 заболеваний. Однако на практике выбирают лишь некоторый класс болезней в существенно меньшем количестве (50–100 наименований). Из многолетней медицинской статистики можно определить вероятность $P(M_i)$ каждого заболевания в интересующем регионе. Эти заболевания должны образовывать полную группу событий, а их вероятности подчиняться условию нормировки:

$$\sum_{i=1}^{n} P(M_i) = 1 .$$

Например, обладая сведениями о числе заболеваний гриппом в г. Минске в данное время года за ряд лет, можно определить $P(M_1)$ — вероятность диагноза «грипп» для этого региона в интересующее время года.

Имеющиеся медицинские статистические данные позволяют лишь выявить распределение возможных диагнозов по их вероятностям. Естественно, без обследования пациента определить точный диагноз нельзя.

Диагностическая деятельность врача начинается с выявления симптомов. Подобно тому как составлялся перечень заболеваний, можно составить список соответствующих им симптомов S_1, S_2, ..., S_j, ..., S_k. В этом перечне, например, S_1 — повышение температуры, S_2 — кашель, S_3 — повышение систолического давления и т.д. Из совокупности клинических данных можно определить, с какой вероятностью симптом S_j проявляется при некотором заболевании M_i, т.е. определить условную вероятность $P(S_j/M_i)$. Например, $P(S_1/M_1)$ — вероятность повышения температуры при гриппе, $P(S_2/M_2)$ — вероятность обнаружения кашля при пневмонии и т.д.

Итак, при использовании вероятностных методов диагностики считаются известными из медицинской статистики распределение диагнозов по их вероятностям при случайной выборке $P(M_i)$ и вероятности проявления симптомов при различных заболеваниях $P(S_j/M_i)$.

Если у больного обнаружен некоторый симптом S_1, то формула Байеса (2.8) позволяет вычислить вероятности наличия у него каждого из заболеваний M_i:

$$P(M_i/S_1) = \frac{P(M_i)\cdot P(S_1/M_i)}{\sum_{i=1}^{n} P(M_i)\cdot P(S_1/M_i)}, \qquad (2.9)$$

где $i = 1, 2, 3, ..., n$.

Затем следует повторить расчет для другого имеющегося симптома S_2 и т.д. Наиболее вероятным будет заболевание, для которого сумма условных вероятностей (2.9), рассчитанных для разных симптомов, будет наибольшей.

Рассмотрим пример, который имеет чисто иллюстративный характер, хотя бы потому, что для упрощения расчетов выбран один симптом и всего три возможных диагноза, а исходные данные взяты произвольно.

Пример 1. Пусть выбранная для диагностики группа заболеваний включает лишь три болезни: M_1 — грипп, M_2 — пневмония, M_3 — острое респираторное заболевание (ОРЗ). Допустим, вероятности этих диагнозов, по данным многолетней статистики, составляют: $P(M_1) = 0{,}3$; $P(M_2) = 0{,}2$; $P(M_3) = 0{,}5$. Следует иметь в виду, что должно выполняться условие нормировки

$$P(M_1) + P(M_2) + P(M_3) = 1.$$

Предположим, что у пациента обнаружен симптом S — шумы в легких. Допустим далее, что вероятность обнаружения шумов в легких при гриппе $P(S/M_1) = 0{,}3$, при пневмонии $P(S/M_2) = 0{,}7$, а при ОРЗ – $P(S/M_3) = 0{,}2$. Для условных вероятностей $P(S/M_i)$ условие нормировки не обязано соблюдаться.

Рассчитаем теперь вероятности этих заболеваний при наличии у больного данного симптома S (шумов в легких) по формуле Байеса (2.9).

Вначале вычислим знаменатель формулы, поскольку он будет общим для всех трех случаев. Суммирование производится по индексу i, который принимает значения 1, 2, 3:

$$\sum_{i=1}^{3} P(M_i) \cdot P(S/M_i) = 0{,}3 \cdot 0{,}3 + 0{,}2 \cdot 0{,}7 + 0{,}5 \cdot 0{,}2 = 0{,}33 .$$

Теперь найдем вероятности наличия у больного заболеваний:

- грипп $P\left(\dfrac{M_1}{S}\right) = \dfrac{0{,}3 \cdot 0{,}3}{0{,}33} = 0{,}27$;
- пневмония $P\left(\dfrac{M_2}{S}\right) = \dfrac{0{,}2 \cdot 0{,}7}{0{,}33} = 0{,}42$;
- ОРЗ $P\left(\dfrac{M_3}{S}\right) = \dfrac{0{,}5 \cdot 0{,}2}{0{,}33} = 0{,}3$.

Отметим, что безусловные вероятности этих заболеваний составляли соответственно $P(M_1) = 0{,}3$, $P(M_2) = 0{,}2$, $P(M_3) = 0{,}5$, а это значит, что обнаружение шумов в легких снизило вероятность диагнозов «грипп» и «ОРЗ», тогда как вероятность диагноза «пневмония» возросла более чем вдвое.

Необходимо выявить следующий симптом и снова учесть его значение с помощью формулы Байеса. При этом расчете входящие в данную формулу вероятности $P(M_i)$ заменяются на полученные вероятности $P(M_i/S)$, что уже отражает результат учета симптома S — шумы в легких. Затем производят учет следующего симптома и так далее до окончательной установки диагноза.

Естественно, в реальных системах диагностики по описанному принципу расчеты с помощью формулы Байеса вручную не производят. Все упомянутые и ряд других расчетных, а также логических операций выполняются с помощью современной вычислительной техники по специальным программам, которые созданы на основе клинического опыта и формализуют логические действия по постановке диагноза.

Подобные подходы применяются на практике не только для диагностики, но и для прогнозирования заболеваний.

Пример 2. Установлено, что в среднем один из 700 детей мужского пола рождается с лишней Y-хромосомой. У таких детей крайне агрессивное поведение встречается в 20 раз чаще, чем у детей с нормальным набором хромосом. Опираясь на эти данные, нужно определить вероятность того, что у ребенка, который ведет себя крайне агрессивно, лишняя Y-хромосома.

Вероятность рождения ребенка с лишней Y-хромосомой $P(M_Y) = 1/700$, значит, вероятность рождения ребенка без нее: $P(M_0) = 699/700$. Обозначим вероятность агрессивного поведения (симптом S) у детей, не имеющих лишней Y-хромосомы, $P(S/M_0) = A$. Тогда по условию вероятность агрессивного поведения у детей с лишней Y-хромосомой $P(S/M_Y) = 20A$. Искомую вероятность того, что ребенок с проявлением агрессивного поведения имеет лишнюю Y-хромосому, найдем по формуле Байеса,

$$P(M_Y/S) = \frac{20A \cdot 1/700}{699A/700 + 20A/700} = \frac{20}{719} \approx 0{,}028.$$

Контрольные вопросы и задания

1. Какие события называют случайными? Дайте классическое и статистическое определение вероятности случайного события.
2. Какие виды случайных событий вы знаете?
3. Сформулируйте теорему сложения вероятностей. Для какого вида событий она применима?

4. Какие события образуют полную группу событий? Чему равна сумма вероятностей таких событий?
5. Сформулируйте теорему умножения вероятностей для независимых событий.
6. Какие события называются зависимыми? Что такое условная и безусловная вероятности события?
7. Сформулируйте теорему умножения вероятностей для зависимых событий.
8. Приведите формулу Байеса, объясните смысл входящих в нее величин.
9. Как используется формула Байеса в вероятностных подходах к задачам диагностики заболеваний?

Глава 3 Случайные величины и их распределение

Величина, которая принимает различные числовые значения под влиянием случайных обстоятельств, называется ***случайной***. Примеры случайных величин: число больных, ожидающих приема у врача; точные значения массы и роста людей и др.

Различают дискретные и непрерывные случайные величины.

Случайная величина называется ***дискретной***, если она принимает только определенные отделенные друг от друга значения (не обязательно целые), которые можно установить и перечислить. Например, число студентов в аудитории может быть только целым положительным числом: 0, 1, 2, 3, 4, ..., 20, ...; относительная частота попадания в цель при 10 выстрелах — 0; 0,1; 0,2; 0,3; ...; 1,0 и др.

Случайная величина называется ***непрерывной***, если она может принимать любые значения внутри некоторого интервала $[a, b]$, который иногда имеет резко выраженные границы. Если они не известны, то считают, что значения случайной величины X лежат в интервале $(-\infty; \infty)$. Непрерывными случайными величинами являются, например, температура, давление, вес и рост людей, размеры форменных элементов крови, pH крови и т.п.

Понятие случайной величины играет определяющую роль в современной теории вероятностей, в рамках которой разработаны специальные приемы перехода от случайных событий к случайным величинам.

Если случайная величина зависит от времени, то можно говорить о случайном процессе.

3.1. Закон распределения дискретной случайной величины

Для того чтобы дать полную характеристику дискретной случайной величине, необходимо перечислить все ее возможные значения и указать их вероятности.

Соответствие между возможными значениями дискретной случайной величины и их вероятностями называется ***законом распределения данной величины.***

Обозначим возможные значения случайной величины X через x_i, а соответствующие им вероятности через p_i. Тогда закон распределения дискретной случайной величины можно задать тремя способами: в виде таблицы, графика или формулы.

1. В табл. 3.1, которая называется *рядом распределения*, указываются все возможные значения дискретной случайной величины X и соответствующие этим значениям вероятности $P(X)$.

Таблица 3.1

X	x_1	x_2	...	x_i	...	x_n
$P(X)$	p_1	p_2	...	p_i	...	p_n

При этом сумма всех вероятностей p_i должна быть равна единице (*условие нормировки*), так как все возможные значения величины X образуют полную систему событий:

$$\sum_{i=1}^{n} p_i = p_1 + p_2 + \ldots + p_n = 1. \qquad (3.1)$$

2. Графически в виде ломаной линии, которую принято называть *многоугольником распределения* (рис. 3.1). Здесь по горизонтальной оси откладывают все возможные значения случайной величины x_i, а по вертикальной оси — соответствующие им вероятности p_i. Полученные точки последовательно соединяют прямыми отрезками. Таким образом, многоуголь-

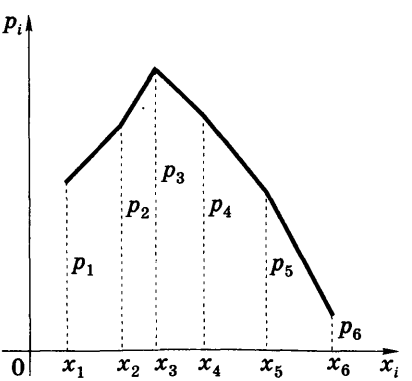


Рис. 3.1. Многоугольник распределения случайной величины X

ник распределения полностью соответствует таблице распределения.

3. Аналитически, т.е. в виде формулы, определяющей зависимость вероятности события от некоторых его характеристик. Например, если известна вероятность p попадания в цель при одном выстреле, то вероятность промаха при одном выстреле будет $q = 1 - p$, а вероятность $P_k(n)$ поражения цели k раз при n выстрелах задается формулой

$$P_k(n) = (1 - p)^{n-k} p^k.$$

Отметим, что на практике распределения дискретной случайной величины чаще всего задают в виде таблицы.

3.2. Распределение непрерывной случайной величины. Плотность распределения вероятности

Для непрерывных случайных величин невозможно применить закон распределения в формах, приведенных выше, так как непрерывная величина имеет бесчисленное («несчетное») множество возможных значений, сплошь заполняющих некоторый интервал. Поэтому составить таблицу, в которой были бы перечислены все ее возможные значения, или построить многоугольник распределения невозможно. Кроме того, вероятность какого-либо ее конкретного значения случайной величины очень мала (близка к 0). Вместе с тем различные области (интервалы) возможных значений непрерывной случайной величины обычно не являются одинаково вероятными, т.е. распределение неоднородно.

Рассмотрим непрерывную случайную величину X, возможные значения которой сплошь заполняют некоторый интервал $[a, b]$ (рис. 3.2).

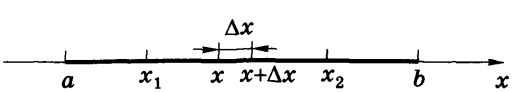


Рис. 3.2. Случайная величина X, непрерывно распределенная на интервале $[a, b]$

Рассмотрим сначала очень малый интервал значений: от x до $(x + \Delta x)$. Малая вероятность dP того, что случайная величина X примет какое-то значение из этого малого интервала $[x, x + \Delta x]$, будет пропорциональна величине этого интервала: $dP \sim \Delta x$, или, введя коэффициент пропорциональности $f(x)$, который может зависеть от величины x, получим

$$dP = f(x)\Delta x. \tag{3.2}$$

Введенная функция $f(x)$ называется ***плотностью распределения вероятностей*** случайной величины X или ***плотностью вероятности*** (плотностью распределения) этой величины. Выражение (3.2) можно рассматривать как дифференциальное уравнение и тогда вероятность того, что случайная величина X примет значение из некоторого широкого интервала (x_1, x_2), равна

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx. \tag{3.3}$$

Графически вероятность $P(x_1 < X < x_2)$ равна площади криволинейной трапеции, ограниченной осью абсцисс, кривой $f(x)$ и прямыми $X = x_1$ и $X = x_2$ (рис. 3.3), что следует из геометрического смысла определенного интеграла (3.3). Кривая $f(x)$ при этом называется ***кривой распределения.***

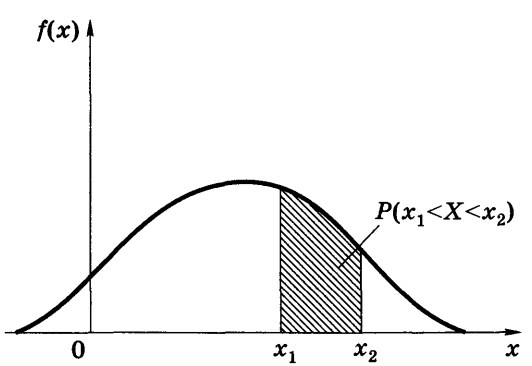


Рис. 3.3. Пример кривой распределения некоторой случайной величины X. Заштрихованная площадь численно равна вероятности принятия ею значений в интервале (x_1, x_2)

Из уравнения (3.3) следует, что если известна функция $f(x)$, то, изменяя пределы интегрирования, можно найти вероятность попадания значений случайной величины в любой интересующий интервал. Поэтому именно задание функции $f(x)$ полностью определяет закон распределения непрерывной случайной величины X.

Для плотности распределения вероятности $f(x)$ должны выполняться следующие условия нормировки:

- если известно, что все значения X лежат в интервале (a, b)

$$\int_a^b f(x)dx = 1, \tag{3.4}$$

- если точные границы интервала для X не известны

$$\int_{-\infty}^{+\infty} f(x)dx = 1. \tag{3.5}$$

Условия нормировки плотности вероятности являются следствием того, что значения случайной величины X достоверно лежат в интервале $[a, b]$ или $(-\infty, +\infty)$, и указывают на то, что площадь фигуры, ограниченной кривой распределения и осью абсцисс, всегда равна 1.

3.3. Числовые характеристики распределения случайных величин

Полную информацию о дискретной и непрерывной случайных величинах дают законы их распределения. Однако во многих практически значимых ситуациях пользуются так называемыми ***числовыми характеристиками распределения*** случайных величин, главное назначение которых — выразить в сжатой форме наиболее существенные особенности их распределения.

Важно, что эти параметры представляют собой конкретные числовые значения, которые можно рассчитать на основании имеющихся экспериментальных данных. Этими оценками занимается так называемая описательная статистика.

В теории вероятностей и математической статистике используется достаточно много различных характеристик; рассмотрим наиболее употребляемые.

3.3.1. Характеристики положения

Характеристиками положения являются математическое ожидание, мода, медиана. Именно они определяют некоторые важные значения случайной величины на числовой оси, которые характеризуют все распределение в целом.

Среди них важнейшую роль играет ***математическое ожидание*** $M(X)$ случайной величины X, которое является вероятностным аналогом ее среднего арифметического значения: $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Для дискретной случайной величины, заданной таблицей распределения 3.1, оно представляет сумму произведений значений случайной величины и соответствующих им вероятностей и вычисляется по формуле

$$M(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n = \sum_{i=1}^{n} x_i p_i \approx \overline{X}. \qquad (3.6)$$

В случае непрерывной случайной величины

$$M(X) = \int_a^b x f(x)dx, \text{ или } M(X) = \int_{-\infty}^{\infty} x f(x)dx, \qquad (3.7)$$

где $f(x)$ — плотность вероятности; $f(x)dx = dP$ — элемент вероятности (аналог p_i) для малого интервала dx.

Пример. Вычислите среднее значение непрерывной случайной величины, имеющей на отрезке $[a, b]$ равномерное распределение.

При равномерном распределении плотность вероятности на интервале $[a, b]$ постоянна, т.е. $f(x) = f_0 = \text{const}$, а вне $[a, b]$ равна нулю. Из условия нормировки (3.4) найдем значение f_0:

$$1 = \int_a^b f_0 dx = f_0 \int_a^b dx = f_0 \Big|_b^b = (b-a)f_0, \text{ откуда } f_0 = \frac{1}{b-a}.$$

Поэтому

$$M(X)\int_a^b \frac{x}{b-a}dx=\frac{1}{b-a}\int_a^b xdx=\frac{1}{2(b-a)}x^2\Big|_a^b=\frac{b^2-a^2}{2(b-a)}=\frac{1}{2}(a+b).$$

Таким образом, в рассматриваемом случае математическое ожидание $M(X)$ совпадает с серединой интервала $[a, b]$ и равно среднему арифметическому значению $\overline{X}$:

$$\overline{X}=M(X)=\frac{1}{2}(a+b).$$

Модой Mo(X) дискретной случайной величины называют ее наиболее вероятное значение (рис. 3.4, *а*), а модой непрерывной величины — значение X, которому соответствует максимальная плотность вероятности $f(x)$ (рис. 3.4, *б*).

Медианой Me(X) распределения случайной величины называют такое ее значение X, которое делит все распределение на две равновероятные части. Другими словами, для случайной величины X одинаково вероятно принять значения меньше $Me(X)$ или больше $Me(X)$:

$$P(X<Me)=P(X>Me)=\frac{1}{2}.$$

Поэтому медиану распределения непрерывной случайной величины можно вычислить из соотношения

$$\int_a^{Me} f(x)dx=0{,}5. \qquad (3.8)$$

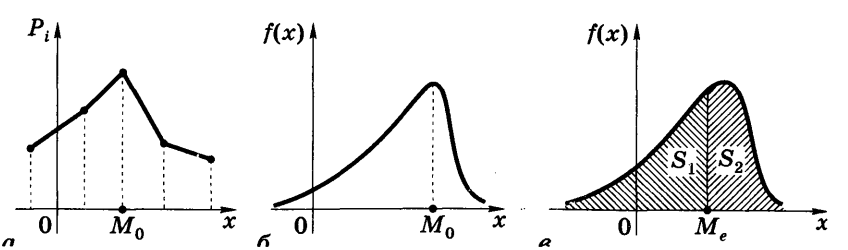


Рис. 3.4. Графическое представление распределений случайной величины X:
а — дискретной с указанием моды; *б* — непрерывной с указанием моды; *в* — непрерывной с указанием медианы

Графически медиана — это значение случайной величины, ордината которой делит пополам площадь под кривой распределения ($S_1 = S_2$) (рис. 3.4, *в*). Этой характеристикой обычно пользуются только для непрерывных случайных величин, хотя формально ее можно определить и для дискретных величин.

Если $M(X)$, $Mo(X)$ и $Me(X)$ совпадают, то распределение случайной величины называют *симметричным*, в противном случае — *асимметричным*.

3.3.2. Характеристики рассеяния

Характеристиками рассеяния являются дисперсия и среднее квадратичное отклонение.

Дисперсия $D(X)$ случайной величины X определяется как математическое ожидание квадрата отклонения случайной величины X от ее математического ожидания $M(X)$:

$$D(X) = M[X - M(X)]^2, \tag{3.9}$$

или

$$D(X) = M(X^2) - [M(X)]^2. \tag{3.9а}$$

Дисперсия *дискретной* случайной величины вычисляется по формуле

$$D(X) = \sum_{i=1}^{n} [x_i - M(X)]^2 p_i,$$

или

$$D(X) = \sum_{i=1}^{n} [x_i^2 p_i] - [M(X)]^2. \tag{3.10}$$

Дисперсия *непрерывной* величины, распределенной в интервале $[a, b]$,

$$D(X) = \int_a^b [x - M(X)]^2 f(x) dx,$$

или

$$D(X) = \int_a^b x^2 f(x) dx - [M(X)]^2. \tag{3.11}$$

Для интервала $(-\infty, \infty)$

$$D(X) = \int_{-\infty}^{\infty} \left[x_i - M(X)\right]^2 f(x)dx$$

или

$$D(X) = \int_{-\infty}^{\infty} x^2 f(x)dx - \left[M(X)\right]^2. \qquad (3.12)$$

Дисперсия характеризует среднюю величину рассеяния, разбросанности значений случайной величины X относительно ее математического ожидания. Само слово «дисперсия» означает «рассеяние».

Однако дисперсия $D(X)$ имеет размерность квадрата случайной величины, что весьма неудобно при оценке разброса в физических, биологических, медицинских и других приложениях. Поэтому обычно пользуются другим параметром, размерность которого совпадает с размерностью X. Это ***среднее квадратичное отклонение*** $\sigma(X)$ случайной величины X:

$$\sigma(X) = \sqrt{D(X)}. \qquad (3.13)$$

Итак, математическое ожидание, мода, медиана, дисперсия и среднее квадратичное отклонение являются наиболее употребляемыми числовыми характеристиками распределений случайных величин, каждая из которых выражает определенное свойство этого распределения.

3.4. Нормальный закон распределения случайной величины

Нормальный закон распределения играет исключительно важную роль в теории вероятностей. Во-первых, это наиболее часто встречающийся на практике закон распределения непрерывных случайных величин. Во-вторых, он является предельным законом, так как к нему при определенных условиях приближаются другие законы распределения.

Нормальный закон распределения характеризуется формулой для плотности вероятности, подобной известной функции Гаусса ($y = e^{-x^2}$), поэтому иногда его называют также ***законом распределения Гаусса***:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{[x-M(X)]^2}{2\sigma^2}}, \qquad (3.14)$$

где x — текущие значения случайной величины X; $M(X)$ и σ — ее математическое ожидание и среднее квадратичное отклонение, которые полностью определяют функцию $f(x)$.

Эта функция подчиняется условию нормировки (3.5).

Таким образом, если случайная величина распределена по нормальному закону, то достаточно знать только два числовых параметра: $M(X)$ и σ, чтобы полностью знать закон ее распределения (3.14).

График функции (3.14) называют ***нормальной кривой распределения (кривой Гаусса)***. Он имеет симметричный вид относительно ординаты $x = M(X)$. Максимальная плотность вероятности, равная $\frac{1}{\sigma\sqrt{2\pi}} \approx \frac{0{,}4}{\sigma}$, соответствует математическому ожиданию $M(X) = X$; по мере удаления от нее плотность вероятности $f(x)$ симметрично спадает, постепенно приближаясь к нулю (рис. 3.5). Изменение значения $M(X)$ в формуле (3.14) не меняет форму нормальной кривой, а приводит лишь к ее сдвигу вдоль оси абсцисс.

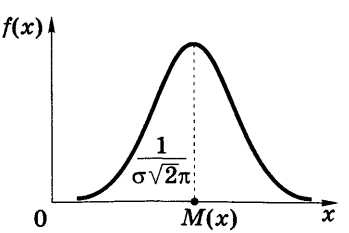


Рис. 3.5. Вид кривой нормального распределения случайной величины X

Величина $M(X)$ называется также центром рассеяния, а среднее квадратичное отклонение σ определяет ширину и высоту кривой распределения (рис. 3.6). С возрастанием σ величина максимума убывает, а сама кривая становится более пологой, растягиваясь вдоль

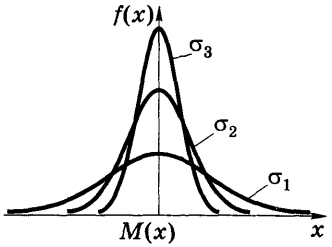


Рис. 3.6. Вид кривой нормального распределения при разных значениях σ ($\sigma_3 < \sigma_2 < \sigma_1$)

оси абсцисс, тогда как при уменьшении σ кривая сужается и вытягивается вверх.

При этом площадь под кривой распределения всегда равна 1, независимо от значений $M(X)$ и σ, подчиняясь условию нормировки (3.5).

Нормальное распределение симметрично, так как

$$M(X) = Mo(X) = Me(X).$$

При нормальном распределении вероятность попадания значений случайной величины X в любой заданный интервал (x_1, x_2) можно вычислить по формуле

$$P(x_1 < X < x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{[x-M(X)]^2}{2\sigma^2}} dx. \qquad (3.15)$$

На практике часто встречается задача нахождения вероятности попадания значений нормально распределенной случайной величины в интервал, симметричный относительно $M(X)$. В частности, рассмотрим важную в прикладном отношении задачу. Отложим вправо и влево от $M(X)$ отрезки, равные σ, 2σ и 3σ (рис. 3.7), и найдем вероятности попадания величины X в соответствующие интервалы:

$$P(M(X) - \sigma < X < M(X) + \sigma) = 0{,}6827 = 68{,}27\ \%; \qquad (3.16)$$

$$P(M(X) - 2\sigma < X < M(X) + 2\sigma) = 0{,}9545 = 95{,}45\ \%; \qquad (3.17)$$

$$P(M(X) - 3\sigma < X < M(X) + 3\sigma) = 0{,}9973 = 99{,}73\ \%. \qquad (3.18)$$

Из формулы (3.18) следует, что значения нормально распределенной случайной величины X с вероятностью $P = 99{,}73$ % лежат в интервале $M(X) \pm 3\sigma$, иначе говоря, в этот интервал попадают практически все возможные значения данной случайной величины. Такой способ оценки диапазона возможных значений случайной величины известен как ***«правило трех сигм»***.

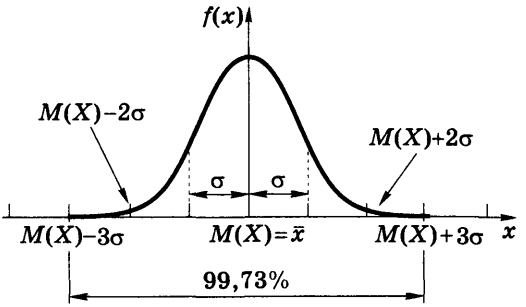


Рис. 3.7. К вероятности попадания значений случайной величины X в интервалы $M(X) \pm \sigma$, $M(X) \pm 2\sigma$, $M(X) \pm 3\sigma$

Пример. Известно, что для человека pH крови является нормально распределенной величиной со средним значением (математическим ожиданием) 7,4 и стандартным отклонением 0,2. Определите диапазон возможных значений этого параметра.

Для ответа на этот вопрос воспользуемся «правилом трех сигм». С вероятностью, равной 99,73 %, можно утверждать, что диапазон значений pH для человека составляет $7{,}4 \pm 3 \cdot 0{,}2$, т.е. 6,8 – 8.

Контрольные вопросы и задания

1. Какие величины называют случайными? Какие случайные величины называют дискретными, а какие — непрерывными? Приведите примеры.
2. Как задают распределение дискретной случайной величины? Что представляет собой условие нормировки?
3. Как задают распределение непрерывной случайной величины? Что представляет собой условие нормировки?
4. Дайте определение и объясните смысл числовых характеристик распределения случайных величин (математическое ожидание, дисперсия, среднее квадратичное отклонение, мода, медиана).
5. Как найти числовые параметры распределения дискретной случайной величины?
6. Как найти числовые параметры распределения непрерывной случайной величины?
7. Каковы свойства нормального распределения непрерывной случайной величины?

Глава 4 Основы математической статистики

Законы теории вероятностей — это математическое выражение реальных закономерностей, которым подчиняются массовые случайные явления. Каждое исследование случайных явлений, выполняемое методами теории вероятностей, прямо или косвенно опирается на экспериментальные данные, на результаты испытаний и наблюдений. Однако детально обследовать все множество интересующих нас объектов по некоторому признаку зачастую невозможно.

Разработка методов исследования массовых случайных явлений на основе анализа экспериментальных данных, полученных при обследовании сравнительно небольшой группы этих явлений, составляет предмет специальной науки — математической статистики.

4.1. Предмет и задачи математической статистики. Генеральная и выборочная совокупность

Фундаментальными понятиями математической статистики являются генеральная совокупность и выборочная совокупность (выборка).

Генеральная совокупность — это множество однородных объектов, подлежащих статистическому исследованию и характеризующихся определенными качественными или количественными признаками. Изучение всех объектов этой совокупности часто невозможно или нецелесообразно по физическим или экономическим причинам.

Например, реально существующая генеральная совокупность — это конкретно выбранная большая популяция людей: все жители Беларуси в данный момент времени или только все мужчины, или женщины, или дети.

Для того чтобы изучить генеральную совокупность по какому-либо из ее количественных признаков X (например, острота зрения, показатели анализа крови и т.д.), следовало

бы изучить все ее объекты по этому признаку и затем обработать полученный массив данных методами теории вероятностей. Однако на практике провести подобное сплошное обследование объектов генеральной совокупности часто физически невозможно и экономически невыгодно. Главная цель данного исследования состоит в установлении закона распределения признака в генеральной совокупности и определении основных количественных характеристик этого распределения. Например, если будет установлено, что изучаемый признак распределен по нормальному закону, то достаточно определить два параметра распределения: его математическое ожидание и дисперсию, чтобы иметь полное представление о распределении этого признака в генеральной совокупности [см. формулу (3.14)]. Для нахождения закона распределения и его основных числовых характеристик достаточно исследовать только часть объектов генеральной совокупности, так называемую выборку.

Совокупность n объектов, отобранных из интересующей нас генеральной совокупности для конкретного статистического исследования, называется ***выборочной совокупностью*** или ***выборкой***.

Исследование выборки дает некоторые приближенные оценочные значения интересующих нас параметров распределения, которые принимают различные значения для разных выборок. Главная цель выборочного метода, основного в математической статистике, — по вычисленным числовым характеристикам выборки как можно точнее определить соответствующие характеристики генеральной совокупности. Это возможно лишь в том случае, если отобранная для работы часть объектов обладает теми же основными свойствами, что и целое. Иначе говоря, выборка должна быть *представительной* (*репрезентативной*), т.е. по возможности полнее «представлять» свойства генеральной совокупности. Это одно из важнейших требований, предъявляемых к выборке, его несоблюдение ведет к грубым ошибкам и обесценивает результаты статистического исследования.

Например, если при изучении заболеваемости населения республики (генеральная совокупность) ишемической бо-

лезнью сердца в качестве выборки взята группа студентов, то результаты будут ошибочны, поскольку свойства такой выборки не соответствуют свойствам генеральной совокупности, как и в случае, если в качестве выборки будут взяты только пациенты кардиологического диспансера. Репрезентативность выборки обеспечивается ее достаточным объемом и определенными правилами ее формирования, которые в данном учебнике рассматриваться не будут.

Из многих задач, решаемых математической статистикой, выделим следующие:

1. Выяснение закона распределения исследуемого признака в выборке.

2. Определение числовых характеристик выборки (методы описательной статистики).

3. Оценка параметров генеральной совокупности по числовым характеристикам выборки: точечные оценки и доверительные интервалы для параметров распределения.

4. Исследование статистической связи между двумя признаками выборочной совокупности (элементы корреляционного анализа).

4.2. Статистическое распределение выборки

Итак, необходимо узнать распределение признака X в генеральной совокупности, но реально исследуется лишь некоторая выборка из нее.

В серии экспериментов, проводимых с выборкой, исследуемая величина X принимает определенные значения. Каждое полученное значение величины X называют ***вариантой***. Эти значения, записанные в том порядке, в котором они были получены в опытах, представляют собой ***простой статистический ряд***. Экспериментальная часть на этом закончена и теперь полученные данные подлежат статистической обработке.

Первый шаг при обработке полученного материала — наведение в нем определенного порядка, позволяющего получить представление о статистическом распределении варианты X в исследуемой выборке. Здесь существует возможность создания либо вариационного, либо интервального ряда.

4.2.1. Вариационный ряд и его графическое представление

Пусть некоторая выборка исследуется по количественному признаку X, который представляет собой дискретную случайную величину. Сначала все полученные числовые данные следует расположить в порядке их возрастания (или убывания), т.е. из простого статистического ряда получить так называемый ранжированный статистический ряд данных. В этом простом ряду варианта x_1 будет встречаться (повторяться) m_1 раз, x_2 — m_2 раза, ..., x_k — m_k раз, при этом $\sum_{i=1}^{k} m_i = n$, т.е. сумма частот равна объему выборки. Далее по этим данным строится статистическое распределение (в медицинской литературе — *вариационный ряд*), которое удобно представить в виде табл. 4.1, включающей в себя:

1) x_i — различные значения варианты, расположенные в ранжированной последовательности (обычно в порядке возрастания);

2) m_i — частота встречаемости варианты, т.е. количество наблюдений (повторений) варианты x_i в простом статистическом ряду;

3) $p_i^* = m_i / n$ — относительные частоты вариант, т.е. отношения частот m_i к объему выборки n; они являются выборочными (эмпирическими) оценками вероятностей появления значений x_i.

Каждая относительная частота указывает долю общего объема выборки, приходящуюся на данное значение варианты x_i.

Таблица 4.1

Варианта, x_i ($x_1 < x_2 < x_3 < ... < x_k$)	x_1	x_2	x_3	...	x_k	Контроль
Частота, m_i	m_1	m_2	m_3	...	m_k	$\sum_{i=1}^{k} m_i = n$
Относительная частота, $p_i^* = \frac{m_i}{n}$	$\frac{m_1}{n}$	$\frac{m_2}{n}$	$\frac{m_3}{n}$	...	$\frac{m_k}{n}$	$\sum_{i=1}^{k} \frac{m_i}{n} = 1$

Под распределением дискретной случайной величины в теории вероятностей понимается соответствие между возможными значениями случайной величины и их вероятностями; в математической статистике — соответствие между наблюдаемыми вариантами x_i и их частотами или относительными частотами.

Для получения наглядного представления о распределении случайной величины в выборке строят соответствующие графики.

Графическим представлением вариационного ряда является полигон частот или полигон относительных частот.

Полигоном частот называют ломаную линию, отрезки которой последовательно соединяют точки (x_1, m_1), (x_2, m_2), ..., (x_k, m_k).

Для построения полигона частот на оси абсцисс откладывают значения вариант x_i, а на оси ординат — соответствующие им частоты m_i. Точки (x_i, m_i) соединяют прямыми отрезками.

Полигоном относительных частот называют ломаную линию, отрезки которой соединяют точки $\left(x_1, \frac{m_1}{n}\right)$, $\left(x_2, \frac{m_2}{n}\right)$; $\left(x_k, \frac{m_k}{n}\right)$.

Пример. Анализируемый показатель X — срок лечения больного при некотором заболевании. Вариационный ряд — распределение больных по срокам лечения (объем выборки $n = 26$ больных) — имеет вид, представленный в табл. 4.2.

Таблица 4.2

Число дней лечения, x_i	17	18	20	22	23	25	Контроль
Число больных с данным сроком лечения (частота), m_i	2	5	4	8	5	2	$\sum_{i=1}^{6} m_i = 26$
Относительная частота, $p_i^* = \frac{m_i}{n}$	0,08	0,19	0,15	0,31	0,19	0,08	$\sum_{i=1}^{6} \frac{m_i}{n} = 1$

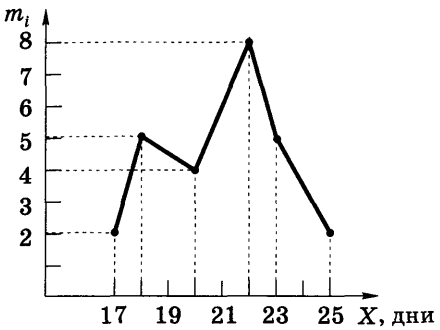


Рис. 4.1. Полигон частот

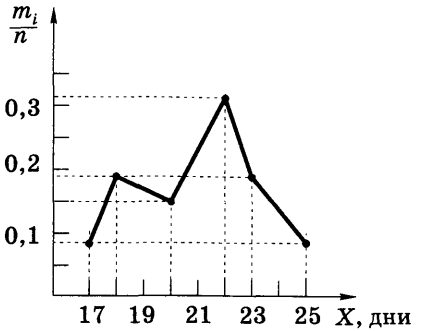


Рис. 4.2. Полигон относительных частот

Соответствующие этому вариационному ряду полигоны частот и относительных частот представлены на рис. 4.1, 4.2.

Подобное графическое представление данных позволяет наглядно оценить распределение случайного признака X и определить его более и менее вероятные значения.

4.2.2. Интервальный ряд и его графическое представление

Интервальный ряд удобно использовать, когда количественный признак X, характеризующий выборку, непрерывен, т.е. может принимать любые значения в некотором интервале. В этом случае статистическое распределение выборки (интервальный ряд) строится следующим образом:

1) область изменения признака ($x_{max} - x_{min}$) разбивают на несколько интервалов обычно равной ширины;

2) число интервалов k (как правило, не менее 5 и не более 25) приближенно определяется следующими эмпирическими формулами:

$$k = \sqrt{n}, \text{ или } k \approx 1 + 3{,}32 \lg n, \tag{4.1}$$

где n — объем выборки;

3) ширина интервалов одинакова:

$$\Delta x = h = \frac{x_{max} - x_{min}}{k}; \tag{4.2}$$

4) вычисляют границы интервалов:

$$x_{\min} = x_0,\ x_1 = x_0 + h,\ x_2 = x_1 + h,\ \ldots,\ x_{\max} = x_k.$$

Поскольку некоторые варианты могут являться границей двух соседних интервалов, то обычно придерживаются следующего правила: к интервалу (a, b) относят варианты, удовлетворяющие неравенству $a \le x < b$, т.е. пограничную варианту относят в правый интервал;

5) для каждого интервала подсчитывают частоты m_i и (или) относительные частоты m_i/n попадания вариант в данный интервал. Нередко используют также плотность относительной частоты:

$$\frac{m_i}{n\Delta x} = \frac{m_i}{nh}, \qquad (4.3)$$

которую можно считать выборочной (эмпирической) оценкой плотности вероятности.

Рассмотренное распределение непрерывной случайной величины X — интервальный ряд — обычно представляется в виде табл. 4.3.

Таблица 4.3

Интервал	Частота, m_i	Относительная частота, $p_i^* = m_i/n$
$x_0 - x_1$	m_1	m_1/n
$x_1 - x_2$	m_2	m_2/n
$x_2 - x_3$	m_3	m_3/n
...	...	...
$x_{k-1} - x_k$	m_k	m_k/n

Графическим представлением интервального ряда является гистограмма частот или относительных частот.

Гистограммой частот называют диаграмму, состоящую из вертикальных прямоугольников, основаниями которых являются интервалы длиной $\Delta x = h$, а высоты равны отношению $\frac{m_i}{\Delta x}$ (плотности частоты). Для построения гистограммы

частот на оси абсцисс откладывают интервалы значений исследуемого показателя (интервалы вариант) и на них строят прямоугольники высотой $\frac{m_i}{\Delta x}$. Площадь i-го прямоугольника равна $\Delta x \frac{m_i}{\Delta x} = m_i$, т.е. количеству вариант (частоте m_i попадания) в i-м интервале. Следовательно, площадь всей такой гистограммы равна сумме частот всех интервалов, иначе говоря, объему выборки.

Гистограмма относительных частот отличается от предыдущей гистограммы тем, что на ней высоты прямоугольников равны отношению $\frac{m_i}{n\Delta x}$, т.е. плотности относительной частоты (эмпирической плотности вероятности). В этом случае площадь i-го прямоугольника равна $\Delta x \frac{m_i}{n\Delta x} = p_i^*$ — относительной частоте вариант, попавших в i-й интервал. Напомним, что $p_i^* = m_i/n$ является примерной оценкой вероятности попадания значений X в выбранный интервал. Площадь гистограммы относительных частот равна сумме относительных частот всех интервалов, т.е. единице.

Пример. Анализируемый показатель X — масса тела новорожденного. Определение массы тела 100 новорожденных показало, что минимальная масса составляет 2,7 кг, максимальная — 4,4 кг. Интервал (2,7 – 4,4) кг разбиваем на 10 равных интервалов ($k = \sqrt{100} = 10$) шириной $h = \frac{4,4 - 2,7}{10} = 0,17$ кг и строим интервальный ряд (табл. 4.4).

Таблица 4.4

Интервал	Масса тела, кг	Частота, m_i	$m_i/n = p_i$	m_i/nh
1-й	2,7–2,87	4	0,04	0,235
2-й	2,87–3,04	8	0,08	0,470
3-й	3,04–3,21	12	0,12	0,700
4-й	3,21–3,38	16	0,16	0,940

Продолжение табл. 4.4

Интервал	Масса тела, кг	Частота, m_i	$m_i/n = p_i$	m_i/nh
5-й	3,38–3,55	21	0,21	1,235
6-й	3,55–3,72	15	0,15	0,880
7-й	3,72–3,89	11	0,11	0,650
8-й	3,89–4,06	7	0,07	0,410
9-й	4,06–4,23	4	0,04	0,235
10-й	4,23–4,40	2	0,02	0,118

Контроль: $k = 10$,

$$\sum_{i=1}^{10} m_i = 4 + 8 + 12 + 16 + 21 + 15 + 11 + 7 + 4 + 2 = 100 = n \text{ (объем выборки)},$$

$$\sum_{i=1}^{10} \frac{m_i}{n} = 0{,}04 + 0{,}08 + 0{,}12 + 0{,}16 + 0{,}21 + 0{,}15 + 0{,}11 + 0{,}07 + 0{,}04 + 0{,}02 = 1.$$

Построенная по данным табл. 4.4 гистограмма относительных частот приведена на рис. 4.3.

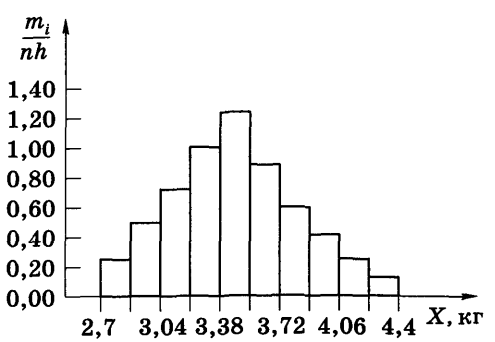


Рис. 4.3. Гистограмма относительных частот массы тела новорожденных

Для представленной выборки интервал наиболее вероятных масс тела новорожденных составляет 3,38–3,55 кг.

Следует отметить, что гистограммой называют и серию прямоугольников, высотами которых являются непосред-

ственно частоты m_i соответствующих интервалов или относительные частоты (в нормированной гистограмме), а также относительные частоты в процентах (процентная гистограмма). Два последних варианта позволяют сравнивать гистограммы, построенные на одних и тех же интервалах, но для различных выборок из одной генеральной совокупности.

Важно, что гистограммы можно использовать для оценки закона распределения признака в генеральной совокупности (в популяции). Соединяя средние точки верхних оснований прямоугольников гистограммы относительных частот плавной линией, можно по данным выборки получить приближенное представление о кривой распределения, т.е. о законе распределения плотности вероятности $f(X)$. Так, по виду гистограммы на рис. 4.3 можно предположить, что анализируемый показатель (масса тела новорожденного) в генеральной совокупности распределен по нормальному закону, т.е. нормальный закон является вероятностной моделью для данного признака в популяции.

Таким образом, обобщив изложенный материал, можно сделать следующие выводы.

1. Если выборка исследуется по количественному признаку X, представляющему дискретную случайную величину, то статистическим распределением выборки является вариационный статистический ряд, представляющий полученные значения признака, записанные в упорядоченном виде с указанием их частот и относительных частот. Графическим представлением вариационного ряда является полигон частот или относительных частот.

2. Если выборка исследуется по количественному признаку X, представляющему непрерывную случайную величину, то статистическим распределением выборки является интервальный статистический ряд. Он включает в себя интервалы значений вариант, частоты попадания вариант в эти интервалы, относительные частоты, а при необходимости — плотности относительных частот для этих интервалов. Для графического представления интервального распределения случайной величины обычно строят *гистограммы частот* или *относительных частот*.

4.3. Методы описательной статистики

Методы описательной статистики — описание выборок, исследуемых по количественному признаку X, с помощью различных числовых характеристик.

Преимущество этих методов заключается в том, что несколько простых и достаточно информативных статистических показателей, если они известны, во-первых, избавляют нас от просмотра сотен, а порой и тысяч значений вариант, во-вторых, позволяют получить более или менее точную оценку характеристик распределения признака в генеральной совокупности.

Показатели, описывающие выборку, разбиваются на несколько групп, которые в большинстве своем аналогичны числовым характеристикам случайных величин в теории вероятностей.

Показатели положения описывают положение вариант выборки на числовой оси. К ним относятся: минимальная и максимальная варианты; выборочное среднее; выборочная мода; выборочная медиана.

Выборочным средним называется среднее арифметическое значение выборки:

$$\overline{x_в} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad (4.4)$$

где x_i — i-я варианта выборки; n — объем выборки.

Выборочная мода $Mo_в$ — варианта, которая чаще всего встречается в исследуемой выборке, т.е. имеет наибольшую частоту.

Выборочная медиана $Me_в$ — это значение варианты, которая делит ранжированный статистический ряд на две равные части по числу попадающих в них вариант, т.е. вероятность оказаться на числовой оси слева и справа от медианы одинакова: $P(X < Me) = P(X > Me) = 0{,}5$.

Если объем выборки n, то выборочная медиана в этом ранжированном ряду занимает среднее место с номером $\frac{n+1}{2}$.

Пример 1. На рис. 4.4 приведено распределение по возрасту заболевших дифтерией (на 10 тыс. населения соответствующего возраста), которое явно не соответствует нормальному. Очевидно, что знание среднего возраста заболевших ($\overline{x_в} \approx 7{,}8$ года) в этом случае менее важно, чем знание возраста, в котором чаще всего возникает заболевание и который представляет собой моду ($Mo_в \approx 4$ года). Именно этот показатель указывает, что главные профилактические меры должны быть сосредоточены в дошкольных учреждениях.

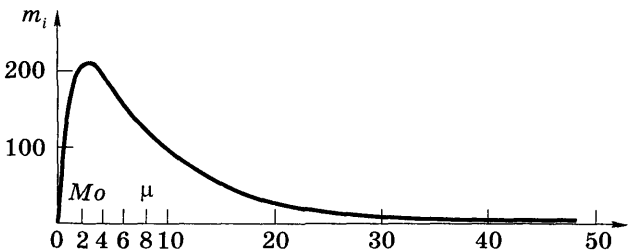


Рис. 4.4. Зависимость частоты заболевания от возраста

Пример 2. Дан ранжированный статистический ряд: 1; 2; 2; 3; 3; 5; 5; 6; 6; 7; 8; 9; 9; $n = 13$. Медиана, т.е. варианта, разделяющая этот ряд на две равные по количеству вариант части, занимает в таком ряду среднее место (здесь седьмое: $\frac{n+1}{2} = 7$) и равна $Me = 5$. Слева и справа от нее количество вариант одинаково (по шесть).

Показатели разброса характеризуют степень рассеяния вариант вокруг их среднего выборочного значения $\overline{x}_в$. Здесь обычно используются:

- выборочная дисперсия $D_в = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$; (4.5)
- среднеквадратичное отклонение $\sigma_в = \sqrt{D_в}$

и стандартное отклонение $S_n = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x_в})^2}{n-1}}$; (4.6)

- размах выборки — разность между максимальной и минимальной вариантами: $x_{max} - x_{min}$;

• коэффициент вариации $v = \frac{\sigma_в}{x_в} \cdot 100\,\%$, (4.7)

используемый для сравнения величин рассеяния двух вариационных рядов (большему значению коэффициента вариации соответствует большее рассеяние).

4.4. Оценка параметров генеральной совокупности по параметрам выборки

Напомним, что главная цель любого статистического исследования — установить закон распределения изучаемого признака в генеральной совокупности и определить основные характеристики этого распределения путем анализа выборки. Иначе говоря, надо определить генеральное среднее $\bar{x}_г = M(X)$, генеральные дисперсию $D_г(X)$, среднее квадратичное отклонение $\sigma_г$, моду $Mo_г$, медиану $Me_г$ и другие характеристики генеральной совокупности, зная соответствующие параметры выборки.

1. *Точечная оценка характеристик генеральной совокупности* заключается в использовании в качестве числовых характеристик генеральной совокупности соответствующих числовых характеристик выборки. Например, в качестве генерального среднего используется выборочное среднее, генеральной дисперсии — выборочная дисперсия и т.д. Такие оценки и называются точечными. Это наиболее простой, но не очень достоверный способ оценки характеристик генеральной совокупности. Его недостаток состоит в том, что не ясно, насколько они отличаются от истинных значений параметров генеральной совокупности. При малых выборках ошибка может быть особенно значительной.

2. *Интервальная оценка параметров генеральной совокупности* является более достоверной. В этом случае определяется интервал, в который с заранее заданной вероятностью попадает истинное значение исследуемого признака. Этот интервал называется ***доверительным***, а вероятность того, что истинное значение оцениваемой величины находится внутри этого интервала, — ***доверительной вероятностью*** γ или

надежностью. В медицинской литературе для этой величины используется также термин «вероятность безошибочного прогноза». Величину $\alpha = 1 - \gamma$, характеризующую вероятность *непопадания* истинного значения исследуемого признака в выбранный доверительный интервал, т.е. вероятность ошибки, называют ***уровнем значимости.*** Значения γ задаются заранее (обычно в медико-биологических исследованиях выбирают значения $\gamma = 0{,}95 = 95\ \%$ или $\gamma = 0{,}99 = 99\ \%$), после чего вычисляют соответствующий доверительный интервал.

Для построения надежных интервальных оценок необходимо знать закон, по которому оцениваемый случайный признак распределен в генеральной совокупности.

Рассмотрим, как находится интервальная оценка генерального среднего $M_г(X)$ в случае малых выборок ($n < 30$), если признак X распределен в генеральной совокупности по нормальному закону. Будем считать, что нам известны выборочная средняя $\bar{x}_в$ и выборочное стандартное отклонение $S_в$. В этом случае интервальной оценкой генерального среднего (математического ожидания) $M_г(X)$ этого признака является доверительный интервал

$$\bar{x}_в - \delta < M_г(X) < \bar{x}_в + \delta, \qquad (4.8)$$

или

$$M_г(X) = \bar{x}_в \pm \delta. \qquad (4.9)$$

Полуширина доверительного интервала δ (точность оценки) рассчитывается по формуле

$$\delta = t_{\gamma,n} \frac{S_в}{\sqrt{n}}, \qquad (4.10)$$

где n — объем выборки; $S_в$ — выборочное стандартное отклонение; $S_в / \sqrt{n}$ — стандартная ошибка выборочного среднего[1]; $t_{\gamma, n}$ — коэффициент Стьюдента, значения которого зависят от заданной доверительной вероятности γ и объема выборки n и определяются по соответствующим таблицам либо содержатся в программных статистических пакетах обработки данных.

[1] В медицинской и биологической литературе эта величина иногда обозначается буквой m и называется *ошибкой репрезентативности*.

Анализ формулы (4.10) показывает, что:

• чем больше доверительная вероятность γ, тем больше коэффициент $t_{\gamma, n}$ и шире доверительный интервал;

• чем больше объем выборки n, тем меньше доверительный интервал и выше точность результатов.

При больших объемах выборки ($n > 30$) коэффициент Стьюдента практически не зависит от n и определяется только значением доверительной вероятности γ, поэтому полуширина доверительного интервала δ при больших выборках определяется соотношениями:

$$\delta = 1{,}96\, S/\sqrt{n} \quad \text{при } \gamma = 95\ \%, \tag{4.11}$$

или

$$\delta = 2{,}58\, S/\sqrt{n} \quad \text{при } \gamma = 99\ \%. \tag{4.12}$$

Существует доверительный интервал и для среднего квадратичного отклонения $\sigma_г$ генеральной совокупности.

Подобные интервальные оценки с заданной надежностью даются и в тех случаях, когда распределение случайного признака в генеральной совокупности отличается от нормального.

Пример. Исследуется состояние дыхательных путей курящих. В качестве характеристики используется показатель функции внешнего дыхания — максимальная объемная скорость середины выдоха в л/с. Предполагая, что в генеральной совокупности данный параметр распределен по нормальному закону, найдите 95%-ный и 99%-ный доверительные интервалы для $\bar{x}_г$ (т.е. $M_г(X)$), характеризующие этих людей. Обследуемая группа — 20 курящих, $\bar{x}_в = 2{,}20$ л/с, $S = 0{,}73$ л/с.

1. Для значений $\gamma = 95\ \%$ и $n = 20$ коэффициент Стьюдента (находим по таблицам критерия Стьюдента) равен $t_{0{,}95;\ 20} = 2{,}09$. Вычисляем полуширину доверительного интервала δ:

$$\delta = t_{\gamma, n}\,(S/\sqrt{n}\,) = 2{,}09 \cdot \frac{0{,}73}{\sqrt{20}} = 0{,}34.$$

Теперь можем сказать, что с доверительной вероятностью $\gamma = 95\ \%$ математическое ожидание $M_г(X)$ генеральной совокупности находится в интервале

$$(2{,}20 - 0{,}34)\ \text{л/с} < M_г(X) < (2{,}20 + 0{,}34)\ \text{л/с},$$

т.е.

$$1{,}86\ \text{л/с} < M_г(X) < 2{,}54\ \text{л/с}.$$

В более компактной форме записи: $M_г(X) = (2{,}20 \pm 0{,}34)$ л/с.

2. Для $\gamma = 99\ \%$ и $n = 20$ $t_{0{,}99;\,20} = 2{,}86$; тогда доверительный интервал для $M_г(X)$ определяется неравенством

$$(2{,}20 - 0{,}47) \text{ л/с} < M_г(X) < (2{,}20 + 0{,}47) \text{ л/с}$$

или

$$1{,}73 \text{ л/с} < M_г(X) < 2{,}67 \text{ л/с},$$

иначе $M_г(X) = (2{,}20 \pm 0{,}47)$ л/с.

Полученные данные подтверждают сделанный ранее вывод: увеличение доверительной вероятности γ расширяет границы доверительного интервала.

Таким образом, формула (4.10) позволяет по заданной доверительной вероятности и объему выборки дать точность оценки для математического ожидания $M_г(X)$ варианты X в генеральной совокупности.

4.5. Определение объема выборки при заданной точности интервальной оценки

Важное практическое значение имеет и обратная задача: определение необходимого объема n выборки по заданной полуширине δ (точности оценки) доверительного интервала для генерального среднего $\overline{x_г} = M_г(X)$. Эта задача решается наиболее просто в случае больших выборок ($n > 30$). Например, при доверительной вероятности $\gamma = 95\ \%$ полуширина доверительного интервала $\delta = 1{,}96\ S / \sqrt{n}$ [см. формулу (4.11)], поэтому необходимый объем выборки

$$n \geq (1{,}96S/\delta)^2 . \qquad (4.13)$$

Пример. Исследователь хочет установить средний уровень гемоглобина у определенной группы населения. Учитывая предварительные данные, он полагает, что этот уровень составляет примерно 150 г/л со стандартным отклонением 32 г/л. Определите, сколько человек он должен обследовать (с какой выборкой он должен работать), чтобы получить точность оценки (полуширину доверительного интервала) $\delta = 5$ г/л при доверительной вероятности $0{,}95 = 95\ \%$.

В случае больших выборок и $\gamma = 95\ \%$ пользуемся соотношением (4.13), в результате находим $n \geq (1{,}96 \cdot 32/5)^2 = 157{,}4$. Таким образом, необходимо обследовать не менее 158 человек.

4.6. Понятие нормы для медицинских показателей

Значения медико-биологических показателей, соответствующие здоровому человеку, часто называют медицинской нормой. Они являются своеобразным ориентиром при оценке состояния здоровья пациента.

Обычно используют два типа норм: *точечную норму* и *нормальный диапазон*, причем при их установлении работают с выборками достаточно большого объема. Точечную норму определяют по значению центра распределения. Нормальные диапазоны в большинстве случаев устанавливаются так, чтобы в них гарантированно попадало 95 % показателей случайно отобранных здоровых людей.

Если соответствующий показатель — случайная величина — распределен по нормальному закону, точечной нормой для него считается $\overline{x}_г = \overline{x}_в$, а нормальный диапазон определяется по формуле (4.11): $\overline{x}_г = \overline{x}_в \pm 1{,}96S$; иногда используют менее точное приближение, заменяя 1,96 на 2: $\overline{x}_г = \overline{x}_в \pm 2S$.

Очень часто нормальные значения некоторого показателя неодинаковы у лиц, живущих в разных географических регионах, у мужчин и женщин, в разных возрастных группах. Поэтому при установлении нормального значения показателя необходимо указывать популяционные группы, к которым он относится.

4.7. Элементы корреляционного анализа

Одной из главных задач корреляционного анализа является выявление связи между случайными величинами X и Y (например, вес и рост человека, частота пульса и артериальное давление и т.п.). Такая зависимость, если она существует, называется ***корреляционной***. Она может быть установлена качественно — по форме корреляционного поля, и количественно — путем вычисления коэффициента корреляции.

Для установления корреляционной зависимости экспериментально определяют соответствующие пары значений при-

знаков X и Y (например, роста и массы тела) для каждого обследованного объекта. Результаты удобно занести в табл. 4.5.

Таблица 4.5

Значения признака X	x_1	x_2	x_3	...	x_n
Значения признака Y	y_1	y_2	y_3	...	y_n

Каждой паре значений (x_i, y_i) на плоскости XOY соответствует одна точка. Количество точек равно объему выборки n. Область на графике $y(x)$, занятая этими точками, образует ***корреляционное поле.***

По форме корреляционного поля можно сделать качественные выводы о характере связи между признаками X и Y. Если форма корреляционного поля близка к кругу (рис. 4.5, *а*), то связи между признаками X и Y нет. Если же корреляционное поле вытянуто (рис. 4.5, *б*, *в*), то корреляционная связь между признаками X и Y существует, причем чем больше вытянуто корреляционное поле, тем сильнее корреляционная связь.

Каждому значению x_i признака X может отвечать несколько разных значений признака Y, которым соответствует некоторое среднее значение $\overline{y_i}$. Зависимость среднего значения $\overline{y_i}$

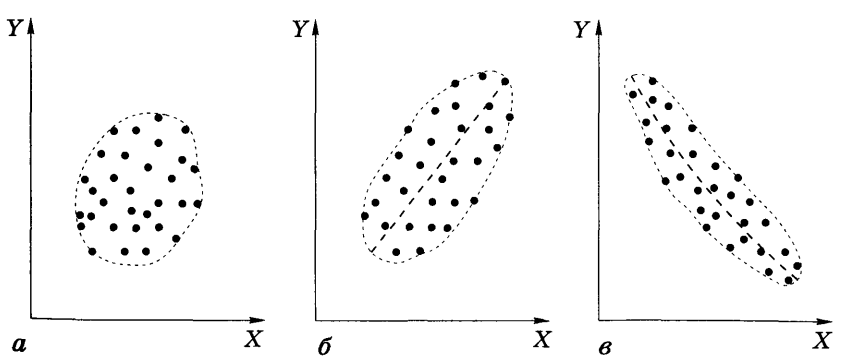


Рис. 4.5. Соответствие формы корреляционного поля степени корреляционной связи между величинами X и Y:
а — связь отсутствует; *б* — связь прямая, средняя;
в — связь обратная, более тесная, чем в случае *б*.
Пунктиром показаны линии регрессии $\overline{Y} = f(X)$

от x_i называется ***эмпирическим уравнением регрессии*** Y **на** X: $\overline{y} = f(x)$. График этой функции называют ***линией регрессии***. Ее обычно строят графически, определяя для каждого значения x_i среднее значение $\overline{y_i}$ соответствующих ему значений признака Y. Через полученные точки $(x_i, \overline{y_i})$ проводят линию регрессии. Если эта линия — прямая (рис. 4.5, *б*), то корреляционная связь между признаками X и Y называется *линейной*, а если нет (рис. 4.5, *в*) — *нелинейной*. Если с увеличением X среднее значение $\overline{y_i}$ возрастает, то корреляционная связь называется *прямой*, а если уменьшается — *обратной*.

При линейной корреляционной зависимости между X и Y степень (силу) связи можно оценить количественно с помощью *коэффициента корреляции* r, который вычисляют по формуле

$$r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \qquad (4.14)$$

где $\sigma_{xy} = \sum_{i=1}^{n} \Delta x_i \Delta y_i$; $\sigma_x = \sqrt{\sum_{i=1}^{n} (\Delta x_i)^2}$; $\sigma_y = \sqrt{\sum_{i=1}^{n} (\Delta y_i)^2}$.

Значения коэффициента корреляции по модулю не превышают 1, но могут быть как положительными, так и отрицательными: $-1 < r < 1$, или $|r| < 1$. При $|r| < 0{,}3$ считают, что связь между X и Y отсутствует или очень слабая; при значениях $0{,}3 < |r| < 0{,}7$ связь умеренная; при $0{,}7 < |r| < 1$ — сильная (тесная). Если $|r| = 1$ — связь функциональная, в этом случае существует функция $Y = f(X)$, жестко связывающая значения Y и X. Если $r > 0$, то связь между признаками X и Y *прямая*, т.е. с увеличением значений одного признака значения другого тоже увеличиваются; а при $r < 0$ связь *обратная* — с увеличением значений одного признака значения другого уменьшаются.

Среднюю ошибку коэффициента корреляции можно найти по формуле

$$m_r = \frac{1 - r^2}{\sqrt{n}}. \qquad (4.15)$$

Видно, что величина этой ошибки тем меньше, чем больше объем выборки n. Значение коэффициента корреляции r

считается *достоверным*, если оно не менее чем в три раза превышает его ошибку, т.е. если $|r| > 3m_r$.

Если выборка имеет достаточно большой объем и репрезентативна, то заключение о тесноте связи между признаками, полученное по данным выборки, можно распространить и на генеральную совокупность.

Контрольные вопросы и задания

1. В чем состоит основная задача статистического исследования? Что такое генеральная совокупность и выборка? В чем заключается требование репрезентативности выборки?
2. Что такое варианта, простой статистический ряд, ранжированный ряд?
3. Каковы этапы создания вариационного ряда (дискретного статистического распределения)? Как он представляется графически?
4. Каковы этапы создания интервального ряда (непрерывного статистического распределения)? Как он представляется графически?
5. Укажите характеристики положения и разброса случайной величины.
6. В чем суть точечной и интервальной оценок параметров генеральной совокупности по данным выборки?
7. Каков порядок нахождения доверительного интервала для параметров генеральной совокупности по данным выборки при заданной доверительной вероятности с использованием коэффициентов Стьюдента?
8. Как определяется необходимый объем выборки при заданной точности интервальной оценки?
9. Как установить значение нормы для медицинских показателей?
10. Какова основная задача корреляционного анализа?
11. Что такое корреляционное поле и какую информацию о корреляционной связи между признаками оно дает?
12. Как построить линию регрессии и определить уравнение регрессии?
13. При каких условиях и как можно рассчитать коэффициент корреляции между признаками?
14. Какую информацию о характере связи дает знание коффициента корреляции?
15. При каком условии значение коэффициента корреляции можно считать достоверным?

Глава 5 Элементы биомеханики

Биологическая механика изучает закономерности движения и деформирования биологических структур под действием факторов окружающей среды и мышечной системы.

Проблемы, решаемые биомеханикой, важны одновременно и для технических, и для медико-биологических наук. Это связано с тем, что в процессе эволюции возникли биологические системы, оптимальные по конструкции в отношении не только выполняемых ими физиологических функций, но и свойств материалов, определяющих их механическое поведение. Поэтому изучение строения и механических свойств различных биологических тканей позволит создать материалы (обычно это композиты[1]), наиболее пригодные для замещения пораженных естественных структур.

5.1. Кинематика и динамика движений человека

Человек и любые другие биологические объекты, являясь частью материального мира, естественно, подчиняются законам механики и физики. Вместе с тем живые объекты отличаются от технических систем и обладают рядом особенностей, которые следует иметь в виду, изучая их механическое поведение. Важнейшей особенностью является способность живых систем адаптироваться к длительным или многократным воздействиям внешних факторов путем функциональной и морфологической перестройки отдельных структур.

С точки зрения биомеханики аппарат движения человека представляет собой управляемую систему подвижно соединенных костей, обладающих определенными размерами, массами, моментами инерции и снабженных мышечными двигателями. Скелет состоит из 206 костей (85 парных и 36 непарных), которые соединены суставами и связками.

[1] Композиты (композиционные материалы) — материалы, неоднородные по составу, в которых именно сочетание разных компонент обеспечивает нужные физико-механические свойства.

Это *пассивная часть* опорно-двигательного аппарата. Поперечнополосатые скелетные мышцы (их более 600) — это его *активная часть*, приводящая в движение костные звенья. Управляет этим костно-мышечным аппаратом движения центральная нервная система.

Само движение осуществляется системой костных рычагов, которые приводятся в движение силой тяги, возникающей при сокращении мышц.

Рычагом называется твердое тело, имеющее ось вращения (точку опоры), к которому приложены силы, создающие моменты сил относительно этой оси. Напомним, что *момент силы* (относительно некоторой оси или точки опоры) равен произведению величины силы и плеча ее действия, а *плечо силы* — это кратчайшее расстояние от оси вращения до линии действия силы, лежащей в плоскости вращения.

В качестве примера рычага рассмотрим предплечье (рис. 5.1, *а*). Здесь сила тяги мышцы F приложена к короткому плечу a рычага (точка опоры O — локтевой сустав), а к его длинному плечу b приложена нагрузка R — сила тяжести удерживаемого груза, приложенная обычно к кисти, а также сила тяжести самого предплечья, приложенная в его центре масс (на рисунке не показана и далее не учитывается). Моменты этих сил равны соответственно Fa и Rb.

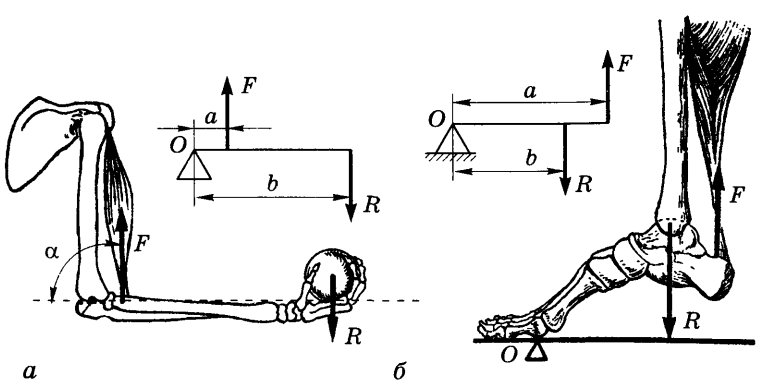


Рис. 5.1. Основные виды рычага:
а — рычаг скорости; *б* — рычаг силы

Для равновесия рычага необходимо равенство моментов противодействующих сил, приложенных к нему:

$$Fa = Rb. \tag{5.1}$$

При движении рычага угловая скорость всех его точек $\omega = \frac{v}{r}$ одинакова (v — линейная скорость точки, r — расстояние от точки до оси вращения), поэтому для точек приложения груза и мышцы к рычагу можем записать

$$\frac{v_{гр}}{b} = \frac{v_{м}}{a}, \tag{5.2}$$

где $v_{гр}$ — скорость движения груза; $v_{м}$ — скорость укорочения мышцы.

Из этих соотношений следует, что $F = R\frac{b}{a}$, $v_{гр} = v_{м}\frac{b}{a}$.

При $a < b$ рычаг проигрывает в силе ($F > R$), но выигрывает в скорости перемещения груза $\left(v_{гр} > v_{м}\right)$ и называется поэтому ***рычагом скорости*** (например, предплечье, рис. 5.1, *а*). При $a > b$ рычаг дает выигрыш в силе ($F < R$) и называется ***рычагом силы***, но он проигрывает в скорости перемещения груза $\left(v_{гр} < v_{м}\right)$. Примером такого рычага служит свод стопы при подъеме на полупальцы (рис. 5.1, *б*). В этом случае опорой O рычага служат головки плюсневых костей; преодолеваемая сила тяжести R тела приложена к таранной кости, а мышечная сила F, осуществляющая подъем тела, передается через ахиллово сухожилие и приложена к выступу пяточной кости и поскольку здесь $a > b$, то $F < R$.

Часто мышечная сила F направлена к рычагу под углом α, отличным от 90° (рис. 5.2). В этом случае условие равновесия рычага принимает вид: $Fa\sin\alpha = Rb$, откуда $F = \frac{Rb}{a\sin\alpha}$, т.е. мышечное усилие F, необходимое для преодоления данной силы сопротивления R, должно быть тем больше, чем под меньшим углом к оси рычага оно направлено.

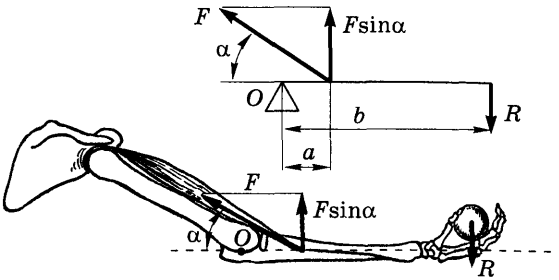


Рис. 5.2. Вариант рычага скорости при $\alpha < 90°$

Одиночное мышечное волокно при сокращении развивает силу порядка 2 мН. В мышцах человека содержится примерно $3 \cdot 10^7$ волокон. Если бы все они сократились одновременно в одном направлении, то создали бы огромное усилие — до **60 000 Н.**

Сила мышцы зависит от количества содержащихся в ней волокон. Поскольку их трудно подсчитать, обычно измеряют поперечное сечение мышцы, которое пропорционально числу волокон, и по нему судят о ее силовых возможностях. При этом нужно определять не геометрическое, а *физиологическое поперечное сечение*, под которым понимают сумму поперечных сечений всех волокон, образующих данную мышцу.

Отношение максимального мышечного усилия, которое может развить мышца, к ее физиологическому поперечному сечению называется ***абсолютной мышечной силой*** и выражается в единицах механического напряжения ($Н \cdot м^{-2} = Па$). Для большинства скелетных мышц человека абсолютная мышечная сила имеет порядок 10^6 Па (100 Н/см^2), а для гладких — 10^5 Па (10 Н/см^2). Сила одних и тех же мышц человека зависит от ряда физиологических условий: возраста, пола, тренированности и т.д.

Перемещение тела человека в пространстве совершается благодаря мышечным сокращениям, при которых мышцы укорачиваются и напрягаются. В *изотоническом* режиме, т.е. при постоянной нагрузке на мышцу, зависимость между скоростью v мышечного сокращения и приложенной силой F описывается ***уравнением Хилла***:

$$v = b\frac{F_{max} - F}{F + a}, \qquad (5.3)$$

где F_{max} — максимальное усилие, развиваемое данной мышцей; a — константа, имеющая размерность силы; b — константа, имеющая размерность скорости.

Из уравнения Хилла (5.3) следует, что при отсутствии нагрузки ($F = 0$) скорость сокращения мышцы максимальная, с увеличением нагрузки она уменьшается и при максимальной нагрузке ($F = F_{max}$) ее укорочение вовсе не происходит ($\upsilon = 0$), что демонстрирует и рис. 5.3.

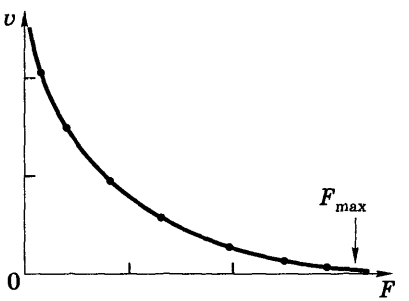


Рис. 5.3. Зависимость скорости υ укорочения мышцы от приложенной к ней нагрузки F

Коэффициент полезного действия (КПД $= A/E$) при сокращении мышцы определяется как отношение совершенной работы A к затраченной энергии E и может достигать 40 – 60 % у разных типов мышц.

Мышечные усилия и внешние силы, действующие на ткани человека, вызывают деформацию этих тканей и создают в них механические напряжения, которые до определенных значений безопасны для организма, но их превышение может приводить к травмам: растяжению мышц, связок, вывихам, переломам и др.

5.2. Виды деформаций и механические характеристики упругих тел

Механические свойства тел проявляются в их реакции на действие внешних сил (нагрузок). Изменения формы и/или размеров образца под действием внешних сил называют ***деформацией***.

Деформации бывают упругие и пластические.

Деформация называется *упругой*, если после снятия нагрузки тело полностью восстанавливает свою исходную форму и размеры; *пластической* (*неупругой*), если размеры тела восстанавливаются не полностью, т.е. имеет место остаточная деформация.

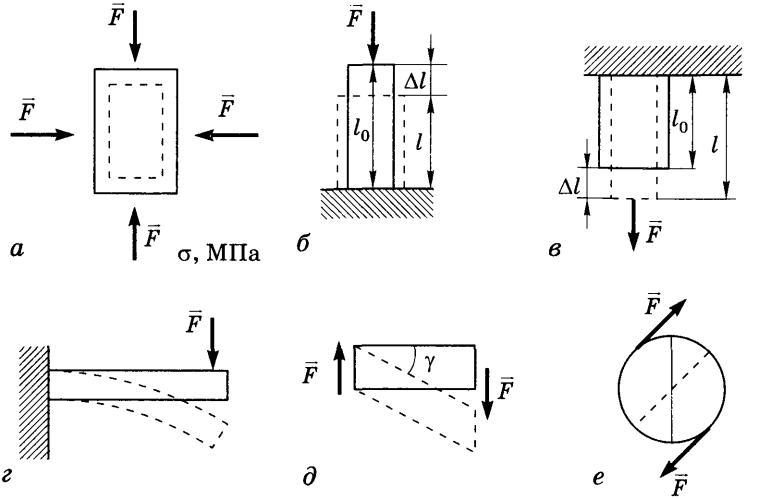


Рис. 5.4. Основные виды деформации твердых тел: *а* — всестороннее сжатие; *б* — одноосное сжатие; *в* — одноосное растяжение; *г* — изгиб; *д* — сдвиг; *е* — кручение

Соответственно *упругость* — это способность тела восстанавливать исходные размеры и форму после снятия нагрузки, а *пластичность* — это способность сохранять остаточную деформацию после снятия нагрузки.

Различают также виды деформации по изменению формы тела и направлению приложенных нагрузок. Они показаны на рис. 5.4 и представлены деформациями *сжатия, растяжения, сдвига, изгиба и кручения*; направление внешних сил (нагрузок) указано стрелками, а пунктирные линии соответствуют деформированному образцу.

Отметим, что разнообразные виды деформации твердого тела могут быть сведены к двум основным: деформации растяжения (или сжатия) и деформации сдвига.

Рассмотрим деформацию ***растяжения (сжатия)***. Ее количественной мерой является *абсолютная* либо *относительная* деформация. Например, если при одноосном растяжении первоначальная длина l_0 тела возросла до значения l, то величина $\Delta l = l - l_0$ называется *абсолютной* деформацией (удлинением) и измеряется в единицах длины. *Относительная*

деформация ε — величина безразмерная и определяется как отношение абсолютной деформации к первоначальной длине образца:

$$\varepsilon = \frac{\Delta l}{l_0}, \text{ или } \varepsilon = \frac{\Delta l}{l_0} \cdot 100\,\%. \qquad (5.4)$$

При деформации в образце возникают упругие силы $F_{упр}$, противодействующие внешней нагрузке и препятствующие его дальнейшей деформации. Деформация тела под действием внешней силы $F_{вн}$ продолжается до тех пор, пока возникающие в нем упругие силы, направленные против $F_{вн}$, не скомпенсируют действие внешней силы: $F_{упр} = F_{вн}$ (в векторном виде $\vec{F}_{упр} = -\vec{F}_{вн}$). Наличие в деформируемом образце упругих сил $F_{упр}$, распределенных по всему его объему, приводит к возникновению в этом образце *механического напряжения* σ, которое равно отношению упругой силы к площади S поперечного, т.е. перпендикулярного действующей на образец силе, сечения образца:

$$\sigma = \frac{F_{упр}}{S}. \qquad (5.5)$$

Еще в XVII веке английский физик Р. Гук установил закон: упругая сила $F_{упр}$, возникающая в образце при упругих деформациях, прямо пропорциональна величине абсолютной деформации:

$$F_{упр} = k\Delta l. \qquad (5.6)$$

Упругая сила всегда противоположна направлению деформации x, поэтому в векторной форме *закон Гука* записывается в виде

$$\vec{F} = -k\vec{x}, \qquad (5.6а)$$

где $x = \Delta l$.

Коэффициент пропорциональности k называют *жесткостью* образца, он зависит как от упругих свойств материала, из которого изготовлен образец, так и от его геометрических размеров; в СИ измеряется в Н/м.

Закон Гука (5.6) часто выражают и в другой форме: механическое напряжение σ, возникающее в образце при упругих

деформациях, прямо пропорционально его относительной деформации ε:

$$\sigma = E\varepsilon. \tag{5.7}$$

Здесь коэффициент пропорциональности E уже не зависит от геометрических размеров образца, а определяется только упругими свойствами материала, из которого он изготовлен, и называется *модулем упругости* (*модулем Юнга*). В СИ он измеряется в паскалях (1 Па = 1 Н/м^2) и его значения для разных материалов приводятся в справочных таблицах.

Подставляя формулы (5.4) и (5.5) в уравнение (5.7) и сопоставляя получившееся выражение с законом Гука (5.6), найдем связь жесткости k образца с его размерами (l_0 и S) и модулем упругости E:

$$k = E\frac{S}{l_0}. \tag{5.8}$$

Видно, что жесткость образца k прямо пропорциональна модулю упругости E, площади поперечного сечения образца S и обратно пропорциональна его длине l_0, т.е. чем короче и толще образец, тем он жестче.

В целом зависимость между механическим напряжением σ в образце и его относительной деформацией ε достаточно сложная и представлена диаграммой растяжения образца на рис. 5.5.

На участке OA зависимость между σ и ε прямолинейная и соответствует закону Гука (5.7). При этом угол наклона α

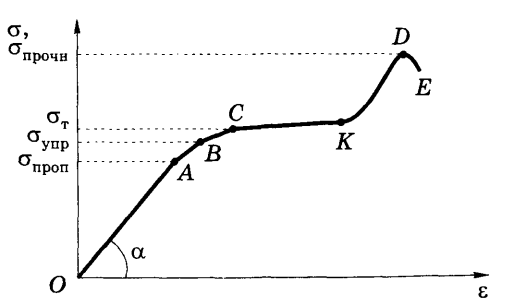


Рис. 5.5. Диаграмма растяжения

определяется модулем упругости E материала образца: $\mathrm{tg}\alpha = E$.

На участке AB деформация остается упругой, но зависимость между σ и ε становится нелинейной, поэтому точке A соответствует *предел пропорциональности* $\sigma_{\text{проп}}$, а точке B — *предел упругости* $\sigma_{\text{упр}}$, поскольку при $\sigma > \sigma_{\text{упр}}$ деформация тела становится уже неупругой (пластической). Закон Гука (5.7) выполняется только на линейном участке OA, т.е. при $\sigma \le \sigma_{\text{проп}}$.

На участке CK относительное удлинение ε образца растет при почти постоянном механическом напряжении, поэтому его называют участком текучести, а напряжение $\sigma_{\text{т}}$, с которого начинается этот участок, — *пределом текучести*.

На участке KD образец опять оказывает сопротивление деформации (т.е. σ растет с увеличением ε) и точке D соответствует *предел прочности* $\sigma_{\text{прочн}}$ — величина механического напряжения в образце, после которого он начинает необратимо разрушаться, так что на участке DE относительная деформация ε растет даже при снижении механического напряжения.

Важной характеристикой материала является *коэффициент Пуассона* μ, который связывает относительные продольную ε и поперечную ε_1 деформации образца:

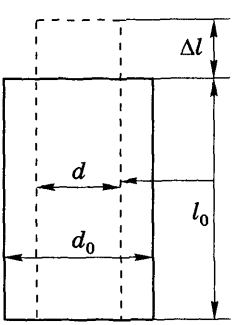


Рис. 5.6. Связь продольной и поперечной деформаций

$$\mu = -\frac{\varepsilon_1}{\varepsilon}. \qquad (5.9)$$

Эти деформации всегда имеют разные знаки, а коэффициент Пуассона всегда положительный. Например, при растяжении образца (рис. 5.6)

$$\varepsilon = \frac{\Delta l}{l_0} > 0, \quad \varepsilon_1 = \frac{\Delta d}{d_0} < 0,$$

так как $\Delta d < 0$, при сжатии — наоборот:

$$\varepsilon < 0, \ \varepsilon_1 > 0.$$

Коэффициент Пуассона μ зависит только от свойств материала, из которого из-

готовлен образец, и определяет относительное изменение объема V образца при деформации:

$$\frac{\Delta V}{V} = \varepsilon(1 - 2\mu). \qquad (5.10)$$

Если коэффициент Пуассона $\mu = 0{,}5$, то $\Delta V = 0$, т.е. вещество несжимаемо.

Рассмотрим теперь ***деформацию сдвига*** (см. рис. 5.4, *д*), которая возникает при действии на тело параллельных противоположно направленных *касательных* сил F, приложенных к разным, но параллельным плоскостям, площадью S каждая. Количественной мерой этой деформации является угол сдвига γ, а отношение $\tau = F/S$ называется касательным (или тангенциальным) механическим напряжением.

Закон Гука для деформации сдвига аналогичен выражению (5.7): касательное (тангенциальное) механическое напряжение τ, возникающее в образце при деформации, прямо пропорционально углу сдвига γ:

$$\tau = G\gamma. \qquad (5.11)$$

Модуль сдвига G зависит только от материала образца и связан с модулем упругости E этого материала и коэффициентом Пуассона μ соотношением

$$G = \frac{E}{2(1 + \mu)}, \qquad (5.12)$$

из которого видно, что значения модуля сдвига всегда меньше значения модуля Юнга данного материала в два и более раза.

Следующим важнейшим свойством деформируемого материала является ***прочность***, характеризующая его способность сопротивляться разрушению при действии внешних сил. Для количественной оценки прочности используют как упомянутый выше *предел прочности* $\sigma_{\text{прочн}}$ материала (его величина зависит от вида деформации), так и величину *разрушающего механического напряжения* — отношение величины разрушающей нагрузки к площади поперечного сечения образца в месте разрушения.

Твердость характеризует сопротивление материала проникновению в него более твердого тела при вдавливании. Количественно твердость определяется как отношение нагрузки F, действующей на вдавливаемое тело (индентор), к площади S поверхности отпечатка, образовавшегося в материале; в СИ измеряется в Н/м^2. Методы измерения твердости отличаются формой и материалом вдавливаемого тела (индентора). Например, при определении твердости по методу Бринелля в образец вдавливается стальной закаленный шарик, по методу Виккерса — алмазная пирамидка определенной формы и т.п. Обозначения твердости: H_B (или HB) — твердость по Бринеллю, H_V (или HV) — твердость по Виккерсу и т.д.

Силы (нагрузки), действующие на твердые тела, обычно сосредоточены в точке либо распределены по некоторой поверхности. Различают также нагрузки статические и динамические. Первые длительно нагружают тело и с течением времени не меняют своей величины, а вторые быстро изменяются во времени по величине и даже по направлению. Примерами динамических нагрузок являются ударные и повторно-переменные (циклические) нагрузки.

5.3. Механические свойства биотканей

Изучение механических свойств биологических тканей необходимо как для выяснения допустимых неразрушающих деформаций этих тканей, так и для совершенствования методов лечения травм, средств защиты человека от неблагоприятных силовых воздействий, протезирования органов и тканей, создания новых материалов, близких по свойствам к биотканям.

Биологические ткани принято делить на *жидкие* (кровь, лимфа, слизистые жидкости, синовиальная жидкость и т.д.), мягкие и твердые, причем последнее деление достаточно условно. К *мягким* относятся ткани, для которых упругие (обратимые) деформации могут составлять десятки и сотни процентов, — кожа, мышечная ткань, ткани легкого и мозга, стенки кровеносных сосудов, дыхательных путей и неко-

торые другие, к *твердым* — кость и зубы. Промежуточное положение занимают сухожилие и суставной хрящ. Для определенности сухожилие отнесем к мягким тканям, а хрящ — к твердым.

5.3.1. Особенности механического поведения биотканей

Выделяют следующие основные особенности механических свойств биологических тканей.

1. Большинство биологических тканей *анизотропно*, т.е. их физические, в том числе и механические, свойства неодинаковы в разных направлениях, что обусловлено неоднородным строением.

2. Почти для всех биотканей характерны следующие временные эффекты при деформировании:

а) при постоянной величине деформации в биотканях со временем происходит *релаксация* (*спад*) *механического напряжения* (рис. 5.7, *а*);

б) при постоянной нагрузке (механическом напряжении) величина деформации со временем увеличивается — это явление называют *ползучестью* (*крипом*) (рис. 5.7, *б*);

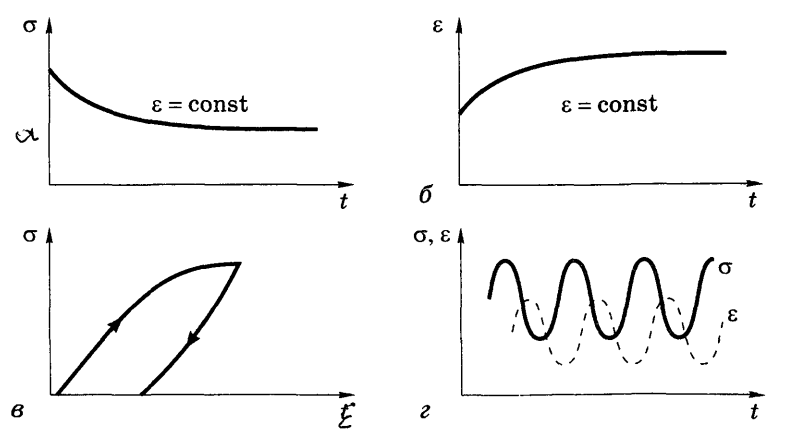


Рис. 5.7. Временны́е эффекты, возникающие при деформировании биотканей:
а — релаксация напряжения; *б* — ползучесть; *в* — петля гистерезиса; *г* — отставание по фазе деформации от напряжения

в) зависимостям механического напряжения от деформации при нагрузке и разгрузке соответствуют разные кривые, в результате образуется так называемая *петля гистерезиса* (рис. 5.7, *в*).

г) при циклическом нагружении биотканей в них возникают колебания механических напряжений и деформаций, при этом деформация ε всегда отстает по фазе от механического напряжения σ (рис. 5.7, *г*);

д) механические характеристики биотканей часто зависят от скорости деформации.

Перечисленные факты — проявление ***вязкоупругого*** поведения биоткани.

Внешние силы, действующие на образец, совершают работу по его деформированию. Энергия, затраченная на деформацию образца, во-первых, частично запасается в объеме деформированной среды в виде потенциальной энергии деформации и, во-вторых, необратимо расходуется на преодоление сил внутреннего трения, переходя в конечном итоге в теплоту.

Если при деформации реализуется только первый случай, то материал образца является *упругой* средой и при снятии нагрузки вся запасенная в нем потенциальная энергия деформации переходит в работу по восстановлению его прежних размеров и формы.

Если реализуется только второй случай, т.е. происходит лишь необратимое рассеивание энергии в образце, то среда называется *вязкой* и все деформации в такой среде необратимы (неупруги).

Если же в разной степени реализуются оба случая, то среда называется *вязкоупругой*. При деформировании такой среды часть работы внешних сил запасается в образце в виде энергии упругой деформации, а оставшаяся часть необратимо рассеивается в виде тепла. После прекращения действия внешних сил происходит лишь частичное восстановление размеров и формы образца и формируется петля гистерезиса (см. рис. 5.7, *в*), площадь которой (между кривыми нагружения и разгрузки образца) определяет долю необратимо затраченной энергии, перешедшей в тепло.

5.3.2. Механические свойства костной ткани

Несмотря на существование различных типов костей в теле человека и животных (длинные трубчатые кости конечностей, плоские кости черепа, короткие кости — позвонки), для них характерны общие механические свойства, которые далее будут рассмотрены на примере компактной и губчатой (спонгиозной) составляющих трубчатых костей.

Костная ткань — один из видов соединительной ткани, ее механические свойства определяются прежде всего составом твердой фазы кости и свойствами ее компонент. ***Твердая фаза кости*** — это органические и минеральные составляющие костных пластинок (ламелл). *Органическая* составляющая представлена коллагеном (40–50 % твердой фазы), а *минеральная* (60–50 %) — преимущественно кристаллами гидроксиапатита $Ca_{10}(PO_4)_6(OH)_2$ и других солей кальция.

Если экспериментально полностью удалить из кости органическую составляющую (коллаген), то кость становится необычайно хрупкой, а если удалить минеральную составляющую, то кость приобретает резиноподобные свойства. Это означает, что костная ткань является прочным материалом лишь при определенном соотношении входящих в нее компонентов. В статических условиях модуль упругости коллагена равен $E_к \sim 10^9$ Па, гидроксиапатита — $E_г \sim 10^{11}$ Па, а для компактной ткани кости $E_{комп} \sim 10^{10}$ Па, что сравнимо с модулями упругости стали и титана: $E_{стали} \approx 2 \cdot 10^{10}$ Па $= 20$ ГПа, $E_{титана} \approx 10^{11}$ Па $= 100$ ГПа.

Минимальные разрушающие напряжения составляют для гидроксиапатита 600–700 МПа, для коллагена 50–100 МПа, а для разных костных тканей принимают промежуточные значения.

Анизотропия механических свойств кости проявляется в существенной зависимости ее упругих свойств от направления нагружения. На рис. 5.8 приведены кривые, полученные при деформировании образца воздушно-сухой компактной костной ткани большеберцовой кости человека при растяжении ее вдоль продольной оси Ox (кривая *1*) и по касательному Oy (кривая *2*) и радиальному Oz (кривая *3*) направле-

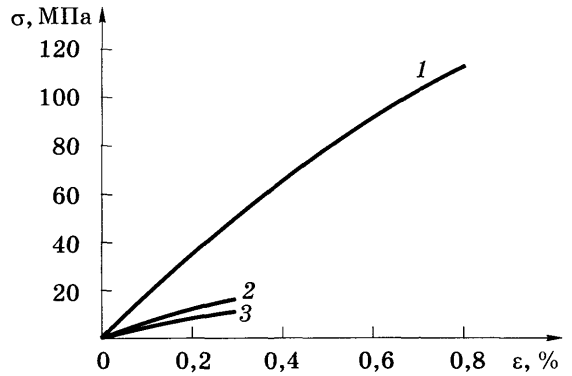


Рис. 5.8. Диаграммы растяжений компактной костной ткани человека вдоль разных направлений: *1* — в продольном (x); *2* и *3* — в поперечных (y и z) направлениях

ниям в плоскости, перпендикулярной Ox. Напомним, что тангенс угла наклона этих кривых к оси абсцисс численно равен модулю упругости и несовпадение этих кривых между собой указывает на значительные различия механических свойств кости в указанных направлениях. Для большеберцовой кости взрослого человека характерны средние значения модулей упругости в этих направлениях: $E_x = 18{,}35$ ГПа, $E_y = 8{,}5$ ГПа, $E_z = 6{,}9$ ГПа, что свидетельствует о значительной анизотропии костной ткани: вдоль продольной оси модуль упругости более чем в два раза выше, чем в поперечных направлениях.

5.3.3. Механические свойства суставного хряща

Суставной хрящ покрывает концевые поверхности трубчатых костей. Он представляет собой пористый, легкопроницаемый материал с низким модулем упругости и во время действия нагрузок выделяет, а при разгрузке поглощает синовиальную жидкость. Его размеры и форма могут изменяться вследствие как кратковременных, так и длительных силовых воздействий, причем после разгрузки исходное состояние суставного хряща полностью восстанавливается, что обеспечивает конгруэнтность (от лат. congruens (congruentis) — сораз-

мерный, соответствующий,совпадающий) суставных поверхностей.

В хряще различают жидкую и твердую фазы. *Жидкая фаза* составляет 60–80 % его массы. В оставшейся *твердой фазе* содержится около 40 % хондроцитов, 25 % протеогликанов и 35 % коллагеновых волокон. *Синовиальная жидкость* выполняет функцию смазки сустава и обеспечивает питание хряща.

С точки зрения механики материалов суставной хрящ является вязкоупругим анизотропным материалом с неоднородными механическими свойствами по суставной поверхности. Он обладает двумя функциональными особенностями:

- под действием нагрузки деформируется легче, чем кость (отношение их жесткостей равно 1:10), благодаря чему хрящ распределяет нагрузки на всю суставную поверхность кости и снижает этим концентрацию в ней механических напряжений;
- обеспечивает чрезвычайно низкий коэффициент трения в суставе (от 0,005 до 0,012).

При нормальном функционировании сустава нагрузка на суставной хрящ колеблется от нуля до значений, в 3–4 раза превышающих вес тела, при этом по краям контактной поверхности могут развиваться достаточно большие деформации растяжения. Следует отметить, что локальная деструкция суставного хряща не восстанавливается и поэтому является опасным повреждением, которое в дальнейшем может привести к дегенерации всего сустава.

Анизотропия механических свойств хряща проявляется в существенном различии разрушающих растягивающих напряжений вдоль и поперек направления коллагеновых волокон. Так, для наружного слоя хряща мыщелков бедренной кости они равны соответственно 25,5 и 9,8 МПа, т.е. различаются в 2,5 раза.

Модуль упругости E суставного хряща, определенный в эксперименте на сжатие, составляет 23–50 МПа, модуль сдвига G — 0,4–4 МПа, причем при дегенеративных изменениях в хряще модуль сдвига увеличивается.

Коэффициент Пуассона хряща равен приблизительно 0,5, поэтому его можно рассматривать как несжимаемый материал.

Большой практический интерес представляет исследование усталостной долговечности хряща, ведь человек делает в среднем около 2 млн шагов в год и тем самым подвергает данную ткань постоянному циклическому нагружению. Установлено, что при циклических нагружениях усталостная долговечность хряща значительно снижается при возрастании уровня механических напряжений в нем, а также с увеличением его возраста.

Таблица 5.1

Вещество	Модуль упругости, МПа	Предел прочности на растяжение или сжатие, МПа	Коэффициент Пуассона
Компактное вещество кости	7000–18 000	120–200	0,2–0,4
Губчатое вещество кости	26–600	3,7–11,4	—
Суставной хрящ	23–50	10–25	0,5

В табл. 5.1 приведены основные механические свойства компактного и губчатого вещества бедренной кости и суставного хряща человека. Широкий диапазон изменений параметров вызван анизотропией свойств этих тканей и разными видами и направлениями их нагружения.

5.3.4. Механические свойства мягких биологических тканей

К мягким биологическим тканям относятся кожа, мышцы, стенки кровеносных сосудов и дыхательных путей, лепестки клапанов сердца, сухожилия, связки и др. Большинство из них имеет общие черты в механическом поведении: способность к большим деформациям — до 200 %, несжимаемость, анизотропность.

Эти свойства обусловлены в первую очередь строением этих тканей. В их состав входят клетки, коллагеновые и эластиновые волокна и основное вещество. Каждая компонента имеет свою ультраструктуру, которая определяет ее механические свойства.

Коллагеновые волокна имеют диаметр 0,2–12 мкм и представляют собой пучки фибрилл диаметром 20–40 нм, в зависимости от вида млекопитающего и ткани. *Эластиновые структуры* в мягких биологических тканях встречаются в двух морфологически различных формах: волокон (кожа, вены, хрящи) и мембран (артерии).

Сравнительная характеристика упругих свойств эластиновых и коллагеновых волокон (табл. 5.2), [26] показывает, что коллагеновые волокна испытывают относительно небольшие деформации, но обладают гораздо большей по сравнению с эластином прочностью и отличаются от него значительно более высоким модулем упругости.

Таблица 5.2

Вещество	Модуль упругости, МПа	Предел прочности на растяжение, МПа	Удлинение при разрыве, %
Эластиновые волокна	0,6	1	200
Коллаген	1000	50–100	50

Характер взаимодействия и количественные соотношения эластиновых и коллагеновых волокон в ткани определяют оптимальное соотношение прочностных и деформационных свойств мягких биологических тканей.

Коллагеновые волокна в мягких тканях уложены волнообразно и распрямляются лишь при значительных деформациях ткани. Поэтому на первоначальных этапах деформации упругие свойства мягких тканей определяет эластин, а затем, после распрямления коллагеновых волокон, упругие свойства этих тканей определяются уже значительно более жесткими коллагеновыми волокнами. Это хорошо видно на

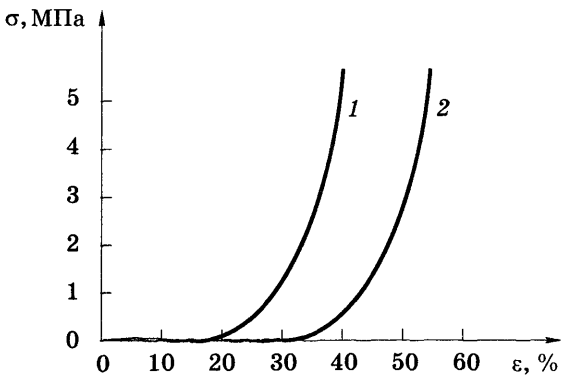


Рис. 5.9. Анизотропия механических свойств кожи при растяжении в поперечном (*1*) и продольном (*2*) направлении

рис. 5.9, где представлены диаграммы растяжения кожи в продольном (вдоль оси тела) и поперечном направлениях.

С увеличением деформации кожи ее упругое сопротивление сначала невелико (обусловлено деформацией эластина), а затем, после распрямления коллагеновых волокон, резко возрастает. Большинство коллагеновых волокон в коже ориентировано в поперечном направлении относительно продольной оси тела, что обусловливает анизотропию механических свойств кожи: она имеет различные модули упругости, разрушающие напряжения, и максимальные деформации в поперечном и продольном направлениях. Так, предел прочности кожи в поперечном направлении (вдоль коллагеновых волокон) примерно в 1,5 раза больше, чем в продольном, а ее максимальное растяжение приблизительно вдвое больше в продольном направлении.

Сухожилие по сравнению с другими мягкими тканями имеет относительно простую микроструктуру, которая определяется однонаправленно расположенными коллагеновыми волокнами. Такая структура связана с главной физиологической функцией сухожилия — с необходимостью передавать усилия в одном направлении: от мышцы к кости. В фазе расслабления коллагеновые волокна сухожилий имеют волнообразную форму, а при передаче усилия они распрямляются и далее почти не деформируются.

Модуль упругости E сухожилий составляет свыше 100 МПа. Сухожилие имеет самый высокий предел прочности на разрыв и наименее растяжимо среди неминерализованных тканей. Механические свойства сухожилий зависят от пола и возраста человека и максимальную прочность они приобретают к 21–25 годам. Значения характерных показателей для некоторых мягких тканей приведены в табл. 5.3.

Таблица 5.3

Ткань	Модуль упругости, МПа	Предел прочности при растяжении, МПа	Удлинение при разрыве, %
Сухожилие	> 100	53	9,7
Кожа	2–40	7,6	78
Аорта в продольном направлении	0,84–3,9	0,7–1	46
Мышца сердца	1	0,11	63,8

Таким образом, механические свойства мягких и твердых биотканей существенно зависят от того, как взаимно организованы волокна, клетки и основное вещество в структуре этих тканей.

Контрольные вопросы и задания

1. Чем отличаются рычаги скорости и силы? Приведите примеры этих рычагов.
2. Что такое изотонический режим работы мышцы? Что устанавливает уравнение Хилла?
3. Дайте определение деформации твердого тела. В чем состоит различие между упругой и пластической деформацией? Перечислите основные виды деформации по изменению формы твердых тел.
4. Что является количественной мерой деформации? Укажите единицы ее измерения. Что называют упругой силой, как она направлена? Дайте определение механического напряжения, укажите его размерность в СИ.

5. Приведите и проанализируйте закон Гука для деформации одноосного растяжения (сжатия). Как связаны коэффициенты жесткости и модуль Юнга, от чего они зависят?
6. Приведите и проанализируйте диаграмму растяжения. Определите пределы пропорциональности, упругости, текучести, прочности деформируемого материала.
7. Приведите и проанализируйте закон Гука для деформации сдвига. Что представляют собой касательное напряжение, модуль сдвига, коэффициент Пуассона, какова связь между ними и модулем Юнга?
8. Дайте определение понятию «твердость материала». Укажите известные вам методы определения твердости, различия между ними, единицы измерения.
9. Чем различаются статические и динамические нагрузки? Приведите примеры.
10. Назовите особенности механических свойств биотканей. В чем суть анизотропии биотканей, какова ее природа?
11. Каковы механические свойства и особенности твердых биологических тканей (костная ткань, суставной хрящ)?
12. Каковы механические свойства и особенности мягких биологических тканей (сухожилия, кожа, ткань кровеносных сосудов)?
13. В чем выражается явление вязкоупругости? Как проявляется вязкоупругое поведение биотканей?

Глава 6 Механические колебания и волны

Колебательное движение может совершаться за счет запаса энергии колебательной системы и без внешних воздействий. Такие колебания называются ***свободными***. При наличии сопротивления движению колебания называются ***затухающими***. Колебательный процесс, возникающий под действием переменных во времени внешних сил, называют ***вынужденными*** колебаниями.

6.1. Гармонические колебания

Простейшим видом колебательного движения являются ***гармонические колебания***, когда колеблющаяся величина изменяется со временем по закону синуса или косинуса:

$$x = A\sin(\omega t + \varphi_0), \tag{6.1}$$

где x — мгновенное смещение тела от положения равновесия; A — амплитуда колебания, т.е. максимальное смещение; $(\omega t + \varphi_0)$ — мгновенная фаза колебания, ω — циклическая частота колебания, t — текущее время, φ_0 — начальная фаза колебания.

Циклическая частота ω, линейная частота ν и период колебания T связаны соотношениями

$$\nu = \frac{1}{T}, \quad \omega = 2\pi\nu = \frac{2\pi}{T}. \tag{6.2}$$

Покажем, что гармонические колебания возникают под действием упругой силы или ей подобной (квазиупругой). Представим материальную точку массой m, закрепленную на пружине жесткости k и расположенную на абсолютно гладкой горизонтальной поверхности (рис. 6.1). Если растянуть пружину на расстояние x, то со стороны пружины на эту точ-

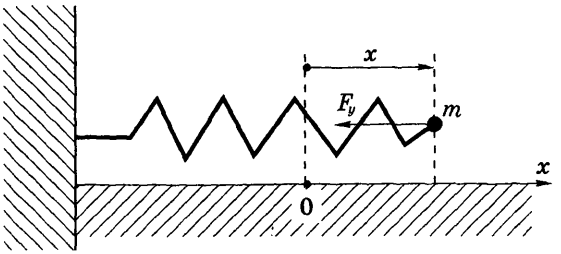


Рис. 6.1. Колебания материальной точки на пружине

ку действует упругая сила $F_{упр}$, пропорциональная по закону Гука смещению x:

$$F_{упр} = -kx.$$

Знак «–» указывает на противоположность направлений смещения и силы упругости.

Для того чтобы установить характер движения, т.е. зависимость $x = f(t)$, *запишем второй закон Ньютона в виде* дифференциального уравнения (поскольку ускорение тела $a = x'' = \dfrac{d^2x}{dt^2}$), учитывая, что в рассматриваемом случае движение определяется только наличием силы упругости:

$$m\frac{d^2x}{dt^2} = -kx\,. \tag{6.3}$$

Разделим левую и правую части уравнения (6.3) на m и обозначим отношение положительных величин k и m через ω_0^2:

$$\frac{d^2x}{dt^2} = -\omega_0^2 x \quad \text{или} \quad x'' + \omega_0^2 x = 0. \tag{6.4}$$

Покажем, что решением этого дифференциального уравнения является гармоническое колебание:

$$x = A_0 \sin(\omega_0 t + \varphi_0). \tag{6.5}$$

Для этого дважды продифференцируем выражение (6.5):

$$v = x' = A_0\omega_0 \cos(\omega_0 t + \varphi_0)\,, \tag{6.6}$$

$$a = x'' = -A_0\omega_0^2 \sin(\omega_0 t + \varphi_0) \tag{6.7}$$

и подставим результат в уравнение (6.4):

$$-A_0\omega_0^2 \sin(\omega_0 t + \varphi_0) + \frac{k}{m} A_0 \sin(\omega_0 t + \varphi_0) = 0$$

или $A_0\left(\frac{k}{m} - \omega_0^2\right)\sin(\omega_0 t + \varphi_0) = 0$.

Это равенство должно выполняться в любой момент времени, что возможно лишь при условии

$$\omega_0^2 = \frac{k}{m}. \qquad (6.8)$$

Следовательно, если в качестве результирующей на тело действует сила упругости, то тело совершает гармонические колебания вида (6.5), а циклическая частота колебания $\omega_0 = \sqrt{k/m}$ определяется только массой тела и упругими свойствами (жесткостью) системы и поэтому называется ***собственной частотой.***

Сравнение выражений (6.5) и (6.6) показывает, что смещение и скорость гармонического колебания различаются по фазе на $\pi/2$, амплитуда скорости равна $\upsilon_{max} = A_0\omega_0$. Скорость тела максимальна при прохождении положения равновесия (смещение равно нулю), а при максимальном смещении (равном амплитуде) скорость равна нулю.

Из формул (6.5) и (6.7) следует, что смещение тела и его ускорение изменяются в противофазе и амплитуда ускорения равна $a_{max} = \upsilon_{max}\omega_0 = A_0\omega_0^2$.

Полная энергия гармонического колебания E определяется суммой кинетической $E_к$ и потенциальной $E_п$ энергии:

$$E = E_к + E_п = \frac{m\upsilon^2}{2} + \frac{kx^2}{2}. \qquad (6.9)$$

Подставляя в формулу (6.9) выражение для скорости из (6.6), а для смещения — из (6.5) и учитывая, что $k = \omega_0^2 m$, получим

$$E = \frac{1}{2} m A_0^2 \omega_0^2 [\cos^2(\omega_0 t + \varphi_0) + \sin^2(\omega_0 t + \varphi_0)] = \frac{1}{2} m A_0^2 \omega_0^2. \qquad (6.9а)$$

Следовательно, полная энергия гармонического колебания не зависит от времени и прямо пропорциональна массе тела, квадрату амплитуды и квадрату частоты колебания. Кинетическая и потенциальная энергии колеблющегося тела непрерывно изменяются во времени, но их сумма всегда остается неизменной.

6.2. Затухающие колебания

В реальных системах всегда существуют некоторые силы сопротивления, препятствующие развитию колебательных процессов. Для установления характера колебательного движения в этом случае будем считать, что наряду с упругой или квазиупругой силой $F_{\text{упр}}$ в системе действует сила трения, пропорциональная скорости и направленная противоположно ей: $F_{\text{тр}} = -r\dfrac{dx}{dt}$.

Тогда учет влияния этих двух сил на характер движения приводит к следующему дифференциальному уравнению (второй закон Ньютона)

$$m\frac{d^2x}{dt^2} = -kx - r\frac{dx}{dt}. \qquad (6.10)$$

Разделив левую и правую части этого уравнения на m, обозначив $r/m = 2\beta$ и сохранив обозначение $k/m = \omega_0^2$, приведем это уравнение к виду

$$\frac{d^2x}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x = 0. \qquad (6.11)$$

Решение этого уравнения имеет вид

$$x = A_0 e^{-\beta t}\sin(\omega t + \varphi_0) \qquad (6.12)$$

и показывает, что смещение тела во времени происходит по гармоническому закону с частотой ω, но амплитуда колебаний уменьшается со временем: $A = A_0 e^{-\beta t}$. Коэффициент

$\beta = r/2m$ ($с^{-1}$) называется ***коэффициентом затухания***, и чем он больше, тем быстрее затухают колебания. По сравнению с преыдущим случаем уменьшается и циклическая частота собственных колебаний ω:

$$\omega = \sqrt{\omega_0^2 - \beta^2}\,. \qquad (6.13)$$

Колебательный процесс может происходить лишь при условии $(\omega_0^2 - \beta^2) > 0$, т.е. при $\omega_0 > \beta$. Если же затухание в системе слишком велико ($\omega_0 < \beta$), то под знаком корня в формуле (6.13) оказывается отрицательная величина, в этом случае колебания невозможны и движение имеет апериодический характер.

Графически затухающее колебание представлено на рис. 6.2, где сплошной линией показана зависимость смещения от времени, а пунктирной — экспоненциальный закон убывания амплитуды.

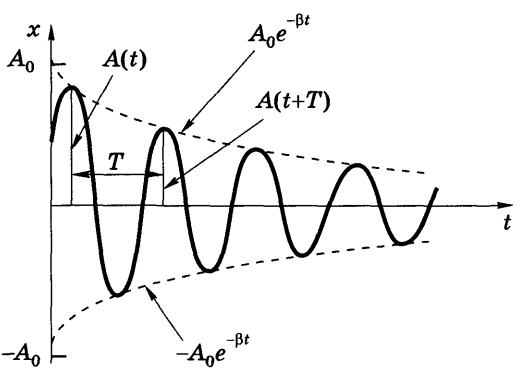


Рис. 6.2. Зависимость смещения x от времени t при затухающем колебании

Однако на практике затухание колебаний удобнее характеризовать ***декрементом затухания*** δ, представляющим собой отношение двух последовательных амплитуд, разделенных периодом колебаний T (см. рис. 6.2):

$$\delta = \frac{A(t)}{A(t+T)}\,. \qquad (6.14)$$

Натуральный логарифм этого отношения называется ***логарифмическим декрементом затухания*** λ; он связан с коэффициентом затухания и периодом:

$$\lambda = \ln\delta = \ln\frac{A_0 e^{-\beta t}}{A_0 e^{-\beta(t+T)}} = \ln e^{\beta T} = \beta T, \text{ или } \lambda = \beta T. \quad (6.15)$$

Удобство использования логарифмического декремента затухания λ для характеристики затухающих колебаний заключается в простоте его экспериментального определения. Если затухающие колебания зарегистрированы в виде соответствующего графика (см. рис. 6.2), то необходимо в любых единицах измерить две амплитуды колебаний, взятые через период, и найти натуральный логарифм их отношения. Определив таким образом λ и зная период T, легко найти и коэффициент затухания β.

6.3. Вынужденные колебания

Рассмотрим случай, когда в системе наряду с силами упругости и трения присутствует некоторая внешняя сила, препятствующая затуханию колебаний. Предположим, что эта вынуждающая сила $F_в$ действует периодически с круговой частотой Ω и зависит от времени по закону

$$F_в = F_0 \sin\Omega t,$$

где F_0 — амплитуда вынуждающей силы.

В этом случае дифференциальное уравнение колебаний, составленное на основании второго закона Ньютона, имеет вид

$$m\frac{d^2x}{dt^2} = -kx - r\frac{dx}{dt} + F_0 \sin\Omega t. \quad (6.16)$$

Сохраняя обозначения $k/m = \omega_0^2$, $r/m = 2\beta$ и обозначив $F_0/m = f_0$, приведем уравнение (6.16) к виду

$$\frac{d^2x}{dt^2} + 2\beta\frac{dx}{dt} + \omega_0^2 x = f_0 \sin\Omega t. \quad (6.17)$$

Решение этого уравнения представляет некоторую функцию, которая графически показана на рис. 6.3. Это решение состоит из двух частей. Одна из них соответствует *неустановившемуся* режиму колебаний, когда их амплитуда зависит от времени. Вторая часть описывает *установившийся* режим колебаний.

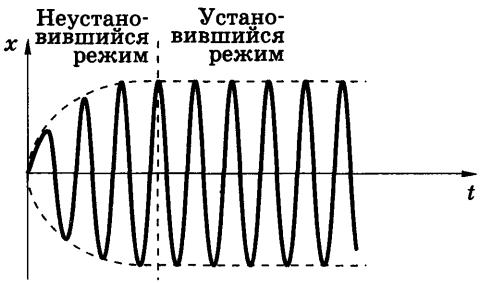


Рис. 6.3. Зависимость смещения x от времени t при вынужденном колебании

В установившемся режиме вынужденных колебаний смещение x подчиняется гармоническому закону и происходит с частотой, равной частоте действия вынуждающей силы:

$$x = A\sin(\Omega t + \varphi). \qquad (6.18)$$

Установившаяся амплитуда A вынужденных колебаний зависит от параметров системы (частоты собственных колебаний ω_0 и коэффициента затухания β) и от характеристик вынуждающей силы (f_0 и Ω): $A = f(\omega_0, \beta, f_0, \Omega)$. Строгое рассмотрение приводит к следующим выражениям для значений A и φ, входящим в формулу (6.17):

$$A = f_0 / \sqrt{(\omega_0^2 - \Omega^2)^2 + 4\beta^2\Omega^2}, \qquad (6.19)$$

$$\operatorname{tg}\varphi = -2\beta\Omega/(\Omega^2 - \omega_0^2). \qquad (6.20)$$

Из формулы (6.19) следует, что амплитуда достигает максимального значения A_{max}, когда знаменатель принимает минимальное значение, что происходит при условии

$$\Omega = \Omega_{рез} = \sqrt{\omega_0^2 - 2\beta^2}. \qquad (6.21)$$

Итак, если частота вынуждающей силы определяется формулой (6.21), то амплитуда вынужденных колебаний будет максимальной.

Явление резкого возрастания амплитуды вынужденных колебаний, происходящее на некоторой частоте действия вынуждающей силы, называется ***резонансом***, а частота, определяемая формулой (6.21), — ***резонансной***. Если бы затухание в системе отсутствовало ($\beta = 0$), то резонанс наступал бы при условии $\Omega = \omega_0$, а амплитуда стремилась бы к бесконечно большим значениям.

6.4. Сложение гармонических колебаний

Результат сложения гармонических колебаний зависит от направления смещений складываемых колебаний, а также от соотношения между их частотами, фазами и амплитудами.

Сложение колебаний с одинаковыми частотами, происходящих вдоль одной прямой. В этом случае складываемые колебания различаются лишь амплитудами A_1 и A_2 и начальными фазами φ_1 и φ_2:

$$\left.\begin{aligned} X_1 &= A_1\sin(\omega t + \varphi_1) \\ X_2 &= A_2\sin(\omega t + \varphi_2). \end{aligned}\right\} \qquad (6.22)$$

Результатом сложения этих колебаний является гармоническое колебание, происходящее с той же частотой:

$$X = X_1 + X_2 = A_1\sin(\omega t + \varphi_1) + A_2\sin(\omega t + \varphi_2) = A\sin(\omega t + \varphi). \qquad (6.23)$$

Амплитуда результирующего колебания A зависит не только от амплитуд исходных колебаний A_1 и A_2, но и от разности их начальных фаз[1] φ_1 и φ_2:

$$A = \sqrt{A_1^2 + A_2^2 + 2A_1A_2\cos(\varphi_2 - \varphi_1)}. \qquad (6.24)$$

Начальная фаза φ результирующего колебания определяется соотношением

$$\operatorname{tg}\varphi = \frac{A_1\sin\varphi_1 + A_2\sin\varphi_2}{A_1\cos\varphi_1 + A_2\cos\varphi_2}. \qquad (6.25)$$

[1] Рассмотрены только когерентные колебания, т.е. колебания, разность начальных фаз которых постоянна во времени.

Таким образом, результирующее колебание будет происходить по гармоническому закону с той же частотой, что и исходные колебания, но амплитуда результирующего колебания существенно зависит от разности начальных фаз исходных колебаний.

Амплитуда *максимальна* и равна $A = A_1 + A_2$, если фазы исходных колебаний совпадают или отличаются на $2k\pi$, т.е. при $\Delta\varphi = 2k\pi$, где $k = 0, \pm1, \pm2, \ldots$.

Амплитуда *минимальна* и равна $A = |A_1 - A_2|$, если исходные колебания находятся в противофазе, т.е. при $\Delta\varphi = \pi(2k + 1)$, где $k = 0, \pm1, \pm2, \ldots$. При $A_1 = A_2$ такие колебания взаимно «погашают» друг друга и $A = 0$.

При других значениях $\Delta\varphi$ амплитуда результирующего колебания определяется выражением (6.24).

Сложение колебаний с разными частотами, происходящих вдоль одной прямой. Складываемые колебания гармонические и происходят с разными частотами:

$$\left.\begin{aligned} X_1 &= A_1\sin(\omega_1 t + \varphi_1) \\ X_2 &= A_2\sin(\omega_2 t + \varphi_2). \end{aligned}\right\} \qquad (6.26)$$

Они представлены графически на рис. 6.4, *а*, а результат их сложения — на рис. 6.4, *б*.

При сложении гармонических колебаний с разными частотами ω_1 и ω_2 (периодами T_1 и T_2) результирующее колебание не будет гармоническим, а будет представлять более сложное, но периодическое движение.

Если складываются гармонические колебания с кратными частотами (например, на рис. 6.4,*а* $\omega_2 = 4\omega_1$), то период результирующего

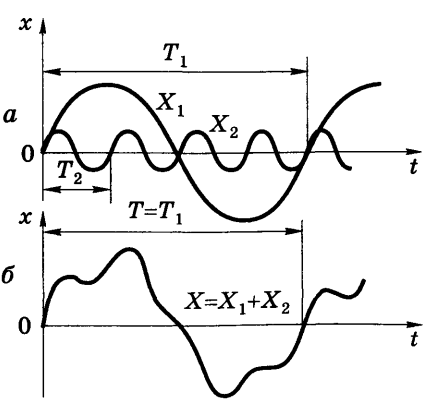


Рис. 6.4. Сложение колебаний с разными частотами $\omega_2 = 4\omega_1$: *а* — исходные гармонические колебания, $T_1 = 4T_2$; *б* — результирующее негармоническое колебание

колебания T совпадает с наибольшим периодом T_1 складываемых колебаний: $T = T_1$, а частота результирующего колебания — с наименьшей частотой: $\omega = \omega_1$.

6.5. Разложение сложных колебаний в ряд Фурье

Рассмотрение рис. 6.4 приводит к выводу, известному как ***теорема Фурье***: любое сложное периодическое движение $x(t) = x(t + T)$ с периодом T (частотой $\nu = 1/T$) можно представить в виде суммы гармонических колебаний (гармоник), частоты которых кратны частоте ν рассматриваемого периодического процесса: $\nu_k = k\nu$. Это утверждение можно записать в виде формулы, представляющей ряд Фурье:

$$x(t) = A_0 + \sum_{k=1}^{\infty} A_k \sin(k\omega t + \varphi_k), \qquad (6.27)$$

где A_k — амплитуды складываемых гармоник; φ_k — начальные фазы гармоник.

Первая гармоника имеет циклическую частоту ω_1, равную циклической частоте сложного колебания $\omega_1 = \omega = 2\pi\nu = 2\pi/T$, вторая — 2ω, третья — 3ω и т.д.

Слагаемое A_0 в формуле (6.27) представляет собой постоянную величину, имеющую смысл постоянной составляющей сложного периодического процесса. На рис. 6.5 представлена периодическая функция $x(t)$, описывающая процессы, где колебания некоторой величины (например, пульсовые изменения кровенаполнения сосуда) происходят на фоне ее среднего постоянного значения (например, среднего уровня кровенаполнения), которое и характеризуется величиной A_0 в формуле (6.27).

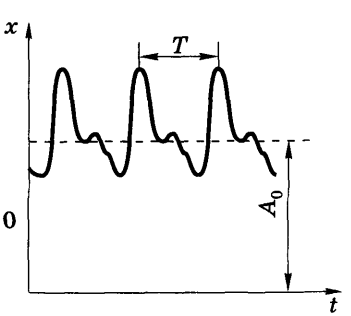


Рис. 6.5. Сложное колебание с постоянной составляющей A_0.

Количество гармоник, входящих в состав сложного колебания (6.27), определяется сложностью исходного колебания $x(t)$ и может

быть очень большим (десятки, сотни и тысячи гармоник). При рассмотрении реальных колебательных процессов следует учесть, что вклад отдельных гармонических составляющих в анализируемое сложное колебание различен, в формулу (6.27) отдельные гармоники входят с разными амплитудами.

График, по оси абсцисс которого отложены частоты гармоник, а по оси ординат соответствующие им амплитуды, представляет собой ***гармонический спектр сложного колебания*** (рис. 6.6).

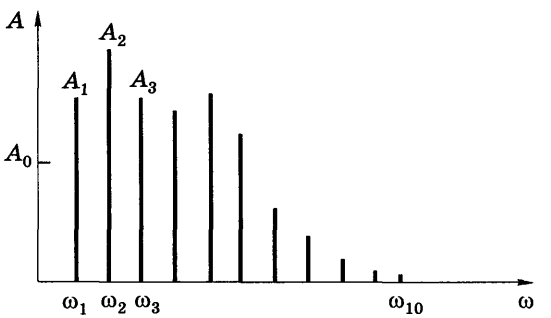


Рис. 6.6. Гармонический спектр сложного колебания

Согласно рис. 6.6, можно сделать вывод, что гармоники, частота которых превышает ω_{10}, имеют очень малую амплитуду и, следовательно, их вклад в колебание, гармонический спектр которого представлен на рисунке, незначителен. Поэтому можно считать, что ряд Фурье для этого случая содержит 10 слагаемых ($k = 1, 2, 3, \ldots, 10$) и вся информация о сложном колебательном процессе заключена в полосе частот от ω_1 (основная частота процесса) до ω_{10}.

6.6. Принципы применения Фурье-анализа для обработки диагностических данных

Многочисленные процессы жизнедеятельности организма носят периодический характер (сердечные сокращения, дыхание, кровенаполнение сосудов и т.д.), поэтому дагностические данные, позволяющие судить о работе ряда органов

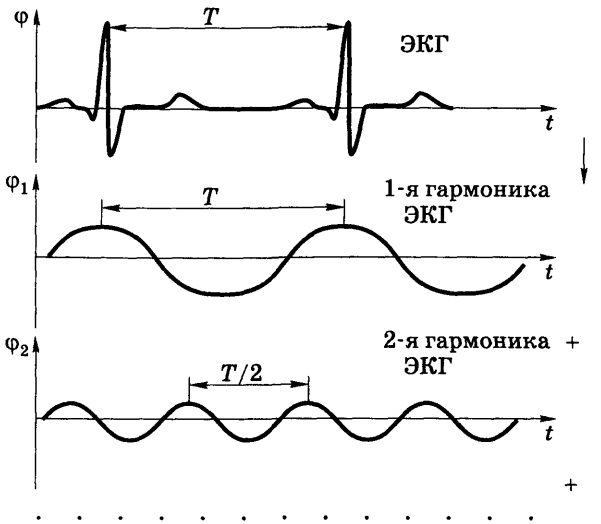


Рис. 6.7. Разложение ЭКГ на отдельные гармонические составляющие

и функциональных систем организма, представляются в виде периодических кривых. Например, электрокардиограмма (ЭКГ) представляет собой зафиксированную на бумажной ленте или на экране монитора сложную периодическую зависимость от времени t биопотенциалов φ, сопровождающих работу сердца (рис. 6.7). Обработка данных ЭКГ может быть произведена с помощью гармонического анализа. С помощью специальных приборов — ***анализаторов*** получают гармонический спектр ЭКГ. Частота первой гармоники в этом спектре соответствует частоте сердечных сокращений у пациента. Она составляет около 1 Гц (период T порядка 1 с). Из реально полученных спектров следует, что гармоники ЭКГ с частотами свыше 150–400 Гц имеют пренебрежимо малую амплитуду и при разложении ЭКГ в ряд Фурье [формула (6.27)] можно ограничиться последней составляющей с частотой 400 Гц. Это означает, что информация об электрической деятельности сердца заключена в частотном диапазоне от 0,5 Гц (минимально возможная частота сердечных сокращений) до 400 Гц (частота гармоники самого высокого порядка).

Применение гармонического анализа для обработки сложных периодических физиологических процессов позволяет с помощью электронной и вычислительной техники автоматизировать диагностику заболеваний и существенно расширить ее возможности.

6.7. Механические волны

Механическая волна представляет собой процесс распространения механических колебаний в упругой среде.

Из-за наличия упругих связей между частицами среды периодические смещения одной из частиц вызывают смещение соседних частиц с таким же периодом. Этот процесс распространяется в упругой среде с некоторой скоростью, определяемой только модулем упругости E и плотностью ρ среды:

$$v = \sqrt{\frac{E}{\rho}}. \qquad (6.28)$$

Волна называется *продольной*, если направление смещения частиц среды совпадает с направлением распространения волны. Если же эти направления взаимно перпендикулярны, то волна называется *поперечной*. Продольные механические волны могут распространяться в любых средах (кроме вакуума), а поперечные — только в твердых телах.

6.7.1. Уравнение волны

Рассмотренные ранее математические выражения, описывающие характер колебаний, определяют смещение как функцию одной переменной — времени. Смещение S частиц среды в волне зависит уже от двух переменных — времени t и пространственной координаты x — и обладает поэтому двойной периодичностью.

Для получения уравнения волны, т.е. аналитического выражения функции двух переменных $S = f(t, x)$, представим, что в некоторой точке M пространства возникают гармонические колебания с круговой частотой ω, амплитудой A и начальной фазой $\varphi = 0$:

$$S_м = A\sin\omega t.$$

Поскольку частицы среды упруго связаны между собой, то колебания от точки M распространяются вдоль оси x со скоростью υ и через некоторое время Δt достигают точки N (рис. 6.8).

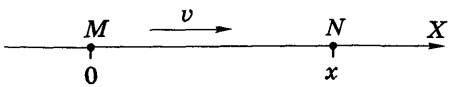


Рис. 6.8. Уравнение волны

Если в среде отсутствует затухание, то смещение в этой точке имеет вид $S_N = A\sin\omega(t - \Delta t)$, так как смещение здесь запаздывает на время Δt относительно точки M. Поскольку $\Delta t = \frac{MN}{\upsilon} = \frac{x}{\upsilon}$, то, подставив это в формулу для S_N, получим уравнение волны

$$S = A\sin\omega\left(t - \frac{x}{\upsilon}\right). \qquad (6.29)$$

Учитывая, что $\omega = \frac{2\pi}{T}$, где T — период, а длина волны $\lambda = \upsilon T$, формулу (6.29) можно записать в виде

$$S = A\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right). \qquad (6.30)$$

Уравнения волны (6.29) или (6.30) позволяют определить смещение любой точки с координатой x в любой момент времени t.

6.7.2. Поток энергии и интенсивность волны. Вектор Умова

Волновой процесс связан с распространением энергии E в пространстве. Количественной энергетической характеристикой этого процесса является ***поток энергии*** Φ — отношение энергии, перенесенной волной через некоторую поверхность, ко времени t, за которое этот перенос совершается. Если перенос энергии осуществляется равномерно, то $\Phi = E/t$. В более общем случае поток представляет производную от энергии по времени — $\Phi = dE/dt$. Единицей измерения потока энергии в СИ является ватт (1 Вт = 1 Дж/с).

Интенсивность (плотность потока энергии) волны I — отношение потока энергии Φ к площади S поверхности, расположенной перпендикулярно направлению распространения волны и представляющей собой поток энергии через единицу поверхности. При равномерном распределении энергии по этой поверхности $I = \Phi/S = E/St$; в общем случае $I = d\Phi/dS$. Измеряется интенсивность в Вт/м^2.

Представим участок среды, в которой распространяется волна, в виде параллелепипеда длиной l. Площадь грани параллелепипеда, перпендикулярной направлению скорости волны v, обозначим через S (рис. 6.9).

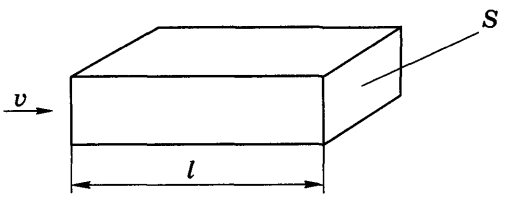


Рис. 6.9. Связь между интенсивностью волны и скоростью ее распространения

Введем объемную плотность энергии колебательного движения w, представляющую количество колебательной энергии в единице объема среды: $w = E/V$. За время t через площадку S пройдет энергия, равная произведению величины объема $V = lS = vtS$ и объемной плотности энергии:

$$E = wvtS. \qquad (6.31)$$

Разделив левую и правую части этой формулы на время и площадь, получим выражение, связывающее интенсивность волны и скорость ее распространения: $I = wv$. Вектор $\vec{I}$, модуль которого равен интенсивности волны, а направление совпадает с направлением ее распространения, носит название ***вектора Умова***:

$$\vec{I} = w\vec{v}\,. \qquad (6.32)$$

Формулу (6.32) можно представить в несколько ином виде. Учитывая, что энергия гармонических колебаний [см. формулу (6.9)] $E = \frac{1}{2}mA_0^2\omega_0^2$ и масса $m = \rho V$, где ρ — плотность

вещества; V — объем вещества, получим выражение для объемной плотности энергии волны:

$$w = \frac{1}{2}\rho A_0^2 \omega_0^2 . \qquad (6.33)$$

Теперь формула (6.32) принимает вид

$$\vec{I} = \frac{1}{2}\rho A_0^2 \omega_0^2 \vec{v} . \qquad (6.34)$$

Итак, интенсивность механической волны прямо пропорциональна скорости ее распространения, квадрату амплитуды колебаний частиц среды и квадрату частоты колебаний.

6.8. Эффект Доплера

Эффект Доплера состоит в ***изменении частоты волн***, воспринимаемых некоторым приемником (наблюдателем), в зависимости от относительной скорости движения источника и наблюдателя.

Когда источник и приемник неподвижны (рис. 6.10, *а*), то, естественно, частота волн, регистрируемых приемником, сов-

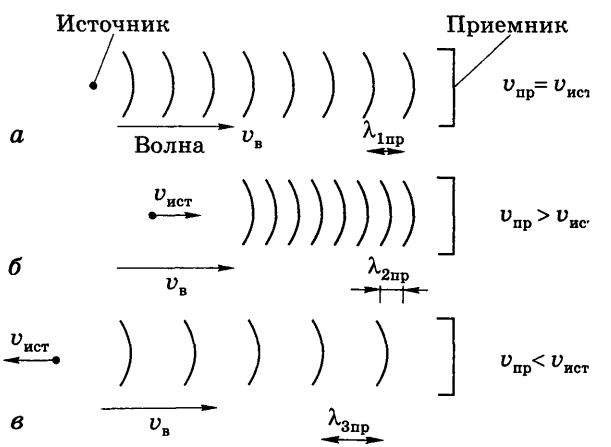


Рис. 6.10. Изменение длины λ принимаемой волны вследствие эффекта Доплера:
а — при $\upsilon_{пр} = \upsilon_{ист}$; *б* — источник приближается к приемнику; *в* — источник удаляется от приемника

падает с частотой волн, испускаемых источником: $\nu_{ист} = \nu_{пр}$. Если источник приближается к неподвижному приемнику с некоторой скоростью $\upsilon_{ист}$, то его движение вызывает «сжатие» волны, т.е. расстояния между максимумами (гребнями) волн уменьшаются (уменьшаются длина волны $\lambda_{пр}$ и ее период, регистрируемые приемником), а частота воспринимаемого волнового процесса увеличивается: $\nu_{пр} > \nu_{ист}$ (рис. 6.10, *б*). При удалении источника (рис. 6.10, *в*) все происходит наоборот.

В общем случае, когда источник и приемник движутся вдоль одной прямой со скоростями $\upsilon_{ист}$ и $\upsilon_{пр}$, частота, воспринимаемая приемником:

$$\nu_{пр} = \frac{\upsilon_в \pm \upsilon_{пр}}{\upsilon_в \mp \upsilon_{ист}} \nu_{ист}, \qquad (6.35)$$

где $\upsilon_в$ — скорость волны в среде.

Верхние знаки, т.е. «+» в числителе этой формулы и «–» в знаменателе, соответствуют сближению источника и приемника, а нижние знаки — их взаимному удалению.

Разность принимаемой и излучаемой частот $\Delta\nu = \nu_{пр} - \nu_{ист}$ называется *доплеровским сдвигом*, он положительный при сближении источника и приемника и отрицательный при их удалении, а по величине сдвига можно определить скорость их относительного движения.

Принципы использования эффекта Доплера для определения скорости движения крови. Представим кровеносный сосуд, на некотором участке которого необходимо определить скорость движения крови $\nu_{кр}$ (рис. 6.11). От источника ультразвука на сосуд направляется пучок ультразвуковых волн с частотой $\nu_{ист}$. Некоторый движущийся со скоростью $\nu_{кр}$ объем крови отражает ультразвуковые волны в различных направ-

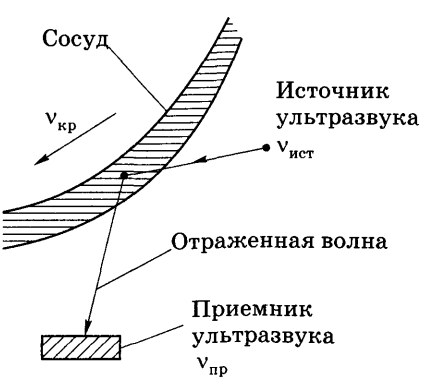


Рис. 6.11. Применение эффекта Доплера для определения скорости движения крови

лениях, в том числе и в направлении приемника ультразвука. Таким образом, этот объем крови становится движущимся вторичным источником ультразвука. Это значит, что в соответствии с эффектом Доплера частота волн, регистрируемых приемником $\nu_{пр}$, будет отличаться от частоты исходного ультразвука. Разница частот (доплеровский сдвиг) зависит от скорости крови:

$$\nu_{пр} - \nu_{ист} = \Delta\nu = f(\upsilon_{кр}).$$

По величине этого частотного сдвига можно определить линейную скорость крови, а по знаку — направление ее движения относительно приемника.

Данный метод является неинвазивным — измерения проводятся без нарушения целостности сосуда. Если площадь поперечного сечения считать известной, то рассмотренным методом можно определить и объемную скорость кровотока. Существенно, что при измерениях нет необходимости определять интенсивность отраженного кровью ультразвука, которая зависит от многих факторов, в том числе и совершенно случайных. Необходимо измерить лишь разницу частот $\Delta\nu$, что современная измерительная техника позволяет выполнить с высокой точностью.

Контрольные вопросы и задания

1. Запишите дифференциальное уравнение гармонических колебаний и выражение для смещения.
2. Запишите выражение для смещения при затухающем колебании. Что такое логарифмический декремент затухания? Как он связан с коэффициентом затухания?
3. От каких параметров зависит частота вынужденного колебания и каким образом?
4. В чем состоит явление резонанса и при каких условиях оно происходит?
5. В чем сущность теоремы Фурье? Что такое гармонический спектр периодического процесса?
6. Запишите уравнение механической волны, укажите смысл входящих в него величин.
7. От каких параметров зависит интенсивность волны и каким образом?
8. В чем состоит эффект Доплера? Как он используется для определения скорости кровотока?

Глава 7 Акустика

Акустика — раздел физики, исследующий упругие волны от самых низких частот до самых высоких (10^{12}–10^{13} Гц), в более узком смысле — учение о звуке.

Механические волны, имеющие частоты от самых низких до 10^{12}–10^{13} Гц, называют ***акустическими***.

В диапазоне частот от 16 Гц до 20 000 Гц воздействие акустических волн на слуховой аппарат человека приводит к формированию слухового ощущения. Поэтому механические волны этого диапазона называют *звуковыми*. При частотах, превышающих 20 кГц, инерционность звуковоспринимающего аппарата не позволяет сформировать слуховое ощущение — эту область частот называют *ультразвуком*. На частотах ниже 16 Гц ощущение звука также не возникает — волны столь низких частот называют *инфразвуком*.

Следует отметить, что частотный диапазон ощущения звука у разных организмов различен — например, дельфины и летучие мыши воспринимают звуки значительно большей частоты, чем человек. Кроме того, частотная граница воспринимаемых звуков зависит от возраста. В процессе старения человека инерционность его слухового аппарата увеличивается, и в пожилом возрасте верхняя граница звукового диапазона становится существенно меньшей 20 кГц.

7.1. Физические и физиологические характеристики звуковых волн

Физические характеристики акустических и, в частности, звуковых волн имеют объективный характер и могут быть измерены соответствующими приборами в стандартных единицах. Возникающее под действием звуковых волн слуховое ощущение субъективно, однако его особенности во многом определяются параметрами физического воздействия.

Скорость акустических волн υ определяется свойствами среды, в которой они распространяются — ее модулем упругости E и плотностью ρ:

$$\upsilon = \sqrt{E/\rho}. \qquad (7.1)$$

Скорость звука в воздухе составляет около 340 м/с и зависит от температуры (с изменением температуры изменяется плотность воздуха). В жидких средах и в мягких тканях организма эта скорость составляет около 1500 м/с, в твердых телах — 3000–6000 м/с.

В формулу (7.1), определяющую скорость распространения акустических волн, не входит их частота, поэтому звуковые волны различной частоты в одной и той же среде имеют практически одинаковую скорость. Исключение составляют волны таких частот, для которых характерно сильное поглощение в данной среде. Обычно эти частоты лежат за пределами звукового диапазона (ультразвук).

Если звуковые колебания представляют периодический процесс, то такие звуки называются ***тонами*** или музыкальными звуками. Они имеют дискретный гармонический спектр, представляя набор гармоник с определенными частотами и амплитудами. Первая гармоника частоты ω называется ***основным тоном***, а гармоники более высоких порядков (с частотами 2ω, 3ω, 4ω и т.д.) — ***обертонами***. ***Чистый*** (или простой) ***тон*** соответствует звуковым колебаниям, имеющим лишь одну частоту. На рис. 7.1 показан спектр сложного тона, в котором представлены четыре гармонических составляющих: 100, 200, 300 и 400 Гц. Величина амплитуды основного тона принята за 100 %.

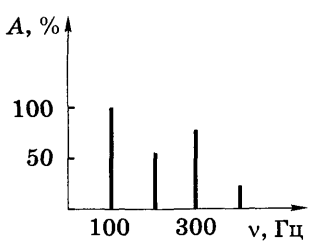


Рис. 7.1. Спектр сложного тона

Непериодические звуки, называемые ***шумами***, имеют сплошной акустический спектр (рис. 7.2). Они обусловлены процессами, в которых амплитуда и частота звуковых колеба-

ний изменяются со временем (вибрация деталей машин, шорох и т.п.).

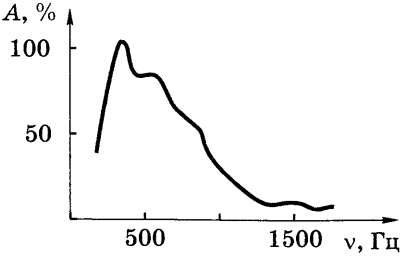


Рис. 7.2. Спектр шума

Интенсивность звука I, как уже отмечалось ранее, представляет собой энергию звуковой волны, приходящуюся на площадку единичной площади за единицу времени, и измеряется в Вт/м^2. Эта физическая характеристика определяет уровень слухового ощущения, который называется ***громкостью*** и является субъективным физиологическим параметром. Связь между интенсивностью и громкостью не является прямо пропорциональной. Пока отметим только, что с увеличением интенсивности возрастает и ощущение громкости. Количественную оценку громкости можно выполнить, сравнивая слуховые ощущения, обусловленные звуковыми волнами от источников с различной интенсивностью.

При распространении звука в среде возникает некоторое добавочное давление, перемещающееся от источника звука к приемнику. Величина этого ***звукового давления*** P также представляет физическую характеристику звука и среды его распространения. Она связана с интенсивностью I соотношением

$$P = \sqrt{2I\rho\upsilon}, \qquad (7.2)$$

где ρ — плотность среды; υ — скорость распространения звука в среде.

Величину $Z = \rho\upsilon$ называют ***удельным акустическим сопротивлением*** или ***удельным акустическим импедансом***.

Частота звуковых гармонических колебаний определяет ту сторону звукового ощущения, которую называют ***высотой звука***. Если звуковые колебания периодичны, но не подчиняются гармоническому закону, то высота звука оценивается ухом по частоте основного тона (первая гармоническая составляющая в ряду Фурье), период которого совпадает с периодом сложного звукового воздействия.

Отметим, что возможность оценки высоты тона слуховым аппаратом человека связана с продолжительностью звучания. Если время звукового воздействия меньше 1/20 с, то ухо не способно оценить высоту тона.

Близкие по частоте звуковые колебания при одновременном звучании воспринимаются как звуки различной высоты в том случае, если относительная разница частот превышает 2–3 %. При меньшей разности частот возникает ощущение слитного звука средней высоты.

Спектральный состав звуковых колебаний (см. рис. 7.1) определяется числом гармонических составляющих и соотношением их амплитуд и характеризует ***тембр*** звука. Тембр, как физиологическая характеристика слухового ощущения, в некоторой степени зависит также от скорости нарастания и изменчивости звука.

7.2. Область слышимости, порог слышимости и порог болевого ощущения

Слуховые ощущения формируются лишь в том случае, когда интенсивность звуковых волн превышает некоторое минимальное значение, называемое ***порогом слышимости***. Для различных частот звукового диапазона этот порог имеет различные значения, т.е. слуховой аппарат обладает *спектральной чувствительностью*.

Зависимость порога слышимости от частоты звука может быть определена с помощью несложного эксперимента. Представим, что имеется некоторый звуковой генератор — устройство, в котором создаются электрические колебания, превращаемые с помощью динамика в звуковые волны различных частот. Интенсивность и частоту звука можно плавно изменять и измерять. Установив определенное значение частоты звука, будем плавно увеличивать его интенсивность начиная с нулевого значения. Минимальное значение интенсивности, которое вызывает у испытуемого ощущение звука, представляет порог слышимости для данной частоты. Результаты этого эксперимента представим графически, откладывая по

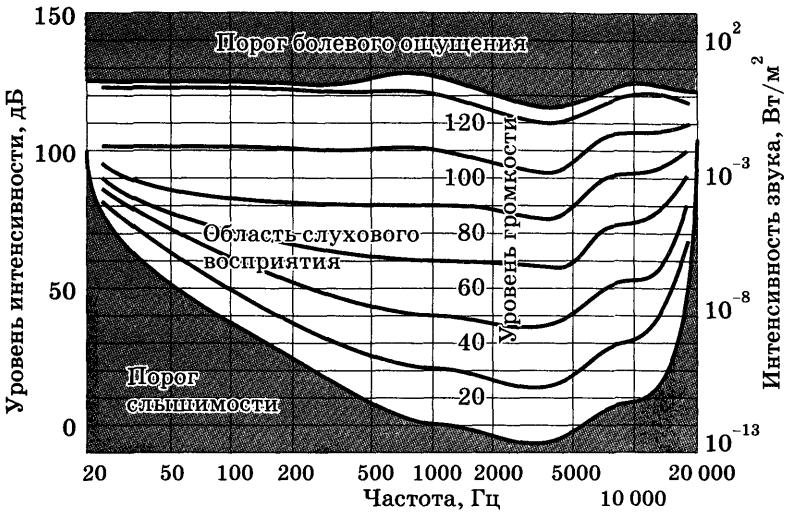


Рис. 7.3. Диаграмма слышимости для уха человека

оси абсцисс частоту, а по оси ординат — пороговое значение интенсивности звука (нижняя кривая на рис. 7.3).

Подобные эксперименты, проведенные на большом числе испытуемых, показывают, что ухо обладает максимальной чувствительностью в области частот 1000–3000 Гц. Здесь пороговое значение интенсивности звука минимально и составляет 10^{-12} Вт/м^2. Отметим, что при интенсивности 10^{-12} Вт/м^2 амплитуда колебаний частиц воздуха в сотни раз меньше диаметра молекул. Нормальный слуховой аппарат формирует слуховое ощущение при крайне слабых энергетических воздействиях.

С увеличением интенсивности звука возрастает и ощущение громкости. Однако звуковые волны с интенсивностью порядка 1–10 Вт/м^2 вызывают ощущение боли. Значение интенсивности, при превышении которого возникает боль, называется ***порогом болевого ощущения***. Как и порог слышимости, он тоже зависит от частоты звука (верхняя кривая на рис. 7.3), хотя и в меньшей степени. Область интенсивностей звука, заключенная между болевым порогом и порогом слышимости

(между верхней и нижней кривыми на рис. 7.3) и соответствующая частотному диапазону 16–20 000 Гц, называется ***областью слышимости.***

7.3. Закон Вебера — Фехнера. Уровни интенсивности и уровни громкости звука

Объективная физическая характеристика звуковой волны — интенсивность определяет ее субъективную физиологическую характеристику — громкость. Количественная связь между ними устанавливается на основе ***закона Вебера — Фехнера***, связывающего степень ощущения и интенсивность вызвавшего его раздражителя: ощущение растет в арифметической прогрессии, если интенсивность раздражителя увеличивается в геометрической прогрессии.

Другими словами, физиологическая реакция (в рассматриваемом случае громкость) на раздражитель (интенсивность звука) не прямо пропорциональна интенсивности раздражителя, а возрастает с ее увеличением существенно слабее — пропорционально логарифму интенсивности раздражителя.

Для пояснения физиологической значимости этого закона вспомним график логарифмической функции $y = k\lg I$ (рис. 7.4).

При малых аргументах I функция $y = k\lg I$ растет довольно быстро с увеличением аргумента. Это означает, что небольшое увеличение малой интенсивности (на величину ΔI) при-



Рис. 7.4. Зависимость между громкостью y звука и его интенсивностью I

водит к значительному увеличению громкости (на величину Δy_1) — как только интенсивность звука немного превысила пороговое значение, уже возникает слуховое ощущение. Если же интенсивность звука велика, то ее дальнейшее увеличение на ту же величину ΔI дает малый прирост громкости (на величину Δy_2) — при большой интенсивности звук хорошо слышен, и дальнейшее увеличение интенсивности звука на возрастание ощущения громкости сказывается существенно слабее.

Для установления количественной связи между интенсивностью и громкостью звука введем ***уровень интенсивности звука*** L — величину, пропорциональную десятичному логарифму отношения интенсивности звука I к интенсивности на пороге слышимости $I_0 = 10^{-12}$ Вт/м^2:

$$L = n \lg \frac{I}{I_0}. \qquad (7.3)$$

Коэффициент n в формуле (7.3) определяет единицу измерения уровня интенсивности звука. Обычно принимают $n = 10$, тогда величина L измеряется в децибелах (дБ). На пороге слышимости ($I = I_0$) уровень интенсивности звука $L = 0$, а на пороге болевого ощущения ($I = 10$ Вт/м^2) $L = 130$ дБ. Если, например, интенсивность звука составляет 10^{-7} Вт/м^2, что соответствует нормальному разговору, то из формулы (7.3) следует, что уровень его интенсивности составляет 50 дБ.

Уровень громкости звука (***громкость***) E связан с уровнем интенсивности L соотношением

$$E = kL, \qquad (7.4)$$

где k — некоторый коэффициент пропорциональности, зависящий от частоты и интенсивности звука.

Если бы пороговое значение интенсивности воспринимаемого звука не зависело от его частоты, то коэффициент k в формуле (7.4) был бы постоянным, уровень громкости совпадал бы с уровнем интенсивности и измерялся в децибелах.

Однако уровень громкости также изменяется с частотой. Например, звук с уровнем интенсивности 20 дБ и частотой 1000 Гц будет восприниматься существенно более громким, чем звук с тем же уровнем интенсивности, но частотой 100 Гц. Одинаковый уровень громкости на этих частотах будет достигнут, если для 1000 Гц уровень интенсивности составляет 20 дБ, а для 100 Гц — 50 дБ. По этим причинам для измерения уровня громкости вводится особая единица, называемая *фоном*.

На частоте 1000 Гц принимают $k = 1$, т.е. уровень интенсивности в децибелах и уровень громкости в фонах совпадают. На других частотах в области слышимости для перехода от децибел к фонам необходимо вводить соответствующие поправки. Этот переход можно осуществить с помощью кривых равной громкости (см. рис. 7.3), полученных на основании физиологических экспериментов. Например, пусть для частоты 200 Гц уровень интенсивности звука составляет 40 дБ. Чему равен уровень громкости этого звука? На рис. 7.3 находим точку с координатами 200 Гц и 40 дБ. Она лежит на кривой, соответствующей уровню громкости 20 фон, следовательно, для данной частоты 20 фон соответствует 40 дБ.

Определение порога слышимости на разных частотах составляет основу ***аудиометрии*** — методов измерения остроты слуха. Сравнивая аудиограммы (кривые, аналогичные представленным на рис. 7.3), полученные для конкретных пациентов, с усредненной нормой, судят о характере и степени развития нарушений слухового аппарата.

7.4. Отражение и поглощение акустических волн

Проходя границу раздела двух сред, акустические волны претерпевают отражение и преломление по законам, аналогичным законам отражения и преломления света (рис. 7.5). ***Коэффициент отражения*** R определяется отношением интенсивности отраженной акустической волны $I_{отр}$ к интенсивности волны, падающей на границу раздела $I_{пад}$: $R = I_{отр}/I_{пад}$.

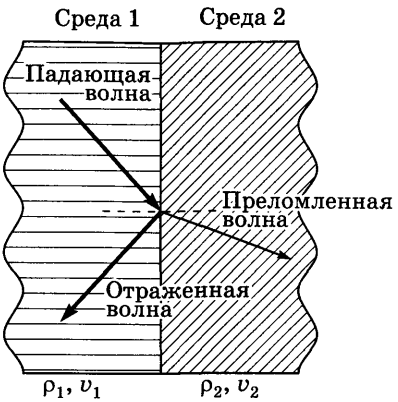


Рис. 7.5. Отражение и преломление акустических волн

Его величина зависит от соотношения плотностей сред (ρ_1 и ρ_2) и скоростей распространения акустических волн в этих средах (v_1 и v_2). Произведение плотности среды и скорости распространения в ней акустических волн называют ***удельным акустическим импедансом*** или ***волновым сопротивлением*** $Z = \rho v$. Так как $v = \sqrt{E/\rho}$ [см. формулу (7.1)], то $Z = \sqrt{E\rho}$.

При нормальном падении коэффициент отражения

$$R = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2, \text{ или } R = \left(\frac{\rho_2 v_2 - \rho_1 v_1}{\rho_2 v_2 + \rho_1 v_1}\right)^2. \qquad (7.5)$$

Из формулы (7.5) видно, что чем больше различаются акустические импедансы сред, тем выше коэффициент отражения. Так, например, на границе раздела воздух — вода отражается свыше 99 % энергии акустической волны и лишь весьма малая часть ее проникает в воду при падении из воздуха. Отражение на границах раздела различных биологических тканей существенно меньше. В частности, при переходе ультразвуковой волны из мышцы в кость коэффициент отражения составляет 30–40 %. Очевидно, что ***коэффициент прохождения*** $D = 1 - R$.

Итак, при падении акустических волн с интенсивностью $I_{\text{пад}}$ на границу раздела двух сред возникают отраженные вол-

ны с интенсивностью $I_{отр} = RI_{пад}$ и волны, входящие во вторую среду, с интенсивностью $I_{вх} = DI_{пад}$.

При распространении акустических волн в среде происходит их поглощение, причем интенсивность $I_{пр}$ волны с увеличением проходимого в среде расстояния x уменьшается по *экспоненциальному закону:*

$$I_{пр} = I_{вх}e^{-kx}, \qquad (7.6)$$

где k — показатель поглощения, зависящий от свойств среды и частоты волны.

Графически эта зависимость представлена на рис. 7.6.

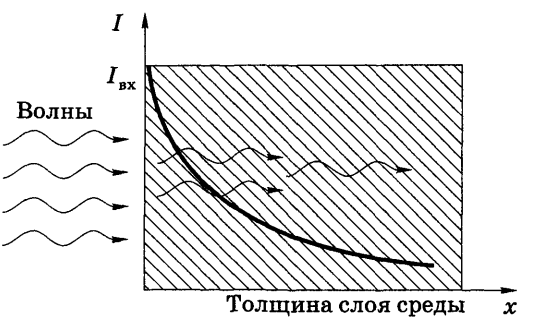


Рис. 7.6. Зависимость интенсивности I акустической волны, прошедшей через среду, от толщины x слоя среды.

Поглощение акустических волн в среде сильно зависит от частоты волны. С увеличением частоты показатель поглощения k увеличивается во всех средах и высокочастотные акустические волны поглощаются существенно сильнее, чем низкочастотные.

7.5. Ультразвук и его медицинское применение. Инфразвук

Механические волны, частоты которых заключены в диапазоне от 20 кГц до $10^9 - 10^{10}$ Гц, называют ***ультразвуком.*** Естественных источников ультразвука в природе практически нет, за исключением некоторых животных (летучие

мыши, дельфины и др.). Поэтому для практического использования необходимо уметь генерировать и регистрировать ультразвуковые волны разных частот.

7.5.1. Получение ультразвука

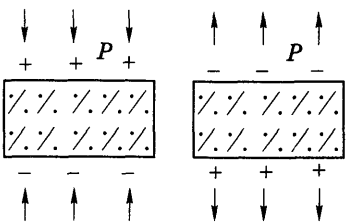


Рис. 7.7. Прямой пьезоэффект

Наиболее распространенным способом получения и регистрации ультразвука является использование обратного и прямого пьезоэффекта. Напомним, что ***прямой пьезоэффект*** заключается в возникновении электрических зарядов на поверхности некоторых кристаллических тел (например, кварца) под действием механического сжатия или растяжения, что сопровождается возникновением электрического напряжения между этими поверхностями (рис. 7.7). При изменении направления деформации под усилием P (сжатие заменяют растяжением) знаки зарядов и электрического напряжения изменяются на противоположные.

Прямой пьезоэффект широко используется в технике, в том числе и в медицинской аппаратуре. На его основе создаются измерительные приборы (датчики), преобразующие давления, вибрации, звуковые колебания и другие механические явления в электрические сигналы, удобные для дальнейшей регистрации. Прямой пьезоэффект используется и для регистрации ультразвука.

Пьезоэффект обратим. Если на упомянутую кварцевую пластинку воздействовать переменным электрическим полем, создаваемым некоторым генератором (рис. 7.8), то возникают механические деформации пластинки — она сжимается и растягивается с частотой действующего электрического поля — возни-

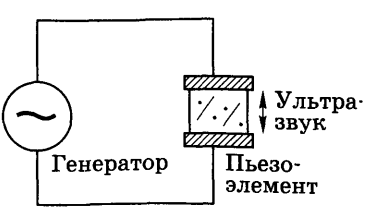


Рис. 7.8. Обратный пьезоэффект

кает ***обратный пьезоэффект, т.е. электрострикция.*** При изменении частоты генерируемых электрических колебаний соответственно изменяется и частота ультразвуковых колебаний.

Для получения ультразвука используют также ***магнитострикцию*** — механическую деформацию тел под действием переменного магнитного поля.

7.5.2. Физические принципы ультразвуковой диагностики

Получение видимых изображений внутренних органов с помощью ультразвука основано на регистрации ультразвуковых волн, отраженных на границах раздела морфологических структур, и последующей обработке полученных электрических сигналов. Методы технической реализации этого принципа достаточно разнообразны.

Наиболее простым является ***эхолокация***, или ***A-метод*** (от англ. amplitude — широта, диапазон, размах). Это одномерный метод исследования, так как он позволяет определять расстояние до отражающих ультразвук объектов лишь в одном заданном направлении (рис. 7.9).

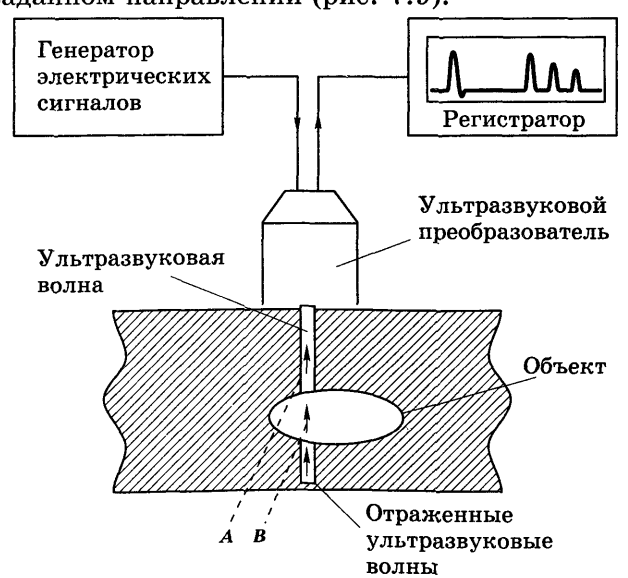


Рис. 7.9. Принципиальная схема эхолокации в *A*-режиме

На поверхности исследуемого объекта (поверхность тела человека) устанавливается ультразвуковой преобразователь, который служит как источником, так и приемником ультразвука.

Источник излучает короткий ультразвуковой импульс, который распространяется в теле человека со скоростью v. На границах раздела различных морфологических структур происходит отражение ультразвука. Приемник воспринимает отраженные ультразвуковые волны и за счет прямого пьезоэффекта преобразует их в электрические сигналы, которые затем усиливаются с помощью усилителя и регистрируются на экране монитора. Отмечаемое на экране время Δt_1 прихода первого отраженного сигнала (рис. 7.10) соответствует времени распространения ультразвука от источника до границы раздела A морфологических структур и обратно (см. рис. 7.9). Значит, глубина расположения этой границы d_1 определяется по формуле $d_1 = v\Delta t_1/2$.

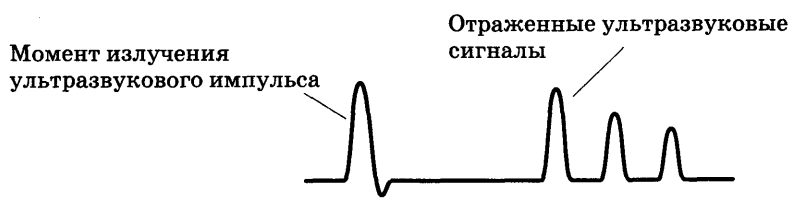


Рис. 7.10. Импульсные сигналы

Аналогично определяются расстояния d_2 (до границы раздела B), d_3 и т.д. Таким образом, полученная эхограмма позволяет судить о глубине расположения границ различных морфологических структур по расстоянию между отметкой начального зондирующего импульса и положением отраженных сигналов.

Эхолокацией, например, определяются продольные размеры глазного яблока и структур глаза (эхоокулометрия), измеряется разница расстояний от поверхности головы до ее внутричерепных структур и выявляются объемные поражения мозга (опухоли, гематомы, инородные тела) (эхоэнцефалография).

Амплитуды зарегистрированных на экране монитора сигналов пропорциональны интенсивности ультразвуковых волн, отраженных границами раздела сред, и определяются разницей в их акустических импедансах [см. формулу (7.5)], а также некоторым поглощением в среде.

Ультразвуковая томография, или ***B-метод*** (от англ. bright — яркость), позволяет получать изображение различных сечений исследуемого органа. При ее реализации ультразвуковой луч перемещается (сканирует) в определенной плоскости, проходящей через исследуемый орган. Отраженные от границ раздела морфологических структур УЗ-волны преобразуются в электрические сигналы, последующая обработка которых позволяет получить на экране монитора изображение органа в выбранном сечении.

M-метод (от англ. motion — движение) является, в сущности, разновидностью уже рассмотренного *A*-метода, но с разверткой по времени. Он используется для оценки характеристик движения некоторых объектов (например, клапанов сердца).

В современных ультразвуковых диагностических устройствах упомянутые методы обычно технически реализованы в едином приборном комплексе.

Частота ультразвуковых волн, обычно используемых для диагностики, составляет 0,5–15 МГц. Чем больше частота, тем выше разрешающая способность метода, т.е. тем более мелкие объекты могут быть рассмотрены и тем выше качество изображения. Это объясняется тем, что минимальный размер объекта d, изображение которого может быть получено, определяется длиной волны λ того излучения, в котором формируется изображение данного объекта. Если $d < \lambda$, то явление дифракции не позволяет получить четкое изображение объекта. Например, при частоте ультразвука $\nu = 1$ МГц и скорости распространения в мышечной ткани $v = 1500$ м/с длина волны составляет

$$\lambda = \frac{v}{\nu} = \frac{1{,}5 \cdot 10^3}{10^6} = 1{,}5 \cdot 10^{-3}\ \text{м} = 1{,}5\ \text{мм}. \qquad (7.7)$$

Это значит, что минимальный размер объекта, изображение которого может быть получено с помощью ультразвука такой частоты, составляет около 1,5 мм. При повышении частоты уменьшается длина волны ультразвука и увеличивается разрешающая способность УЗ-методов.

Однако следует учесть, что с повышением частоты ультразвука увеличивается его поглощение в среде. Так, интенсивность ультразвука частотой 0,8 МГц в мягких тканях уменьшается вдвое на глубине 4–5 см, а для частоты 3 МГц — на глубине 1,5–2 см. Поэтому исследования глубоко расположенных внутренних органов (например, печени) вынуждены проводить на более низких частотах, поскольку высокочастотный ультразвук практически полностью поглотится окружающими орган тканями и отраженный от исследуемого объекта ультразвуковой сигнал будет невозможно зарегистрировать. Если же исследуются органы, расположенные на малом расстоянии от поверхности тела (например, щитовидная железа), то поглощение окружающими тканями сказывается существенно меньше и появляется возможность использовать для диагностики ультразвук более высоких частот.

Ультразвуковая диагностика получила широкое распространение в клинике благодаря высокой разрешающей способности при визуализации исследуемых объектов, возможности проведения многократных исследований, безопасности и в связи с этим отсутствию каких-либо противопоказаний.

Методы ультразвуковой диагностики имеют и свои ограничения в применении. Они обусловлены тем, что на границе раздела мягкие ткани — газ или газ — жидкость происходит практически полное отражение ультразвука. Мощный ультразвуковой сигнал, возникающий за счет отражения на этих границах раздела, делает невозможным исследование структур, которые расположены за ними. Поэтому использование ультразвуковых методов неэффективно для диагностики состояния газонаполненных органов (например, легкие, кишечник).

7.5.3. Взаимодействие ультразвука с биологическими тканями. Терапевтическое и хирургическое применение ультразвука

Ультразвук при распространении в биологических тканях порождает, как и любая механическая волна, локальные периодические перепады давлений в зоне воздействия. При небольших интенсивностях ультразвука эти механические эффекты проявляются в виде своеобразного *микромассажа* тканей и микровибраций на клеточном уровне. При этом возникает и нагрев тканей за счет превращения поглощенной УЗ энергии в тепловую — проявляются *тепловые эффекты.* При высоких интенсивностях ультразвуковых волн может происходить разрушение биомакромолекул и повреждение клеточных мембран. Резкие перепады давлений могут приводить к возникновению в среде кратковременно существующих разрывов (кавитаций). Возможна также *ионизация* и *диссоциация* молекул и целый ряд других первичных физико-химических эффектов.

В ***физиотерапии*** используют ультразвук с частотами 0,8–3 МГц и интенсивностью до 1 Вт/см^2. От частоты ультразвука зависит глубина его проникновения в ткань (чем выше частота, тем меньше глубина проникновения), а при указанной интенсивности в тканях не возникает необратимых разрушений. Происходящие под действием ультразвука микромассаж, локальный нагрев тканей на доли и единицы градуса и другие первичные эффекты повышают интенсивность обмена веществ, стимулируют процессы тканевого дыхания, способствуют улучшению снабжения тканей кровью и лимфой.

В физиотерапии широко применяется метод *фонофореза лекарственных веществ,* где используется сочетательное воздействие на живую биологическую ткань двух факторов: физического (ультразвук) и химического (лекарственные препараты). Лекарственное вещество вносится между поверхностью тела и головкой излучателя. Под действием ультразвука оно проникает в эпидермис, откуда диффундирует в кровь

и лимфу и разносится по всему организму. В этом лечебном методе ультразвук не только выполняет транспортную функцию (введение лекарства через кожу), но и изменяет фармакокинетику.

Возможность фокусировки ультразвуковых волн на весьма малую поверхность (из-за малой длины волны) и получения высокой плотности энергии (энергия пропорциональна квадрату частоты) позволяет использовать ультразвук в ***хирургии*** для рассечения и соединения биологических тканей. Применение ультразвука для разрушения тканей может осуществляться двумя способами. Первый — воздействие на мягкие и костные ткани собственно фокусированным ультразвуком с частотой 1–10 МГц. Второй — наложение ультразвуковых колебаний с частотами 20–75 кГц и амплитудой 10–50 мкм на хирургический инструмент.

При ультразвуковой хирургии уменьшаются кровопотери и болевые ощущения, уменьшается усилие резания. Возможно применение ультразвука для разрушения тромбов в кровеносных сосудах и очищения сосудистых стенок от атероматозных масс.

Ультразвуковые методы применяются для сварки костей (ультразвуковой остеосинтез). При этом промежуток между костными обломками заполняют жидкой пластмассой (например, циакрином), смешанной с костной щебенкой. Под действием ультразвука происходит быстрая диффузия мономера в свариваемые костные участки и его полимеризация, обеспечивающая соединение костей.

После ультразвуковой сварки мягких тканей сварочный шов не препятствует процессам регенерации и рубцевания. Поэтому ультразвуковую сварку применяют при пластике кожи, мышц, повреждениях печени, почек, легких и других органов, а также при операциях на них.

Приведенные примеры иллюстрируют лишь некоторые медицинские применения ультразвука. В последнее время области и методики его использования быстро расширяются и совершенствуются.

7.5.4. Инфразвук

Инфразвук — это механические волны с частотами менее 16 Гц.

Поскольку показатель поглощения акустических волн с уменьшением частоты уменьшается, инфразвук обладает большой проникающей способностью и распространяется на большие расстояния. Инфразвук возникает при работе промышленных установок, транспортных средств и часто сопровождается слышимым шумом, что вызывает определенные трудности в выделении из совокупности акустических волн собственно инфразвука.

Отмечено, что воздействие инфразвуком вызывает у человека и животных комплекс нежелательных ощущений. У людей может наблюдаться головокружение, затруднение дыхания, боли в животе, чувство страха и др.

Первичные механизмы воздействия инфразвука на биологические объекты изучены недостаточно. Предполагается, что в основе этого воздействия лежат резонансные эффекты, так как частоты собственных колебаний тела человека и его отдельных частей и органов соответствуют частотному диапазону инфразвука. Это и создает условия для возникновения резонансных эффектов, определяющих, по-видимому, действие инфразвука на организм.

Контрольные вопросы и задания

1. Какие волны называют звуковыми? От чего зависит скорость звука?
2. Какими физическими характеристиками определяются физиологические характеристики звука?
3. Что называют порогом слышимости? От чего он зависит?
4. Сформулируйте Закон Вебера–Фехнера. Какова связь между уровнями интенсивности и громкости звука? В каких единицах они измеряются?
5. Что такое акустический импеданс? От чего и как зависит коэффициент отражения акустических волн на границе раздела сред?

6. Запишите закон поглощения акустических волн в среде. Как зависит показатель поглощения от частоты волны?
7. Назовите методы получения и регистрации ультразвука. Объясните их физическую основу.
8. В чем сущность методов ультразвуковой диагностики: эхолокации и томографии?
9. Почему при УЗИ органов различной локализации применяют различные частоты ульразвука?
10. Почему отличается фармакотерапевтическое действие одних и тех же лекарств, вводимых с помощью инъекций и фонофореза?

Глава 8 Свойства жидкостей. Поверхностные явления

По физическим свойствам жидкости занимают промежуточное положение между газами и твердыми телами. Молекулы жидкости достаточно плотно упакованы, поэтому жидкости практически не сжимаемы и их плотность близка к плотности твердых тел. Между молекулами жидкости действуют значительные силы взаимодействия, несколько меньшие, чем у твердых тел. Жидкости, подобно твердым телам, сохраняют свой объем, однако, как и газы, принимают форму сосуда, в котором находятся.

В жидкостях имеется «ближний порядок» в расположении соседних молекул, но поддерживается он лишь на расстояниях в несколько молекулярных диаметров, а «дальнего порядка», в отличие от твердых тел, в них нет.

Молекулы жидкости совершают непрерывные беспорядочные поступательные движения, энергия которых прямо пропорциональна ее температуре.

8.1. Поверхностная энергия

Наиболее характерным свойством жидкости является наличие свободной поверхности на границе с газом, что приводит к возникновению явлений особого рода, называемых ***поверхностными***. Они обусловлены особыми физическими условиями, в которых находятся молекулы свободной поверхности жидкости. На каждую молекулу жидкости действуют силы притяжения со стороны окружающих ее молекул, расположенных от нее на расстоянии нескольких нанометров.

Рассмотрим силы, действующие на молекулу внутри объема жидкости и на ее поверхности. На молекулу M_2, расположенную внутри жидкости (рис. 8.1), действуют примерно одинаковые силы со стороны окружающих ее молекул, поэтому их равнодействующая близка к нулю.

Для молекулы M_1, расположенной на поверхности жидкости, силы притяжения со стороны молекул газа очень малы по сравнению с силами притяжения молекул жидкости, поэтому равнодействующая F_1 этих сил отлична от нуля и направлена внутрь жидкости, перпендикулярно ее поверхности. Вследствие этого молекулы поверхностного слоя втягиваются внутрь объема жидкости и поверхностный слой оказывает молекулярное давление на жидкость.

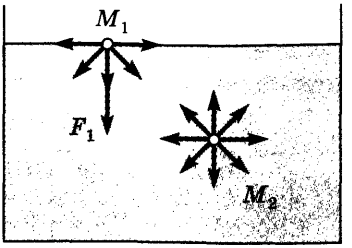


Рис. 8.1. Молекулярные силы, действующие на молекулу M_1 поверхностного слоя и молекулу M_2 внутри объема

Для того чтобы переместить молекулу изнутри объема на поверхность жидкости, необходимо совершить работу против сил молекулярного давления. Аналогичную работу надо совершить и при перемещении молекулы поверхностного слоя в газовую фазу. Следовательно, каждая молекула на поверхности жидкости обладает дополнительной потенциальной энергией по сравнению с молекулами внутри жидкости. Эту потенциальную энергию и называют ***поверхностной энергией.***

Очевидно, что величина поверхностной энергии $W_п$ прямо пропорциональна количеству молекул в поверхностном слое и, следовательно, прямо пропорциональна площади S свободной поверхности жидкости

$$W_п = \sigma S. \qquad (8.1)$$

Коэффициент пропорциональности σ зависит только от свойств жидкости и температуры и называется ***коэффициентом поверхностного натяжения.*** Единицей его измерения в СИ является Дж/м^2 или Н/м.

Поскольку любая физическая система стремится к состоянию с наименьшей возможной в заданных условиях потенциальной энергией, свободная поверхность жидкости стремится принять наименьшую возможную площадь. По этой причине

капля жидкости в состоянии невесомости принимает сферическую форму, что соответствует минимальной поверхности при заданном объеме жидкости.

8.2. Силы поверхностного натяжения

Способность поверхности жидкости сокращаться обусловлена существованием молекулярных сил, стремящихся сократить эту поверхность. Так, на молекулу M_1 поверхностного слоя (см. рис. 8.1) действуют не только молекулы, находящиеся внутри жидкости, но и молекулы, находящиеся на поверхности жидкости и стремящиеся как бы «разорвать» поверхность в точке M_1. Если по поверхности жидкости провести воображаемую линию, то равнодействующая молекулярных сил, действующих на молекулы этой линии по одну из сторон, будет направлена по касательной к поверхности и одновременно перпендикулярна линии возможного разрыва. Эта равнодействующая называется ***силой поверхностного натяжения*** $F_п$, поскольку именно она и стремится сократить поверхность жидкости. Очевидно, что эта сила пропорциональна числу молекул в линии разрыва и, следовательно, ее длине l:

$$F_п = \sigma l, \tag{8.2}$$

где σ — коэффициент поверхностного натяжения, выраженный в Н/м и численно равный силе поверхностного натяжения, действующей на единицу длины рассматриваемой линии.

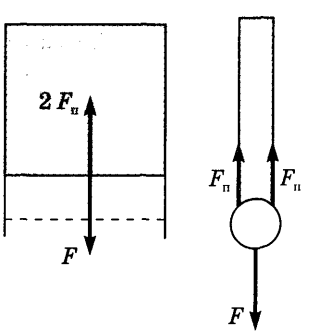


Рис. 8.2. Измерение силы поверхностного натяжения

Эту силу можно измерить, если ограничить поверхность жидкости рамкой, одна из сторон которой может свободно перемещаться (рис. 8.2). На эту подвижную сторону рамки будут действовать силы поверхностного натяжения $2F_п$, так как поверхностей две, стремящи-

еся сократить поверхность. Для их уравновешивания необходимо приложить силу F, направленную противоположно: $F = 2F_{п}$. Измерив силу F, нетрудно вычислить $F_{п}$, а затем и коэффициент поверхностного натяжения σ.

Коэффициент поверхностного натяжения зависит только от природы жидкости и ее температуры. Вдали от критической температуры (при которой исчезает различие между жидкостью и ее насыщенным паром) его значение убывает линейно с увеличением температуры, а при критической температуре обращается в нуль.

Под действием поверхностного натяжения небольшие количества жидких сред стремятся принять шарообразную форму, соответствующую наименьшей величине поверхности. Приближение к шаровой форме достигается тем больше, чем слабее силы тяжести, поскольку у малых капель сила поверхностного натяжения значительно превосходит силу тяжести.

Поверхностное натяжение является важнейшей характеристикой поверхности раздела фаз. Оно непосредственно влияет на образование мелкодисперсных частиц твердых тел и жидкостей при их распылении, а также на слияние капель или пузырьков в эмульсиях, туманах, пенах, на процессы адгезии.

Поверхностное натяжение определяет форму биологических клеток и их частей. Изменение сил поверхностного натяжения влияет на фагоцитоз (захват клетками соседних частиц), пиноцитоз (захват клеточной поверхностью жидкости с содержащимися в ней веществами), на процессы альвеолярного дыхания.

Поверхностное натяжение различных веществ колеблется от 0,01 до 2,0 Н/м, у клеток оно не превышает 0,10 Н/м. Вещества, растворенные в жидкости, способны как понижать, так и повышать поверхностное натяжение. Вещества, понижающие поверхностное натяжение раствора, называются ***поверхностно-активными веществами*** (ПАВ). К ним относится, в частности, сурфактант, снижающий поверхностное натяжение альвеолярных стенок, обеспечивая тем самым возможность дыхания.

Значения коэффициента поверхностного натяжения некоторых жидкостей при температуре 20 °С приведены в табл. 8.1.

Таблица 8.1

Жидкость	σ, Н/м	Жидкость	σ, Н/м
Вода	0,0725	Ртуть	0,47
Спирт	0,022	Желчь	0,048
Сыворотка крови	0,056	Молоко	0,050
Эфир	0,017	Моча	0,066

Измерение коэффициента поверхностного натяжения имеет диагностическое значение в клинике. Например, в норме для мочи человека он равен 66 мН/м, а при появлении в моче желчных пигментов снижается до 56 мН/м.

8.3. Добавочное давление под изогнутой поверхностью жидкости. Формула Лапласа

Молекулы поверхностного слоя втягиваются внутрь объема жидкости, вследствие чего поверхностный слой оказывает молекулярное давление на жидкость. Если поверхность жидкости плоская, то все силы F_1, действующие на каждую молекулу M_1 поверхностного слоя, параллельны друг другу и оказывают на поверхность жидкости молекулярное давление p_0 (рис. 8.3). При этом вклад сил поверхностного натяжения в это молекулярное давление равен нулю, так как на

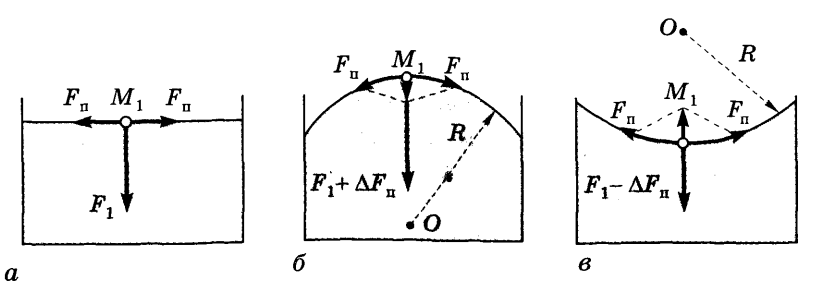


Рис. 8.3. Силы, создающие молекулярное давление под поверхностями: *а* — плоской; *б* — выпуклой; *в* — вогнутой

плоской поверхности эти силы направлены противоположно и компенсируют друг друга.

Если же поверхность жидкости по каким-либо причинам искривлена, то силы поверхностного натяжения, действующие на молекулы M_1, направлены под углом и уже не компенсируют друг друга, а их результирующая направлена к центру кривизны поверхности и оказывает на поверхность дополнительное молекулярное давление Δp, величина которого рассчитывается по ***формуле Лапласа***

$$\Delta p = \sigma\left(\frac{1}{R_1} + \frac{1}{R_2}\right), \tag{8.3}$$

где R_1 и R_2 — радиусы кривизны поверхности жидкости в двух взаимно перпендикулярных сечениях.

Если поверхность жидкости сферическая, то $R_1 = R_2 = R$ и добавочное давление

$$\Delta p = \frac{2\sigma}{R}. \tag{8.3а}$$

Поскольку силы, создающие дополнительное молекулярное давление, направлены всегда к центру кривизны поверхности, то и дополнительному давлению Δp приписывают такую же направленность. В результате молекулярное давление под выпуклой поверхностью жидкости всегда больше, а под вогнутой — меньше, чем под плоской поверхностью:

$$\left.\begin{aligned} p_{\text{вып}} &= p_0 + \Delta p, \\ p_{\text{вогн}} &= p_0 - \Delta p. \end{aligned}\right\} \tag{8.4}$$

Таким образом, молекулярное давление под поверхностью жидкости существенно зависит от кривизны этой поверхности, что наглядно проявляется, в частности, в капиллярных явлениях.

8.4. Явление смачивания

На границе раздела жидкости с твердым телом возникают явления смачивания или несмачивания, обусловленные взаимодействиями молекул жидкости с молекулами твердого тела. Если силы притяжения между молекулами жидкости

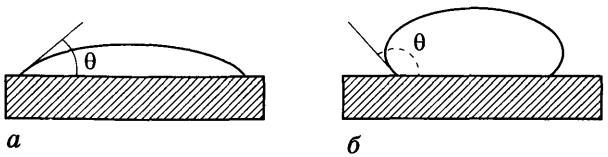


Рис. 8.4. Взаимодействие жидкости с поверхностью твердого тела: *а* — смачивание, краевой угол $\theta < 90°$; *б* — несмачивание, $\theta > 90°$

и твердого тела больше сил притяжения между молекулами жидкости: $F_{ж\text{-}тв} > F_{ж\text{-}ж}$, то жидкость будет растекаться по поверхности твердого тела. Это явление называют ***смачиванием***. Если же силы притяжения между молекулами жидкости и твердого тела меньше сил притяжения между молекулами жидкости: $F_{ж\text{-}тв} < F_{ж\text{-}ж}$, то жидкость не будет растекаться по поверхности твердого тела, а будет стремиться к уменьшению своей поверхности, собираясь в каплю (рис. 8.4). Такое явление называют ***несмачиванием***. Поскольку эти явления определяются относительными свойствам твердого тела и жидкости, то одна и та же жидкость может быть смачивающей для одного твердого тела и несмачивающей для другого. Так, вода смачивает стекло и не смачивает жирную поверхность, а ртуть не смачивает стекло, но хорошо смачивает медь. Поверхности, смачиваемые водой, называют *гидрофильными*, а несмачиваемые ею — *гидрофобными*.

Количественной мерой смачивания служит краевой угол θ, образуемый поверхностью твердого тела и касательной, проведенной к поверхности жидкости в точке их соприкосновения, причем жидкость находится внутри этого угла (см. рис. 8.4). При смачивании $0 \le \theta < 90°$; чем меньше угол θ, тем сильнее смачивание. Если $\theta = 0$, то смачивание называют полным или идеальным. При *идеальном смачивании* капля жидкости растекается по поверхности твердого тела до тех пор, пока не покроет всю поверхность или пока не образуется мономолекулярный слой. К случаю идеального смачивания можно приближенно отнести растекание спирта или воды по чистой поверхности стекла, нефти по воде и пр.

При несмачивании $90° < \theta \leq 180°$ и чем больше угол θ, тем сильнее оно выражено, при $\theta = 180°$ будет *полное несмачивание*. В этом случае жидкость не прилипает к поверхности твердого тела и легко скатывается с нее.

Явление смачивания имеет важное значение и в быту, и в промышленности, и в медицине. Хорошее смачивание необходимо при крашении, склеивании, пайке и других производственных процессах. Несмачивание важно при гидроизоляции, создании непромокаемых материалов. Явления смачивания важны также для продвижения крови по капиллярам, осуществления дыхания и других биологических процессов.

8.5. Капиллярные явления

Смачивание или несмачивание жидкостью стенок сосуда, в котором она находится, влияет на форму свободной поверхности жидкости в этом сосуде. Если размеры поперечного сечения сосуда велики, то форма поверхности определяется главным образом силой тяжести и представляет в этом случае плоскую горизонтальную поверхность и лишь у самых стенок поверхность жидкости искривляется вследствие смачивания или несмачивания. Однако в сосудах малого поперечного сечения (капиллярах) вследствие смачивания или несмачивания их стенок жидкостью происходит существенное искривление поверхности жидкости и образуется так называемый мениск, вогнутый при смачивании жидкостью стенок сосуда и выпуклый при несмачивании, что приводит к капиллярным явлениям.

Явления изменения высоты уровня жидкости в капиллярах по сравнению с уровнем жидкости в широком сосуде называются ***капиллярными явлениями*** (рис. 8.5). Если жидкость смачивает стенки капилляра, то образуется вогнутый мениск радиусом r, молекулярное давление под которым на Δp меньше, чем под плоской поверхностью в широком сосуде, сообщающимся с капилляром. Так как молекулярное давление под плоской поверхностью в широком сосуде на Δp больше, чем в капилляре, то оно выталкивает жидкость в

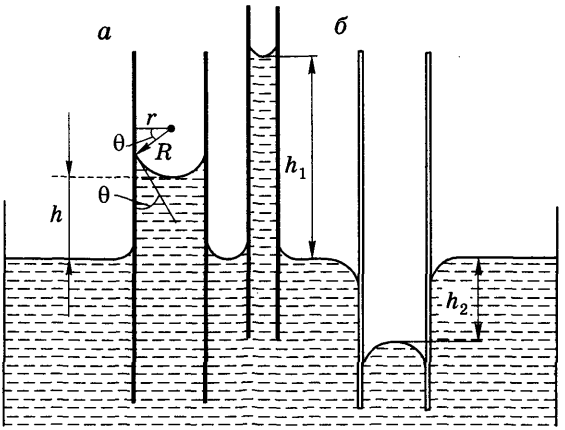


Рис. 8.5. Капиллярные явления:
а — подъем уровня жидкости в капиллярах разного диаметра при смачивании; *б* — опускание уровня жидкости в капилляре при несмачивании

капилляре вверх до тех пор, пока весовое давление образовавшегося столба жидкости высотой h не скомпенсирует добавочное молекулярное давление Δp:

$$\Delta p = \rho g h, \tag{8.5}$$

где ρ — плотность жидкости; g — ускорение свободного падения.

Если жидкость не смачивает капилляр, то образуется выпуклый мениск, молекулярное давление под которым на Δp больше, чем в широком сосуде, и направлено вниз, вытесняя жидкость ниже исходного уровня на глубину h, удовлетворяющую условию (8.5).

Таким образом, в капилляре жидкость поднимается (или опускается) на такую высоту h, при которой гидростатическое давление столба жидкости уравновешивает избыточное молекулярное давление, обусловленное кривизной мениска:

$$h = \frac{\Delta p}{\rho g}. \tag{8.6}$$

Если радиус кривизны мениска равен R, то, подставляя формулу (8.3а) в (8.6), получим

$$h = \frac{2\sigma}{\rho g R} = \frac{2\sigma\cos\theta}{\rho g r}. \qquad (8.7)$$

В последнем соотношении учтена связь между радиусом кривизны мениска R и радиусом капилляра r: $R = r/\cos\theta$, где θ — краевой угол.

Из формулы (8.7) видно, что чем тоньше капилляр и лучше смачивание (меньше θ и больше $\cos\theta$), тем выше поднимается жидкость по капилляру. При идеальном смачивании ($\theta = 0$, $\cos\theta = 1$, $R = r$) высота подъема максимальна:

$$h_{max} = \frac{2\sigma}{\rho g r}. \qquad (8.7а)$$

Если же жидкость не смачивает капилляр, то ее уровень в капилляре опускается на глубину, определяемую формулами (8.7) и (8.7а), т.е. несмачивающая жидкость «выталкивается» из капилляра. Поэтому такая жидкость не может глубоко проникнуть в поры твердого тела. С этим связана, например, непроницаемость для воды перьев птиц, смазанных жиром. Напротив, смачивающая жидкость хорошо проникает в такие поры и удерживается в них. Благодаря этому явлению пористые вещества могут удерживать значительное количество жидкости даже из паров воздуха, что приводит к увлажнению белья, ваты в сырых помещениях, затрудняет сушку гигроскопических тел.

Капиллярные явления весьма распространены, ими обусловлено поднятие воды в почве, по корневой системе растений, движение биологических жидкостей по системе мелких сосудов и канальцев и многие другие явления.

8.6. Газовая эмболия

С поверхностным натяжением связано и явление газовой эмболии, при котором пузырек газа способен затруднить и даже остановить кровоток в мелких сосудах, лишить кровоснабжения какой-либо орган, что может привести

к серьезному функциональному расстройству и даже летальному исходу. Поэтому рассмотрим подробнее поведение пузырька воздуха, находящегося в капилляре с жидкостью.

Пока диаметр газового пузырька меньше диаметра сосуда, он имеет почти сферическую форму и движется вместе с током крови. Если он попадает в мелкий сосуд, диаметр которого меньше диаметра пузырька, то его мениски деформируются под действием динамического давления текущей крови: передний по току крови мениск вытягивается, его радиус кривизны уменьшается, а задний под напором крови уплощается и его радиус кривизны увеличивается. Соответственно, дополнительные молекулярные давления, действующие на эти мениски, будут не одинаковы и направлены навстречу, а их результирующая сила, приложенная к пузырьку, будет направлена против тока крови, противодействуя ему, вплоть до возможной остановки кровотока (рис. 8.6, *а*). Еще более сложная ситуация возникает, когда газовый пузырек попадает на разветвление (бифуркацию) сосудистого русла и полностью останавливает кровоток в дистально расположенных сосудах (рис. 8.6, *б*). Таким образом, попавшие в кровь пузырьки воздуха способны закупорить мелкие сосуды.

Газовая эмболия может возникнуть при ранении крупных вен, где давление крови ниже атмосферного, при неправиль-

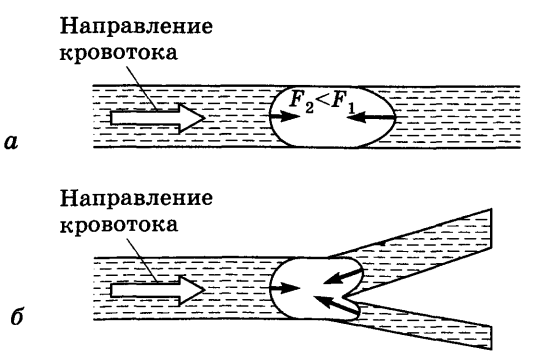


Рис. 8.6. Влияние газового пузырька на кровоток: *а* — в мелком сосуде; *б* — на разветвлении сосудов

но проведенных внутривенных инъекциях и в других ситуациях.

Газовые пузырьки в крови человека и животных могут появиться и при резком снижении внешнего давления на организм, что обусловлено уменьшением растворимости газов (в первую очередь азота) в крови и переходом их из растворенного состояния в газообразное вследствие резкого снижения окружающего давления. С подобной проблемой могут столкнуться водолазы при быстром подъеме с большой глубины на поверхность (кессонная болезнь), летчики и космонавты при разгерметизации кабины или скафандра на большой высоте.

8.7. Роль альвеолярного сурфактанта в процессе дыхания

В легких происходит газообмен кислородом, углекислым газом и азотом между организмом и внешней средой. Структурной единицей легкого, в которой происходит этот газообмен, является *альвеола*. В легких примерно 300 миллионов альвеол, общая поверхность которых достигает 70–80 м^2. Диаметр альвеол составляет всего 75–300 мкм, поверхностное натяжение альвеолярных мембран, выстилающих их изнутри, создает добавочное давление Лапласа, стремящееся уменьшить размер альвеол. При вдохе объем альвеол увеличивается благодаря работе дыхательных мышц, давление в них становится ниже атмосферного и это способствует их заполнению воздухом. После выдоха объем альвеол уменьшается, добавочное давление в них увеличивается, что должно было бы сильнее уменьшить размер альвеол и препятствовало бы следующему акту вдоха. Этого не происходит благодаря специальному веществу — ***сурфактанту***, покрывающему внутреннюю поверхность альвеол. Сурфактант существенно снижает поверхностное натяжение альвеолярных стенок, причем его коэффициент поверхностного натяжения зависит от толщины пленки: на вдохе она тоньше, так как поверхность альвеолы больше, $\sigma = 0{,}05$ Н/м,

а на выдохе эта пленка толще и коэффициент поверхностного натяжения сурфактанта снижается в 10–50 раз, принимая значения $\sigma = 0{,}005$–$0{,}001$ Н/м, поэтому снижается и добавочное давление в альвеолах, обеспечивая возможность последующего вдоха. Таким образом, именно наличие сурфактанта с его уникальными поверхностными свойствами препятствует полному схлопыванию альвеол и обеспечивает возможность дыхания.

8.8. Методы определения коэффициента поверхностного натяжения жидкостей

Капиллярный метод основан на использовании соотношения (8.7) для определения коэффициента поверхностного натяжения жидкости путем измерения высоты h поднятия ее уровня в капилляре известного радиуса r:

$$\sigma = \frac{\rho g h r}{2\cos\theta}. \qquad (8.8)$$

Основной недостаток этого метода — сложность определения величины краевого угла θ. Поэтому использовать этот метод удобно только в тех случаях, когда смачивание близко к идеальному и $\cos\theta \approx 1$.

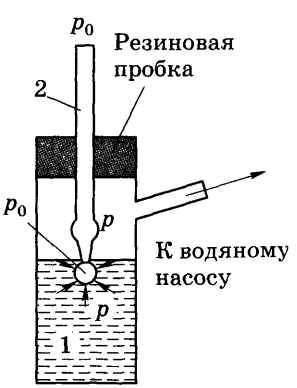


Рис 8.7. Измерение максимального давления в пузырьке воздуха

Метод Ребиндера (метод определения максимального давления в пузырьке). В исследуемую жидкость *1*, находящуюся в сосуде (рис. 8.7), вертикально опускается капиллярная трубка *2*, узкий конец которой диаметром не более 0,5 мм касается поверхности исследуемой жидкости. Другим концом эта трубка сообщается с атмосферным воздухом, поэтому внутри капилляра поддерживается атмосферное давление p_0.

Давление p над исследуемой жидкостью постепенно уменьшают с помощью водяного насоса. Разность давлений $(p_0 - p)$ стремится выдуть пузырек воздуха из капилляра в жидкость, но этому противодействует добавочное давление $\Delta p = 2\sigma/r$, создаваемое силами поверхностного натяжения жидкости в образующемся пузырьке радиусом r и направленное к центру пузырька. Наконец, при некоторой разности давлений $(p_0 - p)$ из капиллярной трубки выдувается в жидкость воздушный пузырек. Разность давлений $(p_0 - p)$, максимальная в этот момент, измеряется U-образным манометром и равна $\rho g h$, где ρ — плотность жидкости в манометре, h — разность ее уровней. Таким образом, в момент выдувания пузырька имеет место равенство

$$\rho g h = \frac{2\sigma}{r}. \qquad (8.9)$$

Здесь неизвестен радиус r выдуваемого пузырька, измерить который крайне затруднительно. Поэтому прибегают к использованию эталонной жидкости, коэффициент поверхностного натяжения σ_0 которой известен и близок к коэффициенту поверхностного натяжения σ исследуемой жидкости. При этом полагают, что радиусы пузырьков, выдуваемых из одного и того же капилляра в обоих случаях будут одинаковы.

Теперь вместо исследуемой жидкости в сосуд наливают эталонную жидкость и измеряют по манометру максимальную разность уровней h_0, при которой пузырек воздуха выдувается в эталонной жидкости и выполняется равенство

$$\rho g h_0 = \frac{2\sigma_0}{r}. \qquad (8.9а)$$

Разделив уравнение (8.9) на (8.9а) и решив его относительно σ, получаем формулу для вычисления поверхностного натяжения исследуемой жидкости:

$$\sigma = \sigma_0 \frac{h}{h_0}. \qquad (8.10)$$

Рассмотренным методом можно определять поверхностное натяжение и на границе раздела двух несмеши-

вающихся жидкостей. В этом случае узкий конец капиллярной трубки должен касаться поверхности раздела этих жидкостей, но при этом следует обязательно учитывать гидростатическое давление $\rho_1 gH$ жидкости, расположенной сверху, толщина слоя которой равна H, а плотность — ρ_1.

Сталагмометрический метод (метод счета капель). Под действием поверхностного натяжения свободная поверхность капель жидкости стремится принять шарообразную форму, соответствующую наименьшей поверхностной энергии и наименьшей площади свободной поверхности. Их форма тем ближе к шаровой, чем меньше вес капель, поскольку для малых капель сила поверхностного натяжения превосходит силу тяжести.

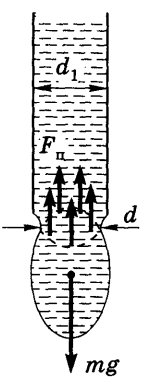


Рис. 8.8. Истечение капли из капилляра

Форма и размер капель, отрывающихся от конца капиллярной трубки, зависят не только от силы поверхностного натяжения, но и от диаметра трубки и плотности вытекающей жидкости. При вытекании жидкости из капиллярной трубки размер капли постепенно растет.

На рис. 8.8 показан процесс образования капли. Перед отрывом капли образуется шейка, диаметр d которой несколько меньше диаметра d_1 капиллярной трубки. По окружности шейки капли действуют силы поверхностного натяжения, удерживающие каплю. По мере увеличения размера капли растет сила тяжести mg, стремящаяся оторвать ее. В момент отрыва капли она равна результирующей силе поверхностного натяжения $F_п = \pi d\sigma$:

$$\pi d\sigma = mg. \qquad (8.11)$$

Таким образом, измеряя массу m одной капли и зная диаметр d шейки капли, можно вычислить коэффициент поверхностного натяжения:

$$\sigma = \frac{mg}{\pi d}. \qquad (8.12)$$

Массу одной капли определяют взвешиванием на аналитических весах определенного отсчитанного количества капель (отсюда и название метода) и последующего вычисления средней массы одной капли.

Метод отрыва кольца. На поверхность исследуемой жидкости помещают кольцо или рамку. Если жидкость смачивает кольцо, то силы поверхностного натяжения F_1 и F_2, действующие на его наружную и внутреннюю поверхности диаметрами D и d, направлены внутрь жидкости, как показано на рис. 8.9, и создают суммарную силу поверхностного натяжения

$$F_{\text{п}} = \sigma\pi(D + d). \qquad (8.13)$$

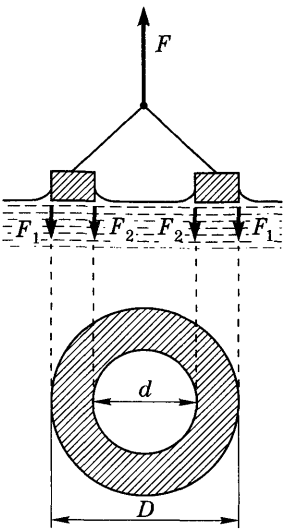


Рис. 8.9. Измерение коэффициента поверхностного натяжения методом отрыва кольца

Чтобы оторвать кольцо от поверхности жидкости, надо приложить направленную вверх силу F, которая скомпенсирует силу тяжести mg кольца и силу поверхностного натяжения $F_{\text{п}}$:

$$F = F_{\text{п}} + mg. \qquad (8.14)$$

Измерив с помощью динамометра или весов силу F отрыва кольца и зная его массу и размеры, из соотношений (8.13) и (8.14) получают выражение для коэффициента поверхностного натяжения жидкости:

$$\sigma = \frac{F - mg}{\pi(D + d)}. \qquad (8.15)$$

Контрольные вопросы и задания

1. Чем обусловлено возникновение потенциальной энергии поверхностного слоя жидкости? От чего зависит эта энергия?
2. Каков физический смысл коэффициента поверхностного натяжения? от чего он зависит? какова его размерность?

3. Чему равны и как направлены силы поверхностного натяжения? Каким образом проявляются эти силы и как их можно измерить?
4. Чем обусловлены молекулярное давление поверхностного слоя на жидкость и добавочное давление под изогнутой поверхностью жидкости? Запишите формулу Лапласа.
5. Как объясняются явления смачивания и несмачивания? Что такое краевой угол?
6. Как рассчитать высоту подъема (опускания) жидкости в капиллярных трубках?
7. В чем сущность газовой эмболии и каковы условия ее возникновения?
8. Какова роль поверхностного натяжения альвеолярного сурфактанта легких в процессе дыхания?
9. Охарактеризуйте методы определения коэффициента поверхностного натяжения жидкостей.

Глава 9 Физические основы гемодинамики

Гемодинамика — наука, изучающая законы движения крови по сосудистой системе. Общие законы течения жидкости изучены гидродинамикой, установлены в рамках классической физики и являются основой для описания гемодинамических процессов в живом организме. Однако сложная организация реальной системы кровообращения, специфические свойства движущейся крови, механические характеристики кровеносных сосудов и ряд других факторов приводят к значительным трудностям в задаче полного количественного описания движения крови в организме. В то же время только на основе количественных закономерностей системы кровообращения возможны глубокое понимание гемодинамических явлений, разработка методов диагностики и лечения ряда заболеваний.

9.1. Основные гидродинамические понятия и законы

Для описания движения жидкости используют два метода. Можно определять, как изменяются во времени траектория и скорость каждой частицы жидкости. Такой метод исследования называют *методом Лагранжа*. Однако можно наблюдать не за отдельными частицами, а за выделенными точками пространства, через которые проходят частицы жидкости. Этот метод называют *методом Эйлера*.

9.1.1. Условие неразрывности струи

Течение жидкости характеризуется линиями тока. Это линии, касательные к которым совпадают с направлением вектора скорости частиц жидкости в данной точке (рис. 9.1). Часть пространства, ограниченная линиями тока,

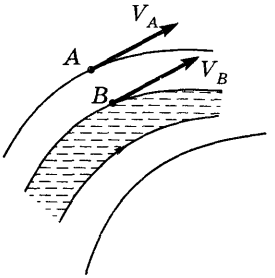


Рис. 9.1. Линии тока при ламинарном течении

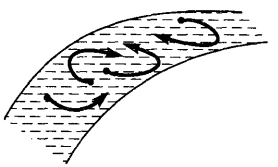


Рис. 9.2. Линии тока при турбулентном течении

называется ***трубкой тока*** (на рис. 9.1 заштрихована). Если при течении жидкости линии тока непрерывны, то такое течение называется ***ламинарным*** (рис. 9.1).

При определенных условиях в движущейся жидкости могут возникать завихрения, скорость ее частиц хаотически изменяется, линии тока претерпевают разрывы, изменяющиеся со временем. Такое движение жидкости называется ***турбулентным*** (рис. 9.2).

Для установления связи между скоростью ламинарного течения жидкости и площадью поперечного сечения участка, через который она протекает, выделим в трубке тока участки с площадью поперечного сечения S_1 и S_2 (рис. 9.3). В пределах этих сечений скорости частиц жидкости одинаковы, направлены перпендикулярно выделенным площадкам и равны по величине v_1 и v_2 соответственно. Объемы жидкости V_1 и V_2, протекающей через выделенное сечение за одно то же время t, одинаковы, так как жидкость практически несжимаема. Это позволяет записать равенство

$$S_1 v_1 t = S_2 v_2 t, \quad \text{или} \quad S_1 v_1 = S_2 v_2, \quad \text{или} \quad Sv = \text{const}. \qquad (9.1)$$

Уравнение (9.1) представляет собой ***условие неразрывности струи***, утверждающее, что при ламинарном течении

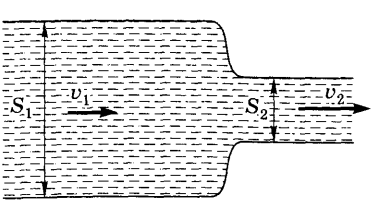


Рис. 9.3. Условие неразрывности струи

несжимаемой жидкости произведение площади сечения участка, через который она протекает, и ее скорости является постоянной величиной для данной трубки тока.

При течении жидкости различают ее линейную и объемную скорость. *Линейная скорость* (υ) — это путь L, проходимый частицами жидкости в единицу времени: $\upsilon = L/t$ (для равномерного течения). *Объемная скорость* (*расход*) Q — это объем жидкости V, протекающий через некоторое сечение за единицу времени t: $Q = V/t$. Объемная и линейная скорости течения жидкости связаны очевидным соотношением: $Q = \upsilon S$, где S — площадь поперечного сечения потока жидкости. Линейная скорость кровотока измеряется в м/с, а объемная — в $м^3/с$, л/мин, мл/мин и др.

Условие неразрывности струи (9.1) выполняется и в реальной гемодинамике. Здесь формулировка этого условия звучит следующим образом: в любом сечении сердечно-сосудистой системы объемная скорость кровотока одинакова: $Q = \text{const}$.

Под площадью сечения сосудистой системы понимают суммарную площадь сечения кровеносных сосудов одного уровня ветвления. Так, в большом круге кровообращения первое (наименьшее по площади) сечение проходит через аорту, второе — через все артерии, на которые непосредственно разветвляется аорта, и т.д. Наибольшую площадь имеет сечение, соответствующее капиллярной сети.

Как следует из условия неразрывности струи, с увеличением площади сечения сосудистой системы скорость кровотока в ее соответствующих участках уменьшается. Так, в покое средняя линейная скорость кровотока в аорте составляет около 0,4–0,5 м/с, а в капиллярах — около 0,5 мм/с. Следовательно, сумма поперечных сечений всех функционирующих капилляров примерно в 800 раз больше площади сечения аорты.

9.1.2. Уравнение Бернулли

Основным количественным соотношением, описывающим течение идеальной жидкости, т.е. абсолютно несжимаемой и невязкой, является уравнение Бернулли, вытекающее из

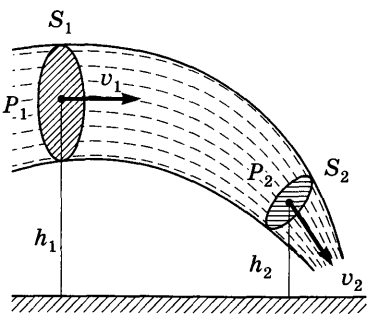


Рис. 9.4. К выводу уравнения Бернулли

закона сохранения энергии в движущейся жидкости. Для его установления рассмотрим трубку тока идеальной жидкости, в которой выделим два сечения площадью S_1 и S_2 (рис. 9.4). Пусть центры этих сечений расположены на высотах h_1 и h_2, отсчитываемых от некоторого уровня. Линейные скорости частиц жидкости в этих сечениях обозначим v_1 и v_2. Силы, обусловливающие течение жидкости, оказывают давление P_1 и P_2 на торцах объема жидкости между выделяемыми сечениями S_1 и S_2. При стационарном течении идеальной жидкости (нет трения) ее полная энергия во всех сечениях одинакова, следовательно,

$$mv_1^2/2 + P_1V + mgh_1 = mv_2^2/2 + P_2V + mgh_2, \qquad (9.2)$$

где m — масса некоторого объема V жидкости, последовательно протекающего через сечения S_1 и S_2.

Первые слагаемые в обоих частях равенства представляют кинетическую энергию жидкости, вторые — потенциальную энергию давления, третьи — потенциальную энергию, обусловленную расположением этой массы жидкости на высотах h_1 и h_2.

Разделив правую и левую части соотношения (9.2) на объем жидкости V и вводя плотность жидкости ρ, получим

$$\rho v_1^2/2 + P_1 + \rho g h_1 = \rho v_2^2/2 + P_2 + \rho g h_2,$$

или

$$\rho v^2/2 + P + \rho g h = \text{const}. \qquad (9.3)$$

Соотношение (9.3) называется ***уравнением Бернулли***, утверждающим, что сумма разнопричинных давлений в любом сечении трубки тока (сосуда) является постоянной величиной. Слагаемое $\rho v^2/2$ представляет *динамическое давление*, обусловленное движением жидкости; P — *статическое дав-*

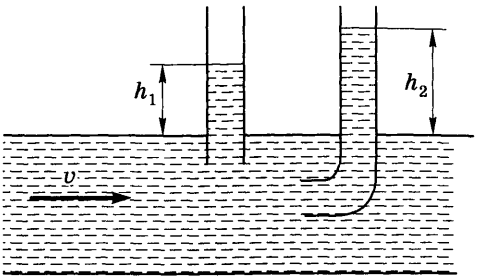


Рис. 9.5. Метод определения скорости движения жидкости

ление, не связанное с движением жидкости (оно может быть измерено, например, жидкостным манометром, ось входного отверстия которого расположена перпендикулярно линии тока жидкости); ρgh — *весовое (гидростатическое) давление.*

Уравнение Бернулли строго выполняется для идеальной, т.е. невязкой, жидкости, поскольку при его выводе не учитывалась работа сил трения. Но его можно успешно применять и для жидкостей с невысокой вязкостью.

Рассмотрим некоторые следствия, вытекающие из уравнения Бернулли:

1. *Способ измерения скорости движения жидкости.*

Представим, что в движущуюся жидкость опущены две трубки малого сечения, причем плоскость поперечного сечения одной из них параллельна направлению скорости движения жидкости v, а другая (трубка Пито) изогнута так, что плоскость сечения изогнутой части перпендикулярна направлению скорости течения (рис. 9.5). Подъем жидкости в прямой трубке на высоту h_1 обусловлен лишь статическим давлением P_c, которое можно определить по формуле $P_c = \rho g h_1$. В трубке Пито подъем жидкости на высоту h_2 обусловлен полным давлением $P_п$ — в данном случае суммой статического P_c и динамического $P_д$ давлений (течение происходит горизонтально и весовое давление не учитывается). Следовательно,

$$P_п = P_c + P_д;$$

$$\rho g h_2 = \rho g h_1 + \rho v^2/2. \qquad (9.4)$$

Из формулы (9.4) находим линейную скорость жидкости

$$v = \sqrt{2g(h_2 - h_1)}. \qquad (9.5)$$

Таким образом, по измеренной разности уровней жидкости в прямой и изогнутой трубках определяется скорость течения жидкости. Этим же методом определяют и скорость самолета относительно воздуха, катера относительно воды и др.

2. *Всасывающее действие струи.*

Рассмотрим течение жидкости по горизонтальной трубе переменного сечения. Выделим два участка с площадью поперечного сечения S_1 и S_2, причем для определенности $S_1 > S_2$ (рис. 9.6). Запишем для данного случая уравнение Бернулли, учитывая, что в горизонтальной трубе $h_1 = h_2$:

$$\rho v_1^2/2 + P_1 = \rho v_2^2/2 + P_2, \qquad (9.6)$$

где v_1 и v_2 — скорости течения жидкости в сечениях S_1 и S_2.

Статические давления P_1 и P_2 в соответствующих сечениях могут быть определены по высотам подъема жидкости h_1 и h_2 в соответствующих трубках. Поскольку $S_1 > S_2$, то $v_1 < v_2$, т.е. в узких местах жидкость течет быстрее. Тогда из уравнения (9.6) следует, что $P_1 > P_2$, т.е. статическое давление в более широкой части трубки большее, чем в ее узкой части. Если сужение значительно, то скорость жидкости в нем v_2 намного превышает v_1, статическое давление P_2 резко уменьшается и может стать ниже атмосферного. В этом случае воздух (или окружающая трубку другая среда) будет засасываться через отверстие в месте расположения сужения. На

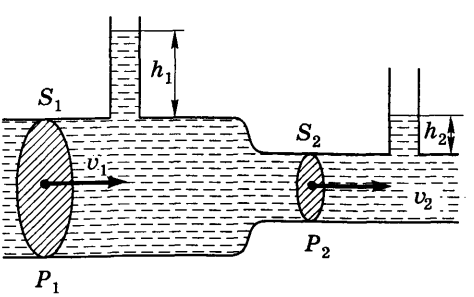


Рис. 9.6. Течение жидкости по горизонтальной трубе переменного сечения

этом принципе устроены водоструйные насосы, ингаляторы, пульверизаторы и др.

3. *Закупорка артерии, артериальный шум* [12].

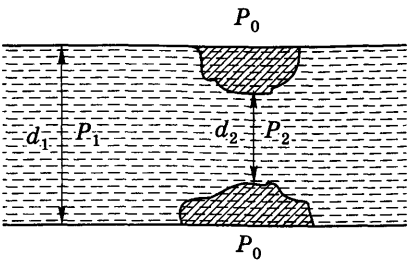


Рис. 9.7. Схематическое изображение участка сосуда

В случае образования атеросклеротической бляшки на некотором участке артерии диаметром d_1 может возникнуть сужение диаметром d_2 (рис. 9.7). Чем меньше диаметр d_2, тем больше скорость кровотока v_2 и меньше статическое давление P_2 в области сужения сосуда. Течение крови по артерии будет происходить до тех пор, пока статическое давление P_2 в области атеросклеротической бляшки будет превышать давление P_0 на сосуд снаружи, которое в обычных условиях равно атмосферному. Итак, нормальный кровоток возможен лишь при условии

$$P_2 - P_0 \geq 0. \qquad (9.7)$$

Это условие будет реализовано, пока диаметр сужения d_2 превышает (или равен) некоторое минимальное значение $d_{\min}$, которое можно вычислить, используя уравнение Бернулли. Например, для сонной артерии диаметром $d_1 = 1$ см в норме скорость крови $v_1 = 0{,}2$ м/с, плотность крови $\rho = 1{,}05 \cdot 10^3$ кг/м^3, разность давлений $P_1 - P_0 = 100$ мм рт.ст. $= 1{,}33 \cdot 10^4$ Па, и расчет дает значение минимального диаметра просвета сосуда $d_{\min} \approx 2$ мм.

Если диаметр сужения станет меньше $d_{\min}$, то внешнее давления P_0 «захлопнет» просвет сосуда в месте расположения атеросклеротической бляшки. Если бы давление P_1 в незакупоренном участке сосуда оставалось неизменным, то кровоток был бы полностью остановлен. Однако в результате работы сердца давление P_1 начнет возрастать и кровь будет с усилием проталкиваться через сужение. При этом сердце работает в условиях повышенной нагрузки. При прослушивании фонендоскопом наличие этих толчков проявляется в виде прерывистого шума, свидетельствующего о нарушении нормального кровотока.

4. *Поведение аневризмы.*

Некоторые патологические процессы могут приводить к локальному снижению прочностных и упругих свойств кровеносных сосудов. В результате на некотором участке сосуда его деформация под действием пульсирующего кровотока становится необратимой — возникает *вздутие сосуда* (аневризма). Схематически такой сосуд показан на рис. 9.8.

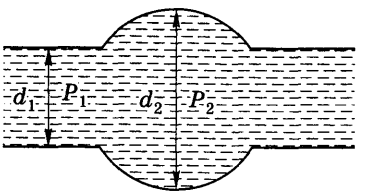


Рис. 9.8. Развитие аневризмы

Скорость кровотока v_2 в месте развития аневризмы диаметром d_2 по условию неразрывности струи будет меньше, чем скорость v_1 в его недеформированной части. На основании уравнения Бернулли (9.6) статическое давление P_2 в месте аневризмы будет больше давления P_1 на участках сосуда нормального сечения. Следовательно, нагрузка на расширенную часть сосуда увеличится и возникшая аневризма под действием повышенного давления будет иметь тенденцию к расширению. Разница давлений $P_1 - P_2$ может быть рассчитана для конкретных сосудов с использованием формулы (9.6). Например, если диаметр сосуда $d_1 = 2{,}5$ см, а диаметр аневризмы $d_2 = 5$ см, то давление в области аневризмы будет больше на $\Delta P \approx 0{,}3$ мм рт.ст. $= 40$ Па. Эта величина мала по сравнению с абсолютным давлением, но все-таки даже это малое избыточное давление стремится расширить аневризму, что приводит к еще большему замедлению скорости кровотока в деформированной части сосуда и дальнейшему повышению статического давления и т.д. В результате возможен разрыв аневризмы.

9.2. Вязкость жидкости. Ньютоновские и неньютоновские жидкости

Между молекулами реальной жидкости всегда существуют силы взаимодействия, которые при ее течении проявляются как силы трения, направленные по касательной к поверхности перемещающихся слоев. Эти силы и определяют ***внутрен-***

нее трение, или ***вязкость жидкости.*** Наличие сил внутреннего трения приводит к тому, что различные слои жидкости движутся с различными скоростями. На рис. 9.9 представлен случай, когда для самого верхнего слоя жидкости помеха его движению минимальна, а для нижнего — максимальна (например, течение в реке).

Таким образом скорость движущихся слоев вязкой жидкости изменяется от слоя к слою в направлении оси Ox, перпендикулярной направлению течения жидкости. Для частного случая, показанного на рис. 9.9, изменение скорости слоев происходит по линейному закону. Количественно величина различий в скоростях движения соседних слоев жидкости характеризуется градиентом скорости dv/dx, называемым также ***скоростью сдвига.*** Сила трения $F_{тр}$ между слоями движущейся жидкости пропорциональна площади соприкосновения S слоев жидкости и градиенту скорости и определяется формулой Ньютона:

$$F_{тр} = \eta S \frac{dv}{dx}. \qquad (9.8)$$

Коэффициент η, зависящий от свойств жидкости и температуры, называют ***коэффициентом внутреннего трения*** или ***динамической вязкостью,*** и часто просто ***вязкостью*** жидкости. Единицей вязкости в Международной системе

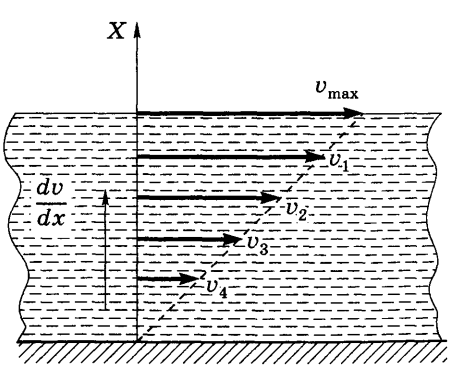


Рис. 9.9. Градиент скорости движения слоев жидкости

единиц является паскаль-секунда (Па·с). Применяется и внесистемная единица вязкости — пуаз (П), 1 Па·с = 10 П.

Вязкость всех жидкостей зависит от температуры и, как правило, с увеличением температуры вязкость уменьшается. Вязкость некоторых жидкостей зависит от градиента скорости, возникающего в текущей жидкости, т.е. от условий течения.

Жидкости, вязкость η которых не зависит от условий течения, называют ***ньютоновскими***. К ним относятся в основном однородные жидкости, чистые растворители, например вода. Вязкость воды при температуре 20 °С составляет 1 мПа · с = 1 сП (сантипуаз).

Если же вязкость жидкости зависит от градиента скорости, то такая жидкость называют ***неньютоновской***. Обычно это жидкости, существенно неоднородные по составу — взвеси, суспензии и т.п. Типичной неньютоновской жидкостью является *кровь*, так как она представляет собой взвесь форменных элементов (эритроцитов, лейкоцитов и др.) в плазме. Это значит, что вязкость крови не одинакова в различных участках сосудистой системы и зависит от величины градиентов скоростей, реализующихся при ее движении.

9.3. Течение вязкой жидкости. Формула Пуазейля

Занимаясь исследованием кровообращения, французский врач и физик Ж. Пуазейль пришел к необходимости количественного описания процессов течения вязкой жидкости вообще. Установленные им для этого случая закономерности имеют важное значение для понимания сущности гемодинамических явлений и их количественного описания.

Не производя строгих математических расчетов, проанализируем, от чего зависит объем V вязкой жидкости, ламинарно протекающей по участку гладкой трубы длиной L и радиусом r (рис. 9.10). Очевидно, этот объем будет прямо пропорционален времени истечения жидкости t и тому перепаду давлений $P_1 - P_2$, который обусловливает ток жидкости

на участке трубы длиной L. Естественно, что объем вытекающей жидкости будет резко возрастать и с увеличением площади поперечного сечения рассматриваемого участка. Теоретические расчеты и непосредственный эксперимент показывают, что $V \sim r^4$. Помехой истечению жидкости является ее вязкость η, поэтому объем $V \sim 1/\eta$. Чем больше длина участка, тем больше потери в скорости протекающей по нему жидкости, значит $V \sim 1/L$. Приведенные соображения, строгое теоретическое рассмотрение и непосредственный эксперимент приводят к ***формуле Пуазейля***:

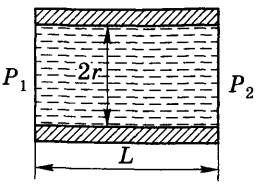


Рис. 9.10. Движение жидкости по участку гладкой трубы

$$V = \frac{\pi r^4 (P_1 - P_2)}{8\eta L} t \,. \qquad (9.9)$$

Разделив обе части этого выражения на время истечения t, получим формулу Пуазейля для объемной скорости течения жидкости:

$$Q = \frac{\pi r^4 (P_1 - P_2)}{8\eta L} \,. \qquad (9.9а)$$

По аналогии с законом Ома для участка электрической цепи это соотношение можно записать в виде более простой формулы Гагена — Пуазейля:

$$Q = \frac{(P_1 - P_2)}{X} \,. \qquad (9.10)$$

Величина X, входящая в это уравнение, называется гидравлическим сопротивлением участка трубы или сосуда:

$$X = \frac{8\eta L}{\pi r^4} \,. \qquad (9.11)$$

Между законами гидродинамики и законами протекания электрического тока по электрическим цепям существует тесная аналогия. Объемная скорость течения жидкости $Q = V/t$ является гидродинамическим аналогом силы электрического тока $I = q/t$. Причиной возникновения электри-

ческого тока является разность электрических потенциалов $\varphi_1 - \varphi_2$ на соответствующем участке цепи, а причиной движения жидкости — разность давлений $P_1 - P_2$ на участке трубы. В законе Ома $I = (\varphi_1 - \varphi_2)/R$ величина R — электрическое сопротивление проводника, аналогом которого в формуле (9.10) является величина $X = 8\eta L/\pi r^4$, представляющая собой гидравлическое сопротивление участка трубы или сосуда.

Если от общих законов истечения вязкой жидкости перейти к задачам гемодинамики, то с помощью уравнения Гагена — Пуазейля можно определить ряд характеристик кровотока. Так, зная объемную скорость кровотока Q и величину гидравлического сопротивления X сосудов, можно найти величину давления крови в любой точке сосудистой системы:

$$P = P_0 - QX. \qquad (9.12)$$

Если P_0 — давление крови в желудочке сердца, а X — общее сопротивление сосудов на участке сосудистой системы между этим желудочком и некоторой точкой, то давление крови P в этой точке определяется формулой (9.12).

Гидравлическое сопротивление X разветвленного участка сосудистой системы может быть определено по аналогии с расчетом общего электрического сопротивления участка электрической цепи, состоящего из набора отдельных резисторов. При последовательном соединении сосудов (рис. 9.11, *а*) общее сопротивление определяется суммой гидравлических сопротивлений их отдельных участков:

$$X = X_1 + X_2 + X_3 + \ldots + X_n, \qquad (9.13)$$

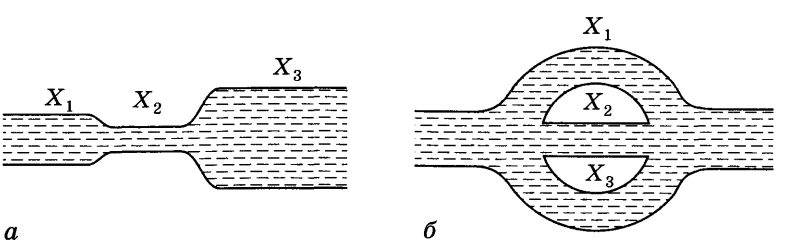


Рис. 9.11. Виды ветвления сосудистого русла:
а — последовательное; *б* — параллельное

а при параллельном ветвлении сосудистого русла (рис. 9.11, *б*) общее сопротивление X находится из уравнения

$$\frac{1}{X}=\frac{1}{X_1}+\frac{1}{X_2}+\frac{1}{X_3}+\ldots+\frac{1}{X_n}\ . \qquad (9.14)$$

Следует отметить, что аналогия в описании электрических цепей и гидродинамических процессов плодотворно используется при моделировании гемодинамических явлений.

9.4. Методы определения вязкости жидкости

Определение вязкости биологических жидкостей, и особенно вязкости крови, имеет существенное диагностическое значение. Разнообразные приборы, применяемые для этой цели, называют ***вискозиметрами***.

9.4.1. Метод Стокса (метод падающего шарика)

Представим цилиндр, заполненный жидкостью плотностью $\rho_ж$, вязкость которой η подлежит определению (рис. 9.12). Если в этой жидкости падает шарик радиусом r, массой m и плотностью ρ, то движение шарика определяется действующими на него тремя силами: силой тяжести $F_т = mg$, силой Архимеда $F_A = 4\pi r^3\rho_ж g/3$ и силой трения $F_{тр}$.

Согласно ***закону Стокса***, сила сопротивления движению шарика пропорциональна его радиусу, скорости движения и вязкости жидкости:

$$F_{тр} = 6\pi\eta r v. \qquad (9.15)$$

Сила трения уменьшает скорость движения шарика и через некоторое время после погружения шарика в жидкость его движение может стать равномерным. При достижении равномерного движения сила тяжести становится равной сумме силы трения и силы Архимеда:

$$4/3\pi r^3\rho g = 4/3\pi r^3\rho_ж g + 6\pi\eta r v. \qquad (9.16)$$

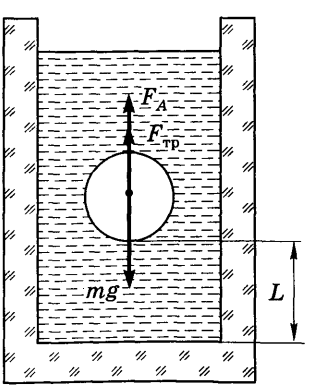


Рис. 9.12. Определение вязкости жидкости методом Стокса

Из этого уравнения определим искомую вязкость:

$$\eta = \frac{2(\rho - \rho_{ж})r^2 g}{9v}. \qquad (9.17)$$

Таким образом, для нахождения вязкости жидкости необходимо знать ее плотность $\rho_{ж}$, а также радиус r и плотность ρ шарика. Скорость шарика v определяется экспериментально, путем измерения времени t, за которое шарик равномерно проходит в жидкости определенное расстояние L: $v = L/t$.

Метод Стокса весьма прост и обладает хорошей точностью. Однако для определения вязкости крови он практически не применяется, так как требуется значительное количество исследуемой крови. Кроме того, в жидкостях, обладающих не очень большой вязкостью, сложно добиться равномерности движения шарика.

9.4.2. Капиллярные методы

Капиллярные методы основаны на применении формулы Пуазейля (9.9). Рассмотрим, например, течение жидкости через капилляр в *вискозиметре Оствальда* (рис. 9.13). Представим U-образную трубку. В одном из ее плеч имеется небольшая полая сфера объемом V, которая капилляром соединяется с резервуаром, расположенным в другом плече. Эта система заполняется жидкостью так, что разность ее уровней составляет величину h. Пусть вначале вискозиметр заполнен эталонной жидкостью, вязкость которой точно известна. В качестве такой жидкости удобно использовать дистиллированную воду.

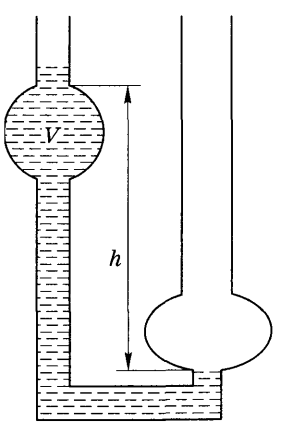


Рис. 9.13. Схема капиллярного вискозиметра

Поскольку при засасывании воды в левое плечо вискозиметра ее уровень здесь выше, чем в правом, то после прекращения всасывания жидкость будет под действием силы тяжести перетекать через капилляр из левого плеча

вискозиметра в правое до наступления равенства уровней. С помощью секундомера легко определить время t_0, за которое жидкость вытекает из полости объемом V. Согласно формуле (9.9), этот объем вытекшей воды:

$$V = \frac{\pi r^4 \rho_0 g h}{8\eta_0 L} t_0, \qquad (9.18)$$

где $\rho_0 gh$ — разница давлений (ρ_0 — плотность воды); η_0 — табличное значение вязкости воды при данной температуре.

Затем вискозиметр заполняют исследуемой жидкостью и измеряют время t истечения такого же объема V исследуемой жидкости:

$$V = \frac{\pi r^4 \rho g h}{8\eta L} t, \qquad (9.19)$$

где ρ — плотность исследуемой жидкости; η — вязкость исследуемой жидкости.

Приравнивая правые части выражений (9.18) и (9.19), получим формулу для определения вязкости исследуемой жидкости:

$$\eta = \eta_0 \frac{\rho t}{\rho_0 t_0}. \qquad (9.20)$$

Итак, при использовании этого метода необходимо знать вязкость эталонной жидкости, плотности исследуемой и эталонной жидкостей при данной температуре, измерить время их истечения через капилляр.

Для определения вязкости проб крови часто используют *вискозиметр Гесса*, в котором определяется не время истечения жидкости из капилляра, а расстояния L_0 и L, на которые перемещаются вода и кровь за одно и то же время. Применение формулы Пуазейля (9.9) для этого случая приводит к следующей расчетной формуле, определяющей вязкость крови η:

$$\eta = \eta_0 \frac{L_0}{L}. \qquad (9.21)$$

Простота капиллярных методов обусловила их широкое распространение в лабораторных исследованиях. Однако они не позволяют исследовать зависимость вязкости от скорости сдвига.

9.4.3. Ротационный метод

Достоинством ротационного метода является возможность определять значения вязкости неньютоновских жидкостей при различных, но контролируемых значениях градиентов скоростей и устанавливать зависимость вязкости от скорости сдвига: $\eta = f(dv/dx)$.

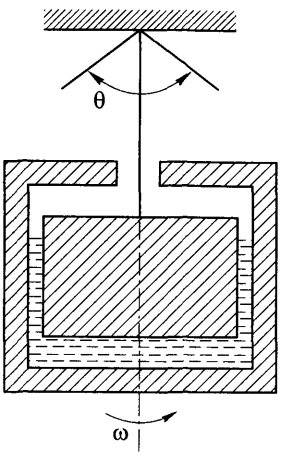


Рис. 9.14. Принцип действия ротационного вискозиметра

Существуют разнообразные ротационные вискозиметры. Рассмотрим принцип действия одного из них. Представим два цилиндра, имеющих общую ось вращения (рис. 9.14). Внутренний цилиндр подвешен на нити, а внешний может вращаться вокруг своей продольной оси с регулируемой угловой скоростью ω. Зазор между цилиндрами заполняется исследуемой жидкостью, например кровью.

За счет вязкости жидкости при вращении наружного цилиндра внутренний цилиндр начинает поворачиваться, достигая равновесия при некотором угле поворота θ. Этот угол можно легко измерить. Чем больше вязкость жидкости и угловая скорость вращения ω, тем больше и указанный угол поворота:

$$\theta = k\eta\omega,$$

где k — постоянная прибора.

При разных значениях угловой скорости вращения ω в жидкости, заполняющей зазор между цилиндрами, реализуются различные градиенты скорости. Для ньютоновских жидкостей значение вязкости не зависит от градиента скорости (и, следовательно, от величины ω), а для неньютоновских жидкостей эту зависимость можно не только наблюдать, но и определять количественно. Таким образом, данные ротационной вискозиметрии позволяют измерять вязкости движущейся крови при различных скоростях сдвига.

9.5. Условия перехода ламинарного течения жидкости в турбулентное

Характер течения жидкости — ламинарный или турбулентный — зависит от плотности жидкости ρ, ее вязкости η, скорости течения v и диаметра трубы d, по которой течет жидкость. Оказывается, что некоторая комбинация этих величин — один безразмерный параметр — может определять условия перехода ламинарного течения жидкости в турбулентное. Таким параметром является ***число Рейнольдса***

$$Re = \frac{\rho v d}{\eta}. \qquad (9.22)$$

Когда число Рейнольдса не превышает некоторого критического значения $Re_{кр}$,течение жидкости ламинарно. Если же $Re > Re_{кр}$, то в потоке жидкости возникают завихрения, т.е. течение становится турбулентным.

Значение критического числа Рейнольдса можно определить экспериментально. Представим, что по гладкой цилиндрической трубе протекает вода с регулируемой и измеряемой скоростью v, которая представляет собой среднюю по сечению трубы скорость течения. Плотность воды, ее вязкость и диаметр трубы известны. Допустим, что труба прозрачна и переход течения жидкости из ламинарного в турбулентное можно определить визуально. Постепенно увеличивая скорость течения, отметим то ее значение $v_{кр}$, при котором в потоке жидкости начинает проявляться турбулентность. Подставив это значение $v_{кр}$ в формулу (9.22), получим величину критического числа Рейнольдса. Для гладких труб $Re_{кр} = 2300$.

Если $Re_{кр}$ известно, то становится возможным для любой жидкости и разных условий ее течения предсказать, будет ее поток ламинарным или турбулентным.

Пример. Вода течет по гладкой трубе диаметром $d = 2$ мм. При какой скорости v ее течение становится турбулентным?

Примем вязкость воды $\eta = 10^{-3}$ Па·с, плотность $\rho = 10^3$ кг/м^3. Подставим эти значения в правую часть формулы (9.22), а в левую — значение критического числа Рейнольдса:

$$2300 = v \cdot 10^3 \cdot 2 \cdot 10^{-3}/10^{-3}.$$

Из полученного уравнения следует, что течение воды в этой трубе становится турбулентным при скорости $\upsilon = 1{,}15$ м/с. С увеличением диаметра трубы и уменьшением вязкости жидкости значение скорости, при которой происходит переход ламинарного течения в турбулентное уменьшается.

9.6. Некоторые особенности движения крови по сосудам

Движение крови в организме в основном ламинарное. Однако при определенных условиях кровоток может приобретать и турбулентный характер. Анализ формулы (9.22) позволяет предсказать эти условия. Действительно, турбулентности могут проявляться в полостях сердца (велико значение d). По-видимому, их наличие здесь физиологически целесообразно, поскольку возникающие завихрения приводят к перемешиванию порций крови, поступавших из малого круга кровообращения в левый желудочек сердца, и, следовательно, способствуют более равномерному обогащению кислородом крови, выталкиваемой затем в большой круг кровообращения. Сравнительно небольшие завихрения могут возникать в аорте и вблизи клапанов сердца (здесь велико и значение скорости движения крови).

При интенсивной физической нагрузке скорость движения крови увеличивается и это может вызвать турбулентности в кровотоке.

Из формулы (9.22) следует также, что с уменьшением вязкости турбулентный характер течения жидкости может проявляться и при меньших скоростях ее движения. Поэтому при некоторых патологических процессах, приводящих к аномальному снижению вязкости крови, кровоток в крупных кровеносных сосудах может стать турбулентным.

Следует иметь в виду, что значение критического числа Рейнольдса 2300 получено для гладких труб и ньютоновской жидкости. Для крови $Re_{кр}$ имеет меньшее значение и по различным литературным данным составляет около 900–1600.

Кроме того, кровеносный сосуд в ряде случаев нельзя моделировать гладкой трубой. Например, при наличии атеросклеротических бляшек в просвете сосудов имеются локальные сужения, приводящие к возникновению турбулентности в течении крови. Наличие турбулентности в кровотоке может быть обнаружено по шумам, прослушиваемым с помощью фонендоскопа.

Турбулентное течение крови по сосудам создает повышенную нагрузку на сердце, что способствует развитию патологических процессов в сердечно-сосудистой системе.

Средняя вязкость крови, измеренная капиллярным вискозиметром, в норме составляет 4–5 мПа·с. При различных патологиях значения вязкости крови могут изменяться от 1,7 до 22,9 мПа·с. Отношение вязкости крови к вязкости воды называют *относительной вязкостью крови*.

Следует подчеркнуть, что приведенные численные значения характеризуют *среднюю вязкость* крови в крупных кровеносных сосудах, точнее, вязкость проб крови вне организма, измеренную капиллярными методами (см. 9.4.2). Неоднородность состава крови, специфика строения и разветвления кровеносных сосудов приводят к довольно сложным изменениям вязкости крови, движущейся по сосудистой системе. Проанализируем основные факторы, влияющие на вязкость крови в живом организме.

Температура. В нормальных условиях температура тела поддерживается постоянной благодаря системе терморегуляции организма, в которой кровь сама играет роль теплоносителя. При повышении температуры должна уменьшаться и вязкость крови. По-видимому, это могло бы несколько уменьшить нагрузку на сердце при развитии в организме патологических процессов, сопровождающихся повышением температуры тела как защитной реакции организма. В переохлажденных участках организма вязкость крови повышается, кровоток затрудняется, ухудшается питание тканей, что ведет к развитию в них патологических процессов. Следует учесть, что изменение температуры может приводить к изменению степени агрегации эритроцитов и вызывать другие

изменения в структуре крови. Поэтому температурные изменения вязкости при патологических процессах достаточно сложны. Температурный фактор необходимо учитывать и при лечебных воздействиях, в частности при использовании гипертермии для лечения ряда заболеваний, т.е. повышения температуры всего тела или отдельных его частей за счет нагревания различными методами.

Гематокрит. Этот показатель представляет собой отношение объема эритроцитов ($V_{эр}$) к объему крови ($V_{кр}$), в котором они содержатся. В норме $V_{эр}/V_{кр} \approx 0{,}4$. Оказалось, что с повышением гематокрита вязкость крови возрастает. Увеличение гематокрита может происходить как из-за увеличения концентрации эритроцитов, их агрегации, так и за счет увеличения их размеров.

Известно, что вязкость венозной крови выше, чем артериальной. Это обусловлено тем, что эритроциты венозной крови содержат углекислый газ и имеют форму, близкую к сферической, тогда как в артериальной крови эритроциты имеют форму тора и соответственно меньший объем. Благодаря этому гематокрит и соответственно вязкость венозной крови выше, чем артериальной.

Скорость сдвига (градиент скорости). Линейная скорость крови и диаметры кровеносных сосудов в различных участках сосудистой системы изменяются очень сильно. Следовательно, существенно отличаются и скорости сдвига в потоке движущейся крови. Поскольку кровь является неньютоновской жидкостью, то и ее вязкость, зависящая от скорости сдвига, будет различной в разных отделах системы кровообращения.

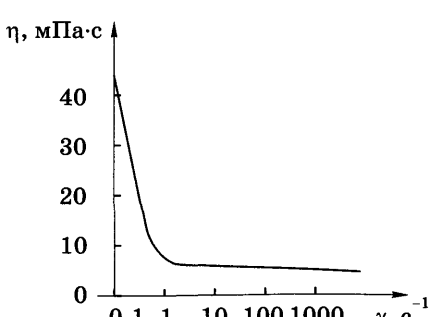


Рис. 9.15. Зависимость вязкости крови от скорости сдвига

Из графика зависимости вязкости крови от скорости сдвига (рис. 9.15) [19], видно, что при скоростях сдвига выше 100 $с^{-1}$

проявление неньютоновского характера движения крови незначительно и ее вязкость соответствует приведенным выше значениям 4–5 мПа·с. Однако при малых скоростях сдвига, меньших 1 $с^{-1}$, эффективная вязкость весьма резко возрастает.



Рис. 9.16. Реальный профиль скорости неньютоновской жидкости [12]

Организация эритроцитов в потоке крови. Существуют довольно сложные и не до конца выясненные механизмы, приводящие к снижению вязкости движущейся крови. Они связаны с перераспределением концентрации эритроцитов в потоке движущейся крови. Если бы по сосуду двигалась однородная ньютоновская жидкость, то скорость ее частиц по оси сосуда была бы максимальной, а у стенок — минимальной. Соединяя концы векторов скорости различных частиц жидкости, получим линию — ***профиль скорости.*** Для ньютоновской жидкости он имеет вид параболы, а для крови, движущейся по сосудам, профиль скоростей существенно «уплощается» (рис. 9.16). Это происходит по нескольким причинам. У стенки сосуда возникают большие градиенты скорости и, следовательно, большие деформации сдвига, которые «выталкивают» эритроциты в область меньших сдвиговых деформаций, т.е. к центру сосуда, где градиент скорости значительно меньше. Концентрация эритроцитов и соответственно вязкость крови возрастают к центру сосуда, что и приводит к «уплощению» профиля скоростей. Одновременно у стенок сосуда образуется тонкий пристеночный слой плазмы крови, не содержащий эритроцитов и поэтому обладающий низкой вязкостью. В итоге эритроциты продвигаются по сосуду как бы в оболочке из плазмы, что уменьшает трение крови о стенки и облегчает движение крови по сосудам.

9.7. Роль эластичности сосудов в системе кровообращения. Пульсовые волны

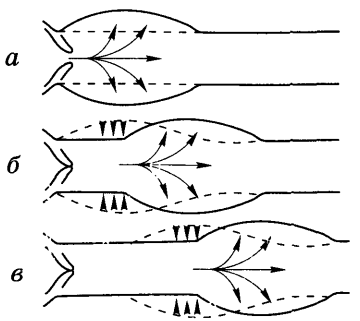


Рис. 9.17. Деформация сосуда при возникновении пульсовой волны: *а* — в начальный момент выброса систолического объема крови в аорту; *б*, *в* — распространение деформации по длине сосуда

При выбросе крови в аорту во время *систолы* часть кинетической энергии систолического объема крови переходит в потенциальную энергию упругой деформации стенок аорты (рис. 9.17, *а*) [37]. Образуется некоторый временный «резервуар», где запасается часть вытолкнутой желудочком крови. В *диастолу* проходит обратный процесс: потенциальная энергия деформированной стенки крупного кровеносного сосуда переходит в кинетическую энергию порции крови, создавая дополнительный фактор, способствующий ее движению. В каком-то смысле эластичный сосуд как бы «дорабатывает» усилие сердца.

Таким образом, выброс крови в аорту сопровождается упругими деформациями ее стенок и периодическими изменениями (колебаниями) давления крови на эти стенки. Их источником является периодический выброс крови в аорту при сокращении желудочка сердца. Распространяющиеся далее по сосудистой системе колебания давления крови, сопровождающиеся деформацией стенок сосудов, называют ***пульсовой волной***. Амплитуда пульсаций уменьшается при распространении волны от аорты к периферии (рис. 9.17, *б*, *в*).

Давление P на стенки кровеносных сосудов в некоторой точке сосудистой системы зависит от ряда параметров: времени t, расстояния от сердца до данной точки x, частоты сердечных сокращений ν, скорости распространения пульсовой волны v: $P = f(x, t, \omega, v, \alpha)$. Это давление можно представить в виде двух слагаемых:

$$P = P_{ср} + P(t), \qquad (9.23)$$

где $P_{ср}$ — давление, обусловленное постоянным средним уровнем кровенаполнения (постоянная составляющая); $P(t)$ — слагаемое, определяемое пульсовыми колебаниями кровотока.

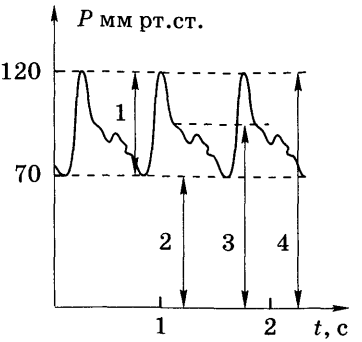


Рис. 9.18. Зависимость давления крови от времени в плечевой артерии

Колебания давления вызывают и изменения объема кровенаполнения. Считая кровеносный сосуд упругим резервуаром, связь между объемом крови V в данном участке сосуда в любой момент времени и давлением можно записать в виде уравнения

$$V = V_0 + k\,P(t), \qquad (9.24)$$

где V_0 — объем полости сосуда при среднем давлении $P_{ср}$; k — коэффициент пропорциональности, характеризующий эластичность сосуда.

Типичная зависимость давления крови от времени в норме в крупном кровеносном сосуде (плечевой артерии) показана на рис. 9.18, где отмечены значения *пульсового* (*1*), минимального, или *диастолического* (*2*), *среднего* (*3*) и максимального, или *систолического* (*4*) давления.

Следует подчеркнуть, что среднее давление $P_{ср}$ определяется не средним значением ординаты графика, а более сложным образом:

$$P_{ср} = \frac{1}{T}\int_0^T P(t)dt\,, \qquad (9.25)$$

где T — период пульсовых колебаний; t — текущее время.

Из рис. 9.18 видно, что пульсовые колебания давления имеют довольно сложную форму и аналитическая запись зависимости $P(x, t, \omega, \upsilon, \alpha)$ затруднена. Однако, как и всякий сложный периодический процесс, они могут быть представлены в виде набора гармонических составляющих (разложение в ряд Фурье). Гармонический анализ пульсовых ко-

лебаний кровотока является одним из важных методов его изучения. Тогда для первой гармонической составляющей давления (P_1) пульсовой волны можно записать достаточно простое выражение:

$$P_1 = P_0 e^{-\alpha x} \sin\omega\left(t - \frac{x}{\upsilon}\right), \qquad (9.26)$$

где P_0 — амплитуда пульсовых колебаний.

Подчеркнем, что коэффициент α зависит от свойств кровеносных сосудов и в формуле (9.26) под этой величиной можно понимать некоторое его эффективное значение. Реально эластичность сосуда уменьшается с увеличением расстояния от сердца к периферии. Морфологически это обусловлено изменением относительного содержания эластина и коллагена в сосудистой ткани. Так, в общей сонной артерии отношение эластина к коллагену 2:1, а в бедренной артерии 1:2. С удалением от сердца увеличивается доля гладких мышечных волокон, которые в артериолах являются уже основной составляющей сосудистой ткани.

Рассмотрим теперь скорость распространения пульсовой волны. В крупных кровеносных сосудах она определяется по ***формуле Моенса — Кортевега***

$$\upsilon = \sqrt{\frac{Eh}{\rho d}}, \qquad (9.27)$$

где E — модуль упругости стенки сосуда; h — толщина стенки; d — диаметр сосуда. Величину ρ можно считать плотностью вещества сосуда.

Как видно из формулы (9.27), с увеличением жесткости сосуда и увеличением толщины его стенки скорость пульсовой волны возрастает. Так, в аорте она равна 4–6 м/с, в артериях мышечного типа — 8–12 м/с. В венах, которые обладают большей эластичностью, скорость пульсовой волны меньше, и, например, в полой вене составляет около 1 м/с. Из этих данных следует, что скорость распространения пульсовой волны намного больше линейной скорости кровотока, в покое не превышающей даже в аорте значения 0,5 м/с.

С возрастом эластичность сосудов человека снижается (модуль упругости растет), а скорость пульсовой волны возрастает. Она растет и с увеличением давления. При повышенном давлении сосуд несколько растягивается, становится более «напряженным» и для его дальнейшего растяжения требуется большее усилие.

Форма пульсовых колебаний и их характеристики являются отражением работы сердца и состояния сосудистой системы. Поэтому их регистрация в различных участках сосудистой системы и последующий анализ имеют диагностическое значение. Некоторые методы регистрации этих процессов будут изучаться при рассмотрении механизмов прохождения электрического тока через живую ткань. Здесь отметим только принцип определения скорости распространения пульсовой волны на некотором участке сосудистой системы, который представлен на рис. 9.19. Верхняя кривая на этом рисунке — электрокардиограмма (ЭКГ), т.е. зависимость биопотенциалов, вызванных работой сердца, от времени. Каждый участок ЭКГ соответствует определенной фазе сокращения сердца. Нижняя кривая — пульсовые колебания, характеризующие изменение давления (а следовательно, и степень кровенаполнения) со временем в определенном участке сосудистой системы.

Начало систолы происходит раньше, чем начало увеличения прилива крови к исследуемому участку сосуда. Для распространения волны давления по сосудистой системе требуется некоторое время Δt, которое может быть определено из сравнения верхней и нижней кривых. Зная из анатомических соображений расстояние по сосуду от сердца до исследуемого

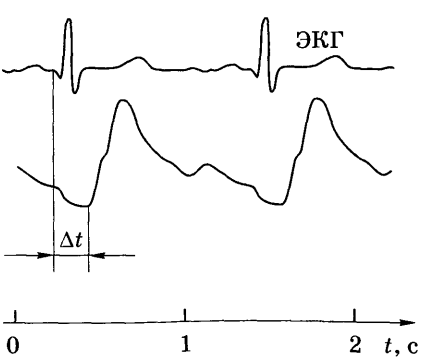


Рис. 9.19. Синхронная запись пульсовой волны и ЭКГ

участка L, можно определить среднюю скорость пульсовой волны: $\upsilon = L/\Delta t$.

При подобных исследованиях регистрируют еще и первую производную от нижней кривой (см. рис. 9.19). Если сама эта кривая отображает изменение объема кровенаполнения в данном участке сосудистой системы, то ее первая производная показывает, как изменяется во времени скорость кровенаполнения.

9.8. Распределение давления и скорости кровотока в сосудистой системе

Само понятие «давление крови» требует некоторых уточнений. В любой точке сосудистой системы давление крови зависит:

- от атмосферного давления;
- гидростатического давления ρgh, обусловленного весом кровяного столба высотой h и плотностью ρ;
- давления, обеспечиваемого насосной функцией сердца.

Разность давлений на внутреннюю $P_в$ и наружную $P_н$ стенки сосуда называют ***трансмуральным давлением*** $P_т$ (рис. 9.20): $P_т = P_в - P_н$. Можно считать, что давление на наружную стенку сосуда равно атмосферному.

Трансмуральное давление является важнейшей характеристикой состояния системы кровообращения, определяя нагрузку сердца, состояние периферического сосудистого русла и ряд других физиологических показателей. Однако необходимо иметь в виду, что это не то давление, которое обеспечивает движение крови от одной точки сосудистой системы к другой. Например, среднее по времени трансмуральное давление в крупной артерии руки составляет около 100 мм рт.ст. ($1{,}33 \cdot 10^4$ Па). В то же время движение крови из восходящей дуги аорты в эту артерию обеспечивается разностью трансмуральных давлений между

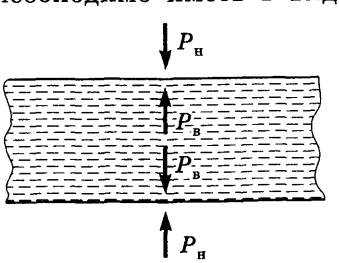


Рис. 9.20. Трансмуральное давление

указанными сосудами, которая составляет 2–3 мм рт.ст. (300 Па).

На величину трансмурального давления существенно влияет сила тяжести, создающая ***гидростатическое (весовое) давление***. Для пояснения этого влияния представим, что происходило бы с кровенаполнением сосудов вертикально расположенного тела человека, если бы его сердце не работало. В этом случае под действием силы тяжести кровь стекала бы в сосуды нижней части тела и верхний его уровень расположился бы в области сердца, где давление равнялось бы атмосферному, т.е. трансмуральное давление было бы равно нулю (рис. 9.21). На некоторой высоте h, отсчитываемой вниз от этого уровня, давление имело бы значение ρgh, где ρ — плотность крови, g — ускорение свободного падения, т.е. определялось бы только гидростатическим давлением.

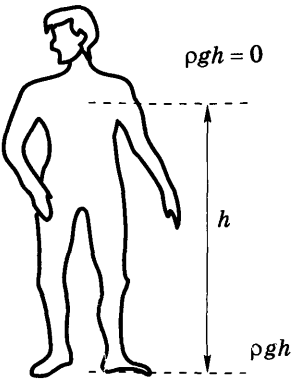


Рис. 9.21. Гидростатическое давление

Очевидно, гидростатическое давление влияет и на распределение крови в сосудистой системе живого человека. В этом случае оттоку крови из верхней части тела вертикально стоящего человека препятствует ряд физиологических механизмов. Кроме очевидной работы сердца, к ним относится рефлекторное сужение венозных сосудов ног в стоячем положении, которое сильно уменьшает способность этих сосудов растягиваться и накапливать кровь, а также способствует венозному возврату крови в сердце.

Если сосудосуживающий эффект ослаблен в результате заболевания или каких-либо внешних воздействий, то при резком вставании человек может впасть в обморочное состояние за счет уменьшения венозного возврата и снижения кровоснабжения головного мозга.

Измеряемое трансмуральное давление из-за воздействия на него гидростатической составляющей может существенно зависеть от выбора участков измерения и взаимного расположения частей тела. Так, трансмуральное давление в артериях

голени может быть намного больше, чем в артериях поднятой руки.

В клинических условиях измерение кровяного давления обычно производят в области плеча, т.е. на уровне сердца. Поэтому гидростатическая составляющая давления в плечевой артерии в этом случае равна нулю.

Движение крови по сосудистой системе происходит за счет превышения давления, обусловленного работой сердца, над атмосферным давлением. Именно градиент указанного давления и является движущей силой кровотока. Распределение этого давления в сосудистой системе показано на рис. 9.22 (верхняя кривая), где видно, что в аорте и крупных артериях падение давления (разница давлений в начале и в конце сосуда) невелико.

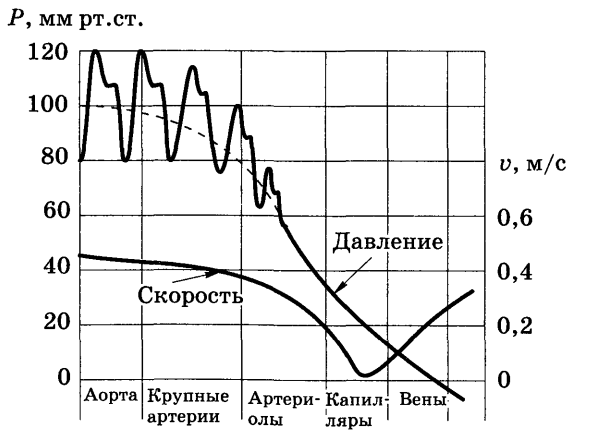


Рис. 9.22. Распределение давления и скорости движения крови в сосудистой системе

В артериолах наблюдается максимальное падение давления, поскольку для совокупности артериол происходит большое увеличение гидравлического сопротивления X [см. формулы (9.10) и (9.11)].

В венах, впадающих в сердце, давление ниже атмосферного. Как уже отмечалось в 9.7, в крупных кровеносных сосудах проявляются пульсовые колебания давления, амплитуда кото-

рых уменьшается с увеличением степени разветвленности сосудистого русла и уменьшением диаметра отдельных сосудов.

Сосудистая система обладает минимальной площадью сечения в области аорты, где наблюдается максимальная амплитуда пульсовых колебаний и наибольшая линейная скорость крови порядка 0,5 м/с (см. рис. 9.22, нижняя кривая). По мере перехода к более мелким кровеносным сосудам суммарная площадь их сечения увеличивается и в соответствии с условиями неразрывности струи (см. 9.1) скорость кровотока в них уменьшается, составляя в капиллярах около 0,5 мм/с. В венозной части сосудистой системы суммарная площадь сечения сосудов уменьшается, что приводит к возрастанию скорости кровотока.

9.9. Некоторые методы определения давления и скорости крови

Прямое измерение давления крови осуществляется введением катетера непосредственно в кровеносный сосуд или полость сердца. Катетер заполняется изотоническим раствором и передает давление крови с вводимого конца на внешний измерительный прибор или устройство автоматизированной обработки данных. Прямая манометрия — практически единственный метод измерения давления в полостях сердца и центральных сосудах. Венозное давление надежно измеряется прямым методом. Основной недостаток прямых измерений очевиден — это необходимость введения измерительных устройств в полость сосуда.

Непрямые измерения давления крови производятся без нарушения целостности сосудов и тканей. В подавляющем большинстве непрямые методы являются ***компрессионными***. Они основаны на уравновешивании давления внутри сосуда измеряемым внешним давлением на его стенку.

Простейшим из таких методов является пальпаторный способ определения систолического артериального давления, предложенный *Рива-Роччи*. При его осуществлении на среднюю часть плеча накладывают компрессионную манжету.

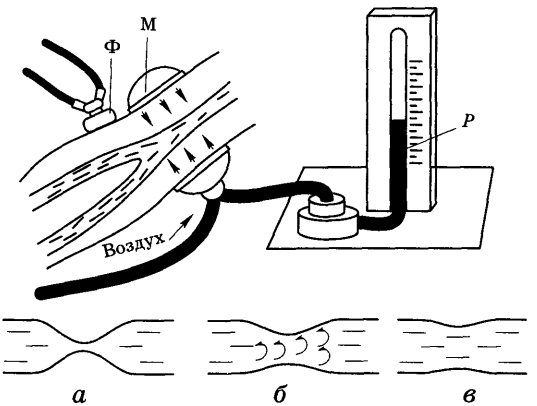


Рис. 9.23. Определение давления крови методом Короткова — Рива-Роччи:
а — тоны Короткова не обнаружены; *б* — систолическое давление; *в* — диастолическое давление

Давление воздуха в манжете измеряется с помощью манометра. Накачиванием воздуха в манжету давление в ней быстро поднимается до значения, превышающего систолическое. Затем воздух из манжеты медленно выпускают, одновременно наблюдая за появлением пульса в лучевой артерии. Зафиксировав пальпаторно появление пульса, отмечают в этот момент значение давления в манжете, которое и принимают за систолическое давление.

Аускультативный метод имеет наибольшее распространение и основан на установлении систолического и диастолического давления по возникновению и исчезновению в артерии особых звуковых явлений — *тонов Короткова*. Так же как и в методе Рива-Роччи, на область плеча накладывается компрессионная манжета М (рис. 9.23), в которую накачивается воздух, создавая давление, большее систолического. В это время (*а*) тоны Короткова не обнаруживаются. При выходе воздуха из манжеты просвет сосуда увеличивается и при равенстве наружного давления систолическому возникают характерные звуки, прослушиваемые с помощью фонендоскопа Ф. В этот момент (*б*) по манометру определяют систолическое давление. Момент исчезновения шумов (*в*) соответствует равенству измеряемого наружного давления диастолическому.

Аускультативный метод реализуется в различных вариантах. В частности, в измерителях давления тоны Короткова могут восприниматься микрофоном, который преобразует звуковые воздействия в электрические сигналы, поступающие на регистрирующее устройство. На цифровом табло регистратора указываются значения систолического и диастолического давления. В некоторых приборах изменения в движении стенок артерии при систолическом и диастолическом давлении (сопровождающиеся возникновением и исчезновением тонов Короткова) определяются с помощью ультразвуковой локации и эффекта Доплера.

Рассмотрим теперь основы некоторых ***физических методов*** определения скорости кровотока.

Метод, основанный на эффекте Доплера. Физическая основа этого метода была рассмотрена в 6.8. Еще раз подчеркнем, что его широкое клиническое применение обусловлено неинвазивностью, высокой точностью и возможностью использования для исследования различных гемодинамических процессов. Поскольку ультразвуковая волна обладает малой длиной и, следовательно, может быть сфокусирована на малую площадь, становится возможным определять скорость крови локально в отдельных узких участках сосудистой системы и даже скорость движения различных слоев крови в отдельном сосуде.

Электромагнитный метод основан на физическом явлении, называемом *эффектом Холла.* Его сущность поясняет рис. 9.24. Пусть в некоторой проводящей электрический ток среде со скоростью v движутся электрические заряды (на рис. 9.24 в направлении слева направо). Если эту среду поместить в магнитное поле с индукцией B,

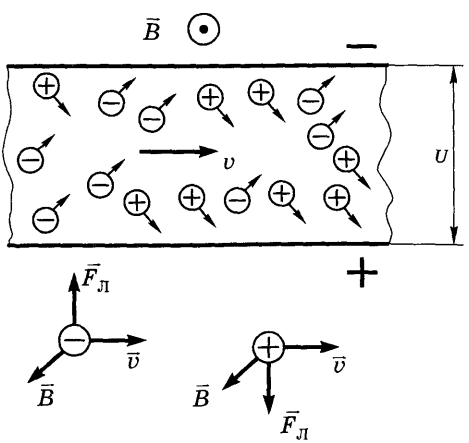


Рис. 9.24. Эффект Холла

направленной перпендикулярно направлению скорости движения зарядов, то на заряды будет действовать сила Лоренца $F_Л = qvB$, где q — величина заряда. Под действием этой силы происходит отклонение положительных зарядов в одну сторону, а отрицательных — в противоположную, как показано на рис. 9.24. Таким образом, из-за происходящего под действием магнитного поля разделения зарядов на поверхности среды возникает разность потенциалов U (холловское напряжение). Ее величина пропорциональна значению силы Лоренца $F_Л$ и скорости движения заряженных частиц v.

Поскольку в крови имеются положительные и отрицательные ионы, этот эффект проявляется и в движущейся по сосуду крови. Из измеряемой разности потенциалов можно определить линейную скорость кровотока.

Методы определения скорости движения крови, основанные на описанном принципе, удобно применять в тех случаях, когда возможен непосредственный контакт электродов соответствующего измерительного устройства со стенками сосуда (при использовании аппаратов искусственного кровообращения, гемосорбции, гемодиализа, в ходе хирургических операций).

9.10. Работа и мощность сердца

В соответствии с законом Пуазейля течение жидкости по сосудам определяется разностью давлений в начале и конце сосудов. В системе кровообращения этот перепад давлений обусловлен в основном работой сердца. Таким образом, по отношению к сосудистой системе сердце можно считать некоторым насосом. В отличие от обычного поршневого насоса, где площадь поршня постоянна, при работе сердца его рабочая поверхность — внутренняя поверхность желудочков — изменяется в различные фазы их сокращения. Сила F, действующая на кровь в желудочках, зависит от давления P в их полости и площади внутренней поверхности S:

$$F = PS.$$

В начале систолы давление меньше, а объем и площадь поверхности полости желудочков — больше. В конце систолы

давление возрастает, а внутренняя поверхность уменьшается. Количественное рассмотрение этих процессов и непосредственные физиологические эксперименты показывают, что при уменьшении объема внутренних полостей сердце развивает меньшую силу. Согласно ***закону Старлинга***, сила, развиваемая при сокращении сердца, пропорциональна начальной длине волокон миокарда.

Не занимаясь рассмотрением сложных явлений, происходящих при сокращении сердца, рассчитаем его работу за одно сокращение исходя из общих физических соображений. Эта работа A складывается из работы левого $A_л$ и правого $A_п$ желудочков:

$$A = A_л + A_п.$$

Работа левого желудочка, выбрасывающего кровь в большой круг кровообращения, естественно, больше, чем работа правого желудочка. Считается, что $A_п = 0{,}2A_л$. Таким образом $A = 1{,}2A_л$.

Работа левого желудочка при выбросе систолического (ударного) объема крови в аорту затрачивается на преодоление сил давления крови в сосудистой системе и на сообщение крови кинетической энергии. Первая составляющая этой работы называется статическим компонентом, а вторая — кинетическим компонентом.

Статический компонент рассчитаем по формуле

$$A_{ст} = P_{ср}V_с,$$

где $P_{ср}$ — давление крови в аорте; $V_с$ — систолический объем.

Считая для человека в покое $P_{ср} = 100$ мм рт.ст. = 13,3 кПа и $V_с = 60$ мл $= 6 \cdot 10^{-5}$ м^3, найдем $A_{ст} \approx 0{,}8$ Дж.

Кинетический компонент может быть рассчитан из формулы для кинетической энергии

$$A_к = \frac{m\upsilon^2}{2} = \frac{\rho V_с \upsilon^2}{2},$$

где m — масса крови в ее систолическом объеме; ρ — плотность крови; υ — линейная скорость крови при выбросе в аорту.

Приняв $\rho = 1{,}05 \cdot 10^3$ кг/м^3, $\upsilon = 0{,}5$ м/с, получим, что $A_к = 0{,}008$ Дж.

Таким образом, работа сердца за одно сокращение определяется по формуле

$$A = 1{,}2\left(PV_c + \frac{\rho V_c v^2}{2}\right). \qquad (9.28)$$

В покое она составляет примерно 1 Дж. Считая, что в среднем сердце сокращается один раз в секунду, определим работу сердца за сутки. Она составляет 86 400 Дж. Для сравнения, такую работу надо совершить, чтобы поднять груз массой в 1 т на высоту около 9 м.

Учитывая, что продолжительность систолы составляет около 0,3 с, и разделив работу сердца за одно сокращение на это время, получим значение средней мощности сердца в покое — 3,3 Вт.

При физической нагрузке работа сердца резко увеличивается, возрастает систолический и минутный объем крови, увеличивается и скорость течения крови в аорте. При этом изменяется и соотношение между статическим и кинетическим компонентами. Как видно из приведенных численных значений, в покое доля кинетического компонента составляет около 1 % общей работы сердца. С повышением нагрузки растут оба слагаемых формулы (9.28), но кинетический компонент растет быстрее, так как он пропорционален квадрату увеличивающейся линейной скорости кровотока и доля кинетического компонента в общей работе сердца может достигать 30 %.

Контрольные вопросы и задания

1. Что такое объемная и линейная скорости движения жидкости? Какова связь между ними?
2. О чем свидетельствует уравнение неразрывности струи? Для каких жидкостей оно применимо?
3. Приведите уравнение Бернулли и охарактеризуйте его составляющие. Для каких жидкостей оно применимо?
4. Основываясь на уравнениях Бернулли и неразрывности струи, проанализируйте изменения давления крови в сосуде при аневризме и при сужении просвета сосуда.

5. Исходя из значений линейной скорости кровотока в аорте и капиллярах оцените соотношение между площадью поперечного сечения аорты и суммарной площадью поперечных сечений капилляров.
6. Укажите особенности течения вязкой жидкости. Приведите формулу Ньютона для сил внутреннего трения в жидкости. Что такое «вязкость» жидкости, в каких единицах она измеряется?
7. В чем различие между ньютоновскими и неньютоновскими жидкостями? Приведите примеры этих жидкостей.
8. Приведите формулу Пуазейля для объема и расхода протекающей жидкости. Как определяется гидравлическое сопротивление сосуда?
9. Охарактеризуйте известные вам методы определения вязкости жидкости, сопоставьте их достоинства и недостатки.
10. Что определяет число Рейнольдса? Запишите его выражение через гидродинамические параметры.
11. Укажите значение вязкости крови в норме и пределы изменения ее значений при патологических процессах.
12. Укажите причины, приводящие к изменению вязкости крови в организме. Сопоставьте вязкость венозной и артериальной крови.
13. Что такое трансмуральное давление? гидростатическое давление? Покажите графически зависимость давления крови от времени в крупных артериях.
14. На каком участке большого круга кровообращения наблюдается наибольшее падение давления крови? Почему?
15. Приведите формулу, описывающую зависимость скорости распространения пульсовой волны от механических свойств и величины просвета сосуда. Укажите приблизительные значения этой скорости в аорте, артериях мышечного типа и в венах.
16. Как и почему изменяется скорость пульсовой волны с возрастом и с повышением артериального давления?
17. В каких участках сосудистой системы течение крови может иметь турбулентный характер и почему? Как обнаруживается турбулентное течение крови?
18. Назовите известные вам методы определения скорости кровотока, укажите их принципиальную основу.
19. Рассчитайте работу сердца в покое за одно сокращение. Какова мощность сердца? Найдите работу сердца за одни сутки.
20. Каково соотношение составляющих работы сердца по преодолению статического давления крови (статический компонент) и по сообщению движения крови (кинетический компонент) в покое и при физической нагрузке?

Глава 10 Физические свойства и функции биологических мембран

Структурной и функциональной единицей живого организма является ***клетка***. Ей присущи все основные жизненные функции. Клетка обменивается с окружающей средой веществом, энергией и информацией. Энергия питательных веществ, поступающих в клетку, расходуется на совершение различного вида работ в ней и на выполнение разнообразных функций. Все эти функции клетки тесно связаны с ее структурой и строением. Изучение структуры клетки позволяет лучше понять ее функциональные возможности.

Роль мембран в строении и функционировании клетки чрезвычайно велика. Как заметил Д. Бернал, только после образования мембраны вокруг клетки мы действительно имеем то, что с полным правом может быть названо организмом. Именно через мембраны происходит обмен веществ в организме. Нарушение структуры мембран ведет к нарушению их функций и, следовательно, к нарушению функционального состояния клетки и организма в целом.

10.1. Строение и физические свойства биологических мембран

Известно, что биологические мембраны состоят в основном из липидов и белков, процентное соотношение которых меняется в зависимости от вида клетки и функционального назначения мембраны.

Липиды в клеточной мембране представлены холестеролом, фосфолипидами и гликолипидами. *Холестерол* содержится главным образом в наружной плазматической мембране клеток млекопитающих. Различные *фосфолипиды* входят в состав всех биологических мембран. Молекулы липидов состоят из двух физически различных частей — полярной головки и неполярных хвостов. Такие молекулы называются амфифильными. Головки *гликолипидов* представляют собой производные сахаров, головки фосфолипидов — остатки

фосфорной кислоты, а хвосты липидов — остатки жирных кислот. Связь между головкой и хвостами осуществляется молекулой глицерина (рис. 10.1).

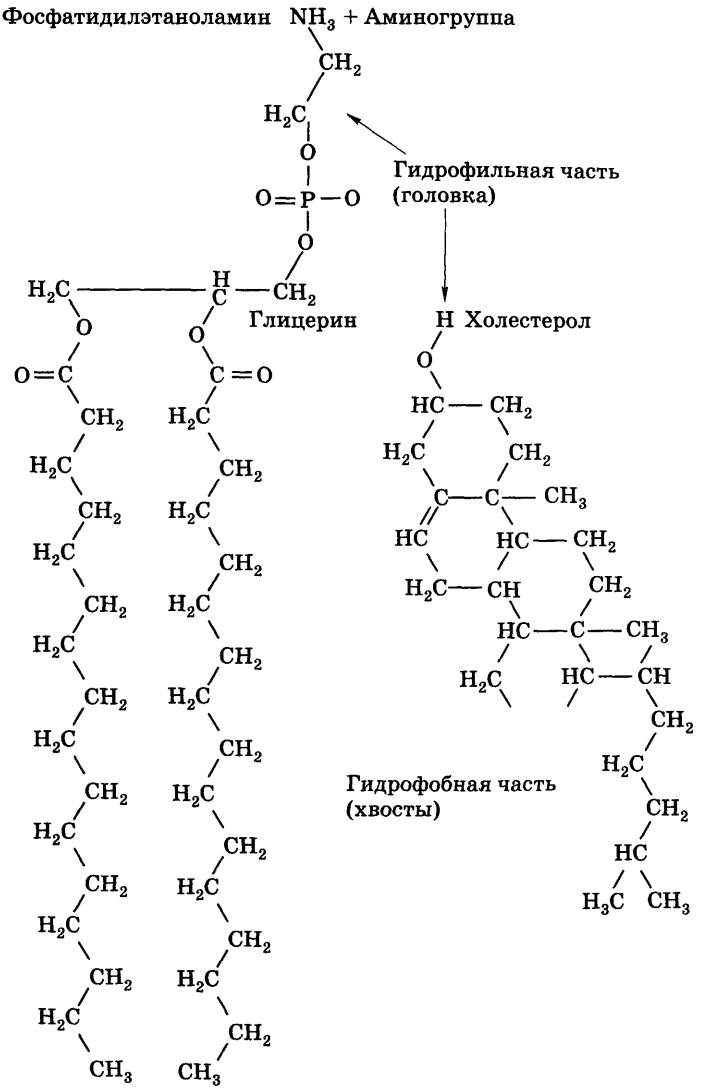


Рис. 10.1. Структурные формулы фосфатидилэтаноламина и холестерола

Головки липидов либо электрически заряжены, либо нейтральны, но имеют дипольный момент, т.е. в любом случае создают вокруг себя электрическое поле. Они хорошо взаимодействуют с полярными растворителями (водой), поэтому головку и называют гидрофильной частью липида. Хвосты, наоборот, неполярны, и хорошо взаимодействуют с неполярными веществами и растворителями, но плохо с водой. Такое поведение неполярных молекул в воде (гидрофобность) обусловлено тем, что им энергетически выгоднее взаимодействовать друг с другом, чем с молекулами воды, вследствие чего они отталкиваются от воды.

В результате амфифильные липиды, нанесенные на поверхность воды, образуют тонкий мономолекулярный слой, в котором головки липидов связаны с водой, а хвосты направлены вверх — подальше от воды (рис. 10.2).

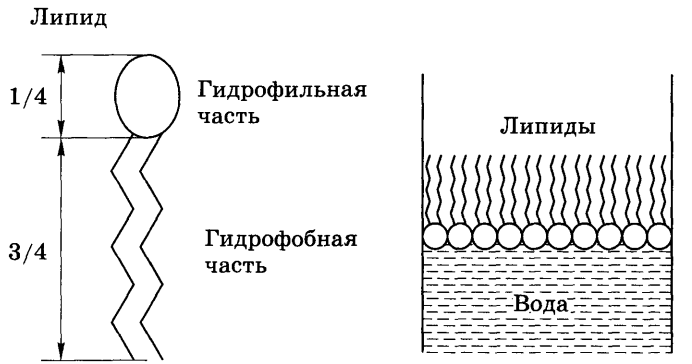


Рис. 10.2. Схема липидной молекулы и поведение липидов на поверхности воды

Если же эти липиды поместить внутрь воды, например, путем интенсивного встряхивания или воздействия ультразвука, то при малых концентрациях они образуют сферические структуры-мицеллы, в которых головки липидов обращены в сторону воды, а хвосты спрятаны внутрь мицелл (рис. 10.3). Подобные образования энергетически выгодны, так как при этом достигается минимум площади контакта гидрофобных участков липидов с водой. По мере увеличения концентрации

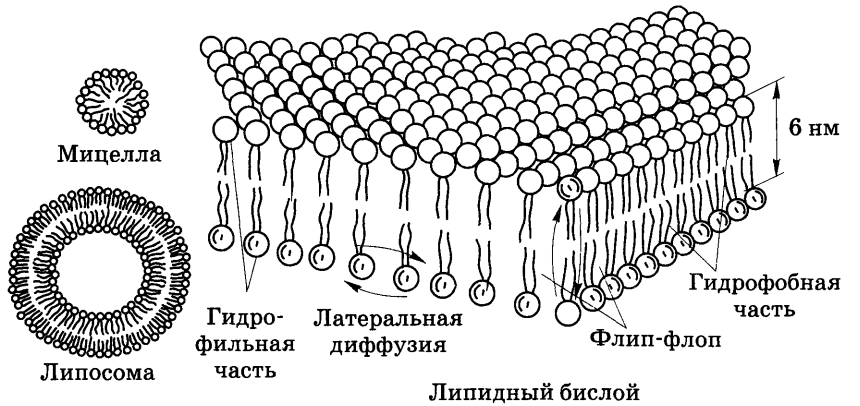


Рис. 10.3. Виды самоорганизации липидов в воде

липидов сначала образуются липосомы (везикулы, пузырьки) из двух слоев липидов, в которых головки обращены в сторону наружной и внутренней воды, а хвосты — друг к другу. Такие образования уже можно рассматривать как искусственные липидные мембраны. И, наконец, при слиянии везикул образуется обширный двойной слой липидов, который и является структурной основой биологической мембраны.

Именно благодаря своей амфифильности липиды мембран самостоятельно организуются в воде в виде двойного слоя, являющегося структурной основой биомембран. При этом полярные головки липидов всегда направлены в сторону воды, а неполярные хвосты — внутрь липидного бислоя, толщина которого составляет 6–7 нм.

Таким образом, самосборка мембран, т.е. образование липидного бислоя со встроенными в него белками, происходит самопроизвольно, вследствие того что такая организация липидов и белков в воде энергетически выгодна.

Гидрофобные взаимодействия липидов и белков в воде играют существенную роль в формировании мембран, во многом определяя процесс их самосборки и представляя собой один из основных факторов их стабилизации.

Кроме липидов, мембраны содержат белки и углеводы.

Белки, связанные с мембранами, подразделяются на два класса: *интегральные* и *периферические*. Интегральные бел-

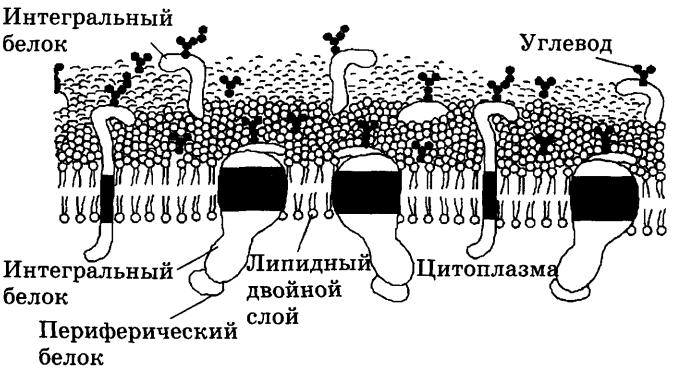


Рис. 10.4. Строение клеточной мембраны

ки погружены в липидный бислой полностью или частично, а периферические находятся на той или другой поверхности мембраны (обычно на ее внутренней стороне).

Углеводы могут быть присоединены либо к белкам, либо к липидам, но в любом случае они неизменно располагаются на внешней, а не на внутренней поверхности мембраны (рис. 10.4) и играют важную роль в межклеточных взаимодействиях. Общая толщина клеточной мембраны 8–9 нм.

Белки имеют на своей поверхности как гидрофильные, так и гидрофобные аминокислотные остатки. Периферические белки можно отделить от мембраны, не нарушая целостности липидного двойного слоя, и после отделения они полностью растворимы в воде, что свидетельствует о гидрофильности их поверхности. Напротив, интегральные белки нельзя удалить, не разрушив двойной слой липидов. Эти белки обычно нерастворимы в воде, так как большая часть находящихся на их поверхности аминокислот гидрофобна. Именно эти участки белка и погружены в гидрофобную зону мембраны, где связываются с хвостами липидов гидрофобными взаимодействиями. Гидрофильные участки поверхности интегрального белка связываются с полярными головками липидов, молекулами воды и периферическими белками.

В заключение отметим, что и липиды, и белки в мембране находятся в постоянном тепловом движении, энергия которого определяется температурой среды T и имеет величину по-

рядка $kT/2$ на одну степень свободы молекулы (здесь k — постоянная Больцмана), поэтому липиды непрерывно меняются местами, перемещаясь вдоль поверхности мембраны со скоростью около 5 мкм/с. Участвуют в этом движении и белки. Это явление хаотического перемещения липидов и белков вдоль поверхности мембраны называется ***латеральной диффузией*** (см. рис. 10.3). Скорость латеральной диффузии белков значительно меньше, чем липидов, из-за их большой массы.

Помимо поступательного движения, мембранные белки и липиды участвуют и во вращательном движении вокруг осей, проходящих через их центр масс, называемом ***вращательной диффузией***. Угловая скорость их вращения весьма велика. При нормальных температурах она составляет для фосфолипидов 10^9 рад/с, а для белков значительно меньше: для родопсина — 10^6 рад/с, для цитохромоксидазы — 10^4 рад/с.

Переход липидов из одного монослоя в другой (движение, называемое *флип-флоп*) маловероятен и происходит очень редко, так как в этом случае полярная головка липида должна пройти через гидрофобную внутреннюю область мембраны, где она нерастворима. Поэтому вероятность таких переходов в 10^{10} раз меньше, чем вероятность латеральной диффузии, а их частота столь мала, что любая данная молекула липида перескакивает из одного монослоя в другой не чаще чем раз в месяц, тогда как вследствие латеральной диффузии липиды меняют свое положение до миллиона раз в секунду.

В организме теплокровных вязкость липидного бислоя мембраны регулируется путем изменения ее липидного состава, например, за счет изменения соотношения между количеством насыщенных и ненасыщенных жирных кислот в мембранных липидах и, кроме того, она растет с увеличением концентрации холестерина в мембранах.

Асимметрия строения мембран проявляется в том, что внутренние и наружные по отношению к клетке стороны любой биологической мембраны всегда имеют разный липидный и белковый состав. Молекулы углеводов располагаются только на внешней стороне мембраны. Асимметричная ориентация ферментативных и транспортных белков в мем-

бране приводит к наличию преимущественного направления активного транспорта веществ через мембрану, что играет исключительно важную роль в функционировании клетки в целом.

10.2. Транспорт молекул и ионов через мембраны

Ряд важнейших процессов в клетке, таких как возбуждение, обмен веществ, поддержание постоянного ионного состава и др., связаны с переносом веществ через клеточную мембрану. Проницаемость мембран для различных веществ — это одно из важнейших свойств мембран, определяющее нормальное функционирование клетки.

Принято различать активный и пассивный перенос (транспорт) молекул и ионов через клеточную мембрану.

Транспорт веществ через клеточную мембрану называется ***пассивным***, если на него не затрачивается метаболическая энергия, запасенная в клетке.

Пассивный транспорт является наиболее распространенным способом переноса молекул и ионов через мембрану. Известны три основных вида *пассивного переноса* вещества через биомембраны:

1) диффузия вещества за счет его растворения в липидном бислое мембраны;

2) диффузия вещества через каналы (поры) в клеточной мембране;

3) облегченная диффузия — перенос вещества через мембрану осуществляется с помощью специальных молекул-переносчиков.

Два первых случая относятся к простой диффузии, когда молекулы диффундирующего вещества движутся через мембрану без образования химического комплекса с другими молекулами.

Как известно, диффузия — это процесс самопроизвольного проникновения вещества из области его большей концентрации в область меньшей концентрации за счет энер-

гии теплового движения. Поэтому основной движущей силой пассивного транспорта является градиент концентрации (точнее, градиент электрохимического потенциала) на мембране.

10.2.1. Диффузия путем растворения в липидном бислое

Этод вид простой диффузии существенно зависит от растворимости диффундирующего вещества в липидном бислое, который по толщине на 75–80 % состоит из неполярных гидрофобных соединений (хвостов липидов) и лишь на поверхности мембраны расположены полярные гидрофильные головки.

В липидах мембран хорошо растворяются неполярные (гидрофобные) вещества, т.е. вещества, не обладающие ни дипольным моментом, ни зарядом, поэтому проницаемость мембраны для неполярных органических соединений очень велика. При этом через мембрану могут проходить даже большие белки, если они не обладают значительным дипольным моментом.

Органические соединения, имеющие полярные группы (гидроксильную, карбоксильную и аминную), значительно хуже растворяются в мембранных липидах, поэтому проницаемость мембран для них существенно меньше. Эти особенности проницаемости клеточных мембран для органических соединений известны как ***правила Овертона***:

1. Проницаемость мембран для органических молекул уменьшается по мере увеличения в них числа полярных групп (гидроксильных, карбоксильных и аминных).

2. Проницаемость мембран для органических молекул возрастает по мере увеличения в них числа неполярных групп (метиловых, этиловых и фенильных).

Многие лекарственные препараты, а также токсины, действующие на внутриклеточные процессы, представляют собой гидрофобные, т.е. неполярные, вещества, благодаря чему они способны быстро проходить в клетку и столь же быстро вызывать соответствующий эффект.

10.2.2. Диффузия через каналы в мембране

Диффузия через каналы в мембране также вид простой диффузии. Неорганические полярные молекулы и ионы в липидах нерастворимы, поэтому они могут проникать через мембраны только при наличии в них специальных каналов (рис. 10.5). Однако число таких каналов сравнительно невелико, поэтому проницаемость мембран для ионов и полярных соединений в сотни раз хуже, чем для неполярных соединений.

Проницаемость каналов весьма избирательна (селективна). В клеточной мембране выделяют каналы натриевые, калиевые, кальциевые, хлорные и др., предназначенные для пропуска только соответствующих ионов. Недавно были открыты и мембранные каналы для воды — *аквапорины*. Все эти каналы играют чрезвычайно большую роль в транспорте полярных соединений и ионов через мембрану и во всей функциональной деятельности клетки. В связи с этим часто говорят, что клеточная мембрана для неполярных соединений служит раствори-



Рис. 10.5. Схема строения ионного канала

телем, а для полярных соединений и ионов — молекулярным ситом.

С функциональной точки зрения все ионные каналы характеризуются *селективностью*, т.е. избирательной проницаемостью, и наличием *воротного механизма*.

Воротный механизм открывает и закрывает канал, изменяя тем самым его проницаемость для иона. Это свойство каналов очень важно для нормального функционирования живой клетки, в частности для установления равновесного мембранного потенциала, генерации и проведения потенциала действия и других процессов.

У одних каналов воротный механизм является *потенциалзависимым*, т.е. они открываются либо закрываются в зависимости от величины мембранного потенциала (например, натриевый и калиевый каналы), другие каналы являются *лиганд-зависимыми* и их «ворота» управляются путем связывания определенных химических молекул-посредников.

10.2.3. Облегченная диффузия

Облегченная диффузия занимает в пассивном транспорте веществ через мембрану особое место. В этом случае вещество тоже перемещается из области высокой его концентрации в область с более низкой концентрацией, но скорость переноса значительно выше, чем при простой диффузии. Этот вид транспорта носит наиболее избирательный характер и обусловлен наличием в мембране специальных веществ — переносчиков, которые образуют с транспортируемыми молекулами химический комплекс и облегчают их перенос через мембрану. Молекула-переносчик может связываться только с определенным ионом или пространственным изомером переносимой молекулы и транспортировать его через мембрану и никогда не будет транспортировать другие ионы или другие изомеры. Например, *D*-глюкоза и многие *L*-аминокислоты проходят через мембрану путем облегченной диффузии, тогда как их соответствующие *L*- и *D*-изомеры вынуждены идти путем простой диффузии.

Известны два вида облегченной диффузии:

1) с помощью подвижного переносчика;

2) эстафетный перенос.

В первом случае молекула переносчика объединяется с транспортируемым веществом на одной стороне мембраны и вместе с ним движется через липидный бислой на другую сторону мембраны. Примером подвижных переносчиков являются, в частности, ионофорные антибиотики, осуществляющие таким путем транспорт определенного вида ионов через мембрану. Например, валиномицин осуществляет транспорт только ионов калия, захватывая их на одной стороне мембраны и транспортируя на другую сторону.

Во втором случае, при эстафетном переносе, молекулы переносчика не совершают челночных движений в мембране, а встраиваются в мембрану друг за другом, образуя как бы мостик в ней. Захватив нужное вещество, крайняя молекула переносчика передает его соседней молекуле, последняя — следующей и так далее по «эстафете», в результате чего вещество транспортируется через мембрану. Примером эстафетного переносчика служит антибиотик грамицидин, транспортирующий таким путем ионы натрия.

Весьма характерной особенностью облегченной диффузии является *насыщение* этого вида транспорта при больших градиентах концентрации переносимого вещества на мембране. В качестве примера рассмотрим зависимость относительной скорости переноса глюкозы от разности ее концентраций на мембране при облегченной диффузии (*D*-глюкоза) и при простой диффузии (*L*-глюкоза) (рис. 10.6). При одинаковой разности концентраций ΔC скорость переноса при облегченной диффузии значительно выше, чем при простой диффузии. При простой диффузии скорость переноса растет прямо пропорционально ΔC. При облегченной диффузии она сначала тоже линейно и очень быстро растет с увеличением ΔC, однако затем рост замедляется и, достигнув максимального значения, больше не меняется, т.е. достигает насыщения. Изменить максимальную скорость переноса при облегченной диффузии можно только путем изменения количества переносчиков в мембране.

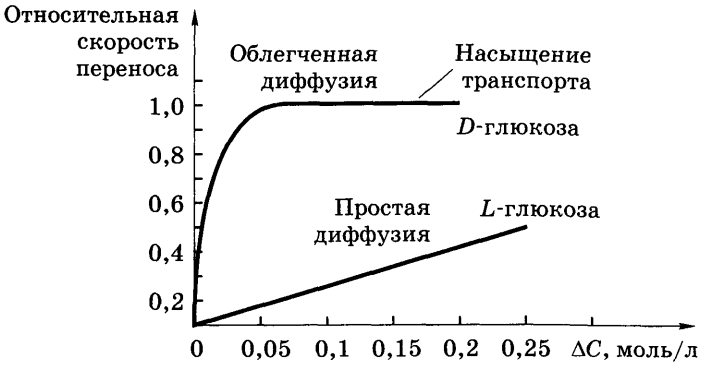


Рис. 10.6. Зависимость скорости переноса глюкозы от разности ее концентраций на мембране при простой (*L*-изомер) и облегченной (*D*-изомер) диффузии

10.3. Математическое описание пассивного транспорта

Рассмотрим теперь математические соотношения, описывающие пассивный транспорт через биомембрану.

С точки зрения термодинамики способность какой-либо физико-химической системы к совершению полезной работы определяется значением ее функции состояния, которая называется ***свободной энергией***. Термодинамический потенциал, представляющий собой свободную энергию одного моля вещества, принято называть ***электрохимическим потенциалом*** μ.

Вся полезная работа A, совершающаяся в 1 моле вещества, осуществляется за счет уменьшения ее электрохимического потенциала (если нет постороннего источника энергии): $A = -\Delta\mu$. Для растворов веществ значение электрохимического потенциала описывается выражением

$$\mu = \mu_0 + RT\ln C + ZF\varphi, \qquad (10.1)$$

где μ_0 — часть химического потенциала 1 моля раствора, определяемая энергией химической связи вещества с растворителем и связанная с растворимостью данного вещества в воде; R — универсальная газовая постоянная; T — абсолют-

ная температура среды; C — молярная концентрация растворенного вещества; Z — заряд растворенных ионов, выраженный в единицах элементарного заряда; F — число Фарадея; φ — электрический потенциал среды.

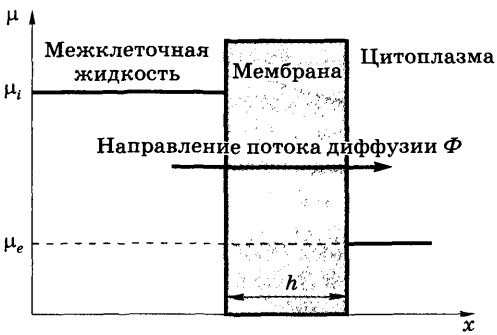


Рис. 10.7. Связь направления потока диффузии и распределения электрохимического потенциала на мембране

Если значения электрохимического потенциала по обе стороны мембраны различны, то система термодинамически неравновесна. При этом на мембране возникает градиент электрохимического потенциала: $d\mu/dx = \Delta\mu/h$, где h — толщина мембраны (рис. 10.7). Любая замкнутая система, предоставленная самой себе, всегда стремится к состоянию термодинамического равновесия, которое характеризуется равенством всех термодинамических, в том числе и электрохимических, потенциалов во всех областях системы, поэтому при равновесии $\mu_e = \mu_i$.

Процесс перехода системы из неравновесного состояния в равновесное в рассматриваемом случае всегда сопровождается пассивным переносом вещества (диффузией) из области большего значения электрохимического потенциала в область с его меньшим значением, пока эти потенциалы не выравняются.

Математически процесс переноса вещества описывается ***уравнением Теорелла***

$$\vec{\Phi} = -CU\frac{\overrightarrow{d\mu}}{dx}, \qquad (10.2)$$

которое показывает, что плотность потока диффузии Φ через поверхность мембраны (т.е. количество вещества, переносимое за 1 с через единицу площади мембраны) прямо пропорциональна молярной концентрации этого вещества C, его

подвижности U и градиенту электрохимического потенциала $d\mu/dx$ на мембране[1].

Из этого уравнения также следует, что по мере выравнивания электрохимических потенциалов по обе стороны мембраны, т.е. по мере уменьшения градиента $d\mu/dx$, уменьшается и поток вещества через мембрану, а при достижении состояния равновесия ($\mu_e = \mu_i$ и, следовательно, $d\mu/dx = 0$) преимущественный поток вещества полностью прекращается: $\Phi = 0$.

Используя выражение (10.1), нетрудно найти производную $d\mu/dx$. Поскольку по обе стороны биомембран растворитель всегда один и тот же — вода, то $\mu_{0i} = \mu_{0e} = \text{const}$. В то же время концентрация C и электрические потенциалы φ по обе стороны мембраны в общем случае различны, поэтому, дифференцируя уравнение (10.1) по координате x, получим

$$\frac{d\mu}{dx} = RT\frac{1}{C}\frac{dC}{dx} + ZF\frac{d\varphi}{dx}. \qquad (10.3)$$

Подставляя это выражение в уравнение (10.2), приходим к уравнению Нернста — Планка, описывающему диффузию ионов через мембрану:

$$\vec{\Phi} = -URT\frac{\overrightarrow{dC}}{dx} - CUZF\frac{\overrightarrow{d\varphi}}{dx}. \qquad (10.4)$$

Первое слагаемое в уравнении (10.4) описывает обычную диффузию, идущую за счет градиента концентрации dC/dx на мембране, а второе слагаемое — электродиффузию, которая обусловлена действием на ионы электрического поля $E = d\varphi/dx$, создаваемого на мембране градиентом электрического потенциала.

При диффузии незаряженных частиц ($Z = 0$) второе слагаемое в уравнении (10.4) обращается в нуль, и пассивный

[1] Равенства (10.2), (10.4) и (10.5) векторные, поэтому знак «–» указывает на то, что направление потока диффузии всегда противоположно направлению вектора градиента. Вектор градиента всегда направлен в сторону наибыстрейшего возрастания той величины, от которой берется производная. Если рассматривать только модули потоков диффузии, то знак «–» в этих уравнениях следует опустить.

транспорт таких веществ описывается обычным законом диффузии — ***законом Фика***:

$$\vec{\Phi} = -D\frac{\overrightarrow{dC}}{dx}. \qquad (10.5)$$

Величина D обычно называется ***коэффициентом диффузии***, и поскольку $D = URT$, то сразу видна его зависимость от подвижности U вещества и абсолютной температуры T среды.

Уравнение Фика (10.5) можно упростить, если приближенно выразить градиент концентрации через толщину h мембраны и абсолютные значения концентрации C_i и C_e вещества по обе ее стороны. Тогда $dC/dx \approx \Delta C/\Delta x = |C_i - C_e|/h$ и для величины плотности потока диффузии получаем уравнение

$$\Phi = p|C_i - C_e| = p\Delta C, \qquad (10.6)$$

где коэффициент $p = D/h$ называют ***коэффициентом проницаемости*** или просто ***проницаемостью*** мембраны.

Следует подчеркнуть, что проницаемость мембраны различна для разных веществ; является одним из важнейших параметров, определяющих диффузию молекул и ионов через мембрану, и зависит от температуры среды и толщины самой мембраны.

Таким образом, пассивный транспорт ионов описывается уравнением Нернста — Планка (10.4), а пассивный транспорт незаряженных частиц — уравнением Фика (10.5) либо (10.6).

10.4. Активный транспорт ионов

Перенос вещества через мембрану в направлении возрастания электрохимического потенциала, осуществляемый клеткой за счет энергии метаболических процессов, называется ***активным транспортом***.

Напомним, что пассивный транспорт веществ всегда идет из области больших значений электрохимического потенциала в область его меньших значений, в результате чего градиент электрохимического потенциала уменьшается. Активный же транспорт идет в противоположном направлении

и ведет к увеличению разности электрохимических потенциалов по обе стороны мембраны. Однако такой процесс не может идти самопроизвольно, т.е. без затрат энергии извне, от других источников. Источником этой энергии в клетке является молекула аденозинтрифосфата (АТФ), а активный транспорт осуществляют специальные белки, катализирующие расщепление молекулы АТФ на аденозиндифосфат (АДФ) и фосфатную группу (Ф): АТФ → АДФ + Ф + *E*. Поэтому они называются ***транспортными АТФ-азами*** или ***насосами***. Освобождаемая при расщеплении АТФ энергия *E* как раз и затрачивается на совершение работы по активному транспорту.

Активный транспорт веществ принято делить на два вида:

1) активный транспорт ионов;

2) активный транспорт органических веществ, главным образом аминокислот и сахаров.

Рассмотрим подробнее *активный транспорт ионов*. В настоящее время известны четыре основные системы активного транспорта ионов в живой клетке. Три из них обеспечивают перенос ионов Na^+, K^+, Ca^{2+}, H^+ через биологические мембраны — это соответственно Na^+–K^+–АТФ-аза, Ca^{2+}–АТФ-аза, H^+–АТФ-аза; четвертый механизм — это активный перенос протонов при работе дыхательной цепи митохондрий.

Поскольку все механизмы активного транспорта совершают работу по переносу соответствующих ионов через мембрану против направления действия электрохимического градиента, то их часто называют насосами или помпами. Следует иметь в виду, что детально механизмы активного транспорта до конца еще не изучены ввиду их большой сложности. Считается, что наиболее сложно устроена протонная помпа, т.е. H^+–АТФ-аза дыхательной цепи митохондрий, а наиболее просто — Ca^{2+}–АТФ-аза.

Механизм активного транспорта рассмотрим в общих чертах на примере Na^+–K^+-насоса. По современным представлениям его работа протекает следующим образом. Под действием ионов Na^+, находящихся на внутренней стороне мембраны, белок-переносчик, т.е. транспортная АТФ-аза, активируется и расщепляет молекулу АТФ на АДФ и Ф с выделением энер-

гии $E = 45$ кДж/моль. При этом происходит присоединение к транспортному белку трех ионов Na^+.

Эти три иона Na^+ затем перемещаются белком-переносчиком на внешнюю сторону мембраны и там освобождаются (рис. 10.8). Вместо них белок-переносчик захватывает на наружной стороне мембраны два иона K^+ и транспортирует их в обратном направлении (внутрь клетки). Отметим, что и Na^+ и K^+ переносятся из области их меньшей концентрации в область их большей концентрации, т.е. в направлении, противоположном их пассивному транспорту.



Рис. 10.8. Схема действия Na^+–K^+-насоса

Поскольку за один цикл действия Na^+ – K^+-насоса из клетки переносится наружу 3 иона Na^+, а внутрь — только 2 иона K^+, то клетка теряет один положительный заряд, ее внутренняя часть заряжается отрицательно, а внешняя — положительно, что приводит к возникновению разности потенциалов на клеточной мембране. Такие насосы, в результате своей работы создающие разность потенциалов на мембране, называют ***электрогенными***. К этой группе относятся все перечисленные выше ионные насосы.

Работа, осуществляемая Na^+–K^+-насосом за один цикл, равна 41,2 кДж/моль. Энергия, освобождаемая при фосфорилировании АТФ в реальных условиях, — 45 кДж/моль, так что практически вся энергия гидролиза АТФ тратится на активный перенос Na^+ и K^+ (с КПД до 92 %).

Известно, что концентрация ионов калия внутри клетки значительно выше, чем в межклеточной жидкости, тогда как распределение ионов натрия противоположно. Такое распределение этих ионов характерно для всех клеток и поддерживается работой $Na^+ - K^+$-насоса, компенсирующего пассивную диффузию этих ионов через мембрану в состоянии покоя, а также восстанавливающего концентрационные градиенты натрия и калия на мембране после генерации потенциала действия.

В заключение отметим, что как пассивный, так и активный транспорт веществ существенно зависят от вязкости липидного бислоя мембран, поскольку с увеличением вязкости снижается подвижность веществ в мембране и уменьшается активность ферментов и транспортных белков.

Контрольные вопросы и задания

1. Охарактеризуйте физические свойства мембранных липидов и их роль в строении и свойствах биологических мембран, роль мембранных белков.
2. Опишите виды движения липидов и белков в мембране (латеральная диффузия, флип-флоп, вращательная диффузия).
3. Каковы виды пассивного транспорта веществ через мембрану? В чем состоят правила Овертона?
4. Какие вещества диффундируют через каналы в мембране? Укажите свойства мембранных каналов.
5. В чем состоит облегченная диффузия? Каковы ее виды и свойства?
6. Каков смысл электрохимического потенциала? Что описывают уравнения Теорелла, Нернста–Планка, Фика? Что такое проницаемость мембран?
7. Что такое активный транспорт? Объясните механизм активного транспорта ионов на примере натрий-калиевого насоса.

Глава 11 Мембранные потенциалы клетки

На мембране, разделяющей цитоплазму и межклеточную жидкость, существует разность электрических потенциалов, которую называют ***мембранным потенциалом.***

Для покоящейся живой клетки эта разность потенциалов называется ***потенциалом покоя клетки.*** Потенциал внутри клетки относительно межклеточной жидкости составляет в покое от –60 мВ до –100 мВ, в зависимости от вида клетки.

В процессе деятельности клетки мембранный потенциал может изменяться, и эти изменения в нервных и рецепторных клетках непосредственно связаны с переработкой и передачей информации, а в мышечных волокнах — с их сокращением.

11.1. Равновесные потенциалы Нернста

Предположим для простоты, что мембрана в покое проницаема только для одного вида ионов, и найдем равновесный мембранный потенциал, возникающий вследствие диффузии этих ионов через мембрану, учитывая, что равновесное состояние достигается при равенстве электрохимических потенциалов по обе стороны мембраны $\mu_i = \mu_e$.

Внутри клетки электрохимический потенциал

$$\mu_i = \mu_{0i} + RT\ln C_i + ZF\varphi_i,$$

а вне ее

$$\mu_e = \mu_{0e} + RT\ln C_e + ZF\varphi_e.$$

Так как с обеих сторон мембраны ионы находятся в одном растворителе — воде, то $\mu_{0i} = \mu_{0e}$ и условие термодинамического равновесия принимает вид

$$RT\ln C_i + ZF\varphi_i = RT\ln C_e + ZF\varphi_e,$$

или

$$ZF(\varphi_i - \varphi_e) = RT(\ln C_i - \ln C_e).$$

Из последнего соотношения получаем уравнение Нернста для равновесного мембранного потенциала:

$$\varphi_i - \varphi_e = -\frac{RT}{ZF}\ln\frac{C_i}{C_e}. \qquad (11.1)$$

Из него следует, что величина равновесного потенциала для любого иона пропорциональна температуре T клетки и логарифму отношения концентраций этого иона внутри и вне клетки. Потенциал среды, окружающей клетку, обычно принимают равным нулю (т.е. $\varphi_e = 0$), поэтому мембранный потенциал представляет собой потенциал внутри клетки по отношению к межклеточной жидкости. Например, для большинства клеток в нормальных условиях концентрация ионов калия внутри значительно выше, чем вне ее, вследствие чего потенциал Нернста по калию отрицательный. Для Na^+ и Cl^-, наоборот, концентрация вне клетки значительно выше, чем внутри, поэтому равновесный потенциал Нернста по натрию положительный, а по хлору — отрицательный (из-за отрицательного заряда этого иона, см. табл. 11.1).

Таблица 11.1

Клетка	Ионы	Межклеточная жидкость, мМ/л	Цитоплазма, мМ/л	Потенциал Нернста, *mV*	Потенциал покоя, φ_0, *mV*
Мышцы млекопитающих	Na^+	142	12	+64	−90
	K^+	4	140	−92	
	Cl^-	120	4	−89	
Мышцы лягушки	Na^+	120	9,2	+67	−90
	K^+	2,5	140	−105	
	Cl^-	120	3,5	−92	
Аксон кальмара	Na^+	460	50	+58	−70
	K^+	10	400	−96	
	Cl^-	540	40	−68	

11.2. Потенциал покоя клетки

Установлено, что наиболее важными ионами, определяющими мембранные потенциалы клеток, являются неорганические ионы K^+, Na^+, Cl^-, а также в ряде случаев Ca^{2+}. Хоро-

шо известно, что концентрации этих ионов в цитоплазме и в межклеточной жидкости различаются в десятки раз.

Из табл. 11.1 видно, что концентрация ионов K^+ внутри клетки в 40–60 раз выше, чем в межклеточной жидкости, тогда как для Na^+ и Cl^- распределение концентраций противоположное. Неравномерное распределение концентраций этих ионов по обе стороны мембраны обеспечивается как их различной проницаемостью, так и сильным электрическим полем мембраны, которое определяется ее потенциалом покоя.

Действительно, в состоянии покоя суммарный поток ионов через мембрану равен нулю, и тогда из уравнения Нернста — Планка следует, что

$$\frac{RT}{C}\overrightarrow{\frac{dC}{dx}} = -ZF\overrightarrow{\frac{d\varphi}{dx}}. \qquad (11.2)$$

Таким образом, в покое градиенты концентрации $\overrightarrow{\frac{dC}{dx}}$ и электрического потенциала $\overrightarrow{\frac{d\varphi}{dx}}$ на мембране направлены противоположно друг другу и поэтому в покоящейся клетке высокая и постоянная разность концентраций основных ионов обеспечивает поддержание на мембране клетки электрического напряжения, которое и называют ***равновесным мембранным потенциалом.***

В свою очередь возникающий на мембране потенциал покоя препятствует выходу ионов из клетки K^+ и чрезмерному входу в нее Cl^-, поддерживая тем самым их концентрационные градиенты на мембране.

Полное выражение для мембранного потенциала, учитывающее потоки диффузии этих трех видов ионов, было получено Гольдманом, Ходжкиным и Катцем:

$$\varphi_м = -\frac{RT}{F}\ln\frac{P_K C_i(K^+) + P_{Na} C_i(Na^+) + P_{Cl} C_e(Cl^-)}{P_K C_e(K^+) + P_{Na} C_e(Na^+) + P_{Cl} C_i(Cl^-)}, \qquad (11.3)$$

где P_K, P_{Na}, P_{Cl} — проницаемость мембраны для соответствующих ионов.

Уравнение (11.3) с высокой точностью определяет мембранные потенциалы покоя различных клеток. Из него следует, что для мембранного потенциала покоя важны не абсолютные величины проницаемостей мембраны для различных ионов, а их отношения, так как, разделив обе части дроби под знаком логарифма, например, на P_K, мы перейдем к относительным проницаемостям ионов.

В тех случаях, когда проницаемость одного из этих ионов значительно больше, чем других, уравнение (11.3) переходит в уравнение Нернста (11.1) для этого иона.

Из табл. 11.1 видно, что мембранный потенциал покоя клеток близок к потенциалу Нернста для ионов K^+ и Cl^-, но значительно отличается от него по Na^+. Это свидетельствует о том, что в покое мембрана хорошо проницаема для ионов K^+ и Cl^-, тогда как для ионов Na^+ ее проницаемость очень низка.

Несмотря на то что равновесный потенциал Нернста для Cl^- наиболее близок к потенциалу покоя клетки, последний имеет преимущественно калиевую природу. Это обусловлено тем, что высокая внутриклеточная концентрация K^+ не может существенно уменьшиться, так как ионы K^+ должны уравновешивать внутри клетки объемный отрицательный заряд анионов. Внутриклеточные анионы представляют собой в основном крупные органические молекулы (белки, остатки органических кислот и т.п.), которые не могут пройти через каналы в клеточной мембране. Концентрация этих анионов в клетке практически постоянна и их суммарный отрицательный заряд препятствует значительному выходу калия из клетки, поддерживая вместе с Na–K-насосом его высокую внутриклеточную концентрацию[1]. Однако основная роль в первоначальном установлении внутри клетки высокой концентрации ионов калия и низкой концентрации ионов натрия принадлежит Na–K-насосу.

[1] Поскольку внутриклеточная концентрация калия поддерживается почти постоянной, то даже относительно небольшие изменения внеклеточной концентрации K^+ могут оказывать заметное влияние на потенциал покоя и на деятельность клетки. Подобные изменения концентрации K^+ в плазме крови происходят при некоторых патологиях (например, при почечной недостаточности).

Распределение ионов Cl^- устанавливается в соответствии с мембранным потенциалом, поскольку в клетке нет специальных механизмов поддержания концентрации Cl^-. Поэтому вследствие отрицательного заряда хлора его распределение оказывается обратным по отношению к распределению калия на мембране (см. табл. 11.1). Таким образом, концентрационные диффузии K^+ из клетки и Cl^- в клетку практически уравновешиваются мембранным потенциалом покоя клетки.

Что касается Na^+, то в покое его диффузия направлена в клетку под действием как градиента концентрации, так и электрического поля мембраны и вход Na^+ в клетку ограничивается в покое только малой проницаемостью мембраны для натрия (закрыты натриевые каналы). Действительно, Ходжкин и Катц экспериментально установили, что в состоянии покоя проницаемости мембраны аксона кальмара для K^+, Na^+ и Cl^- относятся как 1 : 0,04 : 0,45. Таким образом, в состоянии покоя клеточная мембрана малопроницаема только для Na^+, а для Cl^- она проницаема почти так же хорошо, как и для K^+. В нервных клетках проницаемость для Cl^- обычно ниже, чем для K^+, но в мышечных волокнах проницаемость для Cl^- даже несколько преобладает.

Несмотря на малую проницаемость клеточной мембраны для Na^+ в покое, существует, хотя и весьма малый, пассивный перенос Na^+ в клетку. Этот ток Na^+ должен был бы приводить к снижению разности потенциалов на мембране и к выходу K^+ из клетки, что вело бы в конечном итоге к выравниванию концентраций Na^+ и K^+ по обе стороны мембраны. Этого не происходит благодаря работе $Na^+ - K^+$-насоса, компенсирующего токи утечки Na^+ и K^+ и поддерживающего таким образом нормальные значения внутриклеточных концентраций этих ионов и, следовательно, нормальную величину потенциала покоя клетки.

Для большинства клеток мембранный потенциал покоя составляет (–60)–(–100) мВ. На первый взгляд может показаться, что это малая величина, но надо учесть, что толщина

мембраны тоже мала (8–10 нм), так что напряженность электрического поля в клеточной мембране огромна и составляет около 10 млн вольт на 1 м (или 100 кВ на 1 см):

$$E = \frac{\varphi_0}{d} \approx \frac{80\ \text{мВ}}{8\ \text{нм}} = \frac{8 \cdot 10^{-2}\ \text{В}}{8 \cdot 10^{-9}\ \text{м}} = 10^7\ \frac{\text{В}}{\text{м}} = 10^5\ \frac{\text{В}}{\text{см}}. \quad (11.4)$$

Воздух, например, не выдерживает такой напряженности электрического поля (электрический пробой в воздухе наступает при 30 кВ/см), а мембрана выдерживает. Это нормальное условие ее деятельности, поскольку именно такое электрическое поле необходимо для поддержания разности концентраций ионов натрия, калия и хлора на мембране.

Величина потенциала покоя, различная у клеток, может изменяться при изменении условий их жизнедеятельности. Так, нарушение биоэнергетических процессов в клетке, сопровождающееся падением внутриклеточного уровня макроэргических соединений (в частности, АТФ), прежде всего исключает компоненту потенциала покоя, связанную с работой Na^+–K^+–АТФ-азы.

Повреждение клетки приводит обычно к повышению проницаемости клеточных мембран, в результате чего различия в проницаемости мембраны для ионов калия и натрия уменьшаются; потенциал покоя при этом уменьшается, что может вызвать нарушение ряда функций клетки, например возбудимости.

11.3. Возбуждение клетки. Генерация потенциала действия

Способность организма, органа, ткани или клетки отвечать на раздражение активной специфической реакцией (генерацией нервного импульса, сокращением, секрецией и т.п.) называется ***возбудимостью***.

Все клетки возбудимых тканей (к ним относятся нервная и мышечная) при действии раздражителей достаточной силы способны переходить в состояние возбуждения.

Процесс воздействия на живой объект внешних факторов называют ***раздражением*** или ***стимуляцией***. Факторы внешней среды (давление, звук, свет, тепло, холод и т.п.), способные вызвать переход биосистемы в активное состояние, называют ***раздражителями***.

Минимальная сила раздражителя, необходимая для возникновения возбуждения, называется ***порогом возбуждения***. Величина этого порога является количественной мерой возбудимости тканей.

Действие раздражителя в конечном итоге приводит к изменению мембранного потенциала клетки на некоторую величину U, зависящую от силы раздражителя, в результате чего потенциал на мембране изменяется и становится равным $\varphi_м = \varphi_0 + U$, где φ_0 — потенциал покоя клетки. Если под действием раздражителя мембранный потенциал становится выше (положительнее при $U > 0$) потенциала покоя, то такое изменение мембранного потенциала называют ***деполяризацией***, так как разность потенциалов на мембране (т.е. поляризация мембраны) уменьшается по абсолютной величине. Если же мембранный потенциал становится ниже (отрицательнее при $U < 0$) потенциала покоя, то разность потенциалов на мембране увеличивается по модулю и такое состояние называют ***гиперполяризацией***.

Возбуждение клетки происходит только при деполяризации ее мембраны с помощью внешнего раздражителя (стимула) до определенного значения, которое называют ***критическим потенциалом*** $E_{кр}$. Если сила раздражителя ниже порога возбуждения, то $\varphi_м < E_{кр}$ и такой раздражитель называется ***подпороговым***. В этом случае возбуждение клетки не происходит, но развивается так называемый *локальный* ответ, проявляющийся в небольшом, пропорциональном силе раздражителя, изменении мембранного потенциала клетки, не достигающем значения $E_{кр}$.

11.3.1. Генерация потенциала действия

Если сила раздражителя равна порогу возбуждения или превосходит его, то $\varphi_м \geq E_{кр}$ и клетка возбуждается. Такой раздражитель либо стимул называют соответственно ***пороговым*** или ***надпороговым***.

По достижении критического потенциала $E_{кр}$ на мембране дальнейшее изменение мембранного потенциала в возбуждаемой клетке протекает уже независимо от внешнего раздражителя (стимула) и всегда вполне определенным образом. Клетка отвечает на действие такого стимула физическими, физико-химическими и функциональными изменениями, происходящими в ней, при этом обязательно вполне определенным образом изменяется во времени ее мембранный потенциал.

Это изменение во времени мембранного потенциала клетки, происходящее при ее возбуждении, называется ***потенциалом действия***.

Рассмотрим процессы, которые происходят на клеточной мембране при возбуждении начиная с того момента, когда потенциал на ее мембране под действием раздражителя достиг критического значения $E_{кр}$.

Основным фактором, определяющим дальнейшее автоматическое изменение мембранного потенциала клетки, является увеличение проницаемости натриевых каналов. Напомним, что концентрация Na^+ снаружи клетки в десятки раз больше, чем внутри, и, кроме того, в покое потенциал внутри клетки отрицательный, поэтому Na^+ стремится войти в клетку, но в покое этому препятствует низкая проницаемость натриевых каналов. При достижении потенциала $E_{кр}$ натриевые каналы в мембране начинают открываться и ионы Na^+ начинают быстро диффундировать в клетку под действием как градиента концентрации, так и электрического поля. Возникающий при этом поток ионов Na^+ в клетку еще больше увеличивает деполяризацию мембраны, ее проницаемость для натрия резко возрастает и происходит лавинообразное нарастание потока ионов Na^+ внутрь клетки. Благодаря этому потоку в клетку положительных ионов, первоначальный отрицательный потенциал внутри нее быстро повышается, достигает нуля,

а затем становится положительным из-за продолжающегося потока ионов натрия в клетку (рис. 11.1).

По современным представлениям натриевый канал имеет двое ворот, работающих в противофазе и отличающихся своими инерционными свойствами: быстрые m-ворота и более инерционные h-ворота. В покое m-ворота закрыты, а h-ворота открыты. При деполяризации мембраны до $E_{кр}$ легкие m-ворота быстро открываются, а более инерционные h-ворота начинают закрываться и в течение этого времени Na^+ входит в клетку и быстро деполяризует мембрану до положительного значения φ_{max}, при котором h-ворота полностью закрываются.

Значение максимально достигнутого при деполяризации положительного мембранного потенциала φ_{max} всегда несколько меньше равновесного потенциала Нернста для ионов Na^+:

$$\varphi_{max} \leq \varphi_{Na} = -\frac{RT}{F} \ln \frac{C_i\,(Na^+)}{C_e\,(Na^+)} \qquad (11.5)$$

и для разных клеток принимает значения $\varphi_{max} \approx 10 - 40$ мВ (см. рис. 11.1).

Итак, по достижении максимального потенциала φ_{max} натриевые каналы закрываются и поток ионов Na^+ в клетку полностью прекращается. В то же время проводимость калиевых каналов повышается и ионы K^+ под действием как градиента концентрации, так и электрического поля начинают усиленно диффундировать наружу, вынося с собой из клетки положительный заряд и восстанавливая тем самым прежний мембранный потенциал покоя. Этот процесс, называемый ***реполяризацией***, приводит к восстановлению потенциала покоя на мембране.

Последний участок кривой реполяризации для некоторых видов клеток бывает замедлен (см. рис. 11.1, 11.2). Это медленное приближение мембранного потенциала к значению потенциала покоя называют ***следовым потенциалом***, который может быть деполяризационным (если он положительнее потенциала покоя) или гиперполяризационным (отрицательнее потенциала покоя, см. рис. 11.1, *а*).

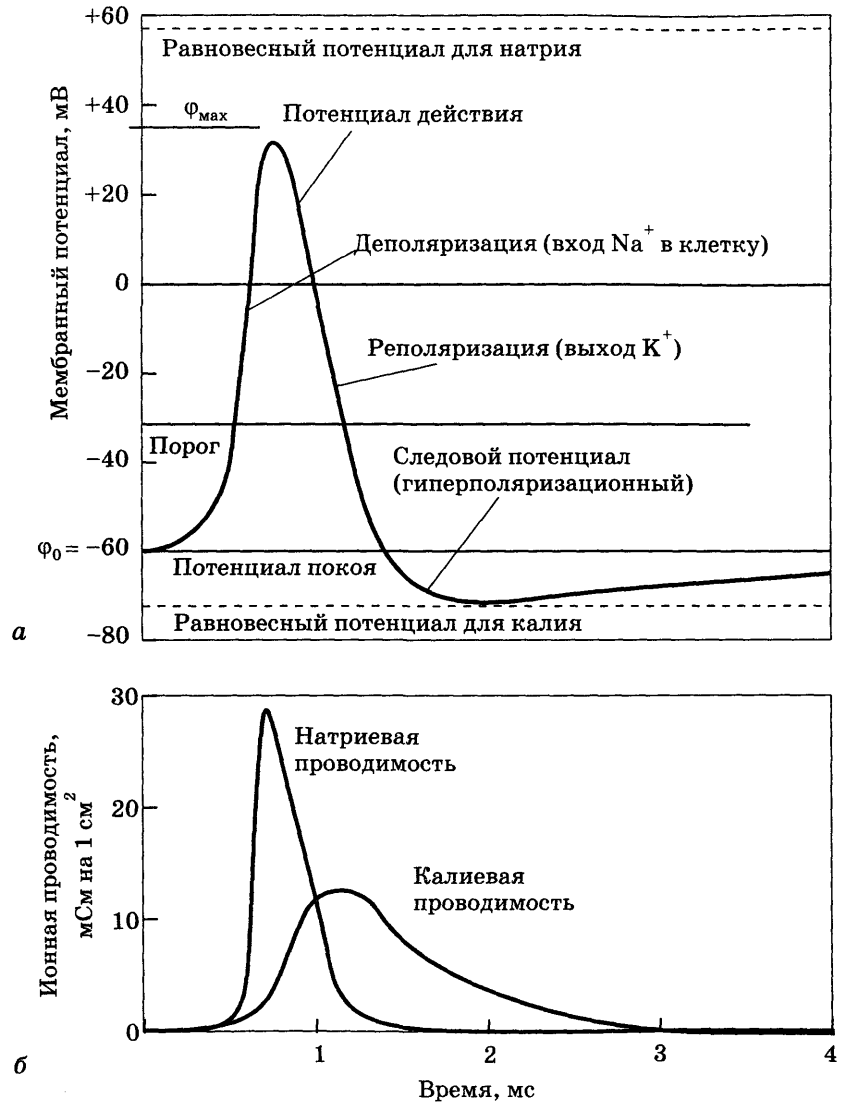


Рис. 11.1. Изменение во времени потенциала и ионной проводимости мембраны при возбуждении:
а — формирование потенциала действия;
б — изменение Na^+ и K^+-проводимости

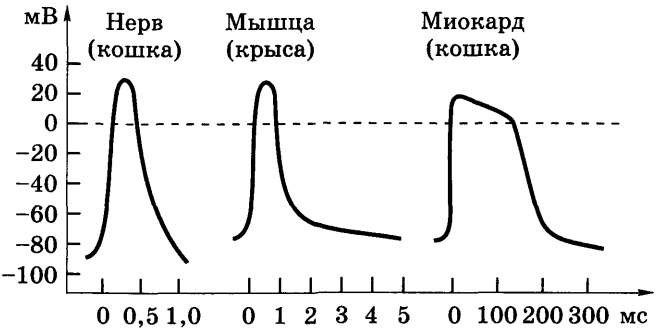


Рис. 11.2. Форма потенциалов действия различных клеток

Формы потенциалов действия разных клеток представлены на рис. 11.2.

Длительность фазы деполяризации во всех случаях мала и для нервных и мышечных клеток составляет 0,5–1 мс. При этом скорость нарастания мембранного потенциала в фазе деполяризации очень велика и достигает значений 300 В/с. Длительность фазы реполяризации существенно зависит от вида клеток: для нервных клеток и клеток скелетных мышц она составляет от 0,5 до 10 мс, а для клеток сердечной мышцы — около 300 мс.

Амплитуда $\varphi_д$ потенциала действия, как видно из рис. 11.1, равна сумме абсолютных значений потенциала покоя и максимально достигаемого потенциала φ_{max} и составляет 90–120 мВ:

$$\varphi_д = \varphi_{max} - \varphi_0 = \varphi_{max} + |\varphi_0|. \qquad (11.6)$$

Таким образом, потенциал действия возникает только при воздействии на клетку порогового или надпорогового раздражителя, который деполяризует мембранный потенциал клетки до значения критического потенциала возбуждения $E_{кр}$. Начиная с этого момента мембранный потенциал изменяется уже независимо от дальнейшего поведения стимула и полностью определяется потоками ионов Na^+ и K^+ через мембрану.

Амплитуда и форма потенциала действия для каждого вида клеток своя и представляет всегда одинаковое и характерное для этих клеток изменение мембранного потенциала во времени.

11.3.2. Механизм порогового действия

Наличие критического потенциала $E_{кр}$, при достижении которого генерируется потенциал действия, является следствием специфической зависимости проводимости натриевых каналов от мембранного потенциала. Кривая потенциал-зависимой проводимости Na^+-каналов, представленная на рис. 11.3, показывает, что Na^+-проводимость мала, если мембранный потенциал ниже (–40)–(–50) мВ, затем очень быстро растет в интервале от –40 до 0 мВ, после чего рост замедляется и канал закрывается.



Рис. 11.3. Зависимость проводимости Na^+-каналов от мембранного потенциала

Из рис. 11.3 видно, что лишь при значениях мембранного потенциала, близких к (–40) – (–50) мВ, проводимость Na^+-каналов начинает повышаться и уже небольшая последующая деполяризация вызывает весьма значительное увеличение их проводимости, что объясняет наличие критического потенциала $E_{кр}$ = (–40) – (–50) мВ и порога возбуждения клетки $U_{пор} = E_{кр} - \varphi_0$. При достижении критического потенциала проводимость натриевых каналов увеличивается и входящий в клетку натриевый ток еще больше деполяризует клетку, поток ионов натрия лавинообразно нарастает и мембрана быстро деполяризуется до значения $\varphi_{max} = 30 - 40$ мВ, при котором натриевые каналы закрываются и начинается фаза реполяризации потенциала действия.

Из рис. 11.1, *б* видно, что проводимость калиевых каналов тоже увеличивается при возбуждении клетки, но максимальная калиевая проводимость достигается позже, чем натриевая. Высокая калиевая проводимость мембраны определяет

быструю реполяризацию мембраны и последующий следовой потенциал.

Ходжкин и Катц экспериментально установили соотношение коэффициентов проницаемости мембраны гигантского аксона кальмара в состоянии покоя: $P_K : P_{Na} : P_{Cl} = 1 : 0{,}04 : 0{,}45$ и в фазе нарастания потенциала действия: $P_K : P_{Na} : P_{Cl} = 1 : 20 : 0{,}45$. Видно, что при возбуждении клетки в наибольшей степени изменяется относительная проницаемость мембраны для Na^+: от 0,04 до 20, т.е. в 500 раз, что сопровождается изменением электрического сопротивления мембраны от 1000 Ом/см^2 в покое до 25 Ом/см^2 при возбуждении, т.е. в 40 раз.

11.3.3. Инактивация Na^+-каналов

Генерация потенциала действия происходит вследствие повышения Na^+-проницаемости мембраны; без входа Na^+ в клетку потенциал действия сформироваться не сможет.

Вместе с тем известно, что при установлении на мембране постоянного уровня деполяризации проводимость мембраны для Na^+ не остается постоянной, а начинает падать примерно через 0,5 мс, а в нервных клетках — через 0,1 мс после установления постоянного уровня деполяризации. Такое быстрое снижение Na^+-проводимости мембраны называется ***инактивацией системы Na^+-каналов***. Как уже отмечалось, это объясняется наличием у Na^+-каналов двух типов ворот: быстрых m-ворот, открывающих канал при $E_{кр}$, и более медленных h-ворот, закрывающих его.

Если до начала действия стимула мембрана была деполяризована на 20–30 мВ, то система Na^+-каналов оказывается полностью инактивирована (успели полностью закрыться h-ворота) и ее не может активировать никакая степень последующей деполяризации. В этом случае возбуждение клетки становится невозможным и потенциал действия сформироваться не может.

11.3.4. Рефрактерность мембран

Используя тестирующие стимулы в различные фазы развития потенциала действия, можно проследить временной ход изменения возбудимости в процессе возбуждения клетки.

Если сразу после начала генерации потенциала действия (при $\varphi_м \geq E_{кр}$) попытаться подействовать на клетку внешним надпороговым раздражителем, то возбуждение не возникает ни при каких значениях силы раздражителя. Такое состояние полной невозбудимости мембраны называют ***абсолютной рефрактерностью***. В нервных клетках абсолютный рефрактерный период длится около 1 мс.

За ним следует ***относительный рефрактерный период***, когда путем значительного надпорогового воздействия все же можно вызвать потенциал действия, хотя его амплитуда и будет ниже нормы. Лишь после окончания этого периода клетка может быть возбуждена обычным пороговым стимулом и потенциал действия будет иметь нормальную амплитуду и форму.

Абсолютная рефрактерность мембран ограничивает максимальную частоту генерации потенциалов действия и определяет максимальную частоту генерации нервных импульсов, передаваемых от рецептора по нервному волокну.

Это же явление ограничивает и максимальную частоту возбуждения клеток при их стимуляции: $\nu_{max} = 1/T_{рефр}$, что приводит к значениям около 500 Гц для нервного волокна, около 200 Гц для скелетных мышц и около 3 Гц для миокардиоцитов.

В заключение отметим, что в результате генерации потенциала действия первоначальные концентрации ионов Na^+ и K^+ в клетке изменяются лишь на 0,0001 %, что практически не меняет величину градиентов концентрации этих ионов на мембране. Благодаря этому клетки могут генерировать тысячи импульсов без «подзарядки» своих концентрационных батарей. Однако ясно, что длительная генерация клеткой потенциалов действия рано или поздно привела бы к значительному уменьшению концентрационных градиентов Na^+ и K^+, в результате чего клетка потеряла бы возможность генерировать полноценный потенциал действия. В действительности же этого не происходит благодаря тому, что слегка измененные после генерации каждого потенциала действия концентрационные градиенты Na^+ и K^+ восстанавливаются за счет работы Na^+–K^+-насоса.

11.4. Подпороговые раздражители и местные потенциалы

При действии на клетку подпорогового раздражителя генерация потенциала действия не происходит, но в месте воздействия раздражителя мембранный потенциал изменяется на некоторую величину U_0, которую называют ***местным потенциалом***. Местные потенциалы могут распространяться вдоль мембраны с затуханием на небольшие расстояния от места воздействия, при этом их амплитуда $U(x)$ уменьшается с расстоянием x по экспоненциальному закону:

$$U(x) = U_0 e^{-\frac{x}{\lambda}}, \qquad (11.7)$$

где λ — характерный для каждой возбудимой ткани параметр, измеряемый в миллиметрах и определяющий расстояние, на котором амплитуда местного потенциала уменьшается в $e = 2{,}7$ раза.

Очень важным свойством местных потенциалов является то, что они способны суммироваться, при этом результирующий мембранный потенциал может превысить порог возбуждения клетки и вызвать генерацию потенциала действия. Дендриты и сома нервных клеток часто деполяризуются как раз до порогового уровня, и тогда от очень небольших различий в интенсивности раздражителя зависит, перейдет ли информация в форму потенциала действия или нет. Так, приход одного или слишком малого числа нервных импульсов к суммирующему нейрону вызывает местный потенциал подпороговой амплитуды, который быстро затухает, не перерастая в потенциал действия. Если же местные потенциалы, вызванные в дендритах и на теле клетки, суммируясь, превышают критический потенциал в той точке, где аксон покидает аксонный холмик, то происходит генерация потенциала действия, т.е. преобразование информации в качественно другую форму. Поэтому электротоническое распространение возбуждения по дендритам может иметь первостепенное значение для передачи и обработки информации в нервных клетках.

Сравнительные характеристики локального ответа (местного потенциала) и потенциала действия приведены в табл. 11.2.

Таблица 11.2

Локальный ответ	Потенциал действия
Не имеет порога возникновения	Возникает только при достижении порога возбуждения
Распространяется лишь на небольшое расстояние, быстро затухает	Распространяется на большое расстояние без затухания
Амплитуда пропорциональна силе раздражителя	Амплитуда не зависит от силы раздражителя, подчиняется закону «все или ничего»
Способен суммироваться	Не способен суммироваться
Возбудимость нервного волокна во время локального ответа повышена	Возбудимость нервного волокна понижена до нуля (до абсолютной рефрактерности)

Видно, что локальный ответ и потенциал действия представляют собой две существенно разные реакции клетки на действие раздражителя.

11.5. Распространение потенциала действия по аксонам

Координированная деятельность организма основана на непрерывном притоке информации из внешней и внутренней среды, ее переработке и выработке ответной реакции организма в целом. Специальные органы или клетки, воспринимающие стимулы, называются ***рецепторами***. Все рецепторы преобразуют энергию стимула в локальный электрический сигнал, который возбуждает нервные импульсы в сенсорном нейроне. Далее эти импульсы передаются в соответствующие отделы нервной системы, где они расшифровываются и используются для выработки нужной ответной реакции.

Нервный импульс представляет собой волну возбуждения, распространяющуюся по нервному волокну. На небольшие

расстояния (в несколько миллиметров) информация может передаваться и в виде местных потенциалов (локального ответа), амплитуда которых быстро затухает с расстоянием. На большие расстояния информация по нервной системе передается без затухания в виде распространяющихся по нервному волокну потенциалов действия.

Сложная информация о действующих на организм раздражителях кодируется в виде групп нервных импульсов, так называемых *разрядов*. Форма, амплитуда и длительность нервных импульсов, проходящих по одному и тому же волокну, одинаковы, а частота их следования и количество в разряде зависят от интенсивности раздражителя. Такой способ передачи информации является наиболее помехоустойчивым, так как в широких пределах не зависит от состояния проводящих волокон.

Распространяющиеся по нервным волокнам потенциалы действия передают информацию от рецепторов к нервным центрам, а от них — к эффекторам (исполнительным органам), индуцируют процессы сокращения мышц и секреции железистых клеток.

Основным элементом нервной системы является ***нервная клетка,*** или ***нейрон***, строение которой схематично представлено на рис. 11.4.

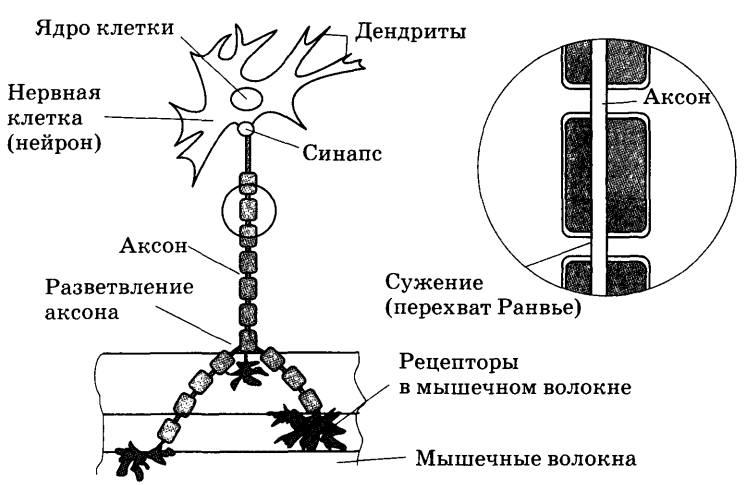


Рис. 11.4. Схема строения нейрона

Нейрон состоит из тела клетки, дендритов (ветвей тела клетки), которые подводят к нему нервные импульсы, и длинного нервного волокна (аксона), по которому нервные импульсы передаются от тела клетки к эффекторам. Волокна, по которым нервные импульсы передаются от рецепторов к телу нервной клетки, называют *афферентными*, а те, по которым нервные импульсы передаются по нервной системе к эффекторам, — *эфферентными*. Диаметр разных аксонов позвоночных животных варьируется от 0,0008 до 0,8 мм, а длина составляет несколько десятков сантиметров и может превышать метр. Отметим, что дендриты сенсорных нейронов также достигают в длину до 90 см.

Общепринятая классификация нервных волокон с указанием их функционального назначения, диаметра, скорости проведения потенциалов действия и длительности последних приведена в табл. 11.3.

Таблица 11.3

Тип волокна	Диаметр волокна, мкм	Скорость проведения, м/с	Длительность потенциала действия, мс	Функция
Aα	12–22	70–120	0,4–0,5	Моторные волокна скелетных мышц, афферентные волокна от мышечных рецепторов
Aβ	8–12	40–70	0,4–0,6	Афферентные волокна от рецепторов прикосновения
Aγ	4–8	15–40	0,5–0,7	Афферентные волокна от рецепторов прикосновения и давления, эфекторные волокна к мышечным веретенам
Aδ	1–4	5–15	0,6–1,0	Афферентные волокна от некоторых рецепторов тепла, давления, боли

Окончание табл. 11.3

Тип волокна	Диаметр волокна мкм	Скорость проведения, м/с	Длительность потенциала действия, мс	Функция
B	1–3,5	3–18	1–2	Преганглионарные вегетативные волокна
C	0,5–2,0	0,5–3	2,0	Постганглионарные вегетативные волокна, афферентные волокна от некоторых рецепторов тепла, давления, боли

Для каждого нервного волокна скорость распространения потенциала действия, или, как ее называют, скорость проведения, постоянна и зависит от функционального назначения аксона, его вида, диаметра и строения. Отметим, что у всех аксонов скорость проведения растет с увеличением их диаметра. У человека скорости проведения различных нервных волокон лежат в интервале от 0,1 до 120 м/с.

Известны два вида аксонов: *безмиелиновые*, не покрытые миелиновой оболочкой, и *миелинизированные*, имеющие такую оболочку. Физические механизмы проведения потенциала действия по этим аксонам несколько различаются.

11.5.1. Распространение потенциала действия по безмиелиновым аксонам

Рассмотрим сначала распространение потенциала действия по нервному волокну, не покрытому миелиновой оболочкой. Возбуждение такого аксона на каком-то участке приводит к полной деполяризации мембраны в этом месте, в результате чего потенциал внутри аксона в этом месте повышается до значения φ_{max}, тогда как в невозбужденных участках он остается отрицательным и равным потенциалу покоя φ_0. Под действием разности потенциалов ($\varphi_{max} - \varphi_0$) между возбужденными и невозбужденными участками аксона в аксоплазме

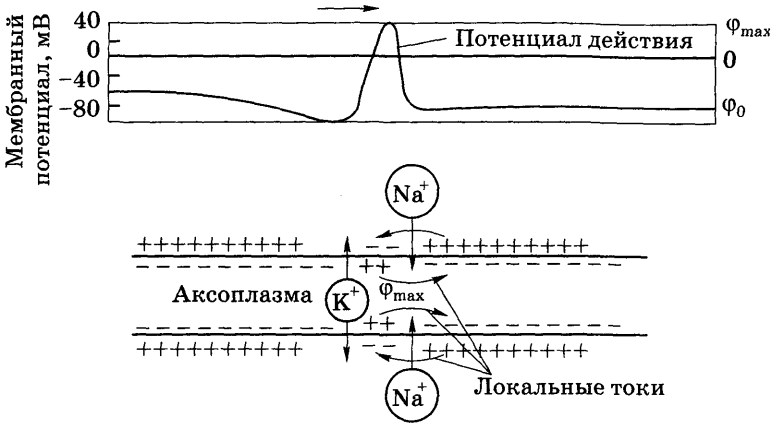


Рис. 11.5. Распространение потенциала действия по безмиелиновому аксону

возникает локальный ионный ток. Подобный локальный ток возникает и на наружной стороне аксона (11.5).

Локальные токи, обусловленные движением ионов, приводят к изменению концентрации зарядов по обе стороны мембраны и вследствие этого к увеличению электрического потенциала в участке аксона, прилегающем к возбужденному. Когда мембранный потенциал на этом участке достигает величины порогового потенциала возбуждения, натриевые каналы в мембране открываются, ионы Na^+ входят в клетку, этот участок аксона возбуждается и в этом месте аксона происходит генерация потенциала действия. Тем временем в ранее возбужденном участке аксона идет процесс реполяризации, обусловленный выходом калия наружу. В результате потенциал действия, т.е. возбуждение, перемещается на соседний участок аксона. Затем в возбуждение вовлекается следующий участок нервного волокна и таким путем возбуждение передается по безмиелиновому аксону с определенной скоростью.

Подчеркнем, что амплитуда распространяющегося по аксону потенциала действия сохраняется постоянной на протяжении всего волокна. Это объясняется тем, что на каждом последовательно возбуждаемом участке безмиелинового аксона всегда происходит генерация потенциала действия.

Увеличение диаметра аксона приводит к снижению электрического сопротивления аксоплазмы и вследствие этого к увеличению силы локальных токов, благодаря чему возрастает скорость изменения мембранного потенциала вдоль аксона, а с нею и скорость проведения нервного импульса. У некоторых беспозвоночных животных в целях быстрейшего сокращения определенных двигательных мышц (в ответ на ту или иную ситуацию) развились гигантские безмиелиновые аксоны диаметром до 1 мм, что обеспечило им скорость проведения нервного импульса до 25 м/с.

11.5.2. Распространение потенциала действия по миелинизированным аксонам

У позвоночных животных в процессе эволюции сформировались аксоны, покрытые миелиновой оболочкой, что обеспечило бóльшие скорости проведения нервных импульсов при значительно меньших диаметрах таких аксонов (см. табл. 11.3).

Строение миелинизированного аксона схематично представлено на рис. 11.6, *а*. Участки аксона длиной в 2–3 мм

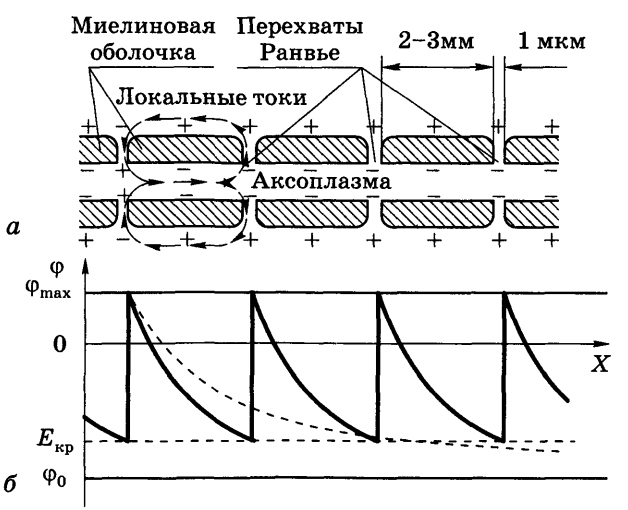


Рис. 11.6. Распространение нервного импульса по аксону: *а* — строение миелинизированного аксона; *б* — изменение мембранного потенциала

покрыты миелиновой оболочкой (шванновской клеткой), а между ними расположены «оголенные» участки мембраны, называемые ***перехватами Ранвье***, длиной лишь около 1 мкм.

На участках аксона, покрытых миелином, мембрана полностью изолирована и не имеет контакта с межклеточной жидкостью, содержащей ионы Na^+, вследствие чего потенциал действия здесь сформироваться не может, даже если потенциал аксоплазмы превышает критический потенциал возбуждения $E_{кр}$. Потенциал действия может сформироваться только в перехватах Ранвье, где мембрана имеет необходимый контакт с межклеточной жидкостью.

Допустим, что один из перехватов Ранвье возбужден, т.е. деполяризован до потенциала φ_{max}. Тогда под действием разности потенциалов ($\varphi_{max} - \varphi_0$) между возбужденным и невозбужденным участками в аксоплазме и на наружной стороне мембраны возникают локальные токи, благодаря которым мембранный потенциал распространяется вдоль аксона как по электрическому кабелю и с большой скоростью. Но по мере удаления от возбужденного участка мембранный потенциал аксона экспоненциально уменьшается (рис. 11.6, *б*), как и любой местный потенциал [см. формулу (11.7)].

Когда этот затухающий мембранный потенциал достигнет ближайшего перехвата Ранвье и повысит там $\varphi_м$ до $E_{кр}$, в мембране этого перехвата открываются Na^+-каналы и генерируется полноценный потенциал действия, что приводит к увеличению мембранного потенциала до максимального значения φ_{max}. Далее весь процесс повторяется, при этом потенциал действия последовательно генерируется лишь в перехватах Ранвье, а на миелинизированных участках аксона происходит его распространение с затуханием, но с большой скоростью, благодаря чему и достигается высокая скорость передачи возбуждения по такому нервному волокну.

Проведение нервного импульса в миелинизированном аксоне называют ***сальтаторным*** (***скачкообразным***), так как потенциал действия генерируется только в перехватах Ранвье и в процессе распространения как бы перескакивает с одного перехвата Ранвье на последующий. Как и в случае безмиелиновых нервных волокон, скорость распространения потенци-

ала действия в миелинизированном аксоне возрастает с увеличением его диаметра, что хорошо видно из табл. 11.3.

Одной из причин высокой скорости проведения потенциала действия в миелинизированных аксонах является высокая плотность Na^+-каналов в перехватах Ранвье, где она примерно в 100 раз выше, чем в безмиелиновых волокнах.

Высокая скорость проведения и малый диаметр миелинизированных аксонов обеспечили позвоночным животным возможность создания большого количества параллельных быстропроводящих нервных путей, что способствовало высокой надежности передачи информации по нервной системе.

Если входящий в клетку натриевый ток уменьшить, например, путем снижения концентрации Na^+ в растворе, окружающем аксон, либо путем инактивации (ухудшения проводимости) Na^+-каналов (введением ядов либо воздействием местноанестезирующего препарата), то скорость проведения потенциала действия снижается и в конечном итоге проведение может быть блокировано.

Если с помощью определенных ядов (например, тетродотоксина) заблокировать натриевые каналы в одном из перехватов Ранвье, то потенциал действия сформируется не на нем, а на следующем перехвате, так как амплитуда затухающего местного потенциала, распространяющегося за счет локальных токов (пунктирная кривая на рис. 11.6, *б*), еще сможет достичь на нем порогового потенциала возбуждения. Но если заблокировать подряд два или более перехватов Ранвье, то потенциал действия на последующих перехватах Ранвье сформироваться не сможет и проведение нервного импульса будет блокировано.

Контрольные вопросы и задания

1. Как формируются мембранные потенциалы покоя клетки?
2. Выведите формулу для равновесного потенциала Нернста.
3. Какие ионы определяют мембранные потенциалы клеток? Запишите уравнение Гольдмана — Ходжкина — Катца. Что оно определяет?

4. Каковы значения мембранных потенциалов покоя разных клеток? Какова при этом напряженность электрического поля в мембране?
5. Как действие раздражителя влияет на мембранный потенциал? При каких условиях клетка может возбудиться?
6. Что такое подпороговые стимулы и местные потенциалы? Каковы их свойства?
7. При каких условиях возникает возбуждение клетки? Какие процессы происходят на мембране при генерации потенциала действия?
8. Охарактеризуйте фазы деполяризации и реполяризации, их длительность и связь с ионными потоками через мембрану. Какова форма и длительность потенциалов действия разных клеток?
9. От чего и как зависит проницаемость натриевых каналов?
10. Что такое рефрактерные периоды? Каковы их виды и длительность для разных клеток?
11. Приведите сравнительную характеристику потенциала действия и локального ответа. Каковы существенные различия между ними?
12. Опишите процесс распространения потенциала действия по безмиелиновому аксону.
13. Охарактеризуйте процесс распространения потенциала действия по миелинизированному аксону.

Глава 12 Электрические поля органов и тканей. Методы их регистрации

Все тела в природе состоят из атомов, которые в свою очередь включают в себя положительно заряженные ядра и отрицательно заряженные электроны. Именно электростатические силы притяжения удерживают атомы и молекулы в устойчивом состоянии, а электрические взаимодействия определяют физическое и химическое строение атомов и молекул, свойства твердых тел, газов, жидкостей и биологических структур. Знание законов взаимодействия электрических зарядов между собой и с внешними электрическими полями необходимо для понимания процессов формирования электрических полей органов и тканей, методов их регистрации и механизмов воздействия внешних электрических полей на биологические объекты.

12.1. Электрическое поле и его характеристики

В природе известны два вида электрических зарядов, которые были условно названы положительными и отрицательными. В СИ единицей электрического заряда является 1 кулон — это заряд, который проходит за 1 с через поперечное сечение проводника при токе в 1 ампер: 1 Кл = 1 А · с. Наименьшим (элементарным) отрицательным зарядом обладает электрон, а таким же по величине, но положительным — протон. Величина элементарного заряда $e = 1{,}6 \cdot 10^{-19}$ Кл.

Одним из важнейших законов природы является ***закон сохранения электрического заряда***: в замкнутой системе тел алгебраическая сумма электрических зарядов всегда остается постоянной при любых взаимодействиях этих тел между собой.

Опыт показывает, что заряды одного знака всегда отталкиваются, а заряды разных знаков притягиваются. Сила взаимодействия точечных зарядов q_1 и q_2, расположенных

на расстоянии r друг от друга, определяется ***законом Кулона***:

$$F = k\frac{q_1 q_2}{\varepsilon r^2}, \qquad (12.1)$$

где $k = 9 \cdot 10^9\ \text{Н} \cdot \text{м}^2/\text{Кл}^2$ — постоянная электрического взаимодействия; ε — относительная диэлектрическая проницаемость среды, в которой находятся заряды.

Способность электрических зарядов взаимодействовать на расстоянии объясняется тем, что каждый из них создает вокруг себя электрическое поле, которое оказывает силовое воздействие на любой внесенный в него другой заряд.

Поэтому, чтобы определить наличие электрического поля в какой-либо точке среды, в нее помещают пробный заряд q_1. Сила F действия поля на этот заряд будет пропорциональна величине внесенного заряда q_1, но отношение $\vec{E} = \dfrac{\vec{F}}{q_1}$ уже не зависит от величины пробного заряда, а является силовой характеристикой электрического поля, которую называют ***напряженностью*** $\vec{E}\,(\text{Н/Кл}, \text{В/м})$. Если напряженность $\vec{E}$ поля известна, то сила, действующая на любой заряд q_1, внесенный в это поле,

$$\vec{F} = q_1\vec{E}. \qquad (12.2)$$

Отметим, что закон Кулона (12.1) применим только для точечных зарядов и шаров, а формула (12.2) — и для любых других зарядов и полей. Из векторного равенства (12.2) следует, что сила, действующая на положительный заряд, направлена так же, как и вектор $\vec{E}$ напряженности поля, а сила, действующая на отрицательный заряд, — противоположно ему.

Электрические поля принято изображать силовыми линиями — направленными линиями, касательные к которым совпадают с направлением вектора напряженности $\vec{E}$ поля в этой точке. Силовые линии электростатического поля всегда направлены от положительных зарядов к отрицательным,

т.е. начинаются на положительных зарядах и заканчиваются на отрицательных.

Электрическое поле называют *однородным*, если его напряженность одинакова по величине и направлению во всех точках пространства: $\vec{E} = \text{const}$. Такие поля изображаются параллельными равноотстоящими друг от друга силовыми линиями.

Пользуясь законом Кулона, найдем напряженность поля, создаваемого точечным зарядом q_0:

$$E = \frac{F}{q_1} = k\frac{q_0}{\varepsilon r^2}. \qquad (12.3)$$

Формула (12.3) справедлива и для поля, создаваемого заряженным шаром радиусом R в окружающем его пространстве, т.е. при $r \geq R$. Такое поле неоднородно, так как его напряженность различна и по величине, и по направлению в разных точках пространства. На рис. 12.1 стрелками показаны силовые линии электрических полей, создаваемых положительным и отрицательным точечными зарядами, а концентрическими окружностями — линии равного потенциала.

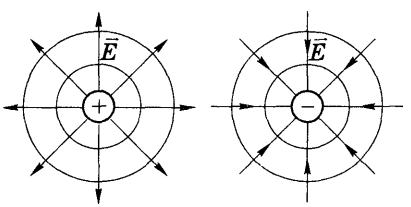


Рис. 12.1. Силовые и эквипотенциальные линии (окружности) электрических полей точечных зарядов

Если электрическое поле создается не одним зарядом, а несколькими, то для определения напряженности результирующего поля следует применить ***принцип суперпозиции полей***: напряженность электрического поля, создаваемого системой зарядов, равна векторной сумме напряженностей полей, создаваемых в этой точке среды каждым из зарядов:

$$\vec{E} = \vec{E}_1 + \vec{E}_2 + \vec{E}_3 + \ldots + \vec{E}_n. \qquad (12.4)$$

Рассмотрим теперь энергетическую сторону взаимодействия поля и заряда. Поскольку на заряд q, помещенный в электрическое поле, действует сила $F = qE$, то он может переместить-

ся вдоль силовой линии на расстояние S и совершить работу $A = \int F(S)dS = q\int E(S)ds$. Это означает, что любой заряд q, находящийся в электрическом поле, обладает потенциальной энергией $W_{пот} = A$, которая пропорциональна величине этого заряда q. Но отношение $\varphi = W_{пот}/q$ уже не зависит от величины заряда, а является энергетической характеристикой поля и называется его ***электрическим потенциалом*** φ. Единицей потенциала в СИ является 1 вольт: $1\ В = 1\ Дж/Кл$. Совокупность всех точек поля, имеющих одинаковый электрический потенциал, образует ***эквипотенциальную поверхность***, которая всегда перпендикулярна линиям напряженности E поля.

Потенциал электрического поля, создаваемого точечным зарядом q_0, определяется формулой

$$\varphi = k\frac{q_0}{\varepsilon r}. \qquad (12.5)$$

В этом случае поверхностями равного потенциала являются сферы радиусом r, тогда как силовые линии идут из центра сферы по радиусам перпендикулярно этим сферам (см. рис. 12.1).

Если электрическое поле создается не одним зарядом, а системой зарядов, то потенциал создаваемого ею электрического поля равен алгебраической сумме потенциалов полей, создаваемых в данной точке среды каждым из зарядов:

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3 + ... + \varphi_n. \qquad (12.6)$$

Очевидно, что нахождение потенциала сложного поля значительно проще, чем вычисление его напряженности, требующей нахождения векторной суммы.

Если электрическое поле перемещает заряд q из одной точки поля в другую, то оно совершает работу

$$A = W_{пот1} - W_{пот2} = q(\varphi_1 - \varphi_2) = qU. \qquad (12.7)$$

Величина $U = \varphi_1 - \varphi_2$ называется ***разностью потенциалов*** или ***напряжением*** между точками 1 и 2 поля и измеряется в вольтах.

Обратим внимание на очень важное *свойство электрического поля*: работа в электрическом поле не зависит от траектории перемещения заряда из точки 1 в точку 2, а определяется только разностью потенциалов между ними [см. (12.7)].

Из формулы (12.7) следует также, что при перемещении заряда по эквипотенциальной поверхности работа не совершается, так как $\varphi_1 = \varphi_2$ и $A = U = 0$.

В общем случае напряженность E электрического поля равна градиенту потенциала с обратным знаком:

$$\vec{E} = -\overrightarrow{\text{grad}\varphi}, \qquad (12.8)$$

поэтому, зная распределение потенциала $\varphi(r)$ в пространстве, можно найти распределение напряженности поля $E(r)$, и наоборот. Знак «–» указывает, что вектор напряженности направлен против градиента потенциала, т.е. в сторону наибыстрейшего уменьшения потенциала. Для однородного поля, силовые линии которого параллельны оси Ox, соотношение (12.8) принимает вид

$$\vec{E} = -\frac{\overrightarrow{d\varphi}}{dx}. \qquad (12.8а)$$

Значения напряженности и потенциала электрического поля, создаваемого точечным или сферическим зарядом q_0, даются формулами (12.3) и (12.5).

Таким образом, электрически заряженная система тел всегда создает вокруг себя электрическое поле. Однако оказывается, что и электрически нейтральная система зарядов тоже может создавать вокруг себя электрическое поле. Простейшей такой системой является электрический диполь.

12.2. Электрический диполь и его поле

Электрический диполь представляет собой электрически нейтральную систему, состоящую из двух равных по величине, но противоположных по знаку точечных зарядов $(+q)$ и $(-q)$, расположенных на некотором расстоянии l друг от друга. Хотя такая система электрически нейтральна, она создает вокруг себя электрическое поле.

Рассчитаем потенциал этого поля в произвольной точке A, расположенной на расстояниях r_1 и r_2 от положительного и отрицательного зарядов диполя соответственно (рис. 12.2). Суммарный потенциал в точке A

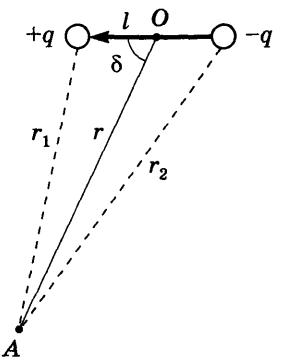


Рис. 12.2. Потенциал поля диполя

$$\varphi_A = \varphi_+ + \varphi_- = k\frac{q}{\varepsilon r_1} - k\frac{q}{\varepsilon r_2} =$$

$$= k\frac{q}{\varepsilon}\left(\frac{1}{r_1} - \frac{1}{r_2}\right) = k\frac{q}{\varepsilon}\frac{(r_2 - r_1)}{r^2},$$

где произведена замена $r_1 r_2 \approx r^2$. Учитывая, что $(r_2 - r_1) \approx l\cos\delta$, получим

$$\varphi_A = k\frac{ql\cos\delta}{\varepsilon r^2} = k\frac{p\cos\delta}{\varepsilon r^2}. \qquad (12.9)$$

Видно, что потенциал φ поля прямо пропорционален величине $p = ql$, которая называется ***дипольным моментом*** и является важнейшей электрической характеристикой диполя. Дипольный момент $\vec{p}$ — это вектор, направленный вдоль оси диполя от $(-q)$ к $(+q)$. Зная этот вектор, можно вычислить потенциал в любой точке поля, пользуясь формулой (12.9). Из нее, в частности, следует, что в плоскости, проходящей через середину диполя перпендикулярно его оси ($\delta = 90°$), потенциал равен нулю, что он положительный со стороны положительного заряда и отрицательный со стороны отрицательного заряда. Примерное распределение эквипотенциальных поверхностей поля диполя приведено на рис. 12.3.

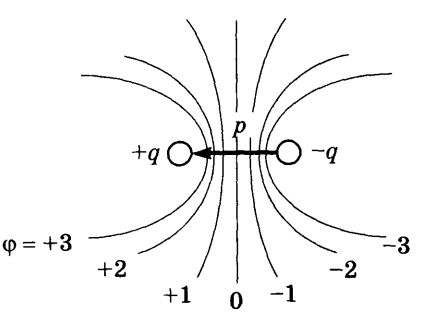


Рис. 12.3. Поле диполя. Линии равного потенциала электрического поля диполя

Итак, зная дипольный момент $\vec{p}$, можно вычислить потенциал поля диполя в любой точке пространства.

На практике, в частности в электрографии, часто приходится решать обратную задачу: необходимо определить величину и направление электрического момента диполя, измеряя разность потенциалов (напряжение) между разными точками (A и B) создаваемого им поля (рис. 12.4).

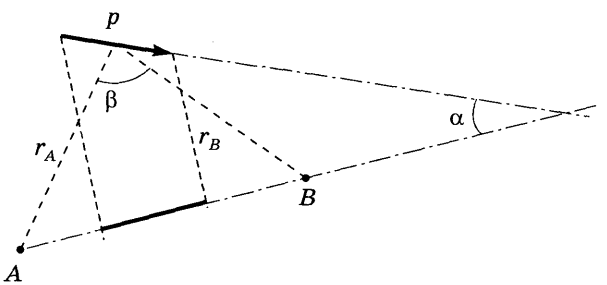


Рис. 12.4. Связь между дипольным моментом $\vec{p}$ и напряжением U_{AB}

В соответствии с расчетом, напряжение $U_{AB} = \varphi_A - \varphi_B$ между двумя равноудаленными от центра диполя точками A и B ($r_A = r_B = r$) прямо пропорционально проекции вектора дипольного момента ($p\cos\alpha$) на линию AB, соединяющую эти точки, и обратно пропорционально квадрату их расстояния от диполя:

$$U_{AB} = 2k\frac{p\sin\frac{\beta}{2}\cos\alpha}{\varepsilon r^2} \sim \frac{p\cos\alpha}{r^2}. \qquad (12.10)$$

Однако знания одной проекции вектора недостаточно. Чтобы полностью определить величину и направление вектора $\vec{p}$, надо знать как минимум две его проекции на разные направления AB и BC (точка C и отрезок BC на рис. 12.4 не показаны), при этом точки регистрации потенциалов A, B и C должны быть равноудалены от центра диполя ($r_A = r_B = r_C$) и находиться с ним в одной плоскости, а отрезки AB и BC должны быть видны из центра диполя под одинаковыми уг-

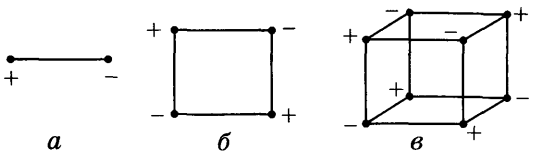


Рис. 12.5. Электрически нейтральные мультиполи:
а — диполь; *б* — квадруполь; *в* — октуполь

лами β. Этим условиям удовлетворяют вершины правильного многоугольника, в центре которого находится диполь $\vec{p}$. В простейшем случае это будут вершины равностороннего треугольника, которые часто и выбирают для регистрации потенциалов.

Отметим, что диполь — простейшая, но не единственная нейтральная система зарядов, создающая вокруг себя электрическое поле. К простым электронейтральным системам относятся также квадруполь (2 положительных и 2 отрицательных равных по модулю заряда), октуполь (4 + 4 заряда) и другие, которые называют одним словом «мультиполи» (рис. 12.5). Потенциалы полей, создаваемых ими, быстро убывают с расстоянием. Так, потенциал поля точечного заряда убывает с расстоянием как r^{-1}, потенциал поля диполя — r^{-2}, квадруполя — r^{-3}, октуполя — r^{-4} и т.д.

Электрическое поле создают не только указанные простейшие мультиполи, но и более сложные электрически нейтральные системы, в том числе биологические ткани и органы.

Из математической физики известно, что любое стационарное электрическое поле, создаваемое электрически нейтральной системой зарядов, можно приближенно представить как сумму электрических полей, создаваемых диполем, квадруполем, октуполем и т.д. Наибольший вклад в это суммарное поле, как правило, вносит диполь, поэтому обычно ограничиваются рассмотрением только дипольной составляющей сложного поля (дипольное приближение), что во многих случаях оказывается полностью оправданным.

Здесь следует отметить, что рассмотренный электростатический диполь хорошо описывает электрическое поле в непроводящих средах, т.е. в диэлектриках, где нет свободных

зарядов и поэтому нет и токов проводимости. Однако многие биоткани — кровь, лимфа, спинномозговая жидкость, мышцы, нервная ткань, и другие — являются хорошими проводниками и в них под действием электрических полей возникают электрические токи.

Поэтому в проводящей среде под действием поля диполя возникнет движение свободных зарядов (электрический ток) и электростатический диполь либо быстро экранируется этими зарядами, либо нейтрализуется.

Электрическое поле, создаваемое живыми возбудимыми биотканями, более правильно рассматривать не как поле электростатического диполя, а как поле, создаваемое токовым диполем. Такое представление в электрокардиографии впервые было развито еще Г. Гельмгольцем, а затем подробно рассмотрено другими исследователями.

Для создания электрического поля в проводящей среде необходим источник тока — токовый генератор, создающий ЭДС и имеющий два полюса: исток (+) и сток (–) зарядов (рис. 12.6). Его внутреннее сопротивление r намного больше сопротивления R внешней цепи (нагрузки), т.е. $r >> R$. По закону Ома для полной цепи сила тока

$$I = \frac{\varepsilon}{r + R},$$

где ε — ЭДС генератора.

Поскольку $r >> R$, то $I \approx \frac{\varepsilon}{r}$. Это означает, что сила тока I, которую создает этот генератор, практически не зависит от

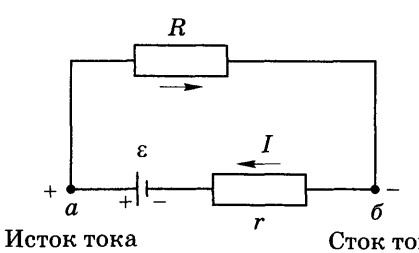


Рис. 12.6. Эквивалентная схема токового электрического генератора и внешней проводящей среды (*а* и *б* — полюса генератора, исток и сток)

сопротивления внешней среды и определяется только параметрами самого токового генератора.

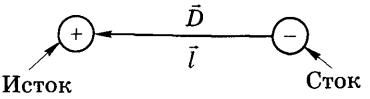


Рис. 12.7. Токовый диполь

При расчете потенциалов, создаваемых таким генератором во внешней проводящей среде, его удобно представить в виде ***токового диполя*** (рис. 12.7) — системы из положительного и отрицательного полюсов (истока и стока электрического тока), расположенных на некотором расстоянии l друг от друга.

Основной параметр токового диполя — его дипольный момент $\vec{D} = I\vec{l}$, равный произведению силы тока I через генератор на расстояние l и направленный от стока (–) к истоку (+) тока, т.е. по направлению тока внутри источника (см. рис. 12.6).

Любой электрически активный элемент биоткани (мембрана клетки, мышечное волокно, аксон и др.) можно рассматривать как элементарный токовый диполь D_i, имеющий свой ***исток*** и ***сток*** тока. Большое число таких малых диполей, расположенных рядом и ориентированных в одном направлении, образуют «двойной слой», на одной стороне которого расположены «истоки», а на другой – «стоки». Общий дипольный момент такого двойного слоя равен векторной сумме моментов составляющих его токовых диполей. Каждый такой диполь можно разложить на три взаимно перпендикулярные компоненты, затем соответствующие компоненты всех малых диполей можно сложить и получить в результате три суммарные компоненты одного эквивалентного токового диполя $\vec{D}$ данной биоткани или органа.

Потенциал в точке A, создаваемый таким диполем, дается формулой

$$\varphi_A = \rho\frac{D\cos\delta}{4\pi r^2} \sim \frac{D\cos\delta}{r^2}, \qquad (12.11)$$

которая аналогична выражению (12.9) для потенциала поля электростатического диполя. Поэтому с математической точки зрения выражения для потенциалов электрических полей, создаваемых электростатическим диполем и токовым

диполем, аналогичны, и формулу (12.10), связывающую разность потенциалов между двумя точками поля A и B и дипольным моментом D, можно получить и в случае токового диполя:

$$U_{AB} = 2\rho \frac{D \sin\frac{\beta}{2} \cos\alpha}{4\pi r^2} \sim \frac{D\cos\alpha}{r^2}. \qquad (12.12)$$

Поэтому, чтобы определить величину и направление вектора токового дипольного момента $\vec{D}$, как и в случае электростатического диполя $\vec{P}$, точки регистрации потенциалов наиболее удобно выбирать в вершинах правильного многоугольника, в центре которого находится диполь $\vec{D}$ или $\vec{P}$; в простейшем случае это будут вершины равностороннего треугольника.

Вместе с тем между электростатическим и токовым диполем есть существенная разница. Токовый диполь возникает только в том случае, когда разные участки возбудимой ткани, например мышцы, находятся в разных состояниях — одни покоятся, а другие возбуждаются, тогда и появляются области истока (+) и стока (–) и возникает токовый диполь, создающий в пространстве определенное электрическое поле. Если же вся мышца находится в одинаковом состоянии (покоится или полностью охвачена возбуждением), то токовый диполь в эти моменты времени равен нулю и не создает разности потенциалов в пространстве. Таким образом, токовый диполь отражает динамику процесса возбуждения в мышцах и других возбудимых тканях.

12.3. Физические основы электрографии органов и тканей

Регистрация электрических полей органов и тканей человека и животных получила общее название ***электрографии***. Хорошо известно, что на каждом участке клеточной мембраны возбудимой ткани существует разность электрических потенциалов и, значит, существует электрическое поле, напряженность которого весьма значительна и составляет в состоянии покоя клетки сотни киловольт на сантиметр (см.

11.2). Эти поля локализованы, прежде всего, в самой мембране клетки, но они есть и вне ее. Электрические поля, создаваемые отдельными элементами клетки, создают электрическое поле этой клетки, а сумма электрических полей клеток, образующих орган или ткань, формирует электрическое поле всего органа или ткани.

Поскольку в это общее электрическое поле вносит свой вклад каждая клетка ткани, то изменение функционального, а значит, и электрического, состояния отдельных клеток отражается на электрическом поле всего органа или ткани. Поэтому регистрация электрических полей органов и тканей позволяет оценивать их функциональное состояние, определять изменения этого состояния под действием различных внешних факторов, выявлять патологические изменения в этих органах и тканях. Все это и определяет то большое значение, которое имеет в клинической медицине регистрация электрических полей тканей и органов в целях диагностики их функционального состояния.

Точное описание электрических полей тканей и органов — ложная математическая и физическая задача. Поэтому используются способы их приближенного описания, дающие вместе с тем достаточно хорошую точность.

Поскольку биоткани и органы в целом электрически нейтральны, то создаваемое ими электрическое поле можно приближенно рассматривать как электрическое поле, образуемое некоторым диполем.

Этот результирующий диполь, описывающий электрическое поле органа, часто называют интегральным (суммарным) электрическим вектором этого органа. Если известны его величина и направление, то фактически известно все об этом поле, так как по формуле (12.9) или (12.11) нетрудно вычислить потенциал электрического поля в любой точке пространства вокруг диполя и затем определить другие характеристики поля.

Если изменяется состояние отдельных клеток ткани (например, часть их возбуждается), то изменяется и электрическое поле этой ткани, и, как следствие, параметры диполя (величина и направление), описывающие это поле. Регис-

трируя изменяющуюся во времени разность потенциалов между определенными точками тела человека (так называемыми *отведениями*), можно для каждого момента времени найти величину и направление электрического диполя, представляющего электрическое поле органа или ткани, определить, как и в какой последовательности изменялось во времени состояние отдельных структурных частей ткани, и оценить ее функциональное состояние в целом. Эта идея и лежит в основе использования электрографии для диагностики функционального состояния возбудимых тканей организма.

Отдельными диагностическими методами электрографии являются: *электрокардиография* (ЭКГ) — регистрация электрических полей, создаваемых сердечной мышцей; *электромиография* (ЭМГ) — регистрация электрических полей скелетных мышц; *электроэнцефалография* (ЭЭГ) — регистрация электрической активности головного мозга и др.

Рассмотрим более подробно метод регистрации электрических полей на примере ЭКГ. Электрическое поле, создаваемое сердечной мышцей, можно приближенно рассматривать как поле, создаваемое некоторым токовым диполем, который называют *интегральным (суммарным) электрическим вектором сердца* или *электрическим генератором сердца*. Для корректной регистрации параметров диполя необходимо регистрировать напряжения между точками, расположенными в вершинах правильного многоугольника, в центре которого находится исследуемый диполь.

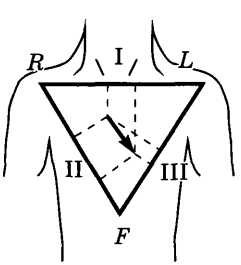


Рис. 12.8. Треугольник Эйнтховена и стандартные отведения

Впервые подобную систему отведений предложил голландский врач В. Эйнтховен. Он предложил рассматривать электрическое поле сердца как поле электрического диполя, расположенного в центре равностороннего треугольника (известного теперь как «треугольник Эйнтховена»), вершинами которого являются правое R и левое L плечо и основание торса F (рис. 12.8). Однако в целях удобства пациентов и персонала для

наложения электродов используются не эти точки, а запястья обеих рук и голень левой ноги, при этом сами конечности играют лишь роль естественных «проводников» примерно равной длины.

Отведения, предложенные Эйнтховеном, принято называть *стандартными*. Электрограммы в этих отведениях определяют изменения во времени проекций интегрального электрического вектора сердца $\vec{D}$ на три стороны треугольника Эйнтховена. Разность потенциалов U_{I} между правой рукой и левой рукой называют I отведением, между правой рукой и левой ногой — II отведением U_{II} и между левой рукой и левой ногой — III отведением U_{III} Эйнтховена.

При этом выполняется ***закон Эйнтховена***: в любой момент времени алгебраическая сумма напряжений в отведениях I и III должна быть равна напряжению в отведении II:

$$U_{\text{I}} + U_{\text{III}} = U_{\text{II}}. \qquad (12.13)$$

На практике, кроме этих трех стандартных отведений, регистрируют еще девять отведений: три усиленных униполярных и шесть грудных.

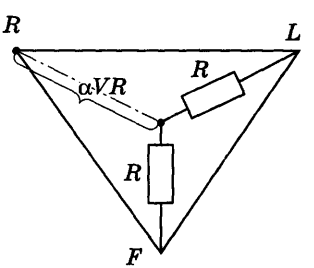


Рис. 12.9. Схема регистрации униполярного усиленного отведения αVR

В *униполярных усиленных отведениях*, обозначаемых как αVR, αVL и αVF, регистрируют разность потенциалов между одной из вершин треугольника Эйнтховена (R, L или F) и усредненным потенциалом двух других его вершин, для чего последние соединяют между собой двумя равными сопротивлениями (рис. 12.9). Униполярные отведения определяют проекции электрического вектора сердца на биссектрисы соответствующих углов R, L, F треугольника.

Однако три стандартных и три усиленных отведения определяют поведение электрического вектора сердца лишь в плоскости треугольника Эйнховена, т.е. во фронтальной плоскости тела, и не дают никакой информации о проекци-

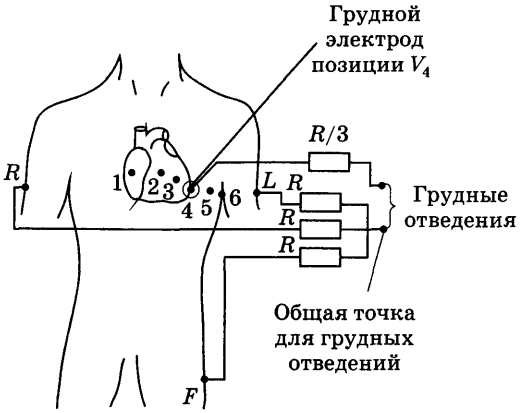


Рис. 12.10. Схема регистрации грудных отведений

ях этого вектора на направление, перпендикулярное этой плоскости. Поэтому, чтобы получить полное представление об электрическом поле сердца, регистрируют еще шесть *грудных отведений*. Они представляют собой разность потенциалов между общей точкой треугольника (она электрически соединена тремя равными сопротивлениями с вершинами треугольника Эйнтховена) и одной из шести фиксированных точек на грудной клетке пациента (рис. 12.10).

Таким образом, при стандартной регистрации ЭКГ записывают 12 отведений: три стандартных отведения I, II, III; три усиленных отведения αVR, αVL, αVF и шесть грудных отведений V_1–V_6.

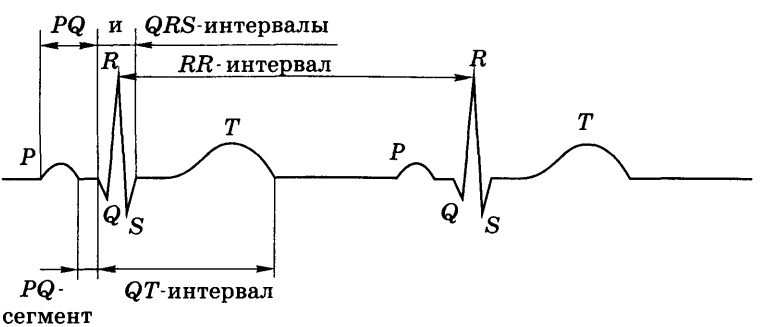


Рис. 12.11. Два последовательных кардиоцикла стандартной ЭКГ

Типичный вид ЭКГ приведен на рис. 12.11, где латинскими буквами обозначены ее основные зубцы. Наименьшую амплитуду (не считая зубцов Q и S, которые могут отсутствовать на электрокардиограмме) обычно имеет зубец P (около 0,1 мВ), а наибольшую — зубец R (до 2 мВ). Длительность основных временных интервалов ЭКГ в норме представлена в табл. 12.1, причем длительность интервалов QT и RR приведена для частоты сердечных сокращений 70 уд./мин.

Таблица 12.1

Интервал ЭКГ	*PQ*	*QRS*	*QT*	*RR*
Длительность, с	< 0,2	< 0,12	0,3–0,4	0,85

Электрическое поле сердца неразрывно связано с физиологическими процессами, происходящими в сердечной мышце. Поэтому, чтобы правильно понимать, как формируется электрокардиограмма и о чем свидетельствуют те или иные ее участки, нужно иметь общие представления о физиологических процессах, происходящих в миокарде в течение сердечного цикла.

Сердце (рис. 12.12) представляет собой орган величиной с кулак; оно имеет несколько заостренную верхушку, обращенную вниз и влево. В верхней части сердца расположены два тонкостенных насоса — правое и левое предсердия, которые наполняют кровью основные насосы сердца — правый и левый желудочки. Они разделены внутри сердца толстой мышечной стенкой, которая называется перегородкой. Из правого желудочка кровь поступает в легкие через легочные артерии, а из левого — в аорту и далее в остальные части тела по большому кругу кровообращения.

В сердце выделяют центр автоматии, расположенный в синусовом узле на правом предсердии, систему специализированных проводящих волокон, по которым возбуждение распространяется по миокарду, и мышечный сократительный аппарат сердца.

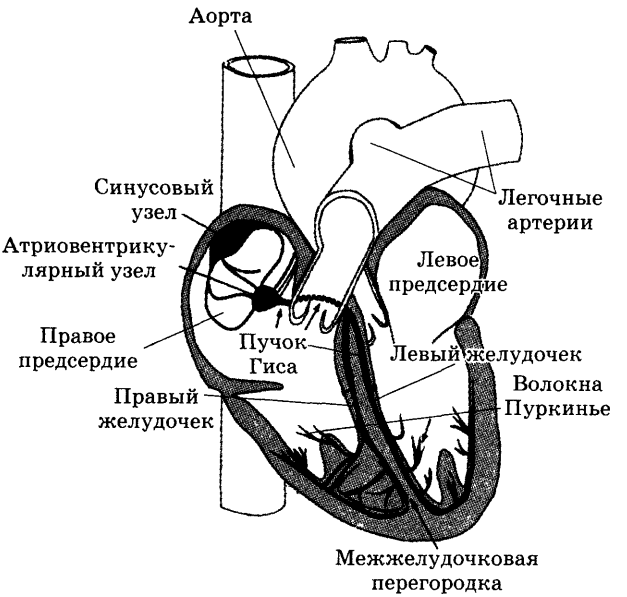


Рис. 12.12. Схематическое изображение сердца

Центр автоматии (пейсмекер), состоящий из специальных клеток, вырабатывает электрические импульсы, запускающие каждый цикл сокращения сердца. Этот электрический импульс запускает волну возбуждения, которая распространяется по проводящей системе предсердий со скоростью около 1 м/с, охватывая мышцу предсердия примерно за 0,08 с (т.е. 80 мс) и вызывая ее сокращение. Этому процессу соответствует первый, небольшой по амплитуде зубец *P* кардиограммы (см. рис. 12.11). Затем волна возбуждения достигает атриовентрикулярного узла, где скорость распространения электрического импульса резко падает (до 0,05 м/с) и происходит задержка передачи возбуждения на желудочки примерно на 0,05 с (сегмент *PQ*). В течение этого времени завершается наполнение желудочков сердца кровью. После окончания атриовентрикулярной задержки электрический импульс быстро распространяется по правой и левой ножкам пучка Гиса проводящей системы, которые у вершины сердца разветвляются на сеть мелких проводящих волокон Пуркинье,

уже непосредственно передающих возбуждение мышечным волокнам правого и левого желудочков, вызывая их сокращение (зубцы Q, R и S). Общая длительность комплекса зубцов QRS в норме составляет около 80 мс. Затем примерно через 0,2 с начинается восстановление (реполяризация) электрического состояния покоя желудочков, чему соответствует зубец T длительностью 120–160 мс, и расслабление мышцы желудочков. Вся систола желудочков (интервал QT) длится 0,3–0,4 с. После паузы (диастолы), длящейся около 0,4 с, пейсмекер запускает новую волну возбуждения сердца.

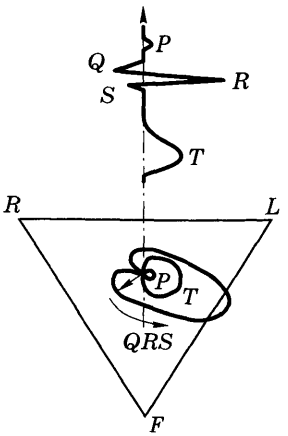


Рис. 12.13. Схема формирования ЭКГ в отведении I

Одновременная регистрация электрокардиограммы в нескольких отведениях позволяет определить поведение интегрального вектора сердца во времени. Оказывается, что в течение сердечного цикла этот вектор непрерывно изменяет свою длину и направление и за один кардиоцикл описывает в пространстве три петли разной амплитуды. Первая петля P имеет наименьший размах и соответствует возбуждению предсердий. Вторая петля QRS имеет наибольшую амплитуду и соответствует возбуждению желудочков, а третья петля T соответствует реполяризации желудочков. Проекции этих трех петель на линии отведений и определяют вид электрокардиограмм в этих отведениях (рис. 12.13).

Следует иметь в виду, что электрическая активность миокарда всегда опережает его механическое сокращение. Так, зубец P возникает на ЭКГ примерно за 20 мс до сокращения предсердий, а комплекс QRS — за 40 мс до начала сокращения желудочков.

Контрольные вопросы и задания

1. Какие виды электрических зарядов вы знаете? Укажите единицы их измерения в СИ и величину элементарного заряда. Сформулируйте закон сохранения электрического заряда.

2. Что определяет закон Кулона и в каких случаях его можно применять?
3. Укажите основную силовую характеристику электрического поля и запишите формулу для ее вычисления в случае, когда поле создается точечным зарядом. В чем заключается принцип суперпозиции электрических полей?
4. Укажите основную энергетическую характеристику электрического поля и запишите формулу для ее вычисления, если поле создается точечным зарядом. Как найти потенциал поля, создаваемого системой зарядов?
5. Как связаны напряженность и потенциал электрического поля?
6. Что такое электрический диполь? Как рассчитать потенциал поля, создаваемого диполем? Приведите соответствующую формулу.
7. Как связано напряжение между двумя точками электрического поля с параметрами диполя, создающего его? Приведите формулу.
8. Что такое «мультиполи»? Охарактеризуйте их. Создают ли они электрические поля и чем они различаются? В чем суть теоремы о мультипольном разложении электрических полей? В чем состоит дипольное приближение?
9. Почему определенные ткани и органы создают собственные электрические поля, от чего они зависят? Что такое интегральный электрический вектор органа? Что такое электрограмма органа и какие виды электрограмм вы знаете?
10. Каковы основные положения теории Эйнтховена? Что такое стандартные отведения, усиленные униполярные отведения? Как они связаны с интегральным электрическим вектором сердца? Что представляют собой грудные отведения и почему их необходимо регистрировать?
11. Нарисуйте примерный вид ЭКГ и укажите, каким физиологическим процессам соответствуют ее основные зубцы и временные интервалы.

Глава 13 Электропроводимость биологических тканей и жидкостей для постоянного тока

Электрический ток обычно определяют как упорядоченное движение свободных носителей зарядов (электронов, ионов и др.). Такой ток называют *током проводимости*. Сила тока и, следовательно, его воздействие на среду, в которой он распространяется, зависят не только от величины приложенной разности потенциалов, но и от свойств самой среды, определяющих ее электропроводимость.

Биологические ткани весьма неоднородны по строению и включают как электропроводящие жидкости (кровь, лимфа, спинномозговая жидкость и др.), так и диэлектрические ткани (кожа, кость, жировая клетчатка).

13.1. Постоянный ток в электролитах

В состав практически любого биологического объекта входят элементы, обладающие свойствами электролитов (тканевая жидкость, цитоплазма клеток и др.). Поэтому перед рассмотрением особенностей прохождения тока через биологические ткани напомним некоторые общие законы прохождения тока через растворы электролитов.

Для возникновения постоянного тока в некоторой среде необходимо осуществление двух условий:

1) наличие свободных электрических зарядов в этой среде;

2) наличие электрического поля, вызывающего направленное движение этих зарядов.

В растворах электролитов свободные электрические заряды (положительные и отрицательные ионы) возникают в результате электролитической диссоциации, а под действием приложенной внешней разности потенциалов U происходит направленное движение ионов в растворе — идет электрический ток. Полная сила тока I через электролит определяется движением как положительных (составляющая I_+), так и отрицательных (I_-) ионов: $I = I_+ + I_-$.

Основной закон прохождения тока по однородным проводникам — закон Ома, согласно которому сила тока I пропорциональна напряжению U на проводнике:

$$I = U/R. \tag{13.1}$$

Сопротивление R однородного проводника, как известно, зависит от его длины l, площади поперечного сечения S и удельного сопротивления проводника ρ:

$$R = \rho l/S. \tag{13.2}$$

Для описания прохождения токов через сложные неоднородные среды используют закон Ома в дифференциальной форме. Для получения соответствующего выражения преобразуем формулу (13.1), подставив в нее выражение (13.2):

$$I = US/\rho l.$$

Учитывая, что $I/S = j$ — плотность тока, а напряженность однородного электрического поля равна $E = U/l$, получим $j = E/\rho$. Обозначив $\gamma = 1/\rho$, окончательно запишем закон Ома в дифференциальной и векторной форме

$$\vec{j} = \gamma \vec{E}. \tag{13.3}$$

Удельная электропроводимость электролитов γ — величина, обратная удельному их сопротивлению ρ, имеет размерность $\text{Ом}^{-1} \cdot \text{м}^{-1}$ или См/м. Напомним, что единицей измерения электропроводимости является сименс (См): 1 См = 1 Ом $^{-1}$.

Установим зависимость удельной электропроводимости γ электролита от его свойств. Скорость υ упорядоченного движения ионов электролита пропорциональна напряженности электрического поля, вызывающего это движение:

$$\upsilon = bE.$$

Коэффициент пропорциональности b зависит от массы иона, его заряда, формы и называется подвижностью ионов. Обозначим ее для положительных ионов b_+, для отрицательных b_-. Подвижность ионов численно равна отношению ско-

ростей v_+ и v_- установившегося движения ионов к величине напряженности E поля, которое вызвало это движение:

$$b_+ = \frac{v_+}{E};\ b_- = \frac{v_-}{E}\left(\frac{\text{м}^2}{\text{В}\cdot\text{с}}\right).$$

Выделим некоторый объем электролита, по которому протекает постоянный электрический ток. Пусть полярность напряжения такова, что положительные ионы движутся через этот объем слева направо, а отрицательные — справа налево (рис. 13.1).

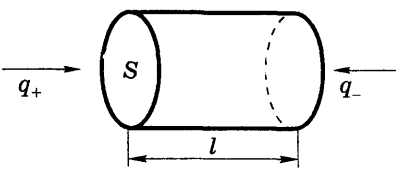


Рис. 13.1. Выделенный объем электролита

Тогда суммарный заряд Q ионов, переносимый за время t через плоскую поверхность площадью S, будет определяться формулой

$$Q = Q_+ + Q_- = q_+ n_+ S l_+ + q_- n_- S l_-,$$

где q_+ и q_- — заряды положительных и отрицательных ионов; l_+ и l_- — расстояние, которое пройдут заряды за время t; n_+ и n_- — концентрации ионов.

Учитывая, что $l_+ = v_+ t = b_+ Et$ и $l_- = v_- t = b_- Et$, получим

$$Q = (q_+ n_+ b_+ + q_- n_- b_-) StE. \qquad (13.4)$$

Разделив правую и левую части формулы (13.4) на площадь S и время t, получим выражение для плотности тока $j = \frac{I}{S} = \frac{Q}{St}$ через электролит:

$$j = (q_+ n_+ b_+ + q_- n_- b_-) E. \qquad (13.5)$$

Сравнивая закон Ома в дифференциальной форме (13.3) и формулу (13.5), получим выражение для удельной электропроводимости электролита:

$$\gamma = q_+ n_+ b_+ + q_- n_- b_-. \qquad (13.6)$$

Если считать, что величина заряда и концентрация положительного и отрицательного ионов одинакова: $|q_+| = |q_-| = q$ и $n_+ = n_- = \alpha n$, где α — коэффициент диссоциации, n — концен-

трация молекул растворенного вещества, то выражение для удельной электропроводимости электролита (13.6) принимает более простой вид:

$$\gamma = \alpha q n (b_+ + b_-). \qquad (13.7)$$

Таким образом, проводимость электролита зависит от заряда ионов, их концентрации и подвижности.

13.2. Особенности электропроводимости биологических тканей

Поскольку в структуру живых тканей входят электролиты, то при прохождении тока через ткань в определенной степени проявляются общие законы прохождения тока через электролиты. Однако биологические ткани содержат и элементы, обладающие выраженными свойствами диэлектриков — клеточные мембраны, которые играют большую роль в формировании механизмов прохождения тока через живую ткань. Если в электролитах под действием приложенной разности потенциалов возникает направленное движение ионов — электрический ток, то в диэлектриках в этом случае возникают явления поляризации — процессы смещения связанных электрических зарядов и образование из-за этого собственного электрического поля, напряженность которого направлена против внешнего поля. Виды поляризации могут быть различными.

Дипольная (*ориентационная*) *поляризация* состоит в том, что под действием внешнего поля отдельные полярные молекулы веществ ориентируются в соответствии с направлением этого поля. В растворах веществ, молекулы которых обладают большим дипольным моментом, дипольная поляризация вызывает внутри диэлектрика значительное уменьшение напряженности внешнего электрического поля. Так, для воды диэлектрическая проницаемость $\varepsilon = 81$.

Для биологических объектов весьма существенна *макроструктурная поляризация*. Она обусловлена тем, что электролиты, содержащиеся в структурных элементах клетки, окружены мембранами. Электрическое поле вызывает пере-

мещение ионов электролита внутри отдельного проводящего слоя, а прохождение ионов через окружающую мембрану затруднено из-за ее низкой проводимости. Таким образом, в структуре ткани возникают образования с установившимся разделением электрических зарядов, которые обладают гигантским (по сравнению с отдельными молекулами) дипольным моментом. Суммарное электрическое поле этих образований направлено против внешнего поля, чем и объясняется высокое значение диэлектрической проницаемости тканей, измеренной в постоянном электрическом поле.

Упрощенно механизм прохождения постоянного тока через ткань представлен на рис. 13.2. Основной тканевый ток определяется движением ионов в тканевой жидкости под действием приложенной разности потенциалов. Внутри клеточных структур происходит разделение зарядов и возникает разность потенциалов противоположного направления (электродвижущая сила поляризации), обусловливающая внутритканевый поляризационный ток. Это приводит к снижению электропроводимости тканей в целом по сравнению с тканевыми электролитами. Следует подчеркнуть, что макроструктурная поляризация, сильно влияющая на электропроводимость тканей, происходит не только на цитоплазматической мембране, но и на отдельных клеточных структурах, имеющих собственные мембраны.

Электрические свойства тканей и органов сильно различаются. Значения удельного сопротивления (ρ) и удельной

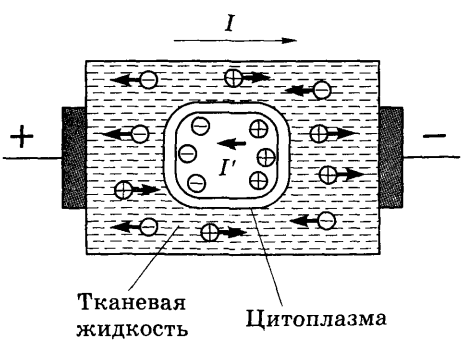


Рис. 13.2. Движение ионов в тканевой жидкости и цитоплазме: I — основной ток; I' — внутритканевый поляризационный ток

электропроводимости (γ) для некоторых тканей и жидкостей приведены в табл. 13.1.

Таблица 13.1

Ткань, электролит	ρ, Ом · м	γ, См/м
Спинномозговая жидкость	0,55	1,8
Кровь	1,66	0,6
Мышечная ткань	2	0,5
Ткань мозговая и нервная	14,3	0,07
Ткань жировая	33,3	0,03
Кожа (сухая)	10	10^{-5}
Кость без надкостницы	10^{7}	10^{-7}

Хорошей электропроводимостью обладают жидкие среды организма: спинномозговая жидкость, кровь, лимфа. Большое сопротивление имеют эпидермис кожи, соединительная ткань (связки, сухожилия) и, особенно, кость без надкостницы. Однако электропроводимость отдельных участков организма определяется не только значениями электропроводимости тканей, входящих в их состав, но и их топографией и функциональным состоянием.

Мягкие ткани внутренних органов человека обладают сравнительно небольшим сопротивлением — десятки и сотни Ом (в расчетах по обеспечению электробезопасности сопротивление тела человека без кожных покровов принимается равным 1000 Ом). В действительности, из-за наличия кожных покровов, имеющих высокое удельное сопротивление, измеряемое на поверхности тела, сопротивление отдельных его участков значительно больше.

Важнейшим фактором, определяющим электрическое сопротивление кожи, является толщина рогового слоя эпидермиса и его состояние. Если неороговевающие слои эпидермиса содержат до 70 % воды, то роговой слой — лишь 10 %, что обусловливает его высокое сопротивление. Однако при выделении пота и при наложении влажных электродных

прокладок роговой слой может впитывать воду, что снижает его сопротивление.

Основной путь электрического тока в коже — протоки потовых желез. Разветвление токов начинается на уровне 4-го слоя эпидермиса, который имеет относительно большое количество межклеточной жидкости, и достигает сильного ветвления на уровне дермы (соединительнотканной основы кожи). В дальнейшем ток распространяется по путям с наименьшим сопротивлением — структурам с максимальной электропроводимостью.

13.3. Некоторые лечебные методы, основанные на использовании постоянного тока

Гальванизация — это воздействие на тело человека постоянным электрическим током при напряжении до 80 В и силе тока до 50 мА. Предельно допустимая плотность тока — 0,1 мА/см^2. При осуществлении гальванизации между электродами и кожей помещают влажные прокладки, снижающие сопротивление кожи и устраняющие прижигающее действие тока под сухими электродами. Источником регулируемого постоянного напряжения является аппарат гальванизации, представляющий, в сущности, выпрямитель переменного тока.

Первичные физические механизмы воздействия постоянного тока на живую ткань обусловлены движением ионов, их разделением и изменением ионных концентраций в разных элементах ткани. Структурные перестройки в ткани, возникающие при токовых воздействиях, изменяют течение локальных биохимических процессов и формируют ответ на воздействие током на органном и организменном уровне, который в некоторых случаях проявляется в достижении положительного терапевтического эффекта.

Следует отметить, что выделение тепла при гальванизации крайне мало. Это объясняется тем, что при всех воздействиях на живую ткань постоянным током его плотность должна быть малой (не выше 0,1 мА/см^2). Ограничения на величину воздействующего постоянного тока налагают сопровождаю-

щие его прохождение явления электролиза. Если они станут существенными, то первичная структура живой ткани может быть нарушена. Поскольку количество теплоты Q определяется законом Джоуля — Ленца: $Q = I^2Rt$ (I — сила тока, R — сопротивление, t — время действия), а сила тока мала, то и тепловой эффект незначителен.

Лекарственный электрофорез представляет собой сочетанное воздействие на организм двух факторов: *физического* — электрического тока и *химического* — ионов лекарственного вещества, поступающих в организм с током через кожу или слизистые оболочки.

Между электродами и поверхностью тела помещают на прокладках раствор медикамента. Содержащиеся в этом растворе лекарственные ионы под действием приложенной разности потенциалов проникают внутрь тканей организма. Ионы лекарственных веществ вводятся с электрода одноименной полярности. Так, положительные ионы (катионы) вводят с анода (+), а отрицательные ионы (анионы) — с катода (–). Считается, что при электрофорезе лекарственные вещества проникают в верхние слои кожи и образуют там кожное депо, из которого в течение 3–20 дней диффундируют в кровь и лимфу и разносятся по всему организму.

Хотя электрический ток при осуществлении электрофореза и выполняет транспортную функцию, его нельзя рассматривать только как средство доставки лекарства в организм. Электрический ток, действуя на ткани, в том числе рецепторы, возбуждает их и приводит в состояние повышенной активности. При этом изменяется фармакокинетика — действие лекарств может усиливаться или ослабляться. Кроме того, вводимые лекарства могут усиливать или ослаблять действие тока. По сравнению с другими способами введения лекарств — инъекционным, внутривенным, ингаляционным и другим — воздействие при электрофорезе носит главным образом местный характер, причем в организм поступает лишь 2–5 % лекарственного вещества, используемого в процедуре. Однако терапевтический эффект может достигаться и при столь малых дозах.

Контрольные вопросы и задания

1. Выведите закон Ома в дифференциальной форме. Каков смысл входящих в него величин?
2. Как связаны удельная электропроводимость и удельное сопротивление электролита? В каких единицах они измеряются?
3. Какие характеристики электролита определяют его электропроводимость?
4. Что такое внутритканевый поляризационный ток и чем он обусловлен?
5. Почему физиотерапевтическое воздействие постоянным током на живую ткань не сопровождается заметным тепловым эффектом?
6. Чем различаются гальванизация и электрофорез?

Глава 14 Переменный ток. Электрический импеданс живой ткани. Основы реографии

Переменными называют токи и напряжения, значения которых изменяются по времени.

Обычно эти изменения имеют периодический характер. Амплитуда и частота напряжения обычно заданы заранее, а ток, возникающий в нагрузке, является следствием приложенного напряжения и электрических свойств нагрузки. Частота тока, как правило, совпадает с частотой напряжения, но амплитуда тока, сдвиг его начальной фазы и выделяющаяся средняя мощность полностью определяются свойствами нагрузки и существенно зависят от частоты.

Исследование закономерностей прохождения токов различных частот через живые биологические ткани позволило предложить ряд методов диагностики состояния органов и функциональных систем организма, а также методов воздействия переменными токами на организм с лечебной целью.

14.1. Основные характеристики переменного тока

Наиболее важное практическое значение имеют переменные токи, изменяющиеся по гармоническому закону:

$$U = U_m \sin\omega t,\ I = I_m \sin(\omega t + \varphi). \qquad (14.1)$$

Значения напряжения U и силы тока I, определяемые в различные моменты времени t формулами (14.1), называют ***мгновенными***, а их максимальные величины U_m и I_m — ***амплитудными***. Циклическая частота $\omega = 2\pi\nu = 2\pi/T$ (ν —линейная частота, T — период) тока и напряжения одинакова.

Разность фаз φ между током и напряжением называется ***фазовым сдвигом*** и зависит от вида нагрузки, включенной в конкретную электрическую цепь.

Усредненные по периоду значения модулей переменных токов $I_{эф}$ и напряжений $U_{эф}$ определяют их действие и называются ***эффективными*** или ***действующими***. Они связаны с амплитудными значениями простыми соотношениями

$$I_{эф} = \frac{I_m}{\sqrt{2}}, \quad U_{эф} = \frac{U_m}{\sqrt{2}}. \qquad (14.2)$$

Средняя мощность P, выделяемая в цепи переменного тока, определяется формулами

$$P = I_{эф}U_{эф}\cos\varphi, \quad P = \frac{1}{2}I_m U_m \cos\varphi \qquad (14.3)$$

и существенно зависит от сдвига фаз φ между током и напряжением, поэтому множитель $\cos\varphi$ часто называют ***коэффициентом мощности***.

Связь между переменными напряжением и током удобно представлять графически в виде векторных диаграмм (рис. 14.1), на которых амплитуды напряжения и тока отображаются векторами соответствующей длины, а угол между ними равен сдвигу фаз φ.

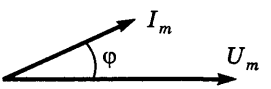


Рис. 14.1. Векторная диаграмма переменного тока и напряжения

Рассмотрим различные виды нагрузок в цепи переменного тока, полагая, что подаваемое напряжение всегда изменяется по закону $U = U_m \sin\omega t$.

14.2. Различные нагрузки в цепи переменного тока

Закон Ома, устанавливающий связь между сопротивлением цепи, постоянным током и напряжением, справедлив и для переменных токов. Только необходимо учитывать, что в данном случае сопротивление цепи и сдвиг фаз между током и напряжением зависят от вида нагрузки и частоты тока.

14.2.1. Активная нагрузка

Если источник напряжения замкнут на омическое сопротивление R (рис. 14.2), то в любой момент времени можно найти силу тока по закону Ома:

$$I = \frac{U}{R} = \frac{U_m}{R}\sin\omega t = I_m \sin\omega t .$$

В данном случае сила тока в цепи изменяется по такому же закону, что и напряжение, фазовый сдвиг $\varphi = 0$, поэтому мощность, выделяющаяся на нагрузке, максимальна (так как $\cos\varphi = 1$) и равна $P_R = I_{эф}U_{эф} = \frac{1}{2}I_m U_m$.

Поскольку на омическом сопротивлении R выделяется наибольшая тепловая мощность, его часто называют ***активным сопротивлением*** или ***активной нагрузкой***. Закон Ома при такой нагрузке выполняется также для амплитудных и эффективных значений тока и напряжения:

$$I_m = U_m/R,\ I_{эф} = U_{эф}/R.$$

Векторная диаграмма токов и напряжений при активной нагрузке имеет простой вид, представленный на рис. 14.3.

14.2.2. Емкостная нагрузка

Пусть переменное напряжение, изменяющееся по гармоническому закону $U = U_m \sin\omega t$, подается на конденсатор

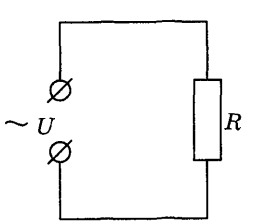


Рис. 14.2. Активная нагрузка в цепи переменного тока

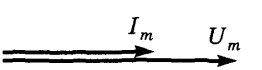


Рис. 14.3. Векторная диаграмма при активной нагрузке

емкостью C (рис. 14.4). Конденсатор начнет заряжаться и перезаряжаться с частотой переменного напряжения, и заряд на конденсаторе будет изменяться по закону

$$q = CU = CU_m \sin\omega t.$$

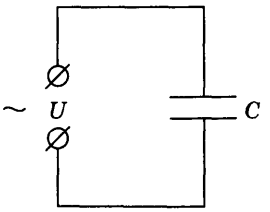


Рис. 14.4. Емкостная нагрузка в цепи переменного тока

Поскольку сила тока в цепи — это производная от заряда по времени $I = \dfrac{dq}{dt}$, то, выполнив дифференцирование, найдем ток в этой цепи:

$$I = C\omega U_m \cos\omega t = I_m \sin\left(\omega t + \frac{\pi}{2}\right), \qquad (14.4)$$

где $I_m = C\omega U_m$.

Величина

$$X_C = \frac{U_m}{I_m} = \frac{1}{\omega C} \qquad (14.5)$$

определяет сопротивление емкости C переменному току частотой ω и называется ***емкостным сопротивлением***. Видно, что X_C уменьшается с ростом частоты переменного тока и стремится к нулю, тогда как для постоянного тока ($\omega = 0$) сопротивление емкости бесконечно велико.

Из формулы (14.4) следует, что ток, идущий через конденсатор, по фазе опережает напряжение на нем на $\varphi = 90°$. Так как $\cos 90° = 0$, то средняя за период мощность, потребляемая чисто емкостной нагрузкой, равна нулю (первые полпериода конденсатор заряжается, а вторые — разряжается и отдает полученную энергию назад в сеть):

$$P_C = \frac{I_m U_m}{2} \cos 90° = 0.$$

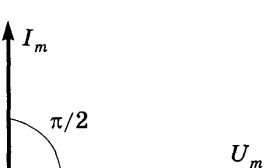


Рис. 14.5. Векторная диаграмма при емкостной нагрузке

Векторная диаграмма токов и напряжений при емкостной нагрузке имеет вид, представленный на рис. 14.5.

14.2.3. Индуктивная нагрузка

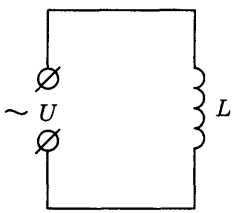


Рис. 14.6. Индуктивная нагрузка в цепи переменного тока

Подключим к сети переменного напряжения $U = U_m \sin\omega t$ катушку индуктивности L с малым активным сопротивлением $R = 0$ (рис. 14.6). Когда по катушке идет переменный ток, в ней возникает ЭДС самоиндукции, которая по закону Ленца противодействует приложенному напряжению:

$$\varepsilon_s = -L\frac{dI}{dt} = -U_m \sin\omega t.$$

Решение этого дифференциального уравнения относительно тока имеет вид:

$$I = -\frac{U_m}{\omega L}\cos\omega t = I_m \sin\left(\omega t - \frac{\pi}{2}\right). \qquad (14.6)$$

Видно, что в цепи с индуктивной нагрузкой ток отстает по фазе от напряжения на 90° (рис. 14.7), поэтому средняя мощность, выделяющаяся на чисто индуктивной нагрузке, равна нулю:

$$P_L = \frac{I_m U_m}{2}\cos 90° = 0.$$

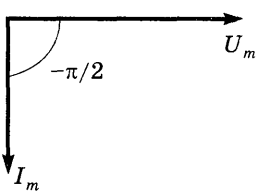


Рис. 14.7. Векторная диаграмма при индуктивной нагрузке

Сопротивление индуктивности переменному току на основании закона Ома

$$X_L = \frac{U_m}{I_m} = \omega L. \qquad (14.7)$$

Видно, что постоянному току ($\omega = 0$) чистая индуктивность L не оказывает сопротивления ($X_L = 0$), а ее сопротивление переменному току растет пропорционально частоте.

14.2.4. Полное сопротивление (импеданс) последовательной цепи переменному току

Рассмотрим цепь переменного тока, содержащую последовательно соединенные нагрузки трех видов (рис. 14.8). Для цепи постоянного тока ее полное сопротивление определялось бы как сумма сопротивлений всех последовательно включенных составляющих.

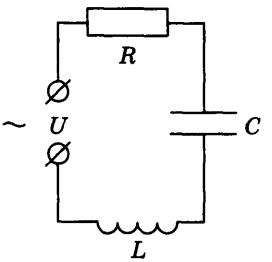


Рис. 14.8. Омическая, емкостная и индуктивная нагрузки в цепи переменного тока

В последовательной цепи переменного тока общим для всех нагрузок цепи является ток, а напряжения на каждом из элементов цепи сдвинуты по фазе относительно тока: напряжение на активной нагрузке совпадает по фазе с током, напряжение на емкости отстает от тока на 90°, а напряжение на индуктивности опережает ток по фазе на 90°.

Поэтому при определении полного сопротивления электрической цепи, представленной на рис. 14.8, необходимо учитывать фазовые соотношения между током и напряжением, зависящие от вида нагрузки.

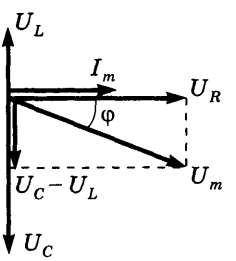


Рис. 14.9. Векторная диаграмма для последовательной цепи переменного тока

Рассчитать ток в такой цепи можно с помощью векторной диаграммы, представленной на рис. 14.9. Из нее видно, что

$$U_m = \sqrt{U_R^2 + (U_C - U_L)^2}\ . \qquad (14.8)$$

Поскольку амплитуды напряжений связаны с амплитудой тока соотношениями $U_R = I_m R$, $U_C = I_m X_C$, $U_L = I_m X_L$, то после подстановки получим

$$U_m = I_m \sqrt{R^2 + (X_C - X_L)^2}\ .$$

Теперь можем определить ***полное сопротивление Z (импеданс)*** последовательной цепи переменному току, учитывая формулы (14.5) и (14.7):

$$Z = \frac{U_m}{I_m} = \sqrt{R^2 + \left(\frac{1}{\omega C} - \omega L\right)^2}. \qquad (14.9)$$

Сдвиг фаз φ между полным током и напряжением в этой цепи определяется из соотношения

$$\mathrm{tg}\varphi = \frac{X_C - X_L}{R} = \frac{\dfrac{1}{\omega C} - \omega L}{R}. \qquad (14.10)$$

Амплитуда тока в этой цепи определяется по закону Ома

$$I_m = \frac{U_m}{Z} = \frac{U_m}{\sqrt{R^2 + \left(\omega L - \dfrac{1}{\omega C}\right)^2}}. \qquad (14.11)$$

Видно, что импеданс Z рассматриваемой цепи зависит не только от параметров нагрузок R, C и L, но и от частоты ω переменного напряжения. На рис. 14.10 приведены графики зависимости сопротивлений разных нагрузок от частоты переменного тока.

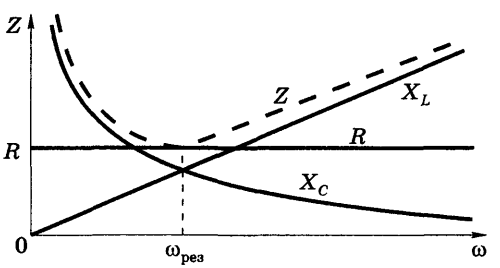


Рис. 14.10. Зависимость активного R, индуктивного X_L, емкостного X_C и полного Z (пунктир) сопротивлений цепи от частоты тока

Импеданс принимает наименьшее значение, равное активной нагрузке R, при $\omega L = \dfrac{1}{\omega C}$. В этом случае сила тока в цепи максимальна и в цепи наступает явление электрического резонанса. Поэтому частота $\omega_{рез} = \dfrac{1}{\sqrt{LC}}$, на которой наблюдается это явление, называется *резонансной частотой* данной цепи. При этом $Z = R$, а сдвиг фаз между током и напряжением $\varphi = 0$, т.е. на этой частоте цепь ведет себя как чисто активная нагрузка R.

14.3. Сопротивление живой ткани переменному току

Измерения полного сопротивления (импеданса) живой ткани, проведенные на разных частотах, показывают, что сопротивление ткани максимально и равно R_1 на постоянном токе ($\omega = 0$), а с увеличением частоты переменного тока импеданс сначала быстро уменьшается, а затем, достигнув некоторого значения Z_2, остается практически постоянным (рис. 14.11). Такая зависимость импеданса от частоты указывает на то, что в живой ткани нет элементов, обладающих индуктивностью, но есть элементы, обладающие свойствами емкости.

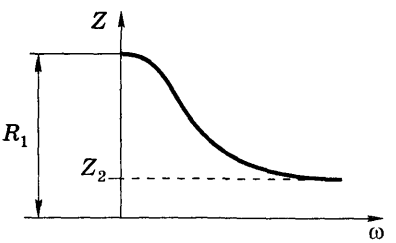


Рис. 14.11. Типичная зависимость импеданса живой ткани от частоты

Простейшая электрическая схема, импеданс которой имеет такую же частотную зависимость, что и импеданс живой ткани, представлена на рис. 14.12. Полное выражение импеданса такой цепи, эквивалентного импедансу живой ткани,

$$Z = \frac{R_1\sqrt{R_2^2 + X_C^2}}{\sqrt{(R_1 + R_2)^2 + X_C^2}},$$

или

$$Z = \frac{R_1\sqrt{1 + R_2^2\omega^2C^2}}{\sqrt{1 + (R_1 + R_2)^2\omega^2C^2}}. \qquad (14.12)$$

Емкостное сопротивление ткани $X_C = 1/\omega C$ определяется ее диэлектрическими структурными составляющими (клеточные мембраны, жировая клетчатка, эпидермис), а величина сопротивлений R_1 и R_2 (причем $R_1 >> R_2$) определяется омическими составляющими проводящих структур биологической ткани (кожи (R_1), тканевой жидкости, крови, цитоплазмы (R_2) и др.

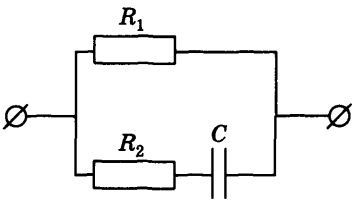


Рис. 14.12. Эквивалентная электрическая схема живой ткани

В представленной на рис. 14.12 эквивалентной схеме постоянный ток может идти только через сопротивление R_1, так как сопротивление емкости C для него бесконечно велико. Но для переменного тока сопротивление емкости уменьшается с увеличением частоты, а с ним уменьшается и полное сопротивление цепи. На очень высоких частотах сопротивление емкости стремится к нулю ($X_C \to 0$) и импеданс стремится к наименьшему значению:

$$Z_2 = \frac{R_1 R_2}{R_1 + R_2}. \qquad (14.13)$$

Следует иметь в виду, что каждая ткань характеризуется своими значениями параметров R_1, R_2 и C в эквивалентной схеме. Например, для кожи сопротивление на постоянном токе очень велико и составляет $R_1 \sim 10^4$–10^6 Ом, а на высоких частотах падает в 10–20 раз. Для мягких кровенаполненных тканей R_1 мало ($R_1 \sim 10^2$ Ом), меньше их емкостного сопротивления на низких частотах, поэтому часто эквивалентные схемы мягких тканей представлены только их активным сопротивлением R_1.

На средних и высоких частотах, для которых сопротивление нижней ветви цепи (см. рис. 14.7), состоящей из R_2 и X_2, будет значительно меньше R_1 и основной ток будет идти по нижней ветви, поэтому на этих частотах импеданс цепи можно приближенно выразить более простой формулой:

$$Z = \sqrt{R_2^2 + \left(\frac{1}{\omega C}\right)^2}. \qquad (14.14)$$

Для постоянного тока ($\omega = 0$) и низких частот эта формула неприменима, но на средних и высоких частотах она дает удовлетворительные результаты, если $R_1 >> R_2 + X_C$. Наименьшее значение импеданса ткани на высоких частотах определяет величина сопротивления R_2.

Зависимость импеданса ткани от частоты переменного тока определяется физиологическим состоянием и морфологическими особенностями ткани, что позволяет использовать измерения их электропроводимости в биологических и медицинских исследованиях. Методы измерения электропроводимости тканей осуществляются при достаточно низких напряжениях (менее 50 мВ) и слабых токах, которые не вносят изменений в физико-химические процессы в тканях и не повреждают их.

При действии повреждающих факторов (повышенная температура, мощный ультразвук, ионизирующие излучения и др.), а также при отмирании ткани происходит увеличение проницаемости мембран, их частичное или полное разрушение, что приводит к уменьшению роли емкостного сопротивления ткани и зависимость ее импеданса от частоты становится слабой. Для мертвой ткани она практически исчезает вообще.

Для иллюстрации на рис. 14.13 представлена частотная зависимость импеданса для трех образцов одной ткани: *1* — образец не подвергался никаким внешним воздействиям; *2* — ткань подвергнута кратковременному нагреванию, приводящему к частичному разрушению клеточных мембран; *3* — образец ткани, подвергнутой длительному кипячению, вызывающему полное разрушение мембран (мертвая ткань).

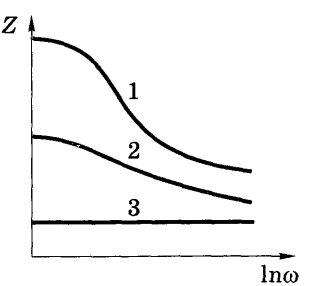


Рис. 14.13. Частотная зависимость импеданса для различных образцов тканей

Из рисунка видно, что сопротивление мертвой ткани практически не зависит от частоты. Поэтому по частотной зависимости импеданса можно оценивать *жизнестойкость* тканей организма.

Это существенно, в частности, для оценки качества трансплантанта при пересадке тканей и органов. По предложению Б.Н. Тарусова, жизнестойкость ткани можно количественно характеризовать с помощью так называемого *коэффициента поляризации* K, представляющего собой отношение импедан-

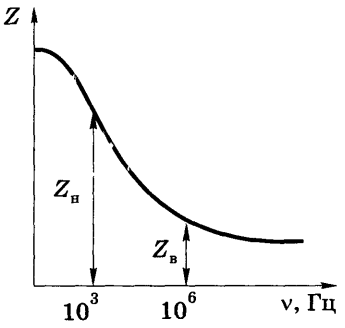


Рис. 14.14. Определение жизнестойкости ткани

са ткани $Z_н$ (рис. 14.14), измеренного на низкой частоте (порядка 10^3 Гц), к ее импедансу на высокой частоте $Z_в$ (10^6–10^7 Гц):

$$K = \frac{Z_н(\nu = 10^3)}{Z_в(\nu = 10^6)}. \qquad (14.15)$$

Для живой ткани этот коэффициент значительно превышает единицу ($K >> 1$) и зависит от способности ткани к обмену веществ. Так, для печени млекопитающих он равен 9–10, что выше, чем для мышц того же организма.

14.4. Физические основы реографии

Плетизмография — это совокупность методов регистрации пульсовых колебаний кровенаполнения исследуемого органа или его участков. Основные факторы, которые определяют характер кровенаполнения, — это работа сердца и сосудистой системы. Для оценки функционирования сердца широко применяется электрокардиография и совокупность других методов, физические основы которых рассматриваются в соответствующих разделах медицинской и биологической физики, а важные сведения о состоянии сосудистой системы отдельных органов дает анализ периодических изменений их кровенаполнения.

Регистрация периодически изменяющегося импеданса тканей лежит в основе важного метода исследования состояния сосудистой системы — ***импедансной плетизмографии (реографии)***.

Это обусловлено тем, что омическое сопротивление тканей сильно зависит от степени их кровенаполнения. Ткани неоднородны по своей структуре, а ток всегда идет по пути с наименьшим сопротивлением и прежде всего — по кровеносным сосудам. Поэтому при увеличении кровенаполнения ткани ее омическая составляющая R импеданса умень-

шается, а при уменьшении кровенаполнения — увеличивается. Таким образом, импеданс ткани периодически изменяется с частотой сердечных сокращений.

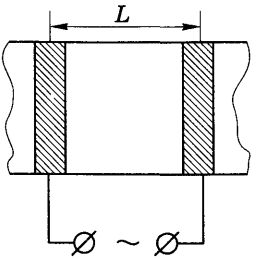


Рис. 14.15. Участок ткани между электродами

Омическое сопротивление R ткани определяется в основном объемом содержащихся в ней электролитов и, прежде всего, крови. Рассмотрим участок ткани между двумя электродами (рис. 14.15). Обозначим расстояние между электродами L, удельное сопротивление исследуемого участка ткани ρ, площадь поперечного сечения S проводника (сосудов, так как именно по ним в основном идет ток), объем проводящей ткани $V = LS$.

Омическое сопротивление проводника определяется формулой $R = \rho L/S$. Выражая площадь через объем $S = V/L$, получим

$$R = \rho L^2/V. \qquad (14.16)$$

Как следует из формулы (14.16), при приливе крови (V увеличивается) активное сопротивление R ткани уменьшается, а при оттоке крови оно возрастает. Поэтому изменение ΔR активного сопротивления органа или ткани связано с изменением ΔV объема его кровенаполнения:

$$\Delta R = -\frac{\rho L^2}{V^2}\Delta V .$$

Разделив это соотношение почленно на выражение (14.16), получим выражение

$$\frac{\Delta R}{R} = -\frac{\Delta V}{V}, \qquad (14.17)$$

которое показывает равенство относительных изменений этих величин. Знак «–» говорит о том, что с увеличением объема кровенаполнения ($\Delta V > 0$) активное сопротивление ткани уменьшается ($\Delta R < 0$).

При технической реализации метода возникает определенное затруднение. На постоянном токе и на токах низких частот проводить измерения нельзя по соображениям безопаснос-

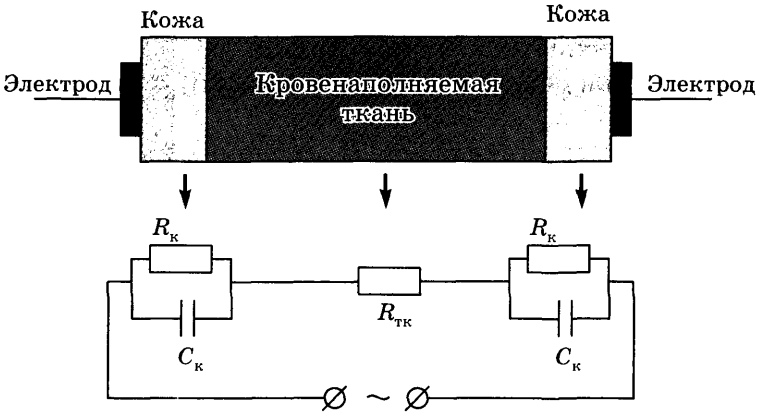


Рис. 14.16. Эквивалентная схема реографических исследований

ти, а на переменном токе сопротивление живой ткани (и сила проходящего через нее тока) определяется полным значением ее импеданса Z, а не только активной составляющей R. Поэтому нужно так реализовать метод, чтобы влияние емкостной составляющей импеданса при реографических исследованиях свести к минимуму и добиться выполнения соотношения $Z \approx R$.

Электрическая цепь, образующаяся при реографических исследованиях, представлена в виде эквивалентной схемы на рис. 14.16. Электроды, подсоединенные проводами к источнику тока, закрепляются на коже пациента через марлевые прокладки, смоченные физраствором или электропроводной пастой. Ток проходит через кожу (под каждым электродом) и исследуемую ткань. Активное сопротивление кожи $R_к$ велико (для чистой сухой кожи оно составляет порядка 1 МОм) и сравнимо с ее емкостным сопротивлением X_C, определяемым частотой тока и емкостью кожи $C_к$. Активное сопротивление кровенаполняемой ткани $R_{тк}$ (которое как раз и изменяется в такт с пульсациями крови в ней) намного меньше ее емкостного сопротивления, поэтому последнее можно не учитывать (поскольку $R_{тк}$ и $C_{тк}$ соединены параллельно).

Если бы для реографии использовался постоянный ток, то из-за большого сопротивления кожи $R_к$ практически невозможно было бы зарегистрировать малые изменения общего высокого сопротивления цепи, вызванные кровенаполнением

ткани $R_{тк}$. Поэтому для реографии применяют переменный ток частотой 40–150 кГц. На таких частотах емкостное сопротивление кожи $X_C = \dfrac{1}{\omega C_K}$ становится много меньше ее активного сопротивления $R_к$ и меньше сопротивления ткани $R_{тк}$, в результате общий импеданс $Z \approx R_{тк}$. В этом случае $\Delta Z = \Delta R$ и соотношение (14.17) принимает вид

$$\frac{\Delta Z}{Z} = -\frac{\Delta V}{V}, \qquad (14.18)$$

который показывает, что на высоких частотах относительное изменение импеданса полностью определяется относительным изменением кровенаполнения органа или ткани, что и лежит в основе метода импедансной плетизмографии (реографии).

Итак, использование в реографии переменного тока частотой 40–150 кГц позволяет выделить из общего сопротивления цепи переменный компонент ее омической составляющей, обусловленный пульсовыми колебаниями кровенаполнения исследуемой ткани.

Для получения более полной диагностической информации и ее дальнейшей интерпретации при реографических исследованиях синхронно с регистрацией изменений импеданса проводится и запись ЭКГ. Пример регистрации указанных показателей представлен на рис. 14.17. Период повторения

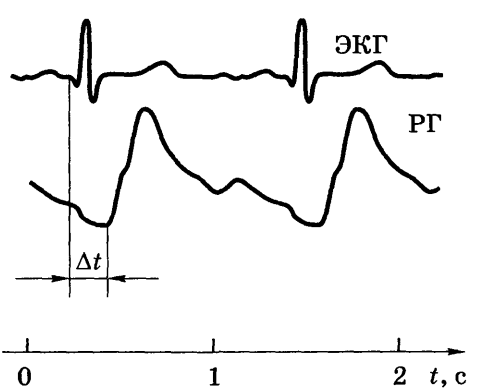


Рис. 14.17. Синхронная запись ЭКГ и РГ

приведенных кривых одинаков, но между ЭКГ и реограммой (РГ) есть сдвиг Δt во времени, определяемый временем распространения пульсовой волны от сердца до исследуемого органа.

Кроме того, часто синхронно записывают еще и продифференцированную реограмму (ее первую производную). Если реограмма РГ отражает зависимость объема кровенаполнения исследуемого участка тела от времени, то дифференцированная реограмма (исходя из физического смысла производной) отображает зависимость скорости кровенаполнения от времени.

Контрольные вопросы и задания

1. Какой ток называют переменным? Что такое амплитудное, мгновенное и эффективное значения переменного тока и напряжения?
2. Как и на что влияет сдвиг фаз между током и напряжением?
3. Что такое векторная диаграмма токов и напряжений? Как и для чего она применяется?
4. Какова связь между переменным напряжением, током и мощностью при активной, емкостной и индуктивной нагрузке?
5. Какова связь между переменным напряжением, током и мощностью в последовательной цепи, состоящей из индуктивности, емкости и активного сопротивления?
6. Что такое импеданс цепи, от чего он зависит, как его вычислить?
7. Как зависит импеданс живой ткани от частоты переменного тока? Приведите и охарактеризуйте эквивалентную схему живой ткани.
8. Какими морфологическими структурами определяются омическая и емкостная составляющие импеданса биологической ткани?
9. Что такое коэффициент поляризации ткани? Что он характеризует?
10. На чем основана импедансная плетизмография (реография)? Какую диагностическую информацию получают посредством этого метода?

Глава 15 Магнитное поле

Магнетизм как способность некоторых железных руд притягивать к себе другие тела был известен с древних времен. Первое объяснение этого явления, основанное на атомистических представлениях о строении вещества, дал еще Лукреций Кар в I в. до н.э. Один из первых известных научных трактатов «О магните, магнитных телах и большом магните — Земле» был издан В. Гильбертом в 1600 г. Однако окончательно физическая природа магнетизма была установлена лишь в XIX в. трудами Эрстеда, Ампера, Фарадея, Максвелла и других ученых. Теория Максвелла объединила электрические и магнитные явления, показала, что это две, хотя и несколько разные, стороны электромагнитных явлений.

Выяснилось, что магнитное поле создают движущиеся электрические заряды, проводники с током и переменные электрические поля, что магнитное поле оказывает силовое воздействие на намагниченные тела и на проводники с током. Магнитными свойствами обладают электроны и ядра атомов, поэтому все тела в той или иной мере имеют магнитные свойства и реагируют на действие внешних магнитных полей.

Магнитные взаимодействия широко используются в современной технике. Установлено также, что магнитные поля оказывают специфические воздействия на живые организмы и магнитные воздействия все шире используются в медицинской практике.

15.1. Общие свойства и характеристики магнитного поля

В настоящее время твердо установлено, что ***магнитное поле*** создается движущимися электрическими зарядами, т.е. электрическим током. Причем это могут быть:

- отдельные движущиеся электрические заряды;
- электрические токи в проводниках;

• молекулярные токи (электроны, движущиеся по атомным и молекулярным орбиталям).

Если эти токи постоянны во времени, то и создаваемые ими магнитные поля будут постоянными, а если токи изменяются во времени, то соответственно будет изменяться и их магнитное поле.

Установлено также, что ***переменное магнитное поле*** может создаваться не только переменным током, но и переменным электрическим полем.

Магнитное поле в свою очередь оказывает силовое действие только:

• на движущиеся электрические заряды;
• проводники с током;
• намагниченные тела.

Максимальная сила, действующая на проводник с током в магнитном поле, всегда пропорциональна произведению силы тока I на длину l проводника: $F_{max} \sim Il$, но различна в разных магнитных полях. Поэтому отношение максимальной силы F_{max}, действующей на проводник, к произведению Il является силовой характеристикой магнитного поля и называется ***магнитной индукцией***: $B = \dfrac{F_{max}}{Il}$. За единицу магнитной индукции в международной системе единиц (СИ) принимается 1 тесла (1 Тл) — это индукция магнитного поля, действующего с максимальной силой в 1 Н на проводник длиной 1 м при силе тока в нем 1 А. Внесистемной единицей магнитной индукции является 1 эрстед $= 10^{-5}$ Тл.

Магнитная индукция B — величина векторная, на графиках ее направление определяет направление линий индукции. ***Линия магнитной индукции*** — направленная линия, касательная к которой в каждой точке совпадает с направлением вектора магнитной индукции в этой точке.

Напомним, что силовые линии электростатического поля начинаются и заканчиваются на электрических зарядах. Однако магнитных зарядов нет, поэтому линии индукции магнитного поля замыкаются сами на себя; такие поля называют ***вихревыми***.

Магнитное поле называют ***однородным***, если величина и направление магнитной индукции одинаковы во всех точках рассматриваемого объема среды: $\vec{B} = \text{const}$. Линии индукции однородного магнитного поля параллельны друг другу и замыкаются через бесконечность.

Величина магнитной индукции dB, создаваемая небольшим элементом dl проводника с током I, расположенным на расстоянии r от точки наблюдения A (рис. 15.1), определяется ***законом Био — Савара — Лапласа:***

$$dB = \mu\mu_0 \frac{Idl\sin\alpha}{4\pi r^2}, \qquad (15.1)$$

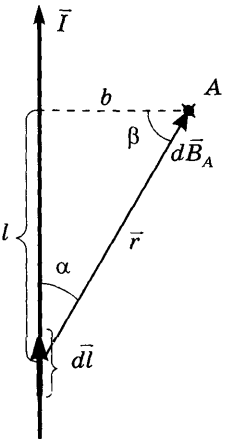


Рис. 15.1. Магнитная индукция, создаваемая проводником с током

где μ — магнитная проницаемость среды; $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м — магнитная постоянная (магнитная проницаемость вакуума); α — угол между направлением тока в элементе dl проводника и направлением на точку наблюдения A.

Полная магнитная индукция B, создаваемая проводником с током I в точке A, находится интегрированием выражения (15.1) по всей длине проводника.

15.1.1. Магнитное поле бесконечного прямого тока

Вычислим магнитную индукцию в некоторой точке A, отстоящей от проводника с током на расстояние b.

Для того чтобы упростить интегрирование выражения (15.1), выразим переменные r, dl и α через расстояние b и угол $\beta = 90° - \alpha$, (см. рис. 15.1):

$$r = \frac{b}{\cos\beta}, \quad \sin\alpha = \cos\beta, \quad l = b\,\text{tg}\beta \rightarrow dl = b\frac{d\beta}{\cos^2\beta}.$$

После подстановки этих соотношений в выражение (15.1) и его интегрирования по углу β в пределах от $(-\pi/2)$ до $(+\pi/2)$ получим выражение для индукции магнитного поля, созда-

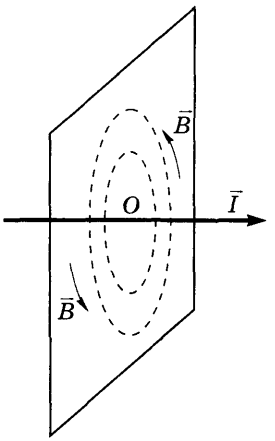


Рис. 15.2. Линии магнитной индукции B (пунктир) прямого тока

ваемого бесконечным прямым проводником с током I на расстоянии b от проводника:

$$B_A = \frac{\mu\mu_0 I}{2\pi b}. \qquad (15.2)$$

Линии магнитной индукции в данном случае имеют вид *концентрических окружностей*, расположенных в плоскостях, перпендикулярных току I, а центром этих окружностей является точка O пересечения тока с плоскостью (рис. 15.2).

Направление линий магнитной индукции определяется ***правилом правого буравчика***: если правый буравчик (винт) вращать так, чтобы он двигался в направлении тока I, то направление вращения рукоятки покажет направление линий магнитной индукции.

15.1.2. Магнитное поле кругового тока

Найдем индукцию магнитного поля в центре кругового тока I (рис. 15.3), для чего проинтегрируем (15.1) по всей длине проводника от 0 до $2\pi r$ и учтем, что здесь $\alpha = \pi/2$, поэтому $\sin\alpha = 1$:

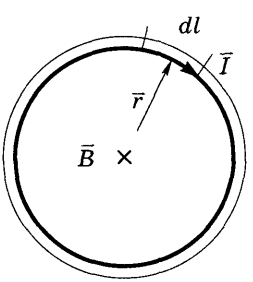


Рис. 15.3. Магнитное поле кругового тока

$$B = \frac{\mu\mu_0}{4\pi}\int_0^{2\pi r}\frac{I}{r^2}dl = \frac{\mu\mu_0 I}{4\pi r^2}2\pi r = \frac{\mu\mu_0 I}{2r}.$$

Итак, в центре кругового тока магнитная индукция равна:

$$B = \frac{\mu\mu_0 I}{2r}. \qquad (15.3)$$

Линии магнитной индукции в этом случае перпендикулярны плоскости проводника и замкнуты сами на себя, их на-

правление внутри кольца тоже определяется ***правилом правого буравчика***: если рукоятку правого буравчика (винта) вращать по направлению кругового тока I, то направление движения буравчика покажет направление линий магнитной индукции внутри кольца, а снаружи их направление будет противоположным (рис. 15.4)

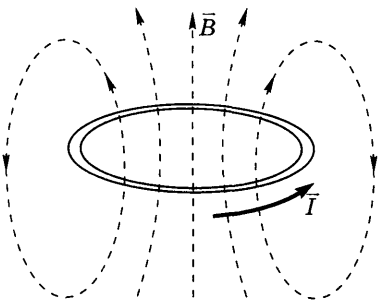


Рис. 15.4. Линии магнитной индукции кругового тока

15.2. Силовое действие магнитного поля

Силовое воздействие магнитного поля на намагниченные тела и железные предметы известно давно и широко используется в практических целях. Действие магнитного поля на электрические заряды и токи известно сравнительно недавно, с XIX в., но нашло широкое применение в физике и многочисленных технических приборах и устройствах. Именно действие магнитного поля на движущиеся электрические заряды и проводники с током лежит в основе работы многих измерительных приборов, электродвигателей, движения поездов на магнитной подушке, МГД-генераторов и многих других технических решений.

15.2.1. Действие магнитного поля на движущийся заряд. Сила Лоренца

На электрический заряд q, движущийся со скоростью υ в магнитном поле индукцией B, действует сила Лоренца, которая определяется выражением

$$F_{\text{Л}} = q\upsilon B\sin\alpha\,, \qquad (15.4)$$

где α — угол между векторами скорости $\vec{\upsilon}$ и магнитной индукции $\vec{B}$.

Направление силы Лоренца, действующей на положительный заряд, определяется ***правилом левой руки***: если распо-

ложить левую руку так, чтобы вектор магнитной индукции входил в ладонь, а четыре пальца левой руки направить вдоль вектора скорости, то оттянутый на 90° большой палец покажет направление силы Лоренца (см. рис. 15.6, *а*).

Если движущийся заряд отрицательный, то направление силы Лоренца будет противоположным.

Свойства силы Лоренца:

- $F_Л$ всегда перпендикулярна плоскости, в которой лежат векторы $\vec{v}$ и $\vec{B}$, и, следовательно, перпендикулярна каждому из них $\vec{F}_Л \perp \vec{v}, \vec{F}_Л \perp \vec{B}$;
- магнитное поле не действует на заряд, движущийся параллельно вектору магнитной индукции, так как в этом случае $\alpha = 0$ и $F_Л = 0$;
- сила Лоренца максимальна, если заряд движется перпендикулярно линиям магнитной индукции ($\sin 90° = 1$):

$$F_{max} = qvB.$$

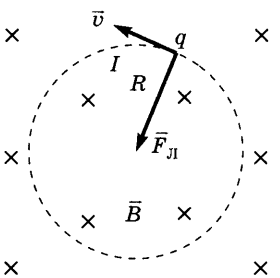


Рис. 15.5. Движение заряда в однородном магнитном поле, где $\vec{v} \perp \vec{B}$

Благодаря этим свойствам силы Лоренца электрический заряд q массой m, влетевший в однородное магнитное поле перпендикулярно линиям индукции B, будет двигаться *по окружности* радиусом R (рис. 15.5), поскольку сила Лоренца будет создавать центростремительное ускорение ($\vec{F}_Л \perp \vec{v}$):

$$qvB = \frac{mv^2}{R}. \qquad (15.5)$$

Находим, что радиус этой окружности прямо пропорционален скорости движения заряда:

$$R = \frac{mv}{qB}, \qquad (15.6)$$

тогда как период и частота вращения заряда по окружности не зависят от скорости и полностью определяются свойствами заряда (q/m) и индукцией B магнитного поля:

$$T = \frac{2\pi R}{v} = \frac{2\pi m}{qB}; \quad \nu = \frac{1}{T} = \frac{qB}{2\pi m}. \qquad (15.7)$$

Если заряд влетает в магнитное поле под углом $0 < \alpha < 90°$, то он будет двигаться по *винтовой траектории* радиусом R_B и шагом h, равными

$$R_B = \frac{m\upsilon}{qB}\sin\alpha\,,\quad h = T\upsilon\cos\alpha = \frac{2\pi m}{qB}\upsilon\cos\alpha\,. \qquad (15.8)$$

Ось этой винтовой траектории параллельна линиям индукции магнитного поля, а период и частота вращения определяются соотношениями (15.7).

15.2.2. Действие магнитного поля на проводник с током. Сила Ампера

На проводник с током I длиной l в магнитном поле индукцией B действует сила Ампера, величина которой определяется выражением

$$F_A = IBl\sin\alpha\,, \qquad (15.9)$$

где α — угол между направлением тока $\vec{I}$ и вектором магнитной индукции $\vec{B}$.

Направление силы Ампера тоже определяется ***правилом левой руки***: если расположить левую руку так, чтобы вектор магнитной индукции входил в ладонь, а четыре пальца левой руки направить вдоль вектора тока, то оттянутый на 90° большой палец покажет направление силы Ампера (рис. 15.6, *б*).

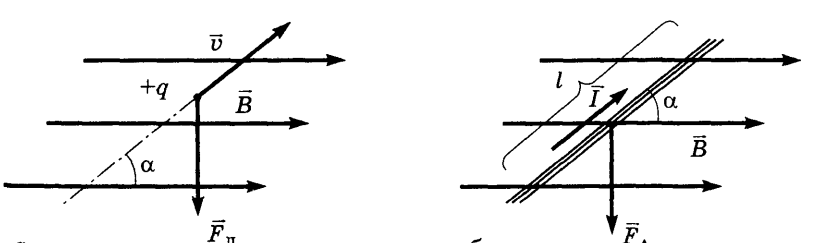


Рис. 15.6. Направление сил действия магнитного поля B:
а — на движущийся электрический заряд — силы Лоренца;
б — на проводник с током — силы Ампера

Свойства силы Ампера:

- она всегда перпендикулярна плоскости, в которой лежат векторы $\vec{I}$ и $\vec{B}$, и, следовательно, перпендикулярна каждому из них: $\vec{F}_A \perp \vec{I}$, $\vec{F}_A \perp \vec{B}$;
- магнитное поле не действует на проводник с током, параллельный вектору магнитной индукции, так как при $\alpha = 0$ и $F_A = 0$;
- сила Ампера максимальна, если проводник с током расположен перпендикулярно линиям магнитной индукции ($\sin 90° = 1$): $F_{max} = IBl$.

Нетрудно показать, что сила Ампера есть векторная сумма всех сил Лоренца, действующих на отдельные заряды, движущиеся в проводнике, поэтому у этих сил много общих свойств.

15.2.3. Действие магнитного поля на рамку с током

Если рамку с током поместить в магнитное поле, то она будет поворачиваться, пока ее плоскость не установится перпендикулярно линиям магнитной индукции $\vec{B}$. Механический момент M, вращающий рамку, создают две противоположно направленные силы Ампера $\vec{F}_1$ и $\vec{F}_3$, действующие на противоположные (вертикальные на рис. 15.7) стороны рамки с током и образующие пару сил, тогда как силы Ампера $\vec{F}_2$ и $\vec{F}_4$, действующие на две другие (верхнюю и нижнюю) стороны рамки, компенсируют друг друга, поскольку направлены противоположно и вдоль одной прямой.

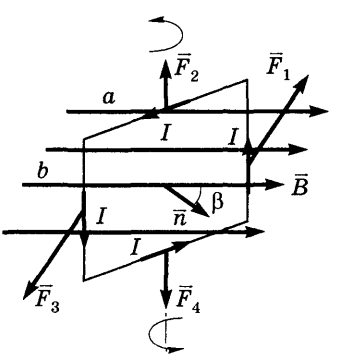


Рис. 15.7. Силы, действующие на рамку с током в магнитном поле

Величина вращающего механического момента зависит от угла β между вектором индукции $\vec{B}$ и нормалью (перпендикуляром) $\vec{n}$ к плоскости рамки (рис. 15.7):

$$M = ISB\sin\beta = p_m B\sin\beta . \qquad (15.10)$$

Произведение силы тока I на площадь рамки S называют ***магнитным моментом*** рамки с током:

$$p_m = IS. \qquad (15.11)$$

Он представляет собой вектор, перпендикулярный плоскости рамки и направленный в соответствии с правилом правого буравчика (на рис. 15.7 — вдоль нормали $\vec{n}$).

Из соотношения (15.10) видно, что механический момент максимален, когда плоскость рамки параллельна линиям магнитной индукции ($\beta = 90°$, $\sin\beta = 1$ и $M_{max} = ISB$). Когда плоскость рамки установится перпендикулярно линиям магнитной индукции ($\beta = 0$), то вращающий момент станет равен нулю: $M = 0$, так как будет $\sin\beta = 0$. Поэтому в магнитном поле рамка с током всегда поворачивается так, чтобы ее плоскость стала перпендикулярна вектору индукции $\vec{B}$.

Благодаря этим свойствам рамку с током используют для обнаружения магнитных полей и для определения направления и величины (измеряя M_{max}) магнитной индукции B.

Следует отметить, что явление вращения рамки с током в магнитном поле широко используется в электродвигателях, электроизмерительных приборах и других технических устройствах.

Сила Ампера определяет также взаимодействие проводников с током: если по двум параллельным проводникам токи идут в одном направлении, то проводники притягиваются, а если в противоположных направлениях — отталкиваются. Это объясняется тем, что каждый проводник с током создает вокруг себя магнитное поле, которое действует силой Ампера на другой проводник с током, притягивая его или отталкивая, в зависимости от направления тока в нем.

15.3. Магнитный поток, закон электромагнитной индукции

Если некоторый плоский контур площадью S находится в однородном магнитном поле индукцией $\vec{B}$, то скалярная величина Φ, равная произведению площади рамки на проекцию вектора магнитной индукции на нормаль к плоско-

сти контура, называется ***магнитным потоком*** через этот контур:

$$\Phi = BS\cos\alpha, \qquad (15.12)$$

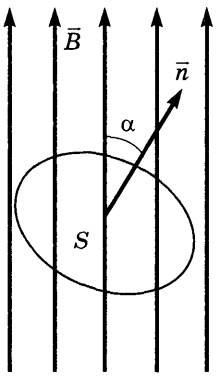


Рис. 15.8. Магнитный поток

где α — угол между нормалью $\vec{n}$ к плоскости контура и вектором $\vec{B}$ (рис. 15.8).

Магнитный поток максимален, если плоскость контура перпендикулярна линиям индукции (α = 0 и cosα = 1), и поток равен нулю, если она параллельна им (α = 90° и cosα = 0).

За единицу измерения магнитного потока в СИ принят 1 вебер: 1 Вб = 1 Тл · м^2 — это магнитный поток, создаваемый магнитным полем индукцией 1 Тл через контур площадью в 1 м^2, расположенный перпендикулярно линиям индукции.

М. Фарадей в 1831 г. экспериментально установил ***закон электромагнитной индукции***: любое изменение магнитного потока через контур ведет к возникновению в нем электрической ЭДС, равной скорости изменения магнитного потока во времени:

$$\text{ЭДС} = -\frac{d\Phi}{dt}. \qquad (15.13)$$

Если контур проводящий и замкнутый, то в нем под действием этой ЭДС возникает индукционный ток

$$I_{\text{инд}} = \frac{\text{ЭДС}}{R} = -\frac{1}{R}\frac{d\Phi}{dt}, \qquad (15.14)$$

где R — электрическое сопротивление контура.

Знак «−» в формулах (15.13) и (15.14) отражает ***правило Ленца***: индукционный ток, возникающий в замкнутом контуре, направлен таким образом, что создаваемое им магнитное поле препятствует изменению магнитного потока, вызвавшего этот индукционный ток.

Из формулы (15.12) следует, что магнитный поток Φ может изменяться во времени тремя способами:

- за счет изменения во времени магнитной индукции B;

• за счет изменения ориентации α контура в магнитном поле;

• в результате изменения площади S, т.е. деформации контура.

Во всех этих случаях в контуре будет возникать ЭДС индукции, тем большая, чем быстрее изменяются во времени эти параметры.

Если проводник помещен в переменное магнитное поле $B = B_0\cos\omega t$, то в нем всегда возникает переменная ЭДС индукции:

$$\text{ЭДС}_i = -SB' = (B_0 S\omega)\sin\omega t = E_0\sin\omega t,$$

под действием которой в объеме проводника возникают переменные индукционные вихревые (замыкающиеся сами на себя) токи (15.14), называемые также ***токами Фуко***. Эти токи вызывают значительный нагрев проводника, что широко используется и в медицине (индуктотермия), и в технических приложениях (индукционные печи и др.).

15.4. Самоиндукция. Индуктивность контура. Энергия магнитного поля

Рассмотрим замкнутый контур, по которому течет ток I (рис. 15.9). Этот ток создает магнитное поле, индукция B которого пропорциональна силе тока I [см. формулу (15.3)]. Соответственно магнитный поток Φ через этот контур также будет пропорционален силе тока в контуре:

$$\Phi = LI. \qquad (15.15)$$

Коэффициент пропорциональности L является характеристикой контура и называется ***коэффициентом самоиндукции*** или ***индуктивностью контура***. Он зависит только от геометрии контура, его размеров и магнитных свойств окружающей его среды. Единицей измерения индуктивности в СИ является 1 генри (Гн) = 1 Вб/А. Это индуктивность контура, в котором ток в 1 А создает магнитный поток через контур в 1 Вб.

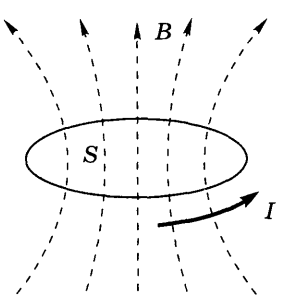


Рис. 15.9. Магнитный поток через контур, создаваемый током в нем

Если ток в контуре постоянный, то и магнитный поток Φ постоянный. Но если сила тока I в контуре изменяется (по любым причинам), то изменяется и магнитный поток Φ через контур, в результате чего в контуре возникает ЭДС самоиндукции, пропорциональная скорости изменения тока в контуре и препятствующая этому изменению (на это указывает знак «–»):

$$\text{ЭДС}_{\text{самоинд}} = -\frac{d\Phi}{dt} = -L\frac{dI}{dt}. \qquad (15.16)$$

При очень быстром замыкании или размыкании цепи величина этой ЭДС может быть очень большой, и вызываемые ею индукционные токи (*экстратоки*) могут выводить из строя приборы (например, электролампы), включенные в эту цепь. Поэтому мощные нагрузки включаются в цепь медленно, через специальные устройства (реостаты и др.).

Магнитное поле, как и любое другое силовое поле, обладает энергией $W_м$, которая зависит от силы тока I, создающего это магнитное поле, и магнитного потока Φ через контур с током, а также с учетом соотношения (15.15) от силы тока и индуктивности L контура:

$$W_м = \frac{1}{2}\Phi I = \frac{1}{2}LI^2 = \frac{\Phi^2}{2L}. \qquad (15.17)$$

Формула (15.17) подобна соответствующей формуле для энергии $W_э$ электрического поля в конденсаторе емкостью C, заряженном до напряжения U ($q = CU$ — заряд конденсатора):

$$W_э = \frac{1}{2}qU = \frac{1}{2}CU^2 = \frac{q^2}{2C}. \qquad (15.18)$$

Таким образом, электрический ток создает магнитное поле, энергия которого определяется силой тока и индуктивностью контура. Любые изменения силы тока в контуре создают в нем ЭДС самоиндукции, препятствующую этим изменениям тока.

15.5. Магнитные свойства веществ

Если внести вещество в магнитное поле индукцией B_0, то внутри этого вещества индукция магнитного поля будет уже другой: $B_{ср} = \mu B_0$. Величина $\mu = \dfrac{B_{ср}}{B_0}$ называется ***относительной магнитной проницаемостью*** этого вещества.

По значению относительной магнитной проницаемости все вещества делятся на три группы:

- диамагнетики — вещества, для которых $\mu < 1$, $B_{ср} < B_0$;
- парамагнетики — вещества, для которых $\mu > 1$, $B_{ср} > B_0$;
- ферромагнетики — вещества, для которых $\mu >> 1$, $B_{ср} >> B_0$.

Для диамагнитных и парамагнитных веществ отклонение значений величины μ от 1 весьма невелико (десятые и сотые доли), тогда как для ферромагнетиков $\mu \sim 1000$ и более.

Объяснение таких разных магнитных свойств веществ состоит в следующем. Молекулярные токи создают магнитные моменты p_m молекул, вследствие чего образуется собственное магнитное поле среды $\vec{B}_{соб}$, прямо пропорциональное векторной сумме магнитных моментов молекул среды: $\vec{B}_{соб} \sim \Sigma \vec{p}_m$. Если эту среду внести во внешнее магнитное поле индукцией B_0, то индукция результирующего поля в среде будет равна векторной сумме:

$$\vec{B}_{ср} = \vec{B}_0 + \vec{B}_{соб}. \qquad (15.19)$$

Молекулы *диамагнетиков* не имеют собственного магнитного момента ($p_m = 0$), но во внешнем магнитном поле в этих молекулах возникает наведенный магнитный момент $p_m \sim B_0$ и, соответственно, намагниченность $\vec{B}_{соб}$, направленные против вектора индукции $\vec{B}_0$. Поэтому индукция результирующего поля в диамагнетике будет меньше внешнего $B_{диа} = B_0 - B_{соб} < B_0$, вследствие чего у диамагнетиков $\mu < 1$.

Диамагнетики, помещенные в неоднородное магнитное поле, выталкиваются в область более слабого поля.

Молекулы *парамагнетиков* имеют собственные магнитные моменты $p_m \neq 0$, но в отсутствие внешнего поля они ориентированы хаотически и их векторная сумма близка к нулю: $\Sigma \vec{p}_m = 0$. Во внешнем поле индукцией B_0 происходит

ориентация магнитных $\vec{p}_m$ молекул вдоль направления $\vec{B}_0$ и возникает намагниченность среды $\vec{B}_{соб} \sim \Sigma \vec{p}_m \sim \vec{B}_0$, направленная параллельно индукции внешнего поля, поэтому результирующее поле в парамагнетиках несколько превосходит внешнее: $B_{пара} = B_0 + B_{соб} > B_0$, следовательно и $\mu > 1$.

Парамагнетики, помещенные в неоднородное магнитное поле, втягиваются в область более сильного поля.

Молекулы *ферромагнетиков* тоже имеют собственные магнитные моменты $p_m \neq 0$, но в отсутствие внешнего поля они ориентированы не хаотически, а содержат крупные области (домены) со спонтанной намагниченностью $\vec{B}_j \sim (\Sigma \vec{p}_m)_j \neq 0$ (рис. 15.10). При внесении ферромагнетика во внешнее поле B_0 происходит ориентация полей $\vec{B}_j$ отдельных доменов вдоль направления $\vec{B}_0$, благодаря чему индукция суммарного поля в ферромагнетике в десятки и сотни раз превосходит индукцию внешнего поля: $B_{фер} = B_0 + \Sigma B_j >> B_0$, поэтому и $\mu >> 1$.

Если теперь убрать внешнее поле, то суммарное магнитное поле доменов, которое в сотни раз больше B_0, будет создавать постоянное магнитное поле ферромагнетика и поддерживать прежнюю ориентацию магнитных полей отдельных доменов, образуя постоянный магнит: $\vec{B}_{фер} = \sum_j \vec{B}_j$.

В силу этих свойств ферромагнетики широко применяются в технике для изготовления постоянных магнитов, для записи информации на ферромагнитные пленки, диски и используются в других устройствах.

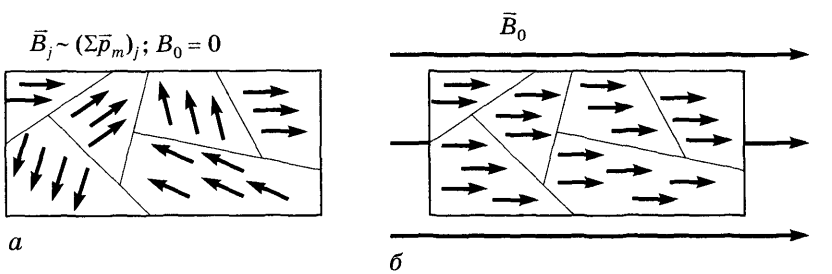


Рис. 15.10. Области спонтанной намагниченности (домены) в ферромагнетиках:
а — в отсутствие внешнего поля; *б* — во внешнем поле B_0

Ферромагнетики, как и парамагнетики, при внесении в неоднородное магнитное поле втягиваются в область более сильного поля.

15.6. Магнитные свойства биотканей

Пассивные магнитные свойства биотканей обусловлены химическими и физическими свойствами веществ, входящих в их состав. Основные химические компоненты сред организма (вода, белки, углеводы, липиды) диамагнитны, поэтому и биологические ткани в основном диамагнитны, их магнитная проницаемость чуть меньше единицы. Однако в организме в небольшом количестве есть и парамагнитные соединения, например свободные радикалы.

Некоторые микроорганизмы способны синтезировать вещества, обладающие ферромагнитными свойствами, и накапливать их в специализированных *органеллах*. Предполагают, что эти органеллы способны ориентироваться во внешнем магнитном поле подобно магнитной стрелке компаса и что подобные ферромагнитные включения присутствуют в тканях высокоорганизованных существ (пчелы, птицы, дельфины и др.).

Ферромагнитные образования имеются и в тканях человека. Они обнаружены в селезенке, печени, надпочечниках. Роль этих образований в протекании физиологических процессов до конца не выяснена.

На действие магнитного поля реагируют все клеточные органеллы, особенно митохондрии. К магнитному влиянию весьма чувствительны нервная ткань, секреторный эпителий эндокринных желез, сердечно-сосудистая система и система крови. Магнитное поле влияет на физико-химические процессы, протекающие в живых тканях. Физической основой такого действия поля могут быть ориентация молекул, изменение пространственного распределения ионов в неоднородном магнитном поле, действие силы Лоренца на движущиеся ионы, возникновение индукционных ЭДС и токов под действием переменных и импульсных магнитных полей и др.

Активные магнитные свойства живых тканей определяются свойствами их собственных магнитных полей, порождаемых электрическими токами, возникающими в тканях в процессе жизнедеятельности.

Например, распространение электрического импульса по аксону — это, в сущности, прохождение электрического тока, который в свою очередь создает магнитное поле в окружающей среде.

О величине магнитных полей, возникающих при функционировании некоторых органов, можно судить по табл. 15.1, в которой для сравнения приведены также соответствующие данные для магнитного поля Земли и физиотерапевтической аппаратуры. Из таблицы видно, что собственные магнитные поля биологических объектов крайне слабые, что обусловливает трудность их регистрации.

Таблица 15.1

Объект	Мозг	Мышца	Сердце	Земля	Терапевтические устройства
Индукция магнитного поля, мТл	10^{-9}	10^{-8}	10^{-7}	$5 \cdot 10^{-2}$	1–100

В основе ***магнитодиагностики*** лежит регистрация собственных магнитных полей биотканей, возникающих в результате физиологических процессов, протекающих в них. Несмотря на малую величину индукции этих магнитных полей, проводится их регистрация с диагностической целью.

Наиболее разработана *магнитокардиография*, поскольку магнитное поле сердца обладает относительно высоким значением магнитной индукции (см. табл. 15.1). Магнитокардиограмма имеет вид, аналогичный электрокардиограмме. Сохранены даже обозначения зубцов. Определенным достоинством магнитокардиографии является получение диагностических данных без непосредственного контакта электродов с поверхностью тела и некоторых новых сведений о функционировании сердца.

При регистрации *магнитоэнцефалограмм* обнаружены новые характеристики активности мозга, не проявляющиеся на электроэнцефалограммах. В частности, оказалось возможным выделить высокочастотную составляющую магнитного поля отдельных нейронов мозга человека.

В настоящее время наибольшее применение магнитных полей в медицине связано с их воздействием на организм при

физиотерапии, а также с использованием сильных градиентных магнитных полей (индукцией 0,5–2,0 Тл) для получения изображения органов и тканей методом ядерного магнитного резонанса (магнитно-резонансная томография).

15.7. Биологическое действие магнитного поля

На молекулярном уровне экспериментально изучено действие переменного магнитного поля на конформацию (изменение пространственной трехмерной структуры) белков сыворотки крови, тканей и ферментов и на показатели энергетического обмена. Исследования выявили отчетливое влияние магнитного поля малой индукции на структуру определенных химических групп белков, активность ферментов, устойчивость липопротеиновых комплексов, резистентность мембраны эритроцитов.

В *постоянном магнитном поле* индукцией 0,01–1 Тл наблюдаются магнитомеханические эффекты в жидкокристаллических структурах, которые могут определять уровень активности ферментов на биомембране и влиять на проницаемость клеточных мембран.

Действию магнитного поля подвержены клеточные и не клеточные компоненты крови, что обусловливает влияние постоянного магнитного поля на свертываемость крови, микроциркуляцию и проницаемость сосудов.

Особенно следует обратить внимание на действие *переменного магнитного поля.* В этом случае, кроме магнитного взаимодействия биомолекул с магнитным полем, происходит взаимодействие биосистем с переменным электрическим полем, индуцируемым в тканях переменным магнитным полем. Напряженность этого электрического поля прямо пропорциональна скорости изменения магнитного поля и, следовательно, его частоте. Так как в тканях имеются свободные заряженные частицы (ионы и электроны), то индуцируемое электрическое поле приводит их в движение, возникает электрический ток, обладающий, как известно, многообразным биологическим действием.

В настоящее время в лечебной практике достаточно широко и активно используются разнообразные магнитные поля

(МП): постоянные, переменные, импульсные, вращающиеся и др. Такое широкое применение обусловлено как большой глубиной проникновения МП в ткани, так и разнообразными механизмами его действия. МП способны влиять на пространственную ориентацию сложных биологически активных веществ (ферментов, нуклеиновых кислот, протеинов и т.д.), подвижность ионов, что может отражаться на кинетике и скорости биохимических реакций, идущих с участием этих веществ. Действие МП может реализовываться и через химические реакции, протекающие по свободнорадикальному механизму, поскольку свободные радикалы парамагнитны и могут взаимодействовать с внешним МП.

Среди *лечебных эффектов* МП наиболее доказанными и значимыми являются: противоотечный, гипотензивный, противовоспалительный, иммуномодулирующий, метаболический, седативный и др. Магнитное поле оказывает нормализующе-стимулирующее действие на функциональное состояние внутренних органов, улучшает микроциркуляцию и транскапиллярный обмен, благоприятно влияет на нейровегетативные процессы.

Считается, что пороговая чувствительность к импульсным магнитным полям составляет 0,1 мТл, к переменным — 3 мТл, к постоянным — 8 мТл.

Медицинской промышленностью выпускается большое количество магнитотерапевтических аппаратов. В качестве источников постоянного магнитного поля используются эластичные магниты-аппликаторы, магнитные клипсы, браслеты и другие изделия на основе постоянных магнитов с индукцией 25–40 мТл.

Источниками переменного и импульсного магнитного поля являются индукторы разнообразной формы, использующие переменные и импульсные токи разных частотных диапазонов: от единиц герц до десятка мегагерц (индуктотермия) — и индукцией от нескольких десятков до нескольких сотен миллитесла.

Одним из перспективных направлений развития магнитотерапии является разработка методов сочетанного воздействия магнитного поля и другого лечебного воздействия, что часто обеспечивает их бóльшую терапевтическую эффектив-

ность и более короткий курс лечения. К таким методам уже можно отнести магнитолазерную терапию, магнитофорез лекарств и др.

Контрольные вопросы и задания

1. Что является источником магнитного поля? На какие объекты магнитное поле может оказывать силовое воздействие?
2. Укажите силовую характеристику магнитного поля и единицы ее измерения. Как графически изображаются магнитные поля?
3. Что определяет закон Био — Савара — Лапласа? Как с его помощью можно вычислить индукцию магнитного поля в точке наблюдения?
4. Как рассчитать величину и определить направление вектора индукции магнитного поля, создаваемого прямым бесконечным током?
5. Как рассчитать величину и определить направление вектора индукции магнитного поля, создаваемого круговым током?
6. Что характеризует сила Лоренца? Как рассчитать ее величину и определить направление?
7. На что действует сила Ампера? Как рассчитать ее величину и определить направление?
8. Что называют магнитным моментом рамки с током? В каких единицах СИ он измеряется?
9. Какое действие оказывает магнитное поле на рамку с током? Приведите выражение для механического момента, действующего на рамку с током в магнитном поле.
10. Что такое магнитный поток? Укажите единицу его измерения в СИ и способы измерения.
11. В чем состоит явление электромагнитной индукции? Запишите этот закон. В чем состоит правило Ленца?
12. В чем состоит явление самоиндукции? Что такое индуктивность контура, от чего она зависит и в каких единицах СИ измеряется? Что такое экстратоки, когда и почему они возникают?
13. Приведите формулы для расчета энергии магнитного поля.
14. Каковы магнитные свойства веществ (диа-, пара- и ферромагнетиков)? Объясните физическую природу различий их свойств.
15. Охарактеризуйте пассивные и активные магнитные свойства биотканей.
16. Дайте характеристику биологическому действию магнитного поля.
17. Каковы основные области применения магнитных полей в лечебных целях?

Глава 16 Основы электростимуляции тканей и органов

Электростимуляция — это дозированное воздействие электрическим током на определенные органы и возбудимые ткани организма с целью обеспечения или улучшения их деятельности. Ее применяют для нормализации функций двигательных нервов и мышц, используя низкочастотные (до 200 Гц) импульсные токи различной формы.

Используемые при электростимуляции токи могут вызывать стимуляцию нервов и сокращения мышц, не управляемые ЦНС, поэтому они могут быть опасны для организма и требуют строгого соблюдения правил техники безопасности. Так, ток частотой 50 Гц и силой 10–15 мА вызывает сокращение мелких скелетных мышц (например, мышц кисти, гортани и др.), и человек не может самостоятельно освободиться от его действия. Такой ток называют *неотпускающим* током. Если полный ток через грудную клетку человека достигает 50–60 мА, то наступает паралич дыхательных мышц (межреберных и мышц, регулирующих движение диафрагмы) и через несколько минут может наступить смерть от удушья. Если же сила тока через грудную клетку достигает 100 мА, может произойти фибрилляция сердечной мышцы и последующая остановка сердца. Поэтому при наложении электродов следует обращать особое внимание на то, чтобы стимулирующий ток не проходил по жизненно важным органам, а его амплитуда была меньше значений поражающего тока $I_{\text{пораж}}$ для стимулируемого органа или ткани. Реально используются токи амплитудой не более 50 мА.

Отметим, что, хотя к пациенту подводится электрическое напряжение, основным фактором, действующим на орган или ткань при электростимуляции, является сила тока, возникающая под действием этого напряжения. Поэтому далее будет рассмотрена именно форма и сила импульсного тока, действующего на ткань при электростимуляции.

16.1. Параметры импульсных токов

Для электростимуляции используют импульсные токи разной формы, в зависимости от вида ткани и лечебной цели воздействия. Наиболее простую форму имеют прямоугольные импульсные токи и именно такие токи обычно дают мультивибраторы — физические электронные устройства, генерирующие импульсы напряжения с заданной частотой.

Импульсные токи прямоугольной формы обычно используются для стимуляции нервной системы, а токи непрямоугольной формы — для стимуляции мышц.

Напомним, что для полной характеристики постоянного тока достаточно указать его величину (I) и направление, для переменного синусоидального тока — его амплитуду (I_0) и частоту (ν или ω). Однако для описания импульсных токов требуется бóльшее количество характеристик.

16.1.1. Прямоугольные импульсные токи

Для полного описания импульсного тока прямоугольной формы (рис. 16.1) необходимо указать его амплитуду I_0 и два временных параметра: длительность импульса $t_и$ (именно она определяет время воздействия тока на живую ткань при электростимуляции) и период повторения импульсов T либо длительность паузы t_0, которые связаны соотношением

$$T = t_и + t_0. \qquad (16.1)$$

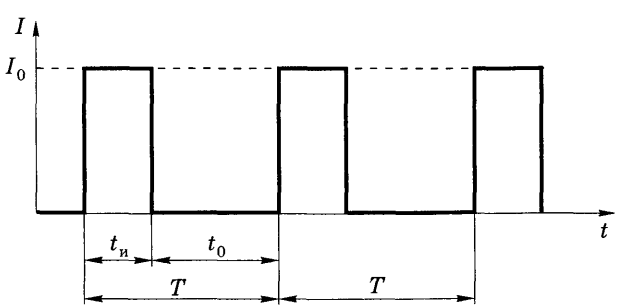


Рис. 16.1. Прямоугольные импульсные токи и их параметры

Зная период T импульсного тока и длительность импульса $t_и$, можно рассчитать другие используемые параметры импульсных токов:

- частоту импульсного тока $\nu = \frac{1}{T}$ (Гц); (16.2)
- скважность $Q = \frac{T}{t_и}$; (16.3)
- коэффициент заполнения $k = \frac{1}{Q} = \frac{t_и}{T}$. (16.4)

Скважность Q — безразмерная величина, которая показывает, во сколько раз период повторения T больше длительности импульса $t_и$. Скважность, как следует из ее определения, всегда больше единицы: $Q = \frac{T}{t_и} = \frac{t_и + t_0}{t_и} = 1 + \frac{t_0}{t_и}$.

Коэффициент заполнения k — величина, обратная скважности, показывает какую долю периода занимает сам импульс.

Если длительность импульса равна длительности паузы $t_и = t_0$, то импульсный ток называют ***симметричным***, при этом скважность $Q = 2$.

Таким образом, для полного описания прямоугольного импульсного тока необходимо указать как минимум три параметра:

1) амплитуду импульсного тока I_0;
2) частоту ν импульсного тока или его период $T = 1/\nu$;
3) скважность Q или длительность импульса $t_и$.

Остальные параметры (длительность паузы, коэффициент заполнения и др.) нетрудно рассчитать по приведенным выше формулам.

16.1.2. Непрямоугольные импульсные токи

Для описания ***непрямоугольных импульсных токов***, представленных на рис. 16.2, необходимо ввести некоторые дополнительные параметры, характеризующие форму этих импульсов. Даже определение прежних параметров — длительности импульса и паузы — теперь неоднозначно, так как зависит от того, на каком уровне их измерять. Поэтому условились в электростимуляции измерять длительность импуль-

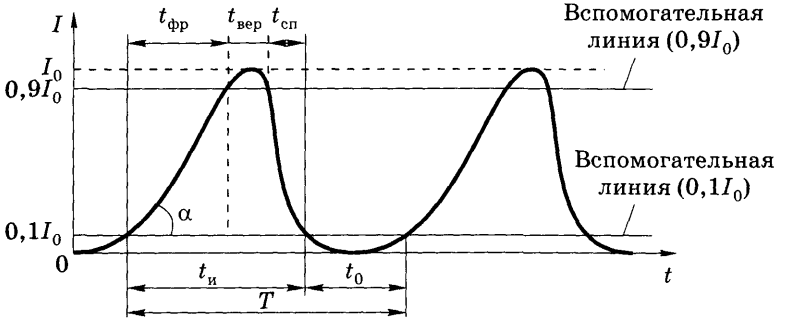


Рис. 16.2. Параметры непрямоугольных импульсных токов

са и паузы на уровне 10 % амплитудного значения тока, т.е. на уровне $0,1I_0$. Проводят вспомогательную линию на этом уровне (см. рис. 16.2), и точки ее пересечения с фронтами импульса позволяют однозначно определить длительность импульса $t_и$, паузы t_0 и периода T. Таким образом, длительность импульса сложной формы — это время, в течение которого сила тока равна или превышает значение $0,1I_0$.

Частота импульсного тока, скважность и коэффициент заполнения вычисляются по формулам (16.2)–(16.4). Однако для полного описания непрямоугольных импульсных токов этих характеристик уже недостаточно и приходится использовать дополнительные параметры:

- *длительность фронта* импульса $t_{фр}$ — это время, в течение которого сила тока в импульсе нарастает от значения $0,1I_0$ до уровня $0,9I_0$;
- *крутизна фронта* импульса определяет скорость нарастания тока во времени от значения $0,1I_0$ до значения $0,9I_0$:

$$Кр = \frac{0,8I_0}{t_{фр}}; \qquad (16.5)$$

- *длительность спада* импульса $t_{сп}$ — это время, в течение которого сила тока в импульсе спадает от значения $0,9I_0$ до значения $0,1I_0$;
- *длительность вершины* импульса $t_{вер}$ — это время, в течение которого сила тока в импульсе равна или превышает значение $0,9I_0$.

Для нахождения этих дополнительных параметров на графике импульсного тока проводят вторую вспомогательную линию на уровне 90 % амплитуды импульса: $0{,}9I_0$ (см. рис. 16.2). Отметим, что сумма длительностей фронта, вершины и спада равна длительности импульса:

$$t_и = t_{фр} + t_{вер} + t_{сп}. \qquad (16.6)$$

Таким образом, для полного описания непрямоугольного импульсного тока надо указать как минимум пять параметров:

1) амплитуду импульсного тока I_0;

2) частоту ν импульсного тока или его период $T = 1/\nu$;

3) скважность Q или длительность импульса $t_и$;

4) крутизну Кр фронта импульса или его длительность $t_{фр}$;

5) длительность спада $t_{сп}$ или длительность вершины $t_{вер}$ импульса.

Остальные параметры импульсного тока можно определить по формулам (16.1)–(16.6).

16.2. Физиологическое обоснование параметров импульсных токов при электростимуляции

Параметры импульсных токов, применяемых для электростимуляции, должны соответствовать как физиологическим особенностям живого организма, так и требованиям техники безопасности. Поэтому в этом разделе будет рассмотрено влияние параметров импульных токов на электростимуляцию органов и тканей человека и определены их допустимые физиологически обоснованные значения.

16.2.1. Амплитуда импульсного тока

Опыт показывает, что физиологический ответ возбудимой ткани на действие электрического тока (генерация потенциала действия в клетках, возникновение нервных импульсов, сокращение мышечных волокон и т.п.) возникает лишь в том случае, когда сила тока равна или превышает определенное значение $I_{пор}$, называемое ***пороговым током***, т.е. при $I_{стим} \geq I_{пор}$.

При этом сила тока не должна превышать безопасных значений

$$I_{\text{пор}} < I_{\text{стим}} < I_{\text{пораж}}. \qquad (16.7)$$

Величина порогового тока $I_{\text{пор}}$ зависит не только от вида ткани, но и от длительности и формы импульса. Однако есть и общая закономерность: с увеличением длительности импульса пороговый ток $I_{\text{пор}}$ уменьшается, как показано на рис. 16.3, а начиная с некоторой длительности $t_{\text{пол}}$ (так называемое «полезное время») уже не изменяется, достигнув своего наименьшего значения R, называемого ***реобазой***. График зависимости порогового тока $I_{\text{пор}}$ от длительности прямоугольного импульса $t_{\text{и}}$, представленный на рис. 16.3, называют ***кривой электровозбудимости*** ткани.

Таким образом, реобаза R представляет собой наименьшее значение порогового тока для данной ткани, наблюдаемое при $t_{\text{и}} \geq t_{\text{пол}}$.

Вторым параметром, характеризующим электровозбудимость ткани, является ***хронаксия*** $t_{\text{хр}}$ — длительность импульса, для которого пороговый ток $I_{\text{пор}}$ вдвое больше реобазы: $I_{\text{пор}}(t_{\text{хр}}) = 2R$.

Каждому импульсному току амплитудой I и длительностью $t_{\text{и}}$ соответствует своя точка на плоскости ($I_{\text{пор}}$, $t_{\text{и}}$), представленной на рис. 16.3. Если эта точка находится ниже кривой электровозбудимости, то такой ток меньше порогового и не сможет возбудить ткань, а если эта точка находится на кривой или выше нее, то $I_{\text{стим}} \geq I_{\text{пор}}$, и ткань возбудится.

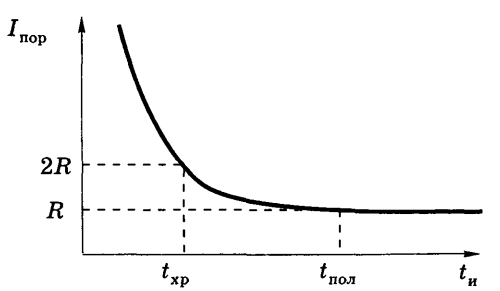


Рис. 16.3. Зависимость величины порогового тока от длительности импульса

Зависимость порогового тока от длительности прямоугольного импульса, представленная на рис. 16.3, приближенно описывается ***уравнением Вейса — Лапика***:

$$I_{\text{пор}} = \frac{a}{t_{\text{и}}} + b\,, \tag{16.8}$$

где a, b — константы, зависящие от вида ткани.

Полезность уравнения (16.8) в том, что, зная константы a и b, можно вычислить величину порогового тока $I_{\text{пор}}$ для любой длительности импульса.

Значения констант a и b в уравнении (16.8) нетрудно связать с определяемыми опытным путем реобазой R и хронаксией $t_{\text{хр}}$. Для длинных импульсов при $t_{\text{и}} > t_{\text{пол}}$ пороговый ток одинаков и равен реобазе R, что соответствует константе b в уравнении Вейса — Лапика. Действительно при больших $t_{\text{и}} \to \infty$ будет $I_{\text{пор}} \to b$, значит, $b = R$ и измеряется в единицах тока (мА).

Если же $t_{\text{и}} = t_{\text{хр}}$, то $I_{\text{пор}} = 2R$, и уравнение Вейса — Лапика принимает вид

$$2R = \frac{a}{t_{\text{хр}}} + R\,,$$

отсюда получаем соотношение

$$a = Rt_{\text{хр}}. \tag{16.9}$$

При некоторых заболеваниях возбудимость нервов и мышц падает и количественной мерой тяжести патологических процессов служат отклонения значений реобазы и хронаксии, полученных экспериментально для исследуемой ткани, от их «нормальных» значений для этой ткани. Эта идея лежит в основе *хронаксиметрии* — метода диагностики состояния нервно-мышечной системы человека путем экспериментального определения реобазы и хронаксии для исследуемой ткани. Эти параметры обычно определяют для импульсных токов прямоугольной формы.

16.2.2. Длительность импульса

Длительность импульса $t_{\text{и}}$ при электростимуляции непосредственно влияет на величину порогового тока, как видно из рис. 16.3 и уравнения Вейса — Лапика (16.8). Наиболее

оптимальным будет воздействие импульсными токами длительностью $t_и \geq t_{пол}$, поскольку в этом случае величина порогового тока минимальна (равна реобазе). При меньших длительностях импульса $t_и$ пороговый ток будет больше и соответственно приходится увеличивать амплитуду стимулирующего тока. Полезное время $t_{пол}$ зависит от вида ткани и может составлять от нескольких единиц до сотен миллисекунд.

16.2.3. Частота импульсного тока

Частота ν стимулирующего импульсного тока может составлять от нескольких долей герца до сотен герц. Максимальная частота возбуждения ткани определяется ее абсолютным рефрактерным периодом, поскольку для нормального ее возбуждения необходимо, чтобы период следования стимулирующих импульсов был больше $T_{рефр}$. ***Абсолютный рефрактерный период*** $T_{рефр}$ — это время, в течение которого клетку нельзя возбудить никаким стимулом. Поэтому физиологически обоснованная максимальная частота возбуждения клеток ткани $\nu_{max} = 1/T_{рефр}$ и имеет примерно следующие значения:

- для нервных клеток $\nu_{max} = 500–1000$ Гц ($T_{рефр} = 1 - 2$ мс);
- скелетных мышц $\nu_{max} = 100–200$ Гц ($T_{рефр} = 5 - 10$ мс);
- сердечной мышцы $\nu_{max} = 3{,}3$ Гц ($T_{рефр} = 300 - 350$ мс).

При бóльших частотах воздействия стимуляция тоже может происходить, но уже не каждый из подводимых импульсов будет вызывать возбуждение ткани: импульсы, попадающие в рефрактерный период, стимуляцию вызвать не смогут. Кроме того, с увеличением частоты уменьшается длительность импульса и соответственно возрастает амплитуда порогового тока. На частотах выше 10 кГц электростимуляция практически не наблюдается при безопасных амплитудах тока.

В медицинских аппаратах, предназначенных для электростимуляции нервов и скелетных мышц, используются обычно импульсные токи низкой частоты (20–150 Гц). Заметим, что в этот диапазон попадает и частота промышленного тока 50 Гц, что и определяет его повышенную опасность для человека.

16.2.4. Крутизна фронта импульса

Еще физиологами XIX в. было установлено, что величина порогового тока $I_{пор}$ зависит не только от длительности воздействия, но и от скорости нарастания (крутизны) фронта импульса, что отражено в ***законе Дюбуа — Реймона***, согласно которому раздражающее действие импульсного тока прямо пропорционально крутизне переднего фронта импульса. Это означает, что с увеличением крутизны фронта импульса (т.е. с уменьшением длительности фронта $t_{фр}$) пороговый ток уменьшается (при той же длительности импульса).

Поскольку наибольшую крутизну имеют прямоугольные импульсы, то им соответствует и наименьший пороговый ток. Если же фронт импульса очень пологий и его крутизна мала, то возбуждение ткани может не произойти вообще, даже при очень большой амплитуде стимула (явление аккомодации).

16.2.5. Длительность спада импульса

Длительность спада $t_{сп}$ импульса не влияет на возбуждение ткани, поскольку возбуждение уже произошло и дальнейшее изменение тока для возбужденной клетки безразлично.

Итак, для электростимуляции нервов и скелетных мышц используются импульсные токи низких частот, обычно менее 200 Гц. Форма импульсных токов может быть разнообразной: прямоугольной, экспоненциальной, однополярной, разных полярностей (рис. 16.4).

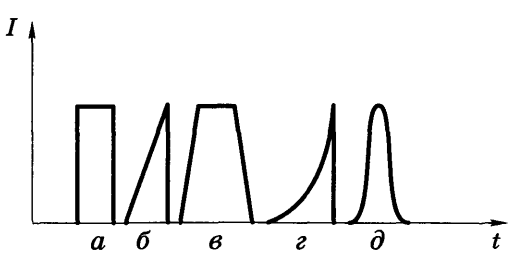


Рис. 16.4. Некоторые виды импульсных токов:
а — прямоугольные; *б* — пилообразные; *в* — трапецеидальные; *г* — экспоненциальные; *д* — колоколообразные

Часто применяют низкочастотные токи, «заполненные» высокочастотной (2–5 кГц и выше) составляющей. Так, для стимуляции мышц и нервных тканей используются синусоидальные модулированные токи (СМТ). В связи с тем что при лечебном воздействии происходят амплитудные пульсации такого тока, этот метод назвали ***амплипульстерапией***. Электротерапия синусоидальными модулированными токами осуществляется аппаратом «Амплипульс-4».

В этом аппарате используются несущие синусоидальные колебания частотой 5 кГц, которые модулируются низкочастотными синусоидальными колебаниями, частота которых ступенчато регулируется в пределах 30–150 Гц. Возможные формы токов, создаваемые этим аппаратом, показаны на рис. 16.5.

Высокочастотная составляющая — это переменный синусоидальный ток частотой около 5 кГц. Поскольку на таких частотах импеданс живой ткани в десятки раз ниже, чем при низких частотах электростимуляции, то высокочастотное заполнение стимулирующего импульса позволяет преодолевать высокое сопротивление кожи и подкожно-жировой клетчатки и оказывать стимуляцию (низкочастотной огибающей тока) внутренних возбудимых тканей при относительно небольших амплитудах тока и напряжения.

Для электростимуляции скелетных мышц с ненарушенной иннервацией обычно применяют СМТ с частотой модуляции

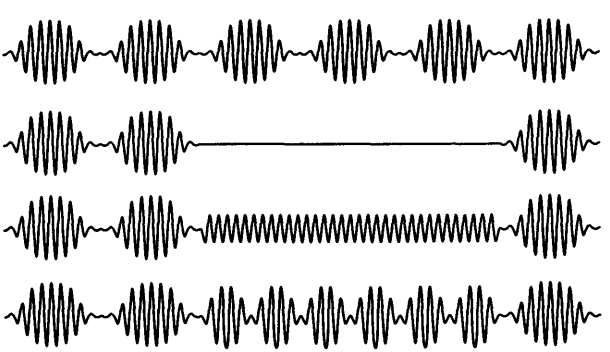


Рис. 16.5. Виды синусоидально-модулированных токов

100–150 Гц и несущей частотой 5 кГц, которые не вызывают раздражения кожи и обеспечивают хорошее тетаническое сокращение мышц, близкое к естественному. При повреждении двигательного нерва применяют импульсные токи большей длительности (до 60 мс) и меньшей частоты следования (до 12 Гц). Для таких состояний более приемлемы экспоненциально нарастающие импульсы тока (см. рис. 16.4, *г*).

В клинической практике применяются аппараты для стимуляции центральной нервной системы («Электросон», «Электронаркон», электростимуляция спинного мозга, зрительных центров головного мозга и др.), нервно-мышечной системы («Амплипульс», «Рефтон» и др.), а также электрокардиостимуляторы.

Таким образом, электростимуляция наиболее эффективна на низких частотах (0,1–200 Гц), где безопасные для организма значения силы тока не превышают нескольких десятков миллиампер. Прогрев живой ткани на этих частотах невозможен из-за малого значения безопасного действующего тока.

С увеличением частоты пороговые токи резко возрастают, вероятность электростимуляции снижается, а на частотах свыше 100 кГц она уже невозможна из-за инерционности ионных каналов.

16.3. Электростимуляция сердечной мышцы

К шоковой электроимпульсной терапии относят ***дефибрилляцию*** — воздействие на сердце мощным одиночным импульсом тока длительностью $t_и = 2 - 5$ мс, которое применяется либо при остановке сердца, либо при нерегулируемой аритмии, обычно на невскрытую грудную клетку. В этом случае один электрод накладывают под лопатку, другой — на грудь в области сердца (или оба на грудную клетку: один электрод с правой, другой — с левой стороны). Напряжение 5–7 кВ, сила тока порядка 1 А. При проведении дефибрилляции на обнаженном сердце рабочее напряжение составляет 1,5–2,5 кВ. Мощный кратковременный импульс тока (разряд конденсаторной батареи) вызывает одновремен-

ное сокращение всех мышечных волокон миокарда и в дальнейшем с большой вероятностью сердце начнет сокращаться самостоятельно.

Асинхронная кардиостимуляция применяется при поражениях водителя ритма сердца (пейсмекер) и проводящей системы сердца. В этом случае роль водителя ритма выполняет кардиостимулятор, который с постоянной частотой, равной примерно 1,2 Гц, вырабатывает электрические импульсы. Последние по системе проводов, вживленных в миокард, разводятся по сердечной мышце, вызывая ее периодическое сокращение. Недостатком этого вида стимуляции является постоянство частоты сердечных сокращений, не зависящей от испытываемой пациентом физической нагрузки.

Кардиосинхронизированная электростимуляция применяется в тех случаях, когда водитель ритма (пейсмекер) вырабатывает электрические импульсы, но малой амплитуды, или же плохо функционируют отдельные участки проводящей системы сердца (атриовентрикулярный узел, пучки Гиса). В этих случаях кардиостимулятор лишь усиливает (при необходимости) электрические сигналы пейсмекера и с помощью специальных электродов разводит его по сердечной мышце, обеспечивая ее нормальное сокращение. В этом случае частота сердечных сокращений задается пейсмекером, а не стимулятором, и может физиологически изменяться, обеспечивая пациенту необходимую частоту сердечных сокращений.

Кардиостимуляторы делают миниатюрными, имплантируемыми под кожу, ток к сердечной мышце подают с помощью вживленных в нее электродов. Источником питания являются специальные имплантируемые вместе с кардиостимулятором литиевые батарейки со сроком работы до 15 лет.

Контрольные вопросы и задания

1. Какие воздействия называют электростимуляцией?
2. Каковы требования к амплитуде токов, воздействующих на организм человека при проведении электростимуляции?
3. Каковы параметры прямоугольных импульсных токов?

4. Укажите основные параметры непрямоугольных импульсных токов.
5. Чем определяются частотные, амплитудные и временные параметры импульсных токов, используемых в электростимуляции?
6. Приведите и охарактеризуйте кривую электровозбудимости ткани.
7. Что такое реобаза и хронаксия? Каковы единицы их измерения? Что такое «полезное время» в электростимуляции?
8. Что описывает уравнение Вейса — Лапика? Как связаны константы этого уравнения с реобазой и хронаксией?
9. В чем состоит закон Дюбуа — Реймона?
10. Какие виды импульсных токов применяются для электростимуляции органов и тканей? Для чего используют «высокочастотное заполнение» импульсных токов?
11. В чем состоит метод дефибрилляции сердечной мышцы? Каковы параметры и особенности этого воздействия?
12. В чем суть электрокардиостимуляции? Каковы ее разновидности?

Глава 17 Использование электромагнитных полей высокой частоты в медицине

Под электромагнитными полями (ЭМП) высоких частот понимаются ЭМП с частотами от 0,1 МГц до 300 ГГц, которые уже не способны возбуждать живую ткань и позволяют воздействовать на нее большими токами и полями, вызывающими ее прогрев или оказывающими на нее специфические нетепловые воздействия. Под действием высокочастотных электромагнитных полей тепло выделяется непосредственно внутри биотканей (эндогенное), в отличие от наружного прогрева при их контакте с более нагретым телом. Нагревание живой ткани на 0,5–2,0 °С приводит к улучшению кровообращения в ней, уменьшению болевых ощущений, активизации процессов обмена и т.п.

Наряду с тепловым действием имеют место и нетепловые эффекты воздействия ЭМП на биологические объекты, которые в последнее время подразделяют на *энергетические* и *информационные*.

17.1. Тепловое действие ЭМП высокой частоты

Известно, что количество теплоты, выделяющееся в проводнике, по которому в течение времени t идет ток силой I, определяется ***законом Джоуля — Ленца***:

$$Q = I^2Rt, \qquad (17.1)$$

где R — сопротивление проводника.

В диапазоне частот 0,5–300 МГц прогрев биотканей определяется как токами проводимости (в проводящих структурах тканей), так и токами смещения (в диэлектрических структурах), причем с увеличением частоты ЭМП вклад токов проводимости уменьшается, а токов смещения — растет. Поскольку ткани организма обладают разной удельной электропроводностью, то в электромагнитном поле они будут нагреваться по-разному и методы избирательного прогрева разных тканей также различаются.

Эффективность любой тепловой процедуры определяется удельной мощностью тепловыделения $q = \frac{Q}{Vt}$, т.е. количеством теплоты, которое выделяется в единице объема живой ткани за единицу времени. Для получения ощутимого нагрева (более 0,1 °С) выделяемая в тканях тепловая мощность должна компенсировать или превышать мощность тепловых потерь живой ткани из-за излучения, испарения, теплопроводности, конвекции и составлять десятки и сотни ватт, а сила тока — единицы ампер.

17.1.1. Диатермия

Метод ***диатермии*** состоит в воздействии на ткань токами частотой 0,5–2,0 МГц и силой 1–2 А. Два электрода площадью S накладываются на тело пациента, и ток, возникающий между ними, проходит последовательно через кожу, подкожно-жировую клетчатку, мышцы и другие ткани, нагревая их.

Удельная тепловая мощность, выделяющаяся при диатермии в единице объема тканей за одну секунду, определяется выражением

$$q = \frac{Q}{Vt} = \frac{I^2 \rho l t}{SlSt} = \frac{I^2}{S^2}\rho = \rho j^2, \qquad (17.2)$$

где $j = \frac{I}{S}$ — плотность тока в ткани; ρ — ее удельное электрическое сопротивление.

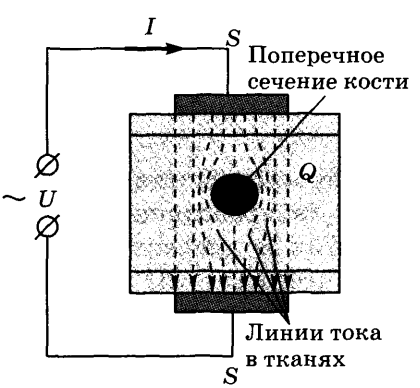


Рис. 17.1. Схема проведения диатермии

Из выражения (17.2) видно, что при одинаковой плотности тока при диатермии лучше прогреваются ткани с высоким удельным сопротивлением ρ, т.е. кожа и подкожно-жировая клетчатка. Кость, несмотря на высокое удельное сопротивление, обычно не испытывает нагрева при такой процедуре, так как линии тока обходят ее по окружающим проводящим тканям (рис. 17.1).

Одним из нежелательных эффектов при диатермии является возможность термических ожогов кожи при плохом ее контакте с электродом, так как в этом случае увеличивается как сопротивление контакта, так и плотность тока (из-за уменьшения площади контакта), а тепловыделение значительно повышается.

17.1.2. Электрохирургия

В современной медицине методы диатермии получили развитие и широкое применение в электрохирургии.

При моноактивной методике один из электродов имеет очень малый размер (активный электрод), а другой (индифферентный, пассивный электрод) — большую площадь соприкосновения с телом. Полный ток через оба электрода одинаков, но плотность тока под малым (активным) электродом в тысячи раз выше, чем под индифферентным (рис. 17.2). В результате под активным электродом ткани нагреваются на десятки и сотни градусов, тогда как под индифферентным электродом прогрев незначительный. Степень нагрева ткани под активным электродом регулируют как изменением силы тока, так и подбором формы и площади активного электрода. При нагреве живой ткани до 60–80 °С происходит электрокоагуляция — денатурация белков и сваривание ткани, а при нагревании до 150–200 °С — ее рассечение (электротомия), при этом вода в тканях вскипает, клетки разрушаются, их содержимое выгорает.

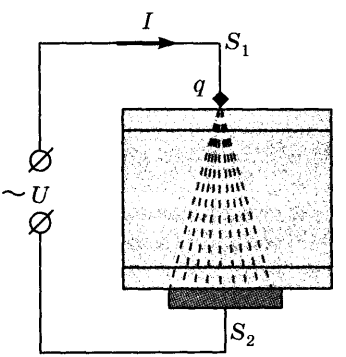


Рис. 17.2. Схема проведения электрохирургической диатермии

Биактивная электрохирургическая методика, при которой оба электрода имеют малую площадь и являются активными, используется для удаления папиллом, бородавок, небольших новообразований и т.п.

17.1.3. Индуктотермия

В этом физиотерапевтическом методе на биоткань воздействуют *высокочастотным магнитным полем* (рис. 17.3), индукция которого изменяется во времени с частотой $\nu = 10 - 20$ МГц:

$$B = B_0 \sin 2\pi\nu t. \qquad (17.3)$$

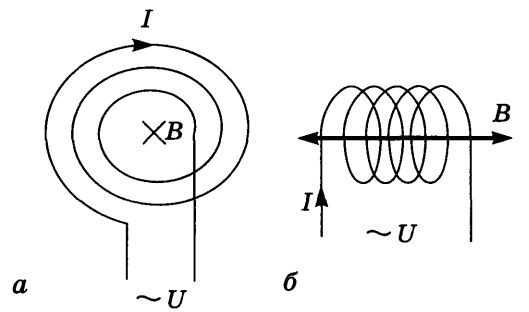


Рис. 17.3. Виды катушек для создания магнитных полей в тканях организма:
а — плоская катушка для наложения на поверхность тела;
б — соленоид для воздействия магнитным полем на конечности

По закону электромагнитной индукции это переменное магнитное поле наводит в биотканях переменную электрическую ЭДС, равную производной от магнитного потока:

$$\text{ЭДС} = \Phi' = SB' = 2\pi\nu SB_0 \cos 2\pi\nu t. \qquad (17.4)$$

Эта ЭДС вызывает образование в тканях *вихревых токов* (токов Фуко) плотностью $j = \dfrac{\text{ЭДС}}{\rho l} = \dfrac{E}{\rho}$, где $E = \dfrac{\text{ЭДС}}{l}$ — напряженность электрического поля, создаваемая в ткани этой ЭДС, а l — длина проводящего контура.

Токи Фуко в соответствии с законом Джоуля — Ленца (17.2) вызывают прогрев проводящих тканей, при этом выделяющаяся в единице объема тепловая мощность обратно пропорциональна удельному сопротивлению ткани:

$$q = \rho j^2 = \frac{E^2}{\rho} = \frac{2\pi^2\nu^2 S^2 B_0^2}{\rho l^2} \sim \frac{\nu^2 B_0^2}{\rho}. \qquad (17.5)$$

Таким образом, при индуктотермии лучше прогреваются ткани с малым удельным сопротивлением ρ, т.е. жидкие проводящие среды (кровь, лимфа, тканевая жидкость), и ткани, богатые сосудами (мышцы, печень, селезенка), слабее прогреваются ткани с высоким удельным сопротивлением.

17.1.4. УВЧ-терапия

УВЧ-терапия заключается в воздействии на ткани тела пациента *электрическим полем ультравысокой частоты* (30–60 МГц) обычно с целью их прогрева. Прогреваемый участок тела человека помещают между двумя образующими конденсатор плоскими электродами ($Э_1$ и $Э_2$), которые подключают к выводам терапевтического контура аппарата УВЧ-терапии. Между электродами создается переменное высокочастотное электрическое поле, которое и вызывает прогрев тканей.

Как уже отмечалось, ткани тела человека можно подразделить на хорошо проводящие электрический ток (проводники) и плохо проводящие (диэлектрики). В поле УВЧ они нагреваются по-разному.

В проводящих тканях прогрев осуществляется токами проводимости и удельная тепловая мощность, выделяющаяся при УВЧ-воздействии, определяется законом Джоуля — Ленца, который здесь удобно выразить через эффективную напряженность $E_{эф}$ электрического поля и удельное электрическое сопротивление ρ ткани:

$$q_{пр} = E_{эф}^2/\rho. \qquad (17.6)$$

Прогрев диэлектриков осуществляется в основном токами смещения, а удельная тепловая мощность, выделяющаяся в них при УВЧ-терапии, определяется формулой

$$q_{диэл} = 2\pi\nu\varepsilon\varepsilon_0 E_{эф}^2 \mathrm{tg}\delta, \qquad (17.7)$$

где ν — частота УВЧ-поля; ε — относительная диэлектрическая проницаемость среды; ε_0 — электрическая постоян-

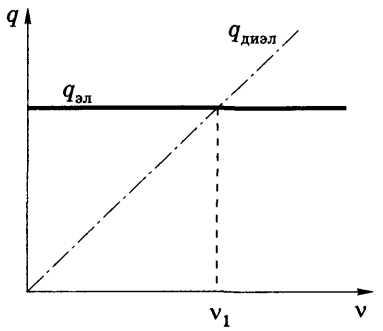


Рис. 17.4. Частотная зависимость тепловой мощности, выделяющейся под действием УВЧ-поля в электролите и диэлектрике

ная вакуума; tgδ — тангенс угла диэлектрических потерь (угол δ дополняет угол сдвига фаз φ между полным током и напряжением до 90°: $\delta + \varphi = \pi/2$).

Наиболее существенное отличие между выражениями (17.6) и (17.7) в том, что прогрев электролитов не зависит от частоты, а прогрев диэлектриков растет с увеличением частоты ЭМП (рис. 17.4), так что для двух конкретно взятых образцов электролита и диэлектрика найдется частота ν_1, на которой их прогрев одинаков. На частотах $\nu < \nu_1$ лучше прогревается электролит, а при $\nu > \nu_1$ — диэлектрик.

На частотах $\nu = 30–60$ МГц, используемых в аппаратах УВЧ-терапии, выделение тепла в диэлектрических тканях живого организма будет больше, чем в его проводящих тканях. Отметим, однако, что любой хороший проводник, в особенности металл, попавший в поле УВЧ, будет интенсивно нагреваться и, кроме того, изменять структуру самого поля (линии напряженности концентрируются около металлических проводников), что может привести к значительному перегреву отдельных участков тела пациента, находящихся возле металлического проводника, и образованию ожогов. Поэтому никаких металлических предметов на теле пациента при проведении УВЧ-терапии не должно быть!

Другая особенность УВЧ-воздействий в том, что пациент по соображениям техники безопасности помещается в так называемый терапевтический контур, который изолирован от опасных для жизни пациента низкочастотных напряжений питания технического контура аппарата УВЧ-терапии. Переменные электромагнитные колебания необходимой частоты ν_0 генерируются именно в техническом контуре

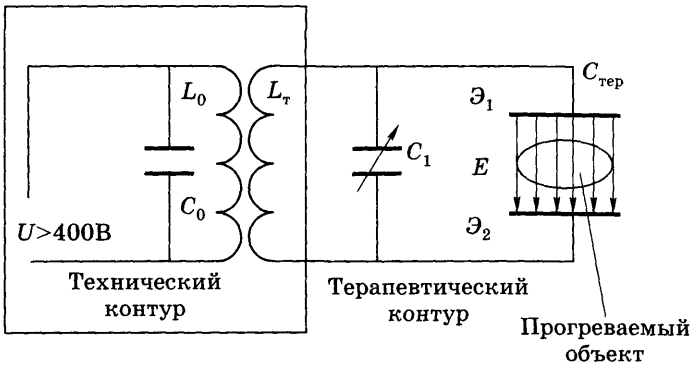


Рис. 17.5. Схема взаимосвязи технического и терапевтического контуров аппарата УВЧ-терапии

(рис. 17.5) и для эффективной передачи энергии ЭМП из технического контура в терапевтический необходимо, чтобы они были настроены на одну частоту, т.е. в резонанс: $\nu_{тер} = \nu_0$. В противном случае амплитуда электромагнитных колебаний в терапевтическом контуре будет слишком мала для получения теплового эффекта.

Собственная частота терапевтического контура определяется только его параметрами:

$$\nu_{тер} = \frac{1}{2\pi\sqrt{L_т(C_{тер} + C_1)}}. \qquad (17.8)$$

Индуктивность катушки $L_т$ фиксирована, но емкость $C_{тер}$ не постоянна, а зависит от расположения электродов $Э_1$ и $Э_2$ у тела пациента, поэтому настройка терапевтического контура в резонанс с техническим осуществляется с помощью дополнительного конденсатора переменной емкости C_1.

17.1.5. Микроволновая терапия

Физиотерапевтические методы, в которых для прогрева пациента используются ЭМП с частотами $\nu = 300–2500$ МГц, называют ***микроволновой (МКВ) терапией***. На этих частотах ЭМП уже не сосредоточено между электродами аппарата, а направленно излучается в пространство в виде электромаг-

нитной волны, длина которой $\lambda = \frac{c}{\nu}$, где $c = 3 \cdot 10^8$ м/с — скорость света в вакууме; в зависимости от используемой частоты λ составляет от 1 м до 12 см.

Электромагнитное излучение (ЭМИ) этих частот хорошо поглощается водой, поэтому при микроволновой терапии хорошо прогреваются водосодержащие ткани организма и хуже — жировая ткань, подкожная клетчатка, костная ткань.

Глубина проникновения электромагнитных волн в биоткани зависит от способности этих тканей поглощать энергию волны и определяется как структурой ткани (главным образом содержанием воды), так и частотой электромагнитных волн. Поглощение ЭМИ биотканями, как правило, растет с увеличением частоты (уменьшением длины волны) излучения, поэтому на практике используются аппараты, излучающие ЭМ-волны дециметровой длины (ДЦВ-терапия, $\nu = 460$ МГц, $\lambda = 65$ см) и сантиметровой длины (СМВ-терапия, $\nu = 2375$ МГц, $\lambda = 12{,}6$ см), причем последние поглощаются сильнее.

Чем сильнее поглощение ЭМИ в тканях, тем меньше глубина его проникновения (и прогрева), под которой обычно понимают расстояние, на котором интенсивность волны убывает ($e = 2{,}72$ раза). В однородных тканях интенсивность волны уменьшается с пройденным расстоянием по экспоненциальному закону. Однако состав живой ткани сложный, поэтому условно считают, что при ДЦВ-терапии средняя глубина проникновения ЭМИ составляет около 9 см, а при СМВ-терапии 3–5 см.

Поскольку при МКВ-терапии в основном прогреваются внутренние водосодержащие ткани и не прогреваются жировая клетчатка и кожа, то доза воздействия должна строго контролироваться во избежание перегрева внутренних органов, которое пациент сразу не сможет почувствовать, так как тепловые рецепторы находятся в коже.

17.2. Нетепловое действие ЭМП высоких частот

Нетепловые эффекты воздействия ЭМП на биологические объекты подразделяют на энергетические и информационные.

Энергетические нетепловые эффекты ЭПМ характеризуются тем, что энергия ответной реакции биологической системы меньше энергии, переданной ей ЭМП.

Напротив, для *информационных* эффектов характерно преобладание энергии ответной реакции организма над энергией ЭМП, вызвавшего эту реакцию.

К энергетическим нетепловым эффектам ЭМП относятся местная дарсонвализация, импульсная УВЧ-терапия, а к информационным — КВЧ-терапия.

Местная дарсонвализация — это метод воздействия на кожу и доступные слизистые оболочки высокочастотным электрическим разрядом ($\nu \sim 100$–500 кГц), который образуется между поверхностью тела пациента и специальным стеклянным полым электродом, заполненным разреженным воздухом или газом. Несмотря на высокое напряжение (10–30 кВ), сила тока невелика (10–15 мА в импульсе, скважность более 100) из-за высокого сопротивления самого стеклянного электрода, составляющего несколько мегаом.

Лечебный эффект при дарсонвализации обусловлен раздражением нервных рецепторов кожи и слизистых оболочек в зоне разряда, вызывающим местное улучшение кровоснабжения ткани и соответственно улучшение ее питания, трофики и др. Поскольку сила тока в разряде незначительна, то тепловыделения в тканях при дарсонвализации практически нет.

Применяется местная дарсонвализация для лечения различных кожных заболеваний, обморожений, ожогов, невралгий и др.

При ***импульсной УВЧ-терапии*** на пациента действует кратковременное (несколько миллисекунд) высокочастотное электромагнитное поле частотой 30–60 МГц, а затем следует в сотни и тысячи раз более длинная пауза, так что усредненный по времени тепловой эффект незначительный. В таких импульсных полях хорошо выражены нетепловые эффекты: ориентация полярных молекул вдоль направления электрической составляющей ЭМП, некоторые изменения структуры белковых молекул, а также явления, оказывающие влияние на функциональное состояние организма. Хотя пиковая мощность ЭМП в момент импульса почти в 40 раз больше

мощности, используемой в непрерывной УВЧ-терапии, средняя тепловая мощность, выделяющаяся в тканях, мала из-за очень длинной паузы (большой скважности), так что и тепловой эффект весьма мал.

Крайне высокочастотная терапия (КВЧ-терапия). Крайне высокими называют частоты ЭМП от $3 \cdot 10^{10}$ до $3 \cdot 10^{11}$ Гц, т.е. 30–300 ГГц, что соответствует длинам волн в воздухе от 10 до 1 мм. Такие электромагнитные волны оказывают воздействие на биологические объекты даже при очень низких информационных уровнях мощности применяемого излучения [30].

В этом диапазоне энергия кванта ЭМП значительно меньше энергии теплового движения биомолекул, меньше энергии электронных и колебательных переходов в молекулах, а также энергии водородных связей, поэтому КВЧ-излучение способно влиять только на вращательные и конформационные переходы в молекулах. Однако давно отмечено, что низкоинтенсивное КВЧ-излучение оказывает значительное влияние на функционирование живых организмов, хотя нагрев тканей не превышает 0,1 К. В настоящее время КВЧ-терапию применяют для лечения широкого круга заболеваний, в том числе язвенной болезни желудка и двенадцатиперстной кишки, псориаза, пневмонии и др. Считается, что при воздействии КВЧ-излучением повышаются защитные силы организма за счет активизации ряда биохимических процессов, формирующих физиологический ответ организма. Хотя механизм воздействия КВЧ-излучений на организм до конца не ясен, предполагают, что он связан с резонансным поглощением этого излучения молекулами воды, для которой наиболее сильный резонанс наблюдается на частоте 50 ГГц.

В последние годы активно исследуется еще более высокочастотный диапазон ЭМИ — терагерцовый (1 ТГц = 10^3 ГГц), которому соответствуют длины волн от 3 мм до 30 мкм. Энергия кванта такого поля уже соответствует энергии колебательных переходов многих биологически важных молекул и газов (N_2, O_2, CO_2, паров H_2O и др.) [3].

В терагерцовом диапазоне находятся резонансные частоты некоторых биологических структур живой клетки. На-

пример, соматическая клетка млекопитающих имеет резонансную частоту 2,39 ТГц, хромосомы различной генной активности — 0,75–15 ТГц, резонансные частоты альвеол легких — 0,3–0,5 ТГц, а эритроцитов крови — 0,5–1 ТГц, но наиболее сильно эти волны поглощает вода, причем ее поглощение растет с увеличением частоты.

Контрольные вопросы и задания

1. Электромагнитные воздействия каких частотных диапазонов в медицине относят к высокочастотным и используют для прогрева биологических тканей?
2. Каковы физические механизмы, ведущие к выделению тепла в проводящих и диэлектрических тканях организма при действии высокочастотных токов и полей?
3. Каковы параметры воздействия при диатермии? Приведите расчет удельной тепловой мощности, укажите наиболее прогреваемые ткани, источники опасности в диатермии.
4. Каковы принципы электрохирургии? Что такое электротомия и электрокоагуляция, моно- и биактивная методики?
5. Каковы параметры и способы подведения воздействия к пациенту при индуктотермии? Что такое вихревые токи Фуко? Какие ткани лучше прогреваются при индуктотермии?
6. Что такое УВЧ-терапия? Укажите параметры и способы подведения воздействия к пациенту. Какие ткани человека лучше прогреваются в поле УВЧ? Приведите соответствующие формулы.
7. Почему при проведении УВЧ-терапии металлические предметы не должны находиться вблизи области воздействия?
8. Каково назначение технического и терапевтического контуров в физиотерапевтической аппаратуре? Какое условие является необходимым для передачи электромагнитной энергии из технического контура в терапевтический?
9. Каковы отличия в параметрах и результатах воздействия на пациента УВЧ-терапии в непрерывном и импульсном режимах?
10. Каков действующий фактор на пациента при микроволновой СМВ- и ДЦВ-терапии? В каких тканях происходит преимущественное выделение тепла? От чего зависит глубина прогрева тканей?
11. Почему требования на дозирование воздействия при МКВ-терапии более жесткие, чем при УВЧ-терапии?

12. В чем состоят нетепловые эффекты высокочастотных электромагнитных воздействий? Какие воздействия относят к энергетическим, а какие — к информационным?
13. Каковы параметры и способ подведения воздействия к пациенту при местной дарсонвализации? Почему при высоком напряжении, действующем на пациента при местной дарсонвализации, сила тока мала?
14. Каков механизм лечебного воздействия? Каков тепловой эффект при местной дарсонвализации?
15. Что такое крайне высокочастотная терапия? Каковы параметры (частота и длина волны) этого воздействия?

Глава 18 Получение и регистрация биофизических данных

Достоверная регистрация физических параметров, проводимая при обследовании пациента, и их правильная интерпретация обеспечиваются четким пониманием врачом физических принципов работы и методики применения различных медицинских приборов, знанием их возможностей, достоинств и недостатков конструкции.

Общая схема получения и регистрации биофизической информации о биообъекте представлена на рис. 18.1, которая отражает все возможные реальные системы, применяемые в медицине для диагностики и обследования.

Измеряемый параметр биообъекта может представлять собой биопотенциал, температуру, давление крови, некие характеристики биоткани и др. Для регистрации электрических сигналов используются электроды, налагаемые на биообъект, а для регистрации неэлектрических параметров — соответствующие измерительные преобразователи (датчики). Поскольку амплитуда биопотенциалов очень мала, их необходимо предварительно усилить, перед тем как подавать на регистрирующий прибор. В ряде случаев полученную с объекта информацию необходимо передать на значительные расстояния по каналам связи (телефонные линии, радиосвязь и т.п.). В этом случае усиленный сигнал поступает сначала в блок дистанционной передачи данных, включающий кодирующее устройство, передатчик, линию связи, а затем в приемник и декодирующее устройство, после чего расшифрованный и при необходимости еще раз усиленный сигнал переходит на регистрирующий прибор и компьютерное устройство обработки и анализа данных.

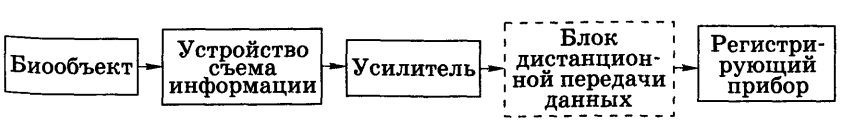


Рис. 18.1. Обобщенная схема регистрации медико-биологической информации

18.1. Регистрация биопотенциалов

Задача регистрации переменных электрических полей, генерируемых живым организмом (биопотенциалов), имеет большое значение для целей диагностики многих заболеваний и определения функционального состояния различных органов.

Для регистрации электрических потенциалов используются ***электроды*** — специальные проводники, закрепляемые на поверхности тела или вводимые внутрь его для регистрации электрических потенциалов (электрограмм) либо для подведения внешнего электрического воздействия к биообъекту (рис. 18.2).

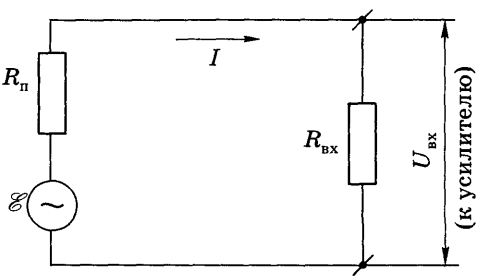


Рис. 18.2. Схема электрической цепи при регистрации биопотенциала

В медицине к электродам предъявляют специальные требования: они должны быстро фиксироваться и сниматься, обладать постоянством электрических параметров, эластичностью, прочностью, не создавать электрических помех, не оказывать раздражающего действия на биоткани.

Биопотенциал характеризуется электродвижущей силой (ЭДС), которая обозначена на схеме $\mathscr{E}$.

При регистрации биопотенциала сила тока в цепи, содержащей ЭДС, определяется законом Ома для полной цепи:

$$I = \mathscr{E} / (R_{вх} + R_п), \qquad (18.1)$$

а напряжение на входе усилителя

$$U_{вх} = IR_{вх}, \qquad (18.2)$$

где $R_{вх}$ — собственное входное сопротивление усилителя либо измерительного прибора; $R_{п}$ — так называемое переходное сопротивление, которое представляет собой сопротивление всех компонентов электрической цепи между источником биопотенциала и входными клеммами усилителя; $U_{вх}$ — измеряемое напряжение, регистрируемое на входе усилителя или регистрирующего прибора.

Подставляя формулу (18.1) в (18.2), получим

$$U_{вх} = \frac{\mathscr{E} R_{вх}}{R_{вх} + R_{п}} = \frac{\mathscr{E}}{1 + R_{п} / R_{вх}}. \qquad (18.3)$$

Из выражения (18.3) следует, что напряжение на входе усилителя (или измерительного прибора) всегда меньше величины измеряемого биопотенциала $\mathscr{E}$ и что это различие уменьшается с уменьшением отношения $R_{п}/R_{вх}$. Поэтому при регистрации биопотенциалов необходимо максимально снижать переходное сопротивление $R_{п}$ и использовать усилители (измерительные приборы) с высоким входным сопротивлением $R_{вх}$. Чтобы различие между $U_{вх}$ и $\mathscr{E}$ не превышало, например, 5 %, необходимо, чтобы входное сопротивление используемого усилителя в 20 раз превышало переходное сопротивление цепи. При более точных измерениях биопотенциалов следует еще больше уменьшать соотношение $R_{п}/R_{вх}$. На практике величина переходного сопротивления может составлять от десятков килоом до нескольких мегаом, поэтому очень важно знать способы его уменьшения. Рассмотрим с этой целью схему переходного сопротивления цепи регистрации ЭДС биопотенциала (рис. 18.3).

Из схемы видно, что основной вклад в переходное сопротивление вносит сопротивление кожи, составляющее единицы мегаом, по сравнению с которым сопротивление внутренних проводящих тканей (единицы кОм) и электродов (единицы Ом) невелико. Поэтому для снижения переходного сопротивления необходимо прежде всего уменьшить сопротивление кожи до десятков и единиц килоом.

Если между электродом и кожей поместить марлевый тампон, смоченный физиологическим раствором, то сопротивле-

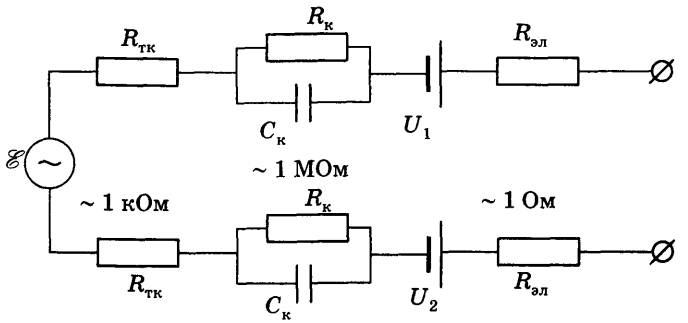


Рис. 18.3. Эквивалентная схема переходного сопротивления

ние кожи уменьшается до десятков килоом. Использование вместо тампона с физраствором специальных токопроводящих гелей и паст позволяет снизить сопротивление кожи до единиц килоом, что уже сравнимо с сопротивлением внутренних проводящих тканей, поэтому дальнейшее неинвазивное снижение переходного сопротивления невозможно.

При регистрации биопотенциалов следует учитывать также возникновение электродных потенциалов U_1 и U_2 при контакте электрода с кожей вследствие диффузии электронов из металла в электролиты кожи. При этом металл электрода всегда заряжается положительно, а электролиты кожи — отрицательно, поэтому электродные потенциалы U_1 и U_2 всегда включены навстречу друг другу (см. рис. 18.3). Полная ЭДС, возникающая в электродной регистрирующей цепи, равна алгебраической сумме ЭДС биопотенциала $\mathscr{E}$ и электродных потенциалов U_1 и U_2:

$$\text{ЭДС} = \mathscr{E} + U_1 - U_2. \qquad (18.4)$$

Величина электродных потенциалов зависит только от материала электродов и составляет десятые доли вольта, что в сотни раз превышает ЭДС биопотенциала (единицы и доли милливольта). Поэтому, во избежание грубых ошибок в измерении $\mathscr{E}$, материалы применяемых электродов должны быть одинаковы и одинаковым образом накладываться на кожу. Тогда U_1 и U_2 будут равны и скомпенсируют друг друга, а полная измеряемая ЭДС будет равна $\mathscr{E}$.

18.2. Регистрация неэлектрических параметров. Датчики

Многие важные физические параметры биообъекта имеют неэлектрическую природу (давление, температура, частота пульса и т.п.) и для их регистрации используют специальные устройства, называемые измерительными преобразователями или датчиками.

Датчик — это устройство, преобразующее измеряемую не электрическую величину в электрический сигнал, удобный для дальнейшего усиления, преобразования, передачи и регистрации. По своему назначению подразделяются на датчики *давления, температуры, частоты пульса, частоты дыхания* и др.

По принципу преобразования измеряемого параметра в электрический сигнал различают генераторные (активные) и параметрические (пассивные) датчики.

18.2.1. Генераторные датчики

Генераторные (активные) датчики под воздействием измеряемого параметра (температуры, давления и др.) генерируют электрическое напряжение или ток. В зависимости от вида физического явления, лежащего в основе такого преобразования, генераторные датчики бывают: пьезоэлектрические, индукционные, фотоэлектрические, термоэлектрические и др.

Пьезоэлектрические датчики преобразуют механическое давление P, оказываемое на них, непосредственно в электрическое напряжение U, прямо пропорциональное давлению $U \sim P$ (см. 7.2). Такие датчики широко используются и в технике, и в медицинских приборах для регистрации давлений, деформаций, вибраций, звуковых и ультразвуковых волн и других механических явлений. Долговечные, точные и надежные пьезодатчики давления, используемые в медицине, выпускаются многими известными фирмами.

В *индукционных генераторных датчиках* используется явление электромагнитной индукции (см. 15.3). Примером такого датчика может служить катушка с вставленным в нее ферромагнитным сердечником (рис. 18.4). При движении сердечника D в катушке возникает ЭДС индукции U, про-

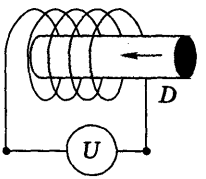


Рис. 18.4. Индукционный датчик скорости перемещения сердечника D

порциональная скорости изменения магнитного потока через катушку, и пропорциональная скорости движения сердечника.

Фотоэлектрические генераторные датчики широко используются для измерения световых потоков и основаны обычно на *фотогальваническом эффекте*, т.е. возникновении ЭДС на облучаемом светом p–n-переходе в полупроводниках. Под действием света в полупроводниках появляются дополнительные свободные электроны и дырки и на p–n-переходе генерируется ЭДС, пропорциональная освещенности полупроводника. К таким датчикам относятся фотодиоды, фототранзисторы, приборы с зарядовой связью (ПЗС-матрицы), широко используемые сейчас как двухмерные твердотельные датчики изображения. Сравнительно высокая скорость отклика и большая чувствительность в ближнем инфракрасном (ИК) и видимом диапазоне спектра обеспечивает этим датчикам широкую сферу применения.

Температурные датчики отличаются особенно большим разнообразием типов, являются одними из самых распространенных и находят широкое применение как в технике, так и в медицине.

К генераторным датчикам температуры относится *термопара*, действие которой основано на ***эффекте Зеебека***: в цепи, состоящей из двух спаянных концами разнородных металлов, возникает ЭДС, величина которой пропорциональна разности температур спаев.

Рассмотрим процессы, происходящие на контакте (спае) двух разнородных металлов A и B (рис. 18.5). Эти металлы

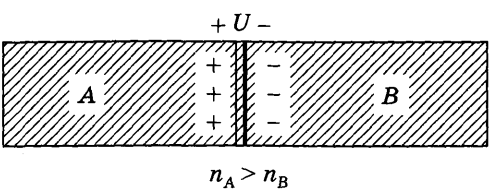


Рис. 18.5. Возникновение контактной разности потенциалов на спае двух разнородных металлов

отличаются концентрацией n в них свободных электронов. Предположим для определенности, что $n_A > n_B$. Вследствие теплового движения электроны будут переходить через контакт (спай) и проникать в соседний металл. Поскольку $n_A > n_B$, то поток электронов из A в B будет больше, чем в обратном направлении, и в области спая в металле B возникнет избыточный отрицательный заряд, а в металле A — такой же по величине положительный заряд. Это разделение зарядов между металлами A и B создает на контакте разность потенциалов U, которая будет препятствовать дальнейшей диффузии электронов из A в B и помогать диффузии электронов в обратном направлении.

Равновесие между этими двумя потоками электронов будет достигнуто при некоторой контактной разности потенциалов, определяемой уравнением Нернста (11.2), только здесь вместо отношения концентраций ионов по обе стороны клеточной мембраны следует брать отношение концентраций электронов по обе стороны контакта (спая):

$$U = \frac{RT}{F}\ln\frac{n_A}{n_B},$$

или

$$U = \frac{kT}{e}\ln\frac{n_A}{n_B}. \qquad (18.5)$$

Эти формулы равнозначны, так как газовая постоянная R и число Фарадея F связаны соответственно с постоянной Больцмана k и элементарным электрическим зарядом e через число Авогадро N_A: $R = kN_A$, $F = eN_A$.

Поскольку для данной пары металлов отношение n_A/n_B постоянно, то контактная разность потенциалов зависит только от температуры спая:

$$U = \alpha T, \qquad (18.6)$$

где $\alpha = \frac{k}{e}\ln\frac{n_A}{n_B}$ — величина, постоянная для данной пары металлов.

Если теперь спаять и вторые концы этих металлических проводников, то в образовавшейся замкнутой цепи возникнет ЭДС, пропорциональная разности (U_1 и U_2 включены

навстречу, в чем можно убедиться, сделав обход по замкнутому контуру) температур спаев, и поэтому называемая термоЭДС:

$$\text{ЭДС} = U_1 - U_2 = \alpha(T_1 - T_2). \qquad (18.7)$$

Для измерения температуры какого-либо объекта один из спаев помещают в этот объект, а второй — в термостат с известной температурой T_2, которая поддерживается постоянной. Измерив высокоомным вольтметром V термоЭДС и зная постоянную этой термопары α и температуру T_2, нетрудно вычислить исследуемую температуру T_1.

Термопары имеют почти линейную зависимость термоЭДС от разности температур спаев и широко используются для измерения температур в интервале от единиц до тысяч градусов Кельвина.

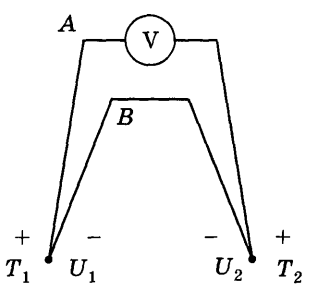


Рис. 18.6. Схема термопары и образование термоЭДС

Эффект Пельтье. Если в цепь (рис. 18.6), вместо вольтметра включить источник постоянного напряжения, то в ней возникнет постоянный ток, при этом один спай будет нагреваться, а другой — охлаждаться, поскольку на первом будет выделяться, а на втором — поглощаться теплота Q, пропорциональная силе тока I в цепи:

$$Q = \Pi I t. \qquad (18.8)$$

Коэффициент Пельтье Π зависит от природы применяемых металлов. Этот эффект был открыт в 1834 году французским физиком Ж. Пельтье и позже назван его именем. В настоящее время он используется в некоторых промышленных холодильных установках, а также в медицинских целях для регулируемого охлаждения или нагревания участка ткани, в который помещается спай двух разнородных проводников.

18.2.2. Параметрические датчики

Параметрические (пассивные) датчики под воздействием измеряемой величины не генерируют электрический сигнал, но изменяют свои электрические параметры — сопротивление, емкость или индуктивность, что и определяет их названия: резистивные (resistance — сопротивление), емкостные, индуктивные и др.

Пассивные датчики всегда необходимо включать в электрическую цепь с внешним источником питания, чтобы по изменению тока или напряжения на элементах этой цепи определить изменения электрических параметров датчика, а по ним — изменения измеряемого параметра.

Резистивные датчики под действием измеряемого параметра изменяют свое электрическое сопротивление R, определив которое можно найти значение измеряемого параметра. Для характеристики резистивных датчиков температуры широко используют параметр ТКС (температурный коэффициент сопротивления):

$$\text{ТКС} = \frac{1}{R_0}\frac{dR}{dT}\ (1/\ldots^\circ).$$

В случае линейной зависимости $R(T)$ эта величина постоянна, а при нелинейной зависимости ТКС зависит от температуры датчика. Чем больше величина ТКС, тем чувствительнее датчик. Положительное значение ТКС указывает на возрастание сопротивления датчика с ростом его температуры, а отрицательный ТКС — на обратную зависимость.

Рассмотрим в качестве примера датчик температуры на основе металлов.

Сопротивление R металлического проводника линейно зависит от его температуры:

$$R = R_0(1+at), \qquad (18.9)$$

где R_0 — сопротивление проводника при 0 °С; a — температурный коэффициент, зависящий только от природы металла; t — температура, °С.

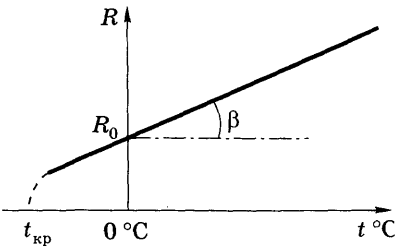


Рис. 18.7. Зависимость сопротивления металлического проводника от температуры

Зависимость $R(t)$ линейна в широком диапазоне температур. Указанная на графике (рис. 18.7) температура $t_{кр}$ соответствует его переходу в сверхпроводящее соотояние ($R = 0$), ее наличие и величина зависят от природы металла. Угол β наклона этой линии к оси температур определяет быстроту изменения сопротивления с температурой, зависит от свойств датчика и определяет его чувствительность: $\mathrm{tg}\beta = \frac{dR}{dt} aR_0$ (Ом/...°), которая зависит как от природы материала (a), так и от сопротивления датчика R_0. Для металлов ТКС = a и практически постоянен в широком диапазоне температур.

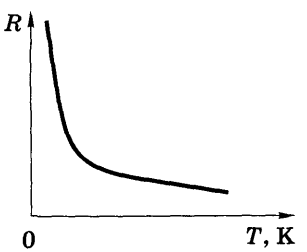


Рис. 18.8. Зависимость сопротивления термистора от его абсолютной температуры

Термисторы — полупроводниковые температурные пассивные датчики. Сопротивление R термисторов, изготовленных из чистых полупроводников (германий или кремний), уменьшается с увеличением их абсолютной температуры T (рис. 18.8) по закону

$$R(T) = Ae^{B/T}, \qquad (18.10)$$

где A (Ом) и B (град) — константы, зависящие от материала полупроводника, используемого в термисторе.

Из графика видно, что эта зависимость нелинейна: с ростом температуры падает и сопротивление R датчика и его ТКС:

$$\text{ТКС} = -\frac{B}{T^2}. \qquad (18.11)$$

Отрицательное значение ТКС указывает на снижение сопротивления R термистора с ростом его температуры, при этом уменьшается и величина ТКС, следовательно, уменьшается чувствительность датчика.

Термисторы, изготовленные из легированных полупроводников, имеют более сложные температурные зависимости $R(T)$.

В *емкостных датчиках* используется зависимость электрической емкости C плоского конденсатора от его геометрических размеров (площади пластин S и расстояния d между ними) и относительной диэлектрической проницаемости ε среды, заключенной между его обкладками (рис. 18.9):

$$C = \frac{\varepsilon\varepsilon_0 S}{d}.$$

При изменении этих параметров (S, d или ε) изменяется емкость C конденсатора и соответственно напряжение на нем: $U = q/C$, изменяется и его сопротивление переменному току [см. формулу (14.5)], что можно зарегистрировать по изменению тока в цепи регистрации и затем определить величину механического или другого фактора, повлиявшего на емкость конденсатора.

Индуктивные параметрические датчики перемещения в простейшем случае состоят из катушки, в которую частично вставлен ферромагнитный сердечник (см. рис. 18.4), являющийся чувствительным элементом датчика. Перемещение сердечника относительно катушки изменяет индуктивность L датчика и его сопротивление переменному току.

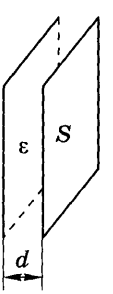


Рис. 18.9. Плоский конденсатор

Для увеличения точности измерений такие датчики часто включают в одно из плеч измерительной мостовой схемы.

Основные *характеристики* датчиков:

1) функция преобразования — зависимость выходной величины (U, R, C, L и др.) от измеряемой входной величины (F, P, T, x и др.), которая описывается некоторой функцией или представлена графиком;

2) динамический диапазон — диапазон измеряемых величин, которые датчик преобразует без существенных искажений;

3) чувствительность — минимальное изменение регистрируемого параметра, которое можно устойчиво обнаружить с помощью данного датчика;

4) время отклика — минимальный промежуток времени, в течение которого показания датчика устанавливаются на уровне, соответствующем значению измеряемой величины. Чем меньше время реакции датчика, тем выше его частотные характеристики.

18.3. Усиление электрических сигналов

Амплитуда биопотенциалов и электрических сигналов датчиков обычно чрезвычайно мала (не более нескольких мВ), поэтому они могут быть зарегистрированы только после их предварительного усиления. При этом очень важно увеличить амплитуду этих электрических сигналов и одновременно сохранить их форму неизменной, поскольку во многих случаях именно она несет диагностическую информацию.

Усилителем электрических сигналов называют устройство, увеличивающее интенсивность подаваемого сигнала за счет энергии постороннего источника. Поступление энергии от источника регулируется управляемым элементом, в качестве которого используются транзисторы, микросхемы и другие элементы.

Основной характеристикой усилителя является ***коэффициент усиления*** K, равный отношению амплитуды сигнала (мощности, напряжения, тока) на выходе усилителя к амплитуде соответствующего сигнала на его входе:

$$K=\frac{P_{\text{вых}}}{P_{\text{вх}}};\ K=\frac{U_{\text{вых}}}{U_{\text{вх}}};\ K=\frac{I_{\text{вых}}}{I_{\text{вх}}}. \qquad (18.12)$$

В зависимости от того, какой из параметров сигнала подлежит усилению, различают усилители мощности, напряжения и тока.

18.3.1. Усилители биоэлектрических сигналов

Основным требованием, предъявляемым к любому усилителю биоэлектрических сигналов, является увеличение амплитуды сигнала без искажения его формы.

Рассмотрим, какими конкретными физическими свойствами должен обладать такой усилитель. Подлежащие усилению биопотенциалы $\mathscr{E}(t)$, генерируемые живым организмом, носят обычно периодический характер и повторяются с некоторой частотой $\nu_0 = \frac{1}{T}$, где T — период повторения исследуемого физиологического процесса.

Согласно теореме Фурье, периодический процесс может быть представлен в виде суммы гармонических колебаний, частоты которых кратны основной частоте ν_0 (или циклической частоте $\omega_0 = 2\pi\nu_0$) изменения биопотенциала:

$$\mathscr{E}(t) = a_0 + a_1\cos(\omega_0 t + \varphi_1) + a_2\cos(2\omega_0 t + \varphi_2) + \ldots + a_n\cos(n\omega_0 t + \varphi_n) \quad (18.13)$$

или в краткой записи

$$\mathscr{E}(t) = a_0 + \sum_{m=1}^{n} a_m\cos(m\omega_0 t + \varphi_m), \quad (18.13а)$$

где a_0 — постоянная составляющая сигнала (во многих случаях она равна нулю); $a_m\cos(m\omega_0 t + \varphi_m)$ — его гармонические составляющие ($m = 1, 2, 3, \ldots, n$), имеющие амплитуду a_m, циклическую частоту $m\omega_0$ и начальную фазу φ_m.

Первая из этих составляющих (при $m = 1$) описывает гармоническое колебание, происходящее с основной частотой ν_0, равной частоте исследуемого физиологического процесса и называется *основным тоном сигнала* (*основной гармоникой*), а остальные составляющие (при $m = 2, 3, 4\ldots$) называются *обертонами*, или соответственно 2-й, 3-й, 4-й и так далее гармониками сигнала. Число n таких гармонических составляющих определяется видом (сложностью) функции $\mathscr{E}(t)$: чем сложнее форма сигнала, тем больше гармоник необходимо для его представления.

Область частот, занимаемая основным тоном и обертонами сигнала (т.е. область от ν_0 до $\nu_n = n\nu_0$), называется ***частотным***

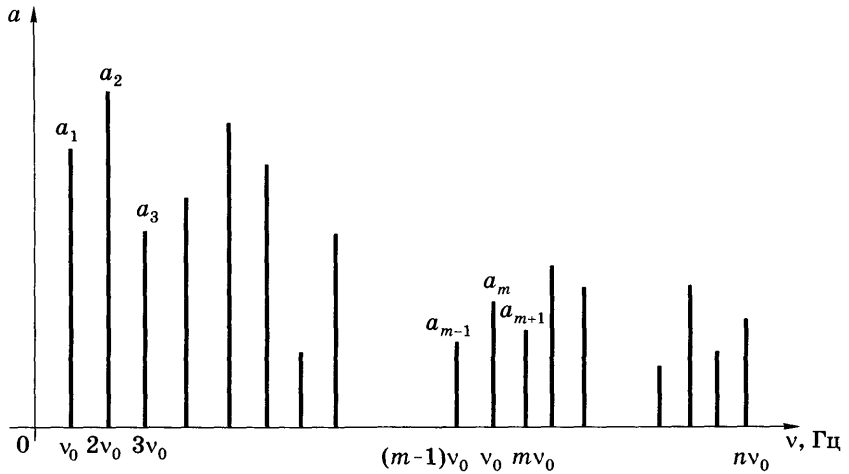


Рис. 18.10. Гармонический спектр сложного периодического сигнала

спектром сигнала, который в совокупности с амплитудами гармоник a_0, a_1, a_2, ..., a_n определяет ***гармонический спектр*** рассматриваемого сигнала (рис. 18.10). Отметим, что гармонический спектр сложного периодического сигнала всегда линейчатый, а непериодического — сплошной.

При прохождении сложного сигнала (18.13) через усилитель каждая из его гармонических составляющих увеличивается по амплитуде в некоторое число раз K_m, равное коэффициенту усиления на частоте этой гармоники и в общем случае различное для разных гармоник.

В процессе усиления могут изменяться и начальные фазы гармоник φ_m. На выходе усилителя происходит суммирование этих усиленных гармонических колебаний, в результате чего выходной сигнал будет представлять собой некую новую периодическую функцию $\mathscr{E}_\beta(t)$:

$$\mathscr{E}_\beta(t) = K_0 a_0 + \sum_{m=1}^{n} K_m a_m \cos(m\omega_0 + \Phi_m), \qquad (18.14)$$

которая может существенно отличаться по форме от входного сигнала $\mathscr{E}(t)$. Из сопоставления выражений (18.13а) и (18.14) видно, что форма периодического сигнала $\mathscr{E}_\beta(t)$ на выходе усилителя будет соответствовать форме сигнала $\mathscr{E}(t)$ на его входе

только в том случае, если коэффициенты усиления разных гармоник K_m будут одинаковы на всех частотах, имеющихся в спектре сигнала, т.е. если $K_0 = K_1 = K_2 = ... = K_n = \text{const} = K$, если начальные фазы гармоник при усилений не изменятся (или будут отличаться на одинаковую величину) $\Phi_m = \varphi_m$. Действительно, только при выполнении этих условий все коэффициенты K_m можно вынести из-под знака суммы:

$$\mathscr{E}_{\text{в}}(t) = K\left\{a_0 + \sum_{m=1}^{n} a_m \cos(m\omega_0 + \varphi_m)\right\} = K\mathscr{E}(t), \quad (18.15)$$

т.е. сигнал на выходе усилителя будет представлять собой усиленный в K раз входной сигнал $\mathscr{E}(t)$.

Таким образом, только в тех случаях, когда коэффициент усиления одинаков на всех частотах, имеющихся в спектре сигнала, и не зависит от амплитуды сигнала ($K = \text{const}$), а начальные фазы гармоник не изменяются в процессе усиления, на выходе усилителя получается сигнал, повторяющий по форме входной, но усиленный в K раз. Усилитель, удовлетворяющий этим условиям, можно считать идеальным.

Однако в действительности у любого реального усилителя коэффициент усиления всегда зависит от частоты и амплитуды входного сигнала, т.е. $K = K(\nu, a_m)$, что приводит соответственно к частотным и амплитудным искажениям усиливаемого сигнала.

Несохранение фазы усиливаемых гармонических составляющих сигнала ведет к фазовым искажениям сигнала.

По этим причинам реальный усилитель может существенно искажать форму усиливаемых сигналов, что недопустимо в медицине, так как именно форма этих сигналов (например, электрограмм) часто имеет определяющее диагностическое значение.

Поскольку идеальных усилителей нет, то надо уметь грамотно выбрать и применить реальный усилитель, чтобы он усиливал полезный сигнал, не искажая его форму. Для этого необходимо знать, как именно зависит коэффициент усиления реального усилителя от частоты и амплитуды входного сигнала. Это обстоятельство заставляет описывать свойства реальных усилителей при помощи характеристик, выражающих зависимость коэффициента усиления от частоты и от амплитуды сигналов.

18.3.2. Частотная характеристика усилителя

Зависимость коэффициента усиления от частоты называют ***частотной характеристикой усилителя.*** У идеального усилителя его частотная характеристика представляет собой прямую, параллельную оси абсцисс, и коэффициент усиления не зависит от частоты: $K(\nu) = \text{const}$. У реального усилителя коэффициент усиления лишь примерно постоянен в некотором диапазоне частот, а вне его быстро падает (рис. 18.11).

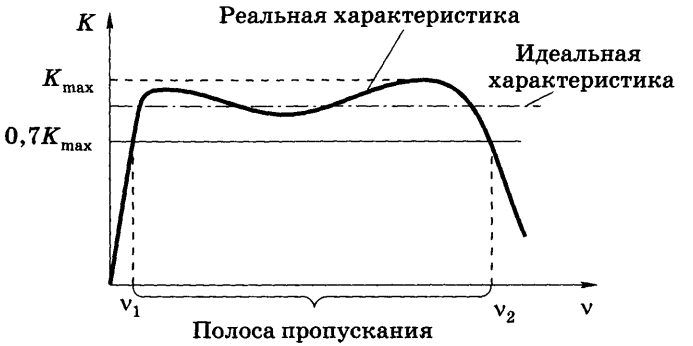


Рис. 18.11. Частотная характеристика усилителя

Это непостоянство коэффициента усиления ведет к частотным искажениям сигнала, однако установлено, что эти искажения еще малы и ими можно пренебречь, если изменения коэффициента усиления не превышают 30 % максимального значения.

В связи с этим область частот от ν_1 до ν_2, в пределах которой коэффициент усиления не ниже (т.е. больше или равен) $0{,}7K_{max}$, называют ***полосой пропускания усилителя.***

Если частотный спектр усиливаемого сигнала полностью попадает в полосу пропускания, то частотные искажения сигнала при усилении незначительны, не влияют на диагностическую ценность кривых и считаются допустимыми.

Если же спектр усиливаемого сигнала хотя бы частично выходит за пределы полосы пропускания усилителя, то частотные искажения будут значительны и такой усилитель не пригоден для усиления данного сигнала.

18.3.3. Амплитудная характеристика усилителя

Второй важнейшей характеристикой любого реального усилителя является его ***амплитудная характеристика,*** представляющая собой зависимость амплитуды сигнала (т.е. напряжения, тока или мощности) на выходе усилителя от амплитуды сигнала на его входе.

Для усилителя напряжения это будет зависимость $U_{вых} = f(U_{вх})$, для усилителя тока $I_{вых} = f(I_{вх})$, а для усилителя мощности $P_{вых} = f(P_{вх})$.

Для идеального усилителя эти зависимости всегда линейны, так как его коэффициент усиления K постоянен и не зависит от амплитуды входного сигнала: $U_{вых} = KU_{вх}$. Амплитудная характеристика идеального усилителя представляет собой прямую линию (прямая *а* на рис. 18.12), тангенс угла наклона которой к оси абсцисс равен коэффициенту усиления:

$$K = \frac{dU_{вых}}{dU_{вх}} = \mathrm{tg}\alpha .$$

Типичной амплитудной характеристике реального усилителя соответствует кривая *б* на рис. 18.12. Она линейна лишь в некотором диапазоне входных напряжений, а именно при

$$U_{вх1} \le U_{вх} \le U_{вх2}. \quad (18.16)$$

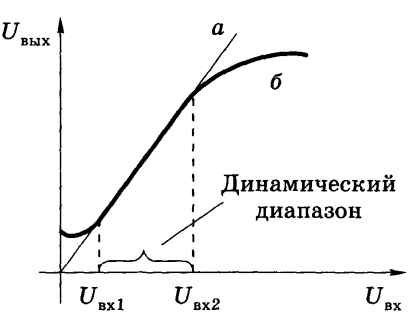


Рис. 18.12. Амплитудные характеристики усилителя

Только в пределах этого диапазона коэффициент усиления постоянный, амплитудные искажения формы сигнала отсутствуют и усилитель подобен идеальному. Поэтому диапазон входных напряжений от $U_{вх1}$ до $U_{вх2}$, в пределах которого амплитудная характеристика линейная, называется ***динамическим диапазоном усилителя*** и выражается в децибелах:

для усилителя мощности $D = 10\lg\frac{P_{вх2}}{P_{вх1}}$,

для усилителя напряжения $D = 10\lg\frac{RU^2_{вх2}}{RU^2_{вх1}} = 20\lg\frac{U_{вх2}}{U_{вх1}}$.

Вне динамического диапазона амплитудная характеристика становится нелинейной, что свидетельствует о непостоянстве коэффициента усиления на этих участках характеристики. Если амплитуды гармоник входного сигнала выходят за пределы этого диапазона, то будут наблюдаться значительные амплитудные искажения формы усиливаемого сигнала.

Таким образом, чтобы реальным усилителем произвести усиление исследуемого сигнала без искажений, необходимо выполнить два условия:

1) частотный спектр усиливаемого сигнала должен полностью попадать в полосу пропускания усилителя;

2) диапазон амплитуд гармоник усиливаемого сигнала должен находиться в пределах динамического диапазона усилителя.

18.3.4. Дифференциальный усилитель

Биопотенциалы, возникающие в организме человека, очень малы по величине, в десятки и сотни раз меньше посторонних электромагнитных сигналов (помех), самопроизвольно поступающих на вход усилителя. В результате этого на выходе обычного усилителя наблюдается очень большой беспорядочно изменяющийся сигнал помехи, из-за которого практически невозможно будет наблюдать слабый исследуемый сигнал. Поэтому для усиления биопотенциалов используются ***дифференциальные усилители***, способные усиливать полезный сигнал и подавлять сигнал помехи, т.е. обладающие высокой помехозащищенностью.

Особенностью дифференциального усилителя является наличие у него двух входов — инвертирующего ($вх_1$) и неинвертирующего ($вх_2$) и одного общего выхода (рис. 18.13). Коэффициенты усиления для этих входов одинаковы по величине, но противоположны по знаку:

$$U_{вых1} = -KU_{вх1}; \quad U_{вых2} = +KU_{вх2}. \qquad (18.17)$$

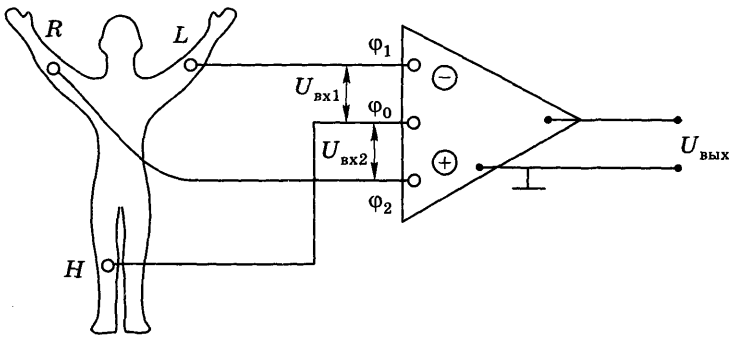


Рис. 18.13. Схема подключения дифференциального усилителя к пациенту при регистрации I отведения ЭКГ

На первый вход поступает напряжение между общей точкой усилителя φ_0 и клеммой φ_1, а на второй — между общей точкой φ_0 и клеммой φ_2.

Если на первый вход подать напряжение $U_{вх1}$, а на второй — напряжение $U_{вх2}$, то на выходе усилителя будет алгебраическая сумма усиленных сигналов, пропорциональная разности напряжений на его входах:

$$U_{вых} = U_{вых2} + U_{вых1} = K(U_{вх2} - U_{вх1}). \qquad (18.18)$$

Поэтому такой усилитель и называют дифференциальным, так как он усиливает не сами сигналы, а их разность (от англ. difference — разность).

Если к обоим входам дифференциального усилителя приложить напряжения, одинаковые и по амплитуде, и по фазе, то на выходе усилителя будет ноль, так как оба сигнала на выходе компенсируют друг друга. Именно это свойство дифференциального усилителя используется для устранения помех и оказывается чрезвычайно полезным при регистрации слабых сигналов на фоне больших помех.

Предположим, что необходимо усилить разность потенциалов между правой и левой рукой пациента (отведение I ЭКГ). На рис. 18.13 показано, что электроды L (левая рука) и R (правая рука) подключены к клеммам φ_1 и φ_2 усилителя соответственно. Третий электрод H, расположеный на правой ноге пациента, служит опорным и подключается к общей клемме усилителя φ_0.

В результате на первый вход усилителя подается напряжение $\varphi_L - \varphi_0$ между левой рукой и опорным электродом (правая нога), а на второй вход — напряжение $\varphi_R - \varphi_0$ между правой рукой и опорным электродом. Напряжение помехи $U_{пом}$ самостоятельно поступает одновременно на оба входа и может значительно превышать напряжение полезного сигнала. Таким образом, на входы усилителя поступают сигналы

$$U_{вх1} = \varphi_L - \varphi_0 + U_{пом}; \; U_{вх2} = \varphi_R - \varphi_0 + U_{пом}. \qquad (18.19)$$

На выходе усилителя получим усиленную в K раз разность входных сигналов

$$U_{вых} = K(U_{вх2} - U_{вх1}) = K(\varphi_R - \varphi_L), \qquad (18.20)$$

которая представляет собой усиленное отведение I ЭКГ, а сигналы помехи взаимно уничтожаются. Отметим, что потенциал φ_0 опорного электрода не влияет на конечный результат, поэтому данный электрод можно накладывать на любую точку тела пациента, но при регистрации ЭКГ его удобнее накладывать на правую ногу пациента, свободную от подключения стандартных отведений.

18.4. Регистрирующие устройства

После усиления полезный сигнал необходимо записать на каком-либо носителе информации или отобразить на шкале (экране) соответствующего прибора.

Регистрирующие приборы подразделяются на аналоговые, цифровые и комбинированные (рис. 18.14).

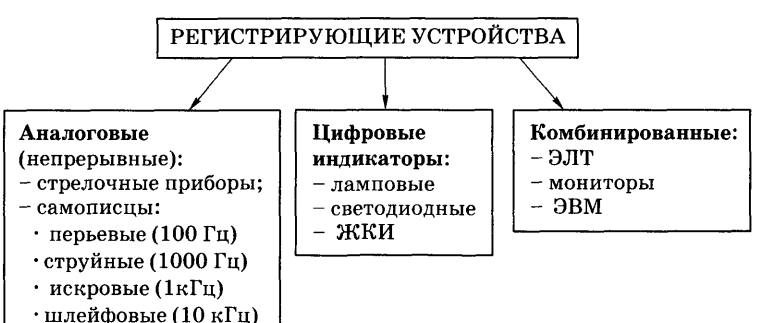


Рис. 18.14. Виды регистрирующих устройств

К ***аналоговым*** относятся приборы, непрерывно отслеживающие регистрируемый параметр в режиме реального времени: стрелочные вольтметры, амперметры, самописцы. Последние в зависимости от способа записи имеют разные частотные характеристики (полосу пропускания): перьевые самописцы из-за трения пера о бумагу имеют наименьшую полосу пропусания — 0,1–100 Гц; струйные и искровые — до 1000 Гц, шлейфовые осциллографы (запись сигнала производится световым лучом, отраженным от колеблющегося зеркальца гальванометра, на фотопленке или фотобумаге) — до 10 кГц.

Цифровые регистрирующие приборы — это разнообразные цифровые индикаторы: ламповые, светодиодные, жидкокристаллические и др. Они хорошо отображают медленно изменяющуюся информацию (температуру, давление, радиационный фон и др.), причем измерения этих параметров проводятся ими не постоянно, а дискретно, через определенные периоды (секунда, минута, 5 минут и т.п.), в зависимости от скорости изменения параметра. В течение этого интервала показания индикатора остаются неизменными до следующего акта измерения.

К ***комбинированным*** приборам относятся электронно-лучевые трубки (экраны телевизоров, осциллографов, мониторы) и высококачественные жидкокристаллические (ЖКИ) индикаторные панели, на которых может отображаться как аналоговая информация в виде непрерывных диагностических кривых, регистрируемых в реальном масштабе времени, так и дискретная информация в виде текста, цифровых значений параметров, как измеренных датчиками, так и вычисленных встроенными микропроцессорами по заранее введенному алгоритму. Комбинированные приборы имеют полосу пропускания в сотни килогерц и выше, а также широкий динамический диапазон. В настоящее время этот тип регистрирующих приборов наиболее быстро развивается и приходит на смену аналоговым. Их основное достоинство — возможность преобразования аналогового сигнала, обычно снимаемого с биообъекта, в оцифрованный сигнал, который можно обрабатывать математически по специальным алгоритмам, передавать по линиям связи и хранить на современных маг-

нитных и электронных носителях информации: жестких дисках ЭВМ, CD- и DVD-дисках, флеш-картах и т.п.

Отметим, что все приборы для регистрации биофизической информации, имеют свои частотные и амплитудные характеристики, и каждый из этих приборов не должен вносить искажения в регистрируемый сигнал.

Полоса пропускания всего тракта регистрации сигнала представляет собой частотный диапазон, общий для полос пропускания всех этих приборов; динамический диапазон тракта регистрации — это диапазон амплитуд входного сигнала, общий для динамических диапазонов всех приборов, задействованных в регистрации сигнала. Поэтому полосы пропускания и динамические диапазоны отдельных приборов в цепи регистрации должны быть согласованы как между собой, так и с частотными и амплитудными параметрами усиливаемых сигналов. Заметим, что в разных пособиях приводятся несколько различающиеся значения диапазонов частотных и амплитудных характеристик электрограмм. Поскольку с увеличением полосы пропускания и динамического диапазона тракта регистрации вероятность искажения регистрируемых сигналов существенно уменьшается, рекомендуется при выборе усиливающей и регистрирующей аппаратуры учитывать наиболее широкополосные версии частотных и амплитудных характеристик основных биоэлектрограмм, приведенные в табл. 18.1.

Таблица 18.1

Электрограмма	ЭКГ	ЭЭГ	ЭМГ
Частотный диапазон, Гц	0,5–400	1–1000	1–10 000
Диапазон амплитуд, мВ	0,1–5	0,01–0,5	0,1–50

Контрольные вопросы и задания

1. Охарактеризуйте обобщенную схему регистрации медико-биологической информации.
2. Приведите схему регистрации биопотенциалов. Покажите влияние переходного сопротивления на точность измерения потенциалов.

3. Каковы составляющие переходного сопротивления цепи и методы его уменьшения?
4. Каковы особенности регистрации неэлектрических параметров? Опишите классификацию датчиков и их основные характеристики. Приведите примеры генераторных и параметрических датчиков.
5. Что представляет собой гармонический спектр периодического сигнала? Чем он определяется? Приведите амплитудные и частотные параметры основных биоэлектрограмм: ЭКГ, ЭЭГ, ЭМГ.
6. Каковы основные требования к усилителям биопотенциалов? Какой усилитель считается идеальным?
7. Приведите частотную и амплитудную характеристики реальных усилителей. Как определяются полоса пропускания и динамический диапазон усилителя? При выполнении каких условий реальные усилители не искажают форму усиливаемых сигналов?
8. Каковы свойства дифференциального усилителя? Почему он усиливает слабые сигналы и не усиливает помехи?
9. Нарисуйте схему подключения дифференциального усилителя к пациенту при регистрации стандартного отведения III Эйнтховена.
10. Расскажите о видах регистрирующих устройств. Каковы особенности их применения и частотные параметры?
11. Как связаны полоса пропускания и динамический диапазон всего тракта регистрации с соответствующими характеристиками отдельных приборов, входящих в схему регистрации?

Глава 19 Общие свойства и поляризация электромагнитных волн

Электромагнитная волна представляет собой распространяющиеся в пространстве переменные электрическое и магнитное поля, изменяющиеся с одинаковой частотой ν (Гц). Они заполняют все мировое пространство, их источниками являются все тела в природе и они взаимодействуют со всеми телами, оказывая на них разнообразные воздействия. Электромагнитное излучение несет информацию как об источнике излучения, так и о среде, через которую она прошла, что широко используется для изучения структуры и строения веществ. Несмотря на свое многообразие, электромагнитные волны имеют много общих свойств.

19.1. Общие свойства и шкала электромагнитных волн

Если электромагнитная волна распространяется со скоростью υ вдоль направления OX, то напряженность E ее электрического поля и индукция B магнитного поля изменяются во времени t по одинаковому закону:

$$E = E_0 \sin\omega\left(t - \frac{x}{\upsilon}\right), \quad B = B_0 \sin\omega\left(t - \frac{x}{\upsilon}\right), \qquad (19.1)$$

где $\omega = 2\pi\nu$ — круговая частота волны (рад/с), $T = 1/\nu$ — период волны; E_0 и B_0 — максимальные (амплитудные) значения электрической напряженности и магнитной индукции волны.

Важной характеристикой является длина волны λ (м) — расстояние, проходимое волной за один период T волны:

$$\lambda = \upsilon T = \frac{\upsilon}{\nu}. \qquad (19.2)$$

Теория электромагнитных волн была развита Максвеллом; в ее основе лежат два экспериментально установленных факта:

1) переменное электрическое поле создает в окружающем пространстве переменное магнитное поле той же частоты;

2) переменное магнитное поле создает в окружающем пространстве вихревое переменное электрическое поле той же частоты.

Максвеллу удалось составить уравнения, описывающие взаимосвязь этих полей и процесс их распространения в среде. Было показано, что при распространении электромагнитной волны ее энергия периодически переходит из электрической формы в магнитную и наоборот. Плотности энергии, т.е. количество энергии в единице объема среды электрического и магнитного полей волны, определяются выражениями

$$w_{эл} = \frac{1}{2}\varepsilon\varepsilon_0 E^2, \quad w_м = \frac{B^2}{2\mu\mu_0}, \tag{19.3}$$

где E и B — мгновенные значения электрической напряженности и магнитной индукции волны.

Максимальные значения этих энергий одинаковы и равны полной плотности энергии электромагнитной волны, что позволяет установить связь между амплитудами электрического и магнитного полей

$$\sqrt{\varepsilon\varepsilon_0}E_0 = \frac{B_0}{\sqrt{\mu\mu_0}}. \tag{19.4}$$

Электромагнитные волны могут распространяться в вакууме, воздухе и любых диэлектрических средах, но в хорошо проводящих средах (металлах) они быстро затухают из-за сильного поглощения.

Максвеллом было показано, что скорость распространения электромагнитных волн определяется только электрическими и магнитными свойствами среды

$$v = \frac{1}{\sqrt{\varepsilon\varepsilon_0\mu\mu_0}},$$

где ε и μ — относительные диэлектрическая и магнитная проницаемости среды, $\varepsilon_0 = 8{,}85 \cdot 10^{-12}$ Ф/м и $\mu_0 = 1{,}43 \cdot 10^{-7}$ Гн/м — электрическая и магнитная постоянные соответственно.

В вакууме $\varepsilon = \mu = 1$ и скорость электромагнитной волны максимальна и равна $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}} = 2{,}98 \cdot 10^8$ м/с независимо от частоты волны. В любых других средах скорость υ электромагнитной волны меньше, чем в вакууме: $\upsilon = \frac{c}{n}$. Величина $n = \frac{c}{\upsilon} = \sqrt{\varepsilon\mu}$ называется ***абсолютным показателем преломления*** среды и показывает, во сколько раз скорость света в данной среде меньше, чем в вакууме. Для всех сред, кроме вакуума и воздуха, $n > 1$, $\upsilon < c$.

Все электромагнитные волны, независимо от частоты, имеют общую природу, однако методы и устройства генерирования и регистрации этих волн, области их технического применения, а также характер воздействия на биологические объекты, в том числе на человека, существенно зависят от частоты (длины волны) электромагнитной волны.

Всю шкалу электромагнитных волн принято делить на шесть крупных диапазонов (табл. 19.1), в пределах которых свойства этих волн, методы их получения, применения и регистрации, степень воздействия на человека примерно одинаковы.

Таблица 19.1

Электромагнитная волна	Длина волны
Радиоволны	$\lambda > 1$ мм
Инфракрасное излучение	760 Нм $< \lambda <$ 1 мм
Видимый свет	400 нм $< \lambda <$ 760 нм (760 нм — красная граница, 400 нм — фиолетовая)
Ультрафиолетовое излучение	80 нм $< \lambda <$ 400 нм
Рентгеновское излучение	10^{-5} нм $< \lambda <$ 80 нм
Гамма-излучение	$\lambda < 10^{-5}$ нм

Следует отметить, что приведенные границы диапазонов условны и в ряде случаев могут перекрывать друг друга. Напомним также, что электромагнитные волны обладают как волновыми, так и корпускулярными свойствами. Волновые

(длина волны λ и частота ν) и корпускулярные (энергия кванта ε и его импульс p) характеристики волны связаны соотношениями $\varepsilon = h\nu$ и $p = h/\lambda$, где $h = 6{,}6262 \cdot 10^{-34}$ Дж · с — постоянная Планка.

В очень небольшом по величине *видимом диапазоне* благодаря зрению человек получает более 90 % информации об окружающем его мире — этот диапазон электромагнитных волн наиболее информативен и важен для человека и других биологических организмов.

Электромагнитные волны *инфракрасного* и *радиодиапазонов* при интенсивностях, существующих в естественной природной среде, не опасны для человека, поскольку энергия электромагнитного кванта этих волн меньше энергии валентных связей молекул. У излучений этого диапазона больше выражены волновые свойства, чем корпускулярные.

Вредное воздействие на биосистемы начинается уже с *ультрафиолетового излучения*, а наиболее опасными для живых организмов являются *рентгеновское* и *гамма-излучения*, поскольку энергия электромагнитных квантов этих волн превышает энергию валентных связей молекул и поэтому способно разрывать эти связи, преобразуя молекулы в ионы и ион-радикалы. Излучение этих диапазонов является ионизирующим и его корпускулярные свойства выражены сильнее, чем волновые.

Несмотря на различия в действии на вещество и на живые организмы, электромагнитным волнам всех диапазонов присущи и общие свойства, одним из которых является поляризация волны.

19.2. Виды поляризации электромагнитных волн

В любой электромагнитной волне векторы напряженности электрического поля $\vec{E}$, индукции магнитного поля $\vec{B}$ и вектор скорости волны $\vec{v}$ образуют взаимно перпендикулярную правую тройку векторов: $\vec{E} \perp \vec{B} \perp \vec{v}$ (рис. 19.1).

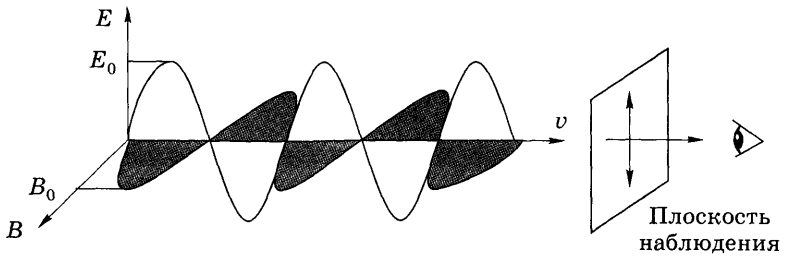


Рис. 19.1. Структура электромагнитной волны

При распространении волны векторы $\vec{B}$ и $\vec{E}$ всегда изменяются в пространстве одинаковым образом, описывая одинаковые по форме кривые, но во взаимно перпендикулярных плоскостях. Поэтому достаточно знать, как изменяется в пространстве один из этих векторов, поведение другого будет аналогичным.

Специальными опытами установлено, что действие электромагнитной волны на глаз, фотоприемники и многие другие объекты оказывает электрический вектор $\vec{E}$ волны, поэтому его называют также световым вектором волны.

Поляризация электромагнитной волны зависит от поведения ее электрического и магнитного полей в пространстве при распространении волны и определяется видом той кривой, которую описывает проекция электрического вектора $\vec{E}$ волны за один период колебания в плоскости наблюдения. Плоскость наблюдения — это плоскость, перпендикулярная направлению распространения волны (см. рис. 19.1), при этом свет должен распространяться к наблюдателю. Ясно, что проекция магнитного вектора $\vec{B}$ волны описывает в плоскости наблюдения такую же по форме кривую, что и вектор $\vec{E}$, но в перпендикулярном ему направлении. Рассмотрим основные виды поляризации.

Линейная (плоская) поляризация: в процессе распространения волны вектор $\vec{E}$ колеблется все время в одной и той же пространственной плоскости (см. рис. 19.1), которую называют плоскостью колебаний вектора $\vec{E}$ (иногда также

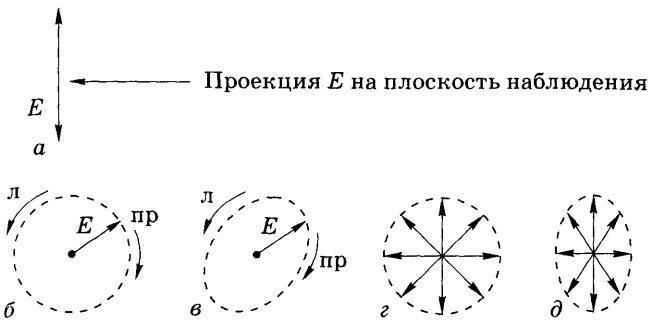


Рис. 19.2. Виды поляризации электромагнитных волн: *а* — линейная (плоская); *б* — круговая; *в* — эллиптическая; *г* — неполяризованная; *д* — частично поляризованная

плоскостью поляризации волны). Проекция вектора $\vec{E}$ на плоскость наблюдения в этом случае будет иметь вид отрезка прямой линии, в связи с чем эту поляризацию и называют линейной или плоской (рис. 19.2, *а*).

Отметим, что плоскость колебания магнитного вектора волны будет перпендикулярна плоскости колебаний вектора $\vec{E}$.

Круговая (циркулярная) поляризация: вектор $\vec{E}$ за один период волны делает полный оборот вокруг направления распространения волны, при этом его длина остается постоянной. В плоскости наблюдения проекция вектора $\vec{E}$ описывает круг, что и определяет название такой поляризации. Если вектор вращается по часовой стрелке (волна должна распространяться к наблюдателю), то поляризация называется правой круговой (пр), а если против часовой стрелки, то левой круговой (л) поляризацией (рис. 19.2, *б*).

Эллиптическая поляризация: вектор $\vec{E}$ за один период волны делает полный оборот вокруг направления распространения волны, но при этом его величина изменяется таким образом, что конец его описывает в плоскости наблюдения эллипс. Если для наблюдателя вектор $\vec{E}$ вращается по часовой стрелке, то поляризация называется правой (пр), а если против — левой (л) эллиптической (рис. 19.2, *в*).

Это виды полной поляризации света. Полностью поляризованный свет можно с помощью специальных поляризационных приборов (компенсаторов и фазовых пластинок) преобразовать без уменьшения интенсивности из одной формы в другую (например линейно поляризованный свет в циркулярно поляризованный, и наоборот).

Естественный свет — свет, испускаемый обычными источниками, т.е. солнцем, пламенем, электрической лампочкой; он является неполяризованным (рис. 19.2, *г*). Такой свет можно рассматривать как совокупность множества линейно поляризованных волн с одинаковыми амплитудами и всеми возможными ориентациями вектора $\vec{E}$ в плоскости наблюдения. По своим поляризационным свойствам такой свет существенно отличается и от линейно поляризованного, и от циркулярно поляризованного света.

Частично поляризованный свет можно рассматривать как совокупность линейно поляризованного и естественного света (рис.19.2, *д*) либо как совокупность множества линейно поляризованных волн с различными направлениями колебаний вектора $\vec{E}$, амплитуда которого неодинакова в различных направлениях. Поэтому концы этих векторов лежат не на окружности, как у естественного света, а на эллипсе. Форма такого эллипса может быть разной: от почти круговой (такой свет по своим свойствам близок к естественному) до сильно вытянутой (близок к линейно поляризованному). Поэтому для количественного описания частично поляризованного света вводится специальный параметр — *степень поляризации p*, который может принимать значения от нуля до единицы:

$$p = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \begin{cases} 0 \text{ — естественный свет (неполяризованный), } I_{\max} = I_{\min}, \\ 1 \text{ — линейно поляризованный свет, } I_{\min} = 0, \\ < 1 \text{ — частично поляризованный свет,} \end{cases} \quad (19.5)$$

где $I_{\max}$ и $I_{\min}$ — интенсивность линейно поляризованных волн вдоль большой и малой осей эллипса.

19.3. Методы получения поляризованного света

Многие важные свойства веществ можно обнаружить и исследовать только в поляризованном свете, однако обычные источники (солнце, пламя, электрические лампочки и другие) излучают свет неполяризованный. Для получения линейно поляризованного света из естественного созданы специальные приборы — поляризаторы, действие которых основано на определенных физических явлениях.

19.3.1. Поляризация света при отражении от диэлектрика

Хорошо известно, что при падении электромагнитной волны на границу раздела двух диэлектриков возникают отраженная и преломленная волны, подчиняющиеся законам отражения и преломления. Поляризация этих волн в общем случае будет различной. На графиках световой вектор $\vec{E}$ изображают точкой, если он ориентирован перпендикулярно плоскости падения, и чертой, если параллелен ей (рис. 19.3).

Если падающий свет неполяризован, то отраженная и преломленная волны будут поляризованы частично. Степень их поляризации будет зависеть от угла падения α: по мере его увеличения степень поляризации отраженного света растет и при определенном угле падения $\alpha_{Бр}$, называемом углом Брюстера (углом полной поляризации), отраженная волна будет

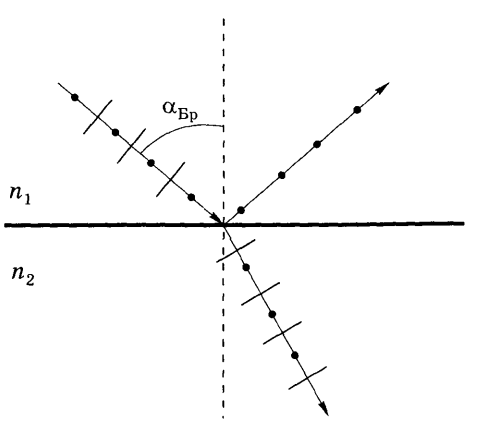


Рис. 19.3. Поляризация света при отражении (явление Брюстера)

полностью линейно поляризована так, что ее электрический вектор перпендикулярен плоскости падения (и параллелен плоскости раздела сред). Угол Брюстера находится из соотношения

$$\operatorname{tg}\alpha_{Бр} = \frac{n_2}{n_1}, \qquad (19.6)$$

где n_1 и n_2 — абсолютные показатели преломления первой и второй сред.

Преломленная волна будет всегда поляризована частично, причем максимальная составляющая ее электрического вектора параллельна плоскости падения. Чтобы при падении естественного света прошедшая волна стала полностью поляризованной, используют стопку плоскопараллельных пластинок. В результате многократного отражения волны одной линейной поляризации (перпендикулярной плоскости падения) полностью отражаются и во вторую среду проходят волны другой линейной поляризации (параллельной плоскости падения), что позволяет получить максимально возможный выход линейно поляризованного света (до 50 %) каждой из этих поляризаций (рис. 19.4).

Заметим, что если падающая волна линейно поляризована параллельно плоскости падения, то при падении под углом Брюстера она не отражается, а полностью проходит во

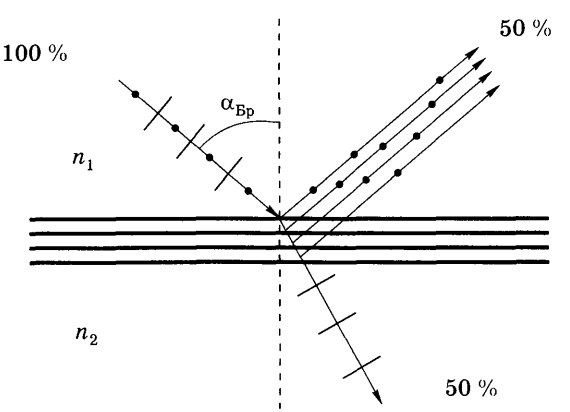


Рис. 19.4. Поляризация стопкой пластин

вторую среду, тогда как волна, поляризованная перпендикулярно плоскости падения, всегда частично отражается и частично проходит во вторую среду.

19.3.2. Некоторые оптические свойства анизотропных сред

Среды, свойства которых одинаковы по всем направлениям, называются ***изотропными***, а среды, свойства которых не одинаковы в различных направлениях, — ***анизотропными***.

Оптическая анизотропия характерна для кристаллов и проявляется в том, что в разных направлениях в кристалле скорость света различна и, кроме того, даже в одном и том же направлении она различна для волн разной поляризации.

Однако в кристаллах всегда есть одно или два направления, в которых скорость света одинакова для волн любой поляризации. Эти направления называют ***оптическими осями кристалла***. Если в кристалле только одно такое направление, то он называется *одноосным*, а если два — *двуосным*. Других видов анизотропных кристаллов в природе не существует. В данном учебнике рассмотрены свойства только одноосных кристаллов.

В одноосных кристаллах, независимо от поляризации падающего на них света, всегда распространяются только две линейно поляризованные волны, плоскости поляризации которых взаимно перпендикулярны.

Для одной из этих плоскополяризованных световых волн скорость распространения одинакова во всех направлениях (как и в изотропной среде); эта волна называется ***обыкновенной*** и характеризуется постоянной скоростью распространения υ_o и соответствующим показателем преломления $n_o = c/\upsilon_o$.

Для второй плоскополяризованной волны скорость распространения υ_e и показатель преломления $n_e = c/\upsilon_e$ различны в разных направлениях, т.е. такая волна называется ***необыкновенной***.

Таким образом, независимо от поляризации волны, падающей на кристалл, в нем будут распространяться одновременно две линейно поляризованные волны — обыкновенная и необыкновенная, у которых плоскости колебаний светового вектора $\vec{E}$ будут взаимно перпендикулярны.

У обыкновенной волны плоскость колебаний электрического вектора $\vec{E}$ всегда перпендикулярна главной плоскости кристалла, а у необыкновенной волны — параллельна ей. ***Главная плоскость кристалла*** — это плоскость, содержащая оптическую ось кристалла и световой луч, распространяющийся в нем.

Поскольку при падении на кристалл естественного света в кристалле всегда распространяются две линейно поляризованные волны — обыкновенная и необыкновенная, то для создания поляризатора необходимо, чтобы из кристалла вышла только одна из этих волн. Для этого используется явление двулучепреломления света или дихроизма поглощения в кристаллах.

19.3.3. Явление двулучепреломления света в кристаллах. Призма Николя

Если на кристалл под некоторым углом α падает естественный свет, то в кристаллах, как правило, возникают два преломленных луча, так как углы преломления обыкновенной и необыкновенной волн будут разными (рис. 19.5) из-за разных значений показателя преломления для этих волн:

$$\sin\beta_o = \frac{\sin\alpha}{n_o}, \quad \sin\beta_e = \frac{\sin\alpha}{n_e}.$$

Явление, называемое ***двулучепреломлением***, широко используется для создания прозрачных поляризационных призм, основная задача которых состоит в том, чтобы пропустить через кристалл только одну из линейно поляризованных

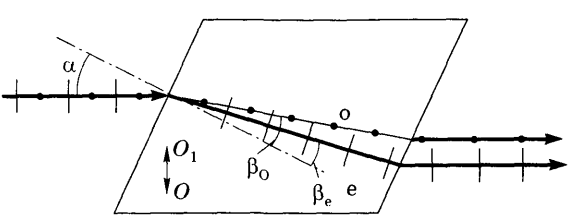


Рис. 19.5. Двулучепреломление света в кристалле: O_1O — направление оптической оси кристалла; o — обыкновенный луч; e — необыкновенный луч

волн и не пропустить другую. Одним из возможных вариантов является призма Николя.

Призма Николя представляет собой призму из исландского шпата, имеющего наибольшую известную разность показателей преломления обыкновенной и необыкновенной волн $n_o - n_e = 0{,}017$.

Призма, вырезанная определенным образом из этого кристалла, распиливается по меньшей диагонали, полируется и затем склеивается по месту распила специальным оптическим клеем — канадским бальзамом (рис. 19.6), показатель преломления которого занимает промежуточное значение между показателями преломления обыкновенной и необыкновенной волн $n_o > n_б > n_e$.

Неполяризованный свет, падающий на призму, распадается при преломлении в кристалле на два линейно поляризованных луча, которые падают на склейку под разными углами (см. рис. 19.6).

Обыкновенный луч падает из оптически более плотной среды на менее плотную ($n_o > n_б$). При таких условиях возможно полное внутреннее отражение света и геометрия призмы рассчитывается именно таким образом, чтобы обыкновенный луч полностью отразился от склейки ($\alpha \geq \alpha_{пред}$). Затем он поглощается черной краской, нанесенной на нижнюю грань призмы.

Необыкновенный луч падает из оптически менее плотной среды на более плотную ($n_e < n_б$) и практически полностью проходит через склейку. В результате через призму Николя проходит всегда только линейно поляризованный свет. Плоскость колебаний вектора $\vec{E}$ этого луча всегда параллельна плоскости, в которой лежит оптическая ось кристалла, поэ-

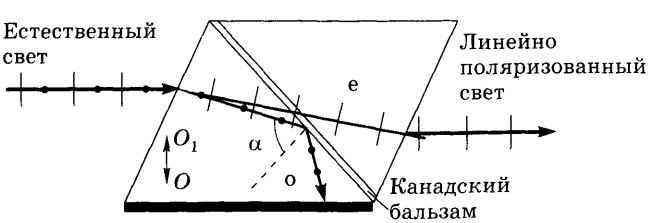


Рис. 19.6. Призма Николя: обыкновенный луч o испытывает полное отражение от склейки, а необыкновенный луч e проходит через нее и выходит из призмы

тому при вращении призмы Николя вокруг падающего луча, плоскость поляризации прошедшей волны тоже будет поворачиваться вместе с призмой.

Таким образом, призма Николя является ***поляризатором*** — устройством, независимо от поляризации падающего света пропускающим только линейно поляризованные волны, плоскость колебания светового вектора которых параллельна определенной плоскости призмы — *плоскости пропускания* поляризатора.

19.3.4. Явление дихроизма поглощения. Поляроиды

Известно, что в поглощающих средах интенсивность I света уменьшается с пройденным расстоянием x по закону Бугера:

$$I = I_0 e^{-kx}, \qquad (19.7)$$

где I_0 — интенсивность падающего света; k — показатель поглощения света в среде.

Для поглощающих анизотропных кристаллов характерно различие в показателях поглощения для обыкновенной и необыкновенной волн: $k_o \neq k_e$. Явление зависимости показателя поглощения света от поляризации волны называется ***дихроизмом поглощения.***

Благодаря этому явлению обыкновенная и необыкновенная волны в поглощающих кристаллах затухают по-разному, поэтому можно подобрать такую толщину кристалла x, что одна из этих волн практически полностью поглотится кристаллом, а другая волна выйдет из него (рис. 19.7).

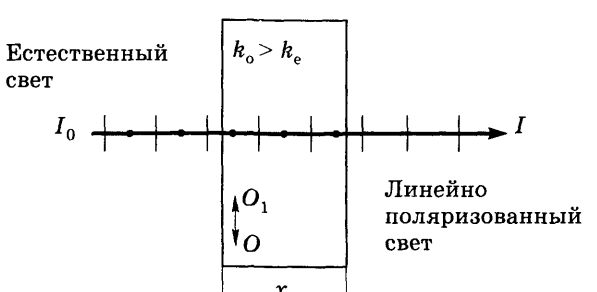


Рис. 19.7. Поляризация света вследствие дихроизма поглощения

Таким образом, независимо от поляризации падающей волны через такой кристалл пройдет только линейно поляризованная волна, плоскость колебаний которой параллельна плоскости пропускания этого поляризатора.

Из природных кристаллов наибольшим дихроизмом (т.е. большой разностью $|k_o - k_e|$) обладает *турмалин*. Пластинка из турмалина толщиной около 1 мм практически полностью поглощает обыкновенный луч, а необыкновенный луч, который поглощается значительно слабее, проходит через пластинку. Наибольшим известным дихроизмом обладает искусственно выращиваемый кристалл *герапатит*, который при толщине всего около 0,1 мм пропускает только линейно поляризованный свет. Эти кристаллы выращиваются на целлулоидной или стеклянной подложке и могут покрывать большие площади, образуя поляризаторы, которые называют ***поляроидами***.

19.4. Прохождение света через поляризатор. Закон Малюса

Если на поляризатор падает естественный свет, то после прохождения поляризатора он становится линейно поляризованным. При этом интенсивность волны уменьшится минимум в два раза:

$$I_{\text{прош}} = \frac{1}{2} I_0 . \qquad (19.8)$$

Из-за отражения света на передней и задней поверхностях поляризатора доля прошедшего света всегда меньше 50 %.

Рассмотрим теперь падение на поляризатор линейно поляризованной волны, причем плоскость колебания светового вектора $\vec{E}_0$ этой волны составляет с плоскостью пропускания AA_1 поляризатора угол φ (рис. 19.8). Поляризатор пропускает только линейно поляризованную волну, плоскость колебаний светового вектора которой параллельна плоскости

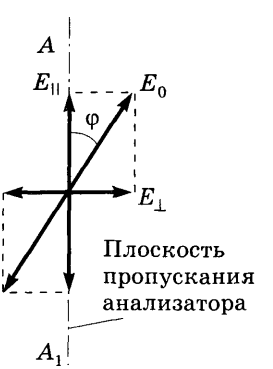


Рис. 19.8. К выводу закона Малюса

пропускания поляризатора. Амплитуда этой составляющей будет равна проекции $\vec{E}_0$ на плоскость пропускания AA_1 (см. рис. 19.8):

$$E_{\text{прош}} = E_0 \cos\varphi .$$

Так как интенсивность волны пропорциональна квадрату ее амплитуды, то после возведения в квадрат последнего соотношения получаем ***закон Малюса***:

$$I_{\text{прош}} = I_0 \cos^2\varphi. \qquad (19.9)$$

Если на поляризатор падает линейно поляризованный свет, то интенсивность прошедшего через него света пропорциональна квадрату косинуса угла между плоскостью пропускания поляризатора (AA_1) и плоскостью колебания вектора $\vec{E}$ падающей волны.

Если эти плоскости параллельны, то свет проходит через поляризатор полностью ($\varphi = 0$, $\cos 0 = 1$ и $I = I_0$), а если они перпендикулярны, то свет через поляризатор не проходит ($\varphi = 90°$, $\cos 90° = 0$, $I = 0$). При других значениях угла φ интенсивность прошедшего света принимает промежуточные значения, которые определяются формулой (19.9).

19.5. Оптическая активность

Оптическая активность — это явление поворота плоскости поляризации линейно поляризованного света при его прохождении через вещество.

Среды, способные поворачивать плоскость поляризации света, называются ***оптически активными.***

Оптическая активность веществ обусловлена тем, что волна, вошедшая в оптически активное вещество, распадается на две когерентные циркулярно поляризованные волны — с правым и левым направлением вращения светового вектора, скорости распространения которых различны. На выходе из вещества эти волны складываются, образуя линейно поляризованную волну, однако плоскость поляризации ее поворачивается на некоторый угол α относительно прежнего положения.

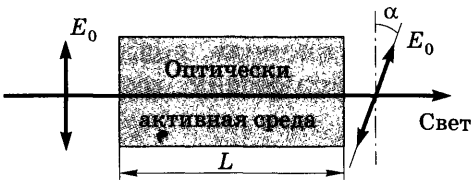


Рис. 19.9. Оптическая активность

Угол поворота α плоскости поляризации прошедшего света всегда прямо пропорционален длине пути L света в веществе (рис. 19.9) и определяется формулами:

- для твердых образцов $\alpha = [\alpha_0]L$, (19.10)

где величина $[\alpha_0]$ — постоянная вращения, зависящая только от природы вещества и длины волны света, град/мм;

- для растворов $\alpha = \alpha_0 CL$, (19.10а)

где α_0 — удельное вращение, ...°/мм · [ед. концентрации]; C — концентрация растворенного оптически активного вещества.

Соотношение (19.10а) является основой для определения концентрации оптически активного вещества путем измерения угла поворота плоскости поляризации света, прошедшего через раствор этого вещества.

У многих органических соединений, в том числе аминокислот, имеются зеркально симметричные пространственные изомеры, которые различаются по направлению вращения плоскости поляризации света, проходящего через них, и обозначаются как D-(правовращающие) и L-(левовращающие) изомеры. Установлено, что многие из этих изомеров по-разному усваиваются живыми организмами. Смесь D- и L-изомеров, не вращающая в итоге плоскость поляризации света, называется ***рацемической***.

19.6. Методы исследования в поляризованном свете

Поляризованный свет широко используется для измерения концентраций оптически активных веществ (поляриметрия), исследования структуры биологических тканей

в поляризованном свете (поляризационная микроскопия), в гигиенических исследованиях, при оптических исследованиях механических напряжений, возникающих в моделях костных тканей.

19.6.1. Оптическая поляриметрия

Оптическая поляриметрия — это совокупность физических методов исследования, основанных на определении угла поворота плоскости поляризации света, прошедшего через оптически активную среду.

К таким средам относятся растворы многих органических соединений — сахара́, кисло́ты, алкалоиды и др. Поляриметрию применяют не только для определения концентрации оптически активных сред, но и как метод исследования структурных превращений молекул, в частности, в молекулярной биофизике.

Измерение угла поворота плоскости поляризации света проводится с помощью специального прибора — *поляриметра*, действие которого основано на использовании закона Малюса. Простейший поляриметр состоит из двух поляризаторов П и А (рис. 19.10), между которыми помещается кювета K с раствором оптически активного вещества.

На поляризатор П падает монохроматический естественный свет, который после прохождения через поляризатор будет линейно поляризован. В отсутствие кюветы плоскость пропускания второго поляризатора А (его называют анализатором) устанавливают перпендикулярно плоскости пропускания первого, так что свет через них не проходит: $I_{\text{прош}} = I_0 \cos 90° = 0$.

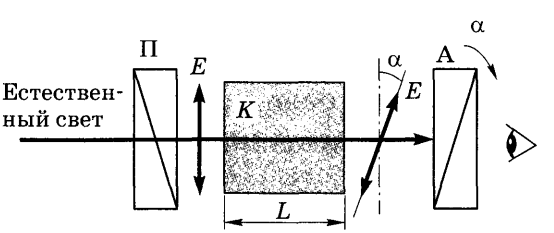


Рис. 19.10. Схема устройства поляриметра

При внесении между поляризаторами кюветы K с раствором оптически активного вещества плоскость поляризации света, падающего на анализатор, поворачивается на некоторый угол $\alpha = \alpha_0 CL$. Чтобы его измерить, анализатор поворачивают вокруг луча до тех пор, пока интенсивность прошедшего света опять не станет минимальной, и измеряют угол поворота анализатора (это и будет угол α) по шкале. Измерив угол α, зная длину кюветы L и удельное вращение α_0 растворенного оптически активного вещества, нетрудно найти его концентрацию C.

Поляриметрия — основной метод контроля концентрации сахара в сахарной промышленности, поэтому поляриметры часто называют ***сахариметрами.***

Поляриметрия находит широкое применение в санитарно-гигиенических, клинических и физиологических исследованиях. Методами поляриметрии определяют наличие и концентрацию углеводов в растительном сырье, белков и аминокислот в растворах, исследуют активность ферментов, расщепляющих углеводы, используют для анализа эфирных масел, алкалоидов, антибиотиков и других оптически активных соединений.

Одним из важных методов исследования строения вещества является ***спектрополяриметрия***, основанная на определении зависимости удельного вращения (или постоянной вращения) α_0 от длины волны падающего света (дисперсия оптического вращения). Эта зависимость для каждого вещества индивидуальна, поэтому спектрополяриметрия используется для проведения качественного анализа прозрачных растворов, а также для уточнения конформации растворенных органических соединений.

19.6.2. Оптический круговой дихроизм

Наряду с вращением плоскости поляризации оптически активные вещества по-разному поглощают световые волны, поляризованные по кругу влево и вправо, т.е. обладают *оптическим круговым дихроизмом*. Спектры кругового дихроизма изучают с помощью специальных приборов, называемых ***дихрографами.***

Методы измерения дисперсии (зависимости от длины волны) оптического вращения и кругового дихроизма широко применяют для изучения структуры многих биологически важных соединений, в частности вторичной и третичной структур молекул белков в растворах при изменении условий среды или при их взаимодействии с другими молекулами. Этими методами исследуются структуры ферментов, пептидных гормонов, мембранных белков, различных белковых комплексов (например, антиген — антитело), сложных комплексов (хроматина, рибосом, вирусов, фагов и др.), процессы нуклеиново-белкового узнавания и др.

19.6.3. Поляризационная микроскопия

Поляризованный свет широко применяется при исследовании оптически анизотропных биологических тканей. Во многих случаях при этом можно установить расположение и строение анизотропных элементов биоструктуры, которые не выявляются при микроскопировании в естественном свете. Оптическая анизотропия наблюдается у костных, мышечных и соединительных (коллагеновых) тканей, у нервных волокон. Само название скелетных мышц (поперечнополосатые) связано с тем, что при микроскопировании в естественном свете волокно кажется состоящим из чередующихся темных A и светлых I участков, что и придает ему поперечную исчерченность. Исследование этих волокон в поляризованном свете показывает, что более темные участки A являются анизотропными, тогда как более светлые I — изотропными, что и является причиной их различия в естественном свете.

Для гистологического исследования структур применяется ***поляризационный микроскоп***. Он аналогичен обычному биологическому микроскопу, но имеет поляризатор перед конденсором и анализатор в тубусе между объективом и окуляром. Предметный столик может вращаться вокруг оптической оси микроскопа. Таким образом, предмет освещается линейно поляризованным светом и рассматривается через анализатор. Если скрестить поляризатор и анализатор, то в отсутствие объекта поле зрения будет темным. При помеще-

нии на предметный столик препарата с изотропной структурой поле зрения останется темным. Если же препарат содержит участки с анизотропной структурой, то они будут выглядеть более светлыми на общем темном фоне из-за изменения поляризации света после прохождения анизотропной среды.

19.6.4. Поляризационно-оптический метод исследования механических напряжений (метод фотоупругости)

При создании в изотропном образце механических напряжений он становится анизотропным. Поэтому явление двойного лучепреломления можно наблюдать и в изотропных веществах, если их подвергнуть деформации. Это явление, получившее название *искусственного двойного лучепреломления*, лежит в основе оптического метода исследования распределения механических напряжений в деформированном твердом теле, который называется ***методом фотоупругости***. Он нашел широкое применение в медицине при изучении механических напряжений и деформации в различных костных тканях и ортопедических конструкциях.

Метод фотоупругости рассчитан на использование модели, выполненной из подходящего прозрачного изотропного материала. Обычно для этого применяют оргстекло, эпоксидную и полиэфирную смолы. Модель, геометрически аналогичная натуре, подвергается таким же условиям нагрузки, как и натура. При просвечивании поляризованным монохроматическим светом нагруженной модели каждый ее элемент ведет себя подобно кристаллу с двойным лучепреломлением. После прохождения возникающих при двойном лучепреломлении обыкновенных и необыкновенных лучей через анализатор возникают условия для интерференции этих когерентных волн, вследствие чего на изображении модели видна система темных и светлых интерференционных полос. Темные полосы получили название *изохром*. По виду изохром можно судить о распределении механических напряжений в модели, так как каждая изохрома соединяет точки, которым соответствует одинаковое механическое напряжение.

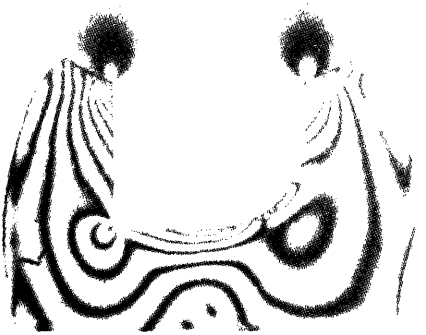


Рис. 19.11. Интерференционная картина распределения механических напряжений в деформированном образце

Исследование напряжений на моделях дает представление о распределении механических напряжений в реальных конструкциях, если модель в механическом отношении подобна образцу.

Таким образом, благодаря явлению фотоупругости можно получить наглядную картину распределения механических напряжений в сложных сооружениях и конструкциях, когда применение вычислительных методов математической теории упругости затруднено или невозможно.

Рассмотрим в качестве примера интерференционную картину, полученную при просвечивании поляризованным светом модели нагруженного зуба с различной геометрией формирования полости (рис. 19.11). Чем больше темных полос сходится в некотором месте образца, тем выше механическое напряжение в этой точке. Из рисунка видно, что в области округленного перехода со стенки полости на ее дно механические напряжения значительно меньше (одна, иногда две темные линии), чем в области прямоугольного перехода (три линии).

Контрольные вопросы и задания

1. Что представляет собой электромагнитная волна? Какова связь между амплитудами напряженности электрического и индукции магнитного полей волны? Чем определяется скорость рас-

пространения волны в среде? Что характеризует абсолютный показатель преломления среды?

2. Укажите основные диапазоны шкалы электромагнитных волн.
3. Какие виды поляризации электромагнитных волн вы знаете и чем они определяются? Что такое степень поляризации?
4. Как изменяется поляризация волны при отражении от диэлектрика? В чем заключается закон Брюстера?
5. В чем состоит явление двулучепреломления света в кристаллах? Каковы свойства обыкновенной и необыкновенной волн?
6. Опишите устройство призмы Николя и прохождение света через нее.
7. В чем состоит явление оптического дихроизма? Что такое поляроиды?
8. Чем различаются прохождения естественного и линейно поляризованного света через поляризатор? В чем заключается закон Малюса?
9. Что такое оптическая активность? Как устроен поляриметр и как с его помощью определить концентрацию раствора?
10. Какие методы использования поляризованного света в медико-биологических исследованиях вы знаете?

Глава 20 Тепловое излучение тел

Все без исключения тела в природе излучают электромагнитные волны, интенсивность и спектральный состав которых существенно зависят от температуры тела. Это излучение называют ***тепловым***. Чем выше температура тела, тем интенсивнее это излучение.

20.1. Основные характеристики теплового излучения

Поток излучения Φ (Вт) — это количество электромагнитной энергии, излучаемое всей поверхностью тела в единицу времени по всем направлениям:

$$\Phi = \frac{E}{t}. \qquad (20.1)$$

Энергетическая светимость R тела (Вт/м^2) — количество электромагнитной энергии, излучаемое с 1 м^2 поверхности тела за 1 с по всем направлениям и во всем спектральном диапазоне, другими словами R — это поток излучения с единицы поверхности тела:

$$R = \frac{\Phi}{S} = \frac{E}{tS}. \qquad (20.2)$$

Спектральная плотность энергетической светимости r_λ (Вт/м^3) — это количество энергии, излучаемое с 1 м^2 поверхности тела за 1 с по всем направлениям в единичном спектральном диапазоне на длинах волн от λ до $\lambda + d\lambda$:

$$r_\lambda = \frac{dR}{d\lambda}. \qquad (20.3)$$

Зависимость спектральной плотности r_λ от длины волны представляет собой *спектр теплового излучения* тела.

Энергетическая светимость R тела и спектральная плотность его энергетической светимости r_λ связаны соотношением

$$R = \int_0^{+\infty} r_\lambda d\lambda. \qquad (20.4)$$

Следовательно, энергетическая светимость любого тела численно равна площади под кривой r_λ, определяющей его спектр излучения.

Монохроматический коэффициент поглощения α_λ — представляет собой отношение потока излучения, поглощенного телом на длине волны λ, к потоку излучения с той же длиной волны, падающему на ту же поверхность тела:

$$\alpha_\lambda = \frac{\Phi_{\text{погл}}(\lambda)}{\Phi_{\text{пад}}(\lambda)}. \qquad (20.5)$$

Зависимость этого коэффициента от длины волны определяет *спектр поглощения тела.*

По типу зависимости коэффициента поглощения α_λ от длины волны все тела в природе можно разделить на три группы (рис. 20.1):

Абсолютно черные тела, коэффициент поглощения которых равен единице на всех длинах волн ($\alpha = 1$).

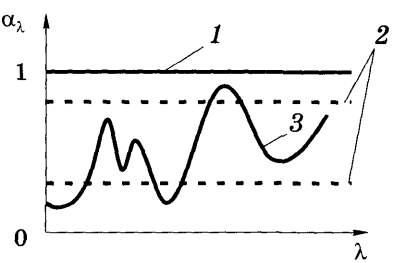


Рис. 20.1. Зависимость коэффициента поглощения разных тел от длины волны: *1* — абсолютно черное тело; *2* — серые тела; *3* — остальные тела

Такое тело полностью поглощает любое падающее на него излучение, поэтому и называется абсолютно черным.

В природе абсолютно черных тел нет, наиболее близкими свойствами обладают сажа, платиновая чернь, поверхность Солнца и др. Поэтому для экспериментального изучения свойств излучения абсолютно черного тела была создана специальная модель, представляющая полый шар из теплопроводящего вещества (обычно медь), в котором проделано небольшое боковое отверстие (рис. 20.2). Излучение попадает в полость через это отверстие и падает на стенку шара под некоторым углом. Часть излучения сразу поглощается, а оставшаяся часть отражается и падает на другой участок внутренней поверхности шара, частично

поглощаясь и частично отражаясь, и так до тех пор, пока все излучение не поглотится стенками шара. В результате отверстие, на которое падает излучение, и есть абсолютно черное тело, поскольку оно полностью поглощает любое падающее на него излучение, независимо от длины волны.

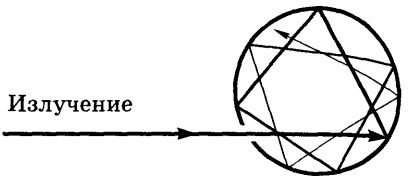


Рис. 20.2. Модель абсолютно черного тела

Вместе с тем при любой температуре внутри полости устанавливается равновесное тепловое электромагнитное излучение, идущее от стенок шара, часть его через отверстие излучается наружу, что позволило экспериментально исследовать спектр излучения абсолютно черного тела при разных температурах.

Серые тела, коэффициент поглощения которых меньше единицы, но одинаков на всех длинах волн ($\alpha = \text{const} < 1$).

Эти тела поглощают излучение не полностью, но одинаково на всех длинах волн. Иногда коэффициент поглощения α этих тел называют *коэффициентом серости*. Идеально серых тел в природе нет, но многие тела в определенных широких спектральных областях поглощают, как серые тела. В частности кожу человека в инфракрасном диапазоне можно считать серым телом с $\alpha = 0{,}9$.

Тела, коэффициент поглощения которых зависит от длины волны ($\alpha_\lambda = \alpha(\lambda) < 1$).

20.2. Законы теплового излучения

Основные законы теплового излучения установили экспериментально и затем обосновали теоретически в XIX в. Г. Кирхгоф, Л. Больцман, Й. Стефан, В. Вин, М. Планк и другие ученые.

Закон Кирхгофа, установленный в 1859 г., показывает, что в условиях термодинамического равновесия для всех тел существует определенная количественная связь между их спектральной плотностью энергетической светимости и монохроматическим коэффициентом поглощения: при одинако-

вой температуре отношение спектральной плотности энергетической светимости к монохроматическому коэффициенту поглощения одинаково для всех тел и равно спектральной плотности энергетической светимости ε_λ абсолютно черного тела:

$$\left[\frac{r_\lambda}{\alpha_\lambda}\right]_1 = \left[\frac{r_\lambda}{\alpha_\lambda}\right]_2 = \ldots = \left[\frac{\varepsilon_\lambda}{1}\right] = \varepsilon_\lambda . \qquad (20.6)$$

Отсюда $r_\lambda = \alpha_\lambda \varepsilon_\lambda$ и вытекают важные *следствия*:

- чем больше тело поглощает в некотором спектральном диапазоне, тем больше оно и излучает в этом диапазоне, так как $r_\lambda \sim \alpha_\lambda$;
- при одинаковой температуре спектральная плотность энергетической светимости любого тела меньше, чем у абсолютно черного тела: $r_\lambda < \varepsilon_\lambda$, так как $\alpha_\lambda < 1$. Другими словами, при любой фиксированной температуре абсолютно черное тело обладает наибольшей излучающей способностью по сравнению с другими телами.

Закон Кирхгофа позволяет определить спектр теплового излучения r_λ любого тела, если известна зависимость его монохроматического коэффициента поглощения тела α_λ от длины волны, поскольку спектр излучения абсолютного черного тела ε_λ хорошо известен.

Закон Стефана — Больцмана: энергетическая светимость абсолютно черного тела прямо пропорциональна 4-й степени его абсолютной температуры:

$$R = \sigma T^4. \qquad (20.7)$$

Коэффициент пропорциональности $\sigma = 5{,}67 \cdot 10^{-8}$ Вт/(м$^2 \cdot$ К^4) и называется *постоянной Стефана — Больцмана*.

С небольшой поправкой этот закон справедлив и для серых тел:

$$R = \alpha\sigma T^4, \qquad (20.7а)$$

где α — коэффициент поглощения серого тела.

Закон смещения Вина определяет зависимость положения максимума спектра излучения абсолютно черного тела

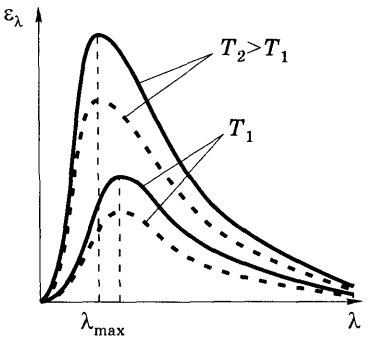


Рис. 20.3. Спектры теплового излучения абсолютно черного тела (сплошная линия) и серого тела (пунктир) при двух разных температурах, $T_2 > T_1$

(рис. 20.3) от температуры: длина волны $\lambda_{макс}$, на которую приходится максимум спектральной плотности энергетической светимости абсолютно черного тела, обратно пропорциональна абсолютной температуре этого тела:

$$\lambda_{max} = \frac{b}{T}, \qquad (20.8)$$

где $b = 2{,}898 \cdot 10^{-3}$ м · К — *постоянная Вина.*

Этот закон справедлив и для серых тел без каких-либо изменений и показывает, что с увеличением температуры максимум спектра теплового излучения смещается в сторону более коротких длин волн.

Из закона Вина следует, что при привычной нам температуре 300 К (27 °С) максимум излучения черных и серых тел соответствует излучению с $\lambda = 9{,}7$ мкм, т.е. средней части инфракрасного диапазона; при температуре 2000 К (нить лампы накаливания) максимум излучения приходится на ближнюю инфракрасную область $\lambda = 1{,}45$ мкм, а максимум теплового излучения Солнца (температура его поверхности примерно 6000 К) приходится на область видимого спектра $\lambda = 480$ нм.

Закон Вина используется также для дистанционного бесконтактного определения температуры поверхности удаленных тел (Солнца, звезд и т.п.) и раскаленных тел путем регистрации их спектра излучения, определения λ_{max} и вычисления температуры тела на основании соотношения (20.8). Температура тела, определенная таким способом, называется *цветовой,* так как определяется по цвету максимума свечения.

Другой способ бесконтактного определения температуры тел основан на выравнивании яркости свечения исследуемого

объекта с регулируемой яркостью нити накаливания специального прибора (пирометра), температура нити которого отображается на шкале прибора. Найденная таким способом температура тела называется *яркостной* или *радиационной*.

Знание спектра излучения абсолютно черного тела при разных температурах является ключевым звеном для определения спектров излучения всех остальных тел по их спектрам поглощения на основании закона Кирхгофа.

Спектр теплового излучения абсолютного черного тела ε_λ был хорошо изучен экспериментально при разных температурах еще в 80-х гг. XIX в. (см. рис. 20.3), но аналитическое выражение удалось получить лишь М. Планку в 1900 г.:

$$\varepsilon_\lambda = \frac{2\pi hc^2}{\lambda^5} \frac{1}{\left[\exp(hc/\lambda kT)\right] - 1}. \qquad (20.9)$$

При выводе этой формулы ему пришлось предположить, что свет излучается и поглощается определенными малыми порциями энергии ($h\nu$), пропорциональными частоте излучения, которые были названы *квантами излучения* или *фотонами*. Это предположение оказалось правильным, позволило Эйнштейну объяснить явление внешнего фотоэффекта и явилось мощным стимулом для дальнейшего развития квантовой механики.

Пользуясь ***формулой Планка*** (20.9), можно теоретически вывести законы Вина и Стефана — Больцмана. Для получения закона Вина формулу Планка (20.9) надо исследовать на экстремум по λ и найти выражение для λ_{max}, а для получения закона Стефана — Больцмана нужно проинтегрировать формулу (20.9) по λ от 0 до ∞ и получить зависимость R от T. В полученные аналитические выражения для λ_{max} и R войдет постоянная Планка h. Таким образом, зная численное значение констант Вина и Стефана — Больцмана, нетрудно вычислить значение постоянной Планка, определяющей наименьший возможный квант действия:

$$h = 6{,}63 \cdot 10^{-34}\ \text{Дж} \cdot \text{с}.$$

Несмотря на малое численное значение, постоянная Планка играет исключительно важную роль в квантовой механике

частиц, атомов и молекул, в теории излучения и поглощения тел, являясь при этом одной из фундаментальных физических констант.

20.3. Излучение Солнца

Наиболее мощным источником теплового излучения, определяющим жизнь на Земле, является Солнце. Как источник теплового излучения Солнце можно считать абсолютно черным телом с температурой поверхности около 6000 К. Поток солнечного излучения, приходящийся на 1 м^2 площади поверхности на границе земной атмосферы, называется ***солнечной постоянной*** и составляет 1350 Вт/м^2.

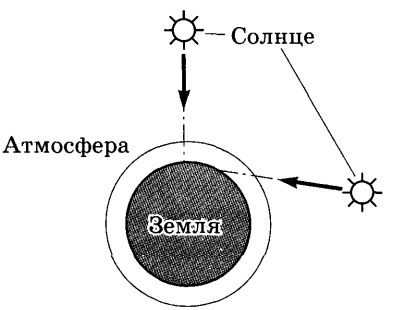


Рис. 20.4. Различие в пути световых лучей в атмосфере для Солнца в зените и на восходе (закате)

Путь, проходимый солнечными лучами в атмосфере, существенно зависит от высоты Солнца над горизонтом и по этой причине может различаться почти в 30 раз (рис. 20.4). Из-за рассеяния и частичного поглощения солнечного излучения в атмосфере плотность потока солнечной радиации уменьшается и на поверхности Земли даже при самых благоприятных условиях составляет не более 1120 Вт/м^2. С удалением от экватора эта цифра уменьшается, например на широте Минска в июне она составляет лишь 930 Вт/м^2.

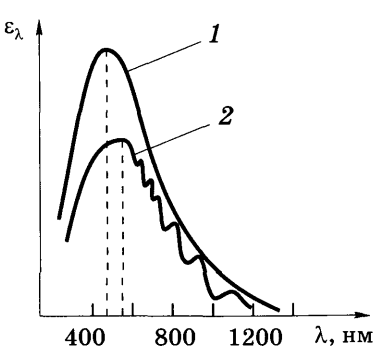


Рис. 20.5. Спектр излучения Солнца:
1 — на границе с атмосферой;
2 — у поверхности Земли

Ослабление потока излучения атмосферой сопровождается изменением его спектрального состава вследствие как поглощения, так и рассеяния. На рис. 20.5 показан спектр солнечного из-

лучения на границе земной атмосферы (кривая *1*) и у поверхности Земли (кривая *2*) при наивысшем стоянии Солнца. Кривая *1* имеет максимум на длине волны около 480 нм и близка к спектру черного тела с температурой около 6000 К, что и позволяет рассматривать ее как температуру поверхности Солнца. Максимум кривой *2* расположен вблизи 550 нм, что соответствует цветовой температуре (по закону Вина) 5230 К. Кроме того, в инфракрасном диапазоне наблюдается множество полос поглощения парами воды и атмосферными газами.

20.4. Инфракрасное излучение и его применение в медицине

Электромагнитное излучение в диапазоне от красной границы видимого света ($\lambda = 0{,}76$ мкм) и до коротковолнового радиоизлучения ($\lambda = 1$ мм) называют ***инфракрасным*** (ИК).

Мощным источником ИК-излучения является Солнце, поскольку около 50 % его излучения лежит в этом диапазоне.

Инфракрасную область спектра условно подразделяют на *ближнюю* к видимой (0,76–2,5 мкм), *среднюю* (2,5–50 мкм) и *дальнюю* (50–1000 мкм).

Если в закон Вина вместо λ_{max} подставить границы ИК-диапазона, то температуры тел, максимум теплового излучения которых приходится на эти границы, будут составлять соответственно от 3800 К до 1,5 К. Это означает, что в обычных условиях практически все тела не только являются источниками ИК-излучения, но и имеют максимальное излучение в этой части электромагнитного спектра.

Лечебное применение инфракрасного излучения основано на его тепловом действии. Наибольший лечебный эффект достигается коротковолновым ИК-излучением, близким к видимому свету. Инфракрасное излучение проникает в тело на глубину около 20 мм, поэтому в большей степени прогреваются поверхностные слои. Терапевтический эффект как раз и обусловлен возникающим температурным градиентом, что активизирует кровоснабжение облученного места и приводит к благоприятным лечебным последствиям.

В лечебных целях дозированную солнечную радиацию применяют как солнцелечение (гелиотерапия), а также как средство закаливания организма. Используют также искусственные источники теплового излучения: лампы накаливания (соллюкс) и инфракрасные излучатели.

20.5. Тепловое излучение тела человека

Температура тела человека поддерживается постоянной благодаря физиологическим механизмам терморегуляции, при этом постоянно происходит теплообмен организма с окружающей средой. Общие для всех сред способы теплообмена — теплопроводность, конвекция, испарение и тепловое излучение (поглощение).

Теплопроводность воздуха незначительна, поэтому этот вид теплоотдачи для человека не существен. Вклад остальных процессов в общий теплообмен зависит от многих факторов: температуры тела человека, состояния окружающей среды (температура, влажность, движение воздуха и др.), одежды (материал, форма, размер, цвет, толщина). Поэтому приведем лишь приближенную оценку вкладов этих процессов для людей, находящихся в покое и в условиях умеренного климата.

Теплопотери вследствие ***конвекции*** (обдувание тела воздухом) существенно зависят от наличия и вида одежды. В условиях умеренного климата на конвекцию приходится не более 15–20 % всех теплопотерь человека.

На ***испарение влаги*** с поверхности тела в условиях покоя приходится около 30 % теплопотерь, а наибольшая их доля (около 50 %) — на ***тепловое излучение*** от открытых частей тела и одежды. При этом основная часть этого излучения лежит в инфракрасном диапазоне, в области длин волн от 4 до 50 мкм.

При вычислении этих потерь сделаем два основных допущения [2].

1. Излучающие тела (кожа человека, одежда) будем считать серыми.

В этом случае для вычисления радиационных тепловых потерь с 1 м^2 поверхности тела можно воспользоваться законом Стефана—Больцмана для серого тела: $R_{изл} = \alpha\sigma T^4$.

Значения коэффициентов поглощения α некоторых серых тел приведены в табл. 20.1 [28].

Таблица 20.1

Поверхность	Кожа человека	Хлопчатобумажная ткань	Шерсть, шелк
Коэффициент поглощения α	0,90	0,73	0,76

Максимум спектральной плотности энергетической светимости тела человека при температуре поверхности кожи 32 °С (305 К), в соответствии с законом Вина приходится на длину волны 9,5 мкм:

$$\lambda_{max} = \frac{b}{T} = \frac{2898 \text{ мкм} \cdot \text{К}}{305 \text{ К}} = 9{,}5 \text{ мкм}.$$

2. Применим закон Стефана — Больцмана к излучению тела человека в неравновесных условиях (температура поверхности тела человека $T_ч$ не равна температуре окружающей среды T_0).

Человек излучает со всей поверхности S тела поток энергии $\Phi_{изл} = \alpha\sigma T_ч^4 S$. Одновременно его тело поглощает часть излучения, попадающего на него от окружающих предметов (стен, потолка и т.п.), равную $\Phi_{погл} = \alpha\sigma T_0^4 S$, поэтому реальные тепловые потери человека составляют лишь разницу этих потоков энергий: $\Phi_{потерь} = \alpha\sigma(T_ч^4 - T_0^4)S$.

Если человек одет, то под $T_ч$ следует понимать температуру поверхности одежды. Поясним роль одежды на примере [28].

Раздетый человек, со средней температурой поверхности кожи 33 °С (306 К) и площадью поверхности тела 1,5 м^2 при температуре окружающей среды 18 °С (291 К), теряет ежесекундно вследствие теплового излучения

$$P_1 = 0{,}9 \cdot 5{,}67 \cdot 10^{-8}(306^4 - 291^4) \cdot 1{,}5 = 122 \text{ Вт}.$$

При той же температуре окружающей среды, но в хлопчатобумажной одежде, температура поверхности которой

24 °С (297 К), его тепловые потери будут значительно меньше:

$$P_2 = 0{,}73 \cdot 5{,}67 \cdot 10^{-8}(297^4 - 291^4) \cdot 1{,}5 = 37 \text{ Вт}.$$

Таким образом, одежда играет роль теплоизолятора: уменьшает тепловые потери при пониженной температуре окружающей среды и предохраняет тело от избыточного перегрева при слишком высокой окружающей температуре.

20.6. Понятие о термографии

Температура различных участков поверхности тела человека неодинакова, ее распределение для здорового человека достаточно характерно и симметрично относительно продольной оси тела. Температура конечностей обычно ниже, чем других частей тела, и зависит не только от температуры окружающей среды, но и от состояния системы кровообращения. Воспалительные процессы, новообразования также значительно изменяют местную температуру.

Регистрация распределения температуры по поверхности тела человека называется ***термографией*** и является диагностическим методом. При термографии регистрируется собственное излучение тела человека, поэтому этот диагностический метод абсолютно безвреден, не имеет противопоказаний и может служить средством массового профилактического обследования населения, поскольку позволяет наблюдать изменения в распределении поверхностной температуры на ранних стадиях развития патологий.

Благодаря сильной зависимости энергетической светимости тела от температуры ($R \sim T^4$), даже небольшое изменение температуры поверхности вызывает значительное изменение излучаемой мощности и надежно фиксируется приборами — *термографами* и *тепловизорами*.

Действительно, продифференцировав уравнение (20.7а): $dR = 4\alpha\sigma T^3 dT$ и разделив почленно полученный дифференциал на (20.7а), получим

$$\frac{dR}{R} = 4\frac{dT}{T}. \qquad (20.10)$$

Это означает, что относительное изменение энергетической светимости больше относительного изменения температуры излучающей поверхности в четыре раза. Если абсолютная температура поверхности тела человека изменится на 3 °С, т.е. приблизительно на 1 %, то энергетическая светимость изменится на 4 %. Современные термографы и тепловизоры способны регистрировать изменения температуры с точностью в несколько сотых долей градуса.

Контрольные вопросы и задания

1. Каковы основные характеристики теплового излучения тел? Укажите связь между ними и единицы измерения.
2. Что такое монохроматический коэффициент поглощения? Как его значение влияет на классификацию тел как тепловых источников?
3. Сформулируйте законы теплового излучения (Кирхгофа, Стефана — Больцмана, Вина), приведите их математическое выражение и укажите область применимости.
4. Приведите формулу М. Планка для теплового излучения. Что она определяет? Каков график этой функции? Каков смысл постоянной Планка?
5. Каким образом из формулы Планка можно получить законы Вина и Стефана — Больцмана и вычислить постоянную Планка?
6. Как измерить температуру удаленных тел? Чем различаются цветовая и яркостная температуры тел, каков принцип их измерения? Как определить температуру поверхности Солнца?
7. Что представляет собой Солнце как тепловой источник? Каков спектр солнечного излучения? Каково влияние атмосферы на этот спектр? Что такое солнечная постоянная?
8. Охарактеризуйте тепловое излучение тела человека, его спектр, положение максимума излучения.
9. Каковы основные механизмы теплопередачи? Как рассчитать теплопотери организма за счет излучения?
10. Как связано относительное изменение температуры поверхности серого тела и относительное изменение его энергетической светимости?
11. Каковы принципы тепловидения и термографии в медицине? Каковы достоинства термографии как диагностического метода?

Глава 21 Оптические спектры атомов и молекул

Электромагнитное излучение, распространяясь в веществе, взаимодействует с его атомами и молекулами, и прошедшее излучение несет информацию о молекулярной структуре вещества и его энергетических состояниях.

При прохождении света через вещество часть энергии излучения поглощается, часть рассеивается, причем на разных частотах (длинах волн) поглощается и рассеивается разное количество электромагнитной энергии, в зависимости от строения и состава вещества. Поэтому, изучая спектры прошедшего и рассеянного излучения, можно определить вид, состав и концентрацию поглощающих и рассеивающих центров в исследуемом веществе.

Спектры испускания и поглощения вещества, лежащие в ультрафиолетовом, видимом и инфракрасном диапазонах, называют ***оптическими спектрами***. Они отражают энергетические состояния его атомов и молекул, поэтому весьма чувствительны к изменению химических связей, их окружения, pH среды, воздействию внешних электрических и магнитных полей. Спектры каждого вещества строго индивидуальны и являются его «паспортом», что позволяет по спектрам излучения или поглощения определить вид вещества, его состав и процентное содержание компонентов смеси, т.е. провести *качественный* и *количественный спектральный анализ*.

Спектральный анализ является одним из важнейших неразрушающих методов исследования как структуры вещества, так и физических и химических процессов, происходящих в этом веществе на атомном и молекулярном уровнях, и широко используется как метод исследования в физике, химии, молекулярной биологии и медицине.

21.1. Поглощение света

Явление поглощения света хорошо изучено и состоит в уменьшении интенсивности световой волны при ее распространении в веществе.

Экспериментально установлено, что при прохождении света через тонкий поглощающий слой его интенсивность изменяется на величину dI, которая прямо пропорциональна толщине dx этого слоя и интенсивности I падающего на него излучения (рис. 21.1):

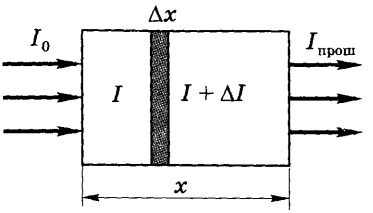


Рис. 21.1. Прохождение света через поглощающее вещество

$$dI = -kIdx, \qquad (21.1)$$

где k — коэффициент пропорциональности, знак «−» указывает на уменьшение интенсивности излучения после прохождения через слой.

Разделяя переменные и интегрируя, получим сначала общее, а затем и частное решение:

$$dI/I = -kdx,$$
$$\int dI/I = -k\int dx,$$
$$\ln I = -kx + \ln C,$$
$$I = Ce^{-kx}.$$

Учитывая, что $I = I_0$ при $x = 0$, находим $C = I_0$. В результате получим закон поглощения света, определяющий зависимость интенсивности I прошедшего излучения от пути x, пройденного светом в веществе:

$$I = I_0 e^{-kx}, \qquad (21.2)$$

который известен как ***закон Бугера.*** Он был открыт французским ученым П. Бугером экспериментально (1729), а позже теоретически выведен И. Ламбертом (1760).

Этот закон справедлив для всех веществ, а их индивидуальные поглощающие свойства полностью определяются *показателями поглощения* k, которые измеряются в обратных метрах ($м^{-1}$) или ($см^{-1}$) и являются одними из важнейших оптических характеристик вещества. Величина показателя поглощения вещества различна на разных длинах волн. Зависимость $k(\lambda)$ или $k(\nu)$ индивидуальна для каждого вещества

и определяет его спектр поглощения и разную цветовую окраску поглощающих веществ.

Для растворов при небольших концентрациях справедлив ***закон Бера*** (1852): показатель поглощения раствора прямо пропорционален концентрации C поглощающего вещества:

$$k = \alpha C, \tag{21.3}$$

где α — удельный (в расчете на единицу концентрации) показатель поглощения вещества, зависящий, как и k, от длины волны.

Подставляя выражение (21.3) в формулу (21.2), получим закон поглощения света для растворов — ***закон Бугера — Ламберта — Бера:***

$$I = I_0 e^{-\alpha C x}. \tag{21.4}$$

Введем теперь две общепринятые безразмерные величины, характеризующие поглощение света образцом:

• *коэффициент пропускания* T — отношение интенсивности света, прошедшего образец, к интенсивности падающего на него света:

$$T = \frac{I}{I_0} = \frac{I_0 e^{-\alpha C x}}{I_0} = e^{-\alpha C x}; \tag{21.5}$$

• *оптическую плотность* D образца, равную десятичному логарифму коэффициента пропускания, взятому со знаком «–»:

$$D = -\lg T = k_1 x = \alpha_1 C x, \tag{21.6}$$

где $k_1 = k \lg e = 0{,}43k$, $\alpha_1 = \alpha \lg e = 0{,}43\alpha$.

Из формул (21.5) и (21.6) видно, что зависимость коэффициента пропускания T от толщины x образца и концентрации C растворенного поглощающего вещества экспоненциальная, тогда как оптическая плотность D образца линейно зависит от этих параметров, что очень удобно при измерениях концентрации веществ.

Кроме того, поскольку $D(\lambda) \sim k(\lambda) \sim \alpha(\lambda)$, они одинаково зависят от длины волны и представляют собой *спектр поглощения вещества*. На практике в качестве спектра поглощения вещества удобнее регистрировать зависимость оптической плотности $D(\lambda)$ от длины волны.

Закон Бугера — Ламберта — Бера лежит в основе *концентрационной колориметрии* — фотометрического метода определения концентрации поглощающего вещества в окрашенных растворах. Для этой цели используются однолучевые или двулучевые фотоэлектроколориметры (ФЭКи).

Принципиальная схема однолучевого ФЭКа представлена на рис. 21.2. Свет от источника *1* формируется линзой *2* в параллельный пучок. Светофильтр *3* пропускает сравнительно узкий спектральный диапазон вблизи длины волны λ. Этот почти монохроматический пучок света сначала проходит через кювету *4* с чистым растворителем и оптический клин *5* и попадает на фотоэлемент *6*. Перемещая поглощающий оптический клин *5*, показания измерительного прибора *7* устанавливают точно на 100 % пропускания. Затем на пути светового пучка вместо кюветы с растворителем помещают кювету *4а* с исследуемым раствором, и измерительный прибор сразу показывает его коэффициент пропускания и оптическую плотность, которые позволяют определить концентрацию поглощающего вещества в растворе.

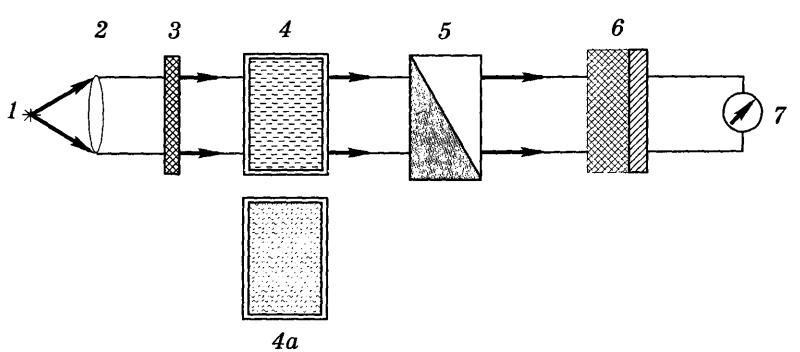


Рис. 21.2. Принципиальная схема однолучевого ФЭКа

21.2. Рассеяние света

При прохождении через среду часть световых лучей отклоняется от первоначального направления на значительные углы, вплоть до 180°. Это явление называют ***рассеянием света***. Оно возникает из-за оптической неоднородности среды, т.е. наличия областей с разными показателями преломления.

Различают два основных вида таких неоднородностей.

Неоднородность состава (гетерогенность) среды, т.е. наличие в однородном веществе инородных частиц. Такие среды называются ***мутными***, к ним относятся взвеси, эмульсии, дым (твердые частицы в газе), туман (капельки жидкости в газе) и т.п. Рассеяние света в мутных средах называют ***рассеянием Тиндаля***, по имени английского физика, впервые (1868) изучившего это явление. Он заметил, что в направлении, перпендикулярном падающему свету (направление *B* на рис. 21.3), рассеянный свет имеет более голубой оттенок, чем свет источника *S*, и распространяется в виде конуса лучей (конус Тиндаля), хорошо видимого на темном фоне, тогда как свет, прошедший прямо сквозь кювету (в направлении *A*), имеет красноватый оттенок.

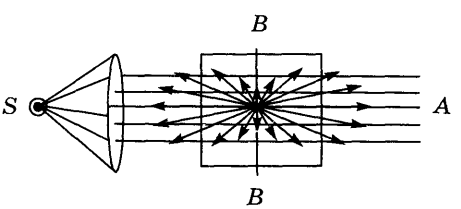


Рис. 21.3. Рассеяние света (стрелками показаны направления рассеянного света)

Оптические неоднородности, возникающие в однородной (гомогенной) среде вследствие флуктуаций плотности, т.е. временных отклонений плотности среды от ее среднего значения из-за непрерывного хаотического теплового движения молекул. Рассеяние света на таких неоднородностях называют ***молекулярным рассеянием*** (рассеянием в чистых средах) или ***рассеянием Рэлея***. Он установил (1871), что при

молекулярном рассеянии, а также при рассеянии в мутных средах с размерами неоднородностей $d < 0{,}2\lambda$ интенсивность рассеянного света обратно пропорциональна четвертой степени длины волны λ — ***закон Рэлея:***

$$I_{\text{рас}} \sim \frac{1}{\lambda^4}. \qquad (21.7)$$

Этот закон означает, что в однородной чистой среде будут рассеиваться преимущественно голубые и фиолетовые лучи, а красные лучи будут лучше проходить сквозь нее. Поэтому спектр рассеянного излучения сильно смещается в сине-голубую область, а прошедшего — в красную.

В чистых однородных средах красный свет рассеивается почти на порядок (в 10 раз) меньше, чем синий, и поэтому виден на значительно более дальних расстояниях. Это свойство молекулярного рассеяния используют в сигнализации: сигнальные огни на высоких башнях, вышках, зданиях, на аэродромах — красные, наиболее ответственный свет светофора — тоже красный. Именно молекулярное рассеяние определяет голубой цвет чистого неба и красный цвет Солнца на восходе и закате, когда его свет доходит до наблюдателя через большую толщу атмосферы и большая часть голубых и фиолетовых лучей в ней рассеивается.

Значительно меньше рассеиваются, а значит, лучше проходят через среду инфракрасные лучи, поскольку их длина волны существенно больше, чем у видимого света. Это позволяет использовать инфракрасные приборы для наблюдения за окружающей обстановкой в условиях ограниченной видимости (в сильном тумане, при задымлении и в других средах).

При рассеянии света в мутной среде (рассеяние Тиндаля), когда размеры неоднородностей превышают $d > 0{,}2\lambda$, интенсивность рассеянного света уже меньше зависит от длины волны, — она обратно пропорционально второй степени длины волны:

$$I_{\text{рас}} \sim \frac{1}{\lambda^2}. \qquad (21.8)$$

В этом случае спектр рассеянного излучения меньше смещен в сине-голубую область, он теряет свою голубизну и ста-

новится белее. Так, чистые облака, чистый дым, плотный туман кажутся нам белесыми в противоположность синему небу чистой атмосферы.

При дальнейшем увеличении размеров неоднородностей зависимость $I_{рас}$ от длины волны становится еще более слабой и на крупных неоднородностях уже практически незаметна.

Интенсивность света, рассеянного под разными углами, не одинакова и отображается специальной диаграммой — ***индикатрисой рассеяния***, которая показывает относительную интенсивность рассеянного излучения в разных направлениях φ. Она, как правило, симметрична относительно оси первичного пучка. При молекулярном рассеянии индикатриса имеет вид, показанный на рис. 21.4, и выражается формулой $I_{рас}/I_0 = (1 + \cos^2\varphi)$.

По мере увеличения размеров частиц распределение интенсивности рассеянного света вперед и назад перестает быть симметричным и начинает преобладать рассеяние вперед (рис. 21.5).

Форма индикатрисы рассеяния зависит от формы и размеров рассеивающих частиц.

Интенсивность излучения, прошедшего через слой рассеивающей среды толщиной x, уменьшается по закону Бугера:

$$I_{прош} = I_0 e^{-\sigma x}, \qquad (21.9)$$

где σ — показатель рассеяния, зависящий от свойств вещества и длины волны падающего излучения, $м^{-1}$ ($см^{-1}$); I_0 — интенсивность падающего излучения.

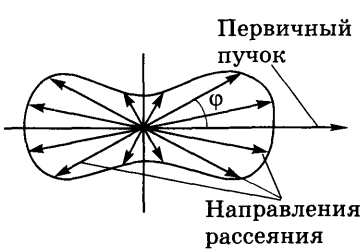


Рис. 21.4. Индикатриса молекулярного рассеяния света

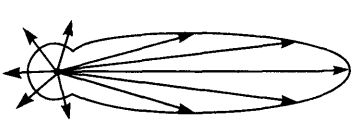


Рис. 21.5. Индикатриса рассеяния на крупных частицах, $d > 0{,}2\lambda$

При одновременном наличии в среде и поглощения и рассеяния интенсивность прошедшего излучения определяется обобщенной зависимостью

$$I_{\text{прош}} = I_0 e^{-(\sigma+k)x}, \qquad (21.10)$$

где $(\sigma + k) = \mu$ — *показатель ослабления.*

Изучая параметры рассеянного излучения (спектр, степень поляризации, индикатрису рассеяния и ряд других), можно определить свойства рассеивающей среды, получить информацию о межмолекулярных взаимодействиях, размерах макромолекул среды, частиц в коллоидных растворах и эмульсиях, определить состав аэрозолей и другие параметры.

Методы исследования рассеянного излучения с целью получения подобной информации называют ***нефелометрией***, а соответствующие приборы — ***нефелометрами***. В настоящее время широко используется зондирование атмосферы специальными лазерными установками (лидарами), которое позволяет по измеренным характеристикам рассеянного атмосферой лазерного излучения определить качественный и количественный состав аэрозолей на расстояниях в несколько километров от источника излучения. Для получения таким способом карты состояния воздушного бассейна над большим городом требуется лишь несколько десятков минут.

21.3. Энергетические состояния атома

Оптические спектры многих веществ были хорошо известны и систематизированы уже в XIX в., но физическая природа их возникновения была неясна.

Впервые правильное объяснение наблюдавшимся спектрам излучения и поглощения атомов дал Н. Бор в начале XX в. Основываясь на планетарной модели атома, выдвинутой Э. Резерфордом, и на квантовых свойствах света, установленных М. Планком, он постулировал следующие положения:

1. Атомная система может находиться только в определенных энергетических состояниях E_1, E_2, E_3, ..., E_n, в которых она не излучает и не поглощает энергию; эти состояния назвали *стационарными*.

2. При переходе атома из одного стационарного состояния с энергией E_n в другое, с энергией E_k, атом может излучать (если $E_n > E_k$) либо поглощать (если $E_n < E_k$) энергию в виде кванта электромагнитного поля, при этом энергия $h\nu$ кванта излучения определяется соотношением (условие Бора)

$$h\nu = |E_n - E_k|. \qquad (21.11)$$

3. Стационарными являются состояния, в которых момент импульса электрона принимает значения, кратные величине $\hbar = h/2\pi$:

$$mvr = nh/2\pi = n\hbar, \qquad (21.12)$$

где h — постоянная Планка, $\hbar = h/2\pi$; m, v, r — масса, скорость и радиус орбиты электрона; $n = 1, 2, 3, 4, \ldots$ — целое положительное число.

Постулаты Бора, хотя и противоречили классической электродинамике, дали возможность вычислить радиусы стационарных орбит и энергию электрона в этих состояниях для простейшего атома — атома водорода и водородоподобных ионов, т.е. имеющих только один электрон, и затем рассчитать их спектры излучения и поглощения.

Теория Бора впервые объяснила и правильно описала наблюдавшиеся спектры излучения и поглощения атома водорода и водородоподобных ионов, указала правильное направление для последующего развития всей квантовой механики, объяснившей явления, происходящие на атомном и молекулярном уровнях.

Таким образом, было установлено, что *испускание* атомами квантов электромагнитного излучения происходит при переходе атомов из более высоких энергетических состояний на более низкие, а *поглощение* квантов электромагнитной энергии сопровождается переходами атомов с нижних энергетических состояний на более высокие в соответствии с условием (21.11).

Поскольку энергия атома прямо зависит от заряда ядра Ze и, следовательно, от номера Z элемента, то набор возможных энергетических состояний у атома каждого химического элемента или у молекулы строго индивидуален, поэтому спектры

испускания и поглощения атомов и молекул также строго индивидуальны. Благодаря этому по ним можно определять химический состав образца. В этом состоит одно из основных направлений практического применения спектрального анализа.

Рассмотрим подробнее энергетические состояния простейших, одноэлектронных атомов и ионов с учетом постулатов Бора.

На электрон, движущийся по круговой орбите радиусом r вокруг ядра с зарядом Ze, действует кулоновская сила $F_{\text{Кул}} = Ze^2/4\pi\varepsilon_0 r^2$, сообщающая электрону необходимое центростремительное ускорение в соответствии со вторым законом Ньютона:

$$F_{\text{Кул}} = ma_{\text{ц.с}}, \tag{21.13}$$

$$\text{т.е. } Ze^2/4\pi\varepsilon_0 r^2 = mv^2/r, \text{ или } Ze^2/4\pi\varepsilon_0 r = mv^2, \tag{21.13а}$$

где ε_0 — электрическая постоянная вакуума; e и m — заряд и масса электрона; r и v — радиус орбиты электрона в атоме и его скорость на этой орбите.

Возведя соотношение (21.12) в квадрат и разделив его на (21.13а), найдем радиусы стационарных орбит электрона в атоме водорода и водородоподобных ионов:

$$r_n = n^2\varepsilon_0 h^2/Ze^2\pi m, \tag{21.14}$$

где $n = 1, 2, 3, 4, \ldots$ — номер орбиты.

При $n = 1$ и $Z = 1$ получим радиус первой, наименьшей боровской орбиты электрона в атоме водорода:

$$r_1 = \varepsilon_0 h^2/e^2\pi m = 5{,}3 \cdot 10^{-11} \text{ м} = 0{,}053 \text{ нм}. \tag{21.14а}$$

Потенциальная энергия электрона в электрическом поле ядра отрицательная и равна

$$E_{\text{пот}} = -e\varphi = -eZe/4\pi\varepsilon_0 r = -Ze^2/4\pi\varepsilon_0 r, \tag{21.15}$$

а кинетическая энергия электрона с учетом (21.13а)

$$E_{\text{кин}} = mv^2/2 = Ze^2/8\pi\varepsilon_0 r. \tag{21.16}$$

Поскольку потенциальная энергия электрона отрицательная и по модулю вдвое превосходит его кинетическую энер-

гию, то полная энергия электрона в атоме всегда отрицательная. Это указывает на то, что электрон не свободен, а связан со своим ядром (энергия этой связи равна модулю полной энергии электрона):

$$E = E_{\text{пот}} + E_{\text{кин}} = -Ze^2/8\pi\varepsilon_0 r. \qquad (21.17)$$

Подставляя в формулу (21.17) выражение (21.14) для радиусов стационарных орбит, получим выражение для полной энергии электрона в стационарных состояниях:

$$E_n = -Z^2e^4m/8\varepsilon_0{}^2h^2n^2 = -Z^2E_0/n^2. \qquad (21.8)$$

Здесь $E_0 = e^4m/8\varepsilon_0{}^2h^2 = 2{,}176 \cdot 10^{-18}$ Дж $= 13{,}6$ эВ (электронвольт, 1 эВ $= 1{,}6 \cdot 10^{-19}$ Дж) — энергия атома водорода в основном состоянии, т.е. энергия электрона на первой боровской орбите атома водорода.

Таким образом, полная энергия электрона в атоме всегда отрицательная и ее величина прямо пропорциональна квадрату заряда ядра (т.е. квадрату порядкового номера элемента в таблице Менделеева) и обратно пропорциональна квадрату номера орбиты. Наибольшую по модулю (но отрицательную по знаку) полную энергию электрон имеет на 1-й, самой близкой к ядру орбите (при $n = 1$): $E_1 = -Z^2E_0$, поэтому это энергетическое состояние является *основным* для электрона, а соответствующий энергетический уровень — самый нижний на энергетической диаграмме. Остальные энергетические состояния (E_2, E_3 и т.д.) расположены выше, в обычных условиях свободны, и электрон может переходить в эти состояния лишь при поглощении необходимых порций энергии, например, в виде квантов электромагнитного излучения в соответствии со вторым постулатом Бора. Набор частот электромагнитного излучения, который может поглощать данный атом, определяет его *спектр поглощения*. Затем, при последующих обратных переходах электрона из верхних, возбужденных энергетических состояний на более низкие (в том числе и на основное E_1) состояния, атом излучает электромагнитные волны соответствующих частот (21.11), образующих его *спектр испускания*.

21.4. Спектр атома водорода

Рассмотрим более подробно энергетические состояния простейшего атома — атома водорода.

В соответствии с постулатами Бора энергия атома водорода может принимать только вполне определенные значения, описываемые формулой (21.18) при $Z = 1$:

$$E_n = -E_0/n^2,\ n = 1, 2, 3, 4... \qquad (21.18а)$$

Положительное целое число n, определяющее радиусы стационарных орбит (21.4) и энергию атома в этих состояниях (21.18) и (21.18а), называют ***главным квантовым числом.***

В первом энергетическом состоянии, при $n = 1$, полная энергия атома наименьшая (а ее модуль — энергия связи электрона — наибольшая: $E_{св} = |E_0| = 13{,}6$ эВ), поэтому это состояние является для электрона основным, и соответствующий ему уровень энергии является самым глубоким на энергетической диаграмме.

В других состояниях (при $n = 2, 3, 4, 5...$) полная энергия больше, а энергия связи с ядром меньше, поэтому электрон долго находиться там не может и примерно через 10^{-9} с переходит на основной, энергетически более выгодный уровень, испуская кванты соответствующей энергии.

Частота ν излучения, возникающего при переходе электрона с энергетического уровня E_k на уровень E_n, определяется правилом частот Бора (21.11) и соотношением (21.18а):

$$\nu = \frac{E_k - E_n}{h} = \frac{E_0}{h}\left(\frac{1}{n^2} - \frac{1}{k^2}\right), \qquad (21.19)$$

где $n = 1, 2, 3, 4, ...; k = n + 1, n + 2, n + 3, ...$

Формула (21.19), правильно описывающая частоты всех линий в спектре излучения атома водорода, была эмпирически установлена Бальмером для видимой части спектра водо-

рода (для $n = 2$ и $k = 3, 4, 5...$) и названа его именем задолго до теории Бора, но лишь теория Бора смогла правильно объяснить происхождение этих линий, что и явилось одним из важнейших доказательств ее правоты.

Таким образом, частоты всех спектральных линий в спектре атома водорода описываются *формулой Бальмера* (21.19). Диаграмма энергетических уровней атома водорода, построенная в соответствии с формулой (21.18а), и схема основных переходов, образующих спектры испускания водорода, представлена на рис. 21.6.

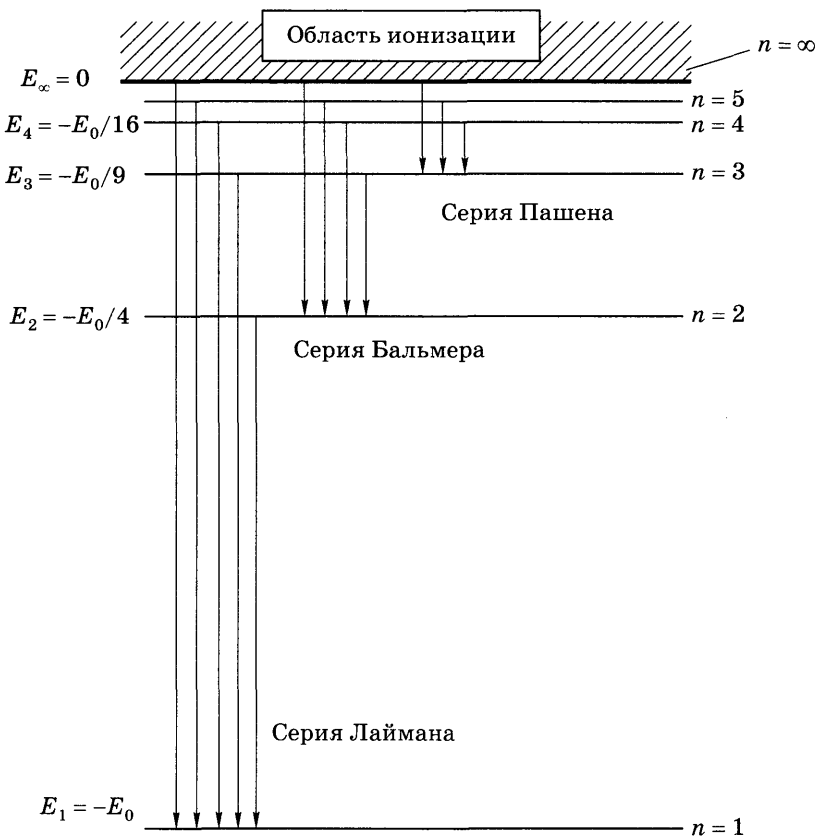


Рис. 21.6. Диаграмма энергетических уровней атома водорода и схема образования основных спектральных серий в излучении

Совокупность спектральных линий, соответствующих переходам с любых верхних энергетических уровней на один и тот же нижний уровень, образует спектральную серию в излучении. В спектре атома водорода выделяют следующие основные ***спектральные серии.***

1. *Серия Лаймана* возникает при переходах электронов со всех верхних уровней ($k = 2, 3, 4 \ldots$) в основное состояние ($n = 1$). Частоты спектральных линий этой серии получают из формулы (21.19):

$$\nu_{\text{Л}} = \frac{E_0}{h}\left(1 - \frac{1}{k^2}\right), \qquad (21.20)$$

где $k = 2, 3, 4, 5 \ldots$

Линии серии Лаймана лежат в ультрафиолетовой области спектра.

2. *Серия Бальмера* соответствует переходам на второй уровень ($n = 2$) со всех верхних уровней ($k \geq 3$). Их частоты определяются формулой Бальмера

$$\nu_{\text{Б}} = \frac{E_0}{h}\left(\frac{1}{4} - \frac{1}{k^2}\right), \qquad (21.21)$$

где $k = 3, 4, 5, 6 \ldots$

Линии этой серии лежат в видимой области спектра.

3. *Серия Пашена* соответствует переходам со всех верхних уровней на третий ($n = 3$):

$$\nu_{\text{П}} = \frac{E_0}{h}\left(\frac{1}{9} - \frac{1}{k^2}\right), \qquad (21.22)$$

где $k = 4, 5, 6, 7 \ldots$

Эта серия лежит в инфракрасной (ИК) области спектра.

Указанные спектральные серии наблюдаются также и в спектрах поглощения атома водорода, где им соответствуют переходы с одного нижнего уровня (1-го — для серии Лаймана, 2-го — Бальмера и 3-го — Пашена) на все более высокие энергетические уровни.

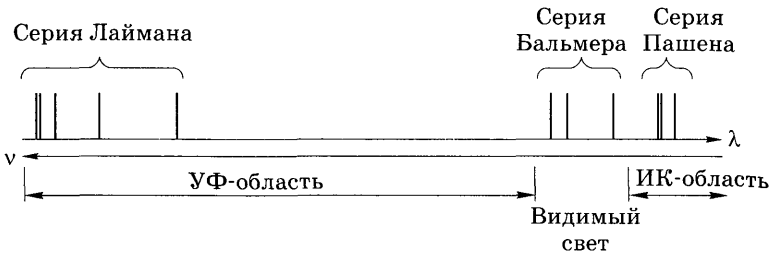


Рис. 21.7. Схема расположения спектральных линий основных серий атома водорода

Как видим, спектры атомов состоят из отдельных узких линий, поэтому называются ***линейчатыми*** (рис. 21.7).

Для атомов характерно совпадение частот (длин волн) линий в спектрах испускания и поглощения.

21.5. Спектры сложных атомов

Атомы называют ***сложными***, если они имеют на своих орбитах два и более электрона. Хотя теория Бора правильно объяснила происхождение спектров атома водорода и водородоподобных ионов, описать так же точно спектры сложных атомов она уже не смогла, но послужила отправной точкой для развития более точной теории атома — ***квантовой механики.***

Согласно квантовой механике, состояние электрона в атомах описывается четырьмя *квантовыми числами*: главным квантовым числом n, орбитальным квантовым числом l, магнитным числом m и спиновым числом m_s[1].

Главное квантовое число n определяет энергию электрона в атоме и может принимать последовательные целые значения от 1 и выше.

Орбитальное (побочное или азимутальное) квантовое число l определяет орбитальный момент импульса электрона

$$p = \sqrt{l(l+1)} \cdot h / 2\pi = \sqrt{l(l+1)} \cdot \hbar \qquad (21.23)$$

[1] Часто используют и другую четверку квантовых чисел: n, l, j, m_j, где числа j и m_j определяют полный механический момент электрона и его проекцию на выделенное направление.

и форму его орбиты (круговая, эллиптичная) и может принимать n целых значений от нуля до $(n-1)$: $l = 0, 1, 2, ..., n-1$.

Магнитное число m может принимать $(2l+1)$ целых значений от $-l$ до $+l$: $m = -l, ..., 0, ..., +l$; оно и определяет проекцию орбитального момента импульса электрона на направление OZ внешнего поля:

$$p_z = mh/2\pi = m\hbar. \qquad (21.24)$$

Спиновое квантовое число m_s принимает лишь два значения ($m_s = \pm 1/2$) и определяет проекцию спина (т.е. собственного механического момента) электрона на выделенное направление.

Электронные состояния в зависимости от значения орбитального числа l обозначаются соответственно как s-состояния ($l = 0$, круговая орбита), p-состояния ($l = 1$, эллиптическая орбита), d-состояния ($l = 2$, эллиптическая более вытянутая орбита), f-состояния ($l = 3$, эллиптическая, еще более вытянутая орбита) и т.д.

По ***принципу Паули*** в одном квантовом состоянии (т.е. состоянии с определенным набором четырех квантовых чисел n, l, m и m_s) в атоме может находиться не более одного электрона. В общем случае число электронов в состоянии с заданным орбитальным числом l и разными значениями m и m_s равно $2(2l+1)$. Поэтому в s-состояниях ($l = 0$) может находиться не более чем 2 электрона ($m = 0$, $m_s = \pm 1/2$), в p- состояниях ($l = 1$) — 6 электронов ($m = 0, \pm 1$, $m_s = \pm 1/2$), в d-состояниях ($l = 2$) — 10 электронов ($m = 0, \pm 1, \pm 2$, $m_s = \pm 1/2$) и т.д.

Основное отличие сложных атомов от атома водорода в том, что энергия электронов в сложных атомах зависит не только от главного квантового числа n, но и от орбитального квантового числа l[1].

Энергия связи электрона в атоме обычно тем больше (т.е. энергетический уровень расположен глубже), чем меньше главное квантовое число n, а при заданном n — чем меньше орбитальное число l. Поэтому по мере возрастания заря-

[1] При воздействии на атом внешних электрических и магнитных полей энергия электронов будет зависеть и от магнитного m, и от спинового m_s квантового числа.

да ядра Z и, следовательно, количества электронов в атоме заполнение электронами энергетических уровней в сложных атомах начинается с нижнего состояния $1s$, в котором может находиться не более 2 электронов. Затем заполняются $2s$-состояния (2 электрона), $2p$-состояния (6 электронов), $3s$-, $3p$- и т.д., так что электронные состояния (конфигурации) многоэлектронных атомов заполняются в следующем порядке:

$$1s^2 2s^2 2p^6 3s^2 3p^6 4s^2 3d^{10} 4p^6 \ldots,$$

где цифра справа вверху указывает максимально возможное количество электронов в данном состоянии.

На самом верхнем энергетическом уровне, занятом электронами, находятся валентные электроны атома, и это состояние для них будет основным энергетическим состоянием. Все более высокие энергетические уровни обычно свободны, а все более низкие — заняты электронами атома.

При поглощении атомом электромагнитной энергии происходят переходы *валентных электронов* атома на свободные, более высокие энергетические состояния (уровни энергии), и совокупность соответствующих частот ν (или длин волн $\lambda = c/\nu$) электромагнитных волн, поглощаемых атомом, определяет его спектр поглощения.

Обратные переходы с верхних энергетических уровней на свободные нижние определяют *спектр испускания атома.* Во всех случаях энергия поглощаемого или излучаемого электромагнитного кванта связана с энергией перехода правилом Бора (21.11): $h\nu = |E_n - E_k|$.

Следует отметить, что в атомах и молекулах переходы между энергетическими уровнями подразделяются на излучательные (оптические) и безизлучательные (неоптические, тепловые) переходы. *Безизлучательные переходы* возможны между любыми уровнями энергии и они тем вероятнее, чем меньше энергия перехода $\Delta E = |E_n - E_k|$.

Оптические спектры атомов определяются, естественно, только излучательными (оптическими) переходами электронов из одних энергетических состояний в другие, а такие переходы возможны лишь между определенными уровнями энергии и определяются специальными правилами отбора

квантовой механики. В частности, оптические переходы разрешены лишь с изменением орбитального квантового числа l на единицу и без изменения спина S электронного состояния атома: $\Delta l = \pm 1$, $\Delta S = 0$.

Переходы, запрещенные этими правилами отбора, могут происходить, но их вероятность значительно (примерно в 10^4–10^6 раз) меньше, чем разрешенных переходов. Поскольку интенсивность спектральных линий прямо пропорциональна количеству оптических переходов в единицу времени, т.е. прямо пропорциональна вероятности этих переходов, то интенсивность спектральных линий, соответствующих разрешенным переходам, велика, а интенсивность спектральных линий, соответствующих запрещенным переходам, обычно мала.

21.6. Спектры молекул

В молекулах, кроме движения электронов вокруг ядер, существуют и другие виды движения, которые тоже квантуются, т.е. принимают дискретные значения и описываются соответствующими квантовыми числами. Это колебательное движение ядер атомов относительно их положения равновесия и вращение всей молекулы в целом вокруг осей, проходящих через ее центр масс. Полная энергия молекулы складывается из энергий этих видов движения:

$$E_{\text{мол}} = E_{\text{эл.дв}} + E_{\text{кол.дв}} + E_{\text{вращ}}, \qquad (21.25)$$

при этом энергия электронного движения значительно больше энергии колебательного движения ядер, а она в свою очередь значительно больше энергии вращения молекулы, т.е. выполняются неравенства

$$E_{\text{эл.дв}} >> E_{\text{кол.дв}} >> E_{\text{вращ}}.$$

Энергия электронного движения описывается квантовыми числами n и l, а колебательного и вращательного движений — квантовыми числами V и j соответственно.

В результате структура энергетических уровней молекулы значительно усложняется по сравнению с энергетической структурой атома. У каждого электронного состояния ($n = 1$,

2, 3, ...) появляются свои колебательные (V = 1, 2, ...) и вращательные (j = 1, 2, 3, ...) подуровни (рис. 21.8). Наличие колебательных и мелких вращательных подуровней приводит к образованию у каждого электронного состояния сплошных зон колебательно-вращательных состояний, причем в обычных условиях лучше «заселены» нижние подуровни этих зон (см. кривые заселенности уровней слева от диаграммы на рис. 21.8).

Переходы молекулы с нижних энергетических уровней на верхние будут сопровождаться поглощением энергии молекулой, а оптические переходы с верхних состояний на нижние сопровождаются испусканием электромагнитных волн.

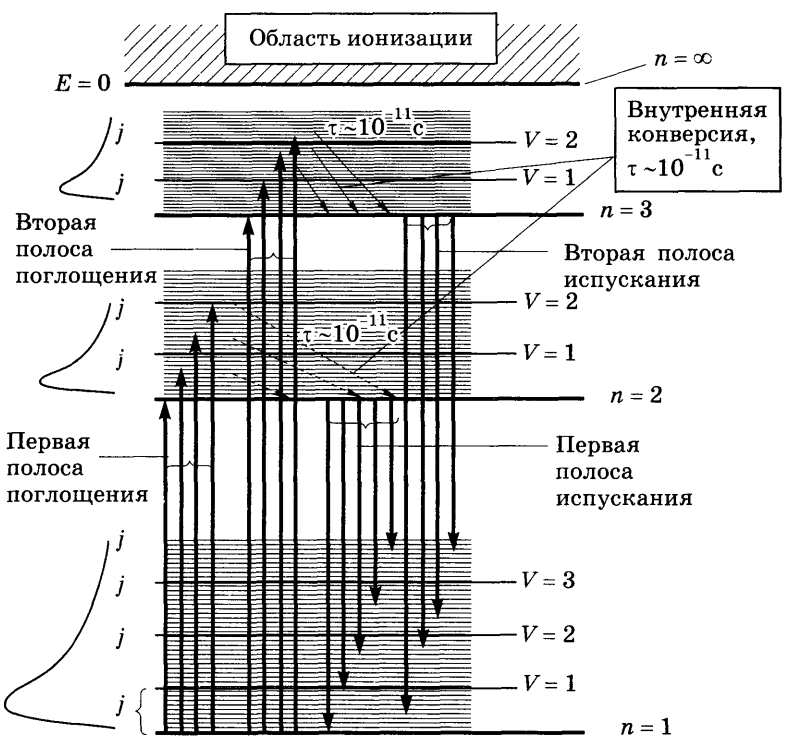


Рис. 21.8. Диаграмма энергетических уровней молекулы (вертикальными стрелками показаны оптические переходы, образующие электронные спектры поглощения и испускания)

Поглощение обычно идет с самых нижних, наиболее заселенных, подуровней основного состояния на любые колебательно-вращательные подуровни верхних электронных состояний, образующих *сплошную зону состояний* (см. рис. 21.8). Поэтому в спектрах поглощения молекул наблюдаются не отдельные линии (как у атомов), а широкие спектральные полосы поглощения.

Первая полоса поглощения соответствует переходам из основного состояния на различные колебательно-вращательные подуровни первого возбужденного электронного состояния.

Вторая полоса поглощения соответствует переходам с основного уровня на колебательно-вращательные подуровни второго возбужденного электронного состояния и т.д.

В возбужденных колебательно-вращательных состояниях молекула долго находиться не может и за время 10^{-11}–10^{-12} с совершает безызлучательные переходы на энергетически более выгодные, самые нижние колебательно-вращательные подуровни этого же электронного состояния, — это явление называется ***внутренней конверсией.***

Вследствие внутренней конверсии испускание электромагнитных квантов происходит при переходе молекулы с самых нижних подуровней возбужденных состояний на любые подуровни нижележащих (главным образом основного) электронных состояний. В результате образуются полосы испускания (люминесценции) — первая, вторая и т.д., которые не совпадают по частоте с полосами поглощения, а смещены относительно них в сторону меньших частот (больших длин волн) — эта закономерность молекулярных спектров известна как ***закон Стокса.***

Рассмотрим подробнее виды молекулярных спектров.

Частоты отдельных линий в спектрах молекул определяются условием Бора:

$$\nu = \frac{|E_{\text{мол.2}} - E_{\text{мол.1}}|}{h} = \frac{\Delta E_{\text{эл}} + \Delta E_{\text{кол}} + \Delta E_{\text{вращ}}}{h}. \quad (21.26)$$

Если изменяются все три вида энергии при переходе молекулы из одного состояния в другое, то спектры называются

электронно-колебательно-вращательными или ***электронными*** (по наибольшей энергии). Они имеют вид достаточно широких спектральных полос, расположенных в ультрафиолетовой (УФ), видимой или ближней инфракрасной (ИК) области спектра (рис. 21.9).

Электронные спектры поглощения большинства биомолекул (белков, нуклеиновых кислот и др.) лежат в УФ-области спектра. Они используются для идентификации исследуемого вещества, исследования кинетики химических реакций, а также для определения концентрации растворенного вещества. Спектры поглощения позволяют получать информацию о свойствах макромолекул и их взаимодействиях с другими молекулами, поскольку эти спектры существенно зависят от ближайшего и внешнего окружения поглощающей молекулы.

Поглощение белков в области длин волн 250–300 нм обусловлено главным образом ароматическими аминокислотами: триптофаном (рис. 21.9), тирозином и фенилаланином, хотя УФ-излучение с длиной волны 250 нм поглощают также гистидин и серосодержащие аминокислоты — цистин, цистеин и метионин.

Электронные спектры испускания называют также *спектрами люминесценции*.

Колебательно-вращательные (колебательные) спектры молекул наблюдаются только в поглощении, когда энергия

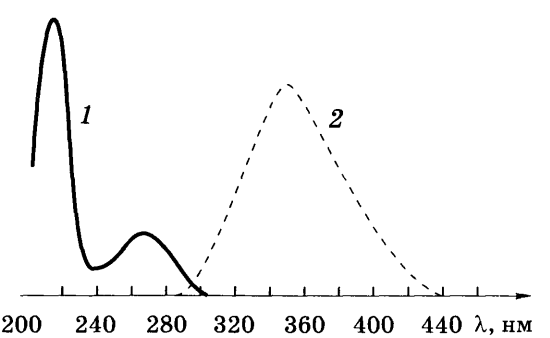


Рис. 21.9. Электронные спектры триптофана в относительных единицах:
1 — поглощения; *2* — флуоресценции

движения электронов не изменяется, т.е. $\Delta E_{эл} = 0$, а изменяются лишь колебательная и вращательная энергии молекулы в пределах основного электронного состояния. Они лежат в инфракрасной области спектра и имеют вид полос, зачастую перекрывающихся между собой.

Колебательные спектры дают информацию об энергиях валентных связей молекулы, энергиях межмолекулярных взаимодействий, о конформационных и других изменениях в структуре молекул и поэтому широко используются при спектральных исследованиях молекул.

Вращательные спектры молекул образуются при переходах между вращательными подуровнями одного и того же электронно-колебательного состояния, т.е. при $\Delta E_{эл} = 0$ и $\Delta E_{кол} = 0$, но $\Delta E_{вращ} \neq 0$, и могут наблюдаться лишь в радиодиапазоне.

В заключение приведем информацию о колебательных спектрах поглощения воды, обусловленных разными видами колебаний групп ОН и расположенных в инфракрасной области спектра (табл. 21.1). Мягкие ткани человека содержат большое количество воды и поэтому поглощают электромагнитное излучение этих диапазонов.

Таблица 21.1

Показатель	Вид колебания группы ОН		
	Валентные	Деформационные	Торсионные
Граница полосы поглощения, мкм	2,7–3,6	5,6–6,7	10–20
Максимум поглощения, мкм	2,9	6,1	13,5
Относительная оптическая плотность, $D_{отн}$	0,7	0,14	2,0

21.7. Люминесценция

Если атом или молекулу каким-либо образом перевести в возбужденное электронное состояние, то они затем самопроизвольно возвращаются в основное электронное состояние. Электромагнитное излучение, сопровождающее такие переходы, называют люминесценцией. Чтобы отличать люминесценцию от теплового излучения тел и рассеянного света, С.И. Вавилов предложил следующее ее определение: ***Люминесценция*** — это излучение, превышающее тепловое при данной температуре и имеющее длительность τ послесвечения много больше периода световых колебаний (т.е. $\tau > 10^{-15}$ с).

Интенсивность люминесценции вещества в сотни раз превышает интенсивность его теплового излучения в том же спектральном диапазоне, поэтому ее часто называют *холодным свечением.*

21.7.1. Виды люминесценции по способу ее возбуждения и длительности послесвечения

По способу возбуждения атомов или молекул можно выделить следующие основные виды люминесценции.

Фотолюминесценция — возбуждение происходит в результате поглощения молекулами или атомами вещества электромагнитной энергии (обычно ультрафиолетового или видимого диапазона).

Катодолюминесценция — возбуждение производится электронным ударом по атомам или молекулам вещества (наблюдается в кинескопах, электронно-лучевых трубках и т.п.).

Электролюминесценция — возбуждение атомов и молекул производится электрическим полем.

Рентгенолюминесценция — возбуждение производится рентгеновскими лучами.

Хемилюминесценция — в результате химической реакции между молекулами A и B образуется их соединение AB^* в возбужденном состоянии, при переходе из которого в основное состояние испускается квант люминесценции $h\nu$:

$$A + B \rightarrow AB^* \rightarrow AB + h\nu.$$

Биолюминесценция — возбуждение молекул происходит в результате биохимических реакций, протекающих в живом организме.

Отметим, что благодаря явлению люминесценции различные виды энергии преобразуются в электромагнитное излучение часто видимого диапазона, и это широко используется на практике (экраны кинескопов, электронно-оптические преобразователи, рентгенолюминесцентные экраны и др.).

Если возбуждение молекул или атомов происходит непрерывно, то и люминесценция тоже будет непрерывной с некоторой интенсивностью I_0, и в этом случае ее называют *стационарной люминесценцией.*

Если возбуждение молекулы или атома внезапно прекратить, то интенсивность люминесценции начнет экспоненциально затухать во времени по закону

$$I_{\text{люм}} = I_0 e^{-t/\tau}, \qquad (21.27)$$

где τ — длительность люминесценции.

Параметр τ называют также длительностью послесвечения или временем жизни возбужденного состояния, он является важнейшей характеристикой люминесцирующего вещества. Данный параметр определяет время, за которое интенсивность люминесценции уменьшается в $e = 2{,}7$ раза. Это время обычно находят из графика, полученного опытным путем (рис. 21.10).

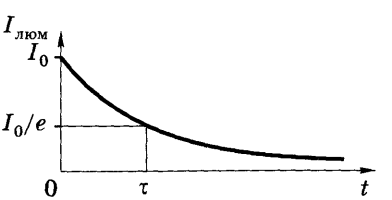


Рис. 21.10. Затухание интенсивности люминесценции во времени

По длительности послесвечения τ люминесценция делится:

- на *флуоресценцию* ($\tau < 10^{-7}$с), т.е. затухание люминесценции происходит очень быстро (для глаза — мгновенно);
- *фосфоресценцию* ($\tau > 10^{-4}$ с), в этом случае затухание идет сравнительно медленно и часто хорошо наблюдается невооруженным глазом.

Флуоресценция возникает в результате разрешенных оптических переходов с возбужденных электронных состояний в основное, вероятность этих переходов очень велика, поэтому время высвечивания мало.

Фосфоресценция происходит в результате оптических переходов, запрещенных правилами отбора квантовой механики, вероятность их очень мала (в 10^6–10^4 раз меньше, чем вероятность разрешенных переходов, т.е. флуоресценции), и соответственно время высвечивания (время жизни в возбужденном состоянии) весьма велико.

21.7.2. Основные характеристики и законы люминесценции

Спектр люминесценции — зависимость интенсивности люминесценции от длины волны или от частоты. Представляется собой спектральные линии или полосы, которые характеризуются длиной волны λ_{max} (или частотой ν_{max}), соответствующей максимуму полосы люминесценции, а также формой и шириной этой полосы $\Delta\lambda$ (или $\Delta\nu$).

Длительность люминесценции τ — это время, за которое интенсивность люминесценции уменьшается в $e = 2{,}7$ раза, т.е. время жизни молекулы в возбужденном состоянии.

Квантовый выход люминесценции γ — это отношение числа квантов люминесценции $n_{люм}$ к количеству квантов $n_{погл}$, поглощенных при возбуждении молекулы: $\gamma = n_{люм}/n_{погл}$. Квантовый выход всегда меньше единицы ($\gamma < 1$) из-за наличия неоптических переходов. Вещество считается хорошо люминесцирующим, если его квантовый выход $\gamma > 0{,}01$, т.е. если $\gamma > 1$ %.

Спектр возбуждения — это зависимость интенсивности фотолюминесценции от длины волны возбуждающего излучения (для многих молекул он совпадает с их спектром поглощения).

Степень поляризации люминесценции определяется при возбуждении вещества линейно поляризованным светом и позволяет оценить скорость вращения люминесцирующей молекулы и микровязкость ее окружения.

Каждое люминесцирующее вещество имеет свой индивидуальный спектр люминесценции, что используется при исследовании структуры веществ с помощью люминесцентного анализа. По спектру люминесценции можно определить как вид, так и концентрацию люминесцирующего вещества в смеси даже при ничтожно малых концентрациях (до 10^{-12} г/л при квантовом выходе всего около 1 %). Люминесцентный анализ чувствительнее абсорбционного (по спектрам поглощения) спектрального анализа более чем в 1000 раз.

Спектры люминесценции очень чувствительны к химическим свойствам окружения (полярности, pH, вязкости и др.) люминесцирующей молекулы, что широко используется в биохимических исследованиях. Люминесценция позволяет изучать изменения конформации макромолекул, а при использовании люминесцентного микроскопа изучать и локализацию люминесцирующих молекул внутри клетки.

Основными законами люминесценции являются:

1. *Закон Стокса*: спектр люминесценции вещества смещен в область более длинных волн относительно его спектра поглощения (см. рис. 21.9).

2. *Закон Вавилова*: квантовый выход и спектр люминесценции сложных молекул не зависят от длины волны возбуждения.

Оба эти закона объясняются наличием внутренней конверсии в сложных молекулах, в результате которой испускание всегда происходит с самых нижних подуровней возбужденных электронных состояний на все подуровни основного состояния, независимо от того, на какие колебательно-вращательные подуровни верхних электронных состояний была возбуждена молекула (см. рис. 21.8). Основное требование к спектральному диапазону волн, используемых для возбуждения люминесценции, — он должен попадать внутрь полосы поглощения этого вещества.

21.7.3. Собственная люминесценция биомолекул

При облучении различных клеток позвоночных и беспозвоночных животных, а также растений и микроорганизмов коротковолновым ультрафиолетом ($\lambda = 250–280$ нм) можно

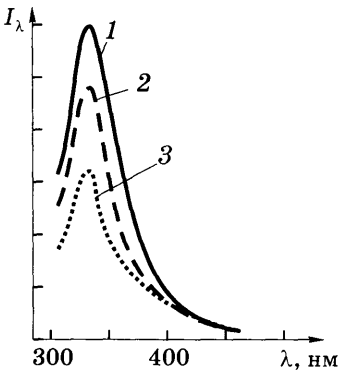


Рис. 21.11. Спектры ультрафиолетовой флуоресценции мышечных волокон (*1*), нейронов (*2*), эритроцитов (*3*)

зарегистрировать их флуоресценцию. Это свечение объясняется тем, что многие ткани организма содержат вещества (витамины, белки и др.), способные флуоресцировать под действием света. Такая флуоресценция биомолекул называется ***собственной***. Собственная ультрафиолетовая флуоресценция живых тканей определяется главным образом свечением белков [39]. Спектры собственной ультрафиолетовой флуоресценции многих клеток имеют сходную форму, хотя одни из них светятся сильнее, другие — слабее (рис. 21.11).

Вместе с тем разные ультраструктурные компоненты одной и той же клетки флуоресцируют неодинаково. Наиболее яркая ультрафиолетовая флуоресценция присуща сократительному аппарату клетки, митохондриям и ядрышкам.

В состав клеток животного происхождения входит ряд веществ, люминесцирующих как в ультрафиолетовой, так и в видимой области спектра. По положению спектров люминесценции биомолекул в шкале длин волн их можно разделить на следующие группы: нуклеиновые кислоты, белки, коферменты, продукты окисления и витамины.

Спектры поглощения *нуклеиновых кислот* определяются входящими в их состав пуриновыми и пиримидиновыми основаниями. Люминесценция нуклеиновых кислот в водных растворах очень слаба: квантовый выход люминесценции ДНК, например, составляет около 10^{-4}, а положение максимума флуоресценции 330–335 нм.

Из *аминокислот* способны люминесцировать только три ароматические аминокислоты: триптофан, тирозин и фенилаланин, максимумы флуоресценции которых в водных растворах приходятся на 348, 303 и 282 нм, а квантовые выходы равны 0,2; 0,21 и 0,04 соответственно. Поэтому собственная

люминесценция белков лежит в УФ-области спектра и полностью определяется люминесценцией этих трех аминокислот. При этом основной вклад в люминесценцию белков принадлежит триптофану, а в его отсутствие — тирозину, и только в отсутствие их обоих проявляется слабая люминесценция фенилаланина.

В клетках человека и животных содержится много *белков*, в состав которых входят триптофан, тирозин и фенилаланин. Это актин, миозин, ферменты типа дегидрогеназ, фосфатаз, оксидаз, некоторые гормоны, ферменты пищеварительной системы, альбумины и глобулины плазмы крови и другие вещества. Собственная ультрафиолетовая флуоресценция всех этих белков определяется, главным образом, триптофаном, что и определяет схожесть спектров люминесценции многих белков и клеток (см. рис. 21.11). Очень важным свойством люминесценции триптофана является смещение максимума его флуоресценции при изменении полярности и жесткости его микроокружения, что позволило разработать флуоресцентный метод анализа структурного состояния нативных триптофансодержащих белков как в растворе, так и в составе клетки. Конформационно-чувствительными оказались и другие параметры флуоресценции этих белков: квантовый выход, длительность флуоресценции и степень поляризации.

В видимой области спектра люминесцируют различные продукты окисления белков и липидов (синяя область с максимумом около 405 нм). Флуоресценция живых тканей в синей и желто-зеленой областях связана с наличием в клетках восстановленной формы пиридиннуклеотидов (НАД · Н и НАДФ · Н) и окисленной формы флавопротеинов (ФП). Эти вещества участвуют в таких процессах, как гликолиз, пентозный цикл, цикл Кребса, окисление жирных кислот и, что особенно важно, клеточное дыхание. Поэтому практически любые сдвиги в клеточном метаболизме отражаются и на динамике флуоресценции НАД · Н и ФП, что может быть выявлено при люминесцентном анализе живых тканей.

Восстановленная форма *никотинамидадениндинуклеотида* (НАД · Н) обладает характерным спектром поглощения, включающим две полосы поглощения в ультрафиолетовой

части спектра с максимумом в области 260 нм и 340 нм, а также собственной синей флуоресценцией с максимумом в интервале 465–480 нм. Переход НАД · Н в окисленное состояние сопровождается потерей одной полосы (при λ = 340 нм) в спектре поглощения и утратой способности флуоресцировать.

Пиридиннуклеотиды в восстановленной форме (НАД · Н, НАДФ · Н) флуоресцируют с максимумом 470 нм в водных растворах и 440 нм при связывании с белками, в последнем случае почти в 6 раз увеличивается и квантовый выход и интенсивность люминесценции.

Флавины — витамин B_2 (рибофлавин), ФМН (флавинмононуклеотид) и ФАД (флавинадениндинуклеотид) в окисленной форме обладают спектрами флуоресценции с максимумом в голубовато-зеленой области спектра при 530 нм (рис. 21.12).

Среди *флавопротеинов* (ФП) наибольший интерес представляют собой производные рибофлавина. В окисленном состоянии флавопротеины обладают характерным спектром поглощения и собственной желто-зеленой флуоресценцией с максимумом в области 530 нм, обусловленной рибофлавином. При восстановлении флавопротеины утрачивают собственную флуоресценцию.

Неодинаковая флуоресценция пиридиннуклеотидов и флавопротеинов в восстановленной и окисленной формах лежит в основе использования флуоресцентного анализа для ко-

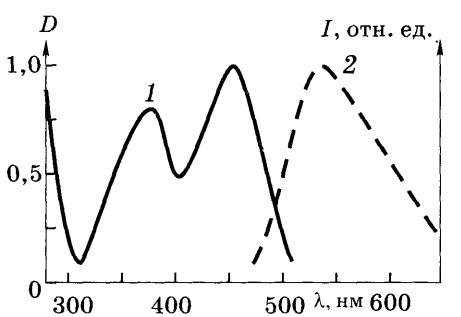


Рис. 21.12. Спектры растворов ФАД (флавинадениндинуклеотид): *1* — поглощение; *2* — флуоресценция

личественной оценки клеточного дыхания живых тканей. Усиление клеточного дыхания сопровождается изменением соотношения восстановленных и окисленных форм компонентов дыхательной цепи в сторону преобладания вторых над первыми. Угнетение дыхания приводит к противоположному эффекту. Поэтому при усиленном дыхании клетки в ней нарастает желто-зеленая и затухает синяя флуоресценция, тогда как при ослаблении дыхания характерно усиление синего свечения НАД · Н и угасание желто-зеленой флуоресценции ФП.

Билирубин, являющийся промежуточным продуктом распада гемоглобина, в водных растворах слабо флуоресцирует в голубовато-зеленой области 500–550 нм, но при взаимодействии с сывороточным альбумином человека его люминесценция возрастает почти в 50 раз, а максимум приходится на 515 нм.

Флуоресценция в красной области присуща прежде всего *порфиринам*, содержащим их цитохромам, пероксидазе, каталазе, гемоглобину, миоглобину и др. Спектры флуоресценции порфиринов обычно имеют две полосы: более интенсивную в красно-оранжевой области 600–650 нм и менее интенсивную в более длинноволновой красной области. Однако в гемопорфиринах флуоресценция «потушена» атомами железа, что ограничивает возможности ее исследования в живых тканях. При обработке ткани некоторыми веществами атом железа отрывается от простетической группы внутриклеточных гемопорфиринов и выявляется характерная красная флуоресценция порфиринов.

Протопорфирин является частью молекулы гемоглобина, поэтому обнаруживается в эритроцитах и в небольших количествах в плазме крови, благодаря чему эритроциты обычно слабо люминесцируют красным светом. При недостаточности железа, отравлении свинцом и других патологических состояниях красная люминесценция протопорфирина увеличивается в десятки и сотни раз. Этим свойством пользуются судебные медики для обнаружения отдельных эритроцитов по их порфириновой флуоресценции после обработки исследуемого объекта серной кислотой.

Значительный интерес представляют порфирины и их люминесцентные свойства в связи со способностью этих соединений избирательно накапливаться в злокачественных опухолях. Помимо диагностики, это свойство порфиринов используется сейчас в фотодинамической терапии опухолей (подробнее см. гл. 23).

21.7.4. Вторичная люминесценция, люминесцентные зонды

Есть много биологически важных соединений, которые сами не люминесцируют и поэтому использовать весьма чувствительные, неразрушающие вещество и хорошо отработанные методы люминесцентного анализа для их изучения напрямую невозможно. В настоящее время при изучении биообъектов наряду с исследованием собственной флуоресценции клеток и тканей широко используется ***метод флуоресцентных зондов.*** Суть этого метода заключается в следующем: существует ряд люминесцирующих веществ (зондов), которые избирательно взаимодействуют с определенными компонентами клетки и отдельными молекулами, причем образующийся комплекс флуоресцентного зонда и взаимодействующего с ним вещества клетки, как правило, имеет более интенсивную флуоресценцию, чем люминесцентный зонд в несвязанном состоянии. Свечение этого комплекса под действием света называют ***вторичной (наведенной)*** или ***зондовой флуоресценцией.*** На рис. 21.13 показана зависимость флуоресценции зонда (АНС) от концентрации сывороточного альбумина человека (САЧ) в водном растворе. В чистой воде (кривая 0) флуоресценция зонда очень слаба (квантовый выход 0,006), но по мере связывания с белком

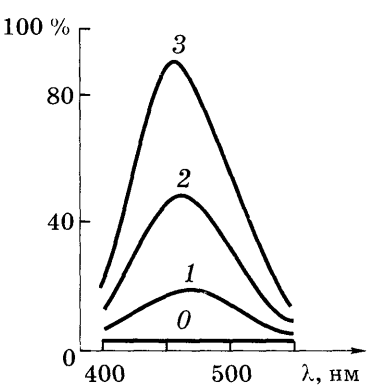


Рис. 21.13. Зависимость относительной нтенсивности люминесценции АНС от концентрации САЧ в растворе

квантовый выход возрастает до 0,3–0,8, т.е. в 50–100 раз (кривые *1–3*). Эта особенность люминесценции зонда АНС используется для экспресс-анализа степени очистки сывороточного альбумина человека при проведении гемо- и лимфосорбции.

Применение вторичной флуоресценции позволило изучить распределение ионов кальция в клетке и различных органах, а также его изменение в процессе жизнедеятельности организма. В качестве зонда на кальций наибольшее распространение нашел *хлортетрациклин*. Он образует сложный комплекс с ионами кальция, который флуоресцирует тем интенсивнее, чем больше кальция адсорбировано на биомембранах. В настоящее время известно более ста различных флуоресцентных зондов, специфически связывающихся с определенными биомолекулами и клеточными субструктурами.

Контрольные вопросы и задания

1. В чем состоит явление поглощения света и каким законам оно подчиняется? Что называется спектром поглощения вещества? В каких единицах измеряется показатель поглощения?
2. Как записывается закон поглощения света для растворов? От чего зависит и в чем измеряется удельный показатель поглощения?
3. Что такое коэффициент пропускания и оптическая плотность образца? Как они зависят от длины волны света и от концентрации растворенного поглощающего вещества?
4. В чем состоит явление рассеяния света и каким законам оно подчиняется? Чем различаются рассеяния Тиндаля и Рэлея? Сформулируйте закон Рэлея.
5. Что такое индикатриса рассеяния и что она характеризует? Укажите области применения нефелометрии.
6. В чем суть основных постулатов теории Бора? Сформулируйте их.
7. Как возникают оптические спектры поглощения и испускания? Почему они индивидуальны для каждого вещества?
8. Приведите схему энергетических уровней атома водорода и объясните формирование спектральных серий, охарактеризуйте их.

9. В чем особенности энергетической структуры молекул? Почему молекулярные спектры не линейчатые, а полосатые?
10. Приведите классификацию молекулярных спектров. В каких областях спектра они наблюдаются и какую информацию о структуре молекулы несут?
11. Что такое люминесценция? На какие виды она подразделяется по способу возбуждения и по длительности послесвечения?
12. Каковы основные характеристики люминесценции, ее законы?
13. В чем заключается люминесцентный анализ и каковы его преимущества по сравнению с абсорбционным спектральным анализом?
14. Какие биомолекулы способны люминесцировать и в каких областях спектра?
15. Что такое люминесцентные зонды и вторичная люминесценция? Как люминесцентный анализ используется в медицинских исследованиях?

Глава 22 Вынужденное излучение. Лазеры

В предыдущей главе были рассмотрены процессы поглощения и самопроизвольного (спонтанного) испускания электромагнитной энергии атомами и молекулами. Вместе с тем А. Эйнштейн еще в 1917 г. указал на возможность так называемого «вынужденного» испускания электромагнитного излучения квантовыми системами.

22.1. Вынужденное излучение и его свойства

Вынужденное излучение рассмотрим подробнее на примере простейшей двухуровневой квантовой системы (рис. 22.1). Нижний уровень E_1 — основной и в этом состоянии находится большинство атомов (или молекул) вещества. Верхний уровень E_2 — возбужденный и в обычном состоянии не заселен.

Облучение такой системы резонансным излучением, т.е. излучением с энергией кванта $h\nu = E_2 - E_1$, равной энергии перехода между уровнями, будет сопровождаться поглощением этих квантов и переходами системы из основного состояния E_1 в возбужденное состояние E_2 (эти переходы на рис. 22.1 показаны черными вертикальными стрелками). Вследствие этого процесса населенность n_2 верхнего уровня будет возрастать, а населенность n_1 нижнего уровня — умень-

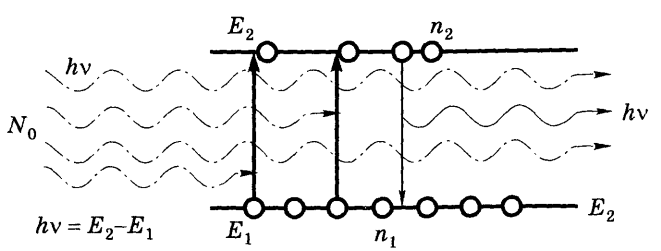


Рис. 22.1. Вынужденные переходы в двухуровневой квантовой системе под воздействием резонансного излучения

шаться. Количество квантов $N_{погл}$, поглощенных системой за единицу времени, пропорционально числу квантов N_0, падающих на квантовую систему в единицу времени, а также населенности n_1 основного состояния: $N_{погл} = BN_0n_1$, где B — коэффициент Эйнштейна.

Квантовая система в состоянии E_2 не равновесна и стремится вернуться в основное состояние E_1. Существуют три возможности такого перехода:

1) *безызлучательный переход*, при котором энергия $\Delta E = E_2 - E_1$ передается в виде кинетической энергии окружающим атомам или молекулами и превращается в тепло; вероятность этого процесса тем меньше, чем больше энергия ΔE и меньше плотность среды;

2) *спонтанное излучение* (*люминесценция*), при котором излучается квант электромагнитной энергии резонансной частоты $\nu = \Delta E/h$, однако направление этого излучения, его поляризация и начальная фаза спонтанны, т.е. произвольны. Вероятность такого перехода тем выше, чем больше энергия ΔE, и определяется только свойствами излучающих атомов и молекул среды;

3) *вынужденное излучение*, возникающее вследствие стимулированного внешним резонансным электромагнитным излучением перехода квантовой системы из возбужденного состояния в основное. Оно сопровождается излучением кванта $h\nu$ резонансного излучения, которое по всем своим характеристикам (частоте, направлению излучения, поляризации и фазе) не отличается от стимулирующего излучения. Вероятность вынужденного излучения пропорциональна интенсивности падающего излучения, т.е. числу квантов N_0, падающих на систему в единицу времени, а также населенности n_2 верхнего уровня: $N_{вын} = BN_0n_2$.

Спонтанное излучение совпадает с падающим лишь по частоте, но отличается от него по направлению, поляризации, фазе (не когерентно с ним). Напротив, вынужденное излучение по всем характеристикам совпадает с падающим (стимулирующим) излучением, поэтому вынужденно испущенный квант невозможно отличить от кванта вынуждающего излучения (табл. 22.1).

Таблица 22.1

Характеристика излучения	Излучение	
	спонтанное	вынужденное
Частота, ν	$\nu = (E_2 - E_1)/h$	$\nu = (E_2 - E_1)/h$
Направление излучения	Произвольное	То же, что и у стимулирующего
Поляризация	Произвольная	То же
Фаза излучения	Произвольная (не когерентно с падающим)	Та же, что и у стимулирующего (когерентно с ним)

22.2. Условия усиления света. Лазеры

Гипотеза А. Эйнштейна о вынужденном излучении долгое время оставалась не доказанной экспериментально, но в 1939 г. советский физик В.А. Фабрикант указал, что за счет вынужденного излучения при определенных условиях можно получить усиление электромагнитного излучения.

Действительно, если на квантовую систему в единицу времени падает N_0 резонансных квантов ($h\nu = E_2 - E_1$), то часть их поглощается $N_{\text{погл}} = BN_0n_1$, а часть таких же квантов вынужденно испускается $N_{\text{вын}} = BN_0n_2$ и количество квантов, направленно прошедших за единицу времени сквозь такую среду, определяется следующим образом:

$$N_{\text{прош}} = N_0 - N_{\text{погл}} + N_{\text{вын}} = N_0 - BN_0n_1 + BN_0n_2 =$$

$$= N_0 + (n_2 - n_1)\,BN_0. \qquad (22.1)$$

Видно, что если $n_2 > n_1$, то и $N_{\text{прош}} > N_0$, т.е. получим усиление падающего излучения, поскольку количество квантов, вышедших из такой среды, превысит количество падающих.

В обычных условиях населенность основного состояния намного больше, чем возбужденного ($n_1 > n_2$), поэтому обычно наблюдается лишь поглощение света.

Если создать в среде инверсную населенность $n_2 > n_1$, то такая среда усиливает проходящее через нее резонансное излучение и называется *активной* или *усиливающей средой*. Впервые создать усиливающую среду удалось в 1954 г. советским физикам Н.Г. Басову и А.М. Прохорову и независимо от них Ч. Таунсу (США), за что все трое впоследствии были удостоены Нобелевской премии по физике.

Создать инверсную населенность за счет поглощения падающего излучения (оптической накачки) в двухуровневой квантовой системе невозможно, так как по достижении равенства населенностей обоих уровней $n_2 = n_1$ наступит динамическое равновесие: количество поглощенных квантов будет равно количеству вынужденно испущенных квантов за то же время. Поэтому создать инверсную населенность при оптической накачке можно только в трех- или четырехуровневой квантовой системе.

Примером трехуровневой активной среды является рубин (рис. 22.2), представляющий собой кристалл корунда Al_2O_3, в котором небольшая (около 0,05 %) часть ионов алюминия Al^{3+} заменена ионами хрома Cr^{3+}, придающими рубину красную окраску и образующими активную среду. Ионы хрома хорошо поглощают оптическое излучение частотой ν_{13}, переходя из основного 1-го состояния на свободный 3-й уровень энергии E_3, с которого затем очень быстро (за 10^{-12} с) безызлучательно переходят на метастабильный уровень E_2. Вероятность спонтанных переходов с этого уровня в основное состояние E_1 очень мала, что способствует накоплению на нем большого количества частиц и созданию инверсной населенности ($n_2 > n_1$) в канале 2–1. Для этого на

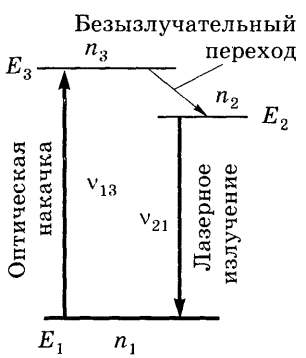


Рис. 22.2. Схема трехуровневой активной среды

уровень E_2 необходимо перевести более половины первоначального количества ионов хрома. В этом состоянии среда способна усиливать излучение частотой ν_{21}, которая для рубина соответствует длине волны $\lambda = 694$ нм.

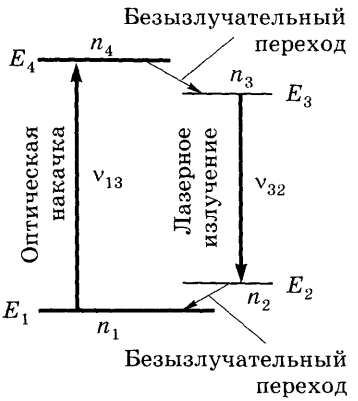


Рис. 22.3. Схема четырехуровневой активной среды

Значительно меньшие энергетические затраты необходимы для создания инверсной населенности в четырехуровневой активной среде, примером которой являются ионы неодима Nd^{3+} (рис. 22.3). В такой среде поглощение излучения накачки осуществляется в канале 1–4, затем происходит быстрый безызлучательный переход с 4-го на 3-й уровень, а поскольку 2-й уровень практически всегда пустой (благодаря быстрым безызлучательным переходам с него на 1-й уровень), то $n_3 > n_2$ и в канале 3–2 практически сразу возникает инверсная заселенность, среда способна усиливать излучение частотой ν_{32}, что для неодима соответствует $\lambda = 1060$ нм (ближний ИК-диапазон).

Таким образом, активная среда способна усиливать излучение с определенной (резонансной) длиной волны. Однако чтобы создать оптический квантовый генератор, необходимо иметь источник такого резонансного излучения и объединить его с усиливающей средой.

Первый квантовый генератор видимого диапазона был создан на рубине в 1960 г. в США Т. Мейманом и назван ***лазером*** (от англ. light amplification by stimulated emission of radiation (laser) — усиление света за счет вынужденного излучения). С тех пор были разработаны сотни различных лазеров (в советской физической литературе их называли оптическими квантовыми генераторами (ОКГ), которые широко применяются в настоящее время в технике, военном деле и в медицине.

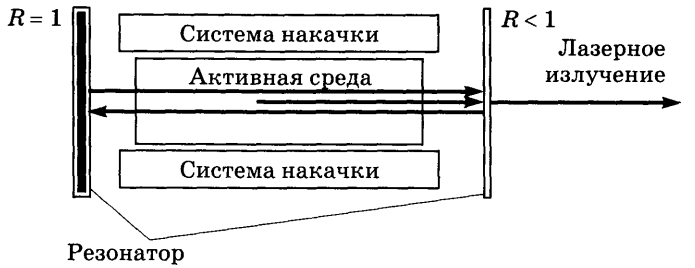


Рис. 22.4. Принципиальная схема лазера

Любой лазер (ОКГ) состоит из трех обязательных *компонент* (рис. 22.4):

- активная среда, которая при определенных условиях может быть переведена в усиливающее состояние;
- система накачки, обеспечивающая инверсную населенность между рабочими уровнями энергии активной среды, переводя ее тем самым в усиливающее состояние;
- резонатор — специальная система зеркал, формирующая обратную оптическую связь, необходимую для генерации излучения.

Принцип формирования лазерного излучения состоит в следующем: система накачки переводит активную среду в усиливающее состояние, в котором она способна усиливать излучение резонансной частоты, соответствующей переходу между уровнями с инверсной заселенностью. Первоначальным, или как говорят «затравочным», излучением нужной частоты становится спонтанное излучение (люминесценция), возникающее в результате самопроизвольного перехода системы из возбужденного в основное состояние. Из всего набора хаотично направленных спонтанных излучений наибольшее усиление испытывает излучение, распространяющееся перпендикулярно зеркалам резонатора, поскольку благодаря отражениям от зеркал оно многократно проходит через активную среду и соответственно многократно усиливается.

Резонатор обычно состоит из двух параллельных друг другу зеркал, из которых одно отражает резонансное излучение полностью (коэффициент отражения $R = 1$), а другое — час-

тично ($R < 1$). Именно излучение, выходящее из резонатора через это полупрозрачное зеркало, и представляет собой лазерное излучение.

Основные *свойства* лазерного излучения:

- высокая направленность — определяется свойствами резонатора, так как максимально усиливается лишь излучение, распространяющееся перпендикулярно его зеркалам, вследствие чего формируется узконаправленный параллельный лазерный пучок;
- высокая монохроматичность — определяется прежде всего длиной волны и шириной линии люминесценции (лазерного перехода) активной среды, а в пределах ширины этой линии — свойствами резонатора;
- высокая когерентность — является прямым следствием когерентных свойств вынужденного излучения;
- высокая спектральная плотность мощности — отношение интенсивности лазерного излучения к ширине спектральной линии излучения. По данному параметру лазерное излучение превосходит все другие известные виды излучений.

Длина волны генерируемого лазером излучения определяется прежде всего *свойствами активной среды*, а в пределах ее линии люминесценции — *свойствами резонатора*.

В зависимости от используемой активной среды современные лазеры генерируют излучение в различных спектральных диапазонах: от инфракрасного до ультрафиолетового.

В зависимости от *агрегатного состояния активной среды* лазеры бывают газовые, твердотельные, полупроводниковые, жидкостные.

По виду *используемого способа создания инверсной населенности* (системы накачки) лазеры бывают с оптической накачкой (твердотельные и жидкостные), накачкой электронным ударом (газовые), химической накачкой (эксимерные), накачкой электрическим током (полупроводниковые), лазерной накачкой (некоторые жидкостные лазеры).

По *временным параметрам генерируемого излучения* лазеры подразделяются на непрерывные, импульсные и частотные, с частотой следования импульсов излучения от единиц до тысяч герц.

Непрерывные лазеры в зависимости от вида и назначения имеют мощность излучения от единиц милливатт до нескольких киловат.

Импульсные лазеры характеризуются энергией излучения в импульсе, длительностью импульса излучения от микросекунд (10^{-6} с) до фемтосекунд (10^{-15} с) и пиковой мощностью от нескольких мегаватт (10^{6}) до десятков и сотен тераватт (10^{12}). Они могут работать как в режиме одиночных импульсов, так и в частотном режиме.

22.3. Применение лазерного излучения в медицине

Уникальные свойства лазерного излучения — монохроматичность, когерентность, высокая степень поляризации, малая расходимость, широкий диапазон временных параметров и соответственно большой диапазон мощности излучения — обеспечили широкое применение лазеров в различных областях медицины.

При этом следует учитывать, что глубина проникновения излучения в биоткани существенно зависит от длины волны излучения: лучше всего проникает излучение красного и ближнего инфракрасного (ИК) диапазонов (630–2700 нм), поскольку оно меньше поглощается и рассеивается биотканями. Излучение зеленого (630–650 нм) и фиолетового ($\lambda < 405$ нм) диапазонов спектра хорошо поглощается кровью, а в ультрафиолетовом диапазоне (200–340 нм) лежат полосы поглощения многих белков. В среднем и дальнем ИК-диапазонах (2,7–3,6 мкм, 5,6–6,7 мкм и 10–20 мкм) лежат полосы поглощения воды и соответственно мягких биотканей.

Установлено, что низкоинтенсивное (~10 мВт) красное излучение гелий-неонового лазера ($\lambda = 633$ нм) обладает биостимулирующими свойствами и широко используется для лечения многих патологий: дерматозов, трофических язв, гнойных ран, ожогов и др. В настоящее время для этих целей используют и низкоинтенсивное излучение полупроводниковых лазеров красного и ближнего ИК-диапазонов. Используют также воздействие такого излучения на точки акупунктуры.

Импульсное излучение твердотельных лазеров на рубине ($\lambda = 694$ нм) широко используется в офтальмологии при проведении бескровных хирургических операций на сетчатке глаза (приваривание отслоившейся сетчатки, коагуляция лопнувших сосудов и т.п.), а также в микрохирургии переднего отдела глаза.

Лазеры ультрафиолетового диапазона, излучение которых хорошо поглощается роговицей, применяются в операциях по коррекции кривизны роговицы (для устранения близорукости и сильной дальнозоркости) путем испарения части ее вещества сфокусированным лазерным излучением.

В хирургии мягких тканей в основном используется инфракрасное непрерывное излучение газовых CO_2-лазеров средней мощности (10–30 Вт), которые позволяют проводить как рассечение тканей на нужную глубину, так и их коагуляцию с возможностью обеспечения биологической сварки тканей, в том числе сварки стенок сосудов и т.п.

Лазерное излучение средней мощности непосредственно используется и для удаления (термического разрушения) опухолей, особенно поверхностной локализации. В последние годы в онкологии используют лазерное излучение небольшой интенсивности (~1 Вт) в сочетании с фотосенсибилизаторами для лечения опухолей методами фотодинамической терапии.

Лазерное излучение используют также для внутриполостных воздействий на стенки и слизистую полых органов, сосудов, применяя эндоскопическую технику.

В косметологии также используют лазеры разных спектральных диапазонов и мощностей в зависимости от объекта и цели воздействия.

В современной медицине области применения лазерное излучения постоянно расширяются.

Контрольные вопросы и задания

1. Какие переходы могут происходить в двухуровневой квантовой системе под воздействием резонансного излучения?
2. От чего зависит количество поглощаемых системой в единицу времени квантов падающего излучения?

3. Какие переходы называются спонтанными? Каковы их основные характеристики? От чего зависит вероятность спонтанных переходов?
4. Какое излучение называется вынужденным? Каковы его основные свойства? От чего зависит вероятность вынужденных переходов?
5. При каких условиях возможно усиление электромагнитного излучения?
6. Что такое лазер? Каковы основные компоненты лазерных систем?
7. Каковы основные свойства лазерного излучения, чем они определяются?
8. Приведите классификацию лазеров по состоянию активной среды, режиму и способу накачки, длительности и мощности излучения.
9. Укажите области применения лазеров в медицине.

Глава 23 Элементы фотобиологии

Фотобиологическими называются процессы, которые начинаются с поглощения света биологическими системами, а заканчиваются определенной физиологической реакцией организма.

Фотобиология исследует воздействие света на организмы и его роль в жизни растений, животных и человека, изучает фотосинтез, фоторецепцию, в том числе механизмы зрения, реакцию организмов на периодическую смену дня и ночи, проявляющуюся в колебаниях интенсивности физиологических процессов. Дозированное воздействие на организм инфракрасным, видимым или ультрафиолетовым излучением с лечебными целями уже давно применяется в медицине, но в последние годы получили развитие методы фотодинамической терапии опухолей и других заболеваний, основанные на использовании фотосенсибилизаторов и новых источников, в том числе лазерных, оптического излучения.

23.1. Фотобиологические процессы

По воздействию на организм фотобиологические процессы можно разделить на позитивные (полезные) и негативные (вредные) для человека.

К ***позитивным*** фотобиологическим процессам у человека и животных относятся: зрение; фотопериодизм — регуляция суточных и годовых циклов жизни животных путем циклического воздействия «темнота–свет»; образование витамина *D* из его провитаминов.

К ***негативным*** фотобиологическим процессам относят:

- фототоксические эффекты (световые повреждения глаз и кожи, которые проявляются в форме помутнения хрусталика, эритемы, пигментации кожи); бактерицидное действие (имеет и положительное значение); канцерогенное действие ультрафиолетового излучения;

• фотоаллергические эффекты — увеличение чувствительности организма к некоторым аллергенам под действием света.

Несмотря на большое разнообразие фотобиологических процессов, в любом из них всегда можно выделить следующие *первичные стадии*, благодаря которым формируется отклик организма на световое воздействие:

1) поглощение кванта света отдельной молекулой организма;

2) преобразование полученной энергии внутри этой же молекулы (переход ее в возбужденное состояние и др.);

3) передача (миграция) всей поглощенной энергии или ее части от одной молекулы к другой;

4) образование первичных фотохимических продуктов — ионов, радикалов или ион-радикалов;

5) темновые превращения первичных фотохимических продуктов, заканчивающиеся образованием новых стабильных продуктов. Световое воздействие на этой и последующих стадиях уже не требуется;

6) биохимические реакции с участием образовавшихся фотопродуктов (например, инактивация ферментов и белков, окисление липидов и т.п.);

7) общефизиологический ответ на действие света (зрение, эритема, ожог, пигментация и др.).

Одной из важнейших характеристик любого фотобиологического процесса является его ***спектр фотобиологического действия*** — зависимость фотобиологического эффекта от длины волны действующего света.

Знание спектра действия для какого-либо фотобиологического процесса позволяет, в зависимости от стоящей задачи, выбрать область спектра, наиболее эффективно вызывающую этот процесс либо, наоборот, не вызывающую его, если он нежелателен.

В качестве примера рассмотрим порядок определения спектра возникновения ультрафиолетовой эритемы.

Сначала для каждой длины волны была определена минимальная эритемная доза (МЭД), т.е. доза облучения D (Дж) = $= I$ (Вт) $\cdot$ t (с), вызывающая минимально обнаруживаемую

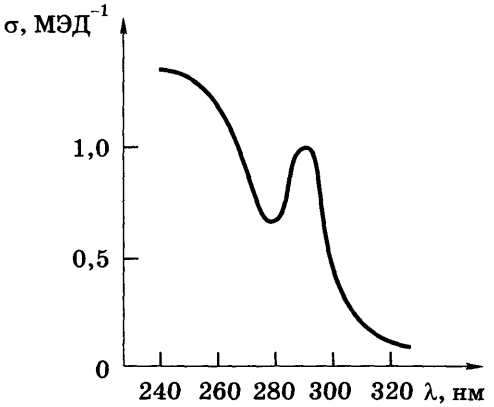


Рис. 23.1. Спектр фотобиологического возникновения УФ-эритемы

эритему (покраснение). Чем больше эта доза, тем меньше эффективность действия излучения $\sigma = 1/$ МЭД. Поэтому спектр действия в данном случае представляет собой зависимость σ от длины волны света (рис. 23.1).

С увеличением дозы эффективность фотобиологических процессов возрастает.

Очень важным *законом*, широко используемым в приложениях, является следующий: форма спектра фотобиологического действия соответствует спектру поглощения $k(\lambda)$ вещества, ответственного за рассматриваемое фотобиологическое действие.

Этот закон позволяет путем сравнения спектров фотобиологического действия со спектрами поглощения отдельных веществ, входящих в биосистему, выяснить физико-химические механизмы возникновения интересующего фотобиологического процесса.

Например, было установлено, что кривая гибели бактерий под действием света УФ-диапазона имеет максимум в области 265 нм, а ее форма близка к спектру поглощения нуклеиновых кислот. Это и позволило сделать вывод, что гибель бактерий под действием УФ-излучения обусловлена именно повреждением их ДНК.

Подобным путем удается выяснить природу акцепторов квантов света в биосистеме, а затем делать предположения и выводы относительно фотохимических реакций, протекающих в организме.

23.2. Типы фотохимических реакций

Для фотобиологии в основном типичны фотохимические реакции с участием молекул, находящихся в возбужденном электронно-колебательном состоянии, синглетном либо триплетном. Состояние называется ***синглетным***, если суммарный спин электронов на внешней оболочке атома или молекулы равен нулю, и ***триплетным*** — если суммарный спин равен единице.

Известны следующие основные типы одноквантовых фотохимических реакций органических молекул.

Фотораспад — под действием света происходит разрыв отдельных химических связей и распад молекулы на радикалы, ионы или нейтральные молекулы. Фотораспад наблюдается, например, при облучении аминокислот, белков и нуклеиновых кислот большими дозами ультрафиолета.

Фотоперегруппировки, в ходе которых один изомер или таутомер молекулы превращается в другой. (Фототаутомеризация — перенос атома водорода из одной части молекулы в другую.) Например, именно фотохимическая реакция цис-транс-изомеризации ретиналя лежит в основе механизма зрения практически всех живых организмов, в том числе и человека.

Фотоприсоединение — присоединение к возбужденной светом молекуле других молекул. В зависимости от того, какая молекула присоединяется, различают:

- фотодимеризацию: $A^* + A \rightarrow (A^* \ldots A) \rightarrow A_2$;
- фотооксидирование: $A^* + O_2 \rightarrow (A^* \ldots O_2) \rightarrow OAO$;
- фотогидратацию: $A^* + H_2O \rightarrow HAOH$.

Такие димеры, или комплексы, обычно неустойчивы и могут самопроизвольно распадаться с выделением энергии

в виде кванта света, что широко используется для создания усиливающей среды в эксимерных лазерах:

$$A^* + A \rightarrow (A^*...A) \rightarrow A + A + h\nu.$$

Фотоперенос электронов, при котором возбужденная молекула A^* отдает свой электрон другой, невозбужденной молекуле B:

$$A^* + B \rightarrow A^+ + B^- \text{ — фотоокисление,}$$

или $A^* + C \rightarrow A^- + C^+$ — фотовосстановление (при получении лишнего электрона).

Фотоперенос протона — присоединение к возбужденной молекуле протона от невозбужденной кислоты или, наоборот, отдача возбужденной кислотой своего протона основанию:

$A^* + BH \rightarrow AH^+ + B^-$ — фотоприсоединение H^+;

$AH^* + B \rightarrow A^- + HB^+$ — фотоотдача H^+.

В обоих случаях фотореакция обусловлена изменением протонно-донорных или протонно-акцепторных свойств возбужденных молекул.

Промежуточным этапом многих фотохимических реакций является образование свободного радикала, т.е. молекулы, несущей один или более неспаренных (химически активных) электронов на внешней орбите.

23.3. Фотосенсибилизированные фотобиологические процессы

Существует ряд веществ, наличие которых в организме повышает его чувствительность к ультрафиолетовому и видимому свету. В фотобиологии такие вещества называют ***фотосенсибилизаторами***, а в медицине их часто относят к ***фототоксическим*** или ***фотоаллергическим веществам***, в зависимости от того, какие процессы в организме они вызывают.

По механизму действия фотосенсибилизаторы делятся на два типа.

Фотосенсибилизаторы 1-го типа сами претерпевают химические превращения под действием света. В медицине вещества такого типа используются, в частности, для лечения

псориаза, в связи с чем они получили наименование *псоралленов*. В основе их терапевтического действия лежит реакция присоединения их к ДНК клеток под действием УФ-излучения.

Другим примером фотосенсибилизаторов 1-го типа являются *фотоаллергены*. При УФ-облучении эти вещества способны химически присоединяться к белкам, в результате чего образуется аллерген, вызывающий затем аллергические реакции организма в виде эритемы, экземы, отека кожи, а в некоторых случаях и астмы.

Фотосенсибилизаторы 2-го типа, поглотив квант света, переходят в метастабильное возбужденное состояние, в котором могут находиться сравнительно долго. При столкновении с молекулярным кислородом O_2 эта возбужденная молекула отдает ему свой избыточный запас энергии и возвращается в исходное состояние. Поглотивший энергию кислород O_2 переходит в возбужденное синглетное состояние, в котором он примерно в 100 раз более эффективно, чем невозбужденный кислород, окисляет липиды, белки и другие биомолекулы, оказывая тем самым биологическое воздействие на организм. В связи с этим эффективность действия фотосенсибилизаторов 2-го типа полностью зависит от присутствия кислорода.

Типичным представителем сенсибилизаторов 2-го типа является *протопорфирин*. При некоторых заболеваниях он накапливается в организме в аномально высоких концентрациях. Протопорфирин хорошо поглощает синий свет, после чего эффективно возбуждает молекулярный кислород, последний окисляет ненасыщенные липиды, вызывая резкое увеличение проницаемости биомембран. Одновременно инактивируются многие белки, ферменты, полимеризуется мембранный белок спектрин и происходят другие явления. У больных, находящихся на дневном свету, развивается эритема и отек, поэтому таким больным приходится находиться на желтом свету, который не поглощается протопорфирином. В последнее время для лечения таких заболеваний предложен терапевтический метод, основанный на приеме внутрь бета-каротина, способного эффективно тушить (быстро переводить в основное состояние) возбужденный кислород.

23.4. Основы фотодинамической терапии

В последние годы в медицине активно развивается новое направление — **фотодинамическая терапия** (ФДТ), основанная на использовании фотосенсибилизаторов в диагностике и лечении заболеваний.

Метод фотодинамической терапии основан на введении в организм пациента специальных фотосенсибилизаторов, которые затем возбуждаются световым излучением определенного спектрального диапазона и стимулируют образование синглетного кислорода, вызывающего гибель клеток облученных тканей.

Этот щадящий для организма метод хорошо зарекомендовал себя в лечении злокачественных новообразований, очагов воспалений и других патологий. На сегодняшний день основная область применения ФДТ — онкология, но есть и другие важные области ее применения: лечение длительно незаживающих гнойных ран и трофических язв; антибактериальная терапия — при фотодинамическом воздействии погибают даже штаммы бактерий, устойчивые к антибиотикам и др. ФДТ опухолей показана и в тех ситуациях, когда хирургическая операция невозможна из-за тяжелых сопутствующих заболеваний.

При лечении и диагностике опухолевых заболеваний методом ФДТ в организм больного вводится специальный фотосенсибилизатор (ФС), который накапливается в больших количествах в патологических тканях. Освещение такой ткани излучением определенной длины волны вызывает люминесценцию ФС и показывает тем самым области патологических изменений, что и используют для целей *диагностики*. С *лечебной* целью пораженные участки ткани, насыщенные фотосенсибилизатором, облучают видимым излучением средней мощности и определенного спектрального диапазона (должен попадать в полосу поглощения фотосенсибилизатора), обычно используют лазер мощностью 1–2 Вт. Облученный ФС переходит в возбужденное активное состояние и стимулирует переход обычного (триплетного) кислорода в синглетный

кислород, который является мощным окислителем и губительно действует на окружающие ткани. В результате патологические клетки, накопившие ФС, погибают. Отметим, что для успешной реализации этого метода необходимо наличие кислорода в облучаемых тканях.

Метод ФДТ позволяет избежать общего токсичного воздействия на организм, например, как при химиотерапии опухолей, и других побочных эффектов.

Внедрение метода ФДТ в медицинскую практику сравнивают по значимости с внедрением антибиотиков в медицину. К его преимуществам относят одновременность проведения диагностического и лечебного воздействия; избирательность (поражается только опухолевая ткань); возможность удаления опухолей в труднодоступных зонах, используя световоды; возможность многократного повторения лечебного процесса; отсутствие тяжелых местных и системных осложнений и др.

Основное ограничение метода ФДТ — небольшая глубина проникновения лазерного излучения в ткани, которая составляет лишь несколько миллиметров. Используемые сейчас в клинике фотосенсибилизаторы (Фотофрин, Фотогем, Фотосенс и др.) имеют спектр фотодинамического воздействия в области 620–690 нм, тогда как максимальная проницаемость мягких тканей приходится на дальнюю красную и ближнюю ИК-области (750–1500 нм).

В последнее время большой интерес в качестве фотосенсибилизаторов в ФДТ вызывают производные ряда хлорина, которые имеют полосу поглощения в длинноволновой области видимого спектра и высокую скорость выведения из организма (2 суток), что может радикально изменить технологию ФДТ, сократив длительность пребывания в стационаре и даже заменив его амбулаторным лечением.

В настоящее время для проведения ФДТ внутриполостных, внутритканевых и поверхностных образований используются промышленная лазерная установка оптической мощностью до 2,5 Вт, которая комплектуется набором волоконно-оптических катетеров для различных локализаций, а также

видеофлуоресцентное устройство, предназначенное для проведения как диагностики, так и фотодинамической терапии с использованием излучения мощных полупроводниковых светодиодов.

Контрольные вопросы и задания

1. Какие процессы называются фотобиологическими? Как они подразделяются по воздействию на живые организмы? Каковы их первичные стадии?
2. Что представляет собой спектр фотобиологического действия и какую информацию он содержит?
3. Укажите основные типы фотохимических реакций.
4. Какие вещества называют фотосенсибизаторами? Укажите их типы и механизмы действия на организм.
5. Каковы принципы проведения фотодинамической терапии, области ее медицинского применения?
6. Укажите основные достоинства и ограничения метода фотодинамической терапии.

Глава 24 Биофизические основы зрения

Общую проблему зрения можно разделить на три более простые:

1. Как формируется изображение внешнего мира на сетчатке глаза?

2. Каким образом энергия света, поглощенного зрительным пигментом, превращается в фоторецепторный сигнал и нервный импульс?

3. Как перерабатывается полученная информация в зрительных центрах мозга и как возникает субъективный образ внешнего мира?

Решение третьего вопроса является предметом нейрофизиологии.

24.1. Строение глаза

Глазное яблоко имеет почти шаровидную форму диаметром около 24 мм и окружено тремя оболочками (рис. 24.1). ***Склера*** — наружная непрозрачная белковая оболочка — имеет защитное значение, придает глазу форму и в его передней части переходит в прозрачную выпуклую ***роговицу***, через которую в глаз проходят лучи света. Через заднюю стенку склеры проходят кровеносные сосуды и глазной нерв.

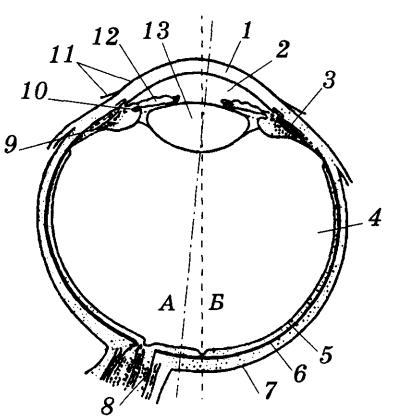


Рис. 24.1. Глаз человека: *1* — роговица; *2* — передняя камера; *3* — цилиарная мышца; *4* — стекловидное тело; *5* — сетчатка; *6* — сосудистая оболочка; *7* — склера; *8* — зрительный нерв; *9* — цилиарное тело; *10* — задняя камера; *11* — конъюнктива; *12* — радужная оболочка; *13* — хрусталик; *А* — главная оптическая ось глаза; *Б* — зрительная ось глаза

К склере прилегает ***сосудистая оболочка***, которая состоит из сети мелких кровеносных сосудов, питающих внутренние структуры глаза. Ее внутренняя поверхность выстлана слоем пигментных клеток, поглощающих свет и препятствующих рассеянию света в глазу (тем самым улучшается острота зрения). В передней части глаза сосудистая оболочка переходит в ***радужную***, в которой имеется круглое отверстие — ***зрачок***. От пигментации радужной оболочки, содержащей пигментные клетки, зависит цвет глаз.

Зрачок играет роль глазной диафрагмы, его расширение или сужение регулирует количество света, падающего на сетчатку глаза. Непосредственно за зрачком располагается прозрачный ***хрусталик***, имеющий форму двояковыпуклой линзы. Он состоит из нескольких пластинчатых слоев, отличающихся друг от друга по радиусу кривизны и показателю преломления (среднее значение 1,44). Хрусталик эластичен, его кривизна может изменяться усилием специальной (цилиарной) мышцы, благодаря чему обеспечивается аккомодация глаза — его способность создавать на сетчатке четкое изображение предметов, находящихся на различном расстоянии от глаза.

Между роговицей и хрусталиком расположена ***передняя камера*** глаза, заполненная водянистой влагой — жидкостью, близкой по оптическим свойствам к воде. Пространство позади хрусталика (задняя камера) заполнено прозрачной желеобразной массой — ***стекловидным телом***.

Самая внутренняя оболочка глаза — ***сетчатка***, или ретина, состоит из светочувствительных клеток и радиально расходящихся разветвлений зрительного нерва. Она преобразует световые сигналы в нервные импульсы и производит, таким образом, первичную обработку светового изображения, преобразуя его в поток нервных импульсов, поступающих в зрительные центры мозга.

Место вхождения зрительного нерва в глазное яблоко называется ***слепым пятном***, так как здесь нет светочувствительных клеток. Недалеко от слепого пятна находится ***желтое пятно***, выстланное очень плотно упакованными колбочками. В его центральной части имеется углубление,

называемое ***центральной ямкой***. Желтое пятно, и в особенности центральная ямка, являются наиболее чувствительными местами сетчатки (при дневном зрении), в них достигается наибольшая острота зрения и лучше всего различаются цвета. Поэтому при рассмотрении предмета человек непроизвольно старается расположить глаз таким образом, чтобы изображение рассматриваемого объекта попадало в область желтого пятна, поскольку в этом случае детали предмета различаются наиболее отчетливо.

24.2. Оптическая система глаза

Для того чтобы глаз хорошо видел рассматриваемый предмет, на сетчатке должно быть построено его четкое действительное оптическое изображение. Эту задачу решает оптическая система глаза. Ее образуют в основном роговица и хрусталик, в незначительной степени — передняя камера глаза.

Оптические свойства глазных сред приведены в табл. 24.1.

Таблица 24.1

Структурный элемент глаза	Показатель преломления	Радиус кривизны		Оптическая сила, дптр
		передней поверхности	задней поверхности	
Роговица	1,376	7,7	6,8	40–43
Влага передней камеры	1,333	6,8	~10	4
Хрусталик неаккомодированный	1,414	10,0	6,0	20
Хрусталик максимально аккомодированный	1,424	5,3	5,3	33
Стекловидное тело	1,336	~6	–	–

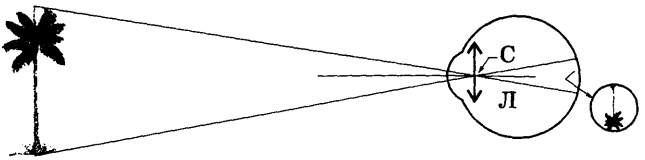


Рис. 24.2. Схема построения изображения дерева на сетчатке глаза: С — оптический центр хрусталика; Л – «приведенный глаз»

Прямая линия, проходящая через центры кривизны всех четырех сферических поверхностей, образующих роговицу и хрусталик, называется ***главной оптической осью*** глаза, а прямая, проходящая через центр хрусталика и центральную ямку на сетчатке, — ***зрительной осью*** глаза (см. рис. 24.1). Угол между этими осями составляет примерно 5°, при этом зрительная ось впереди глаза отклонена назально (к носу).

Суммарная оптическая сила глаза может изменяться в пределах 60–73 дптр. Наибольший вклад принадлежит роговице — приблизительно 40 дптр. Оптическая сила хрусталика может изменяться от 20 до 33 дптр, в зависимости от усилия, создаваемого цилиарной мышцей. При расслабленной мышце преломляющая сила хрусталика около 20 дптр, а при ее максимальном напряжении — 33 дптр.

Сложную оптическую систему глаза удобно заменить одной тонкой собирающей линзой — ***приведенным глазом***, расположенным примерно в центре хрусталика и имеющим переменную оптическую силу в 60–73 дптр.

На сетчатке глаза всегда создается действительное, уменьшенное и перевернутое изображение рассматриваемых предметов (рис. 24.2).

24.2.1. Аккомодация глаза

Способность глаза четко видеть предметы, находящиеся от него на различных расстояниях, называется ***аккомодацией***. Изображение предмета на сетчатке будет четким только в том случае, если расстояние d от предмета до центра хрусталика,

расстояние f от центра хрусталика до сетчатки и фокусное расстояние F глаза удовлетворяют формуле тонкой линзы

$$\frac{1}{d}+\frac{1}{f}=\frac{1}{F}. \qquad (24.1)$$

Из этой формулы следует, что при изменении расстояния d до предмета должно соответственно измениться либо расстояние до изображения f, либо фокусное расстояние линзы F. Первый случай реализуется в фотоаппаратах, где F объектива неизменно. Для глаза неизменным является расстояние от центра хрусталика до сетчатки $f = 17$ мм. Поэтому получение четкого изображения на сетчатке глаза при изменении положения предмета возможно только за счет изменения его фокусного расстояния F. Таким образом, аккомодация глаза осуществляется путем изменения кривизны хрусталика цилиарной мышцей до тех пор, пока на сетчатке не образуется четкое изображение предмета. В результате в соответствии с формулой (24.1) изменяется фокусное расстояние F глаза.

Так как изменение кривизны поверхностей хрусталика может происходить только в определенных пределах, то для всякого глаза существуют границы расположения предмета, в пределах которых глаз может его отчетливо видеть, называемые *дальней* и *ближней точками аккомодации* (четкого видения). Нормальный глаз в ненапряженном состоянии аккомодирован на бесконечно удаленные (более 20 м) предметы, т.е. его дальняя точка аккомодации находится в бесконечности. С приближением предмета к глазу размеры его изображения на сетчатке увеличиваются, что позволяет лучше рассмотреть детали предмета. Ближняя точка аккомодации находится на расстоянии 10–12 см от глаза, но при этом напряжение мышцы, деформирующей хрусталик, значительно возрастает и она быстро утомляется. ***Расстояние наилучшего зрения*** — это оптимальное расстояние при рассматривании предметов, чтении и т.п. На этом расстоянии аккомодация осуществляется без значительного утомления цилиарной мышцы. Для нормального глаза оно составляет около 250 мм.

Для близорукого глаза оба предела аккомодации находятся ближе, т.е. дальняя точка находится на конечном расстоянии (несколько метров) от глаза, а расстояние наилучшего зрения меньше 250 мм. Для дальнозоркого глаза ближняя точка аккомодации отдаляется и расстояние наилучшего зрения превышает 250 мм.

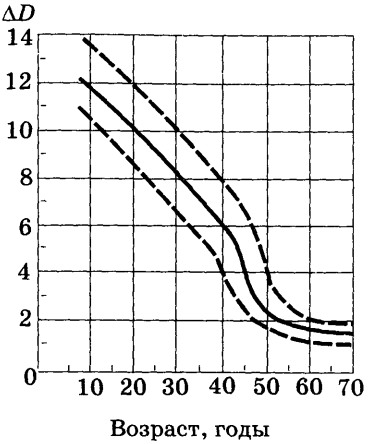


Рис. 24.3. Зависимость диапазона ΔD аккомодации глаза от возраста человека в норме (сплошная линия). Пунктиром показана величина стандартного отклонения [37]

Диапазоном аккомодации называется интервал изменения преломляющей силы глаза (в диоптриях) при изменении его аккомодации от дальней до ближней точки. Самый широкий диапазон — в молодом возрасте (14 дптр), а с возрастом хрусталик постепенно теряет свою упругость, цилиарная мышца ослабевает и диапазон аккомодации существенно снижается (рис. 24.3) — развивается старческое зрение (пресбиопия).

24.2.2. Недостатки рефракции глаза и их коррекция

В норме при аккомодации глаза на удаленные предметы ($d = \infty$) задний фокус глаза располагается на сетчатке (рис. 24.4, *а*) — такой глаз называется ***эмметропическим***. При *близорукости* (*миопии*) преломляющая сила глаза избыточна, и задний фокус глаза находится перед сетчаткой, поэтому лучи от удаленного предмета фокусируются перед сетчаткой, а затем расходятся, образуя на ней размытое пятно (рис. 24.4, *б*). Поэтому такой глаз четко видит предметы, расположенные не далее определенной точки, называемой дальней точкой аккомодации (составляет лишь несколько

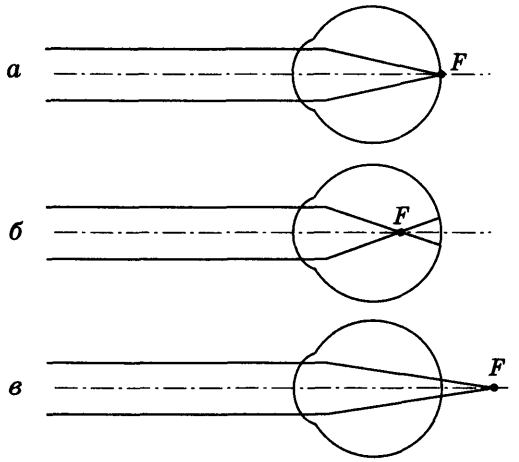


Рис. 24.4. Расположение заднего фокуса глаза при расслабленном хрусталике:
а — в норме; *б* — при близорукости; *в* — при дальнозоркости

метров), а ближняя точка аккомодации и расстояние наилучшего зрения у такого глаза меньше, чем в норме.

В случае *дальнозоркости* (*гиперметропии*) при расслабленной цилиарной мышце задний фокус глаза оказывается за сетчаткой (рис. 24.4, *в*), а на ней создается нечеткое изображение, но этот недостаток глаз компенсирует за счет некоторого напряжения цилиарной мышцы. С возрастом происходит уплотнение и повышение ригидности хрусталика, а также ослабление этой мышцы, поэтому создаваемое мышцей напряжение становится недостаточным для четкого наблюдения близко расположенных предметов и расстояние наилучшего зрения становится больше нормы (>250 мм).

При обоих видах отклонения рефракции глаза от нормы зрение может быть подкорректировано помещением перед глазом соответствующих линз (очков): при коррекции близорукости пациенту подбирают необходимую рассеивающую линзу (рис. 24.5, *а*), а для коррекции дальнозоркости — собирающую (рис. 24.5, *б*). Эти линзы возвращают задний фо-

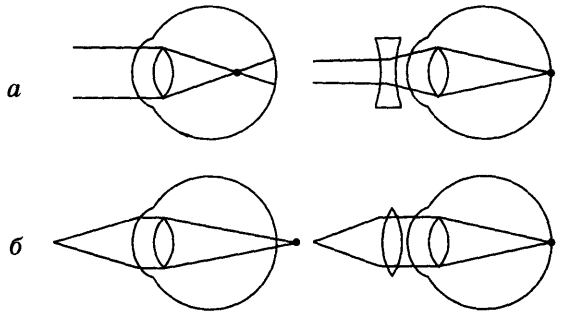


Рис. 24.5. Компенсация с помощью линз:
а — близорукости; *б* — дальнозоркости

кус глаза на сетчатку и позволяют расширить пределы аккомодации глаза.

Дефекты зрения различаются не только по виду, но и по степени. Чем дальше находится задний фокус глаза от сетчатки, тем выше степень аметропии (отклонения рефракции глаза от нормы). Степень аметропии измеряют в диоптриях, она равна преломляющей силе линзы, исправляющей дефект зрения, т.е. помещающей фокус на сетчатку, причем при близорукости аметропия отрицательная, при дальнозоркости — положительная.

Более сложным недостатком рефракции глаза является *астигматизм*, при котором световые лучи, лежащие в горизонтальной и вертикальной плоскостях, имеют разные, не совпадающие фокусы. Вследствие этого изображение предмета на сетчатке глаза всегда нечеткое: если глаз аккомодирован на горизонтальные линии, то размыты вертикальные и наоборот. Компенсируют астигматизм путем подбора цилиндрических линз, которые также могут быть положительными и отрицательными.

24.2.3. Острота зрения

Если глаз рассматривает некоторый предмет AB, то угол φ, образованный лучами, идущими от его крайних точек A и B через центр хрусталика, называется *углом зрения* φ этого предмета (рис. 24.6). Под таким же углом видно изображение $A'B'$ этого предмета на сетчатке из центра хрусталика.

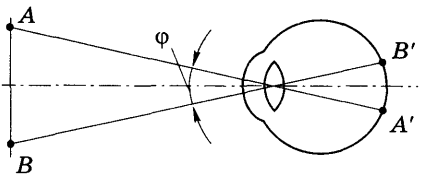


Рис. 24.6. Угол зрения

С уменьшением размера предмета AB уменьшается и угол зрения, под которым он виден.

Острота зрения определяется наименьшим углом зрения, при котором глаз способен видеть раздельно две точки предмета.

Две точки предмета A и B воспринимаются глазом раздельно, если их изображения на сетчатке приходятся на разные колбочки. Так как минимальное расстояние между колбочками в центральной ямке $d = 5$ мкм, а расстояние от сетчатки до центра хрусталика $f = 17$ мм, то минимальный угол зрения, при котором две точки еще видны раздельно, в норме равен примерно одной угловой минуте: $\varphi_{min} = \frac{d}{f} = 3 \cdot 10^{-4}$ рад $= 1'$.

Минимальный угол зрения φ_{min} (в угловых минутах) человека можно определить, используя специальные таблицы с символами разного размера, где сразу указана острота зрения γ, вычисляемая по формуле

$$\gamma = \frac{1'}{\varphi_{min}}. \qquad (24.2)$$

Для нормального глаза $\varphi_{min} = 1'$, поэтому в норме острота зрения равна единице, а с увеличением наименьшего угла зрения φ_{min} острота зрения уменьшается.

Острота зрения определяет способность глаза различать малые детали предметов, и чем ближе к глазу расположен предмет, тем лучше видны его мелкие детали. Минимальное расстояние $Z_{гл}$ между двумя точками предмета, различаемыми невооруженным глазом на расстоянии наилучшего зрения, называют ***пределом разрешения глаза***. В норме (острота

зрения $\gamma = 1$, $\varphi_{min} = 1' = 3 \cdot 10^{-4}$ рад и расстояние наилучшего зрения $d = 250$ мм) предел разрешения глаза составляет примерно 75 мкм:

$$Z_{гл} = d\varphi_{min} = 250\,мм \cdot 3 \cdot 10^{-4} = 75 \cdot 10^{-3}\,мм = 75\,мкм. \quad (24.3)$$

Размер изображения на сетчатке при этом равен 5 мкм, т.е. среднему расстоянию между двумя колбочками в центральной ямке.

Отметим, что на остроту зрения влияет очень много факторов, например уровень освещенности среды, уровень цветового и яркостного контраста между рассматриваемым объектом и фоном, диаметр зрачка, возраст человека, особенности рефракции глаза.

При нормальном освещении острота зрения максимальна, когда изображение предмета попадает на центральную ямку сетчатки, где плотность колбочек наибольшая. В сумерках, когда колбочки уже «отключаются», острота зрения максимальна при попадании света на тот участок сетчатки, где наибольшая плотность палочек, т.е. примерно под углом 20° от центральной ямки.

24.3. Биофизические основы фоторецепции

В данном параграфе рассматриваются биофизические механизмы преобразования световой энергии, поглощенной фоторецепторными клетками сетчатки глаза, в серию нервных импульсов, передающих информацию по зрительному нерву в мозг, свойства дневного и сумеречного зрения, механизмы цветного зрения и световой адаптации глаза.

24.3.1. Строение сетчатки и светочувствительных клеток

Светочувствительными элементами в сетчатке являются палочки и колбочки. ***Палочки*** (примерно 125 млн) расположены по всей поверхности сетчатки, причем на пери-

ферии их концентрация выше, чем в центральной части. Палочки ответственны за черно-белое (сумеречное) зрение и обладают большей чувствительностью, чем колбочки.

Колбочки (около 6,5 млн) отвечают за цветное (дневное) зрение, они концентрируются в основном в центральной части сетчатки, особенно много их в желтом пятне и вокруг него. Здесь колбочки наиболее тесно прижаты друг к другу и каждая из них связана с отдельным волокном зрительного нерва, который передает в мозг ее сигналы, что определяет высокую разрешающую способность глаза в этой области сетчатки. На периферии сетчатки одно волокно зрительного нерва связано с целой группой фоторецепторов, разрешающая способность глаза здесь значительно ниже, но выше чувствительность.

Строение сетчатки человека (рис. 24.7) достаточно сложно. Обычно в ней различают десять слоев, причем светочувствительные и нервные клетки находятся в самом внутреннем слое сетчатки. Более того, палочки и колбочки ориентированы к свету своими внутренними сегментами, не содержащими зрительного пигмента и прозрачными для видимого света.

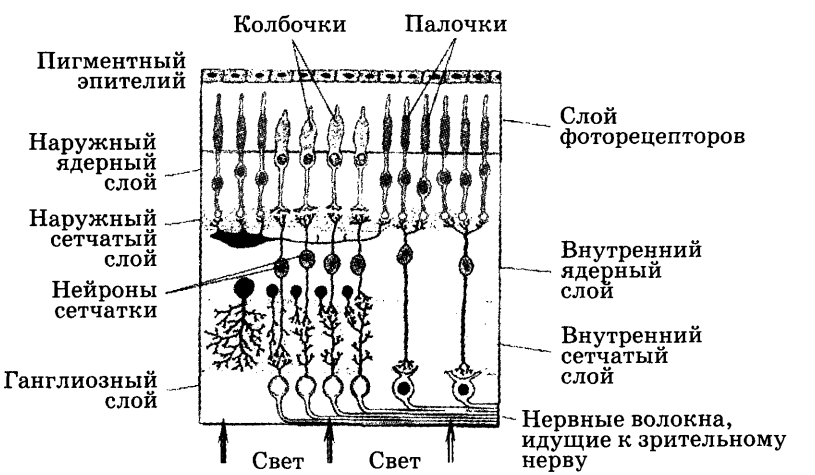


Рис. 24.7. Схема строения сетчатки глаза

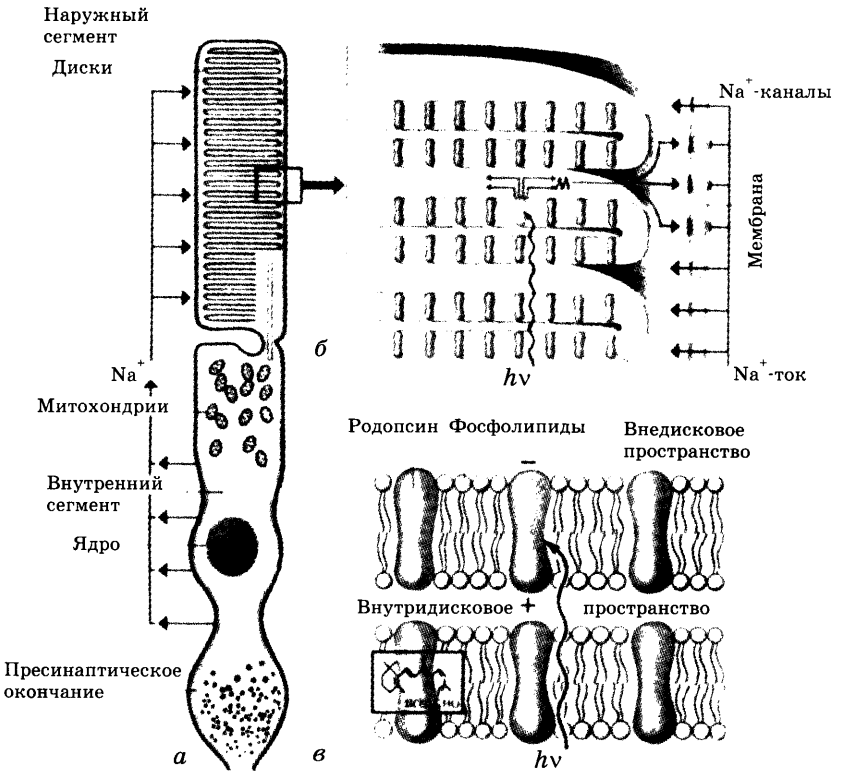


Рис. 24.8. Процессы в фоторецепторе (палочке):
а — схема строения палочки; *б* — выход медиатора из дисков и блокирование Na^+-каналов в мембране палочки;
в — поглощение светового кванта $h\nu$ родопсином, встроенным в мембрану диска

Зрительный пигмент палочек — родопсин, поглощающий световые кванты, находится в мембране дисков, расположенных в наружном сегменте палочек (рис. 24.8). Мембрана дисков образована бислоем фосфолипидов, а внутри дисков находятся ионы кальция Ca^{2+} [24].

Родопсин представляет собой бесцветный белок опсин, с которым химически связана хромофорная, т.е. светочувствительная, группа, так называемый *ретиналь*, представляющий собой альдегид витамина А.

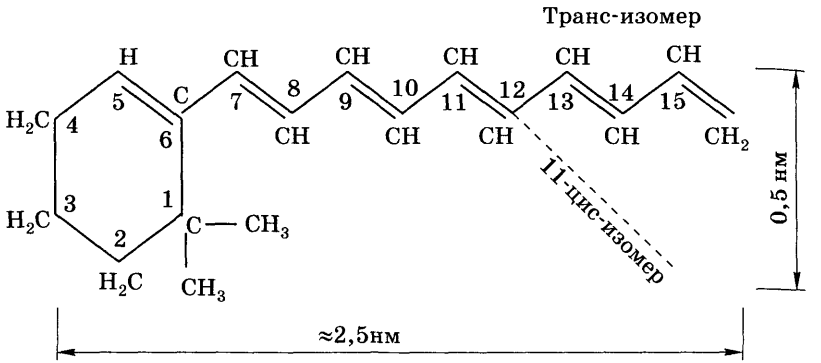


Рис. 24.9. Ретиналь

Ретиналь может иметь различные пространственные конфигурации (изомеры), однако с белком опсином может связываться только один изомер — так называемый 11-цис-ретиналь (рис. 24.9).

Следует отметить, что эта форма зрительного пигмента присуща всем без исключения организмам от низших до высших.

24.3.2. Преобразование световой энергии в нервные импульсы

Следует отметить, что потенциал покоя палочек и колбочек положительный, поскольку мембраны их наружных сегментов в покое хорошо проницаемы для ионов Na^+, тогда как мембраны клеток других возбудимых тканей в состоянии покоя наиболее проницаемы для ионов K^+. Веществом, которое в покое поддерживает открытыми натриевые каналы, является циклический гуанозин-монофосфат (цГМФ), который, по некоторым данным, при световом облучении связывается специальными молекулами [24].

Основной фотохимической реакцией, которая происходит при поглощении кванта света, является ***фотоизомеризация*** ретиналя. Поглотив световой квант, молекула 11-цис-ретиналя распрямляется и за 10^{-12} с превращается в другой изомер — транс-ретиналь. Транс-ретиналь уже не помещается в центре связывания белка опсина, вследствие чего происходит

разрыв родопсина на опсин и транс-ретиналь, протекающий через ряд промежуточных стадий.

После распада родопсина на опсин и ретиналь проницаемость дисковой мембраны изменяется и из диска начинают выходить ионы кальция Ca^{2+}, способствующие связыванию цГМФ[1]. В результате проницаемость мембраны палочки для натрия снижается и на ее мембране развивается фоторецепторный потенциал, пропорциональный количеству поглощенных квантов. Амплитуда этого фоторецепторного потенциала не превышает 5 мВ. Он распространяется по мембране палочки, доходит до пресинаптического окончания и стимулирует выброс содержащегося в нем нейромедиатора в синаптическую щель, что приводит к возбуждению биполярной нервной клетки. Потенциал действия, сформированный этой нервной клеткой, передается далее в виде нервного импульса по зрительному нерву в зрительные центры головного мозга.

Частота следования нервных импульсов возрастает с увеличением интенсивности падающего света по логарифмическому закону.

Таким образом, вся *цепочка превращения энергии света* в нервный импульс выглядит следующим образом: квант света $h\nu$ → поглощение его 11-цис-ретиналем → цис-транс-конформация ретиналя → разрыв родопсина на опсин и ретиналь → увеличение проницаемости дисковой мембраны для Ca^{+2} → диффузия Ca^{+2} к цитоплазматической мембране палочки → блокирование Na^{+}-каналов → образование фоторецепторного потенциала на мембране палочки → возбуждение нейронов сетчатки → поступление нервных импульсов в зрительные центры головного мозга.

Зрительный пигмент родопсин, распавшийся под действием света на опсин и ретиналь, может затем опять восстанавливаться под действием специального фермента ретиналь-изомеразы, который переводит транс-ретиналь в 11-цис-ретиналь, после чего происходит присоединение последнего к опсину с образованием родопсина. Этот процесс может осу-

[1] По некоторым данным, при световом облучении цГМФ связывается также специальными ферментами, что увеличивает рецепторный потенциал на мембране.

ществляться и на свету, и в темноте. Однако на свету процесс регенерации родопсина ускоряется, так как переход транс-ретиналя в 11-цис-ретиналь осуществляется в этом случае и под действием света, благодаря обратимости фотореакции цис-транс-конформации ретиналя.

24.3.3. Цветное зрение

Процессы в колбочках полностью аналогичны тем, что происходят в палочках. Отличие лишь в том, что в колбочках зрительным пигментом является уже не родопсин, а ***йодопсин***, который представляет собой другой белок, содержащий тот же хромофор: 11-цис-ретиналь.

В сетчатке имеются три вида колбочек с разными спектрами поглощения и спектрами фотобиологического действия. Эти три вида колбочек отличаются друг от друга строением белка-матрицы йодопсина, с которым связана хромофорная группа 11-цис-ретиналь. Энергия связи ретиналя с белком в этих трех видах йодопсина разная, поэтому в сетчатке имеются три типа колбочек — с максимумами поглощения на 440, 540 и 590 нм, тогда как родопсин палочек имеет максимум поглощения на 510 нм (рис. 24.10 и 24.11). Монохроматический свет возбуждает один (иногда два) вид колбочек, а свет сложного спектрального состава — все три типа колбо-

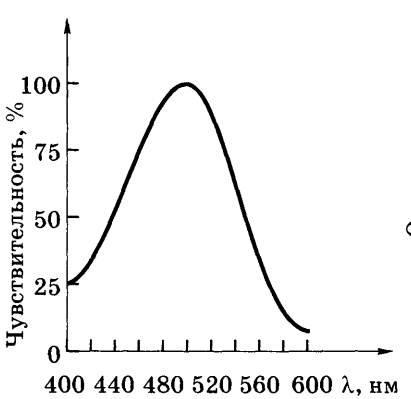


Рис. 24.10. Спектр фотобиологического действия палочек

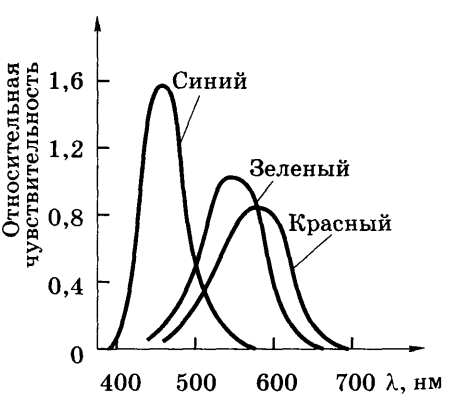


Рис. 24.11. Спектры фотобиологического действия колбочек

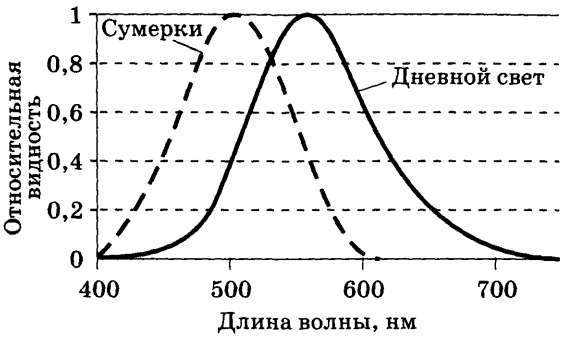


Рис. 24.12. Кривые видности для дневного и сумеречного зрения

чек, но в различной степени, в зависимости от формы этого спектра, что и приводит в итоге к возникновению оттенков цветового ощущения.

Таким образом, цветовое зрение обусловлено наличием в сетчатке трех видов колбочек, отличающихся спектрами фотобиологического действия.

При отсутствии в сетчатке хотя бы одного из видов колбочек глаз теряет способность правильно различать цвета (дальтонизм). Чувствительность глаза к излучениям разных длин волн видимого диапазона характеризуется *кривой видности*. Суммарная кривая спектральной чувствительности глаза для случая яркого освещения, т.е. цветового зрения, имеет максимум в желто-зеленой части спектра при $\lambda = 555$ нм (рис. 24.12).

В целом чувствительность колбочек к свету более слабая, чем у палочек, поэтому цветовое зрение имеет место при хорошей освещенности объектов, а при слабой освещенности колбочки «отключаются» и работает только палочковый аппарат сетчатки.

Поэтому при сумеречном зрении кривая видности смещается в сторону коротких волн с максимумом около 510 нм (см. рис. 24.10) в соответствии со спектром поглощения палочек. Поскольку в сумерках колбочки уже «не работают», то изменяется и цветовосприятие («ночью все кошки серы»). Так как палочки чувствительны в основном в сине-голубой

области спектра, то красный цвет ($\lambda = 630–760$ мкм) при низких уровнях освещенности будет выглядеть практически черным или темно-серым.

Отметим, что область зрительного восприятия глаза не обрывается резко на длинах волн 400 и 760 нм. В условиях темновой адаптации глаз может видеть (в очень слабой степени) интенсивные инфракрасные лучи с длинами волн до 950 нм и ультрафиолетовые лучи в диапазоне 400–380 нм.

24.3.4. Адаптация глаза

При резком изменении освещенности рассматриваемого объекта глаз некоторое время не может правильно различать его детали и оттенки, поскольку ему необходимо время, чтобы «подстроить» свою чувствительность под условия освещенности объекта наблюдения. Процесс приспособления глаза к тому или иному уровню освещенности называется ***адаптацией***. При повышении освещенности объекта происходит *световая*, а при понижении — *темновая* адаптация глаза. Световая адаптация происходит достаточно быстро, в пределах минуты, тогда как темновая адаптация длится десятки минут (до часа), при этом чувствительность глаза при переходе от яркого света к темноте возрастает примерно в 10^7 раз.

Глаз является одним из наиболее чувствительных оптических приборов и способен воспринимать световые потоки в интервале $10^{-17}–10^{-5}$ Вт.

Известно два основных механизма адаптации глаза:

1) изменение диаметра зрачка глаза;

2) изменение концентрации родопсина в палочках и йодопсина в колбочках.

При изменении освещенности объекта *диаметр зрачка* изменяется за 0,3 с в диапазоне от 2 до 8 мм (в 4 раза), при этом освещенность сетчатки глаза изменяется в $4^2 = 16$ раз.

Основной механизм адаптации — это изменение *концентрации зрительного пигмента в сетчатке* — родопсина и йодопсина. Под действием света эти соединения быстро распадаются и, соответственно, быстро падает и чувствительность глаза, поэтому адаптация глаза к повышенной освещенности

происходит достаточно быстро. Обратный процесс — адаптация к плохой освещенности — идет гораздо медленнее, поскольку он связан с восстановлением зрительных пигментов под действием специального фермента — ретиналь-изомеразы. На свету процесс регенерации родопсина идет быстрее, чем в темноте, благодаря обратимости фотореакции цис-транс-конформации ретиналя.

Если освещенность сетчатки постоянна, то устанавливается динамическое равновесие между процессом распада зрительного пигмента под действием света и восстановлением его с участием ферментов. При снижении освещенности сетчатки уменьшается количество распадов зрительного пигмента и равновесие сдвигается в сторону более высокой концентрации родопсина и йодопсина, чувствительность глаза повышается. При увеличении освещенности сетчатки равновесие сдвигается в противоположную сторону и чувствительность глаза снижается.

Рассмотренные механизмы адаптации позволяют глазу в широких пределах изменять светочувствительность фоторецепторов с изменением освещенности сетчатки и обеспечивают глазу способность воспринимать световые потоки в широком интервале интенсивностей.

Абсолютный порог чувствительности глаза человека соответствует примерно 60–150 квантам сине-зеленого света, падающим в глаз. Причем только 5–15 из них поглощаются непосредственно молекулами зрительного пигмента, остальные рассеиваются или поглощаются другими структурами сетчатки. Все кванты света, достигающие сетчатки при пороговой интенсивности, поглощаются отдельными палочками, которые обладают гораздо большей чувствительностью к свету, чем колбочки.

Поскольку чувствительность палочек к свету значительно выше, чем колбочек, то при *темновой адаптации* основное световосприятие обусловлено палочковой системой сетчатки (сумеречное зрение). Поэтому в темноте слабые световые стимулы лучше различаются не центральной ямкой, а окружающей ее частью сетчатки. Максимальная темновая адаптация достигается достаточно медленно (40—60 мин), поскольку

ферментам необходимо время, чтобы «наработать» большое количество родопсина из ретиналя и опсина, распавшихся на свету.

Световая адаптация протекает намного быстрее темновой, в течение нескольких секунд, поскольку при высокой освещенности распад родопсина в палочках идет очень интенсивно и его концентрация резко снижается. В этих условиях в процессе зрения участвуют почти исключительно колбочки. Дневное зрение характеризуется высокой остротой зрения и хорошим восприятием *цвета*, так как колбочки способны различать цвета, а палочки этой способностью практически не обладают.

Информация об интенсивности света кодируется частотой нервных импульсов, поступающих в зрительные центры головного мозга. Однако если интенсивность света остается постоянной, то частота нервных импульсов постепенно уменьшается. Поэтому, если каким-либо способом зафиксировать глаз на освещенном предмете, он через 1–2 с вообще перестанет видеть предмет, так как прекратится поступление нервных импульсов в головной мозг (временна́я адаптация). В реальных условиях этого не происходит благодаря наличию специальных мышц, периодически сдвигающих глазное яблоко на угол до 20°. Продолжительность такого скачка равна сотым долям секунды, а интервал между ними, т.е. продолжительность фиксации глаза на предмете, составляет обычно 0,2–0,5 с. В результате этих движений изображение предмета перемещается на другие светочувствительные клетки сетчатки и информация продолжает поступать в головной мозг.

Частота нервных импульсов, поступающих в мозг, пропорциональна логарифму интенсивности падающего в глаз света, благодаря чему к зрению применим известный психофизический закон Вебера — Фехнера (см. гл. 7): зрительное ощущение E пропорционально логарифму интенсивности I падающего в глаз света:

$$E = k\ln I + C, \qquad (24.4)$$

где k — коэффициент пропорциональности, а значение константы C зависит от адаптации глаза.

Выясним некоторые следствия этого закона. Если человек осматривает освещенный фон интенсивностью I, на котором есть более светлое пятно интенсивностью $I_1 = = I + dI$, то минимальную обнаруживаемую человеком разность интенсивностей dI пятна и фона называют *разностным порогом*, а его отношение к интенсивности I фона, т.е. dI/I — *дифференциальным порогом.*

Продифференцировав (24.4), получим выражение для чувствительности глаза

$$dE = k\frac{dI}{I},$$

а затем выражения для дифференциального и разностного порогов:

$$\frac{dI}{I} = \frac{1}{k}dE, \qquad dI = \frac{1}{k}I \cdot dE. \tag{24.5}$$

Соотношения (24.5) показывают:

• что *дифференциальный порог* dI/I *постоянен*, не зависит от яркости фона и определяется только чувствительностью dE глаза;

• разностный порог dI пропорционален как чувствительности dE глаза, так и яркости I фона: чем яркость выше, тем выше и разностный порог, следовательно, тем хуже способность глаза к различению градаций яркости.

Эта особенность зрения учитывается в оптических приборах: наиболее точные сравнения освещенностей двух объектов (например, выравнивание освещенности полей в поляриметре и др.) проводятся при минимальных яркостях фона.

Контрольные вопросы и задания

1. Опишите строение глаза. Каковы оптические свойства глазных сред?
2. Каково назначение оптической системы глаза и какими структурными элементами она создается? Какова оптическая сила глаза, что такое приведенный глаз?

3. Каково условие четкого видения предметов? Каков механизм аккомодации глаза? Что такое диапазон аккомодации, от чего он зависит?
4. Каковы основные недостатки рефракции глаза и как их корректируют?
5. Что такое минимальный угол зрения? острота зрения? предел разрешения глаза?
6. Каково строение сетчатки глаза? Назовите виды светочувствительных клеток, их количество и распределение по сетчатке.
7. Опишите строение палочки. Какова роль родопсина и ретиналя? В чем различие палочек и колбочек как светочувствительных клеток сетчатки?
8. Опишите механизм образования фоторецепторного потенциала на мембране палочки или колбочки.
9. Каков путь передачи информации об освещенности палочки в зрительные центры головного мозга?
10. Чем обусловлены различия сумеречного и дневного зрения, в чем они проявляются? Приведите спектры фотобиологического действия палочек и колбочек, кривую видности глаза при дневном и сумеречном освещении.
11. Что такое адаптация глаза? Каковы основные механизмы ее реализации?
12. Укажите количественную связь между зрительным ощущением и интенсивностью падающего в глаз света, а также влияние яркости фона на разностный и дифференциальный пороги зрения.

Глава 25 Приборы для исследования микрообъектов

Предел разрешения глаза человека в норме составляет около 1 угловой минуты, что позволяет различать на расстоянии наилучшего зрения детали предметов размером не менее 75 мкм. Более мелкие детали предмета невооруженным глазом рассмотреть невозможно. Поэтому изучение микрообъектов нельзя проводить без специальных приборов, создающих увеличенное изображение микрообъекта в десятки, сотни и даже сотни тысяч раз. Без таких приборов немыслима современная наука, в том числе биология и медицина. Так, без оптического микроскопа нельзя изучить гистологию, патологическую анатомию и многие другие науки.

25.1. Линза

Главными оптическими элементами, создающими увеличенные изображения предметов, являются сферические линзы. Знание оптических свойств линз, методов и условий построения ими четких изображений позволяет лучше понять принципы и особенности формирования изображений лупами и микроскопами, рассчитать создаваемое ими увеличение.

25.1.1. Оптические характеристики линз

Линзой называется прозрачное тело, ограниченное двумя сферическими (или плоской и сферической) поверхностями и отличающееся от окружающей среды показателем преломления. Линза называется *тонкой*, если ее толщина мала по сравнению с радиусами кривизны R_1 и R_2 ее поверхностей.

Прямая O_1O_2, проходящая через центры кривизны поверхностей линзы (рис. 25.1), называется ***главной оптической осью***. На этой оси расположен и ***оптический центр*** O линзы, находящийся в геометрическом центре ее средней части. Толщина тонкой линзы не учитывается, поэтому ее оптичес-

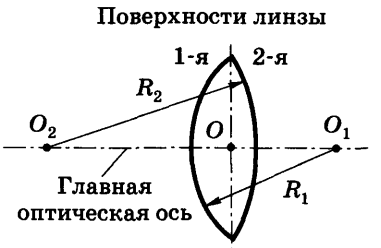


Рис. 25.1. Схема двояковыпуклой линзы

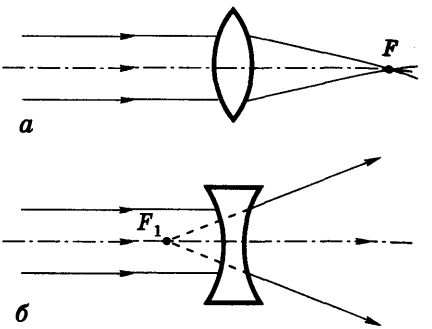


Рис. 25.2. Фокусы линзы: *а* — действительный у собирающей линзы; *б* — мнимый у рассеивающей линзы

кий центр — это точка пересечения самой линзы с ее главной оптической осью. Любая прямая, проходящая через оптический центр O, является ***побочной оптической осью*** линзы, и все лучи света, идущие через центр линзы, не преломляются. Другие лучи, идущие не через центр линзы, после прохождения линзы изменяют свое направление, т.е. преломляются.

Если параллельный пучок света после прохождения линзы сходится в одной точке F, то линзу называют ***собирающей*** (рис. 25.2, *а*), а точку F — ее ***действительным фокусом.***

Если же параллельный пучок света после прохождения линзы расходится так, что кажется исходящим из одной точки F_1, то линза называется ***рассеивающей***, а точка F_1 — ***мнимым фокусом*** этой линзы (рис. 25.2, *б*).

Фокус линзы, лежащий на главной оптической оси, называется ***главным фокусом*** линзы, а все остальные ее фокусы (лежащие на других оптических осях) называются ***побочными***. При этом все фокусы линзы лежат в одной плоскости, которая называется фокальной плоскостью линзы. Она перпендикулярна главной оптической оси линзы и проходит через ее главный фокус.

Расстояние F от центра линзы до ее главного фокуса называется ***фокусным расстоянием линзы.***

Если в передний фокус собирающей линзы поместить точечный источник света, то после прохождения линзы все лучи

пойдут параллельно оптической оси линзы, что используется на практике для получения параллельных пучков света от точечных источников.

Основной количественной характеристикой преломляющей способности линзы служит ее ***оптическая сила*** — величина, равная отношению показателя преломления среды, окружающей линзу, к фокусному расстоянию линзы в этой среде:

$$D = \frac{n_0}{F}. \qquad (25.1)$$

В СИ оптическая сила линзы измеряется в диоптриях, при этом фокусное расстояние должно измеряться в метрах: 1 дптр = 1 $м^{-1}$. Оптической силой в одну диоптрию обладает линза с фокусным расстоянием $F = 1$ м, находящаяся в вакууме или воздухе ($n = 1$). Оптическая сила собирающих линз положительная, а рассеивающих — отрицательная.

Оптическая сила линзы зависит от показателей преломления вещества линзы и окружающей ее среды, а также от радиусов кривизны ее поверхностей и может быть вычислена по формуле

$$D = \frac{n_0}{F} = (n - n_0)\left(\frac{1}{R_1} - \frac{1}{R_2}\right), \qquad (25.2)$$

где n_0 — показатель преломления среды, окружающей линзу; n — показатель преломления вещества линзы; R_1 и R_2 — радиусы кривизны ее первой и второй (по ходу луча) поверхностей, которые считаются положительными, если их центры расположены справа от линзы (за линзой по ходу луча), и отрицательными — если слева от нее (перед линзой).

Например, на рис. 25.1 для первой поверхности $R_1 > 0$, а для второй $R_2 < 0$, поэтому формула (25.2) для двояковыпуклой линзы примет вид

$$D = \frac{n_0}{F} = (n - n_0)\left(\frac{1}{R_1} + \frac{1}{|R_2|}\right).$$

Если линза находится в воздухе, то $n_0 = 1$ и формула (25.2) несколько упрощается:

$$D = \frac{1}{F} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right). \qquad (25.3)$$

Если $F > 0$, то фокус действительный и линза собирающая ($D > 0$), а если $F < 0$, то фокус мнимый и линза рассеивающая ($D < 0$).

Отметим, что каждая линза имеет передний (перед линзой по ходу луча) и задний (за линзой) фокусы. Если по обе стороны линзы одинаковая среда, то переднее F_1 и заднее F_2 фокусные расстояния линзы равны. Если же по обе стороны линзы среды разные (например, воздух и водянистая влага по разные стороны роговицы глаза), то переднее F_1 и заднее F_2 фокусные расстояния не равны и отличаются в n_1/n_2 раз:

$$\frac{F_1}{F_2} = \frac{n_1}{n_2}, \qquad (25.4)$$

где n_1 и n_2 — показатели преломления соответствующих сред.

25.1.2. Построение изображения в линзе

Чтобы получить изображение точки, создаваемое линзой, необходимо найти точку пересечения исходящих из нее лучей после их преломления в линзе. Для этого удобно использовать любую пару из трех лучей, представленных на рис. 25.3:

- луч, параллельный главной оптической оси, после преломления в линзе проходит через ее задний фокус F_2;
- луч, проходящий через передний главный фокус F_1 линзы, после преломления идет параллельно главной оптической оси;
- луч, идущий через оптический центр O линзы, не преломляется, т.е. после линзы идет в том же направлении.

Пересечение этих трех лучей, исходящих из точки A, дает точку A_1, являющуюся изображением точки A. Опуская из точки A_1 перпендикуляр на главную оптическую ось, найдем точку B_1 — изображение точки B.

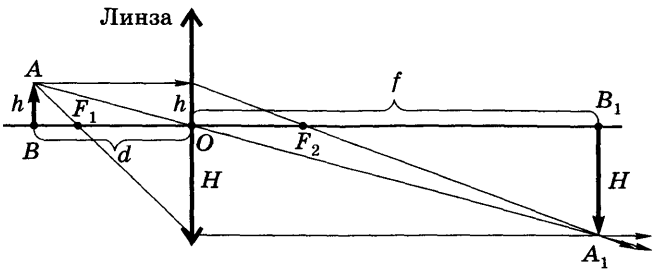


Рис. 25.3. Построение действительного изображения в собирающей линзе

Если лучи, исходящие из точки A, после прохождения линзы пересекаются, то изображение A_1 называется ***действительным***. Действительное изображение всегда *перевернутое* (обратное) и находится с предметом всегда по разные стороны линзы (рис. 25.3). Его можно наблюдать и регистрировать, поместив в плоскости пересечения лучей экран или фотопленку. Если же лучи, исходящие из точки A, после прохождения линзы расходятся, то пересекаются их продолжения, и такое изображение A_1 называется ***мнимым***, — оно всегда *прямое* и расположено по ту же сторону линзы, что и предмет. Такое изображение нельзя непосредственно наблюдать на экране или зафиксировать на фотопленке, но его можно наблюдать глазом.

25.1.3. Формула тонкой линзы

Если предмет находится на расстоянии d от линзы с фокусным расстоянием F, а его изображение — на расстоянии f от линзы (см. рис. 25.3), то эти величины связаны соотношением, которое называют ***формулой тонкой линзы***:

$$\frac{1}{d}+\frac{1}{f}=\frac{1}{F}. \qquad (25.5)$$

Расстояние d от предмета до линзы всегда положительное, а расстояние f от линзы до изображения положительное для действительных изображений и отрицательное для мнимых.

Фокусное расстояние F собирающих линз положительное, а рассеивающих — отрицательное.

Если предмет находится на бесконечности (т.е. удален на большое расстояние $d \to \infty$), то его изображение будет расположено в фокусе линзы ($f = F$), а по мере приближения предмета к линзе его изображение удаляется от линзы и $f > F$.

Линейным увеличением Г линзы называют отношение линейных размеров H изображения к линейным размерам h предмета; оно может быть выражено соотношением, которое нетрудно получить, пользуясь рис. 25.3 и формулой (25.5):

$$\Gamma = \frac{H}{h} = \frac{f}{d} = \frac{F}{|d - F|}. \qquad (25.6)$$

Видно, что увеличение линзы тем больше, чем меньше знаменатель этой дроби, т.е. чем ближе предмет расположен к переднему фокусу линзы.

Собирающая линза может создавать как действительные, так и мнимые изображения, в зависимости от расположения предмета относительно линзы. Действительные изображения формируются в тех случаях, когда предмет расположен от линзы на расстоянии d, большем фокусного расстояния F, т.е. при $d > F$, а мнимые — при $d < F$, когда предмет располагается между передним фокусом и линзой.

Собирающая линза может давать три вида действительных изображений, а также мнимое увеличенное изображение:

- действительное уменьшенное ($\Gamma < 1$), если $d > 2F$;
- действительное равное ($\Gamma = 1$), при $d = 2F$;
- действительное увеличенное ($\Gamma > 1$), если $F < d < 2F$;
- мнимое увеличенное ($\Gamma > 1$), если $d < F$.

Рассеивающая линза всегда дает только мнимые уменьшенные прямые изображения. Это прямо следует из соотношений (25.5) и (25.6): поскольку $d > 0$ всегда, а фокусное расстояние F такой линзы отрицательное, то изображение получается мнимым ($f < 0$): $\frac{1}{f} = -\frac{1}{|F|} - \frac{1}{d} < 0$ — и уменьшенным: $\Gamma = \frac{|F|}{d + |F|} < 1$.

25.2. Лупа

Лупой называется собирающая линза, используемая для получения увеличенных мнимых изображений предмета.

Лупу обычно размещают непосредственно перед глазом, а рассматриваемый предмет помещается вблизи ее переднего фокуса, но ближе к линзе ($d \approx F$, но $d < F$). Прямое мнимое изображение, создаваемое лупой, располагают от нее, как правило, на расстоянии наилучшего зрения $f = d_0 = 0{,}25$ м. Лупа фактически увеличивает угол зрения, под которым рассматривается изображение предмета. Последнее наблюдается под бóльшим углом, чем сам предмет при рассматривании невооруженным глазом, что и позволяет видеть более мелкие детали предмета (рис. 25.4).

Линейное увеличение лупы определяется формулой

$$\Gamma = \frac{A_1B_1}{AB} = \frac{f}{d} \approx \frac{d_0}{F}. \qquad (25.7)$$

Чем короче фокусное расстояние лупы, тем больше ее увеличение. Однако при больших увеличениях уменьшается поле зрения и возрастают искажения (аберрации) изображений. Поэтому наиболее распространены лупы с 2–25-кратным увеличением, что соответствует ее фокусным расстояниям от 12,5 до 1,0 см. Если предел разрешения невооруженного глаза составляет 75 мкм, то при использовании 25-кратной лупы он составит 75 мкм/25 = 3 мкм.

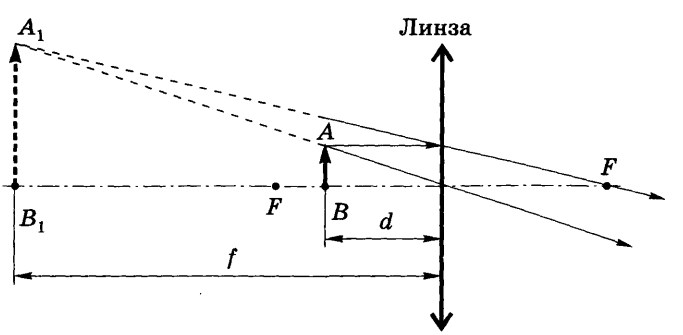


Рис. 25.4. Мнимое увеличенное изображение, создаваемое лупой

25.3. Микроскоп

Для получения бóльших увеличений применяют ***микроскоп***, который представляет собой сочетание двух собирающих линз: объектива и окуляра, расположенных на значительном расстоянии друг от друга.

Ход лучей и построение изображения в микроскопе поясняет рис. 25.5.

Предмет AB помещают вблизи переднего фокуса объектива (причем $d_1 > F_1$), который создает его действительное увеличенное перевернутое изображение A_1B_1. ***Окуляр*** — это лупа, через которую рассматривают промежуточное изображение A_1B_1, наблюдая при этом его увеличенное мнимое изображение A_2B_2. Таким образом, оптический микро-

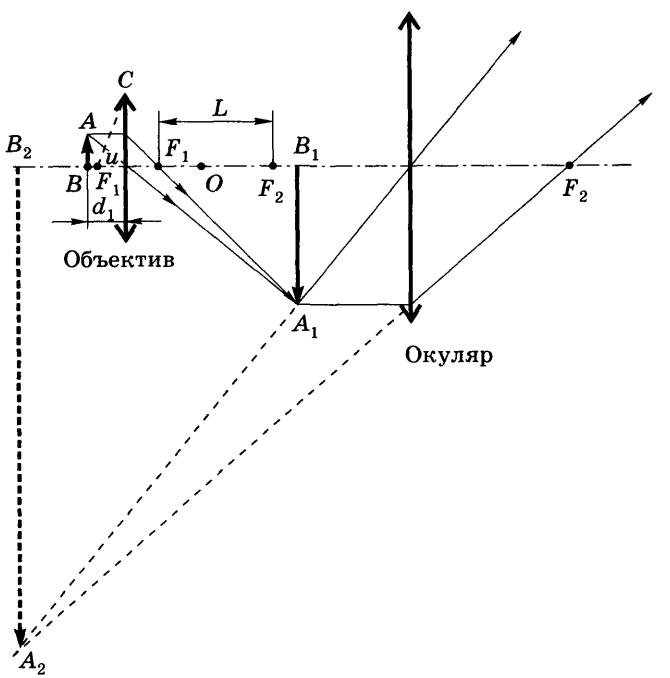


Рис. 25.5. Ход лучей и построение изображения в микроскопе: AB — объект; A_1B_1 — промежуточное изображение; A_2B_2 — конечное мнимое увеличенное изображение

скоп дает мнимое увеличенное и перевернутое изображение объекта.

Нетрудно показать, что увеличение микроскопа равно произведению увеличений объектива и окуляра:

$$\Gamma_м = \frac{A_2B_2}{AB} = \frac{A_2B_2}{A_1B_1}\frac{A_1B_1}{AB} = \Gamma_{ок}\Gamma_{об} = \frac{d_0L}{F_2F_1}. \qquad (25.8)$$

В формуле (25.8) учтено, что увеличение окуляра (лупы) $\Gamma_{ок} = d_0/F_2$, а увеличение объектива $\Gamma_{об} = A_1B_1/AB = F_1B_1/F_1 \approx L/F_1$, где d_0 — расстояние наилучшего зрения (0,25 м), L — оптическая длина тубуса микроскопа, т.е. расстояние между задним фокусом объектива и передним фокусом окуляра (см. рис. 25.5).

Формула (25.8) для увеличения микроскопа получена из соображений геометрической оптики и позволяет, казалось бы, создавать микроскопы с любым увеличением, варьируя фокусные расстояния F_1, F_2 и оптическую длину тубуса L. Например, при $F_1 = 2$ мм, $F_2 = 18$ мм и $L = 180$ мм получаем $\Gamma_м = 1250$.

Однако на практике получить такие большие увеличения сложно, поскольку наблюдению очень малых объектов препятствует явление дифракции света, которое геометрическая оптика учесть не способна. *Дифракция* — это явление отклонения света от прямолинейного распространения. Наблюдается оно при размерах Z препятствий (в данном случае — деталей рассматриваемых объектов), сравнимых с длиной волны λ видимого света, т.е. при $Z \sim \lambda$.

Волновую теорию микроскопа, учитывающую дифракционные явления, создал немецкий физик Э. Аббе. Он показал, что предел разрешения Z (наименьшее расстояние между двумя различимыми точками объекта) оптического микроскопа определяется как длиной волны λ излучения подсветки, так и свойствами микроскопа (его числовой апертурой $A = n\sin u$) и формой наблюдаемых объектов.

Формула Аббе для предела разрешения микроскопа имеет вид:

• для объектов продолговатой (линейной) формы

$$Z = \frac{0{,}5\lambda}{n\sin u}, \qquad (25.9)$$

• для объектов округлой формы

$$Z = \frac{0{,}61\lambda}{n\sin u}, \qquad (25.9а)$$

где n — показатель преломления среды, заполняющей пространство между объектом и линзой объектива (для сухого объектива — это воздух, $n = 1$); u — апертурный угол, образуемый главной оптической осью микроскопа и крайним лучом F_1C, еще попадающим в объектив из его переднего фокуса F_1 (см. рис. 25.5).

В современных микроскопах апертурный угол достаточно велик ($u \approx 70°$), $\sin 70° = 0{,}94 \approx 1$, поэтому предел разрешения «сухих» ($n = 1$) микроскопов примерно равен половине длины волны света, используемого для подсветки объекта ($Z \approx 0{,}5\lambda$), поэтому в синем и фиолетовом свете ($\lambda \approx 400$ нм) будут видны почти вдвое более мелкие детали ($Z \approx 200$ нм) объекта, чем в красном ($\lambda \approx 700$ нм, $Z \approx 350$ нм).

Улучшить разрешающую способность, т.е. уменьшить Z, микроскопа можно внесением между объектом и объективом иммерсионной жидкости с показателем преломления $n = 1{,}5$–$1{,}6$. Однако объектив в этом случае должен быть специально рассчитан на применение такой иммерсии, лишь в этом случае предел разрешения микроскопа уменьшится в n раз [см. формулу (25.9)] и соответственно возрастет его разрешающая способность.

Таким образом, даже при использовании иммерсии предел разрешения оптических микроскопов не может быть меньше 125 нм, а для «сухих» микроскопов — 200 нм.

Дальнейшее увеличение разрешающей способности возможно только за счет уменьшения длины волны подсветки, т.е. путем использования ультрафиолетового излучения.

Оптика такого микроскопа должна быть прозрачной для ультрафиолета, а изображение наблюдать непосредственно глазом уже нельзя. Поэтому создают не мнимое, а действительное увеличенное изображение предмета путем перемещения окуляра вправо (см. рис. 25.5) так, чтобы промежуточное изображение A_1B_1 оказалось левее переднего фокуса F_2 окуляра. В этом случае окуляр (и микроскоп в целом) будет давать действительное увеличенное изображение предмета. В плоскости этого изображения помещают фотопленку или люминесцирующий экран, преобразующий ультрафиолетовое изображение в видимое, которое рассматривают глазом или фотографируют. При использовании подсветки с $\lambda = 250–300$ нм удается улучшить разрешающую способность (уменьшить предел разрешения Z) микроскопа в 1,3–1,5 раза.

Дальнейшее увеличение разрешающей способности путем уменьшения длины волны подсветки уже невозможно, так как УФ-излучение с $\lambda < 200$ нм сильно поглощается не только оптикой микроскопа, но и воздухом.

Многообещающим был бы переход к рентгеновскому диапазону с его короткими длинами волн ($\lambda = 80–10^{-5}$ нм) и проникающим излучением, что сразу на несколько порядков уменьшило бы предел разрешения и позволило рассматривать структуру не только крупных молекул, но и отдельных атомов. Однако здесь существенное препятствие заключается в том, что рентгеновские лучи не преломляются ни одним веществом и поэтому пока не удалось создать линзу для рентгеновских лучей и построить рентгеновский микроскоп.

25.4. Электронный микроскоп

Создание ***электронного микроскопа*** стало возможным после установления волновых свойств микрочастиц, в том числе и электронов. Отправной точкой послужило установление корпускулярно-волнового дуализма света, что привело в итоге к установлению волновых свойств микрочастиц и развитию волновой (квантовой) механики.

Еще И. Ньютон полагал, что свет — это поток частиц, чем объяснял прямолинейность его распространения. Опыты Френеля, Юнга и других ученых по дифракции и интерференции света убедительно доказали, что свет — это волна, характеризующаяся частотой ν, периодом T и длиной волны $\lambda = c/\nu$.

Однако в 1900 г. М. Планк при объяснении законов теплового излучения пришел к выводу, что тела испускают и поглощают световую энергию определенными маленькими порциями — *световыми квантами*, или *фотонами*. Причем энергия светового кванта пропорциональна частоте излучения: $\varepsilon = h\nu$, где $h = 6{,}63 \cdot 10^{-34}$ Дж · с — постоянная Планка.

Используя эти представления, А. Эйнштейн в 1905 г. объяснил все ранее установленные законы внешнего фотоэффекта, которые волновая теория света не могла объяснить на протяжении почти 20 лет.

Таким образом, была доказана двойственная природа света: с одной стороны, свет — это электромагнитная волна с частотой ν и длиной волны $\lambda = c/\nu$, а с другой — поток частиц (фотонов) с энергией ε и импульсом p[1], причем эти характеристики волны связаны через постоянную Планка:

$$\varepsilon = h\nu; \quad p = h/\lambda. \tag{25.10}$$

В 1923 г. французский физик Л. де Бройль предположил, что возможно и обратное: микрочастице с импульсом $p = mv$ (т.е. имеющей массу m и скорость v) соответствует некий волновой процесс с длиной волны

$$\lambda_{Бр} = \frac{h}{p} = \frac{h}{mv}, \tag{25.11}$$

которая затем была названа длиной волны де Бройля.

Удобными частицами для проверки этой гипотезы оказались электроны, так как их масса хорошо известна, а кинетическая энергия зависит от ускоряющего напряже-

[1] Импульс фотона $p = mc = mc^2/c = h\nu/c = h/\lambda$. Здесь учтены соотношения $E = mc^2 = h\nu$, $\lambda = c/\nu$.

ния U, приложенного между анодом и катодом электронной пушки:

$$\frac{1}{2}mv^2 = eU. \qquad (25.12)$$

Умножая обе части уравнения (25.12) на $2m$ и извлекая затем квадратный корень, найдем зависимость импульса и длины волны де Бройля электронов от ускоряющего напряжения:

$$p = mv = \sqrt{2meU}; \qquad (25.13)$$

$$\lambda_{Бр} = \frac{h}{mv} = \frac{h}{\sqrt{2meU}}. \qquad (25.14)$$

Опыты по дифракции электронов на тонкой золотой фольге, игравшей роль дифракционной решетки, доказали наличие у электронов волновых свойств, причем длина волны, вычисленная по наблюдаемой дифракционной картине, соответствовала формуле (25.11) длины волны де Бройля.

Так было установлено, что малые движущиеся частицы обладают волновыми свойствами, и начала развиваться волновая (квантовая) механика микрочастиц.

Поскольку электроны — заряженные частицы, на них можно воздействовать электрическими и магнитными полями, изменяя их траекторию. Это позволило создать ***электронные линзы***, способные фокусировать электронные пучки и создавать увеличенные электронные изображения.

Таким образом, наличие у электронов заряда и волновых свойств с малой длиной волны $\lambda_{Бр}$ позволило создать электронный микроскоп (рис. 25.6), предел разрешения которого в тысячи раз меньше, чем у оптического микроскопа.

Все пространство электронного микроскопа находится в высоком вакууме. Катод *электронной пушки 1* является источником электронов, которые разгоняются под действием высокого напряжения U анода до необходимой скорости и длины волны де Бройля [формула (25.14)]. Этот пучок элек-

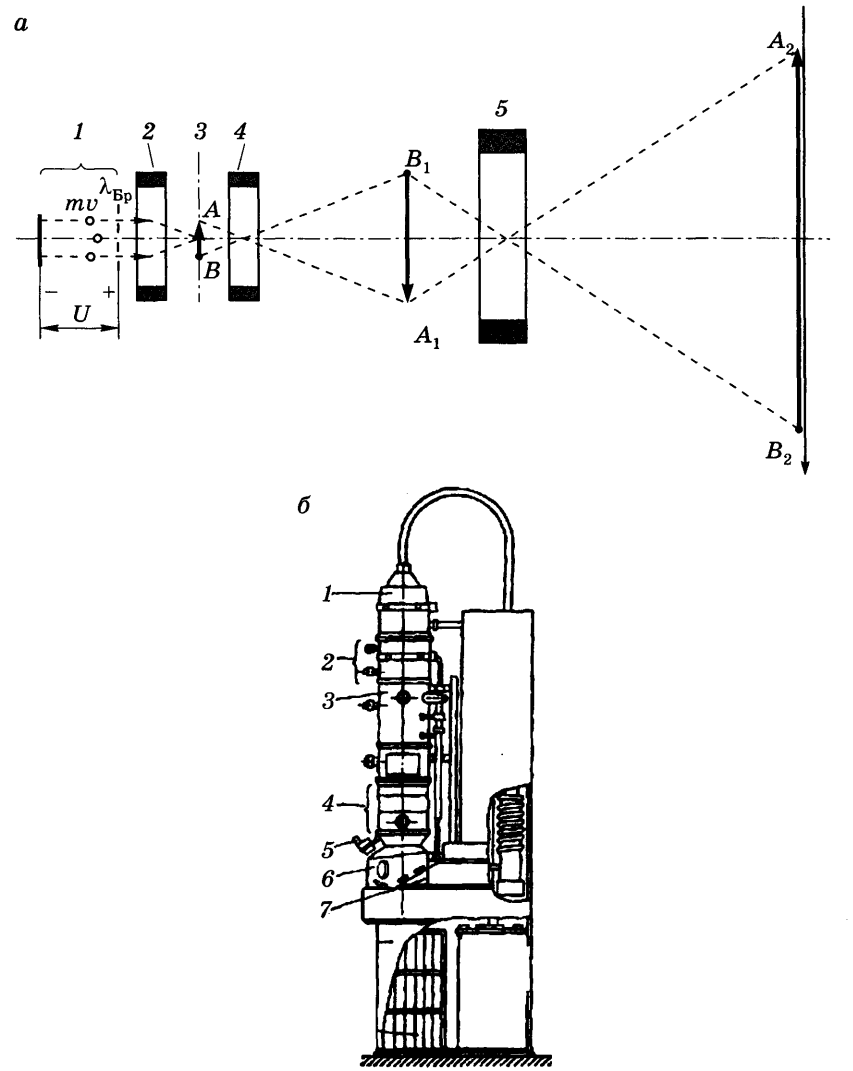


Рис. 25.6. Упрощенная схема (*а*) и внешний вид (*б*) просвечивающего электронного микроскопа (ПЭМ):
1 — электронная пушка; *2* — конденсорные линзы; *3* — линза объектива; *4* — проекционные линзы; *5* — световой микроскоп, дополнительно увеличивающий изображение, наблюдаемое на люминесцентном экране; *6* — тубус со смотровыми окнами; *7* — люминесцентный экран

тронов собирается специальной электронной линзой — *конденсором* на исследуемом образце AB и после рассеивания на нем электроны проходят последовательно через две (иногда и более) электронные линзы *4* и *5*, поочередно создающие увеличенные электронные изображения A_1B_1 и A_2B_2 (см. рис. 25.6). Последнее изображение проецируется на люминесцентный экран *7*, преобразующий электронное изображение в видимое, которое можно наблюдать глазом и фотографировать.

Напряжение в электронной пушке, ускоряющее электроны, достигает 100–200 кВ, регулируется ступенчато и отличается высокой стабильностью: за 1–3 мин оно изменяется не более чем на 1–2 миллионные доли от исходного значения. Величина ускоряющего напряжения определяет толщину объекта, которую можно «просветить» электронным пучком. В 100-киловольтных микроскопах изучают объекты толщиной от одного до нескольких сотен нанометров.

Предел разрешения электронного микроскопа также определяется формулой Аббе:

$$Z_{эл} = \frac{0{,}5\lambda_{Бр}}{A} = \frac{h}{2A\sqrt{2emU}}. \qquad (25.15)$$

Числовая апертура электронного микроскопа обычно мала $A = \sin u \approx 0{,}001$, но за счет использования высокого ускоряющего напряжения U в несколько десятков киловольт (до 200 кВ) и соответственно малой длины волны де Бройля предел разрешения электронного микроскопа может составлять единицы и доли нанометра (до 0,1 нм), что в сотни и тысячи раз меньше, чем у оптического микроскопа.

Основным недостатком электронного микроскопа является сложность приготовления образцов (реплик объекта), которые должны точно передавать форму объекта и быть тонкими, чтобы не поглотить полностью электроны.

Биологические объекты обычно предварительно высушивают или заливают смолой, затем разрезают микротомом на тонкие слои и исследуют.

В настоящее время широко используются ***растровые электронные микроскопы*** (РЭМ), которые могут работать не только на просвет, но и в отраженных электронных пучках. В последнем случае толщина образца жестко не ограничена, так как сканируется лишь поверхность образца.

В растровых микроскопах электронный луч, излучаемый пушкой *1*, фокусируется конденсором *2* в точку на образце, и с помощью специальных магнитных линз этот сфокусированный электронный луч сканирует образец *3* (как бы построчно «прочитывает» его) и на специальном экране (мониторе) отображается увеличенный рельеф поверхности микрообъекта.

Кроме того, при торможении электронов на поверхности образца возникает характеристическое рентгеновское излучение (см. гл. 27), которое дает информацию как о структуре, так и о элементном составе каждой точки поверхности образца. Это позволяет выяснить распределение различных атомов по исследуемой поверхности.

Растровые электронные микроскопы предназначены для исследования массивных объектов. Предел их разрешения хуже, чем у просвечивающих электронных микроскопов, и составляет от 5 до 20 нм, а ускоряющие напряжения — до 30–50 кВ.

25.5. Сканирующая зондовая микроскопия

Сканирующие зондовые микроскопы представляют собой сравнительно новый (1981 г.) класс приборов, предназначенных для исследования поверхности образца и ее локальных характеристик.

Процесс построения изображения основан на сканировании («построчном считывании») рельефа поверхности образца специальным сверхтонким зондом-иглой, при этом расстояние между зондом и поверхностью выдерживается постоянным. Компьютер регистрирует величину вертикального перемещения зонда над каждой точкой поверхности, после чего вычисляет и показывает на мониторе трехмерное изображение поверхности с разрешением в единицы нанометра.

В зависимости от физического способа, каким поддерживается неизменным расстояние между зондом и образцом, сканирующие зондовые микроскопы подразделяются на туннельные микроскопы и атомно-силовые микроскопы.

В *сканирующем туннельном микроскопе* (*СТМ*) расстояние между иглой-зондом и электропроводящим образцом выдерживается таким, чтобы обеспечить постоянство туннельного тока через воздушный зазор между зондом и образцом, предполагая, что при этом будет выдерживаться постоянной и величина этого зазора. Такой микроскоп позволяет получить информацию о топографии (форме) проводящей поверхности и электрических свойствах образца.

В *атомно-силовом микроскопе (АСМ)* исследуются в основном диэлектрические образцы, в том числе и биологические объекты, которые могут исследоваться в естественной для них среде — жидкой или воздушной. Расстояние между зондом (здесь его называют *кантилевером*) АСМ и поверхностью образца поддерживается таким, чтобы сила межатомного взаимодействия между кончиком кантилевера и атомами поверхности образца была неизменной, что обеспечивает практически постоянную величину зазора между ними.

Атомно-силовой микроскоп (рис. 25.7) устроен так, что при сканировании образца в горизонтальной плоскости XY

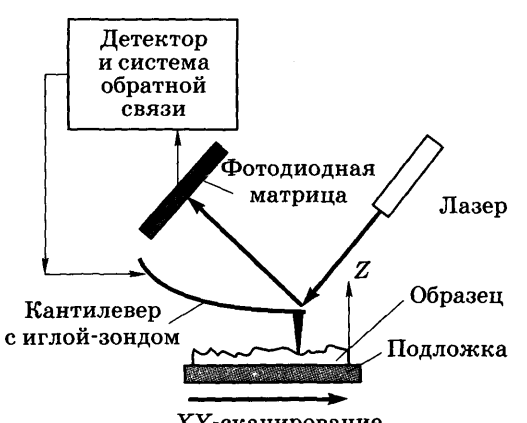


Рис. 25.7. Принципиальная схема АСМ

изменения рельефа поверхности приводят к вертикальным Z-смещениям зонда, которые считываются следящей системой микроскопа (лазерный луч, отраженный от поверхности кантилевера, попадает на фотодиодную матрицу, по сигналу от которой определяется величина Z-смещения зонда).

Компьютер проводит математическую обработку полученных результатов и по специальной программе вычисляет рельеф поверхности образца, т.е. зависимость $Z(X, Y)$, и отображает на дисплее либо трехмерное изображение исследуемой поверхности, либо ее сечения заданными плоскостями. Предел пространственного разрешения зондовых микроскопов зависит от кривизны кончика иглы-зонда и может достигать 0,1 нм, что сопоставимо с размерами отдельных атомов

В настоящее время сканирующие зондовые микроскопы нашли применение практически во всех областях естественных наук: в физике, химии, биологии, медицине и др. Перспективным полагают совмещение сканирующих зондовых микроскопов с другими традиционными и современными методами исследования и создание на их основе принципиально новых приборов.

Контрольные вопросы и задания

1. Что называют линзой? Каковы их виды и свойства? Что такое главная оптическая ось, передний и задний фокусы, оптическая сила?
2. Как построить изображение предмета линзой? Что такое действительные и мнимые изображения?
3. Приведите формулу тонкой линзы. Как рассчитать линейное увеличение, даваемое линзой?
4. Какие изображения и при каких условиях могут создавать собирающая и рассеивающая линзы?
5. Что называют лупой? Какое изображение она дает, как его построить? Какое увеличение дает лупа?
6. Постройте ход лучей и изображение в оптическом микроскопе.
7. От чего зависит и как вычисляется увеличение микроскопа?
8. Чем определяется предел разрешения микроскопа? Приведите формулу Аббе. Каков предел разрешения оптических микроскопов?

9. В чем состоит корпускулярно-волновой дуализм? Как связаны корпускулярные и волновые характеристики света?
10. В чем заключаются волновые свойства микрочастиц? Что такое длина волны де Бройля?
11. На каких физических явлениях основан электронной микроскоп?
12. Как определяется предел разрешения электронного микроскопа?
13. Приведите принципиальную схему электронного микроскопа.
14. Каков принцип действия сканирующих зондовых микроскопов? Что они исследуют? Каковы их разновидности?
15. Каков предел разрешения сканирующего атомно-силового микроскопа? Каковы его преимущества при исследовании биообъектов?

Глава 26 Электронный парамагнитный и ядерный магнитный резонанс

Явления ***электронного парамагнитного резонанса*** (ЭПР) и ***ядерного магнитного резонанса*** (ЯМР) широко используются в современной физике, химии, биологии и медицине при исследовании процессов, протекающих с участием парамагнитных молекул и ядер. Кроме того, ядерный магнитный резонанс является физической основой наиболее мощного современного метода получения изображений органов и тканей человека — ***магнитно-резонансной томографии*** (МРТ).

Метод ЭПР приобрел большое значение в химии и биологии прежде всего благодаря способности обнаруживать и идентифицировать свободные радикалы в химических и биологических системах. При этом с высокой точностью определяются не только вид и концентрация свободных радикалов, но и кинетика биохимических реакций, протекающих с образованием свободных радикалов как в промежуточных, так и в конечных стадиях реакции.

26.1. Свободные радикалы в биологических системах

Известно, что в соответствии с принципом Паули в каждом квантовом состоянии молекулы может находиться не более двух электронов, спины которых должны быть противоположно ориентированы (скомпенсированы). ***Спин*** — это внутреннее свойство электрона, которое проявляется в наличии у него собственного механического момента $\vec{J}$, т.е. электрон представляет собой как бы «закрученный» волчок. Для устойчивых молекул обычно характерно четное число электронов и каждая пара электронов на любом энергетическом уровне имеет противоположно направленные, или, как говорят, скомпенсированные (спаренные) спины.

Однако есть соединения, у которых число электронов нечетное и тогда у одного из валентных электронов спин не бу-

дет скомпенсирован. Такая же ситуация возникает и в том случае, если у стабильного соединения отнять либо, наоборот, добавить ему один электрон. Тогда спин одного из электронов тоже будет не скомпенсирован.

Молекула или ее часть, имеющая неспаренный электрон, называется ***свободным радикалом.***

С точки зрения химии наличие в молекуле неспаренного электрона есть не что иное, как наличие у нее свободной валентности. Поэтому свободные радикалы очень активны в химическом отношении. Они легко вступают в химические связи с другими молекулами и химическими соединениями, что влияет на протекание многих процессов в биологических системах.

Наиболее важную роль в биологических системах играют следующие виды радикалов (радикал часто обозначается точкой над соответствующей химической группой):

- свободные радикалы воды: $\dot{O}H$ — гидроксильный, $H\dot{O}_2$ — перекисный, O_2^- — супероксид;
- свободные радикалы органических молекул, образующиеся при действии ионизирующей и ультрафиолетовой радиации:

$$AH + h\nu \rightarrow AH^* \rightarrow A\dot{H}^+ + e^- \rightarrow \dot{A} + H^+ + e^-,$$

где e^- — сольватированный электрон, а образующийся радикал обозначен точкой сверху.

Эти свободные радикалы играют важную роль в возникновении радиационного повреждения тканей и органов, а также при УФ-ожогах;

- свободные радикалы хинонов участвуют в окислительно-восстановительных реакциях организма;
- свободные радикалы липидов могут образовываться в определенных условиях при окислении их жирных кислот. Наличие свободных радикалов в липидах биологических мембран ведет к нарушению их проницаемости для ионов и других молекул, что приводит к развитию той или иной патологии в организме. Примером таких патологий может служить развитие УФ-эритемы кожи, световых ожогов глаз и др.

Основное физическое отличие свободных радикалов от других молекул состоит в том, что свободные радикалы парамагнитны, т.е. обладают собственным магнитным моментом, тогда как стабильные молекулы его не имеют, т.е. они диамагнитны. Именно это различие в магнитных свойствах и позволяет обнаруживать свободные радикалы среди диамагнитных молекул.

Основным физическим методом изучения свободных радикалов в биологических системах является *электронный парамагнитный резонанс* (ЭПР). Метод ЭПР получил большое распространение в биологии и медицине именно благодаря его способности определять наличие и вид свободных радикалов в биологических системах *in vivo*, исследовать кинетику биохимических реакций с их участием и др.

При этом очень важно, что метод этот неинвазивный, безвредный и позволяет исследовать процессы, протекающие в живых организмах, не внося никаких изменений в эти процессы.

26.2. Электронный парамагнитный резонанс

Явление *ЭПР*, открытое в 1944 г. советским физиком Е.К. Завойским, обусловлено тем, что молекулы парамагнитных веществ обладают собственным магнитным моментом и благодаря этому могут взаимодействовать с внешним магнитным полем. Такие взаимодействия ведут к изменениям в структуре энергетических уровней молекулы, что можно зарегистрировать специальной спектральной аппаратурой.

Парамагнитные свойства свободных радикалов, а также других парамагнитных молекул обусловлены наличием у них неспаренных электронов. Известно, что электрон обладает собственным механическим моментом $\vec{J}$ (спином), а поскольку электрон электрически заряжен, это приводит к наличию у него и соответствующего магнитного момента:

$$\vec{\mu}_s = -\frac{e}{m_e}\vec{J}, \qquad (26.1)$$

где e и m_e — заряд и масса электрона соответственно.

Этот неспаренный электрон одновременно движется по своей орбите в атоме или молекуле и, как любой круговой ток, создает орбитальный магнитный момент. В результате полный магнитный момент электрона в атоме или молекуле

$$\vec{\mu}_e = -g\frac{e}{2m_e}\vec{J}. \qquad (26.2)$$

Множитель g, называемый *фактором Ланде*, зависит от вида свободного радикала и для большинства изученных радикалов принимает значения от 2,0000 до 2,0060.

Если бы в молекуле все электроны были спарены, то их магнитные моменты были бы попарно противоположны и скомпенсировали друг друга. Если атом или молекула имеет неспаренный электрон, подобно свободным радикалам, то его магнитный момент не скомпенсирован, и молекула в целом будет иметь соответствующий магнитный момент, т.е. обладать парамагнитными свойствами.

В отсутствие внешнего магнитного поля этот факт никак не проявляется, поскольку энергия электрона в атоме или молекуле в этом случае не зависит от величины и ориентации этого магнитного момента в пространстве и имеет вполне определенную величину E_0.

Если же поместить такую молекулу во внешнее магнитное поле индукцией $\vec{B}$, то ее магнитный момент начнет взаимодействовать с этим внешним полем, стремясь сориентироваться относительно него определенным образом. При этом, согласно квантовой механике, магнитный момент электрона в простейшем случае может занять лишь два возможных положения: либо по направлению поля ($\vec{\mu}_e \uparrow\uparrow \vec{B}$), либо противоположно ему ($\vec{\mu}_e \downarrow\uparrow \vec{B}$). Если говорить более точно, то квантово-механическое рассмотрение требует, чтобы проекция механического момента электрона на направление внешнего магнитного поля принимала лишь определенные значения:

$$J_B = \pm\frac{1}{2}\frac{h}{2\pi} = \frac{h}{2\pi}m_s, \qquad (26.3)$$

где $m_s = \pm\frac{1}{2}$ — магнитное спиновое квантовое число электрона.

Энергия взаимодействия магнитного момента электрона с внешним магнитным полем индукцией B определяется величиной

$$E_{\text{вз}} = (\vec{\mu}_e)_B B = -g\frac{e}{2m_e}J_B B = \pm\frac{1}{2}g\mu_{\text{Б}}B\,. \qquad (26.4)$$

Постоянная величина $\mu_{\text{Б}} = \dfrac{eh}{4\pi m_e} = 9{,}274\cdot 10^{-24}\ \text{А}\cdot\text{м}^2$ называется *магнетоном Бора*.

Теперь полная энергия электрона в молекуле в магнитном поле

$$E = E_0 + E_{\text{вз}} = E_0 \pm \frac{1}{2}g\mu_{\text{Б}}B\,. \qquad (25.5)$$

Знак «+» в этом уравнении соответствует ориентации магнитного момента электрона против направления поля, а знак «−» — вдоль поля. Таким образом, внесение молекулы во внешнее магнитное поле индукцией $\vec{B}$ приводит к «расщеплению» прежнего энергетического уровня E_0 на два новых уровня энергии:

$$E_1 = E_0 - \frac{1}{2}g\mu_{\text{Б}}B\ (\uparrow\uparrow), \quad E_2 = E_0 + \frac{1}{2}g\mu_{\text{Б}}B\ (\downarrow\uparrow).$$

Величина этого расщепления прямо пропорциональна величине внешнего магнитного поля:

$$\Delta E = E_2 - E_1 = g\mu_{\text{Б}}B\,. \qquad (26.6)$$

Таким образом, основной энергетический уровень парамагнитных ионов и молекул во внешнем магнитном поле расщепляется на два энергетических состояния E_1 и E_2, отличающихся противоположным направлением спина валентного электрона (рис. 26.1).

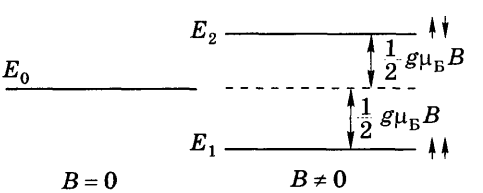


Рис. 26.1. Расщепление электронного уровня парамагнитной молекулы в магнитном поле

Большая часть молекул (ионов) при этом будет находиться в нижнем энергетическом состоянии E_1, энергетически более выгодном.

Если теперь облучать такую систему электромагнитным полем резонансной частоты (энергия кванта равна энергии перехода: $h\nu = \Delta E$)

$$\nu = \frac{E_2 - E_1}{h} = \frac{1}{h} g\mu_{\text{Б}} B, \qquad (26.6\text{а})$$

то будут наблюдаться как переходы снизу вверх, т.е. с уровня E_1 на E_2, соответствующие поглощению кванта электромагнитного поля, так и переходы сверху вниз (с уровня E_2 на E_1), соответствующие испусканию такого же кванта.

И те и другие переходы сопровождаются переориентацией спина электрона и соответственно магнитного момента электрона.

Поскольку на нижнем энергетическом уровне E_1 всегда находится больше молекул, чем на верхнем, то поглощение будет преобладать над испусканием и в целом будет наблюдаться явление резонансного поглощения электромагнитного излучения частоты (26.6а).

Явление резонансного поглощения электромагнитного излучения атомными или молекулярными парамагнитными системами, находящимися в магнитном поле, которое сопровождается переориентацией спина электрона, называют ***электронным парамагнитным резонансом*** (ЭПР).

Установка для наблюдения ЭПР (рис. 26.2) действует следующим образом.

Электромагниты *1* создают постоянное магнитное поле индукцией $\vec{B}$, величина которого может изменяться по желанию экспериментатора (за счет изменения тока в обмотках электромагнита). В это поле помещается исследуемый образец *2*. Источник *3* облучает образец электромагнитным полем фиксированной частоты ν (обычно СВЧ-диапазона), приемник *4* определяет интенсивность прошедшего излучения, и система регистрации *5* записывает ЭПР-спектр.

Поскольку частота используемого в установке электромагнитного излучения строго фиксирована, то выполнения усло-

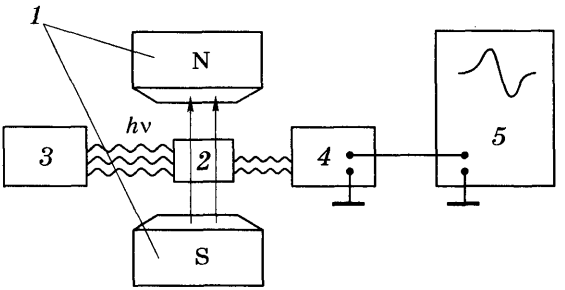


Рис. 26.2. Схема установки для наблюдения ЭПР:
1 — полюса электромагнита; *2* — исследуемый образец;
3 — источник СВЧ-излучения; *4* — приемник излучения;
5 — регистратор ЭПР-спектра

вия (26.6) резонансного поглощения добиваются изменением индукции B магнитного поля: увеличивая ток в обмотках электромагнитов, увеличивают индукцию B налагаемого на образец магнитного поля и, следовательно, величину расщепления ΔE основного энергетического состояния парамагнитных молекул. Когда величина этого расщепления становится равной энергии электромагнитного кванта $h\nu$, наблюдается резкое возрастание коэффициента поглощения K образца (рис. 26.3), что и фиксируется приемником излучения.

Однако по техническим причинам удобнее регистрировать не зависимость $K(B)$, изображенную на рис. 26.3, а зависимость ее первой производной dK/dB от B (рис. 26.4).

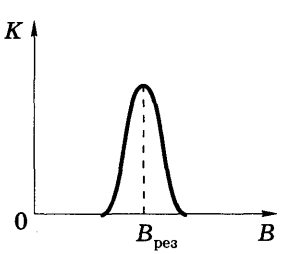


Рис. 26.3. Зависимость коэффициента поглощения образца от индукции магнитного поля

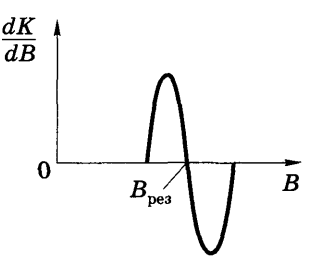


Рис. 26.4. Вид ЭПР-спектра

Спектр ЭПР позволяет получить следующую информацию:

1. Каждый свободный радикал характеризуется своим значением фактора Ланде, что позволяет уверенно определять вид свободных радикалов по значению $B_{рез}$, т.е. по положению линий в ЭПР-спектре. Действительно, частота электромагнитного излучения известна и строго фиксирована, поэтому, зная $B_{рез}$, можно вычислить g-фактор Ланде исследуемого радикала из условия (26.6):

$$g = \frac{h\nu}{\mu_Б B_{рез}}, \qquad (26.7)$$

и определить тем самым вид этого свободного радикала (парамагнитной молекулы).

2. Площадь S под линией ЭПР позволяет судить о количестве N свободных радикалов (парамагнитных молекул) в образце:

$$N = \alpha S,$$

где α — коэффициент пропорциональности, известный для данной ЭПР-установки.

3. Форма ЭПР-линий также несет информацию о физических свойствах среды, в которой находится исследуемый радикал: чем шире линия, тем более вязкая среда, и наоборот. Таким способом оценивают вязкость липидного бислоя и других клеточных структур.

Взаимодействие электронов между собой и с магнитным полем ядра проявляется в виде *расщепления линий ЭПР* на несколько близких компонент. В спектрах ЭПР имеются два типа таких расщеплений:

- электронное расщепление — обусловлено наличием у атома или молекулы не одного, а двух и более неспаренных электронов;
- сверхтонкое расщепление — вызвано взаимодействием магнитного момента электрона $\vec{\mu}_e$ не только с внешним магнитным полем $\vec{B}$, но и с магнитным полем ядра атома.

С помощью метода ЭПР изучают поведение в биологических системах не только свободных радикалов, но и других биоло-

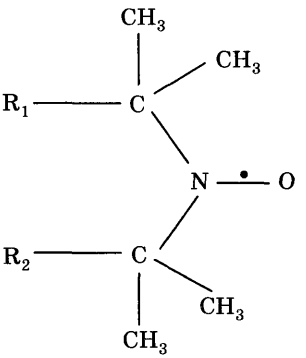


Рис. 26.5. Структурная формула иминоксильного радикала

гически важных молекул — белков, липидов, нуклеиновых кислот, предварительно превратив их в парамагнитные зонды, поскольку сами они обычно непарамагнитны.

Парамагнитный зонд представляет собой соответствующий белок или липид, в котором одна из групп замещена на устойчивый известный свободный радикал, называемый *парамагнитной* (*спиновой*) *меткой*.

Основу многих спиновых меток составляет иминоксильный радикал (рис. 26.5), парамагнитные свойства которого определяются неспаренным электроном, локализованным в группе N–O. Форма сигнала ЭПР от такого зонда зависит от вращательной подвижности той химической группы или макромолекулы в целом, в состав которой он входит. Чем сильнее заторможено такое вращение, тем шире сигнал ЭПР. На интенсивность сигнала влияет также полярность окружающей среды (растворителя).

Метод ЭПР обладает достаточной чувствительностью и избирательностью для непосредственного наблюдения свободных радикалов, возникающих при протекании многих биологически важных реакций: окислительно-восстановительных, радиационного поражения, фотосинтеза и др.

26.3. Ядерный магнитный резонанс

Явление ядерного магнитного резонанса (ЯМР) широко используется в физике, химии, биологии и медицине для определения строения органических соединений, липидных и белковых структур, а также для получения компьютерных изображений органов и тканей человека и других живых организмов.

Это явление обусловлено наличием собственного магнитного момента у ядра атома. Существование такого магнитного момента связано с тем, что протоны и нейтроны, состав-

ляющие ядра атомов, обладают собственными магнитными моментами $\vec{\mu}_p$ и $\vec{\mu}_n$. Полный магнитный момент ядра ${}^{A}_{Z}X$ равен векторной сумме магнитных моментов всех его нуклонов:

$$\vec{\mu}=\sum_{i=1}^{Z}\vec{\mu}_{pi}+\sum_{i=1}^{A-Z}\vec{\mu}_{ni}\ . \qquad (26.8)$$

Для большинства естественных ядер магнитные моменты нуклонов скомпенсированы, так как они попарно направлены в противоположные стороны, в результате чего суммарный магнитный момент таких ядер равен нулю. Однако у ядер с нечетным числом нуклонов полный магнитный момент не равен нулю. К ним относятся: ${}^{1}H$, ${}^{2}H$, ${}^{13}C$, ${}^{19}F$, ${}^{31}P$ и др.

Если такое парамагнитное ядро поместить во внешнее магнитное поле индукцией $\vec{B}$, то магнитный момент $\vec{\mu}$ ядра будет стремиться сориентироваться вполне определенным образом относительно направления внешнего магнитного поля, в простейшем случае либо вдоль поля, либо против него. В первом случае полная энергия ядра уменьшится, а во втором — увеличится на $\frac{1}{2}g_{я}\mu_{я}B$, где $g_{я}$ — ядерный g-фактор, величина которого для ядер одного вида всегда одна и та же,

$\mu_{я}=\frac{eh}{4\pi m_p}=5{,}049\cdot10^{-27}\ \text{А}\cdot\text{м}^2$ — *ядерный магнетон*, отличающийся от магнетона Бора тем, что масса электрона заменена на массу протона (поэтому он меньше на три порядка). Это приводит, как и в случае ЭПР, к расщеплению прежнего энергетического уровня E_0 ядра на два новых подуровня E_1 и E_2 (см. рис. 26.1). Величина расщепления подуровней энергии ядра линейно растет с увеличением индукции внешнего магнитного поля:

$$\Delta E=E_2-E_1=g_{я}\mu_{я}B.$$

Если теперь эти ядра, находящиеся во внешнем магнитном поле $\vec{B}$, облучить электромагнитным излучением резонансной частоты

$$\nu_0=\frac{\Delta E}{h}=\frac{1}{h}g_{я}\mu_{я}B, \qquad (26.9)$$

то, как и в случае ЭПР, произойдет его поглощение.

Явление резонансного поглощения электромагнитного излучения парамагнитными ядрами, находящимися в магнитном поле, которое сопровождается переориентацией спина ядра, называют ***ядерным магнитным резонансом*** (ЯМР).

Схема установки для наблюдения ЯМ-резонанса полностью аналогична схеме ЭПР-спектроскопии (см. рис. 26.2). Она отличается лишь конструкция излучателя и детектора электромагнитного поля, так как резонансная частота ν_0 в ЯМР-спектроскопии на 2–3 порядка меньше частоты, применяемой в ЭПР (это обусловлено тем, что ядерный магнетон $\mu_я$ на три порядка меньше магнетона Бора $\mu_Б$). В установках для ЯМР-спектроскопии частота электромагнитного излучения тоже строго фиксирована, а резонансного поглощения добиваются путем изменения величины внешнего магнитного поля $\vec{B}$.

ЯМР-спектр представляет собой зависимость коэффициента поглощения K образца от величины внешнего поля B (см. рис. 26.3).

Информация ЯМР-спектров та же, что и в методе ЭПР:

1. Площадь S под линией поглощения в спектре ЯМР пропорциональна количеству N исследуемых ядер в образце: $N = \alpha S$.

2. Ширина линий ЯМР несет информацию о физических свойствах среды, в которой находятся исследуемые ядра. Как правило, с увеличением вязкости этой среды линии в ЯМР-спектре расширяются, что позволяет количественно оценивать микровязкость биологических систем (мембран и т.д.) по спектрам ЯМР. По этой же причине в ЯМР-спектрах твердых тел наблюдаются широкие линии поглощения, а в спектрах жидкостей — узкие линии.

На рис. 26.6 приведены спектры ЯМР на ядрах водорода (протонах) для воды и льда (тот же образец, но заморожен). Таким путем оценивают вязкость липидных слоев в биологических мембранах, вязкость цитоплазмы и других биологических структур.

В формуле (26.9) величины h и $\mu_я$ — константы, ядерный фактор $g_я$ для данного типа ядер постоянен, казалось бы, одинаковые ядра (например, протоны) должны иметь одну и ту же линию в спектре ЯМР, независимо от того, в какие

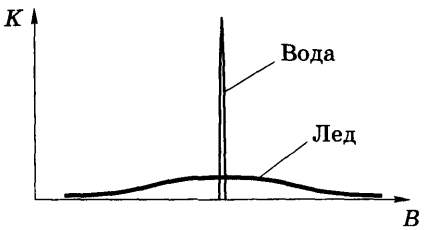


Рис. 26.6. Сравнительные ЯМР-спектры воды и льда

химические соединения эти ядра входят. На самом деле это не так. Дело в том, что на ядро атома воздействует не только внешнее магнитное поле $\vec{B}_0$, но и магнитное поле $\vec{B}'$, создаваемое движущимися по орбитам электронами и ионами, поэтому частота линий поглощения ЯМР смещается относительно ожидаемой частоты ν_0 (26.9) на величину $\Delta\nu$, зависящую от того электронного и ионного окружения, в котором находится изучаемое ядро:

$$\nu = \frac{1}{h} g_я \mu_я (B_0 + B') = \frac{g_я}{h}\mu_я B_0 + \frac{g_я}{h}\mu_я B' = \nu_0 + \Delta\nu = \nu_0\left(1 + \frac{\Delta\nu}{\nu_0}\right). \quad (26.10)$$

Величина внутреннего магнитного поля B', в котором находится ядро, зависит от химического строения вещества, поэтому относительный сдвиг $\Delta\nu / \nu_0$ частоты электромагнитного излучения или индукции резонансного поля B'/B_0 называют ***химическим сдвигом*** δ. Поскольку он очень мал, его умножают на 10^6 и измеряют в миллионных долях (м.д.):

$$\delta = \frac{\Delta\nu}{\nu_0} \cdot 10^6 = \frac{B'}{B_0} \cdot 10^6. \quad (26.11)$$

Из-за наличия химического сдвига разным химическим группам, содержащим одни и те же парамагнитные ядра (например, протоны OH, CH_2, CH_3 и т.д.), соответствуют разные резонансные линии в спектре ЯМР:

$$B_{рез} = B_0 + B' = B_0\left(1 + \frac{B'}{B_0}\right). \quad (26.12)$$

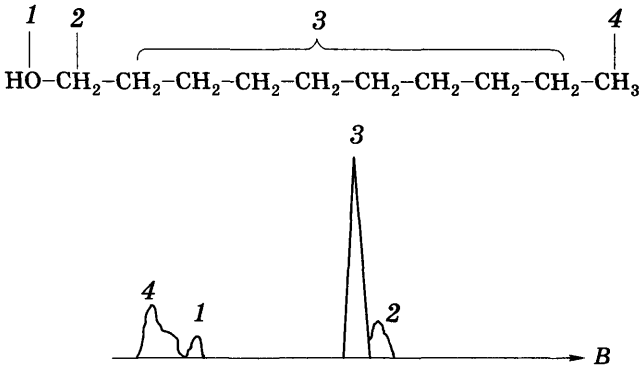


Рис. 26.7. Структурная формула деканола и его ЯМР-спектр. Цифрами указаны группы протонов и соответствующие им ЯМР-линии

Это хорошо видно на примере ЯМР-спектра деканола (рис. 26.7).

Благодаря химическому сдвигу в спектрах ЯМР, метод ядерного магнитного резонанса является важнейшим методом определения структуры и строения молекул в биологии и физической химии.

Поскольку биологические соединения — это соединения органические, то в них очень много ядер водорода ^{1}H (протонов). Это хорошо с точки зрения интенсивности ЯМР-линий, но из-за химического сдвига количество линий очень велико, что затрудняет их идентификацию, т.е. расшифровку ЯМР-спектров. Поэтому ЯМР-спектры биологических объектов часто регистрируют на ядрах изотопов ^{13}C и ^{31}P, которые тоже входят во все важные биологические соединения, но в значительно меньших количествах. Наиболее распространенные в естественных условиях изотопы углерода ^{12}C и фосфора ^{30}P не обладают магнитным моментом, поэтому на них невозможно наблюдать явление ЯМР.

Основные преимущества таких спектров в том, что ЯМР-линии получаются более узкими, а химические сдвиги увеличиваются, что улучшает расшифровку ЯМР-спектров на ядрах ^{13}C и ^{31}P по сравнению с протонным резонансом. Однако содержание этих изотопов в естественных соединениях неве-

лико, в результате чего сильно уменьшается интенсивность ЯМР-линий и для их регистрации необходимо использовать более чувствительные и более дорогие ЯМР-установки, что является их недостатком.

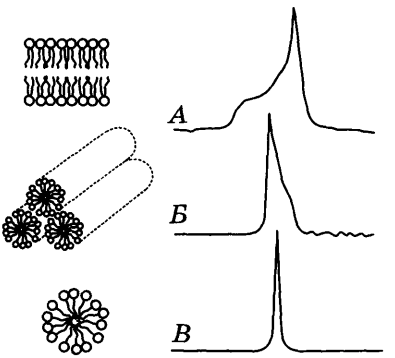


Рис. 26.8. Спектры фосфорного ЯМР при разной упаковке липидов

Например, в молекулу фосфолипида входит только один атом фосфора и одна из особенностей фосфорного ЯМР состоит в том, что форма ЯМР-линий зависит от упаковки липидов и от условий окружения, в которых находится парамагнитное ядро ^{31}P. В качестве примера на рис. 26.8 приведены спектры фосфорного ЯМР для бислоя липидов *А*, гексагональной структуры липидов *Б* и мицелл *В*.

Основной недостаток метода ЯМР — малая чувствительность. Она примерно в 10^3 раз меньше, чем у метода ЭПР, и соответственно в 10^5 раз меньше, чем у флуоресцентных методов спектрального анализа.

26.4. Принципы магнитно-резонансной томографии

Магнитно-резонансная томография (***МРТ*** или ***ЯМР-интроскопия***) — современный и наиболее информативный метод получения изображений тела пациента в различных сечениях, основанный на регистрации сигналов ядерного магнитного резонанса от протонов биологических тканей.

Количество протонов в разных молекулах и тканях различно, значит, в разных тканях плотность протонов (их количество в единице объема) тоже разная, поэтому, измерив методом ЯМР-резонанса распределение плотности протонов в исследуемом объекте в разных пространственных направлениях (вдоль осей *OX*, *OY*, *OZ*), можно вычислить концентрацию протонов в каждой точке объекта, получая таким образом его протонное изображение.

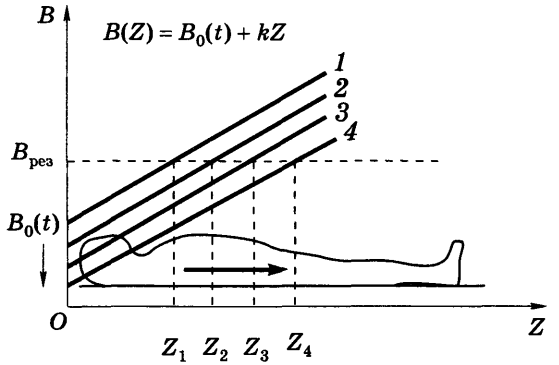


Рис. 26.9. Упрощенная схема регистрации распределения плотности протонов (интенсивности ЯМР-сигналов) вдоль направления OZ

Упрощенно принцип регистрации распределения протонов в выбранном направлении представлен на рис. 26.9. Исследуемая область тела пациента помещается в сильное магнитное поле, индукция которого линейно изменяется в выбранном направлении OZ, например возрастает по закону $B(Z) = B_0 + kZ$, где B_0 — постоянная составляющая магнитного поля. Такие поля называют *градиентными*, так как они равномерно изменяются в заданном направлении и поэтому имеют не нулевой градиент: $\text{grad}B = dB/dZ = k$.

В таком магнитном поле величина расщепления энергетических уровней протонов будет линейно изменяться с координатой Z и при облучении объекта радиоизлучением фиксированной частоты поглощать это излучение и давать сигнал ЯМР-поглощения будут только протоны, имеющие одну и ту же определенную координату Z, т.е. находящиеся в одной плоскости $Z_1 = \text{const}$, перпендикулярной оси OZ. При этом условие резонанса принимает вид

$$h\nu = g_я \mu_я (B_0 + kZ), \qquad (26.13)$$

откуда

$$Z_{рез} = \frac{1}{k}\left(\frac{h\nu}{g_я \mu_я} - B_0\right). \qquad (26.14)$$

Постепенно изменяя постоянную составляющую магнитного поля B_0 (например, уменьшая ее, как на рис. 26.9) или меняя зондирующую частоту ν, можно последовательно зафиксировать интенсивность ЯМР-сигналов от сечений образца плоскостями с координатами Z_1, Z_2, Z_3, Z_4 и т.д., регистрируя таким образом распределение плотности протонов $\rho(Z)$ в образце вдоль направления OZ. Затем создаются градиентные магнитные поля вдоль направлений OX и OY и аналогичным образом регистрируется распределение $\rho(X)$ и $\rho(Y)$ плотности протонов вдоль этих направлений.

Компьютер на основании полученных данных вычисляет распределение плотности протонов $\rho(X, Y, Z)$ в объекте, определяя таким образом его «протонное» изображение. Затем врач-оператор задает компьютеру координаты интересующего его сечения исследуемого объекта и на экране дисплея отображается распределение плотности протонов в этом сечении, которое представляет собой черно-белое или псевдоцветное изображение анатомических структур, попавших в это сечение. Подобное МРТ-изображение называют *взвешиванием по плотности протонов*.

Более контрастные и информативные изображения дают «взвешивания» по временам *продольной* T_1 или *поперечной* T_2 *релаксации* магнитных моментов протонов. В постоянном магнитном поле магнитные моменты протонов вынуждены ориентироваться вдоль направления этого поля. Если теперь специальным внешним кратковременным воздействием отклонить магнитные моменты протонов от направления вектора индукции внешнего поля $\vec{B}$ (оно направлено, например, вдоль оси Z), то они начнут возвращаться в прежнее положение, при этом продольная компонента Mz восстанавливает свою начальную величину M_0 с постоянной времени T_1 (время продольной релаксации), а поперечные компоненты Mx и My экспоненциально затухают с постоянной времени T_2 (время поперечной релаксации):

$$\left.\begin{aligned} Mz &= M_0(1 - e^{t/T1}), \\ Mx &= M_0 e^{-t/T2}\cos(\omega_0 t), \\ My &= M_0 e^{-t/T2}\sin(\omega_0 t). \end{aligned}\right\} \qquad (26.15)$$

Наличие функций $\cos(\omega_0 t)$ и $\sin(\omega_0 t)$ у поперечных составляющих связано с тем, что вектор магнитного момента M_0 протонов вращается (*ларморова прецессия*) с резонансной частотой ω_0 (*ларморова частота*), описывая коническую поверхность вокруг направления поля $\vec{B}$, поэтому его проекции Mx и My на плоскость XOY изменяются с этой частотой по гармоническому закону, одновременно затухая по амплитуде.

Специальной зондирующей последовательностью радиочастотных импульсов определяют времена релаксации T_1 и T_2 для каждого Z-сечения образца, перпендикулярного направлению индукции магнитного поля.

Аналогичным образом регистрируют времена релаксации вдоль направлений OX и OY, предварительно создавая градиентные магнитные поля вдоль этих направлений. После этого компьютор по специальной программе вычисляет распределение T_1 и T_2 по объему образца. Каждому значению времени релаксации присваивается определенное значение серой (или цветной) шкалы, и компьютер строит черно-белое (или псевдоцветное) изображение объекта в заданном врачом сечении. Врач-оператор может регулировать контраст получаемого изображения, изменяя используемый диапазон серой или цветной шкалы.

Время продольной релаксации T_1 (его называют также временем *спин-решеточной релаксации*) всегда больше времени поперечной T_2 (*спин-спиновой*) релаксации. Важно, что длительность обоих процессов существенно зависит от плотности среды и физико-химических свойств ближайшего окружения исследуемых протонов. Поэтому даже небольшие изменения в структуре и свойствах этого окружения ведут к заметным изменениям времен релаксации T_1 и T_2 и видны на МРТ-изображениях.

В табл. 26.1 представлены значения времен релаксации протонов для разных тканей при резонансной частоте 20 МГц.

Показатели времени существенно различаются для разных тканей, что обеспечивает высокий уровень контрастности МРТ-изображения и позволяет хорошо видеть структуры головного и спинного мозга, распознавать патологические процессы в мышцах, сухожилиях и суставах, поражения

костного мозга при заболеваниях крови, злокачественные и доброкачественные новообразования костей и мягких тканей даже без использования контрастных веществ (рис. 26.10).

Таблица 26.1

Ткань	T_1, мс	T_2, мс
Цельная кровь	900	200
Мышца	500	35
Жир	200	60
Вода	3000	3000

ЯМР-томография дает уникальную информацию о состоянии кровеносных сосудов, позволяет четко детализировать анатомические структуры средостения и паренхимы лег-

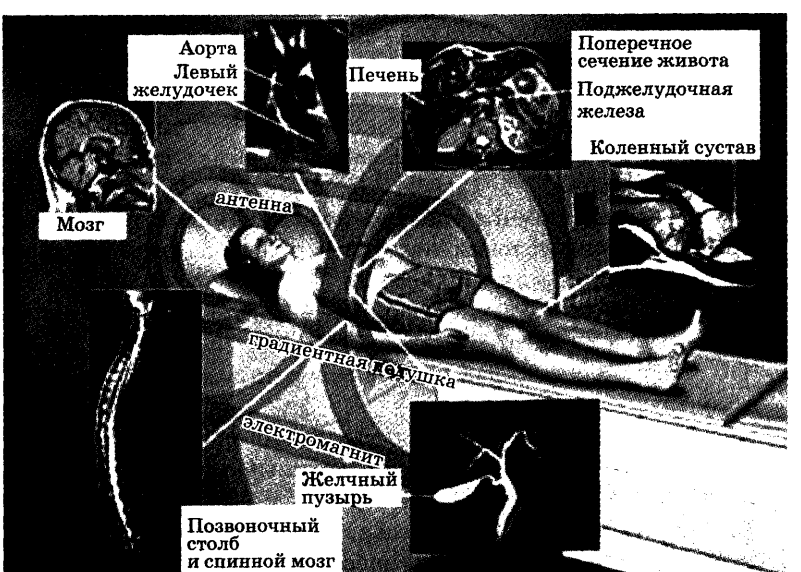


Рис. 26.10. МРТ-обследование

ких, органов брюшной полости. Перспективно также развитие МРТ с применением магнитно-контрастных веществ. Важным преимуществом метода магнитно-резонансной томографии является возможность получения изображений органов в разных плоскостях без изменения положения пациента в пространстве.

В магнитно-резонансной томографии используются сильные магнитные поля (индукцией около 0,5–2 тесла) и низкоинтенсивное зондирующее электромагнитное излучение радиочастотного диапазона, которые не представляют опасности для организма человека. Поэтому МРТ практически безвредна, не имеет возрастных ограничений и не требует предварительной подготовки пациента. Вместе с тем специфика этого метода диагностики не позволяет проводить МРТ у пациентов с кардиостимулятором, а также при наличии в теле пациента ферромагнитных предметов.

Контрольные вопросы и задания

1. Каковы физические отличия свободных радикалов от стабильных соединений? Укажите основные виды свободных радикалов в живых организмах.
2. Что такое магнитный момент электрона? Чем он определяется?
3. Как влияет магнитное поле на энергетический спектр парамагнитных молекул?
4. В чем состоит явление электронного парамагнитного резонанса? При каком условии наступает этот резонанс? Приведите схему установки для наблюдения ЭПР.
5. Что представляет собой спектр ЭПР? Какую информацию можно получить из него?
6. Что представляют собой спиновые метки и зонды? Для чего их используют?
7. Каковы магнитные свойства нуклонов? Как образуется магнитный момент ядра атома? Какие атомные ядра называют парамагнитными? Приведите примеры.
8. Как влияет магнитное поле на энергию атомного ядра? Приведите формулу для величины энергетического расщепления.
9. В чем состоит явление ядерного магнитного резонанса? При каком условии он наступает?

10. Что представляет собой ЯМР-спектр? Какую информацию он дает?
11. В чем состоит и чем вызван химический сдвиг в спектрах ЯМР?
12. В чем отличия и особенности протонного и фосфорного ЯМР? Какие еще ядра используют в ЯМР-исследованиях биообъектов?
13. Как используется явление ЯМР в медицине? На каких атомных ядрах? Почему?
14. Каковы принципы магнитно-резонансной томографии? Каковы ее возможности? Какие физические факторы действуют на пациента при МРТ-обследовании?

Глава 27 Рентгеновское излучение

Использование ионизирующих излучений в медицине началось с открытия в 1895 г. В. Рентгеном лучей, названных *X*-лучами, а позже рентгеновскими лучами. Выяснение природы и свойств данных лучей наряду с открытием естественной радиоактивности послужило стимулом для установления закономерностей радиоактивных превращений химических элементов, для развития рентгенологии, радиобиологии, радиационной медицины и других современных методов рентгеновских и радионуклидных исследований. Разработка практических методик воздействия ионизирующих излучений на биологические объекты, методов лечения заболеваний, методов радиационного контроля окружающей среды и ряда других подобных проблем невозможна без глубокого понимания чисто физических закономерностей, лежащих в основе взаимодействия ионизирующих излучений с веществом.

Рентгеновское излучение представляет собой электромагнитные волны с длиной волны от 80 до 10^{-5} нм. Границы этого диапазона несколько условны, так как длинноволновое рентгеновское излучение перекрывается коротковолновым ультрафиолетовым, а коротковолновое — длинноволновым гамма-излучением. Рентгеновские лучи коротких длин волн называют *жесткими*, а длинноволновые — *мягкими*.

27.1. Тормозное рентгеновское излучение

Согласно законам электродинамики, при ускорении или торможении любого движущегося электрического заряда возникает электромагнитное излучение. Причем частота этого излучения тем больше, чем больше модуль ускорения заряда. Если заряженная частица, обладающая большой кинетической энергией, испытывает резкое торможение, то возникает электромагнитное излучение рентгеновского диапазона. В медицинской практике источником рентгеновского излу-

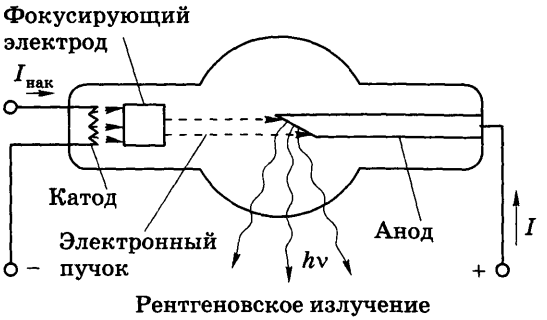


Рис. 27.1. Схема рентгеновской трубки

чения является ***рентгеновская трубка*** (рис. 27.1), которая представляет собой вакуумированный стеклянный баллон с двумя основными электродами: катодом (−) и анодом (+). Спираль катода разогревается электрическим током $I_{нак}$ (по цепи накала) и, благодаря термоэлектронной эмиссии, является источником свободных электронов. Между катодом и анодом (его называют также антикатодом) создается высокое постоянное напряжение U, которое сообщает электронам большую кинетическую энергию и заставляет электроны двигаться на анод. При попадании на анод электроны испытывают резкое торможение, их кинетическая энергия частично переходит в энергию рентгеновского излучения, а оставшаяся часть — в тепло, нагревающее анод.

Опишем процесс возникновения тормозного рентгеновского излучения количественно. Электрон, имеющий заряд e, под действием разности потенциалов U между катодом и анодом движется ускоренно и приобретает кинетическую энергию, равную eU. Эта кинетическая энергия при ударе об анод переходит частично в энергию кванта излучения $h\nu$ (h — постоянная Планка, ν — частота) и частично в тепло Q:

$$eU = h\nu + Q. \qquad (27.1)$$

Для одного электрона невозможно предсказать, какая часть его энергии перейдет в излучение, а какая в тепло — это процесс случайный. Однако поток электронов, падающих на анод, порождает рентгеновские кванты разных энергий,

вследствие чего спектр тормозного рентгеновского излучения оказывается сплошным. В этом спектре существует коротковолновая граница, определяющая минимальную длину волны λ_{min} (наибольшую возможную частоту ν_{max}) рентгеновского излучения при заданном напряжении U на трубке. Наибольшая частота ν_{max} в спектре тормозного рентгеновского излучения наблюдается в случае, когда вся кинетическая энергия электронов переходит в излучение. Тогда в формуле (27.1) $Q = 0$ и для коротковолновой границы спектра получаем соотношение

$$h\nu_{max} = \frac{hc}{\lambda_{min}} = eU \Rightarrow \lambda_{min} = \frac{hc}{eU}. \qquad (27.2)$$

В формулу (27.2), кроме напряжения U, входят постоянные величины: h — постоянная Планка, c — скорость света в вакууме, e — заряд электрона. Вычислив значение постоянной hc/e, формулу (27.2) можно представить в более удобном для практического использования виде:

$$\lambda_{min} = 1{,}23/U, \qquad (27.3)$$

дающем значения λ_{min} в нанометрах (нм), если напряжение U выражено в киловольтах.

Спектры тормозного рентгеновского излучения представлены на рис. 27.2 и 27.3. По оси ординат отложена спектральная плотность потока Φ_λ рентгеновского излучения, т.е.

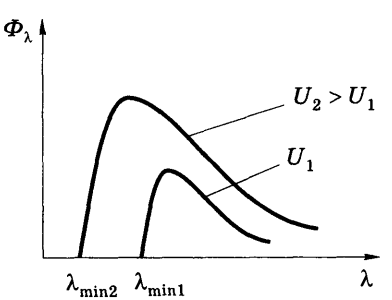


Рис. 27.2. Влияние напряжения в R-трубке на спектр тормозного излучения

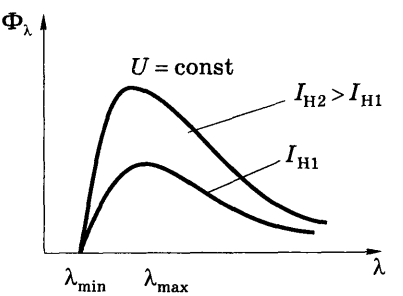


Рис. 27.3. Влияние тока в R-трубке на спектр тормозного излучения

отношение потока энергии $d\Phi(\lambda)$, излучаемого рентгеновской трубкой в узком интервале длин волн от λ до $\lambda + d\lambda$, к ширине этого интервала $d\lambda$.

Отметим, что максимум тормозного излучения приходится на длину волны, примерно в 1,5 раза большую коротковолновой границы:

$$\lambda_{max} \approx 1{,}5\lambda_{min}.$$

Полный поток Φ (Вт) тормозного рентгеновского излучения равен интегралу от Φ_λ по всем излучаемым длинам волн:

$$\Phi = \int_{\lambda_{min}}^{\infty} \Phi_\lambda d\lambda \qquad (27.4)$$

и численно равен площади, заключенной между кривой Φ_λ и осью абсцисс на графиках рис. 27.2 и 27.3.

Согласно формулам (27.2) и (27.3), с повышением напряжения между анодом и катодом рентгеновской трубки ($U_2 > U_1$) в спектре тормозного рентгеновского излучения уменьшается λ_{min} и излучение становится более жестким. Поэтому регулировка жесткости излучения в рентгеновских аппаратах осуществляется изменением напряжения между катодом и анодом рентгеновской трубки.

Полный поток Φ (Вт) тормозного рентгеновского излучения зависит от силы тока I и напряжения U в рентгеновской трубке и рассчитывается по формуле

$$\Phi = kIU^2Z, \qquad (27.5)$$

где $k = 10^{-9}$ ($В^{-1}$) — коэффициент пропорциональности; Z — номер атома вещества анода в периодической системе элементов.

Регулировать поток рентгеновского излучения можно как путем изменения напряжения U, так и путем изменения тока I в рентгеновской трубке [см. формулу (27.5)]. При изменении напряжения поток излучения будет изменяться быстрее, пропорционально U^2, но одновременно будет изменяться и жесткость (спектральный состав) излучения (см. рис. 27.2), что не всегда желательно.

Если же увеличивать только силу тока I в трубке, то поток излучения растет медленнее, но спектр и жесткость излучения при этом не изменяются (рис. 27.3).

Изменить силу тока I в рентгеновской трубке при фиксированном напряжении U можно путем изменения тока $I_{нак}$ накала катода. С его увеличением температура катода возрастает, увеличивается число электронов, покидающих катод за счет термоэлектронной эмиссии, и соответственно увеличивается сила тока в трубке. Таким образом, изменение тока накала катода позволяет регулировать поток рентгеновского излучения без изменения его спектрального состава (жесткости).

27.2. Характеристическое рентгеновское излучение

При высоких напряжениях на рентгеновской трубке на фоне сплошного спектра тормозного рентгеновского излучения проявляются резкие линии. Такое излучение называется ***характеристическим*** и не может быть объяснено тормозными эффектами. Оно объясняется следующим образом.

При больших энергиях электроны, бомбардирующие анод, преодолевают кулоновские силы отталкивания от наружных электронных оболочек атомов вещества анода и взаимодействуют с электронами, заполняющими внутренние оболочки. Происходит выбивание электрона, находящегося на внутренней оболочке атома. Этот эффект называется ***внутренней ионизацией*** или ***эффектом Оже*** и приводит к образованию свободного места на одной из внутренних электронных оболочек атома. При появлении такой вакансии, например в ближайшей к ядру K-оболочке, туда начнут переходить электроны с более удаленных L-, M-, N-оболочек и др. При этом будут испускаться кванты излучения с энергией $h\nu_{12} = E_2 - E_1$, равной разности энергий атома на исходной (L-, M-, N-оболочках и др.) и конечной K-оболочке. Энергия этих квантов очень велика и соответствует рентгеновскому диапазону. Спектр возникающего таким образом характеристичес-

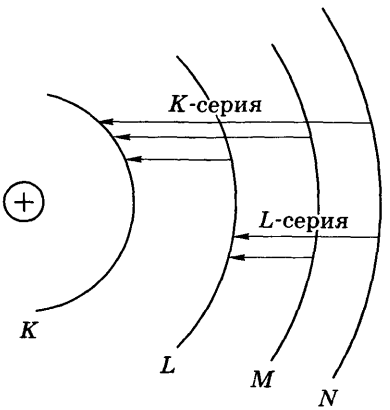


Рис. 27.4. Схема электронных переходов, образующих K- и L-серии в характеристическом излучении

кого рентгеновского излучения состоит из групп отдельных узких линий, соответствующих переходам электронов между внутренними оболочками бомбардируемых атомов. Напомним, что спектры излучения оптического диапазона обусловлены переходами валентных электронов из возбужденных верхних состояний в их основное состояние.

В линейчатом спектре характеристического излучения выделяют группы спектральных линий, которые называют K-, L-, M-сериями. Они образуются при переходах электронов с со всех более удаленных от ядра оболочек соответственно на K-, L-, M-оболочку и т.д. (рис. 27.4).

Частоты ν линий характеристического излучения зависят от порядкового номера элемента Z по ***закону Мозли***:

$$\sqrt{\nu} = A(Z - B), \qquad (27.6)$$

где A и B постоянные, зависящие от вида серии.

В отличие от оптических спектров, которые зависят от характера химических связей между атомами вещества, *характеристические рентгеновские спектры* атомов не зависят от химического окружения и для данного вида атомов они всегда одинаковы. Эта их особенность лежит в основе рентге-

новского спектрального анализа, который позволяет по частотам характеристического излучения вещества определить его элементный состав на основании соотношения (27.6).

27.3. Первичные эффекты взаимодействия рентгеновского излучения с веществом

Рассмотрим взаимодействие квантов рентгеновского излучения с атомами и молекулами вещества, в котором они распространяются. Очевидно, что результат этого взаимодействия зависит в первую очередь от энергии кванта излучения и здесь принято выделять следующие случаи.

1. *Энергия кванта меньше энергии ионизации атома* ($h\nu < A_и$). Тогда на атомах вещества происходит рассеяние квантов без изменения их частоты (рис. 27.5, *а*). Такое взаимодействие называется ***когерентным рассеянием***. Оно характерно для длинноволнового (мягкого) рентгеновского излучения. Когерентное рассеяние приводит к изменению структуры падающего пучка рентгеновских лучей: если падающий на вещество пучок был параллельным, то появляются кванты, которые распространяются в других направлениях.

2. *Энергия кванта несколько превышает энергию ионизации* ($h\nu \geq A_и$). В этом случае происходит ***фотоэффект*** (рис. 27.5, *б*) — рентгеновской квант поглощается веществом, а его энергия расходуется на выбивание электрона из атома (ионизацию) и сообщение ему кинетической энергии $E_к$ в соответствии с уравнением Эйнштейна для фотоэффекта

$$h\nu = A_и + E_к. \qquad (27.7)$$

3. *Энергия кванта значительно превышает энергию ионизации атомов вещества* ($h\nu >> A_и$). В этом случае лишь часть энергии $h\nu$ падающего кванта идет на ионизацию атома и на сообщение выбитому электрону кинетической энергии $E_к$, а основная часть идет на образование фотона меньшей энергии $h\nu'$ и другого направления распространения:

$$h\nu = A_и + h\nu' + E_к. \qquad (27.8)$$

Это явление называется ***некогерентным рассеянием***, поскольку частота рассеянного рентгеновского излучения уменьшается, или ***эффектом Комптона*** (рис. 27.5, *в*). Вторичные рентгеновские кванты, обладая энергией $h\nu' > A_и$, могут еще вызывать ионизацию вещества, в котором они распространяются. Электроны отдачи, получившие большую кинетическую энергию, могут также ионизировать соседние атомы вещества путем соударения.

Отмеченные первичные акты взаимодействия рентгеновского излучения и вещества могут порождать ряд вторичных процессов. В частности, возбуждение атомов и молекул вызывает ***рентгенолюминесценцию***, что используется для создания люминесцирующих экранов, позволяющих наблюдать рентгеновские изображения. Если происходит отрыв электронов с внутренних оболочек атомов, то возникает характеристическое рентгеновское излучение (рис. 27.5, *г*).

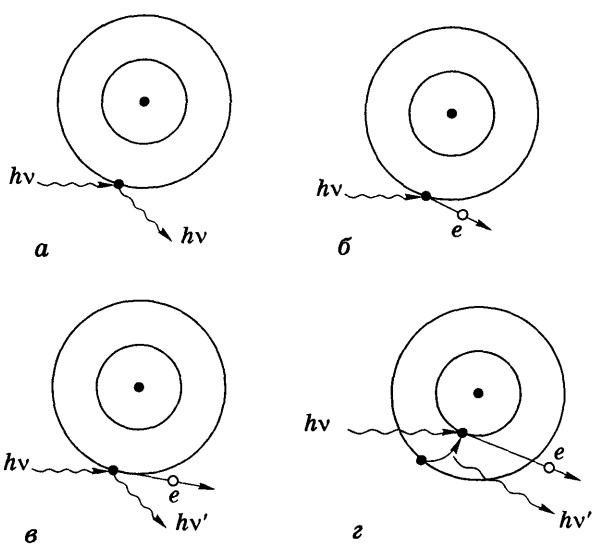


Рис. 27.5. Схемы первичных механизмов взаимодействия рентгеновского излучения с веществом:
а — когерентное рассеяние; *б* — фотоэффект; *в* — некогерентное рассеяние; *г* — образование характеристического рентгеновского излучения при некогерентном рассеянии

Таким образом, жесткие рентгеновские лучи за счет фотоэффекта, эффекта Комптона, а также вторичных процессов вызывают значительную ионизацию вещества, с которым они взаимодействуют. Подобная ионизация ведет к образованию свободных радикалов в биологических тканях, что весьма отрицательно влияет на кинетику биохимических процессов в этих тканях.

27.4. Ослабление рентгеновского излучения в веществе

Рассмотренные первичные эффекты взаимодействия рентгеновского излучения с веществом обусловливают уменьшение интенсивности данного излучения. Это происходит за счет двух процессов: *поглощения*, когда энергия рентгеновских квантов полностью или частично расходуется на структурные перестройки в веществе, и *рассеяния*, когда рентгеновские кванты изменяют свое первичное направление распространения (рис. 27.6). Так, если на вещество падает параллельный пучок рентгеновских лучей интенсивностью I_0, то при прохождении слоя толщиной x интенсивность излучения, распространяющегося в прежнем направлении, принимает значение I. Ослабление интенсивности излучения в данном случае происходит по тому же экспоненци-

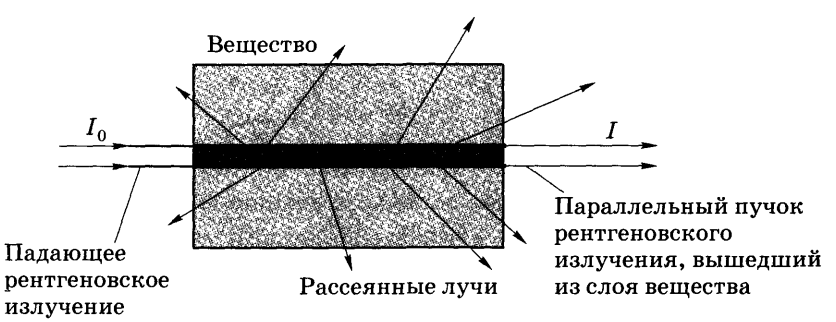


Рис. 27.6. Прохождение пучка рентгеновского излучения через вещество

альному закону, что и ослабление интенсивности световых лучей:

$$I = I_0 e^{-\mu x}, \qquad (27.9)$$

где μ — линейный показатель ослабления, характеризующий убыль интенсивности рентгеновских лучей на единице толщины слоя за счет поглощения ($\mu_п$) и рассеяния ($\mu_р$):

$$\mu = \mu_п + \mu_р.$$

Линейный показатель ослабления прямо пропорционален плотности вещества ($\mu \sim \rho$), тогда как массовый показатель ослабления (μ_m), представляющий собой отношение линейного показателя к плотности ρ вещества ($\mu_m = \mu/\rho$), уже не зависит от плотности и определяется только порядковым номером Z атомов этого вещества и длиной волны λ рентгеновского излучения:

$$\mu_m = k\lambda^3 Z^3, \qquad (27.10)$$

где k — коэффициент пропорциональности.

При ослаблении потока тормозного рентгеновского излучения веществом следует учитывать непрерывность спектра этого излучения и тот факт, что длинноволновое излучение поглощается значительно сильнее, чем коротковолновое [см. формулу (27.10)].

На практике для оценки проникающей способности рентгеновского излучения используется понятие *«слой половинного ослабления $d_{1/2}$»* — это толщина слоя поглотителя, которая ослабляет интенсивность падающего излучения в два раза (рис. 27.7). Связь между слоем половинного ослабления и показателем ослабления μ можно установить из формулы (27.9):

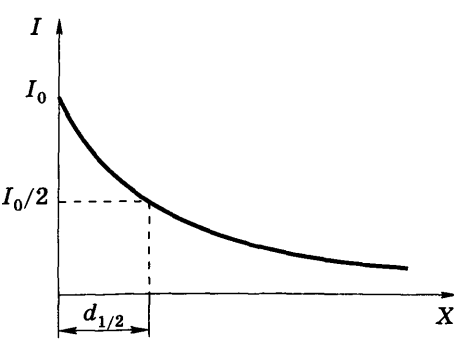


Рис. 27.7. Ослабление интенсивности рентгеновского излучения при прохождении через вещество

при $x = d_{1/2}$ интенсивность прошедшего излучения $I = I_0/2$:

$$\frac{I_0}{2} = I_0 e^{-\mu d_{1/2}} \Rightarrow e^{-\mu d_{1/2}} = \frac{1}{2} \Rightarrow e^{\mu d_{1/2}} = 2 \Rightarrow$$

$$\Rightarrow \mu d_{1/2} = \ln 2 = 0{,}69. \qquad (27.11)$$

Таким образом, $d_{1/2} = \dfrac{\ln 2}{\mu} = \dfrac{0{,}69}{\mu}$.

Например, для рентгеновского тормозного излучения, возникающего при напряжении на рентгеновской трубке 60 кВ, слой половинного ослабления составляет 10 мм воды или 1 мм алюминия.

Если рентгеновское излучение широкого спектрального состава (например, тормозное) проходит через слой вещества, то его спектральный состав изменяется — излучение становится более жестким, так как мягкое излучение поглощается сильнее, а более короткие рентгеновские лучи поглощаются меньше и лучше проникают через слой вещества. Поэтому второй слой половинного ослабления окажется толще первого. Например, при указанных условиях для воды он составит 15,3 мм, а третий слой — уже 20 мм, при этом спектр прошедшего излучения сужается и в нем остается лишь коротковолновая часть спектра. Этот эффект используется для создания пучков моноэнергетических рентгеновских лучей в процессе фильтрации рентгеновского излучения. Так, при напряжении на трубке 80 кВ и фильтре в виде пластинки алюминия толщиной 20 мм излучение становится почти моноэнергетическим. Слой половинного ослабления при этом составляет 7 мм алюминия.

27.5. Физические принципы рентгенодиагностики

Массовый показатель ослабления рентгеновских лучей веществом сильно зависит от состава вещества и длины волны, что прямо следует из выражения (27.10). При просвечивании тела человека рентгеновским излучением на люминесцирующем экране или фотопленке, расположенным за ним, наблюдается теневое изображение внутренних органов. Для

рентгенодиагностики используется излучение, для которого первичные процессы взаимодействия с веществом обусловливаются в основном фотоэффектом.

Разные ткани и органы человека по-разному поглощают рентгеновские лучи. Так, изображение костей очень отчетливо проявляется на фоне изображения мягких тканей. Значительно сложнее различаются изменения в структуре мягкой ткани, так как показатели ослабления рентгеновского излучения разных ее участков отличаются незначительно, поэтому очаги патологии выявляются трудно. Эту проблему решают путем введения в исследуемые органы *рентгеноконтрастных веществ*, содержащих элементы с большим Z, которые сильно поглощают рентгеновское излучение и тем самым хорошо контрастируют исследуемый орган или ткань при рентгеновском обследовании.

Так, при осмотре желудка, кишечника и т.п. применяют сульфит бария ($Z_{Ba} = 56$) в виде кашеобразной массы, заполняющей эти полые органы. При исследовании сосудистого русла часто используют нетоксичные для организма соединения иода ($Z_I = 53$).

Рентгенодиагностика создает значительную лучевую нагрузку на организм, особенно при ***рентгеноскопии***, когда изображение достаточно долго рассматривают на люминесцирующем экране. При ***рентгенографии*** время экспозиции для получении рентгеновского снимка составляет доли секунды, поэтому лучевая нагрузка в сотни раз меньше.

Существенное снижение дозы облучения при рентгенодиагностике и расширение диагностических возможностей достигается за счет применения усилителей рентгеновского изображения, телевизионных систем, чувствительных датчиков и цифровых методов обработки изображений.

Важной задачей рентгенодиагностики является получение изображений внутренних органов человека в различных сечениях. Эта задача решена с созданием ***рентгеновской компьютерной томографии*** (КТ).

От традиционных теневых методов рентгенодиагностики она принципиально отличается тем, что вычисляет изображения органов в узких заданных поперечных сечениях тела

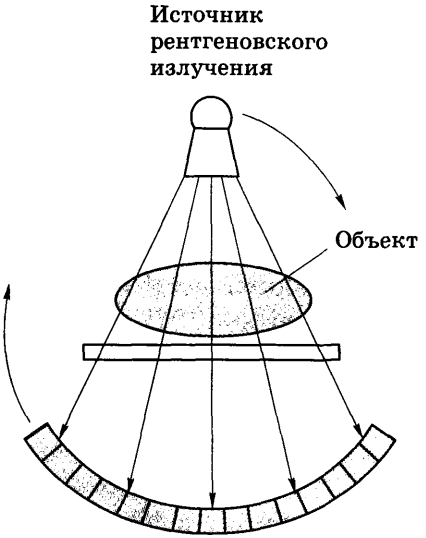


Рис. 27.8. Линейка детекторов, преобразующих рентгеновское излучение в электрические сигналы

человека путем компьютерной обработки электрических сигналов, поступающих от детекторов рентгеновского излучения. Один из многообразных приемов просвечивания и детектирования излучения для такого исследования представлен на рис. 27.8, а сам рентгеновский компьютерный томограф изображен на рис. 27.9.

Расходящийся веерный пучок рентгеновских лучей, проходя через объект, попадает на линейку детекторов, число которых составляет от 300 до 700 и более (чем их больше, тем выше разрешающая способность метода). Система излучатель — детекторы вращается вокруг объекта, создавая рентгеновские «электронные снимки» объекта под разными углами, через 1–2 градуса поворота. При исследовании система совершает обычно полный оборот и производит за это время 300–400 включений.

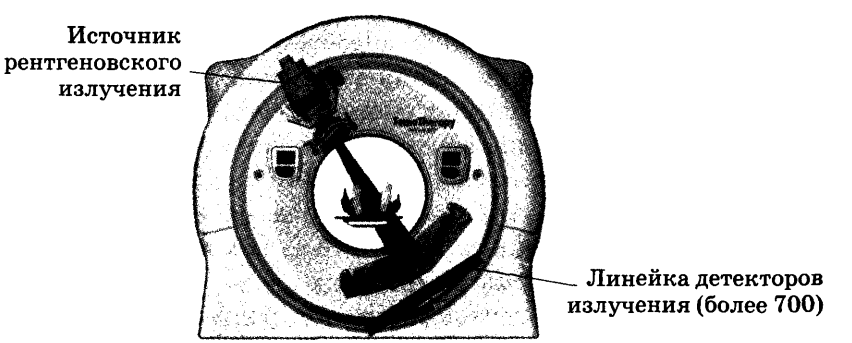


Рис. 27.9. Схема рентгеновского компьютерного томографа

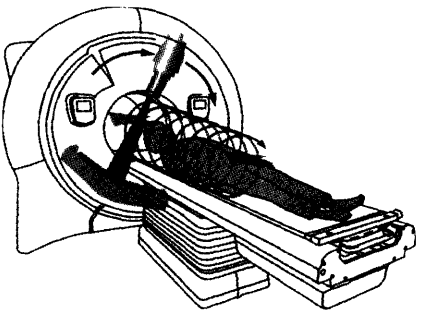


Рис. 27.10. Схема «спирального» КТ-сканирования

Зарегистрированные линейкой детекторов многочисленные рентгеновские проекции органа в заданной плоскости поступают в компьютер, который обрабатывает их по специальной программе и вычисляет рентгеновскую плотность в каждой точке исследованного сечения, после чего выводит получившееся изображение в серой шкале на экран монитора. Это изображение при необходимости можно контрастировать, увеличивать и сохранять в памяти компьютера либо на внешнем цифровом носителе.

При обследовании пациента обычно подробно изучают определенную область его тела, поэтому получают КТ-изображения нескольких параллельных сечений этой области. Для ускорения процедуры получения последовательности таких изображений применяют «спиральное» сканирование: стол с пациентом равномерно движется при одновременном вращении системы излучатель — датчик (рис. 27.10).

Рентгеновская компьютерная томография позволяет получать послойное изображение органов и различать структуры, имеющие отличия лишь на 0,1 % по показателю поглощения и не менее 2 мм по размерам деталей. Время исследования для получения одного сечения может достигать нескольких секунд.

Контрольные вопросы и задания

1. Каков механизм возникновения тормозного рентгеновского излучения? Почему спектр тормозного излучения сплошной? Как определить его коротковолновую границу?
2. Как осуществляется регулировка интенсивности и жесткости рентгеновского излучения в рентгеновских аппаратах? Как и от чего зависит поток тормозного рентгеновского излучения?

3. Сравните спектры теплового излучения и тормозного рентгеновского излучения. В чем их сходства и различия?
4. Как возникает характеристическое рентгеновское излучение? Каковы его особенности? Сформулируйте закон Мозли.
5. В чем различия механизмов образования оптических спектров атомов и спектров характеристического рентгеновского излучения?
6. Опишите основные механизмы взаимодействия рентгеновского излучения с веществом. Почему жесткое рентгеновское излучение, которое в меньшей степени поглощается веществом, более вредно по биологическому действию, чем мягкое?
7. Как изменяется поток рентгеновского излучения после прохождения через слой вещества? Что такое линейный и массовый показатели ослабления, слой половинного ослабления? Как они связаны между собой?
8. На чем основано использование рентгеновского излучения в диагностике? Опишите методы контрастирования кровеносного русла, ЖКТ и других мягких тканей.
9. Сравните физические принципы ультразвуковой и рентгеновской диагностики, укажите их достоинства и недостатки.
10. Каковы принципы и возможности рентгеновской компьютерной томографии?
11. Каковы способы снижения дозы облучения пациента при рентгенодиагностическом обследовании?

Глава 28 | Радиоактивность

Ядра многих атомов (радионуклиды) способны к самопроизвольным превращениям, в результате которых возникают ядра других атомов и ионизирующие излучения в виде квантов электромагнитной энергии и элементарных частиц. Эти излучения и частицы, взаимодействуя с биологическими тканями, оказывают на них сильное воздействие, вызывая структурные и функциональные изменения. Рассмотрим законы радиоактивности и их медико-биологическое применение.

28.1. Основные характеристики ядер

Ядра атомов состоят из элементарных частиц, называемых ***нуклонами***. Последние, в свою очередь, подразделяются на *протоны* и *нейтроны*. Протон имеет массу, близкую к одной атомной единице массы (1 а.е.м.), и заряд $e = +1{,}6 \cdot 10^{-19}$ Кл, равный по величине и противоположный по знаку заряду электрона. Нейтрон не имеет заряда, его масса примерно совпадает с массой протона (1,001 а.е.м.).

Электрический заряд q ***ядра элемента*** определяется числом протонов, входящих в его состав, и в единицах элементарного заряда равен порядковому номеру Z элемента в таблице Менделеева: $q = Ze$.

Массовое число ядра A всегда целое и определяется количеством нуклонов в ядре: $A = Z + N$, где Z — число протонов; N — число нейтронов. Символическая запись ядра элемента: ${}^{A}_{Z}X$. Нижний индекс Z у элемента означает порядковый номер, верхний A — массовое число.

Радиус ядра R зависит от массового числа и может быть вычислен (в метрах) по приближенной формуле

$$R = 1{,}5 \cdot 10^{-15} \sqrt[3]{A} \, . \qquad (28.1)$$

Нуклоны в ядре связаны особыми ядерными силами, превышающими силы электростатического отталкивания между протонами. Они обеспечивают ядру достаточную устойчивость и обладают следующими специфическими *свойствами*:

• короткодействие — действуют только на расстояниях порядка размеров самого ядра ($< 10^{-14}$ м);

• сильнодействие — сила взаимодействия нуклонов в ядре на несколько порядков выше, чем силы любых известных в природе взаимодействий;

• зарядовая независимость — силы ядерного взаимодействия не зависят от электрического заряда нуклона, поэтому силы взаимодействия двух протонов такие же, как и двух нейтронов или нейтрона и протона;

• насыщение связей — каждый нуклон взаимодействует только с ограниченным числом окружающих его нуклонов (обычно не более 6).

Устойчивость атомных ядер зависит от общего числа нуклонов в ядре, а также от соотношения числа нейтронов и протонов N/Z. Наиболее устойчивы ядра с относительно небольшим числом нуклонов и значением $N/Z = 1$. При увеличении общего числа нуклонов в ядре и превышении числа нейтронов над числом протонов ($N/Z > 1{,}6$) устойчивость ядра ослабляется и увеличивается вероятность его самопроизвольного распада.

Действие ядерных сил обусловливает наличие определенной энергии связи $\varepsilon_{св}$ нуклонов в ядре. Энергия отдельно взятых нуклонов $\sum m_н c^2$ всегда больше, чем энергия ядра $M_я c^2$, которое они образуют. Разность этих энергий превращается в ***энергию связи нуклонов***:

$$\varepsilon_{св} = \sum m_н c^2 - M_я c^2 = (\sum m_н - M_я)c^2, \qquad (28.2)$$

где $M_я$ — масса ядра; $m_н$ — средняя масса нуклона; c — скорость света в вакууме.

Разность масс ядра и составляющих его нуклонов $\Delta M = (\sum m_н - M_я)$ называют ***дефектом массы***, эта величина и определяет энергию связи ядра.

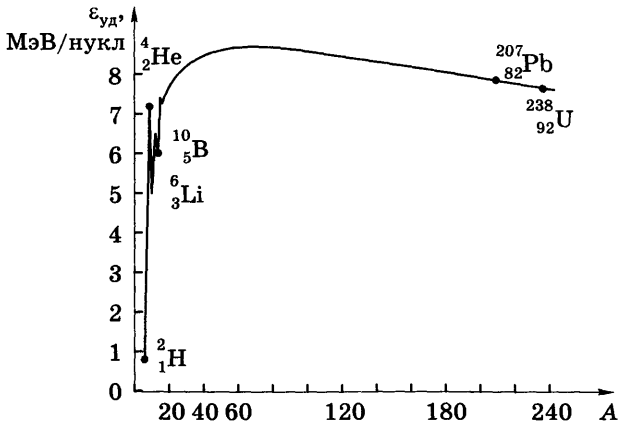


Рис. 28.1. Зависимость удельной энергии связи нуклонов в ядре от массового числа

Удельная энергия связи $\varepsilon_{уд}$ представляет собой среднее значение энергии связи, приходящееся на один нуклон: $\varepsilon_{уд} = \varepsilon_{св}/A$ (рис. 28.1). Ее максимальное значение (8,7 МэВ/нуклон) характерно для ядер с числом нуклонов $A = 50–60$. С изменением числа A у более легких элементов она уменьшается резко, а у более тяжелых — плавно. Отсюда возникают две возможности высвобождения и использования ядерной энергии.

1. *Ядерная реакция деления*, при которой ядро тяжелого элемента (например, $^{235}_{92}$U) делится с образованием более устойчивых ядер и превращением части энергии связи нуклонов в кинетическую энергию продуктов деления, которая затем преобразуется в теплоту в ядерных энергетичеких установках. На реакции деления основана работа всех современных атомных электростанций (АЭС) и других промышленных ядерных энергоустановок.

2. *Реакция синтеза* ядер легких элементов, сопровождающаяся образованием более устойчивого ядра большей массы и выделением значительного количества энергии. Реакция синтеза может идти только при высоких температурах (несколько миллионов градусов), необходимых для преодоления кулоновского отталкивания ядер, и реализуется в термоядерном взрыве и в термоядерных исследовательских реакторах.

28.2. Виды радиоактивного распада

Радиоактивностью называют самопроизвольный распад неустойчивых ядер с образованием других ядер и испусканием элементарных частиц. *Естественная* радиоактивность характерна для ядер, существующих в природных условиях, *искусственная* — для ядер, искусственно полученных в результате ядерных реакций. Вид распада определяется строением ядра, соотношением и общим количеством протонов и нейтронов, характером ядерных сил.

Исходное ядро принято называть ***материнским***, а ядро, образовавшееся в результате его распада, — ***дочерним***.

При любых видах распада, естественно, выполняются законы сохранения электрического заряда, массового числа, полной энергии и импульса.

Если образовавшееся в результате распада дочернее ядро оказывается не в основном, а в возбужденном энергетическом состоянии, то через доли секунды оно переходит в основное состояние, отдавая избыток энергии в виде γ-излучения. При этом энергия γ-кванта, испускаемого данным дочерним ядром, всегда одинакова, что позволяет уверенно определять вид образовавшихся ядер путем измерения энергии γ-квантов, сопровождающих радиоактивный распад.

Альфа-распад (α-распад) сопровождается испусканием α-частицы (ядра атома гелия ${}^{4}_{2}\mathrm{He}$). При этом ядро элемента X, имеющего порядковый номер Z (число протонов) и массовое число A, превращается в ядро нового элемента Y по схеме

$$ {}^{A}_{Z}X \rightarrow {}^{A-4}_{Z-2}Y + {}^{4}_{2}\alpha , \qquad (28.3) $$

вытекающей из законов сохранения электрического заряда и массового числа. Как и любой вид распада, α-распад может сопровождаться испусканием γ-квантов. Такие процессы, например, имеют место при α-распаде плутония:

$$ {}^{239}_{94}\mathrm{Pu} \rightarrow {}^{235}_{92}\mathrm{U} + {}^{4}_{2}\alpha + \gamma . \qquad (28.4) $$

Кинетическая энергия α-частицы, испускаемой при α-распаде ядра, имеет вполне определенное значение, т.е. α-излучение имеет линейчатый энергетический спектр.

Бета-распад (β-распад) подразделяется на три вида:

1. *Электронный,* или β^-*-распад.* В этом случае ядро радионуклида испускает электрон, образуя ядро нового элемента с порядковым номером на единицу большим исходного:

$$ {}^{A}_{Z}X \rightarrow {}^{A}_{Z+1}Y + {}^{0}_{-1}\beta + {}^{0}_{0}\tilde{\nu}\,. \qquad (28.5) $$

Детальное изучение этого распада показало, что вместе с электроном из ядра вылетает еще одна электрически нейтральная частица ничтожно малой массы, которую назвали *антинейтрино* ($ {}^{0}_{0}\tilde{\nu} $). Может возникнуть вопрос: откуда при таком распаде появляется электрон, если в ядре его нет? Противоречия здесь нет, поскольку электрон образуется при внутриядерном превращении нейтрона в протон:

$$ {}^{1}_{0}n \rightarrow {}^{1}_{1}p + {}^{0}_{-1}\beta + {}^{0}_{0}\tilde{\nu}\,. \qquad (28.5а) $$

Энергия, выделяющаяся при электронном распаде, распределяется в основном между электроном и антинейтрино, при этом случайным образом, поэтому кинетическая энергия β-частиц, испускаемых при распаде, принимает все возможные значения — от 0 до E_{max}, в связи с этим кинетическая энергия β-частиц имеет сплошной спектр. На рис. 28.2 пока-

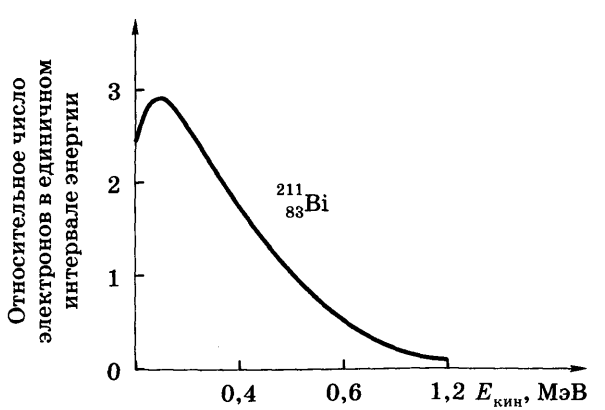


Рис. 28.2. Энергетический спектр β-частиц при распаде $ {}^{211}_{83}\mathrm{Bi} $

зано распределение электронов по энергиям (энергетический спектр β-частиц) при распаде висмута $^{211}_{83}\mathrm{Bi}$.

Для каждого конкретного радионуклида существует свое характерное значение максимальной энергии β-частицы E_{max}.

2. *Позитронный*, или β^+-*распад*, сопровождается испусканием позитрона (элементарной частицы, равной по массе и модулю заряду электрона, но с положительным электрическим зарядом). При этом образуется ядро нового элемента, порядковый номер которого на единицу меньше исходного, и нейтрино:

$$^{A}_{Z}X \rightarrow {}^{A}_{Z-1}Y + {}^{0}_{+1}\beta + {}^{0}_{0}\nu . \qquad (28.6)$$

Энергетический спектр β^+-частиц сплошной. Позитрон образуется в результате внутриядерного превращения протона в нейтрон:

$$^{1}_{+1}p \rightarrow {}^{1}_{0}n + {}^{0}_{+1}\beta + {}^{0}_{0}\nu . \qquad (28.6а)$$

Примером такого распада может служить превращение фосфора в кремний

$$^{30}_{15}\mathrm{P} \rightarrow {}^{30}_{14}\mathrm{Si} + {}^{0}_{+1}\beta + {}^{0}_{0}\nu .$$

Нейтрино и антинейтрино весьма сложно обнаружить экспериментально, поскольку из-за малой массы и отсутствия заряда они обладают колоссальной проникающей способностью, тем не менее их существование подтверждено опытным путем. Различие между нейтрино и антинейтрино состоит в том, что у первой частицы спин направлен вдоль движения, а у второй — против него.

3. *Электронный захват* (*е-захват*). В этом случае электрон захватывается ядром с одной из внутренних оболочек своего атома, в результате чего протон ядра превращается в нейтрон с выделением нейтрино:

$$^{1}_{1}p + {}^{0}_{-1}\beta \rightarrow {}^{1}_{0}n + {}^{0}_{0}\nu . \qquad (28.7)$$

Примером *е*-захвата может служить превращение бериллия в литий:

$$^{7}_{4}\mathrm{Be} + {}^{0}_{-1}\beta \rightarrow {}^{7}_{3}\mathrm{Li} + {}^{0}_{0}\nu .$$

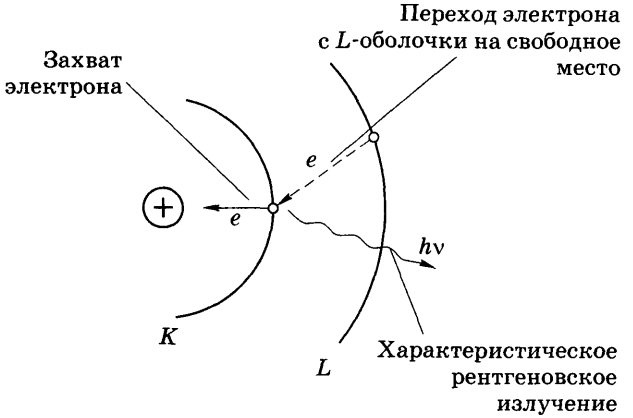


Рис. 28.3. Электронный захват (e-захвата)

Поскольку при данном виде радиоактивного распада освобождается место для электрона на внутренней оболочке атома, то электроны из более удаленных оболочек переходят на это место, поэтому e-захват всегда сопровождается характеристическим рентгеновским излучением. Это отличает e-захват от позитронного распада, поскольку в обоих случаях образуются одинаковые дочерние ядра из одинаковых исходных ядер. На рис. 28.3 представлен случай захвата ядром электрона с K-оболочки (K-захват) и возникновения характеристического излучения при переходе электрона с L- на K-оболочку.

Следует иметь в виду, что образовавшиеся при радиоактивных распадах дочерние ядра могут подвергаться дальнейшим радиоактивным превращениям до тех пор, пока не образуется стабильное ядро.

28.3. Простейшие ядерные реакции. Получение радионуклидов

Ядерной реакцией называют взаимодействие ядра атома с другими ядрами или элементарными частицами с образованием ядра другого элемента и других элементарных частиц.

При всех ядерных реакциях выполняются законы сохранения суммарного электрического заряда, массового числа, полной энергии и импульса.

Основные методы получения радионуклидов:

1. *Бомбардировка ядер атомов заряженными элементарными частицами.* Первая такая ядерная реакция была осуществлена Э. Резерфордом при бомбардировке ядер, азота альфа-частицами:

$$^{14}_{7}\mathrm{N} + {}^{4}_{2}\alpha \rightarrow {}^{17}_{8}\mathrm{O} + {}^{1}_{1}p\,.$$

При этом ядра азота превращались в ядра кислорода с выделением протона. Для осуществления ядерных реакций бомбардировкой положительно заряженными частицами (α-частицами, протонами, дейтронами) их необходимо ускорить до довольно высоких энергий, чтобы заряженная частица преодолела силы электростатического отталкивания от ядра и была «захвачена» короткодействующими ядерными силами. Образующиеся в результате этого новые ядра чаще всего неустойчивы и испытывают радиоактивный распад того или иного вида.

2. *Нейтронная активация.* Радионуклиды в относительно больших количествах могут быть получены в атомных реакторах путем облучения ядер стабильных элементов нейтронами (нейтронная активация). Поглощая нейтрон, стабильное ядро превращается в радионуклид того же элемента:

$$^{A}_{Z}X + {}^{1}_{0}n \rightarrow {}^{A+1}_{Z}X. \qquad (28.8)$$

Таким образом, можно получить широко применяющийся в медицине радиоактивный кобальт

$$^{59}_{27}\mathrm{Co} + {}^{1}_{0}n \rightarrow {}^{60}_{27}\mathrm{Co}\,,$$

который затем испытывает электронный распад

$$^{60}_{27}\mathrm{Co} \rightarrow {}^{60}_{28}\mathrm{Ni} + {}^{0}_{-1}\beta + \gamma + {}^{0}_{0}\tilde{\nu}\,.$$

Образующееся ядро атома никеля оказывается в возбужденном состоянии. Его переход в основное состояние сопро-

вождается испусканием γ-кванта, и это γ-излучение используется в лучевой терапии для воздействия на злокачественные новообразования.

3. *Выделение радионуклидов из отработанных стержней ядерного реактора.* Описанные ядерные реакции следует отличать от *цепных реакций* деления тяжелых ядер, которые используются для получения энергии в ядерных реакторах. В последнем случае после поглощения нейтрона ядром оно делится с испусканием нескольких нейтронов и выделением энергии. Примером может служить наработка $^{239}_{94}\mathrm{Pu}$ в урановом реакторе. В ядерный реактор большой мощности (типа РБМК Чернобыльской АЭС) загружается около 190 т топлива в виде природного урана, состоящего в основном из $^{238}_{92}\mathrm{U}$, слабо обогащенного ядрами $^{235}_{92}\mathrm{U}$. Превращение некоторой части урана в плутоний происходит в несколько этапов. При облучении $^{238}_{92}\mathrm{U}$ нейтронами образуется $^{239}_{92}\mathrm{U}$:

$$^{238}_{92}\mathrm{U} + {}^{1}_{0}n \rightarrow {}^{239}_{92}\mathrm{U} + \gamma,$$

который путем электронного распада превращается сначала в радионуклид нептуния

$$^{239}_{92}\mathrm{U} \rightarrow {}^{239}_{93}\mathrm{Np} + {}^{0}_{-1}\beta + {}^{0}_{0}\tilde{\nu}, \qquad (28.9)$$

а затем в радиоактивный плутоний

$$^{239}_{93}\mathrm{Np} \rightarrow {}^{239}_{94}\mathrm{Pu} + {}^{0}_{-1}\beta + {}^{0}_{0}\tilde{\nu}. \qquad (28.9а)$$

Образовавшийся $^{239}_{94}\mathrm{Pu}$ имеет большой период полураспада и при попадании в окружающую среду, например вследствие аварии реактора, вызывает ее длительное и крайне опасное радиационное загрязнение из-за своей α-радиоактивности.

Рассмотрим примеры β^--распада основных радиоактивных элементов, обусловивших значительное радиационное заражение среды после аварии на Чернобыльской АЭС.

В первое время после аварии основной вклад в суммарную радиоактивность давал радиоактивный иод, распадающийся с испусканием гамма-квантов с энергией 0,72 МэВ:

$$^{131}_{53}\mathrm{I} \rightarrow {}^{131}_{54}Xe + {}^{0}_{-1}\beta + \gamma + {}^{0}_{0}\tilde{\nu}. \qquad (28.10)$$

Максимальная энергия электронов, образующихся при этом распаде, равна 0,81 МэВ. Поскольку скорость распада иода велика (за 8 суток распадается примерно половина ядер), то он имел наибольшую начальную активность и легко обнаруживался по интенсивному γ-излучению и большому числу электронов, испускаемых за единицу времени. Через несколько месяцев после аварии практически весь иод распался.

Распад радиоактивного $^{137}_{55}Cs$ сопровождается испусканием γ-квантов с энергией 0,66 МэВ и максимальной энергией электронов 1,18 МэВ:

$$^{137}_{55}Cs \rightarrow {}^{137}_{56}Ba + {}^{0}_{-1}\beta + \gamma + {}^{0}_{0}\tilde{\nu}\,. \qquad (28.11)$$

Скорость распада у $^{137}_{55}Cs$ существенно ниже, чем у $^{131}_{53}I$, половина ядер цезия-137 распадается примерно за 30 лет.

Распад радиоактивного $^{90}_{38}Sr$ не сопровождается испусканием γ-квантов, поэтому этот изотоп гораздо сложнее обнаружить, чем радиоактивный иод или цезий:

$$^{90}_{38}Sr \rightarrow {}^{90}_{39}Y + {}^{0}_{-1}\beta + {}^{0}_{0}\tilde{\nu}\,. \qquad (28.12)$$

Образующееся в результате этого распада ядро иттрия также неустойчиво и после электронного распада превращается в стабильное ядро циркония:

$$^{90}_{39}Y \rightarrow {}^{90}_{40}Zr + {}^{0}_{-1}\beta + {}^{0}_{0}\tilde{\nu}\,. \qquad (28.12а)$$

Радиационное заражение среды радионуклидами, распадающимися с испусканием γ-квантов, обнаруживается наиболее просто, так как γ-излучение имеет высокую проникающую способность, распространяется в воздухе на сотни метров и легко улавливается приборами.

Радиационное заражение среды радионуклидами, не испускающими при распаде γ-кванты, в частности стронцием $^{90}_{38}Sr$, обнаружить существенно сложнее, поскольку регистрация β-излучения более сложная задача, чем обнаружение γ-излучения.

28.4. Основной закон радиоактивного распада. Активность радионуклидов

Предсказать время, в течение которого отдельно взятое ядро радиоактивного атома распадется, невозможно, можно лишь сделать заключение о вероятности данного события. Если же имеется большая совокупность ядер, то легко вывести статистический закон, выражающий зависимость числа нераспавшихся ядер от времени. Он устанавливается на основании следующих соображений.

Будем считать, что за малый промежуток времени dt распадется число ядер dN. Очевидно, это число пропорционально промежутку времени dt и общему числу нераспавшихся к данному моменту радиоактивных ядер N. Тогда дифференциальное уравнение, описывающее процесс распада, приобретает следующий вид:

$$dN = -\lambda N dt. \qquad (28.13)$$

Знак минус указывает на убыль числа нераспавшихся со временем ядер, коэффициент $\lambda(c^{-1})$ зависит от вида ядер и называется ***постоянной распада.***

Введем начальное условие для этого уравнения, обозначив число нераспавшихся ядер в исходный момент времени через N_0 ($N = N_0$ при $t = 0$). Решение дифференциального уравнения (28.13) с учетом этого начального условия имеет вид

$$N = N_0 e^{-\lambda t}. \qquad (28.14)$$

Следовательно, ***закон радиоактивного распада*** показывает, что число нераспавшихся ядер атомов убывает со временем экспоненциально. Скорость распада определяется постоянной радиоактивного распада λ, входящей в показатель экспоненты в формуле (28.14), и зависит только от вида ядер. На практике чаще используют ***период полураспада*** T — время, за которое распадается половина начального количества радиоактивных ядер. Период полураспада легко найти из графика зависимости $N(t)$ (рис. 28.4).

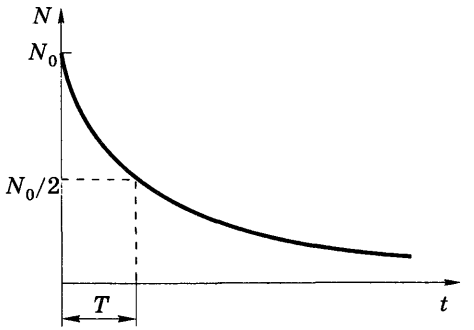


Рис. 28.4. Зависимость числа нераспавшихся ядер N от времени t

Для того чтобы найти аналитическую связь между постоянной распада λ и периодом полураспада T, подставим в уравнение (28.14) $t = T$. Тогда $N_0/2 = N_0 e^{-\lambda T}$, отсюда получаем $2 = e^{\lambda T}$, или $\lambda T = \ln 2$, и окончательно: $T = \dfrac{\ln 2}{\lambda} \approx \dfrac{0{,}69}{\lambda}$.

Отметим, что постоянная распада λ измеряется в единицах, обратных времени, а период полураспада определяют в секундах, часах, годах и т.д.

Кроме рассмотренных параметров λ и T, распад радионуклидов характеризуют также ***средним временем жизни*** τ — промежутком времени, за который число нераспавшихся ядер убывает в e раз. Среднее время жизни τ связано с λ или T соотношением

$$\tau = \frac{1}{\lambda} = \frac{T}{0{,}69} = 1{,}44T. \qquad (28.15)$$

28.4.1. Активность. Единицы измерения активности

Важной характеристикой, определяющей скорость распада радионуклидов, является их ***активность***, равная числу актов распада за единицу времени:

$$A = -\frac{dN_{\text{расп}}}{dt} = -\frac{dN}{dt}. \qquad (28.16)$$

Подставляя в формулу (28.16) выражения (28.13) и (28.14), получим

$$A = -\frac{dN}{dt} = \lambda N = \lambda(N_0 e^{-\lambda t}) = A_0 e^{-\lambda t}, \qquad (28.17)$$

где $A_0 = \lambda N_0$ — начальная активность препарата.

Таким образом,

$$A = \lambda N = 0{,}69\frac{N}{T}. \qquad (28.18)$$

Активность радионуклидов прямо пропорциональна числу ядер, еще нераспавшихся на данный момент. Видно также, что активность радиопрепарата тем выше, чем меньше период его полураспада.

Со временем активность убывает по тому же экспоненциальному закону, что и число нераспавшихся ядер:

$$A = A_0 e^{-\lambda t}. \qquad (28.19)$$

За единицу активности в международной системе единиц принимают беккерель (Бк) — активность, при которой за 1 с происходит 1 распад. Более крупная единица активности — 1 резерфорд (Рд): 1 Рд = 10^6 Бк. Часто используется внесистемная единица активности — кюри (Ки), соответствующая активности 1 г радия $^{226}_{88}\mathrm{Ra}$: 1 Ки = $3{,}7 \cdot 10^{10}$ Бк.

Активность типичных расфасовок радиоактивных веществ, применяемых для лабораторных исследований, составляет обычно микро- и милликюри.

Установим связь между массой радионуклидов m и их активностью A. Число нераспавшихся ядер N равно их массе m, деленной на массу одного атома m_A (массой электронов можно пренебречь):

$$N = \frac{m}{m_\mathrm{A}} = \frac{m}{M}N_\mathrm{A},$$

где M — молярная масса; N_A — число Авогадро.

Подставляя это выражение в формулу (28.18), найдем связь активности A с массой m, молярной массой M и периодом полураспада T радионуклида:

$$A = \frac{0{,}69N}{T} = \frac{0{,}69mN_A}{TM} = \frac{0{,}69 \cdot 6{,}02 \cdot 10^{23} \cdot m}{TM} =$$

$$= 4{,}17 \cdot 10^{23} \frac{m}{TM}. \qquad (28.20)$$

Подставив в данное выражение массу m радионуклида в граммах и его атомную массу M в г/моль, а значение периода полураспада T в секундах, получим активность в беккерелях. Наоборот, по известной активности радионуклида можем найти его массу:

$$m = 2{,}4 \cdot 10^{-24} ATM.$$

Например, определим активность 1 г $^{226}_{88}\mathrm{Ra}$ ($T = 1600$ лет = $= 5{,}04 \cdot 10^{10}$ с):

$$A = \frac{4{,}17 \cdot 10^{23}}{226 \cdot 5{,}04 \cdot 10^{10}\,\text{с}} = 3{,}7 \cdot 10^{10}\ \text{Бк} = 1\ \text{Ки}.$$

Поскольку активность препарата прямо пропорциональна его массе, то для характеристики радиационной загрязненности объекта (например, твердых продуктов питания) вводят *удельную массовую активность* $A_m = A/m$ (Бк/кг) — отношение активности продукта к его массе.

Если рассматриваемый объект — жидкость или газ (например, вода или воздух), то содержание радионуклидов в нем характеризуют *удельной объемной активностью* $A_V = A/V$ (Бк/м^3, Бк/л, Ки/л) — отношением активности радионуклидов в объекте к его объему V.

Удельная поверхностная активность применяется для оценки радиационной загрязненности поверхностей и представляет собой отношение активности радионуклидов, содержащихся на поверхности образца, к площади S этой поверхности: $A_S = A/S$. Системная единица ее измерения Бк/м^2 слишком мала, поэтому в публикуемых картах радиоактивной загрязненности больших территорий данные приводятся в Ки/км^2.

Понятие удельной активности широко используется для оценки степени загрязненности радионуклидами окружающей среды, продуктов питания, воды, техники и других объектов.

28.4.2. Активность радионуклидов в живом организме

Воздействие на биологические объекты γ-квантов и частиц, возникающих при распаде радионуклидов, зависит прежде всего от количества распадов за единицу времени, т.е. от их активности в живом организме.

Уменьшение количества находящихся в организме радионуклидов обусловлено не только их радиоактивным распадом, но и биологическим выведением из организма. Количественно процессы выведения характеризуют постоянной биологического выведения $\lambda_б$ либо (чаще) *периодом биологического полувыведения* $T_б$, представляющим собой время, за которое количество радионуклидов в организме (или органе) уменьшается вдвое только вследствие выведения из него, причем $T_б = \ln 2/\lambda_б$.

Дифференциальное уравнение, описывающее убыль dN радионуклидов в живом организме за малое время dt, принимает вид

$$dN = -(\lambda_р + \lambda_б)Ndt \text{ , или } dN = -\lambda_{эф}Ndt \text{ ,} \qquad (28.21)$$

где $\lambda_{эф} = \lambda_р + \lambda_б$ — эффективная постоянная распада, равная сумме постоянных радиоактивного распада $\lambda_р$ и $\lambda_б$ биологического выведения.

Решение дифференциального уравнения (28.21) приводит к закону уменьшения количества нераспавшихся радионуклидов в живом организме, аналогичному формуле (28.14):

$$N = N_0 e^{-\lambda_{эф} t} . \qquad (28.22)$$

Вводя *эффективный период биологического полувыведения* $T_{эф}$ — время, за которое количество нераспавшихся радионуклидов в организме уменьшается вдвое, и учитывая, что $\lambda_{эф} = \dfrac{\ln 2}{T_{эф}} = \dfrac{0{,}69}{T_{эф}}$, получим

$$\frac{\ln 2}{T_{эф}} = \frac{\ln 2}{T_р} + \frac{\ln 2}{T_б} \quad \Rightarrow \quad \frac{1}{T_{эф}} = \frac{1}{T_р} + \frac{1}{T_б} .$$

Следовательно,

$$T_{эф} = \frac{T_б \cdot T_р}{T_б + T_р} . \qquad (28.23)$$

Эта формула показывает, что $T_{эф}$ всегда будет меньше наименьшей из величин $T_р$ и $T_б$. Если $T_р$ и $T_б$ значительно различаются (в десятки или сотни раз), то $T_{эф}$ будет практически равно наименьшему из них. Например, для ^{131}I период радиоактивного распада составляет $T_р = 8$ суток, а период биологического полувыведения $T_б = 138$ суток, т.е. $T_р << T_б$, и в соответствии с (28.23) $T_{эф} = 7{,}56$ суток, т.е. практически совпадает с $T_р$, оставаясь меньше его.

Если наоборот, $T_р >> T_б$, то $T_{эф}$ определяется практически только периодом биологического полувыведения $T_б$. Это характерно, например, для ^{137}Cs, у которого период полураспада 30 лет, а период полувыведения около 70 суток.

Если первоначально в организм поступило N_0 ядер радионуклида, то его активность в организме с учетом закона (28.22) будет изменяться по закону

$$A = -\frac{dN}{dt} = \lambda_{эф} N = \frac{0{,}69}{T_{эф}} N_0 e^{-\lambda_{эф} t}. \qquad (28.24)$$

Если выразить число исходных ядер N_0 через массу m радионуклида, его молярную массу M и число Авогадро N_A: $N_0 = \frac{m}{M} N_A$, и учесть, что $\lambda_{эф} = \frac{\ln 2}{T_{эф}}$, то получим

$$A = \frac{0{,}69m}{T_{эф}M} N_A e^{-\frac{\ln 2}{T_{эф}} t} = \frac{0{,}69m}{T_{эф}M} N_A \cdot 2^{-\frac{t}{T_{эф}}}. \qquad (28.25)$$

Видно, что активность радионуклида в организме прямо пропорциональна его массе m, обратно пропорциональна $T_{эф}$ и экспоненциально уменьшается со временем.

28.5. Взаимодействие ионизирующего излучения с веществом

Фотонное излучение и частицы при распространении в веществе взаимодействуют с атомами и молекулами вещества, изменяя его структуру и состояние воздействующих частиц. Кинетическая энергия частиц расходуется на ионизацию ато-

мов среды и другие структурные перестройки, не только ионизационной природы. Физические процессы, происходящие на первичных этапах взаимодействия, приводят к резкому изменению локальных биохимических реакций, формирующих организменный ответ на воздействие ионизирующих излучений.

28.5.1. Характеристики взаимодействия с веществом

Рассмотрим общие параметры, характеризующие взаимодействие ионизирующего излучения с веществом.

Линейная плотность ионизации. Основная часть энергии заряженных частиц и γ-квантов, взаимодействующих с веществом, идет на ионизацию его атомов и молекул. Степень ионизации зависит от свойств излучения, структуры объекта и расстояния l, которое квант излучения или частица проходит в веществе. Например, α-частицы и протоны создают максимальную ионизацию в конце пробега — перед моментом полной растраты своей энергии. На пути распространения этих частиц выделим произвольный тонкий слой вещества толщиной dl и допустим, что в нем образовалось число пар ионов dn (рис. 28.5). Количественной характеристикой ионизирующего эффекта является *линейная плотность ионизации* $i = dn/dl$ (пар ионов/см), равная отношению числа ионов одного знака (или числа пар ионов), образованных на пути dl, к величине этого пути. Подчеркнем, что линейная плотность ионизации зависит от пути, пройденного частицей в веществе: $i = f(l)$.

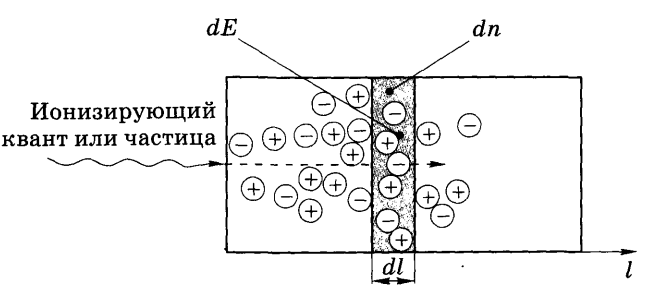


Рис. 28.5. Образование ионов в веществе при действии ионизирующего излучения

Линейная передача энергии (ЛПЭ). Энергия, передаваемая частицей веществу, расходуется не только на его прямую ионизацию, но и на другие структурные перестройки, характер и степень развития которых также зависят от вида излучения и состава вещества. Чем больше энергия dE, переданная излучением слою вещества толщиной dl, тем сильнее разрушительное действие ионизирующего излучения. Это воздействие ионизирующего излучения на вещество характеризуется ЛПЭ $= dE/dl$. В СИ единицей ее измерения является Дж/м, но часто эту величину измеряют в килоэлектрон-вольтах на микрометр воды (кэВ/мкм). Для незаряженных частиц ЛПЭ не применяется, но используются значения ЛПЭ вторичных заряженных частиц, образующихся при ионизации вещества.

Для одних и тех же частиц зависимости линейной плотности ионизации и ЛПЭ от расстояния, проходимого в веществе, обычно очень близки.

Средний линейный пробег частицы определяется средним значением расстояния, которое частица проходит в веществе до тех пор, пока ее кинетическая энергия не сравняется со средней кинетической энергией атомов и молекул вещества. Для электромагнитных излучений (γ-квантов, рентгеновских квантов) это определение не имеет смысла, поэтому прохождение их через вещество характеризуют ***слоем половинного ослабления*** $d_{1/2}$, после прохождения которого количество квантов, идущих в прежнем направлении, уменьшается в два раза.

28.5.2. Особенности взаимодействия с веществом различных частиц

Альфа-частицы обладают высокой ионизирующей способностью. По мере продвижения α-частицы в вещество плотность ионизации возрастает и, достигнув максимума почти в конце пробега, затем резко падает (рис. 28.6). Как видно из рисунка, линейная плотность ионизации в воздухе изменяется от $2 \cdot 10^4$ пар ионов/см в начале пробега до $8 \cdot 10^4$ пар ионов/см в его конце.

Суммарное число ионов, образованных в веществе за счет полного пробега α-частицы, может быть получено интегри-

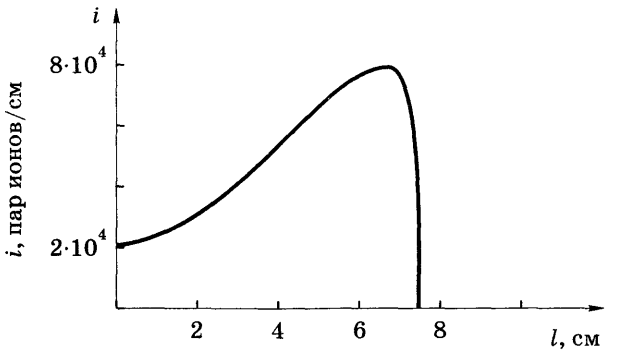


Рис. 28.6. Зависимость линейной плотности ионизации воздуха от расстояния, проходимого α-частицей

рованием зависимости $i(l)$ и численно равно площади под кривой на рис. 28.6. Средний линейный пробег α-частицы зависит от ее энергии и свойств вещества. В воздухе он равен нескольким сантиметрам, в жидкостях и биологических тканях — 10–100 мкм. Несмотря на малую глубину проникновения α-частиц в биологическую ткань, их разрушительное действие из-за высокой ионизирующей способности весьма значительно при попадании внутрь организма.

Кроме ионизации, при взаимодействии α-частиц с веществом возникает ряд других эффектов: возбуждение атомов, возникновение характеристического рентгеновского излучения, радиолюминесценция и т.д. Реже имеет место взаимодействие α-частиц с ядрами атомов, входящих в состав вещества. В этом случае возможны ядерные реакции.

Бета-частицы, обладающие меньшим зарядом и существенно меньшей массой по сравнению с α-частицами, отличаются и меньшей ионизирующей способностью. В воздухе линейная плотность ионизации, создаваемая потоком электронов, может быть рассчитана по формуле

$$i = k(c/v)^2,$$

где k — коэффициент пропорциональности ($k \approx 46$ пар ионов/см); c — скорость света в вакууме; v — скорость бета-частицы.

При распространении в веществе β-частицы изменяют свои первичные направления за счет рассеяния на электронах атомов вещества. Глубина проникновения β-излучения в мягкие ткани организма 10–15 мм.

Кроме ионизации, при торможении электронов в веществе может возникать тормозное рентгеновское излучение.

Позитроны являются античастицами по отношению к электронам, потому при попадании в вещество они при встрече с электронами *аннигилируют* с образованием двух γ-квантов с энергиями не менее 0,51 МэВ каждый (энергия покоя электрона) и летящих в противоположных направлениях:

$$ {}^{0}_{+1}e + {}^{0}_{-1}e \rightarrow 2\gamma. \qquad (28.26) $$

Гамма-излучение, представляя собой весьма короткие электромагнитные волны, является более жестким, чем рентгеновское излучение, но их первичные механизмы взаимодействия с веществом во многом сходны. Гамма-квант обладает значительно большей энергией, чем квант рентгеновского излучения, поэтому основные первичные механизмы его взаимодействия с веществом — фотоэффект и некогерентное рассеяние, при которых образуются рассеянные вторичные γ-кванты высоких энергий и электроны, ионизирующие среду. Кроме того, высокоэнергетичные γ-кванты ($h\nu > 1{,}02$ МэВ) в силовом поле ядер атомов могут вызывать образование электронно-позитронных пар:

$$ \gamma \rightarrow {}^{0}_{+1}e + {}^{0}_{1}e. \qquad (28.27) $$

Это процесс, обратный приведенной реакции аннигиляции.

Поскольку при распространении в веществе γ-излучение вызывает ряд вторичных процессов, экспоненциальный закон ослабления его интенсивности в данном случае выполняется приближенно, особенно для γ-квантов высоких энергий. Проникающая способность γ-излучения весьма высока — в воздухе средняя длина пробега квантов превышает десятки и даже сотни метров, в воде и мягких тканях — несколько метров, поэтому тело человека пронизывается ими практически полностью.

Нейтроны, несмотря на отсутствие заряда, обладают значительной кинетической энергией и также оказывают на вещество ионизирующее действие, в основном при столкновении с ядрами атомов и за счет других процессов. Таковым может быть деление ядер при захвате нейтронов, из-за которого возникают радиоактивные осколки, заряженные частицы и γ-излучение. Могут происходить ядерные реакции с выделением протонов и α-частиц.

Из-за отсутствия заряда нейтроны глубоко проникают в вещество, а защита от нейтронного облучения — сложная задача. Быстрые нейтроны необходимо сначала замедлить, для чего используются легкие ядра водородосодержащих веществ (например, вода H_2O, тяжелая вода D_2O и др.). Затем применяют вещества, активно поглощающие нейтроны (например, кадмиевые поглотители).

28.6. Принципы радионуклидных методов диагностики

В основе радионуклидной диагностики лежит использование радионуклидов или меченых ими химических соединений. Все радионуклидные методики связаны с введением радиофармацевтических препаратов в организм больного или в извлеченные из организма ткани и жидкости. Вводимые в организм радионуклиды обычно являются источниками γ-излучения, которое регистрируется специальными детекторами.

В составе ***радиодиагностических приборов*** выделяют три основных блока:

- *детектор*, преобразующий излучение от объекта в электрические импульсы;
- *электронный блок*, обеспечивающий необходимые преобразования импульсов детектора (усиление и др.) и их передачу на регистрирующее устройство;
- *регистрирующее устройство*, отображающее диагностическую информацию в той или иной форме; к нему могут прилагаться дополнительные устройства для обработки и анализа информации.

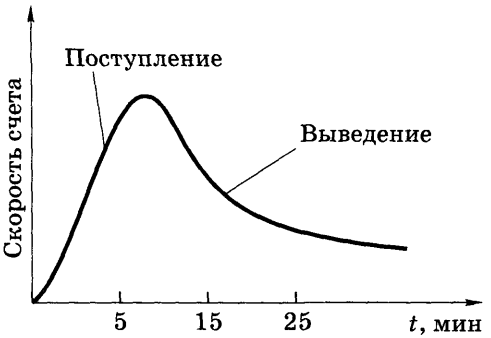


Рис. 28.7. Гамма-хронография почки

Методы радионуклидной диагностики условно можно разделить на три группы: радиографию, гамма-топографию и позитронно-эмиссионную томографию.

Радиография (гамма-хронография) — метод функционального исследования, в основе которого лежит изучение зависимости интенсивности γ-излучения органа (после введения радиофармацевтического препарата) от времени. Посредством таких исследований можно судить о движении крови по сосудам и камерам сердца, функционировании печени, почек, легких и т.д.

Рассмотрим пример, связанный с широко распространенным *радиографическим исследованием почек* (рис. 28.7). Радиоактивный препарат, введенный внутривенно, приносится с кровью в артерии почек. Это обусловливает быстрое появление γ-излучения в области почек (возрастающая часть кривой). Затем содержимое канальцев стекает и в течение нескольких минут практически создается равновесие между поступлением и убыванием радиоактивного препарата, чему соответствует вершина кривой. В дальнейшем отток начинает преобладать над поступлением и количество γ-квантов, поступающих на детектор за единицу времени, уменьшается — это вызывает снижение кривой. Анализ полученных таким методом данных дает подробную информацию о функционировании почки.

Гамма-топография, представляя собой клинико-анатомическое исследование, дает информацию о локализации, вели-

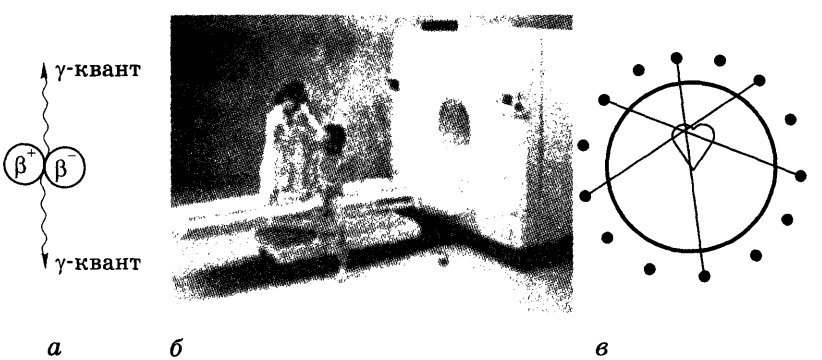
чине и положении органа, распределении в нем функционирующей ткани. Она основана на изучении пространственного распределения радионуклидов, накапливающихся в органе. С помощью данного метода распределение радионуклидов в организме получают в виде изображения.



Рис. 28.8. Гамма-топограмма печени, пораженной циррозом

В качестве примера возможностей гамма-топографии рассмотрим исследование печени. После внутривенного введения радиоактивного препарата печень здорового человека накапливает его достаточно равномерно. Если же часть функционирующей паренхимы печени заменена соединительной тканью (например, при циррозе), то в этих местах будет накапливаться существенно меньше радионуклидов, чем в здоровой части. Очагам поражения на гамма-топограмме будет соответствовать сниженная интенсивность гамма-излучений (рис. 28.8) [21].

Позитронно-эмиссионная томография (ПЭТ) — это новейший метод медицинской визуализации, основанный на применении радиофармпрепаратов (РФП), подверженных позитронному (β^+) распаду (рис. 28.9). РФП вводится в исследуе-



а *б* *в*

Рис. 28.9. Схема аннигиляции позитрона и электрона (*а*), общий вид позитронно-эмиссионного томографа (*б*) и расположение датчиков (•) схемы регистрации (*в*)

мый орган, где быстро распадается. Позитрон, испущенный в результате β^+-распада радиоизотопа, проходит в тканях не более 1–3 мм и сталкивается с электроном, в результате чего происходит их аннигиляция, при которой образуются два гамма-кванта с энергией $h\nu \geq 0{,}51$ МэВ, летящие в противоположные стороны:

$$\beta^+ + \beta^- = 2\ \gamma\text{-кванта}.$$

Блок системы регистрации томографа выделяет только те γ-кванты, которые одновременно зарегистрированы противоположно расположенными детекторами, определяя так называемую «линию ответа» (см. рис. 28.9, *в*). Затем компьютер суммирует все линии ответа от пар детекторов и реконструирует изображение области, испускающей β^+-частицы (позитроны), показывая тем самым функционирующую часть исследуемого органа.

На сегодняшний день в ПЭТ применяются позитрон-излучающие изотопы элементов с малым периодом полураспада (высокой активностью): ^{11}C ($T = 20{,}4$ мин), ^{13}N (10 мин), ^{15}O (2 мин), ^{18}F (110 мин). Фтор-18 обладает оптимальными характеристиками для использования в ПЭТ: наибольшим периодом полураспада и малой энергией позитронов.

Изображения, создаваемые ПЭТ, отражают функциональные процессы жизнедеятельности органов и тканей человека на молекулярном и клеточном уровне, включая метаболизм глюкозы и утилизацию кислорода, оценку кровотока и перфузии, оценку концентрации и сродства специфических рецепторов. Многие заболевания диагностируются при помощи ПЭТ намного раньше, чем на КТ и МРТ, поскольку функциональные изменения на клеточном уровне предшествуют морфологическим изменениям. ПЭТ является единственной методикой для оценки метаболических процессов *in vivo*.

В настоящее время ПЭТ применяется для диагностики главным образом в онкологии, неврологии и кардиологии.

В *онкологической* ПЭТ-диагностике наиболее широко используется радиофармпрепарат 18-фтордезоксиглюкоза (18F-ФДГ), что основано на особенности метаболизма злокачественных клеток: более интенсивном потреблении глюкозы по

сравнению с нормальными клетками. Это позволяет визуализировать злокачественные новообразования и метастазы на фоне здоровых тканей как «горячие» очаги и проводить исследования в режиме «все тело», позволяющем не только выявить первичный опухолевый очаг и определить степень его злокачественности, но и оценить распространенность метастатического поражения других тканей.

В *неврологии* метод ПЭТ применяется для выявления функциональных изменений головного мозга при его сосудистых заболеваниях, деменции, а также для дифференциальной диагностики очаговых образований.

В *кардиологии* ПЭТ применяется для определения метаболизма глюкозы в миокарде и жизнеспособности кардиомиоцитов. Результаты ПЭТ имеют принципиальное значение в определении показаний для кардиохирургического вмешательства, его объема и оценки проведенного лечения.

Таким образом, ПЭТ открывает перед врачами новые возможности:

- ранней диагностики различных заболеваний до появления структурных изменений, что существенно улучшает прогноз;
- проведения мониторинга терапии с ранней оценкой эффективности проводимого лечения, что позволяет сократить длительность лечения и в ряде случаев избежать госпитализации больного;
- замены серии исследований одним, наиболее информативным;
- прогнозирования результатов хирургического лечения;
- выявления отдаленных метастазов, что может коренным образом повлиять на тактику лечения.

28.7. Физические основы лучевой терапии

Лучевая терапия — клиническая дисциплина, изучающая применение ионизирующих излучений для лечения болезней. Методы лучевой терапии являются одними из ведущих, применяющихся в комплексном лечении больных со злокачественными новообразованиями.

При использовании лучевой терапии необходимо наиболее интенсивно воздействовать излучением на очаг патологии при минимальном поглощении энергии в окружающих его здоровых тканях и органах. Это весьма серьезная проблема, и некоторые приемы ее разрешения для каждого конкретного случая лечения предусматривают оптимальный выбор излучения с определенными физическими свойствами и способом воздействия.

Рассмотрим лишь некоторые особенности излучений, применяемых для лучевой терапии.

Рентгеновское излучение низких и средних энергий применяют при воздействии на поверхностно расположенные образования или при лечении заболеваний, не требующих высокоинтенсивного облучения. Это вызвано тем, что в ткани интенсивность такого рентгеновского излучения достаточно резко убывает и в 10 см от поверхности обычно не превышает 20 % интенсивности падающего пучка рентгеновских лучей (рис. 28.10).

Гамма-излучение вызывает ионизацию молекул биоткани на любом расстоянии от поверхности тела. Это позволяет разрушать глубоко расположенные опухоли, облучая их

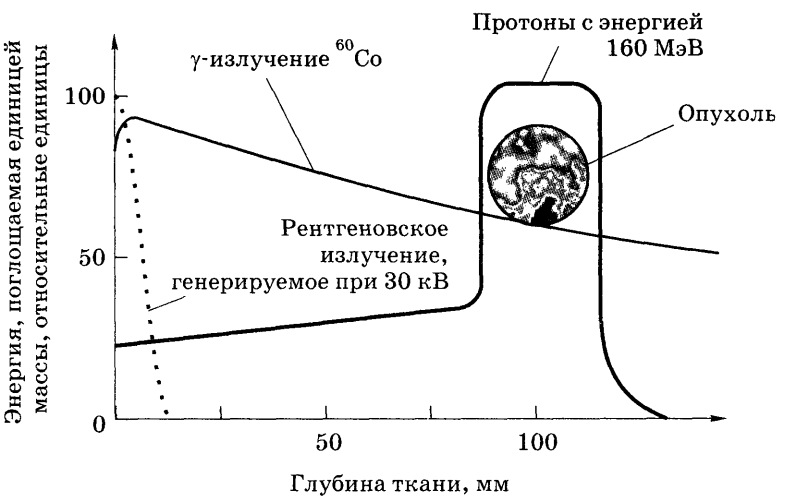


Рис. 28.10. Зависимость относительной величины поглощенной энергии некоторых излучений от глубины проникновения в ткани

по разным направлениям, сходящимся в области опухоли, одновременно снижая риск повреждения кожи и окружающих здоровых тканей. Источником излучения в γ-терапии обычно являются радионуклиды $^{60}_{27}Co$, создающие при своем распаде, наряду с электронами, высокоэнергетичные γ-кванты.

Электронные пучки с энергией 25 МэВ создают максимальную ионизацию на глубине 1–3 см и используются для облучения неглубоко лежащих патологических очагов. До столь значительных энергий электронные пучки разгоняются в ускорителях.

Ускорители заряженных частиц используются в лучевой терапии не только для сообщения высокой энергии частицам, непосредственно воздействующим на объект. В ускорителях электронов (бетатронах) за счет резкого торможения электронов, попадающих на мишень, получают жесткое тормозное рентгеновское излучение с высокой энергией квантов (десятки МэВ), которые используются для облучения. При этом максимальная ионизация тканей происходит на глубине 4–6 см, а ткани, расположенные ближе к поверхности тела, получают не более половины максимальной дозы.

Протоны, как и α-частицы, дают максимум ионизации в конце своего пробега в веществе. Это позволяет при протонном облучении подвести высокую энергию к облучаемому очагу без существенного облучения окружающих его нормальных тканей. Однако для обеспечения глубины проникновения в ткань на 11–14 см протонные пучки должны обладать энергией 120–160 МэВ.

Альфа-частицы из-за малого пробега в веществе могут использоваться в лучевой терапии при контакте с организмом или введении внутрь его. Характерным примером является радоновая терапия. Вода, содержащая $^{222}_{86}Rn$ и продукты его распада, используется для воздействия на кожу при приеме радоновых ванн.

Нейтроны с энергией 10–15 МэВ дают распределение производимой ионизации по глубине ткани, сходное с аналогичными эффектами при рентгеновском облучении. Максимум ионизирующего эффекта располагается на поверхности облучаемого участка тела.

Контрольные вопросы и задания

1. Какими параметрами характеризуются ядра атомов?
2. Каковы свойства ядерных сил?
3. Как определяется энергия связи нуклонов в ядре? Как зависит удельная энергия связи от массового числа?
4. На чем основано получение энергии в ядерных реакторах деления и синтеза?
5. Назовите основные виды радиоактивного распада.
6. Почему распад одних радионуклидов всегда сопровождается γ-излучением, а других — никогда?
7. Каковы энергетические спектры α- и β-частиц и γ-квантов, сопровождающих радиоактивный распад ядер?
8. Приведите примеры распада основных радиоизотопов, определяющих радиационное заражение среды после аварии на ЧАЭС.
9. Выведите основной закон радиоактивного распада. Что такое постоянная распада? период полураспада? время жизни радионуклида? Как они связаны между собой и в каких единицах измеряются?
10. Что такое активность радиопрепарата и в каких единицах она измеряется? Каковы соотношения между этими единицами?
11. Как изменяется активность во времени? Как она зависит от массы радионуклидов и периода их полураспада?
12. Как изменяется во времени активность радионуклидов в живом организме? Что характеризуют радиационный период полураспада, период биологического полувыведения и эффективный период? Как они связаны между собой? Какой из них наименьший?
13. Охарактеризуйте основные параметры, описывающие взаимодействие ионизирующего излучения с веществом, укажите единицы их измерения.
14. Каковы особенности взаимодействия с веществом α- и β-частиц, γ-квантов, нейтронов?
15. Какие радионуклидные методы диагностики вы знаете? На каких физических принципах они основаны?
16. Каковы физические основы лучевой терапии? Каковы особенности воздействия на организм рентгеновского и γ-излучения, α- и β-частиц, нейтронов, протонов?

Глава 29 Дозиметрия ионизирующих излучений

Биологические последствия воздействия ионизирующих излучений на клеточном, тканевом, органном и организменном уровнях определяются энергией воздействия, конкретным видом излучения и структурой объекта.

Для количественной оценки биологических эффектов, возникающих при действии ионизирующих излучений на биологические объекты, необходимо ввести объективные и доступные измерению физические характеристики излучения, от которых зависит степень развития этих эффектов.

Определение физических характеристик излучения с целью установления и прогнозирования радиационных воздействий на организм и представляет задачу ***дозиметрии*** — области прикладной физики, изучающей воздействие ионизирующих излучений на объекты живой и неживой природы, а также приборы и методы измерения этих величин.

29.1. Дозы излучения

В качестве физических эффектов, приводящих к установлению дозиметрических параметров, обычно используют ионизацию среды, возникающую под воздействием частиц и фотонных излучений, а также поглощение их энергии. Указанные эффекты по-разному проявляются в излучениях различной природы и разнообразных структурах объектов, которые подвергаются радиационным воздействиям.

29.1.1. Экспозиционная доза излучения

Наиболее просто характеризовать радиационные воздействия количественным выражением того основного эффекта, который ионизирующие излучения производят в веществе — степенью его ионизации, т.е. суммарным зарядом ионов, образовавшихся в единице массы вещества.

Экспозиционная доза излучения X представляет собой отношение суммарного заряда dQ ионов одного знака, образовавшихся в некоторой массе воздуха dm под действием рентгеновского или γ-излучения, к величине этой массы воздуха:

$$X = \frac{dQ}{dm}. \qquad (29.1)$$

В формуле (29.1) масса dm считается настолько малой, что распределение по ней заряда dQ равномерно. Если в любой массе m заряд Q распределен равномерно, то

$$X = \frac{Q}{m}. \qquad (29.2)$$

Подчеркнем, что экспозиционная доза вводится только для воздуха, рентгеновского и γ-излучения. Она характеризует радиационную обстановку вокруг объекта — степень ионизации окружающей воздушной среды электромагнитным излучением.

Из формул (29.1) и (29.2) видно, что системная единица измерения экспозиционной дозы 1 Кл/кг. На практике применяется и внесистемная единица — 1 рентген (Р) — это доза рентгеновского или γ-излучения, которая вызывает образование $2{,}08 \cdot 10^9$ пар ионов в 1 $см^3$ воздуха при нормальных условиях.

Установим связь между этими единицами. При экспозиционной дозе в 1 Р образуется заряд: $Q = nq = 2{,}08 \cdot 10^9 \times \times 1{,}6 \cdot 10^{-19} = 3{,}33 \cdot 10^{-10}$ Кл, где n — число пар ионов; q — заряд одного иона. Этот заряд образован в конкретной массе воздуха: $m = \rho V = 1{,}29 \cdot 10^{-6}$ кг ($\rho = 1{,}29$ кг/$м^3$ — плотность воздуха; $V = 1$ $см^3 = 10^{-6}$ $м^3$ — объем).

Следовательно,

$$1\ \text{Р} = \frac{3{,}33 \cdot 10^{-10}}{1{,}29 \cdot 10^{-6}} = 2{,}58 \cdot 10^{-4}\ \text{Кл/кг}$$

или 1 Кл/кг = 3876 Р.

Если экспозиционная доза X получена за время t, то *средняя мощность* этой дозы

$$\dot{X} = \frac{X}{t}. \qquad (29.3)$$

Если экспозиционная доза меняется со временем, то скорость этого изменения характеризует *мощность экспозиционной дозы* ($\dot{X}$), представляющей производную от дозы X по времени:

$$\dot{X} = \frac{dX}{dt}. \qquad (29.3а)$$

Системная единица измерения мощности экспозиционной дозы — 1 А/кг, внесистемные — 1 Р/ч, 1 мР/ч, 1 мкР/с и др. Эту величину измеряют многие дозиметры, и если мощность $\dot{X}$ дозы постоянна, то экспозиционная доза, полученная объектом за время τ, определяется произведением: $X = \dot{X}\tau$.

Если мощность дозы не постоянна, а зависит от времени $\dot{X} = f(t)$, то экспозиционная доза за время τ определяется интегрированием:

$$X = \int_0^{\tau} \dot{X}(t)dt. \qquad (29.4)$$

Для γ-излучающих радиоактивных препаратов установлена простая зависимость мощности экспозиционной дозы от расстояния r до препарата активностью A. От точечного источника γ-кванты вылетают с равной вероятностью по всем направлениям (рис. 29.1). Общее число этих квантов пропорционально числу актов распада за единицу времени, т.е. активности, а число квантов, попадающих на поверхность единичной площадки S, находящейся на расстоянии r от источника, обратно пропорционально квадрату расстояния r. Таким образом

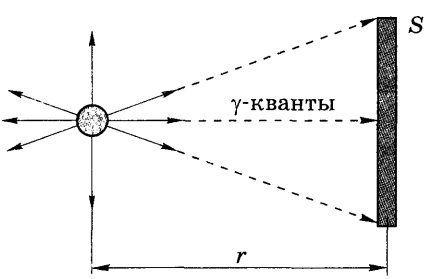


Рис. 29.1. Распространение γ-квантов от точечного источника

$$\dot{X} = K_{\gamma}\frac{A}{r^2}, \qquad (29.5)$$

где K_{γ} — γ-постоянная, характерная для каждого радионуклида и измеряемая в $\frac{А \cdot м^2}{кг \cdot Бк}$ или во внесистемных единицах

$\frac{\text{Р} \cdot \text{см}^2}{\text{ч} \cdot \text{мКи}}$. Следует отметить, что это соотношение применимо лишь для радионуклидов, при распаде которых возникают γ-кванты.

Итак, экспозиционная доза характеризует степень радиационного воздействия рентгеновского и гамма-излучения на воздух.

29.1.2. Поглощенная доза

Воздействие ионизирующего излучения на биологический объект определяет та часть излучения, которая поглощается объектом. Существенное биологическое действие оказывают ионизирующие кванты и частицы, которые возникают при распаде радионуклидов, содержащихся в объекте (внутреннее облучение). Поэтому для характеристики радиационного воздействия на организм вводится физическая величина, называемая ***поглощенной дозой ионизирующего излучения D*** — отношение энергии dE, переданной ионизирующим излучением веществу массой dm, к величине этой массы:

$$D = \frac{dE}{dm}. \qquad (29.6)$$

Если поглощенная энергия E равномерно распределена по массе m облучаемого объекта, то

$$D = \frac{E}{m}. \qquad (29.6а)$$

Единицей поглощенной дозы в СИ является грей (Гр): 1 Гр = 1 Дж/кг; часто используется несистемная единица поглощенной дозы — 1 рад, причем между ними простая связь: 1 Гр = 100 рад (рад — от англ. **r**adiation **a**bsorbed **d**ose — поглощенная доза излучения).

Отметим, что разовая доза в 6 Гр вызывает смерть всех облученных людей, не получивших специализированного медицинского лечения (доза 100-процентной летальности).

Мощность поглощенной дозы определяется как $\dot{D} = \frac{dD}{dt}$ и измеряется в Гр/с, рад/с или производных от них единицах.

Средняя мощность дозы D, полученной за время t, рассчитывается как $\dot{D} = \frac{D}{t}$.

Экспериментальное определение поглощенной дозы представляет весьма сложную задачу и во многих случаях невозможно. Поэтому оценить поглощенную дозу можно по легко измеряемой экспозиционной дозе, поскольку при внешнем облучении объекта поглощенная доза пропорциональна экспозиционной:

$$D = fX, \tag{29.7}$$

Коэффициент f зависит от ряда факторов, прежде всего от структуры облучаемого объекта, энергии квантов и используемых единиц измерения.

Численные значения коэффициента f оказываются очень удобными, если измерять дозы во внесистемных единицах: X — в рентгенах (Р), D — в радах. Тогда коэффициент f принимает следующие простые значения:

- для воздуха f = 0,88 рад/Р;
- для воды и мягких биотканей f = 1,0 рад/Р;
- для костной ткани f = 1 – 4,5 рад/Р.

Этот коэффициент довольно сильно зависит от энергии квантов, уменьшаясь с увеличением их энергии.

Значение коэффициента f для воздуха не сложно установить. Для этого рассчитаем энергию, поглощенную 1 см^3 воздуха при экспозиционной дозе X = 1 Р. В этом случае в 1 см^3 воздуха массой $m = 1{,}29 \cdot 10^{-6}$ кг образуется $2{,}08 \cdot 10^9$ пар ионов. Известно, что на образование одной пары (на ионизацию 1 молекулы воздуха) расходуется энергия 34 эВ = $= 34 \cdot 1{,}6 \cdot 10^{-19}$ Дж. Тогда поглощенная доза

$$D = \frac{E}{m} = \frac{34 \cdot 1{,}6 \cdot 10^{-19} \cdot 2{,}08 \cdot 10^9 \text{ Дж}}{1{,}29 \cdot 10^{-6} \text{ кг}} = 88 \cdot 10^{-4} \text{ Гр} = 0{,}88 \text{ рад}.$$

Таким образом, при экспозиционной дозе X = 1 Р поглощенная доза в воздухе составит 0,88 рад. Следовательно, для этого случая коэффициент f = 0,88 рад/Р.

29.1.3. Эквивалентная доза

Биологический эффект воздействия ионизирующих излучений зависит не только от величины энергии, поглощенной в веществе. На степень радиационного поражения существенно влияет вид излучения, поскольку при одной и той же поглощенной дозе α-, β- и γ-излучения обладают различной ионизирующей способностью. Поэтому при одинаковой поглощенной дозе степень поражения организма будет разной в зависимости от вида воздействовавшего ионизирующего излучения.

Для сравнения биологических эффектов, вызываемых различными видами излучения, используется понятие ***относительной биологической эффективности излучения (ОБЭ)***. Для нахождения ОБЭ необходимо выбрать некоторое излучение за эталонное и определить поглощенную дозу, вызывающую определенный биологический эффект. Затем воздействовать на такой же объект ионизирующим излучением другого вида и определить для него поглощенную дозу, вызывающую тот же биологический эффект, что и эталонное излучение.

В качестве эталонного излучения выбирают рентгеновские лучи с энергией фотонов 180–200 кэВ. С выбором биологического эффекта и его количественной оценкой дело обстоит сложнее. Для этого используется, например, критерий $ЛД_{50/30}$ — критерий полулетальной дозы для животных. Это такая доза, которая вызывает гибель половины (50 %) подвергшихся облучению животных определенного вида в течение 30 суток после облучения. Кроме того, могут быть выбраны и другие радиобиологические эффекты: хромосомные аберрации, частота возникновения лейкемии, нарушения в формуле крови и т.п.

Таким образом, относительная биологическая эффективность излучения определяется соотношением

$$ОБЭ = \frac{\text{Поглощенная доза рентгеновского излучения (180–200 кэВ), вызывающая биологический эффект}}{\text{Поглощенная доза ионизирующего излучения другого вида, вызывающая тот же биологический эффект}}$$

Пример. Допустим, что облучение стандартным рентгеновским излучением некоторого количества экспериментальных животных вызвало летальный исход для половины из них в течение 30 суток при поглощенной дозе 5 Гр. Воздействие на такое же количество аналогичных животных α-частицами обусловило тот же биологический эффект (гибель половины из них за то же время) в дозе 0,25 Гр. Тогда ОБЭ для α-частиц равна 5/0,25 = 20.

Регламентированное значение ОБЭ представляет собой *коэффициент качества* (k) излучения, учитывающий различия в степени радиационного воздействия излучений различного вида.

Коэффициент качества для разных ионизирующих излучений имеет следующие значения:

- для рентгеновского, γ- и β-излучения $k = 1$;
- для α-частиц с энергией, меньшей 10 МэВ, $k = 20$;
- для протонов $k = 10$;
- для нейтронов с энергией 0,1–5 МэВ $k = 3–10$.

Чтобы учесть разную относительную биологическую эффективность излучения, вводится ***эквивалентная доза.*** При одинаковой эквивалентной дозе, полученной живым организмом, биологический эффект воздействия тоже будет одинаковым, независимо от вида ионизирующего излучения.

Эквивалентная доза H ионизирующего излучения представляет собой произведение поглощенной дозы D и коэффициента качества излучения:

$$H = kD. \qquad (29.8)$$

В СИ единицей эквивалентной дозы является зиверт (Зв) — такая эквивалентная доза, при которой произведение поглощенной дозы в биологической ткани на коэффициент качества равно 1 Дж/кг. Из определения эквивалентной дозы следует, что зиверт и грей имеют одинаковую размерность (Дж/кг), но они представляют разные дозы!

Если в формулу (29.8) подставлять поглощенную дозу в греях, то после умножения на k получим эквивалентную дозу в зивертах. Если же поглощенная доза измерена в радах, то соответствующей внесистемной единицей эквивалентной дозы является бэр (биологический эквивалент рада), 1 бэр = 0,01 Зв.

Мощность эквивалентной дозы $\dot{H}$ — отношение приращения эквивалентной дозы dH за интервал времени dt к величине этого интервала:

$$\dot{H}=\frac{dH}{dt}. \qquad (29.9)$$

Если эквивалентная доза прямо пропорциональна времени, то мощность дозы определяется отношением дозы ко времени:

$$\dot{H}=\frac{H}{t}.$$

Единицы измерения мощности эквивалентной дозы — Зв/с, мЗв/ч, бэр/с, мбэр/ч и т.д.

Если мощность дозы меняется во времени $\dot{H}=f(t)$, то эквивалентная доза облучения за время τ находится интегрированием:

$$H=\int_0^{\tau}\dot{H}(t)dt.$$

Именно полная эквивалентная доза определяет степень радиационного поражения органа.

29.1.4. Эффективная эквивалентная доза

Различные органы и ткани имеют разную радиочувствительность, поэтому реакция организма в целом на воздействие ионизирующих излучений зависит от того, какой орган получил какую дозу. Со случаями локального облучения отдельных участков тела сталкиваются, например, при лучевой терапии. Учет радиочувствительности отдельных органов необходим и при оценке доз внутреннего облучения, поскольку отдельные нуклиды специфично накапливаются в различных органах (^{90}Sr — в основном в костной ткани; ^{131}I — в щитовидной железе и т.д.), вызывая их преимущественное облучение.

Эффективная эквивалентная доза $H_{эф}$ представляет собой сумму произведений эквивалентных доз, полученных отдельными органами организма H_i, и соответствующих им коэффициентов радиационного риска w_i:

$$H_{эф}=\sum_i w_i H_i\,. \qquad (29.10)$$

Коэффициент радиационного риска w_i (взвешивающий фактор) представляет собой отношение вероятностного риска летального исхода при облучении i-го органа в некоторой эквивалентной дозе к риску смерти от равномерного облучения всего организма в той же эквивалентной дозе.

Пример. Если легкие человека подверглись облучению в эквивалентной дозе 1 Зв, то риск (вероятность) смерти от рака легких составит $P_л = 2 \cdot 10^{-3}$. Если же все органы человека получили равномерно 1 Зв, то вероятность летального исхода составит $P_о = 1{,}65 \cdot 10^{-2}$. Коэффициент радиационного риска для легких

$$w_л = \frac{P_л}{P_о} = \frac{2 \cdot 10^{-3}}{1{,}65 \cdot 10^{-2}} \approx 0{,}12 .$$

Итак, коэффициент w_i определяет вклад радиационного поражения данного органа в риск неблагоприятных последствий для организма при равномерном облучении (табл. 29.1).

Сумма всех коэффициентов радиационного риска равна единице: $\sum_i w_i = 1$.

Таблица 29.1

Орган	w_i
Гонады	0,25
Молочная железа	0,15
Красный костный мозг	0,12
Легкие	0,12
Щитовидная железа	0,03
Поверхность кости	0,03
Все другие органы вместе	0,30

29.1.5. Коллективная доза

Эквивалентная доза характеризует последствия радиационного воздействия для отдельных органов, эффективная эквивалентная доза — соответствующие последствия для организма в целом. Для оценки радиационных последствий на уровне большого контингента населения вводится поня-

тие ***коллективной дозы***, или ***коллективной эффективной эквивалентной дозы*** S. Эта величина является объективной оценкой масштаба радиационного поражения большого коллектива людей и представляет собой сумму эффективных эквивалентных доз $H_{i\,эф}$, полученных различными индивидуумами:

$$S = \sum_i H_{i\,эф}\, N_i, \qquad (29.11)$$

где N_i — число лиц в данной группе, получивших дозу $H_{i\,эф}$.

Единица измерения коллективной дозы — человеко-зиверт.

Введение коллективной дозы дает возможность применять статистические методы для определения числа явных и слабо выраженных последствий облучения. Определенная индивидуальная доза вызывает развитие некоторого заболевания данного индивидуума лишь с какой-то вероятностью. При большом числе облученных лиц предсказанное число заболеваний должно осуществляться в действительности.

Для населения, продолжающего жить в условиях длительного хронического облучения, возможен расчет ***ожидаемой коллективной эквивалентной дозы*** на определенный период. Точность такого прогноза (и, следовательно, прогнозирования состояния здоровья населения) зависит от полноты учета характеристик радиационного заражения, закономерностей изменения радиационной обстановки, рациона питания, гигиенического режима и других факторов. Выполнение подобных прогнозов представляет достаточно сложную задачу.

29.2. Принципы расчета доз внутреннего облучения

Радионуклиды, попавшие внутрь организма, создают повышенную опасность его радиационного поражения. Это определяется рядом факторов.

Во-первых, радионуклиды, излучающие α- и β-частицы во внешней среде и находящиеся вне организма, не представляют большой опасности для его внутренней среды, так как эти излучения обладают низкой проникающей способностью. Но, попадая внутрь организма, в основном при дыхании и потреблении воды и пищи, эти радионуклиды своим высоко-

ионизирующим излучением вызывают в нем мощные радиационные повреждения.

Во-вторых, многие радионуклиды избирательно накапливаются в отдельных органах и возникающие при их распаде ионизирующие излучения отдают свою энергию сравнительно небольшому объему ткани, вызывая в нем значительные локальные повреждения.

В-третьих, радионуклиды, обладающие большим периодом полураспада и полувыведения, однажды попав в организм, оказывают свое разрушительное действие на него в течение длительного времени.

Дозиметрической величиной, определяющей степень биологических последствий воздействия ионизирующих излучений на органном уровне, является эквивалентная доза. Она непосредственно не измеряется, а определяется расчетным путем. Основными данными для расчета получаемых доз внутреннего облучения являются сведения об удельной активности воздуха, воды и продуктов питания.

29.2.1. Связь активности радионуклидов и эквивалентной дозы внутреннего облучения

Эквивалентная доза облучения H определяется произведением поглощенной дозы D и коэффициента качества: $H = kD$. Напомним, что поглощенная доза определяется энергией E ионизирующего излучения, поглощенной единицей массы вещества: $D = E/m$. Энергия, поглощенная веществом, $E = E_1 N_р$, где E_1 — энергия, выделяемая при одном акте распада, $N_р$ — число распадов. Таким образом, для эквивалентной дозы в органе массы m получаем выражение

$$H = \frac{k}{m} E_1 N_р.$$

Значения энергии одного распада E_1 для конкретного радионуклида известны из справочной литературы. Однако для биологического действия излучений, возникающих при распаде, важно еще и то, в каком конкретном органе происходит распад радионуклида и вид излучения. Поэтому для каждого радионуклида в справочной литературе приводятся данные

(МэВ на распад), характеризующие эффективную энергию одного распада в соответствующем органе $E_{эф}$. В величине $E_{эф}$ уже учитывается коэффициент качества, что упрощает проведение расчетов.

Итак, для эквивалентной дозы внутреннего облучения в органе получаем

$$H = \frac{1}{m} E_{эф} N_р. \qquad (29.12)$$

Число распадов $N_р$ определяется активностью A радионуклидов в организме

$$dN_р = A(t)dt,$$

и за время τ в организме произойдет количество распадов, равное

$$N_р = \int_0^\tau A(t)dt. \qquad (29.13)$$

Подставляя это выражение в формулу (29.12), найдем эквивалентную дозу, полученную органом массой m за время τ при внутреннем облучении:

$$H = \frac{E_{эф}}{m} \int_0^\tau A(t)dt \qquad (29.14)$$

Если активность A радионуклидов в организме не меняется со временем (очень большой период полураспада и полувыведения), то эквивалентная доза растет прямо пропорционально времени τ нахождения радионуклида в организме:

$$H = \frac{1}{m} E_{эф} A\tau. \qquad (29.15)$$

Для практического применения формулы (29.14) необходимо знать закон изменения активности радионуклидов $A(t)$ в организме. В общем случае этот закон может быть достаточно сложным, поскольку необходимо учитывать процессы физического распада радионуклидов, их выведение из организма, участие в метаболизме и иметь количественную характеристику их поступления в организм.

29.2.2. Расчет эквивалентной дозы внутреннего облучения при однократном поступлении радионуклидов в организм

При однократном поступлении радионуклидов в организм их активность изменяется по закону (28.24)

$$A(t) = A_0 e^{-\lambda_{эф} t}, \qquad (29.16)$$

где $A_0 = \lambda_р N_0$ — начальная активность поступившего в организм радионуклида; N_0 — начальное количество ядер радионуклида; $\lambda_р$ — постоянная радиационного распада; $\lambda_{эф}$ — эффективная постоянная, определяющая скорость исчезновения радионуклида из организма [см. (28.22)].

Для нахождения эквивалентной дозы, полученной организмом за время τ, подставим выражение (29.16) в (29.14):

$$H = \frac{E_{эф}}{m} A_0 \int_0^{\tau} e^{-\lambda_{эф} t}\, dt\,. \qquad (29.17)$$

Производя интегрирование и учитывая, что $\lambda_{эф} = 0{,}69/T_{эф}$, окончательно получим

$$H = \frac{E_{эф} A_0 T_{эф}}{0{,}69\, m}\left(1 - e^{-\frac{0{,}69\tau}{T_{эф}}}\right) = H_{max}\left(1 - e^{-\frac{0{,}69\tau}{T_{эф}}}\right). \quad (29.18)$$

Здесь через H_{max} обозначена максимальная эквивалентная доза, которая может быть получена органом массой m за неограниченное время от некоторого радионуклида с начальной активностью A_0:

$$H_{max} = 1{,}44 \frac{E_{эф} A_0}{m} T_{эф}\,. \qquad (29.19)$$

Видно, что максимальная эквивалентная доза прямо пропорциональна не только начальной активности A_0 радионуклида, попавшего в организм, но и эффективному периоду $T_{эф}$ полуисчезновения радионуклида из организма. Эквивалентная доза, полученная за некоторое время τ, также существенно зависит от $T_{эф}$ (рис. 29.2).

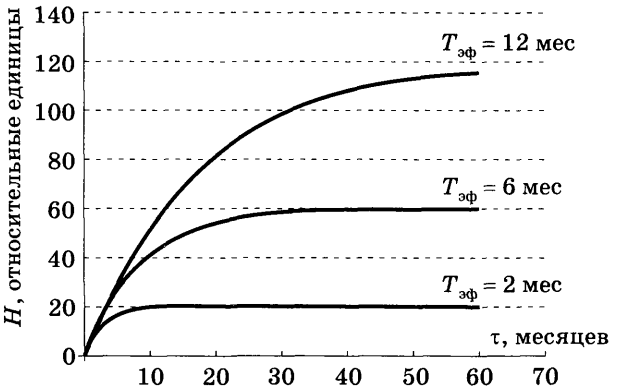


Рис. 29.2. Зависимость эквивалентной дозы H от времени τ, прошедшего после однократного попадания радионуклидов активностью A_0 в организм, при разных значениях $T_{эф}$

Из анализа выражения (29.18) следует, что за время $\tau = 5T_{эф}$ организм получит 97 % максимальной дозы H_{max}, а за время $\tau = 7T_{эф}$ – 99,3 % H_{max}. Это также хорошо видно из графиков, приведенных на рис. 29.2.

Таким образом, чем меньше $T_{эф}$, тем меньшую эквивалентную дозу внутреннего облучения получит организм от попавшего в организм радионуклида начальной активностью A_0.

29.3. Детекторы ионизирующих излучений

Детекторами ионизирующих излучений называют устройства, предназначенные для обнаружения излучений и частиц, определения состава излучения и измерения его энергетического спектра.

Для реализации детекторов используют разнообразные эффекты, возникающие при взаимодействии излучения с веществом.

Следовые детекторы позволяют определять траекторию частицы и длину ее пробега в веществе. Так, в *камере Вильсона*, заполненной перегретым паром, движущаяся частица оставляет след в виде мелких капелек жидкости на центрах конденсации, например на ионах, образующихся на пути дви-

жения заряженной частицы. Капельки достигают видимых размеров и могут быть сфотографированы. Природу и свойства исследуемых частиц устанавливают по величине пробега и искривлению следов частицы под действием магнитного поля. В *пузырьковой камере* след частицы образуется за счет вскипания перегретой жидкости вдоль траектории частицы. Этот след отмечается цепочкой пузырьков пара и может быть сфотографирован. В *эмульсионных камерах* регистрация частиц производится по следу, который они оставляют, проникая в толстослойные фотографические эмульсии.

Действие широко распространенных ***ионизационных детекторов*** основано на измерении числа и параметров электрических импульсов, возникающих из-за движения ионов, образовавшихся в газе или полупроводниковых материалах под действием излучений. Например, детектор, используемый в *счетчике частиц Гейгера* (рис. 29.3), представляет собой цилиндрическую трубку, наполненную газом при давлении 100–200 мм рт.ст. По оси трубки укрепляется на изоляторах анод в виде тонкой нити. Вторым электродом (катодом) служит корпус счетчика. Зачастую это устройство заключается в оболочку из стекла. Между катодом и анодом газоразрядной трубки создается разность потенциалов. При попадании частицы в детектор в газе образуется некоторое количество ионов. Под действием разности потенциалов между катодом и анодом возникает импульс тока, протекающего по сопротивлению R. Амплитуда этого импульса пропорциональна числу ионов, участвующих в его образовании, и зависит от напряжения.

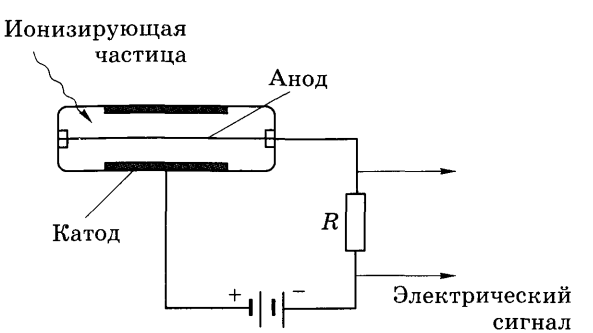


Рис. 29.3. Принципиальная схема ионизационного детектора

Не производя детального анализа работы этого детектора, отметим, что по амплитуде электрического сигнала можно судить об энергии ионизирующих частиц, а по числу импульсов — об их количестве.

Детектирование фотонного излучения и частиц ионизационными детекторами производится по степени ионизации в определенном объеме вещества. Рентгеновские и γ-кванты, имеющие большую длину пробега в газе, в небольшом его объеме редко производят ионизацию. Они преимущественно выбивают электроны из атомов вещества стенок трубки, которые, попадая в газ, его ионизируют.

Сцинтилляционные детекторы (рис. 29.4) основаны на регистрации вспышек света, возникающих при попадании на вещество (сцинтиллятор) ионизирующих излучений. При взаимодействии со сцинтиллятором ионизирующие излучения порождают его видимую люминесценцию. С помощью чувствительных фотоэлектрических устройств (фотоумножителей) световые сигналы пропорционально преобразуются в электрические и обрабатываются с помощью электроизмерительной техники. В детекторах небольших размеров сцинтилляторы наносятся непосредственно на катод фотоумножителя. В ряде случаев между сцинтиллятором и фотокатодом помещают световод. Амплитуда электрического сигнала фотоумножителя пропорциональна интенсивности световой вспышки, которая, в свою очередь, пропорциональна энергии частицы. Поэтому, выбрав для каждого вида излучений

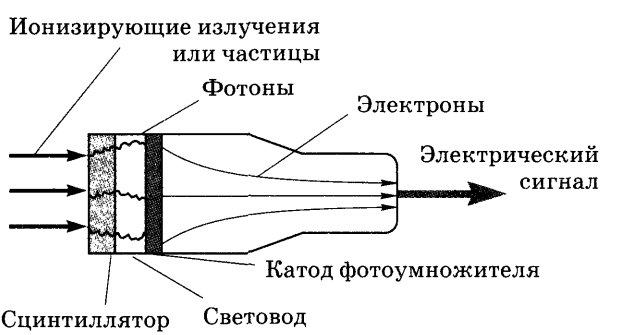


Рис. 29.4. Принципиальная схема сцинтилляционного детектора

соответствующие сцинтилляторы, можно определять энергетические спектры излучений.

Принцип действия фотографических детекторов основан на том, что степень почернения дозиметрической фотопленки в некотором диапазоне почернений пропорциональна экспозиционной дозе.

Рассмотрим теперь ***особенности детектирования*** различных ионизирующих частиц (рис. 29.5).

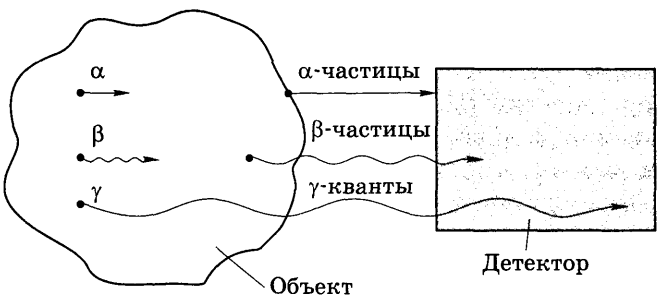


Рис. 29.5. Особенности детектирования частиц и γ-квантов

Наиболее сложна регистрация α-*частиц*. Это связано с малой длиной их пробега в веществе — порядка нескольких микрометров. Поэтому стенки детектора должны быть крайне тонкими, чтобы α-частицы могли попасть в рабочее вещество, вызвать его ионизацию и обусловить последующее формирование электрического импульса. Естественно, что при исследовании радиоактивности различных объектов α-излучение может быть зарегистрировано лишь с их тонкого поверхностного слоя. Поэтому необходима специальная подготовка образцов для исследования (изготовление тонких срезов, растворение, нанесение на подложку и др.).

Детектирование β-*частиц*, обладающих большей длиной пробега, несколько проще. Однако надо иметь в виду, что длина пробега β-частиц зависит от их энергии. Поэтому обнаружить β-частицы, возникающие при распаде $^{137}_{55}\mathrm{Cs}$, проще, чем при распаде $^{90}_{38}\mathrm{Sr}$, поскольку максимальная энергия последних намного меньше, и данные β-частицы практически полностью поглощаются миллиметровыми слоями вещества.

Наиболее просто детектируется γ-*излучение*, возникающее при радиоактивном распаде, так как γ-кванты обладают большой длиной пробега. Они могут быть зарегистрированы даже в том случае, когда излучающие их радионуклиды расположены глубоко внутри исследуемого образца.

Так, при распаде $^{131}_{53}I$ за единицу времени возникает большое число γ-квантов высокой энергии (активность высока из-за малого периода полураспада). Поэтому детектирование γ-излучения этого радионуклида, находящегося в окружающей среде или внутри организма, не представляет технических сложностей. Регистрация γ-квантов, образующихся при распаде $^{137}_{55}Cs$, также не вызывает особых проблем.

Для оценки степени радиационного заражения среды и организма важно знать не только общую активность, но и радионуклидный состав радиационного заражения, т.е. знать, какие конкретно радионуклиды его обусловливают и каково количественное соотношение их содержания в среде и организме. Для γ-активных радионуклидов эта проблема решается измерением энергетического спектра γ-излучения — зависимости его интенсивности от энергии γ-квантов. Каждому радиоактивному превращению таких радионуклидов соответствует свой набор γ-квантов вполне определенных энергий — *энергетический спектр*.

Приборы, с помощью которых измеряются спектры гамма-излучения, называют γ-***спектрометрами***.

29.4. Дозиметрические приборы

Дозиметрические приборы — устройства для измерения доз ионизирующих излучений или активности объектов, содержащих радионуклиды.

Они подразделяются на дозиметры и радиометры.

Дозиметры — приборы, предназначенные для измерения дозы (прежде всего экспозиционной). Дозиметр состоит из детектора и электронно-измерительного устройства. Электронно-измерительное устройство осуществляет преобразование сигнала детектора в форму, удобную для последующей регистрации.

Рассмотрим принцип устройства распространенных *конденсаторных дозиметров* (рис. 29.6). Основной их частью

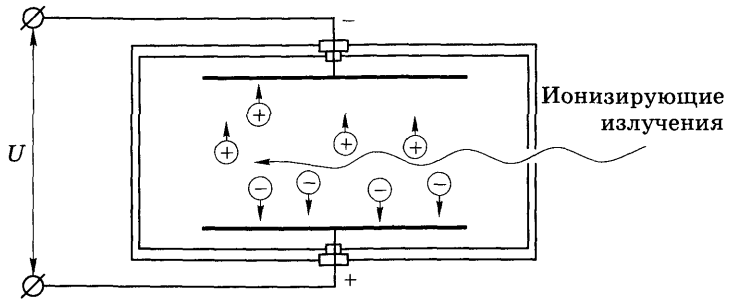


Рис. 29.6. Принципиальная схема конденсаторного дозиметра

является ионизационная камера, заполненная воздухом при атмосферном или пониженном давлении, рабочим объемом V. В камере располагаются электроды, представляющие, в простейшем случае, пластины плоского конденсатора. Эти электроды перед началом использования дозиметра заряжаются до разности потенциалов U_1 и приобретают заряд q_1. Под действием ионизирующего излучения в камере образуются ионы. Из-за имеющейся разности потенциалов в камере возникает ток и разность потенциалов между пластинами конденсатора уменьшается от своего первоначального значения U_1 до U_2, а заряд — до значения q_2. Изменение заряда связано с изменением разности потенциалов через емкость конденсатора C:

$$q_1 - q_2 = \Delta q = C(U_1 - U_2).$$

Поскольку уменьшение заряда конденсатора обусловлено суммарным зарядом ионов одного знака, образованных в воздухе действием ионизирующего излучения, то экспозиционная доза

$$X = \frac{\Delta q}{m} = \frac{C(U_1 - U_2)}{\rho V} = k(U_1 - U_2) = k\Delta U,$$

где m — масса воздуха; ρ — плотность воздуха; V — постоянный объем воздуха в камере. Эти постоянные величины объединены в коэффициент k, устанавливаемый при градуировке прибора.

Таким образом, экспозиционная доза пропорциональна изменению напряжения ΔU и определяется по двум значени-

ям напряжения — при зарядке и после облучения в течение некоторого времени. Градуируют приборы сразу в единицах экспозиционной дозы.

Индивидуальные конденсаторные дозиметры исполняются в виде миниатюрных устройств (например, в виде цилиндрического конденсатора размером с авторучку) и позволяют определять индивидуальные дозы, полученные персоналом, работающим в условиях внешнего радиационного облучения.

Измерение мощности экспозиционной дозы основано на определении с помощью соответствующих электронных схем и регистрирующих устройств силы тока I в ионизационной камере, полупроводнике (для полупроводникового детектора) или тока фотоэлектрического приемника (для сцинтилляционных детекторов). Это следует из того, что мощность экспозиционной дозы (см. 29.2)

$$\dot{X} = \frac{dX}{dt} = \frac{dQ}{mdt} = \frac{I}{m}.$$

Измерители мощности экспозиционной дозы градуируются чаще всего в мР/ч или мкР/с.

Подчеркнем, что дозиметры предназначены для измерения экспозиционной дозы в воздухе, позволяют контролировать уровень фона рентгеновского и γ-излучения и не могут применяться для контроля степени радиационной загрязненности пищевых продуктов и организма.

Радиометры — приборы для измерения активности. Поскольку активность определяется числом актов распада за единицу времени, то в радиометрах осуществляется счет электрических импульсов, вызванных попаданием частиц на детектор за определенное время.

Рассмотрим подробнее распространенный способ определения удельной объемной или массовой активности продуктов питания. Детекторы радиометра и помещенные под ними исследуемые объекты располагаются в камере с толстыми свинцовыми стенками, так как даже при отсутствии в измеряемом объекте радионуклидов на детектор радиометра попадают ионизирующие частицы из-за имеющегося естественного фона, от которого необходимо экранироваться. До начала измерений радиоактивности объекта в камере необходимо измерить начальную фоновую активность с пустой измерительной кюве-

той. Допустим, что за время t_1 счетчик радиометра показывает число импульсов фоновой активности N_1. Затем измерительная кювета, имеющая строго фиксированный объем, заполняется исследуемым продуктом. Допустим, что за время измерений t_2 счетчик показывает теперь число импульсов N_2. Если времена t_1 и t_2 одинаковы, то активность продукта тем больше, чем больше разница $N_2 - N_1$. Расчет удельной объемной активности (A_v) производится по формуле

$$A_v = \frac{N_2/t_2 - N_1/t_1}{P},$$

где коэффициентом P учитывается объем измеряемого объекта и чувствительность радиометра к излучению смеси радионуклидов в пробе.

Отметим, что описанная методика определения активности не является единственной. Модификаций приборов и методов довольно много.

Радиометрия внутреннего облучения наиболее эффективна для γ-излучающих радионуклидов. Их идентификацию, накопление, выведение и определение индивидуальной дозы внутреннего облучения осуществляют с помощью ***γ-спектрометров излучения человека (СИЧ)***.

Наиболее распространенные установки СИЧ многоцелевого назначения состоят из стальной защитной комнаты, набора сцинтилляционных спектрометрических детекторов, регистрирующей аппаратуры, кресла и носилок для помещения исследуемого пациента. Импульсы от детекторов поступают на многоканальный анализатор, с помощью которого выделяются сигналы, соответствующие γ-квантам определенной энергии. По энергии γ-квантов определяют вид радионуклидов, а по количеству импульсов судят об их концентрации в организме.

29.5. Естественный радиационный фон, фоновое облучение человека

Естественным радиационным фоном называют излучение, создаваемое космическими лучами и естественными радиоактивными веществами, содержащимися в окружающей среде и теле человека.

Рассмотрим компоненты фоновых источников ионизирующих излучений и их вклад в формирование эффективной эквивалентной дозы фонового облучения человека.

Важным источником внешнего облучения является попадающее в атмосферу Земли космическое излучение, которое подразделяется на первичное и вторичное. *Первичные космические лучи* состоят, в основном, из протонов и α-частиц высоких энергий (до 10^{14} МэВ), попадающих в земную атмосферу из космического пространства и проникающих до высоты около 20 км над уровнем моря. В результате их взаимодействия с ядрами атомов, входящих в состав земной атмосферы, образуется вторичное космическое излучение, достигающее поверхности Земли и содержащее практически все известные элементарные частицы.

Оценки эффективной эквивалентной дозы, получаемой человеком на уровне моря за счет воздействия космических лучей, дают ее значение около 0,31 мЗв в год. Эти данные (как и приводимые ниже) относятся к среднемировым и опубликованы Научным комитетом ООН по действию атомной радиации (НКДАР).

В земной коре содержится ряд долгоживущих радионуклидов, которые при своем распаде порождают ионизирующие излучения. Продукты их распада также радиоактивны, причем их период полураспада меньше, а активность больше, и вклад в естественный радиоактивный фон еще выше. Кроме того, под действием космических лучей в атмосфере образуются космогенные радионуклиды. В результате более 60 радионуклидов, содержащихся в биосфере Земли, увеличивают фоновую дозу внешнего облучения в среднем до 0,65 мЗв в год. Естественно, что этот фон существенно отличается в различных участках земного шара и зависит от состава горных пород, почвы, наличия вулканических выбросов и т.п. Средняя мощность фоновой экспозиционной дозы на территории Беларуси составляет в норме 10–12 мкР/ч.

Внутреннее облучение человека создается радионуклидами, поступающими в организм с воздухом, водой и пищей. Среди естественных радионуклидов наибольшее значение имеют ^{40}K, ^{14}C, ^{226}Ra, ^{220}Rn, ^{222}Rn. Согласно оценкам НКДАР, эффективная эквивалентная доза внутреннего об-

лучения от естественных источников земного происхождения вдвое выше, чем внешнего, и составляет 1,35 мЗв в год. Таким образом, среднемировая суммарная эквивалентная доза от естественных источников радиации составляет 2 мЗв в год.

Производственная деятельность человека (извлечение из недр земли полезных ископаемых, производство минеральных удобрений, сжигание топлива и др.) увеличивает в окружающей среде количество естественных радионуклидов и источников ионизирующих излучений. Возникает ***техногенно измененный естественный радиационный фон***, под которым понимают дозу облучения, обусловленную указанной хозяйственной деятельностью. В него не включают увеличение дозы облучения, происходящее из-за испытаний ядерного оружия, работы предприятий в ядерной энергетике, а также применение в медицинских целях рентгеновских и радионуклидных методов диагностики и терапии.

Примеры конкретных значений доз, получаемых по некоторым причинам при облучении всего тела, приведены в табл. 29.2.

Среднемировые значения индивидуальной эквивалентной дозы облучения всего тела из-за производства медицинских процедур оцениваются в 0,4 мЗв/год, что составляет около 20 % естественного фона. Естественно, что этот показатель сильно отличается в различных странах.

Таблица 29.2

Источник излучения	Средняя мощность эквивалентной дозы облучения всего тела, мкЗв/ч
Полет на самолете на высоте 8 км	1,35
Цветной телевизор с ЭЛТ на расстоянии: 5 см от экрана 250 см от экрана	 100 $2{,}5 \cdot 10^{-3}$
Район мощной тепловой электростанции радиусом 20 км	$0{,}6 \cdot 10^{3}$

Контрольные вопросы и задания

1. Что такое экспозиционная доза излучения? Каковы системные и внесистемные единицы ее измерения? Выведите связь между этими единицами.
2. Как связаны расстояние до γ-источника, активность γ-излучающих радионуклидов и мощность экспозиционной дозы?
3. Что такое поглощенная доза, ее мощность? Укажите системные и внесистемные единицы измерения, связь между ними.
4. Что такое относительная биологическая эффективность излучения? Что такое эквивалентная доза? в каких единицах измеряется?
5. Как рассчитывается эффективная эквивалентная доза? Что она характеризует?
6. Что такое коэффициент радиационного риска? Какие органы и почему обладают наибольшим значением коэффициента радиационного риска?
7. Что определяет коллективная эффективная эквивалентная доза?
8. Составьте и решите дифференциальное уравнение, описывающее изменение активности радионуклидов в организме при их однократном поступлении.
9. Получите формулу для расчета эквивалентной дозы в органе при однократном поступлении радионуклидов в организм.
10. Какие факторы и как влияют на величину эквивалентной дозы, получаемой со временем от радионуклида, однократно поступившего в организм?
11. Каков принцип работы ионизационного детектора излучений?
12. Что такое энергетические спектры γ-излучений? Каково их значение для определения радионуклидного состава радиационного загрязнения?
13. Почему α-частицы регистрировать сложнее, чем γ-кванты?
14. Чем различаются по своему назначению дозиметры и радиометры?
15. Какую информацию и как получают при обследовании пациента с помощью спектрометров излучений человека (СИЧ)?
16. Из каких основных составляющих складывается естественный радиационный фон? Какова в норме его мощность на территории Беларуси?
17. Какова среднегодовая эквивалентная доза от естественных источников радиации? от медицинских процедур?

ЛИТЕРАТУРА

1. *Бегун П.И.* Биомеханика: учебник для вузов / П.И. Бегун, Ю.А. Шукейло. СПб.: Политехника, 2000.

2. *Бергельсон Л.Д.* Мембраны, молекулы, клетки / Л.Д. Бергельсон М.: Наука, 1982.

3. *Бецкий О.В.* Терагерцовые волны и их применение. Биомедицинские технологии / О.В. Бецкий [и др.]. Биомедицинские технологии и радиоэлектроника. 2005. № 8. С. 40–48.

4. Биофизика / Ю.А. Владимиров [и др.]. М.: Владос, 2003.

5. *Блинов Н.Н.* Основы рентгенодиагностической техники: учеб. пособие / Н.Н. Блинов [и др.]; под ред. Н.Н. Блинова. М.: Медицина, 2004.

6. *Боголюбов В.М.* Общая физиотерапия: учебник / В.М. Боголюбов, Г.Н. Пономаренко. М.: Медицина, 2003.

7. *Владимиров Ю.А.* Биофизика / Ю.А. Владимиров [и др.]. М.: Медицина, 1983.

8. *Глазер Р.* Очерк основ биомеханики / Р. Глазер; пер. с нем. М.: Мир, 1988.

9. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике / В.Е. Гмурман. М.: Высш. шк., 2001.

10. *Гмурман В.Е.* Теория вероятностей и математическая статистика / В.Е. Гмурман. М.: Высш. шк., 1998.

11. *Губанов Н.И.* Медицинская биофизика / Н.И. Губанов, А.А. Утепбергенов. М.: Медицина, 1978.

12. *Давид Р.* Введение в биофизику / Р. Давид; пер. с франц. М.: Мир, 1982.

13. *Добрецов Г.Е.* Флуоресцентные зонды в исследовании клеток, мембран и липопротеинов / Г.Е. Добрецов. М.: Наука, 1989.

14. *Дубовский В.И.* Биомеханика / В.И. Дубовский, В.Н. Федорова. М.: Владос-пресс, 2003.

15. *Зациорский В.М.* Биомеханика двигательного аппарата человека / В.М. Зациорский, А.С. Аруин, В.Н. Селуянов. М., 1981.

16. *Каро К.* Механика кровообращения / К. Каро [и др.]; пер. с англ. М.: Мир, 1981.

17. *Катц Б.* Нерв, мышца и синапс / Б. Катц. М.: Мир, 1968.

18. Лазеры в клинической медицине / под ред. проф. С.Д. Плетнева. М.: Медицина, 1981.

19. *Левтов В.А.* Реология крови / В.А. Левтов, С.А. Регирер, Н.Х. Шадрина. М.: Медицина, 1982.

20. *Ливенцев Н.М.* Курс физики: учебник для вузов / Н.М. Ливенцев. М.: Высш. шк., 1978.

21. *Линдербратен Л.Д.* Медицинская радиология (основы лучевой диагностики и лучевой терапии): учебник / Л.Д. Линдербратен, И.П. Королюк. М.: Медицина, 2000.

22. *Лобоцкая Н.Л.* Высшая математика / Н.Л. Лобоцкая, Ю.В. Морозов, А.А. Дунаев. Минск: Вышэйш. шк., 1987.

23. *Медик В.А.* Статистика в медицине и биологии. В 2 т. / В.А. Медик, М.С. Токмачев, Б.Б. Фишман. М.: Медицина, 2001.

24. *Островский М.А.* Зрительная рецепция — проблемы на стыке наук / М.А. Островский // Наука в СССР. 1981. № 1 С. 71–76.

25. *Педли Т.* Гидродинамика крупных кровеносных сосудов / Т. Педли; пер. с англ. М.: Мир, 1983.

26. Проблемы прочности в биомеханике / под ред. акад. И.Ф. Образцова. М.: Высш. шк., 1988.

27. *Пуриня Б.А.* Биомеханика крупных кровеносных сосудов человека / Б.А. Пуриня, В.А. Касьянов. Рига: Зинатне, 1980.

28. *Ремизов А.Н.* Медицинская и биологическая физика: учебник для вузов / А.Н. Ремизов, А.Г. Максина, А.Я. Потапенко. М.: Дрофа, 2003.

29. *Рубин А.Б.* Биофизика. В 2 т. / А.Б. Рубин. М.: Высш. шк., 1987.

30. *Самойлов В.О.* Медицинская биофизика: учебник для вузов / В.О. Самойлов. СПб.: СпецЛит, 2007.

31. *Титомир Л.И.* Электрический генератор сердца / Л.И. Титомир. М.: Наука, 1980.

32. *Торнуев Ю.В.* Электрический импеданс биологических тканей / Ю.В. Торнуев [и др.]. М.: Медицина, 1990.

33. *Улащик В.С.* Магнитотерапия / В.С. Улащик // Применение магнитных полей в медицине: материалы междунар. науч.-практ. конф. Оренбург. 25–26 октября 2000 г., БелЦНМИ. Минск, 2001.

34. Физика визуализации изображений в медицине. В 2 т. / под ред. С. Уэбба; пер. с англ. М.: Наука, 1981.

35. *Федорова В.Н.* Краткий курс медицинской и биологической физики / В.Н. Федорова. М.: Изд. инст. мед.-соц. реабилитации, 1996.

36. Физика и биофизика: учебник для вузов / В.Ф. Антонов [и др.]; под ред. проф. В.Ф. Антонова. М.: ГЭОТАР-Медиа, 2009.

37. Физиология человека. В 4-х т. / под ред. Р. Шмидта и Г. Тевса. М.: Мир, 1985.

38. *Цедерс Э.Э.* Механические свойства кровеносных сосудов человека в зависимости от их локализации / Э.Э. Цедерс, Б.А. Пуриня // Механика полимеров. Рига, 1975.

39. *Черницкий Е.А.* Спектральный люминесцентный анализ в медицине / Е.А. Черницкий, Е.И. Слобожанина. Минск: Наука и техника, 1989.

40. *Юнкеров В.И.* Математико-статистическая обработка данных медицинских исследований / В.И. Юнкеров, С.Г. Григорьев. СПб: Элби, 2002.

41. *Ясногородский В.Г.* Электротерапия / В.Г. Ясногородский. М.: Медицина, 1987.

По вопросам приобретения книг обращайтесь:

Республика Беларусь

ООО «Новое знание»
220050, а/я 79, Минск,
пр. Пушкина, д. 15а
Тел./факс: (10-375-17) 211-50-38
E-mail: nk@wnk.biz
http:// wnk.biz

Российская Федерация

Отдел оптовых продаж «ИНФРА-М»:
127282, Москва, ул. Полярная, д. 31в, стр. 1
Тел. (495) 380-4260; факс (495) 363-9212
E-mail: books@infra-m.ru

Отдел «Книга–почтой»:
Тел. (495) 363-4260 (доб. 232, 246)

Учебное издание

Высшее образование

Лещенко Вячеслав Григорьевич
Ильич Генрих Казимирович

МЕДИЦИНСКАЯ И БИОЛОГИЧЕСКАЯ ФИЗИКА

Учебное пособие

Ведущий редактор С.В. Исаенко
Редактор Е.Г. Сазончик
Художник обложки С.В. Ковалевский
Компьютерная верстка Е.Ю. Гурченок
Корректор К.А. Степанова

Оригинал-макет подготовлен ООО «Новое знание»

Подписано в печать 25.11.2011.
Формат 60×90 $^{1}/_{16}$. Бумага офсетная. Гарнитура Школьная.
Печать офсетная. Усл. печ. л. 35,0. Уч.-изд. л. 26,96.
Тираж 1700 экз. Заказ № 5281.

ТК 164000-10907-251111

Общество с ограниченной ответственностью «Новое знание».
ЛИ № 02330/0552555 от 08.04.2009.
Ул. Шаранговича, 7-213б, Минск, Республика Беларусь.
Почтовый адрес: а/я 79, 220050, Минск, Республика Беларусь.
Телефон/факс: (10-375-17) 211-50-38
E-mail: nk@wnk.biz http://wnk.biz

Научно-издательский центр ИНФРА-М
127282, Москва, ул. Полярная, д. 31в
Тел.: (495) 380-05-40, 380-05-43. Факс: (495) 363-92-12
E-mail: books@infra-m.ru http://www.infra-m.ru

Отпечатано в ОАО ордена «Знак Почета»
«Смоленская областная типография им. В. И. Смирнова».
214000, г. Смоленск, проспект им. Ю. Гагарина, 2.


ISBN 978-5-16-005338-7
Российская Федерация
9 785160 053387
ISBN 978-985-475-456-7
Республика Беларусь
9 789854 754567