

Федеральное агентство по образованию

САНКТ-ПЕТЕРБУРГСКИЙ  
ГОСУДАРСТВЕННЫЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

---

*В.И. Прошин В.Г. Сидоров*

**МЕТРОЛОГИЯ, СТАНДАРТИЗАЦИЯ  
И СЕРТИФИКАЦИЯ:  
МЕТОДЫ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ**

*Рекомендовано Учебно-методическим объединением  
по университетскому политехническому образованию  
в качестве учебного пособия для студентов высших учебных заведений,  
обучающихся по направлению подготовки  
140400 «Техническая физика»*

Санкт-Петербург  
Издательство Политехнического университета  
2007

УДК 389.001:681.3.06

ББК 30.10я73

П 847

**Рецензенты:**

Доктор физико-математических наук, профессор СЗТУ В. И. Соколов

Доктор технических наук, профессор, ведущий научный сотрудник

ФТИ им. А.Ф.Иоффе РАН И. А. Линеичук

*Прошин В. И., Сидоров В. Г.* **Метрология, стандартизация и сертификация. Методы обработки результатов измерений:** учеб. пособие. СПб.: Изд-во Политехн. ун-та, 2007. 139 с.

Пособие соответствует государственному образовательному стандарту «Метрология, стандартизация и сертификация» по направлениям 140400 «Техническая физика», 210100 «Электроника и микроэлектроника».

Изложены общие вопросы обработки и представления результатов измерений, элементы теории погрешностей. Особое внимание уделено графическому анализу результатов измерений. Подробно изложена работа с компьютерной программой Origin, предназначенной для графического представления и анализа экспериментальных результатов.

Рекомендовано в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлению подготовки 140400 «Техническая физика». Пособие может быть использовано студентами, специализирующимися в различных областях экспериментальной физики, а также аспирантами, инженерами и научными работниками.

Табл. 5. Ил. 35. Библиогр.: 8 назв.

Печатается по решению редакционно-издательского совета Санкт-Петербургского государственного политехнического университета.

ISBN 5-7422-1485-5

© Санкт-Петербургский государственный  
политехнический университет, 2007

## ОГЛАВЛЕНИЕ

Условные обозначения.....	5
Введение.....	7
<b>1. Оценка результатов и погрешностей измерений.....</b>	<b>10</b>
1.1. Виды погрешностей.....	10
1.2. Обработка результатов многократных измерений.....	13
1.2.1. Законы распределения погрешностей.....	13
1.2.2. Нормальное и прямоугольное распределения.....	19
1.2.3. Композиция распределений.....	23
1.2.4. Точечные статистические оценки математического ожидания и дисперсии.....	26
1.2.5. Прямые неравноточные измерения.....	30
1.3. Доверительный интервал.....	33
1.4. Исключение промахов из серии измерений.....	37
1.5. Методы обработки косвенных измерений.....	39
1.5.1. Косвенные измерения при линейной зависимости между аргументами.....	40
1.5.2. Косвенные измерения при нелинейной зависимости между аргументами.....	44
1.6. Суммирование составляющих погрешности.....	46
1.6.1. Суммирование случайных составляющих погрешности.....	47
1.6.2. Суммирование систематических погрешностей.....	49
1.6.3. Суммирование систематических и случайных погрешностей.....	51
1.7. Погрешности и выбор методики эксперимента.....	52
1.8. Представление окончательного результата.....	54
Контрольные вопросы.....	55
<b>2. Аппроксимация функциональных зависимостей физических величин..</b>	<b>57</b>
2.1. Простейшие виды аппроксимаций экспериментальных зависимостей.....	58

2.2. Определение параметров аппроксимирующей функции методом наименьших квадратов.....	66
Контрольные вопросы.....	76
<b>3. Использование программы Microcal Origin для графического отображения экспериментальных данных.....</b>	<b>77</b>
3.1. Главное окно Origin.....	77
3.2. Работа с таблицей Origin.....	80
3.3. Построение графиков на основе данных таблицы.....	83
3.4. Форматирование графика.....	89
3.5. Представление полученных результатов.....	92
Контрольные вопросы.....	95
<b>4. Аппроксимация зависимостей в Microcal Origin.....</b>	<b>97</b>
4.1. Подготовительные операции.....	98
4.2. Аппроксимация линейными по параметрам функциями при обращении к меню Analysis.....	101
4.3. Аппроксимация линейными по параметрам функциями при обращении к меню Tools.....	103
4.4. Аппроксимация линейной по параметрам функцией произвольного вида.....	107
4.5. Аппроксимация наиболее часто применяемыми нелинейными по параметрам функциями.....	108
4.6. Проверка статистической однородности двух групп данных.....	113
4.7. Расширенный контроль над процессом нелинейной аппроксимации.....	115
Контрольные вопросы.....	122
Заключение.....	123
Литература.....	126
Приложение. Пример построения графика по экспериментальным данным и анализ полученной зависимости.....	127
Предметный указатель.....	137

## УСЛОВНЫЕ ОБОЗНАЧЕНИЯ

МНК — метод наименьших квадратов (с. 68)

СИ — средство измерений (с. 11)

СКО — среднее квадратичное отклонение (с. 18)

$\tilde{A}, \tilde{M}, \tilde{D}$  — оценки величины  $A$ , математического ожидания и дисперсии, соответственно (с. 27, 28)

$\text{cov}_{XY}$  — корреляционный момент или ковариация случайных величин  $X$  и  $Y$  (с. 41)

$D(X), DX$  — дисперсия случайной величины  $X$  (с. 18)

$F$  — интегральная функция распределения, или функция распределения (с. 14)

$M(X), MX$  — математическое ожидание (среднее значение) случайной величины  $X$  (с. 17)

$N, n, n_j$  — число наблюдений в выборке (с. 31)

$P$  — вероятность (с. 14)

$P_\delta$  — доверительная вероятность (с. 34)

$p$  — дифференциальная функция распределения, или плотность распределения вероятностей; плотность вероятностей (с. 14)

$q$  — уровень значимости (с. 34)

$S_{\bar{x}}$  — среднее квадратичное отклонение среднего арифметического значения серии измерений (с. 29)

$S_{\tilde{\sigma}_X}$  — СКО оценки  $\tilde{\sigma}_X \equiv S_X$  (с. 29)

$t(P, \nu)$  —  $P$ -квантиль распределения Стьюдента (с. 36)

$w$  — вес результата измерений (с. 32)

$X, Y, Z$  — обозначение физических величин

$X_{ист}$  — истинное значение физической величины (с. 10)

$x, x_i, y, y_i, z, z_i$  — результаты отдельных измерений величин  $X, Y, Z$

$\bar{x}, \bar{y}, \bar{z}$  — средние арифметические значения величин  $X, Y, Z$  (с. 28)  
 $\bar{\bar{x}}$  — взвешенное среднее (с. 31)  
 $z_p$  — квантиль; квантильный множитель (с. 34, 35)  
 $\alpha_r$  — начальный момент  $r$ -го порядка (с. 17)  
 $\Delta x, \Delta y, \Delta z$  — абсолютная погрешность измерений (с. 10)  
 $\delta_x, \delta_y, \delta_z$  — относительная погрешность измерений (с. 10)  
 $\varepsilon$  — эксцесс распределения (с. 30)  
 $\mu_r$  — центральный момент  $r$ -го порядка (с. 17)  
 $\nu$  — число степеней свободы (с. 35)  
 $\rho_{XY}$  — коэффициент корреляции величин  $X$  и  $Y$  (с. 41)  
 $\sigma_X \equiv +\sqrt{D(X)}$  — среднее квадратичное отклонение (СКО) случайной величины  $X$  (стандартное отклонение) (с. 18)  
 $\tilde{\sigma}_X \equiv S_X$  — оценка среднего квадратичного отклонения величины  $X$  (с. 29)  
 $\Phi$  — функция Лапласа (с. 22, 34)

## ВВЕДЕНИЕ

Дисциплина «Метрология, стандартизация и сертификация» охватывает широкий круг проблем, в основе решения которых лежат измерения. Без надежной измерительной информации невозможны ни научные открытия, ни обеспечение высокого качества выпускаемой продукции, ни высокая эффективность и конкурентоспособность экономики страны в целом. В решении всех этих задач большая роль принадлежит метрологии — науке об измерениях, методах и средствах обеспечения их единства и требуемой точности. В настоящем пособии рассмотрен раздел метрологии, посвященный методам оценивания погрешностей измерений и аппроксимации экспериментальных зависимостей. Обсуждается так же использование для графического отображения зависимостей и их аппроксимации одной из наиболее эффективных программ такого класса — *Microcal Origin*.

Результат любого измерения имеет малую ценность, если не указано, с какой погрешностью выполнено это измерение.

Несмотря на то что правила численной оценки величины погрешностей измерений достаточно детально разработаны, они порой не известны даже опытным экспериментаторам, не говоря уже о рядовых инженерах-практиках. Ошибки в оценке погрешностей измерений зачастую приводят к прямому экономическому ущербу. То есть вопрос о необходимости изучения и практического применения теории погрешностей носит не только академический характер. Анализ погрешностей является существенной частью любого научного эксперимента, поэтому изучение теории погрешностей должно занимать важное место в курсах обучения экспериментальным наукам. Однако преподаванию этого предмета в вузах, к сожалению, не всегда уделяется должное внимание. Конечно, теория погрешностей, основанная на теории вероятностей, сложна и для ее освоения требуются определенные затраты труда. Тем не менее знание основ этой теории и

практическое освоение математических методов анализа результатов измерений обязательны для инженеров всех специальностей.

Уже на самых первых этапах обучения при выполнении лабораторных работ необходимо показать студенту, что погрешность любых измерений есть объективно существующая реальность и что инженер и тем более физик-экспериментатор должны владеть основами математической и графической обработки результатов измерений, уметь оценивать величину погрешности измерений и знать способы ее уменьшения, если это необходимо.

Экспериментатор должен ясно представлять, насколько надежна каждая цифра в полученном им результате. В учебной литературе, посвященной вопросам обработки результатов измерений, на наш взгляд, недостаточное внимание уделяется методам графической обработки экспериментальных данных, и практически отсутствует учебная литература по использованию компьютерных программ для этой цели. Поэтому студенты нуждаются в руководстве, в котором доступно объяснялись бы основные идеи и понятия, приводились формулы для обработки результатов измерений, оценок погрешностей и пр.

В данном пособии рассмотрены основные приемы математической обработки результатов измерений и методы оценки погрешностей измерений, представлены способы графического анализа результатов измерений, позволяющие упростить обработку результатов и сделать эти результаты более наглядными. На конкретных примерах продемонстрированы преимущества графической обработки результатов измерений: простота, наглядность, возможность быстрой обработки почти любой закономерности. Представление и анализ экспериментальных результатов с помощью компьютерной программы для построения и обработки научных графиков Microsoft Origin логично дополняют ручную графическую обработку.

В первой главе пособия рассмотрены общие вопросы математической обработки и представления результатов измерений. Кратко изложены элементы теории погрешностей, дана классификация погрешностей, статистические оценки параметров распределений, способы суммирования распределений и погрешностей, методы оценки погрешностей при прямых и косвенных измерениях.



Во второй главе пособия подробно рассмотрен метод графического анализа экспериментальных результатов, даны графические и аналитические способы линеаризации экспериментальных зависимостей и способы определения параметров этих зависимостей, детально рассмотрен один из наиболее широко используемых методов аппроксимации экспериментальных зависимостей — метод наименьших квадратов.

В третьей и четвертой главах представлены возможности одной из лучших компьютерных программ Microsoft Origin, используемой для графического представления и аналитической аппроксимации экспериментальных результатов.

Третья глава посвящена начальному этапу освоения программы Origin. Продемонстрировано назначение различных элементов главного и других окон программы и подробно рассмотрен порядок работы с программой: заполнение таблиц экспериментальными данными, выбор имен столбцов и функциональных связей между столбцами, построение графиков, управление свойствами графической страницы, форматирование графиков и представление результатов для печати и презентации.

Четвертая глава посвящена многочисленным вариантам аппроксимации экспериментальных данных различными линейными и нелинейными по параметрам функциями с помощью программы Origin. Приведен порядок обращения к программе с названными целями, даны примеры аппроксимации экспериментальных данных с помощью наиболее часто встречающихся в экспериментальной практике функций, показаны возможности проверки статистической однородности различных групп экспериментальных данных и контроля процесса аппроксимации.

Рассмотренный в пособии материал дополнен приложением, в котором приведены примеры работы с экспериментальными данными.

Цель учебного пособия — сформировать у читателя общее представление и получить практические навыки по обработке результатов измерений, полученных в физическом эксперименте, причем, как правило, без обращения к дополнительной литературе.

Для углубленного изучения отдельных вопросов в конце пособия приведен список рекомендуемой литературы.

## 1. ОЦЕНКА РЕЗУЛЬТАТОВ И ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЙ

Измерения играют очень важную роль как в физических исследованиях, так и в инженерной практике. При недостаточно грамотной обработке результатов измерений может не состояться научное открытие или не будут обеспечены требования сложного технологического процесса.

В первой главе кратко рассмотрены общие вопросы обработки и представления результатов измерений.

### 1.1. Виды погрешностей

Результат измерения некоторой физической величины отличается от ее истинного значения, то есть содержит погрешность. Величина погрешности измерения определяет степень доверия к результату и должна быть обязательно оценена. Измеренное значение имеет ценность только в том случае, если для него указана погрешность.

Количественно погрешность  $\Delta x$  определяется как разность между измеренным  $x$  и истинным  $X_{\text{ист}}$  значениями физической величины:

$$\Delta x = x - X_{\text{ист}}. \quad (1.1)$$

Погрешность, определяемая в соответствии с (1.1), называется *абсолютной погрешностью*. Используется также представление погрешности в виде *относительной погрешности*  $\delta_x$ , определяемой как отношение абсолютной погрешности к истинному значению измеряемой величины:

$$\delta_x = \frac{x - X_{\text{ист}}}{X_{\text{ист}}}. \quad (1.2)$$

Относительная погрешность выражается безразмерным числом либо в долях измеряемой величины, как в (1.2), либо в процентах. Относительная погрешность дает лучшее представление о качестве измерений, чем абсолютная.

Соотношения (1.1) и (1.2) используются при теоретическом анализе погрешностей. На практике для нахождения погрешностей измерений применить их невозможно, поскольку истинное значение измеряемой величины  $X_{ист}$  всегда неизвестно. При практических измерениях получают лишь оценки значений для физической величины и для ее погрешности. Выполняется это на основе всей информации, полученной при анализе результатов, методики и конкретной реализации метода измерений.

Составляющие погрешности измерений классифицируют по разным признакам. Обязательные составляющие любого измерения — метод измерений, средства измерений и человек — участник процесса измерений. Все эти три составляющие вносят в результат измерения свои погрешности. Соответственно по источнику происхождения *погрешности* делятся на *методические, инструментальные* и *личные (субъективные)*, а по характеру проявления — на *систематические* и *случайные*. При применении современных средств измерений можно значительно снизить вероятность личных погрешностей, обусловленных неправильными действиями оператора, выполняющего измерения.

*Систематической* называют погрешность, которая остается постоянной или закономерно изменяется при повторных измерениях одной и той же величины в одних и тех же условиях. Систематическая погрешность либо завышает, либо занижает значение измеряемой величины. Исходя из причин возникновения систематических погрешностей, различают систематические погрешности, присущие средствам измерений, — *инструментальные погрешности* и погрешности метода измерений — *методические погрешности*.

Систематическая погрешность стандартизованного измерительного прибора не превышает пределов допускаемой для приборов данного типа суммарной погрешности, в состав которой входит как систематическая, так и случайная погрешности. Для уменьшения инструментальных систематических погрешностей при измерениях необходимо использовать приборы, соответствующие по допустимым погрешностям задачам эксперимента и прошедшие своевременную метрологическую поверку, то есть сравнение с образцовыми средствами измерений.

Методические погрешности обусловлены несовершенством метода измерений, и их учет или устранение являются непростой задачей. Выявление методических погрешностей требует детального анализа всех факторов, сопровождающих измерение физической величины. К факторам, позволяющим уменьшить методические погрешности, можно отнести:

- а) правильное понимание физики исследуемого явления, выражающееся, прежде всего, в соответствии используемой модели явления объекту и условиям эксперимента;
- б) соответствие выбранной методики измерений цели измерительной задачи;
- в) соответствие характеристик средств измерений требованиям методики;
- г) высокую экспериментальную культуру исследователя.

К методическим погрешностям также можно отнести погрешности, обусловленные неидеальностью объекта исследования. Они связаны с отклонением реального объекта от математической модели, для которой построена теория изучаемого физического явления, и, соответственно, с тем, насколько применимы формулы теории для анализа явлений, возникающих в реальном объекте.

Систематическую погрешность в принципе можно предсказать и либо скорректировать методику измерений так, чтобы исключить систематическую погрешность, либо исправить результат измерения путем введения поправки на величину установленной систематической погрешности. Однако универсальных методов исключения систематических погрешностей не существует и никогда нельзя быть абсолютно уверенным в полном устранении систематической погрешности. Необнаруженная систематическая погрешность может существенно исказить результат измерения.

*Случайной погрешностью* измерения называют погрешность, величину и знак которой при повторных измерениях одной и той же величины в одних и тех же условиях невозможно предсказать. Случайная погрешность возникает как результат совместного влияния различных нестабильных неконтролируемых воздействий либо на измерительное устройство, либо непосредственно на измеряемую физическую величину. В результате величина погрешности при многократных измерениях изменяется случай-

ным образом. Таким образом, по своей природе случайные погрешности не могут быть исправлены введением поправок. Однако это не означает, что случайные погрешности не подчиняются каким-либо закономерностям. Присущие случайной погрешности закономерности, носят статистический характер и проявляются при анализе большого количества результатов измерений, проведенных в одинаковых условиях.

Если при многократных измерениях результаты не отличаются друг от друга, это свидетельствует о том, что измерения проводятся недостаточно чувствительными для выявления случайных погрешностей средствами измерений, а общая погрешность измерения определяется погрешностью, вносимой средствами измерений.

Деление погрешностей на систематические и случайные очень важно, так как их оценка требует принципиально разного подхода. Для оценки случайных погрешностей используются методы математической статистики. Увеличивая количество повторных измерений, можно значительно уменьшить случайную ошибку результата. Чем больше объем статистики, тем точнее и надежнее результат. Исключить систематическую составляющую погрешности результата увеличением количества измерений невозможно. Общих рецептов исключения систематических погрешностей не существует. Еще раз отметим, что основной путь устранения постоянной систематической погрешности — всесторонний анализ методики измерений, адекватность предлагаемого теоретического описания изучаемому физическому процессу, выявление несоответствия условий эксперимента и условий, для которых построена теория. Систематическая погрешность может быть выявлена при сопоставлении результатов измерения данной физической величины с результатами, полученными другим методом или с использованием более точных средств измерений.

## **1.2. Обработка результатов многократных измерений**

### ***1.2.1. Законы распределения погрешностей***

Случайные погрешности результата измерения оценивают статистическими методами, проводя многократные измерения физической величины в одинаковых условиях. Из теории вероятностей известно, что полно-

стью свойства случайной величины  $X$  описываются либо интегральной функцией распределения  $F(x)$ , либо плотностью распределения вероятностей  $p(x)$ . Здесь и далее при анализе случайных величин будем обозначать прописной буквой случайную величину  $X$  как совокупность всех ее возможных значений, а строчной — частное значение  $x$  этой случайной величины.

*Интегральной функцией распределения* (рис. 1.1), или просто *функцией распределения*  $F(x)$  называют функцию, значение которой равно вероятности того, что случайная величина в  $i$ -м опыте примет значение  $x_i$ , меньшее  $x$ :

$$F(x) = P\{x_i < x\}. \quad (1.3)$$

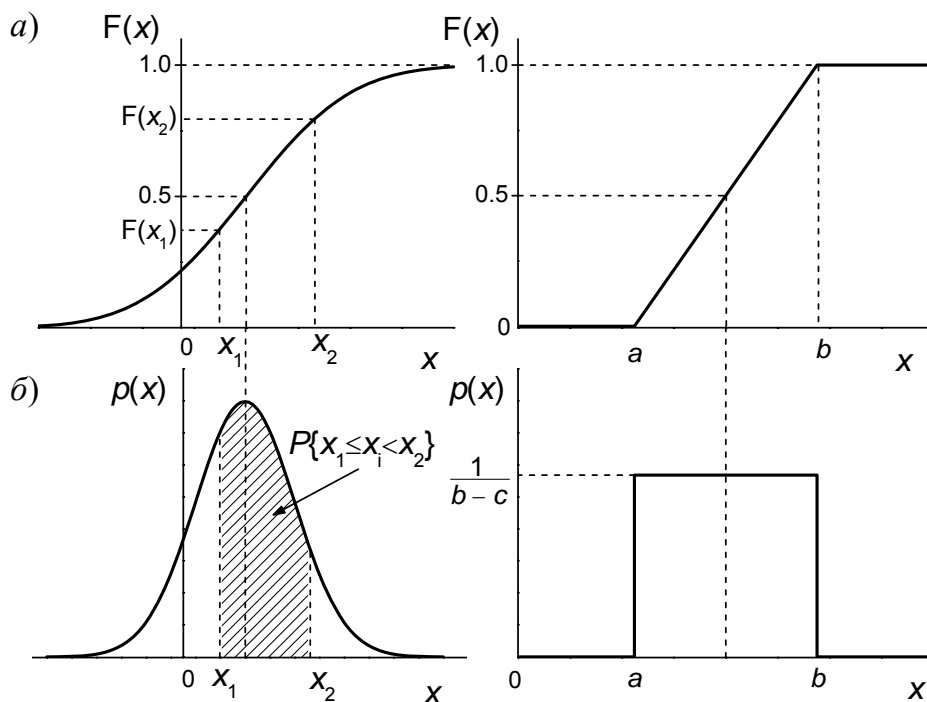


Рис. 1.1. Интегральные (а) и дифференциальные (б) функции нормального (слева) и равномерного, или прямоугольного (справа) распределений для непрерывных случайных величин

Функция распределения — неубывающая функция, изменяющаяся от нуля до единицы:  $F(-\infty) = 0$ , а  $F(+\infty) = 1$ . Вероятность нахождения случайной величины  $X$  в диапазоне  $x_1-x_2$  определяется следующим образом:

$$P\{x_1 \leq x_i < x_2\} = F(x_2) - F(x_1). \quad (1.4)$$

Наряду с интегральной функцией распределения  $F(x)$  для непрерывной случайной величины широко применяется *дифференциальная функция распределения*  $p(x)$ , чаще называемая *плотностью вероятностей*:

$$p(x) = \frac{dF(x)}{dx}, \quad F(x) = \int_{-\infty}^x p(x)dx. \quad (1.5)$$

Вероятность попадания случайной величины в заданный интервал  $[x_1, x_2]$  равна площади под кривой  $p(x)$  на указанном интервале:

$$P\{x_1 \leq x_i < x_2\} = \int_{x_1}^{x_2} p(x)dx. \quad (1.6)$$

Плотность вероятности всегда неотрицательна и подчиняется условию нормировки

$$\int_{-\infty}^{+\infty} p(x)dx = 1. \quad (1.7)$$

Плотность вероятности — величина размерная. Размерность  $p(x)$  обратна размерности случайной величины. Графики плотности вероятностей  $p(x)$  более наглядно показывают разницу между различными распределениями, чем графики соответствующих функций распределения  $F(x)$  (см. рис. 1.1).

Дискретные случайные величины принимают либо конечное, либо бесконечное, но счетное число значений. То есть можно пронумеровать все возможные значения  $x_i$ . Для задания распределения вероятностей случайной дискретной величины необходимо указать вероятности каждого возможного значения этой случайной величины  $P_i = P(x_i)$ . Для дискретной случайной величины

$$F(x) = \sum_{x_i < x} P(x_i); \quad \sum_i P_i = 1. \quad (1.8)$$

Таким образом,  $F(x)$  для дискретной случайной величины является ступенчатой функцией со скачками высотой  $P_i$  в точках  $x_i$  (рис. 1.2).

При практической работе с экспериментальными данными, содержащими случайную погрешность, пользуются обобщенными характеристиками закона распределения, так называемыми *моментами случайных величин*. Моменты представляют собой некоторые средние значения случайных величин.

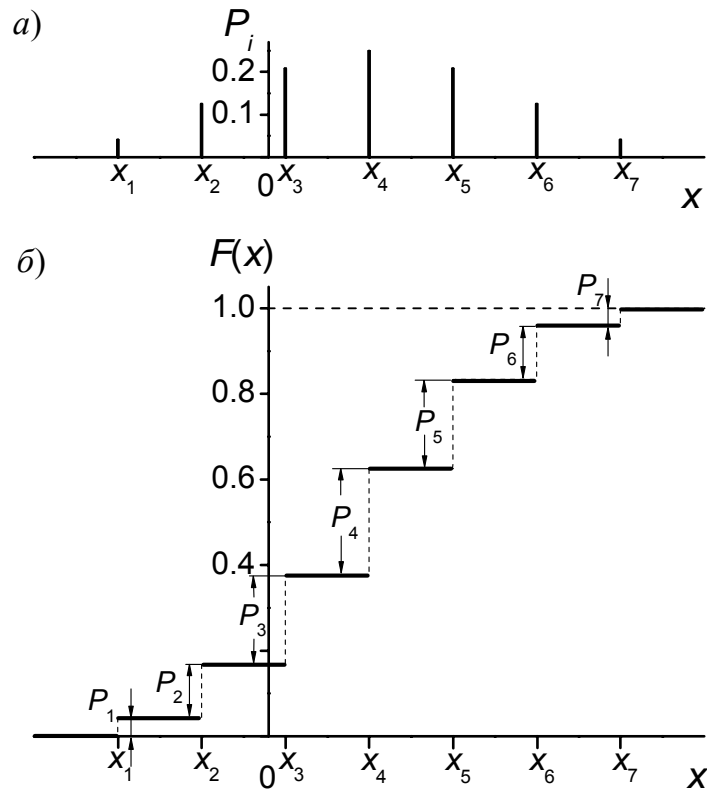


Рис. 1.2. Пример распределения вероятностей (а) и интегральной функции распределения (б) дискретной случайной величины

Если усредняются величины, отсчитываемые от начала координат, то моменты называют *начальными*, а если от центра распределения, то *центральными*. Поскольку моменты являются просто числами, обычно пользоваться ими удобнее, чем функциональными зависимостями — функциями распределения.

Начальный и центральный моменты  $r$ -го порядка  $\alpha_r$  и  $\mu_r$  определяются следующим образом:

а) для дискретной случайной величины  $X$

$$\alpha_r = \sum_{i=1}^N x_i^r P(x_i), \quad (1.9)$$

$$\mu_r = \sum_{i=1}^N (x_i - \alpha_1)^r P(x_i); \quad (1.10)$$

б) для непрерывной случайной величины  $X$

$$\alpha_r = \int_{-\infty}^{+\infty} x^r p(x) dx, \quad (1.11)$$



$$\mu_r = \int_{-\infty}^{+\infty} (x - \alpha_1)^r p(x) dx. \quad (1.12)$$

Наиболее часто для характеристики распределений используют начальный момент 1-го порядка  $\alpha_1$  и центральный момент 2-го порядка  $\mu_2$ .

Первый начальный момент  $\alpha_1$  называют *математическим ожиданием* или *средним значением случайной величины*  $X$  и обозначают  $M(X)$  или  $MX$  (здесь  $X$  — символ случайной величины, для которой вычисляется математическое ожидание, но не аргумент). Для дискретной случайной величины  $X$

$$\alpha_1 \equiv M(X) = \sum_{i=1}^N x_i P(x_i), \quad (1.13)$$

где  $x_1, x_2, \dots, x_i, \dots, x_N$  — возможные значения случайной величины,  $P(x_i)$  — вероятность того, что случайная величина принимает значение  $x_i$ ,  $N$  — полное число возможных значений случайной величины.

Для непрерывной случайной величины  $X$  математическое ожидание имеет вид

$$\alpha_1 \equiv M(X) = \int_{-\infty}^{+\infty} xp(x) dx. \quad (1.14)$$

Математическое ожидание определяет положение центра распределения случайной величины и обладает следующими свойствами. Если  $c = \text{const}$ ,

$$M(c) = c, \quad (1.15)$$

$$M(cX) = cM(X). \quad (1.16)$$

Математическое ожидание суммы нескольких случайных величин равно сумме их математических ожиданий:

$$M(X + Y + Z + \dots + V) = M(X) + M(Y) + M(Z) + \dots + M(V), \quad (1.17)$$

$$M(c + X) = c + M(X). \quad (1.18)$$

Отсюда следует, если измеренное значение физической величины  $X$  содержит случайную погрешность  $\Delta$  и  $M(\Delta) = 0$ , а действительное значение этой величины равно  $a$ , то

$$M(X) = M(a + \Delta) = M(a) = a. \quad (1.19)$$

Для произведения случайных величин формула

$$M(X \cdot Y \cdot Z \cdots V) = M(X)M(Y)M(Z) \cdots M(V) \quad (1.20)$$

справедлива только при условии, если отсутствует корреляция между этими величинами, а также при выполнении более жесткого условия — взаимной независимости случайных величин. Независимые случайные величины некоррелированы, но обратное утверждение не всегда справедливо.

*Второй центральный момент*  $\mu_2$  называется дисперсией случайной величины  $X$  и обозначается  $D(X)$  или  $DX$ . Он характеризует разброс значений случайной величины относительно центра распределения. Для дискретной случайной величины  $X$

$$\mu_2 \equiv D(X) = \sum_{i=1}^N [x_i - M(X)]^2 P(x_i) = M\left([X - M(X)]^2\right). \quad (1.21)$$

Для непрерывной случайной величины  $X$

$$\mu_2 \equiv D(X) = \int_{-\infty}^{+\infty} [x - M(X)]^2 p(x) dx = M\left([X - M(X)]^2\right). \quad (1.22)$$

Обычно в качестве меры разброса значений случайной величины используют положительное значение квадратного корня из дисперсии

$$\sigma_X = +\sqrt{D(X)}. \quad (1.23)$$

Эту величину называют *средним квадратичным отклонением* (СКО), или *стандартным отклонением*, а при оценке точности измерений — *средней квадратичной погрешностью*. Буквой  $\sigma$  без индекса будем обозначать СКО в общем смысле, а  $\sigma_X$ ,  $\sigma_Y$  и др. примем как обозначения для СКО случайных величин  $X$ ,  $Y$  и др. Обе величины  $\sigma_X$  и  $D(X)$  являются мерой разброса значений случайной величины  $X$  относительно среднего. Использовать СКО  $\sigma_X$  удобнее, чем  $D(X)$ , так как размерность  $\sigma_X$  совпадает с размерностью случайной величины и их значения легко сопоставлять.

Дисперсия обладает следующими свойствами:

$$D(c) = 0, \quad (1.24)$$

$$D(cX) = c^2 D(X) = c^2 \sigma_X^2, \quad (1.25)$$

$$D(X + Y + \dots + V) = D(X) + D(Y) + \dots + D(V) = \sigma_X^2 + \sigma_Y^2 + \dots + \sigma_V^2. \quad (1.26)$$

Частный случай:

$$D(c + X) = D(X) = \sigma_X^2. \quad (1.27)$$

Формула (1.26) для дисперсии суммы случайных величин справедлива только при взаимной независимости суммируемых случайных величин.

Причем в этом случае (1.26) справедлива при любых распределениях случайных величин  $X, Y, \dots, V$ .

Для вычисления дисперсии часто полезна следующая формула, которая может быть записана исходя из (1.21) и (1.22):

$$D(X) = M\left([X - M(X)]^2\right) = M(X^2) - [M(X)]^2. \quad (1.28)$$

Еще раз подчеркнем, что математическое ожидание  $M(X)$  случайной величины  $X$  и ее дисперсия  $D(X)$  не являются функциями  $x$  — значений, которые может принимать величина  $X$ . Это числа, которые характеризуют свойства распределения случайной величины:  $M(X)$  определяет положение центра распределения, а  $D(X)$  — величину разброса значений  $x$  относительно этого центра.

### ***1.2.2. Нормальное и прямоугольное распределения***

В практике измерений чаще всего используют нормальное и равномерное распределения.

Нормальное распределение (распределение Гаусса) занимает центральное место в теории обработки результатов измерений. Его особое положение определяется тем, что согласно центральной теореме теории вероятностей случайная величина будет иметь нормальное распределение, если является суммой бесконечно большого числа взаимно независимых бесконечно малых случайных величин с любыми распределениями. Случайную погрешность измерений можно рассматривать как результат аддитивного наложения большого числа не зависящих друг от друга случайных возмущений. На практике распределение случайных погрешностей результатов измерений будет близко к нормальному уже при числе независимых случайных возмущений (частных погрешностей) большем 4–5, если среди частных погрешностей отсутствуют отдельные доминирующие источники погрешностей. При грамотно поставленном эксперименте частные погрешности обычно соизмеримы между собой. Именно поэтому случайные погрешности измерений, как правило, можно считать имеющими нормальное распределение.

Случайная величина называется распределенной нормально, если ее функция распределения и плотность вероятностей выглядят, как представлено на рис. 1.1, и описываются функциями

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(y-a)^2}{2\sigma^2}} dy, \quad (1.29)$$

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}. \quad (1.30)$$

Из выражений (1.29) и (1.30) следует, что нормальное распределение определяется двумя параметрами —  $a$  и  $\sigma$ . Рассмотрим их более подробно.

Используя формулы (1.14), (1.22) и (1.30), вычислим математическое ожидание — среднее значение, и дисперсию — разброс относительно среднего, для нормального распределения случайной величины  $X$ :

$$M(X) = \int_{-\infty}^{+\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = a, \quad (1.31)$$

$$D(X) = \int_{-\infty}^{+\infty} (x-a)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \sigma^2 = \sigma_X^2. \quad (1.32)$$

Отсюда следует, что параметры  $a$  и  $\sigma$  в нормальном распределении плотности вероятностей величины  $X$  имеют значения математического ожидания и среднего квадратичного отклонения.

Рассмотрим основные свойства нормального распределения, которые демонстрирует рис. 1.3. Кривые плотности вероятности имеют колоколообразную форму, симметричную относительно максимума при  $x = a$ , с «хвостами», простирающимися в обе стороны до бесконечности. Параметр  $\sigma$  равен расстоянию от оси симметрии кривой до точки перегиба. При изменении  $a$  кривая смещается вдоль оси  $x$ , не изменяя свою форму. При увеличении  $\sigma$  (среднего квадратичного отклонения) кривая распределения становится более плоской. Чем больше  $\sigma$ , тем значительнее разброс результатов измерений относительно математического ожидания величины.

Если перенесем начало координат в центр распределения  $a$  и отложим по оси абсцисс случайную погрешность  $\Delta x = x - a$ , получим кривую нормального распределения погрешностей:

$$p(\Delta x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\Delta x^2}{2\sigma^2}}. \quad (1.33)$$

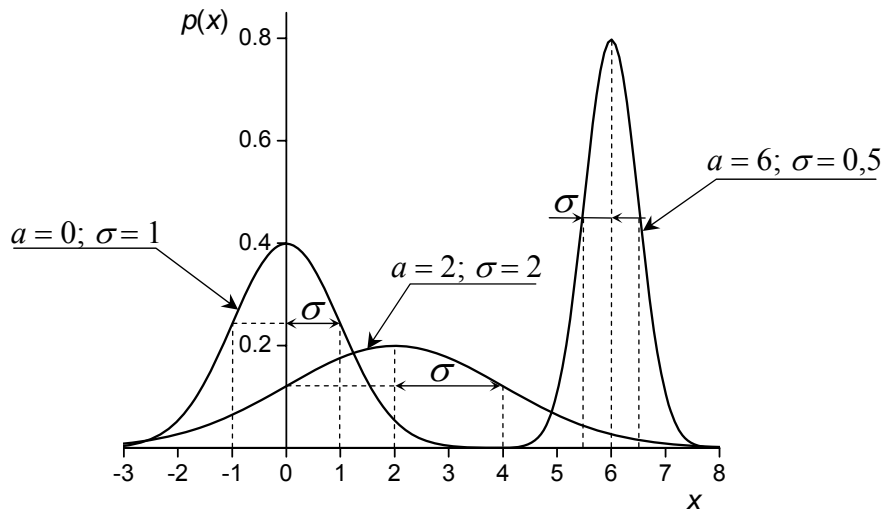


Рис. 1.3. Кривые плотности вероятности для нормального распределения с различными значениями математического ожидания  $a$  и среднего квадратичного отклонения  $\sigma$

Можно выделить следующие особенности нормального распределения погрешностей.

1. Кривая нормального распределения случайных погрешностей симметрична относительно оси ординат ( $a = 0$ ). Это означает, что погрешности, одинаковые по величине, но противоположные по знаку, при большом числе измерений встречаются одинаково часто, если они распределены по нормальному закону.
2. Из характера кривой следует, что при нормальном законе распределения малые погрешности встречаются чаще, чем большие.
3. Качество измерений тем выше, чем меньше средняя квадратичная погрешность  $\sigma$ , поскольку разброс результатов уменьшается.

Выделим еще одно важное свойство нормального распределения. Часть площади под кривой плотности вероятностей, равная 99.73 %, лежит в пределах трех средних квадратичных отклонений (по  $3\sigma$  в каждую сторону от центра распределения), в пределах  $\pm 2\sigma$  находится 95.45 % площади,  $\pm\sigma$  — 68.27%.

При расчетах, особенно с использованием таблиц функций распределения, удобнее пользоваться нормированными случайными величинами. Нормирование и центрирование случайной величины  $X$  выполняют следующим образом:

$$z = \frac{x - M(X)}{\sqrt{D(X)}}, \quad (1.34)$$

где  $z$  — нормированная (стандартизованная) случайная величина, соответствующая  $x$ .

Проводя нормирование таким способом, начало координат переносят в центр распределения, а значения случайной величины откладывают по оси абсцисс в долях СКО, то есть единицей измерения по этой оси выступает  $\sigma$ .

Для нормального распределения

$$z = \frac{x - a}{\sigma}. \quad (1.35)$$

Соответственно получим:

для нормированной и центрированной функции нормального распределения

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{y^2}{2}} dy, \quad (1.36)$$

для плотности нормированного и центрированного нормального распределения

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad (1.37)$$

для математического ожидания

$$M(Z) = \int_{-\infty}^{+\infty} z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 0, \quad (1.38)$$

для дисперсии

$$D(Z) = \int_{-\infty}^{+\infty} [z - M(Z)]^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1. \quad (1.39)$$

Нормированное нормальное распределение также называют *стандартным нормальным распределением*. Значения нормированной функции

распределения можно найти, воспользовавшись таблицами для функции Лапласа

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{y^2}{2}} dy, \quad (1.40)$$

которая связана с функцией  $F(z)$  соотношением

$$F(z) = 0,5 + \Phi(z). \quad (1.41)$$

Практически в любой книге, посвященной теории вероятностей или математической статистике, приведены таблицы значений функции Лапласа. Для функции Лапласа справедливы следующие равенства:  $\Phi(-\infty) = -0,5$ ;  $\Phi(0) = 0$ ;  $\Phi(+\infty) = 0,5$ ;  $\Phi(-z) = -\Phi(z)$ .

Другое часто используемое при анализе погрешностей измерений распределение — прямоугольное, или равномерное.

Случайная величина  $X$  называется равномерно распределенной на участке  $[a, b]$ , если ее плотность вероятностей в интервале  $[a, b]$  постоянна,

а вне интервала равна нулю (см. рис. 1.1). Поскольку  $\int_{-\infty}^{+\infty} p(x) dx = 1$ , то

$$p(x) = \frac{1}{b-a}, \quad (1.42)$$

а математическое ожидание и дисперсия таковы:

$$M(X) = \int_a^b x \frac{1}{b-a} dx = \frac{a+b}{2}, \quad (1.43)$$

$$D(X) \equiv \sigma_X^2 = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}. \quad (1.44)$$

В частном случае равномерного распределения, симметричного относительно оси ординат с границами интервала  $[-\theta, +\theta]$ ,

$$M(X) = 0, \quad D(X) = \frac{\theta^2}{3}, \quad \sigma_X = \frac{\theta}{\sqrt{3}}. \quad (1.45)$$

### 1.2.3. Композиция распределений

Как правило, погрешность результата измерения складывается из ряда составляющих, обусловленных различными факторами. Важно понимать, что при суммировании случайных погрешностей закон распределения ре-

зультатирующей погрешности может существенно изменить свою форму. Поэтому трансформацию законов распределения необходимо учитывать при оценке погрешности результата измерения. Именно знание закона распределения погрешности позволяет рассчитать ее величину.

Оценивание суммарной погрешности связано с построением композиции распределений. Композицией называют закон распределения для суммы независимых случайных величин. Если случайные величины  $X$  и  $Y$  имеют распределения  $p_1(x)$  и  $p_2(y)$ , плотность вероятностей распределения суммы этих величин есть свертка плотностей слагаемых:

$$p(z) = \int_{-\infty}^{+\infty} p_1(\xi)p_2(z - \xi)d\xi = \int_{-\infty}^{+\infty} p_1(z - \xi)p_2(\xi)d\xi. \quad (1.46)$$

Приведем примеры изменения формы законов распределения в результате их композиции (рис. 1.4).

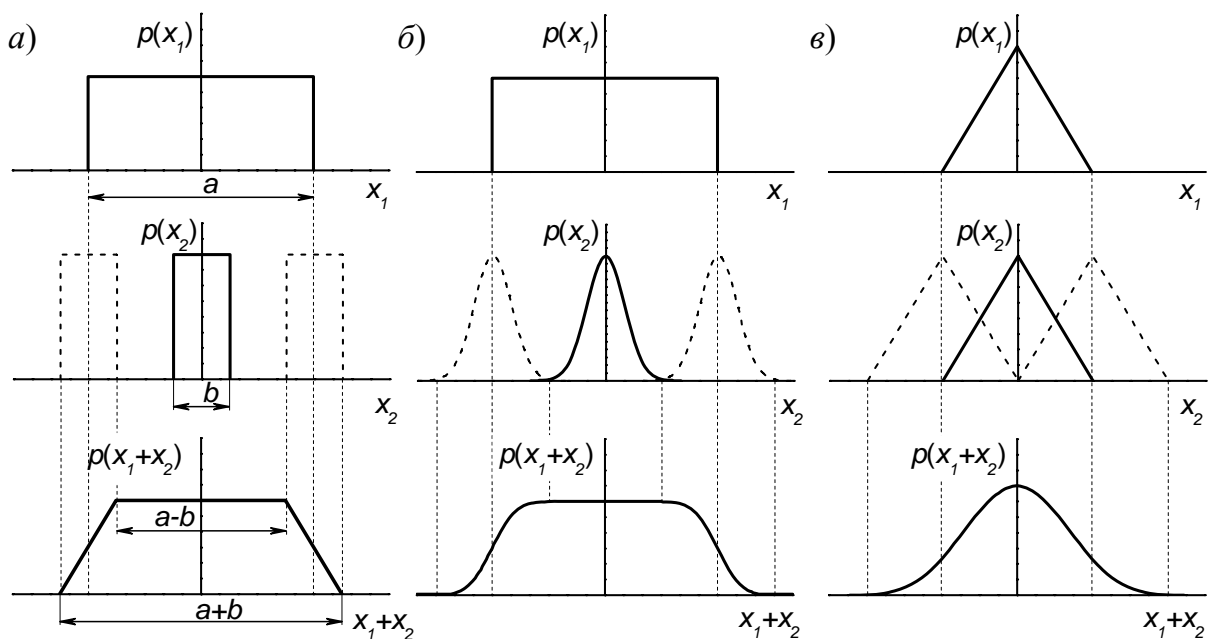


Рис. 1.4. Изменение формы законов распределения при образовании композиций

При суммировании двух равномерно распределенных погрешностей результирующая погрешность имеет трапецеидальное распределение (см. рис. 1.4, а). Эту трансформацию прямоугольного распределения можно представить как размывание резко ограниченных концов более широкого



распределения на ширину более узкого, что и продемонстрировано на рисунке. Отсюда следует, что композиция двух прямоугольных распределений одинаковой ширины есть треугольное распределение. Подобным образом формируется композиция прямоугольного и нормального распределений (см. рис. 1.4, б). На рис. 1.4, в представлена композиция двух треугольных распределений одинаковой ширины, которая, в свою очередь, эквивалентна композиции четырех одинаковых прямоугольных распределений. При этом результирующая кривая близка к кривой нормального распределения и в масштабе рис. 1.4, в сливается с ней.

Следует отметить, что графики в среднем ряду на рис. 1.4 построены без соблюдения относительного масштаба кривых по оси ординат, который определяется требованием равенства единице площадей под каждой кривой. Действительно, полные вероятности для распределений должны быть равны единице в соответствии с нормировкой (1.7).

Единственный закон распределения, который сохраняет свою форму для композиции случайных величин, — это нормальный закон распределения. Если две независимые случайные величины  $X$  и  $Y$  распределены по нормальному закону каждая с координатами центров распределения  $a$  и  $b$  и дисперсиями  $\sigma_X^2$  и  $\sigma_Y^2$ , то их сумма  $Z = X + Y$  будет также распределена по нормальному закону с плотностью вероятностей

$$p(z) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_X^2 + \sigma_Y^2}} \exp\left(-\frac{[z - (a + b)]^2}{2(\sigma_X^2 + \sigma_Y^2)}\right). \quad (1.47)$$

Формулу (1.47) можно обобщить и на композицию любого конечного числа независимых слагаемых, каждое из которых распределено по нормальному закону.

При нелинейной зависимости результата измерения от прямо измеряемых физических величин закон распределения погрешностей результата формируется в результате более сложной трансформации законов распределения погрешностей прямо измеряемых величин, что, несомненно, необходимо учитывать при оценке суммарной погрешности результата измерений.

#### **1.2.4. Точечные статистические оценки математического ожидания и дисперсии**

Рассмотренные нами функции распределения описывают поведение непрерывных случайных величин, тогда как на практике результаты измерений и их случайные погрешности являются, как правило, дискретными величинами. При этом всегда в нашем распоряжении лишь некоторое число значений — результаты независимых повторных измерений физической величины одним и тем же методом при неизменных условиях измерений. Результаты ограниченного ряда наблюдений  $x_1, x_2, \dots, x_n$  представляют собой *выборку* из теоретически бесконечной *генеральной совокупности*. Строго говоря, выборкой называют конечную последовательность из  $n$  независимых одинаково распределенных случайных величин,  $n$  — объем выборки,  $x_1, x_2, \dots, x_n$  — элементы выборки. Генеральной называют совокупность всех мыслимых наблюдений, которые могли бы быть сделаны при данном комплексе условий. По значениям величин, представленным в выборке, необходимо дать заключение о виде распределения случайной величины, принадлежащей данной генеральной совокупности, и оценить параметры предполагаемого закона распределения.

Чтобы воспользоваться функцией распределения вероятностей для анализа погрешностей, необходимо найти оценки, то есть примерные значения параметров этой функции. Оценка, выраженная одним числом, то есть точкой на числовой оси, называется точечной. Необходимо обратить внимание на то, что в отличие от самих параметров распределения их точечные оценки так же, как и результаты отдельных наблюдений, являются случайными величинами, поскольку значения оценок будут разными при использовании различных выборок (разных серий измерений). Поскольку точечные оценки — случайные величины, то они также подчиняются некоторым законам распределения. Поэтому при нахождении оценок встает задача выбора метода статистической обработки, дающего наилучшую оценку параметров искомого распределения. Точечная оценка, полученная по экспериментальным данным, считается в статистическом смысле наилучшей, если она является состоятельной, несмещенной и эффективной.

Оценку будем обозначать той же буквой, что и оцениваемую величину, но с волнистой чертой сверху (тильдой).

Оценка  $\tilde{A}$  называется *состоятельной*, если при увеличении числа наблюдений она *стремится по вероятности* к истинному значению оцениваемой величины  $A$

$$\tilde{A}(x_1, x_2, \dots, x_n) \xrightarrow[n \rightarrow \infty]{P} A. \quad (1.48)$$

Сходимость по вероятности означает, что вероятность отличия  $\tilde{A}$  от  $A$  на любое конечное число  $\varepsilon$ , стремится к нулю при  $n \rightarrow \infty$ . То есть для состоятельной оценки ее дисперсия стремится к нулю при увеличении объема выборки.

*Несмещенной* называется оценка, математическое ожидание которой равно оцениваемой величине, то есть для несмещенной оценки при любом  $n$  должно выполняться условие

$$M[\tilde{A}(x_1, x_2, \dots, x_n)] = A \quad (1.49)$$

*Эффективной* считают ту из нескольких возможных несмещенных оценок, которая имеет наименьшую дисперсию.

На практике не всегда удастся удовлетворить одновременно всем этим требованиям. В таком случае выбору оценки должен предшествовать ее критический анализ со всех перечисленных точек зрения. Например, требование несмещенности не всегда целесообразно. Оценка с небольшим смещением и малой дисперсией может оказаться предпочтительнее несмещенной оценки с большой дисперсией.

Для определения оценок применяются два основных метода: *метод моментов* и *метод максимального правдоподобия*. При использовании метода моментов необязательно знать вид распределения случайной величины. Если предполагаемый закон распределения зависит только от двух параметров (например, нормальный или равномерный законы), то достаточно оценить два первых момента распределения. За оценку истинного значения измеряемой физической величины принимают оценку первого начального момента (математического ожидания), а в качестве характеристики рассеивания результатов измерений — оценку второго центрального момента (дисперсии). При необходимости более подробного описания особенностей распределения случайных погрешностей можно

использовать моменты более высоких порядков. Так, третий центральный момент характеризует асимметрию распределения.

Метод максимального правдоподобия требует предварительного предположения о виде закона распределения при некоторых неизвестных параметрах, входящих в аналитическое выражение этого закона. В качестве оценки неизвестных параметров по выборке принимаются такие их значения, которые максимизируют функцию правдоподобия, представляющую собой вероятность наблюдения именно данной выборки. Простейшим вариантом метода максимального правдоподобия является оценивание по методу наименьших квадратов. Этот метод обсуждается в главе 2.

В методе моментов точечная оценка математического ожидания результата измерения, которая принимается за истинное значение измеряемой величины, — среднее арифметическое значение элементов выборки:

$$\tilde{M}(X) = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (1.50)$$

При любом законе распределения среднее арифметическое значение является состоятельной и несмещенной оценкой для генерального среднего. По критерию наименьших квадратов среднее арифметическое значение является и наиболее эффективной оценкой, то есть

$$\sum_{i=1}^n (x_i - X_{\text{ѐпѐ}})^2 = \min \quad \text{ї ѓѐ} \quad X_{\text{ѐпѐ}} = \bar{x}. \quad (1.51)$$

Состоятельная и несмещенная точечная оценка дисперсии определяется как среднее арифметическое квадратов отклонений  $(x_i - \bar{x})^2$ :

$$\tilde{D}(X) = \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2. \quad (1.52)$$

Появление в знаменателе формулы (1.52)  $(n-1)$  вместо  $n$  в соответствии с (1.21) и (1.50) есть следствие того, что среди усредняемых величин  $(x_i - \bar{x})^2$  независимыми являются лишь  $(n-1)$  значений, так как  $\bar{x}$  определяется из тех же  $n$  значений элементов выборки.

Для характеристики случайного разброса экспериментальных значений величины  $X$  в качестве оценки среднего квадратичного отклонения  $\sigma_X$  используют величину  $S_X$  (оценки СКО различных величин принято обозна-

чать буквой  $S$  вместо  $\tilde{\sigma}$ ). Эту оценку  $S_X$  называют стандартным отклонением выборки и вычисляют по формуле

$$\tilde{\sigma}_X \equiv S_X = +\sqrt{\tilde{D}(X)} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}. \quad (1.53)$$

В отличие от величины  $\tilde{D}(X)$ , которая является несмещенной оценкой дисперсии, величина  $S_X$  — смещенная оценка  $\sigma_X$  из-за нелинейной операции — извлечения корня (1.53). Оценка  $S_X$  будет несмещенной лишь в пределе при  $n \rightarrow \infty$ . Однако, поскольку смещение величины  $S_X$  не превышает 8 % при  $n = 4$  и меньше 3 % при  $n \geq 10$ , что, как правило, значительно меньше случайной погрешности  $S_X$ , то в практике измерений для оценки разброса результатов пользуются формулой (1.53).

Как отмечалось ранее, оценки математического ожидания ( $\bar{x}$ ) и среднего квадратичного отклонения ( $S_X$ ) являются случайными величинами, так как рассчитываются по случайным выборкам из генеральной совокупности. Разброс значений этих оценок, рассчитанных по разным выборкам, в свою очередь, характеризуют средним квадратичным отклонением среднего арифметического (называемым стандартным отклонением среднего и обозначаемым  $S_{\bar{x}}$ ) и средним квадратичным отклонением величины  $S_X$  (обозначаемым  $S_{\tilde{\sigma}_X}$ ).

$$S_{\bar{x}} = \frac{S_X}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2}. \quad (1.54)$$

Обратите внимание на важный результат, выражаемый формулой (1.54). Погрешность оценки математического ожидания  $\bar{x}$ , определяемая  $S_{\bar{x}}$ , с увеличением числа измерений  $n$  уменьшается как  $\sqrt{n}$ . По мере увеличения числа измерений случайные отклонения значений в сторону завышения результата измерений относительно истинного и в сторону занижения все лучше компенсируют друг друга, и оценка среднего значения  $\bar{x}$  приближается к истинному значению физической величины  $X_{\text{ист}}$ .

Среднее квадратичное отклонение для  $S_X$  находят по формуле

$$S_{\tilde{\sigma}_X} = \frac{S_X \sqrt{\varepsilon - 1}}{2\sqrt{n}} = S_{\bar{x}} \frac{\sqrt{\varepsilon - 1}}{2}, \quad (1.55)$$

где  $\varepsilon = \mu_4 / \mu_2^2$  — эксцесс распределения, характеризующий островершинность распределения,  $\mu_2, \mu_4$  — центральные моменты 2-го и 4-го порядков (1.12). Эксцесс характеризует долю площади под кривой распределения, приходящуюся на ее «хвосты». Чем больше «размыто» распределение, тем больше значение  $\varepsilon$ . Из (1.55) следует, что относительная погрешность определения СКО  $S_X$  может быть оценена так:

$$\frac{S_{\tilde{\sigma}_X}}{S_X} = \frac{1}{2} \sqrt{\frac{\varepsilon - 1}{n}}. \quad (1.56)$$

Относительная погрешность зависит только от значения эксцесса и числа наблюдений. Значения  $\varepsilon$  могут лежать в диапазоне от 1 до  $\infty$ . Для распределений, встречающихся в практике измерений,  $\varepsilon \geq 2$ . Так, для прямоугольного распределения  $\varepsilon = 1.8$ , для треугольного  $\varepsilon = 2.4$ , для нормального  $\varepsilon = 3$ . Таким образом, погрешность оценки  $S_X$  при обычном числе измерений весьма значительна. При  $n = 10$   $S_{\tilde{\sigma}_X} / S_X \geq 20\%$  и медленно уменьшается с увеличением  $n$ , поэтому обычно уже вторая значащая цифра в значении  $S_X$  ненадежна.

### 1.2.5. Прямые неравноточные измерения

До сих пор мы рассматривали задачу определения параметров закона распределения случайной величины по ее выборке, полученной при измерениях в одинаковых условиях. Предполагалось, что закон распределения для каждого элемента выборки одинаков и, соответственно, погрешность измерения каждого элемента выборки также одинакова. Такие измерения называют *равноточными*. Для ряда равноточных измерений  $x_1, x_2, \dots, x_n$  все получаемые значения принадлежат одной и той же генеральной совокупности и поэтому характеризуются одинаковой дисперсией и имеют одинаковые средние квадратичные отклонения  $\sigma_1 = \sigma_2 = \dots = \sigma_n = \sigma$ . Как отмечено ранее, при этом наиболее надежной оценкой результата равноточных измерений является среднее арифметическое всех наблюдений (1.50).

Однако практический интерес представляет и случай, когда одна и та же физическая величина измеряется или разными наблюдателями, или с

применением разных измерительных средств и методов измерений, или когда для дальнейшего анализа используются усредненные значения, полученные при равноточных измерениях, но с различным числом измерений в усредняемых сериях (выборках). Встает задача получения наиболее достоверного результата на основе такого неоднородного набора данных.

Пусть все выполненные  $N$  наблюдений распределены в  $r$  группах, каждая группа с индексом  $j$  содержит  $n_j$  наблюдений и  $N = \sum_{j=1}^r n_j$ . Измерения

внутри группы равноточные. Группа может состоять и из одного значения, например может быть представлена средним для данной группы. При этом должна быть известна стандартная погрешность данного среднего. Распределения случайных погрешностей измерений для разных групп, вообще говоря, различны. Это приводит к отличию стандартных погрешностей среднего для групп:  $\sigma_{\bar{x}_1} \neq \sigma_{\bar{x}_2} \neq \dots \neq \sigma_{\bar{x}_r}$ , поскольку случайные погрешности получаемых значений принадлежат разным генеральным совокупностям. Такие измерения называют неравноточными. Совместная обработка совокупности неравноточных данных возможна в том случае, если отличия значений  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_r$  несущественны, то есть находятся в границах случайных погрешностей.

Для получения из неравноточного ряда измерений наиболее вероятного значения величины  $X$  нельзя использовать просто среднее арифметическое всех полученных  $N$  значений. Более точные измерения должны иметь большее влияние на окончательный результат, должны иметь больший вес. Наиболее вероятное значение измеряемой величины при неравноточных измерениях равно так называемому взвешенному среднему, которое принято обозначать  $\bar{\bar{x}}$ . Такое обозначение подчеркивает тот факт, что, как правило, группа представлена ее средним значением  $\bar{x}_j$ .

$$\bar{\bar{x}} = \frac{\sum_{i=1}^N \left( \frac{1}{\sigma_{x_i}^2} x_i \right)}{\sum_{i=1}^N \frac{1}{\sigma_{x_i}^2}} = \frac{\sum_{j=1}^r \left( \frac{1}{\sigma_{\bar{x}_j}^2} \bar{x}_j \right)}{\sum_{j=1}^r \frac{1}{\sigma_{\bar{x}_j}^2}}, \quad (1.57)$$

или

$$\bar{\bar{x}} = \frac{\sum_{i=1}^N w_{x_i} x_i}{\sum_{i=1}^N w_{x_i}} = \frac{\sum_{j=1}^r w_{\bar{x}_j} \bar{x}_j}{\sum_{j=1}^r w_{\bar{x}_j}}, \quad (1.58)$$

где  $\sigma_{x_i}$  — СКО отдельных измерений  $x_i$ .

Если группа представлена ее средним значением, используется СКО среднего  $\sigma_{\bar{x}_j}$ . Величины

$$w_{x_i} = \frac{1}{\sigma_{x_i}^2}, \quad w_{\bar{x}_j} = \frac{1}{\sigma_{\bar{x}_j}^2}, \quad (1.59)$$

называют весом результата  $x_i$  и весом среднего  $\bar{x}_j$  соответственно.

На практике вместо  $\sigma_{x_i}$  и  $\sigma_{\bar{x}_j}$  используют оценки СКО —  $S_{x_i}$  и  $S_{\bar{x}_j}$ .

В (1.57) и (1.58) либо проводится суммирование по всем  $N$  наблюдениям во всех  $r$  группах (при этом  $\sigma_{x_i}$ , а значит, и  $w_{x_i}$  сохраняют свои значения в пределах каждой группы, но изменяются от одной группы наблюдений к другой), либо проводится обработка каждой группы измерений отдельно, а затем рассчитывается взвешенное среднее с использованием  $\bar{x}_j$  и  $\sigma_{\bar{x}_j}$ .

Дисперсия взвешенного среднего рассчитывается по формулам

$$\sigma_{\bar{\bar{x}}}^2 = \frac{1}{\sum_{i=1}^N \frac{1}{\sigma_{x_i}^2}} = \frac{1}{\sum_{i=1}^N w_{x_i}} = \frac{1}{\sum_{j=1}^r \frac{1}{\sigma_{\bar{x}_j}^2}} = \frac{1}{\sum_{j=1}^r w_{\bar{x}_j}}. \quad (1.60)$$

Хорошо иллюстрирует смысл понятия «вес измерения» следующий пример. Выполнено  $r$  серий равноточных измерений величины  $X$  с одинаковыми СКО отдельных измерений  $\sigma_X$  во всех сериях. Причем серии состоят из различного числа измерений  $n_1, n_2, \dots, n_j, \dots, n_r$ . Как известно, СКО для средних арифметических серий измерений  $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_j, \dots, \bar{x}_r$  равны соответственно  $\sigma_X/\sqrt{n_1}, \sigma_X/\sqrt{n_2}, \dots, \sigma_X/\sqrt{n_j}, \dots, \sigma_X/\sqrt{n_r}$ . Согласно формуле (1.59) вес среднего значения  $j$ -й серии пропорционален числу измерений в этой серии  $n_j$



$$w_{\bar{x}_j} = \frac{n_j}{\sigma_X^2}, \quad (1.61)$$

а взвешенное среднее при этом равно среднему арифметическому всех результатов измерений

$$\bar{\bar{x}} = \frac{\sum_{j=1}^r n_j \bar{x}_j}{n_1 + n_2 + \dots + n_j + \dots + n_r} = \frac{1}{N} \sum_{j=1}^r n_j \bar{x}_j = \frac{1}{N} \sum_{i=1}^N x_i. \quad (1.62)$$

Важно помнить, что рассмотренная нами совместная обработка групп измерений, полученных различными методами, допустима только в том случае, если различие в средних значениях для групп незначимо. При этом доверительные интервалы (см. раздел 1.3) полученных средних значений должны перекрываться. Это означает, что измеряется одна и та же физическая величина и измерения не отягощены большими систематическими погрешностями.

### 1.3. Доверительный интервал

Точечные оценки не дают всей желаемой информации об измеряемой физической величине. Необходимо иметь критерий, характеризующий точность сделанной оценки. Состоятельность и несмещенность оценки дают лишь уверенность в том, что эта оценка лучше по сравнению с другими возможными.

В какой-то степени среднее квадратичное отклонение среднего арифметического  $\sigma_{\bar{x}}$  характеризует точность оценки, поскольку служит мерой разброса значений случайной величины  $\bar{x}$ . Однако при разных распределениях случайной погрешности вероятность попадания истинного значения физической величины в интервал  $\bar{x} \pm \sigma_{\bar{x}}$  будет разной. На практике принято указывать интервал, включающий в себя точечную оценку, между границами которого с заданной вероятностью находится истинное значение физической величины. Такой интервал называется доверительным интервалом, а задаваемая вероятность — доверительной вероятностью. Доверительную вероятность  $P_d$ , как правило, принимают равной 0.95. Если требуется более высокая степень надежности результата, выбирается дру-

гое стандартное значение  $P_d = 0.99$ . В соответствии с данным определением доверительной вероятности

$$P\{x_i < X_{\text{ист}} < x_{\hat{a}}\} = P_{\hat{a}} = 1 - q, \quad (1.63)$$

где  $x_n, x_b$  — нижняя и верхняя границы доверительного интервала ( $x_n, x_b$ ) соответственно;  $P_d$  — доверительная вероятность, с которой доверительный интервал покрывает истинное значение  $X_{\text{ист}}$  измеряемой физической величины;  $q$  — уровень значимости.

Подход к нахождению интервальной оценки для значения физической величины при известном законе распределения погрешности основан на понятии квантиля. Квантилем  $x_p$  порядка  $P$  называется такое значение случайной величины  $X$ , для которого интегральная функция распределения равна  $P$ , то есть

$$F(x_p) = \int_{-\infty}^{x_p} p(x) dx = P. \quad (1.64)$$

Вероятность того, что  $X$  лежит в интервале  $(x_n, x_b)$  равна:

$$P_{\hat{a}} = F(x_{\hat{a}}) - F(x_i). \quad (1.65)$$

Для симметричного распределения

$$F(x_{\hat{a}}) = \frac{1 + P_{\hat{a}}}{2} \quad \text{и} \quad F(x_i) = \frac{1 - P_{\hat{a}}}{2}. \quad (1.66)$$

При этом правая часть в (1.65) будет иметь вид

$$F(x_{\hat{a}}) - F(x_i) = \frac{1 + P_{\hat{a}}}{2} - \frac{1 - P_{\hat{a}}}{2} = P_{\hat{a}}. \quad (1.67)$$

Если погрешности результатов отдельных измерений  $x_i$  распределены нормально и стандартное отклонение  $\sigma_x$  для них известно, то и  $\bar{x}$  распределены по нормальному закону с известным СКО  $\sigma_{\bar{x}} = \sigma_x / \sqrt{n}$ . Стандартное отклонение среднего  $\sigma_{\bar{x}}$  можно считать известным и равным его точечной оценке  $S_{\bar{x}}$ , если последняя выполнена по большому числу измерений ( $n \geq 30$ ) либо при проведении данных измерений, либо проведенных ранее при отработке методики измерений. При этом значения  $x_i = \bar{x} - z_p \sigma_{\bar{x}}$  и  $x_{\hat{a}} = \bar{x} + z_p \sigma_{\bar{x}}$  определяются по таблицам значений функции Лапласа  $\Phi(z)$ , а доверительный интервал  $\pm z_p \sigma_{\bar{x}}$  удовлетворяет условию

$$P\{\bar{x} - z_P \sigma_{\bar{x}} < X_{\text{ср}} < \bar{x} + z_P \sigma_{\bar{x}}\} = P_{\text{д}} = 2\Phi(z_P), \quad (1.68)$$

где  $z_P$  — аргумент функции Лапласа  $\Phi(z)$ , отвечающий вероятности  $P_{\text{д}}/2$ .

В данном случае  $z_P$  называется квантильным множителем, поскольку  $z_P$  — нормированная в соответствии с (1.35) величина. Равенство (1.68) справедливо, поскольку интегральная функция распределения  $F(x)$  связана с функцией Лапласа соотношением  $F(x) = 0,5 + \Phi(x)$ . Используемые в практике обработки измерений значения  $z_P$  приведены в таблице 1.1.

**Таблица 1.1**

**Значения квантильных множителей нормального распределения  $z_P$  для заданных значений доверительной вероятности  $P_{\text{д}}$**

$P_{\text{д}}$	0.90	0.95	0.96	0.97	0.98	0.99
$z_P$	1.65	1.96	2.06	2.17	2.33	2.58

Если распределение результатов измерений подчиняется нормальному закону, но число измерений невелико и  $\sigma_{\bar{x}}$  неизвестно, для определения доверительного интервала среднего приходится использовать оценку СКО  $S_{\bar{x}}$ . Поскольку  $S_{\bar{x}}$  — случайная величина, плотность распределения вероятностей среднего арифметического уже не описывается нормальным распределением. Нормированная случайная величина

$$t = \frac{\bar{x} - M(X)}{S_{\bar{x}}} \quad (1.69)$$

подчиняется распределению Стьюдента  $p(t, \nu)$  с числом степеней свободы

$$\nu = n - 1, \quad (1.70)$$

где  $n$  — число усредняемых значений.

Симметричная колоколообразная кривая плотности вероятности для распределения Стьюдента похожа на кривую нормального распределения. Однако при малых  $n$  она намного шире кривой стандартного нормального распределения. При увеличении  $n$  кривая распределения Стьюдента все больше приближается к кривой плотности вероятности нормального распределения и в пределе при  $n \rightarrow \infty$  переходит в нее.

Таким образом, когда число измерений невелико ( $n \leq 30$ ) и погрешности отдельных значений распределены по нормальному закону, довери-

тельный интервал рассчитывается по формуле

$$\Delta_P = t(P, \nu) S_{\bar{x}}, \quad (1.71)$$

где  $t(P, \nu)$  —  $P$ -квантиль распределения Стьюдента с числом степеней свободы  $\nu$ .

При уменьшении числа измерений  $n$ , по которым находится оценка  $S_{\bar{x}}$  для  $\sigma_{\bar{x}}$ , значение  $t(P, \nu)$  при малых  $n$  быстро возрастает. При  $n = 5$   $\Delta_{0,95} = 2.78 S_{\bar{x}}$ , а при  $n = 4$   $\Delta_{0,95} = 3.18 S_{\bar{x}}$ . Для нормального распределения  $\Delta_{0,95} = 1.96 S_{\bar{x}}$ . Однако при  $n \geq 8$  отличие квантилей распределения Стьюдента от квантилей нормального распределения составляет уже менее 20 % (табл. 1.2).

**Таблица 1.2**  
**Значения квантилей распределения Стьюдента  $t(P, \nu)$  для заданных значений доверительной вероятности  $P_d$  и числа измерений  $n$**

$t$	$n$									
	2	3	4	5	7	10	15	20	30	$\infty$
$t_{0,90}$	6.31	2.92	2.35	2.13	1.94	1.83	1.76	1.73	1.70	1.64
$t_{0,95}$	12.7	4.30	3.18	2.78	2.45	2.26	2.15	2.09	2.04	1.96
$t_{0,99}$	63.68	9.93	5.84	4.60	3.71	3.25	2.98	2.86	2.75	2.58

Если функция распределения для  $\bar{x}$  неизвестна, но стандартное отклонение  $\sigma_{\bar{x}}$  задано, то доверительный интервал можно построить, воспользовавшись неравенством Чебышева. При любом законе распределения случайной величины  $X$  вероятность  $P$  попадания частного значения  $x$  этой величины в интервал  $M(X) \pm \gamma_P \sigma_X$  не меньше вероятности, определяемой неравенством Чебышева:

$$P\{M(X) - \gamma_P \sigma_X \leq x \leq M(X) + \gamma_P \sigma_X\} \geq 1 - \frac{1}{\gamma_P^2} = P_a. \quad (1.72)$$

Соответственно, коэффициент  $\gamma_P$  вычисляется по заданной доверительной вероятности  $P_d$

$$\gamma_P = \frac{1}{\sqrt{1 - P_a}}. \quad (1.73)$$

Если распределение случайных погрешностей можно считать симметричным относительно  $M(X)$ , то доверительный интервал сужается:

$$P\{M(X) - \gamma_P \sigma_X \leq x \leq M(X) + \gamma_P \sigma_X\} \geq 1 - \frac{4}{9} \cdot \frac{1}{\gamma_P^2} = P_{\bar{\alpha}}, \quad (1.74)$$

$$\gamma_P = \frac{2}{3\sqrt{1 - P_{\bar{\alpha}}}}. \quad (1.75)$$

На практике интервалы, полученные с помощью неравенства Чебышева, оказываются излишне широкими. Очень часто доверительные интервалы строят на основе распределения Стьюдента без проверки допустимости этого. То, что при этом, как правило, не возникает явных противоречий с наблюдаемыми значениями физической величины, косвенно свидетельствует о близости реальных распределений к нормальному.

#### 1.4. Исключение промахов из серии измерений

Большие ошибки, возникающие при измерениях, называют грубыми погрешностями или промахами. Промахи необходимо выявить и исключить из последующей обработки результатов измерений, поскольку принадлежность всех результатов, входящих в серию измерений, к одной и той же генеральной совокупности — одно из условий правомерности статистической обработки. Если серия из небольшого числа измерений содержит промах, то может быть сильно искажено как среднее значение измеряемой величины, так и границы доверительного интервала.

Если причина существенного отличия результата от остальных в серии ясна, например кратковременное нарушение условий эксперимента, сбой в работе приборов, такой результат правомерно отбросить. Однако исключение просто резко отличающегося от других измерений значения недопустимо без использования объективных статистических критериев обнаружения промахов.

Для выявления промахов задаются вероятностью  $q$  (уровнем значимости) того, что сомнительный результат может принадлежать распределению вероятностей случайной погрешности. Обычно выбирают  $q = 0.01$ . Затем по таблицам квантилей для соответствующего распределения находят

нормированное граничное отклонение от среднего значения  $z_{гр}$  для выбранного уровня значимости. Резко выделяющийся результат  $x_{max}$  считают промахом, если

$$\frac{|x_{max} - \bar{x}|}{S_X} \geq z_{\alpha\delta}, \quad (1.76)$$

где  $S_X$  — оценка СКО серии измерений.

Если серия измерений принадлежит нормальному распределению, то при числе измерений  $n \geq 20$ , когда можно считать  $S_X = \sigma_X$ , для нахождения величины  $z_{гр}$  используют значения квантильных множителей нормального распределения (табл. 1.1), а при  $n < 20$  используют табл. 1.2 квантилей распределения Стьюдента. В первом случае при  $q = 0,01$   $z_{гр} = 2.58$ , а во втором, например при  $n = 10$ , —  $z_{гр} = 3.25$ .

Если в серии измерений присутствует аномально далеко отстоящее от остальных значение, вычисление оценки  $S_X$  для расчета по (1.76) производится после исключения из серии измерений этого значения, похожего на промах. В противном случае оценка  $S_X$  значительно возрастает, увеличивая граничные значения доверительного интервала  $\pm z_{гр} S_X$ . После расчета границ промахов результат наблюдений, который предварительно был отброшен, но впоследствии оказался внутри вычисленных границ, возвращают в выборку. Наблюдения, оказавшиеся за границами найденного интервала, признаются промахами и исключаются из дальнейшей обработки результатов измерений. Процедура расчета  $z_{гр}$  повторяется с использованием только оставшихся значений столько раз, сколько необходимо для исключения всех промахов.

Для результатов измерений, распределенных по нормальному закону, часто используют критерий трех сигм. По этому критерию значение считается промахом, если  $|x_{max} - \bar{x}| \geq 3\sigma_X$ . Данный критерий надежен при  $n \geq 20$ . При этом он соответствует  $q < 0.003$ .

В общем случае границы промахов при выбранном значении  $q$  зависят от числа измерений в серии и от вида распределения случайных погрешностей. Для нормального распределения  $|x_{max} - \bar{x}| \geq 3\sigma_X$  — хороший критерий, а для прямоугольного распределения промахом является уже  $|x_{max} - \bar{x}| > \sigma_X \sqrt{3}$ .

Таким образом, нельзя произвольно отбрасывать более или менее резко отличающиеся от других результаты измерений. В сомнительных случаях делают дополнительные измерения, но не взамен сомнительных, а в дополнение к ним, и проводят описанную выше статистическую обработку для всей дополненной группы измерений. Нужно ориентироваться на то, что если оставлено сильно отклоняющееся значение и оно изменяет  $\bar{x}$  на величину, малую по сравнению с  $S_x$ , то это не искажает заметно оценку искомой величины, но увеличивает погрешность результата.

### 1.5. Методы обработки данных косвенных измерений

В большинстве экспериментов интересующую исследователя физическую величину  $Q$  не измеряют непосредственно. Измеряют некоторые другие величины  $q_1, q_2, \dots, q_m$ , а искомое значение величины  $Q$  вычисляют по известной заранее функциональной зависимости

$$Q = f(q_1, q_2, \dots, q_m), \quad (1.77)$$

где  $Q$  — измеряемая величина, то есть величина, значение которой определяется;  $q_i$  — физические величины, называемые измеряемыми аргументами, значения которых получаются в результате прямых измерений.

Каждая величина  $q_i$  — результат многократных измерений. Задача заключается в том, чтобы, зная средние квадратичные погрешности прямо измеряемых величин  $q_i$ , вычислить среднюю квадратичную погрешность физической величины  $Q$ .

Погрешности косвенных измерений величины  $Q$ , как случайные, так и систематические, зависят от погрешностей измерений аргументов  $q_1, q_2, \dots, q_m$ , а вклад погрешностей отдельных величин  $q_i$  в результирующую погрешность величины  $Q$  зависит от вида функциональной зависимости и соотношения между собой независимых переменных  $q_i$ .

По виду функциональной зависимости  $f$  искомой физической величины косвенные измерения этой величины делятся на две основные группы — линейные и нелинейные.

При линейных косвенных измерениях уравнение измерений имеет вид

$$Q = \sum_{i=1}^m b_i q_i, \quad (1.78)$$

где  $b_i$  — постоянные коэффициенты при аргументах  $q_i$ .

При любых других функциональных зависимостях (1.77) физической величины ее измерение относится к нелинейным косвенным измерениям.

Для линейных косвенных измерений математический аппарат статистической обработки измеренных значений детально разработан. При нелинейной зависимости величины от прямо измеряемых аргументов не всегда возможно найти несмещенные, эффективные и состоятельные оценки измеряемой величины.

### **1.5.1. Косвенные измерения при линейной зависимости между аргументами**

Рассмотрим простейший вариант, когда косвенно измеряемая физическая величина является суммой или разностью величин, определяемых прямыми измерениями. Пусть

$$Z_{\text{эпò}} = X_{\text{эпò}} \pm Y_{\text{эпò}} \quad (1.79)$$

При однократном измерении величин  $X$  и  $Y$  результат может быть записан в соответствии с определением погрешности (1.1):

$$z - \Delta_Z = (x - \Delta_X) \pm (y - \Delta_Y), \quad (1.80)$$

где  $z, x, y$  — измеренные значения физических величин;  $\Delta_Z, \Delta_X, \Delta_Y$  — случайные погрешности величин  $z, x, y$  (могут быть как положительными, так и отрицательными).

Следовательно,

$$\bar{z} = \bar{x} \pm \bar{y}, \quad \Delta_Z = \Delta_X \pm \Delta_Y, \quad (1.81)$$

где  $\bar{z}, \bar{x}, \bar{y}$  — средние значения результатов измерений для соответствующих величин, выступающие как оценки этих величин при многократных измерениях.

В соответствии с (1.21), (1.22), (1.26), определяемая случайными погрешностями, дисперсия  $Z$  равна

$$D(Z) = D(\Delta_Z) = D(\Delta_X \pm \Delta_Y) = M[(\Delta_X \pm \Delta_Y)^2] = M(\Delta_X^2 + \Delta_Y^2 \pm 2\Delta_X\Delta_Y).$$

Таким образом,



$$D(Z) = D(X) + D(Y) \pm 2M(\Delta_X \Delta_Y). \quad (1.82)$$

Как видим, выражение для дисперсии суммы/разности двух случайных величин кроме суммы дисперсий этих величин содержит удвоенное математическое ожидание произведения погрешностей. Математическое ожидание произведения случайных погрешностей называют корреляционным моментом или ковариацией  $\text{cov}_{XY}$  случайных величин  $X$  и  $Y$ :

$$\text{cov}_{XY} = M[(x - M(X))(y - M(Y))] = M(\Delta_X \Delta_Y). \quad (1.83)$$

Заметим, что  $\text{cov}_{XX} = \sigma_X^2$  — дисперсия случайной величины  $X$ . Нормированный корреляционный момент называют коэффициентом корреляции  $\rho$

$$\rho_{XY} = \frac{M(\Delta_X \Delta_Y)}{\sigma_X \sigma_Y}. \quad (1.84)$$

При использовании коэффициента корреляции уравнение (1.82) принимает вид

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 \pm 2\rho_{XY}\sigma_X\sigma_Y. \quad (1.85)$$

Оценки коэффициента корреляции между случайными величинами  $X$  и  $Y$  делают по серии результатов наблюдений исходных величин:

$$\tilde{\rho}_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (1.86)$$

где  $n$  — число пар точек  $(x_i, y_i)$ . Возможные значения коэффициента корреляции лежат в интервале от  $-1$  до  $+1$  и характеризуют уровень линейной зависимости между случайными величинами. При наличии корреляции ( $\rho_{XY} \neq 0$ ) между случайными величинами  $X$  и  $Y$  можно выявить прямую, в окрестности которой группируются точки  $(x_i, y_i)$ . При положительной корреляционной связи  $\rho_{XY} > 0$  случайное увеличение/уменьшение одной величины с большей вероятностью вызывает согласованное изменение другой величины в том же направлении, чем в противоположном. Если  $\rho_{XY} < 0$ , корреляция называется отрицательной. При этом случайное увеличение одной величины с большей вероятностью сопровождается уменьшением другой. Чем ближе  $\rho$  к  $\pm 1$ , тем теснее будут ложиться точки вдоль прямой линии, определяющей взаимозависимость величин  $X$  и  $Y$ . Коэффициент

корреляции  $|\rho_{XY}|=1$  только в том случае, когда между  $x_i$  и  $y_i$  существует линейная функциональная зависимость.

Корреляция между погрешностями случайных величин может возникнуть, когда они измеряются одновременно, поскольку неконтролируемые изменения влияющих факторов могут приводить к возникновению взаимозависимых погрешностей результатов измерений. Если величины измеряются в разное время, то нет оснований ожидать появления корреляции между погрешностями данных измерений.

В общем случае линейной функциональной зависимости физической величины (1.78) результат косвенных измерений рассчитывают по формуле

$$\tilde{Q} = \sum_{j=1}^m b_j \tilde{q}_j, \quad (1.87)$$

где  $\tilde{Q}$  — оценка искомой физической величины;  $\tilde{q}_j$  — оценка результата измерений аргумента  $q_j$ , получаемая, как правило, посредством обработки результатов многократных прямых измерений.

При несмещенности и состоятельности оценок  $\tilde{q}_j$  полученная оценка результата измерений  $\tilde{Q}$  будет также несмещенной и состоятельной. Если результаты  $\tilde{q}_j$  обладают минимальной дисперсией, то есть являются эффективными, то и оценка результата измерений  $\tilde{Q}$  тоже будет эффективной.

Оценку СКО результата косвенных измерений  $Q$  рассчитывают по формуле

$$S(\tilde{Q}) = \sqrt{\sum_{j=1}^m b_j^2 S^2(\tilde{q}_j) + \sum_{l=1}^m \sum_{\substack{k=1 \\ k \neq l}}^m \tilde{\rho}_{kl} b_k b_l S(\tilde{q}_k) S(\tilde{q}_l)}, \quad (1.88)$$

где  $S(\tilde{q}_j)$  — оценка СКО результата измерений  $j$ -го аргумента;  $\tilde{\rho}_{kl}$  — оценка коэффициента корреляции между погрешностями аргументов  $q_k, q_l$ .

Несмещенная оценка  $\tilde{\rho}_{kl}$  рассчитывается по формуле

$$\tilde{\rho}_{kl} = \frac{1}{n(n-1)S(\tilde{q}_k)S(\tilde{q}_l)} \sum_{i=1}^n (q_{k_i} - \tilde{q}_k)(q_{l_i} - \tilde{q}_l), \quad (1.89)$$

где  $q_{k_i}, q_{l_i}$  —  $i$ -е результаты прямых измерений  $k$ -го и  $l$ -го аргументов;  $n$  — число пар прямых измерений аргументов  $q_k, q_l$ .

Учет корреляционных связей на практике довольно сложен. Коэффициент корреляции, рассчитываемый по формуле (1.89), обладает плохой статистической устойчивостью. Особенно большие трудности возникают при выявлении слабых корреляционных зависимостей. Поэтому на практике все погрешности в зависимости от значения полученной оценки  $\rho$  условно делят на сильно ( $0,7 \leq |\rho| \leq 1$ ) и слабо ( $0 \leq |\rho| < 0,7$ ) коррелированные. При сильной корреляции считают, что  $\rho = \pm 1$  в зависимости от знака коэффициента корреляции. При слабой корреляции полагают, что  $\rho = 0$ .

Если погрешности результатов измерений аргументов  $q_j$  распределены по нормальному закону, то и случайные погрешности результата измерений  $Q$  при линейной зависимости между аргументами подчиняются нормальному закону распределения. При этом доверительные границы случайной погрешности результата измерений при небольшом числе в сериях измерений аргументов вычисляются по формуле

$$\Delta_P = t(P_{\alpha}, \nu_{\hat{y}_0}) S(\tilde{Q}), \quad (1.90)$$

где  $t(P_{\alpha}, \nu_{\hat{y}_0})$  — коэффициент Стьюдента, соответствующий доверительной вероятности  $P_{\alpha}$  и эффективному числу степеней свободы  $\nu_{\hat{y}_0} = n_{\min} - 1$ , где  $n_{\min}$  — минимальное из  $n_j$ . Здесь  $n_j$  — число измерений, сделанных для получения оценки значения  $j$ -го аргумента.

При некоторых условиях формулой (1.90) можно пользоваться и в том случае, если распределения погрешностей аргументов отличны от нормального. При этом число аргументов должно быть больше четырех, распределения их погрешностей должны быть *одномодалными* (кривые плотности вероятностей имеют только один максимум) и среди погрешностей не должно быть резко выделяющихся. При таких условиях распределение погрешностей результата  $Q$  также близко к нормальному.

При большом числе измерений (более 25–30), выполненных при измерении каждого аргумента, оценку средней квадратичной погрешности можно считать равной СКО. При этом доверительную границу случайной погрешности результата косвенных измерений рассчитывают по формуле

$$\Delta_P = z_P S(\tilde{Q}), \quad (1.91)$$

где  $z_P$  — квантиль нормального распределения, соответствующий выбранной доверительной вероятности  $P_d$ ,  $S(\tilde{Q})$  — СКО результата косвенных измерений.

### **1.5.2. Косвенные измерения при нелинейной зависимости между аргументами**

При обработке косвенных измерений в случае нелинейной зависимости измеряемой величины  $Q = f(q_1, q_2, \dots, q_m)$  от прямо измеряемых аргументов  $q_i$  используют метод линеаризации. Нелинейная функция  $f$  разлагается в ряд Тейлора, и выделяется линейная часть ряда. Каждый из аргументов  $q_j$  измеряется с некоторой погрешностью  $\Delta q_j$ . Поскольку реальные погрешности малы по сравнению с измеряемой величиной, можно заменить  $dq_j$  на  $\Delta q_j$ :

$$Q = f(q_1, q_2, \dots, q_m) = f(\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_m) + \sum_{j=1}^m \frac{\partial f}{\partial q_j} \Delta q_j + \tilde{R}, \quad (1.92)$$

где  $\frac{\partial f}{\partial q_j}$  — первая частная производная от функции  $f$  по аргументу  $q_j$ , вычисленная в точке  $(\tilde{q}_1, \tilde{q}_2, \dots, \tilde{q}_m)$ ;  $\tilde{q}_j$  — оценка результата измерения аргумента  $q_j$ ;  $\Delta q_j = q_j - \tilde{q}_j$  — отклонение результата отдельного измерения  $j$ -го аргумента от его средней арифметической оценки;  $\tilde{R}$  — остаточный член

$$\tilde{R} = \frac{1}{2} \sum_{k,l=1}^m \frac{\partial^2 f}{\partial q_k \partial q_l} \Delta q_k \Delta q_l. \quad (1.93)$$

Метод линеаризации применим, если остаточным членом  $\tilde{R}$  можно пренебречь. Это возможно, если

$$\tilde{R} < 0,8 \sqrt{\sum_{j=1}^m \left( \frac{\partial f}{\partial q_j} \right)^2 S^2(\tilde{q}_j)}, \quad (1.94)$$

где  $S(\tilde{q}_j)$  — оценка СКО случайной погрешности результата измерения аргумента  $q_j$ .

При этом абсолютная погрешность косвенного измерения величины  $Q$  вычисляется по формуле

$$\Delta Q = \sum_{j=1}^m \frac{\partial f}{\partial q_j} \Delta q_j. \quad (1.95)$$

Следовательно, после линеаризации функциональной зависимости обработка данных должна проводиться так же, как при линейной зависимости (1.78). При этом роль коэффициентов  $b_j$  выполняют частные производные  $\frac{\partial f}{\partial q_j}$ :

$$s(\tilde{Q}) = \sqrt{\sum_{j=1}^m \left( \frac{\partial f}{\partial q_j} \right)^2 s^2(\tilde{q}_j) + \sum_{l=1}^m \sum_{\substack{k=1 \\ k \neq l}}^m \tilde{\rho}_{kl} \frac{\partial f}{\partial q_k} \frac{\partial f}{\partial q_l} s(\tilde{q}_k) s(\tilde{q}_l)} \quad (1.96)$$

Если применим метод линеаризации и погрешности аргументов распределены по нормальному закону, то и погрешность результата имеет нормальное распределение. В этом случае доверительный интервал для погрешности результата строят обычным образом: либо по формуле (1.90), либо по (1.91) в зависимости от числа измерений с заменой коэффициентов линейной зависимости  $b_j$  на  $\frac{\partial f}{\partial q_j}$ . Если число аргументов больше четырех,

и распределения их погрешностей одномодальны, а среди погрешностей нет резко выделяющихся, то результат также можно считать распределенным нормально.

В общем случае при нелинейной зависимости (1.77) функция распределения погрешностей результата измерений неизвестна и может отличаться от нормальной, даже если погрешности аргументов распределены нормально. При строгом подходе в этом случае доверительный интервал для случайной погрешности не может быть построен обычным образом, то есть с использованием распределения Стьюдента. Для интервальной оценки возможно использовать неравенство Чебышева с рассмотренными в разделе 1.3 оговорками.

В качестве примера рассчитаем погрешности результата по известным погрешностям измеряемых аргументов для некоторых простых нелинейных зависимостей.

Для функции

$$Q = q_1 \cdot q_2 \cdots q_m \quad (1.97)$$

относительная погрешность  $\delta Q$  очень просто выражается через относительные погрешности аргументов  $\delta q_j$ . Действительно, если прологарифмировать уравнение (1.97), то получим

$$\ln Q = \ln q_1 + \ln q_2 + \dots + \ln q_m. \quad (1.98)$$

После дифференцирования (1.98) и замены дифференциалов малыми приращениями получаем

$$\frac{\Delta Q}{Q} = \frac{\Delta q_1}{q_1} + \frac{\Delta q_2}{q_2} + \dots + \frac{\Delta q_m}{q_m}, \quad (1.99)$$

или

$$\delta Q = \delta q_1 + \delta q_2 + \dots + \delta q_m, \quad (1.100)$$

то есть относительная погрешность произведения просто равна сумме относительных погрешностей сомножителей.

Отметим, что прием логарифмирования и последующего дифференцирования функции  $Q = f(q_1, q_2, \dots, q_m)$  в целях вывода формулы для  $\delta Q = F(q_1, q_2, \dots, q_m, \delta q_1, \delta q_2, \dots, \delta q_m)$  удобно использовать во многих случаях. С помощью этого приема для зависимости более общего вида

$$Q = k q_1^\alpha q_2^\beta q_3^\gamma, \quad (1.101)$$

где  $k$  — постоянный множитель, получим

$$\delta Q = \alpha \delta q_1 + \beta \delta q_2 + \gamma \delta q_3. \quad (1.102)$$

Следовательно, для средних квадратичных отклонений, представленных в относительной форме, в соответствии с (1.25), (1.26) имеем:

$$\frac{\sigma_Q}{Q} = \sqrt{\alpha^2 \left(\frac{\sigma_{q_1}}{q_1}\right)^2 + \beta^2 \left(\frac{\sigma_{q_2}}{q_2}\right)^2 + \gamma^2 \left(\frac{\sigma_{q_3}}{q_3}\right)^2}. \quad (1.103)$$

Другие примеры СКО, рассчитанных по формуле (1.96) в предположении некоррелированности аргументов ( $\rho = 0$ ), представлены в табл. 1.3.

## 1.6. Суммирование составляющих погрешности

Как правило, результирующая погрешность измеренного значения физической величины складывается из нескольких аддитивных состав-

ляющих. При суммировании составляющих погрешности возникают две основные задачи, а именно:

- 1) суммирование составляющих одного вида (случайных или систематических);
- 2) суммирование случайной и систематической составляющих погрешности.

Таблица 1.3

**Примеры формул для вычисления СКО  
результата косвенных измерений  $Q = f(x,y)$**

Функция	$\sigma(Q)$	
	Абсолютное значение	Относительное значение
$ax$	$a\sigma_x$	$\sigma_x/\bar{x}$
$ax \pm by$	$\sqrt{a^2\sigma_x^2 + b^2\sigma_y^2}$	$\frac{\sqrt{a^2\sigma_x^2 + b^2\sigma_y^2}}{a\bar{x} \pm b\bar{y}}$
$x \cdot y$	$\sqrt{\bar{y}^2\sigma_x^2 + \bar{x}^2\sigma_y^2}$	$\sqrt{\left(\frac{\sigma_x}{\bar{x}}\right)^2 + \left(\frac{\sigma_y}{\bar{y}}\right)^2}$
$x/y$	$\frac{\sqrt{\bar{y}^2\sigma_x^2 + \bar{x}^2\sigma_y^2}}{\bar{y}^2}$	$\sqrt{\left(\frac{\sigma_x}{\bar{x}}\right)^2 + \left(\frac{\sigma_y}{\bar{y}}\right)^2}$
$x^\alpha$	$\alpha(\bar{x})^{\alpha-1}\sigma_x$	$\alpha \frac{\sigma_x}{\bar{x}}$
$a \exp(bx^\alpha)$	$ab\alpha(\bar{x})^{\alpha-1} \exp\left[b(\bar{x})^\alpha\right] \cdot \sigma_x$	$b\alpha(\bar{x})^{\alpha-1} \sigma_x$
$\sin x$	$\cos \bar{x} \cdot \sigma_x$	$\text{ctg} \bar{x} \cdot \sigma_x$

**1.6.1. Суммирование случайных составляющих погрешности**

Правило суммирования случайных погрешностей основано на формуле для оценки суммарного СКО, являющейся обобщением выражения (1.85):

$$S_\Sigma = \sqrt{\sum_{i=1}^m S_i^2 + 2\sum_{i<j}^m \rho_{ij} S_i S_j}, \quad (1.104)$$

где  $S_\Sigma$  — оценка СКО суммарной погрешности;  $S_i, S_j$  — оценки СКО составляющих погрешности;  $m$  — число составляющих погрешности.

Если составляющие суммарной случайной погрешности распределены по нормальному закону и для каждой составляющей известно не СКО, а доверительный интервал  $\Delta_i$  и соответствующая доверительная вероятность  $P_{дi}$ , то для каждой составляющей можно рассчитать оценку СКО:

$$S_i = \frac{\Delta_i}{z(P_{дi})}, \quad (1.105)$$

где  $z(P_{дi})$  — квантиль нормального распределения, соответствующий доверительной вероятности  $P_{дi}$ .

Если значения  $P_{д}$  одинаковы для всех составляющих, то получим:

для коррелированных составляющих ( $\rho_{ij} = \pm 1$ )

$$S_\Sigma = \sqrt{\sum_{i=1}^m S_i^2 \pm 2 \sum_{i<j}^m S_i S_j} = \sum_{i=1}^m \pm S_i = \frac{1}{z(P_{д})} \sum_{i=1}^m \pm \Delta_i, \quad (1.106)$$

где знак « $\pm$ » означает, что для составляющих с положительной корреляцией нужно считать величины  $S_i$  и  $\Delta_i$  имеющими знак «+», а для составляющих с отрицательной корреляцией — знак «-»;

для независимых составляющих ( $\rho_{ij} = 0$ )

$$S_\Sigma = \sqrt{\sum_{i=1}^m S_i^2} = \frac{1}{z(P_{д})} \sqrt{\sum_{i=1}^m \Delta_i^2}. \quad (1.107)$$

Поскольку при суммировании составляющих с нормальными распределениями результирующая погрешность также будет распределена нормально, то для суммарной погрешности доверительный интервал

$$\Delta_\Sigma = z(P_{д}) S_\Sigma. \quad (1.108)$$

С учетом (1.106) и (1.107) получим интервальную оценку:

для коррелированных составляющих

$$\Delta_\Sigma = \sum_{i=1}^m \pm \Delta_i; \quad (1.109)$$

для некоррелированных составляющих

$$\Delta_\Sigma = \sqrt{\sum_{i=1}^m \Delta_i^2}. \quad (1.110)$$



Суммирование погрешностей по формуле (1.109) называют арифметическим, а по формуле (1.110) — геометрическим. Поскольку реальные значения коэффициентов корреляции по абсолютному значению могут находиться в пределах от нуля до единицы, арифметическое суммирование обычно дает завышенное значение суммарной погрешности, а геометрическое — заниженное. Действительное значение находится между этими значениями.

Мы рассмотрели случай нормально распределенных составляющих погрешности. В общем случае для нахождения интервальной оценки суммарной погрешности необходимо по известным законам распределения суммируемых составляющих рассчитать результирующий закон распределения и соответствующий ему квантильный множитель. Можно без установления результирующего закона распределения воспользоваться неравенством Чебышева (1.72), (1.74), но при этом мы наверняка получим неоправданно завышенную оценку.

### ***1.6.2. Суммирование систематических погрешностей***

Суммарную систематическую погрешность определяют, оценивая ее составляющие, называемые элементарными систематическими погрешностями. Как правило, элементарные систематические погрешности характеризуют безусловными границами  $\pm\theta_j$ , такими что элементарная систематическая погрешность при  $i$ -м измерении  $\vartheta_{ji}$  всегда меньше или равна  $\theta_j$ , то есть  $|\vartheta_{ji}| \leq \theta_j$ .

Границы для каждой составляющей оценивают на основе анализа причин их возникновения и учета метрологических нормативных данных измерительных приборов. Если часть элементарных составляющих систематической погрешности устраняется введением поправок, то в качестве элементарных погрешностей выступают погрешности определения этих поправок. Для оценки границ полной систематической погрешности используют квазистатистические методы суммирования, условно рассматривая элементарные систематические погрешности как случайные величины. При выборе закона распределения для элементарной систематической погрешности на практике руководствуются следующими правилами: если

известна оценка границ погрешности  $\pm\theta_j$ , то ее распределение считают равномерным; если известна оценка СКО, то распределение считают нормальным. Применение этих правил позволяет статистически суммировать элементарные систематические погрешности и обычно приводит к осторожным, но не слишком завышенным оценкам систематической погрешности результата измерения.

При равномерном законе распределения элементарных систематических погрешностей границу суммарной систематической погрешности находят по формуле

$$\theta_P = k_P \sqrt{\sum_{j=1}^m \theta_j^2}, \quad (1.111)$$

где  $k_P$  — коэффициент, зависящий от доверительной вероятности  $P_d$ , числа элементарных составляющих систематической погрешности  $m$  и соотношения между ними.

Значение  $k_P$  определяется из характеристик распределения для суммарной погрешности, построенного как композиция прямоугольных распределений элементарных погрешностей на соответствующих интервалах  $\pm\theta_j$ . Величина  $\theta_j$  — граница  $j$ -й неисключенной систематической погрешности.

При доверительных вероятностях  $P_d = 0.90$  и  $P_d = 0.95$ , обычно используемых при измерении, значение поправочного коэффициента  $k_P$  мало чувствительно к изменению числа элементарных составляющих погрешности  $m$ . На практике при  $P_d = 0.90$  принимают  $k_{0.90} = 0.95$ , а при  $P_d = 0.95$   $k_{0.95} = 1.1$ . Для больших значений  $P_d$  используют таблицы значений при разных  $m$ .

При малом числе составляющих ( $m \leq 4$ ) после нахождения границы  $\theta_P$  по формуле (1.111) необходимо сравнить ее с арифметической границей  $\theta_a$ , вычисляемой как сумма границ:

$$\theta_a = \sum_{j=1}^m \theta_j. \quad (1.112)$$

В качестве окончательной границы рекомендуется принять меньшее из значений  $\theta_P$ ,  $\theta_a$ . Для малого числа составляющих граница  $\theta_a$  обычно незначительно превышает  $\theta_P$ .

При большом числе слагаемых ( $m \geq 4$ ) результирующая погрешность имеет практически нормальное распределение. Оценка дисперсии этого распределения равна сумме дисперсий слагаемых, имеющих прямоугольные распределения:

$$S_{\theta}^2 = \sum_{j=1}^n \frac{\theta_j^2}{3}. \quad (1.113)$$

Задавшись доверительной вероятностью  $P_d$ , получим  $\theta_p$  как границу доверительного интервала для нормального распределения:

$$\theta_p = z_p S_{\theta}, \quad (1.114)$$

где  $z_p$  — квантиль нормального распределения.

### ***1.6.3. Суммирование систематических и случайных погрешностей***

Строгого решения задача суммирования систематических и случайных погрешностей не имеет. На практике руководствуются следующими правилами.

Если граница суммарной неисключенной систематической погрешности  $\theta_p$  и оценка результирующего СКО случайной погрешности  $S$  удовлетворяют неравенству

$$\theta_p < 0,8S, \quad (1.115)$$

то пренебрегают систематической составляющей погрешности и учитывают только случайную погрешность результата. При этом доверительные границы

$$\Delta_p = t(P, \nu)S, \quad (1.116)$$

где  $t(P, \nu)$  — коэффициент Стьюдента, зависящий от доверительной вероятности  $P_d$  и числа проведенных измерений  $n$  ( $\nu = n - 1$ ).

Если

$$\theta_p > 8S, \quad (1.117)$$

то пренебрегают случайной составляющей погрешности и погрешность результата характеризуют лишь границами суммарной неисключенной систематической погрешности:

$$\Delta_p = \theta_p. \quad (1.118)$$

При этом погрешность, возникающая из-за пренебрежения одной из со-

ставляющих погрешности при выполнении условия (1.115) или (1.117), по оценкам не превышает 15 %.

Если ни одно из перечисленных условий не выполняется, то есть

$$0,8 \leq \frac{\theta_P}{S} \leq 8, \quad (1.119)$$

необходимо учитывать обе составляющие погрешности. При этом

$$\Delta_P = K(t(P, \nu)S + \theta_P). \quad (1.120)$$

Коэффициент  $K$  зависит от отношения  $\theta_P/S$  и доверительной вероятности  $P_d$ . Обычно используют усредненные значения:  $K = 0.8$  при  $P_d = 0.95$  и  $K = 0.85$  при  $P_d = 0.99$ .

### 1.7. Погрешности и выбор методики эксперимента

Во многих случаях правильно выполненный учет погрешностей измерений помогает в выборе методики эксперимента. Вообще говоря, рационально так планировать эксперимент, чтобы ни одна из составляющих результирующей погрешности значительно не превышала остальные.

Уменьшать СКО результата  $\sigma_{\bar{x}}$ , улучшая точность измерений, то есть уменьшая  $\sigma$ , и проводя многократные отсчеты с последующей статистической обработкой, целесообразно только до тех пор, пока случайная погрешность превышает систематическую. Обычно сравнивают доверительный интервал  $\Delta_P$  для случайной погрешности при заданной доверительной вероятности  $P_d$  с оценкой величины систематической погрешности  $\Delta_c$ . На практике считается достаточным, с учетом невысокой точности оценки величины  $\sigma$ , достичь выполнения условия

$$\Delta_P \leq \Delta_c/3. \quad (1.121)$$

Дальнейшее уменьшение  $\sigma_{\bar{x}}$  не может привести к значительному снижению погрешности результата.

Не оправдано стремление экспериментатора к использованию как можно более точной аппаратуры. Не рационально уменьшать случайную погрешность аппаратуры (средства измерений)  $\Delta_{СИ}$  более чем до

$$\Delta_{\tilde{N}\tilde{E}} \leq \Delta_P/3. \quad (1.122)$$

Поскольку суммарная случайная погрешность  $\Delta_{\Sigma} = \sqrt{\Delta_P^2 + \Delta_{\tilde{N}\tilde{E}}^2}$ , то при вы-

полнении неравенства (1.122) вклад  $\Delta_{\text{СИ}}$  в  $\Delta_{\Sigma}$  не будет превышать 10 %. При использовании более точной аппаратуры требуются бóльшие экономические затраты, а иногда и бóльшие затраты времени на измерения, причем точность измерений при этом не повышается.

При косвенных измерениях для снижения погрешности результата необходимо особое внимание уделять, прежде всего, тем прямо измеряемым аргументам, которые в соответствии с формулами (1.88) и (1.96) дают наибольший вклад в окончательную погрешность.

Рассмотрим такой пример. Измеряется величина  $Q = X/Y$  (или  $Q = XY$ ). Прямоизмеряемые величины  $X$  и  $Y$  имеют относительные погрешности  $\delta_X = 2\%$  и  $\delta_Y = 10\%$ . Таким образом,  $\delta_Q = \sqrt{2^2 + 10^2} \approx 10.2\%$ . Снижение погрешности измерения значения  $X$  вдвое (до  $\delta_X = 1\%$  приведет к  $\delta_Q = \sqrt{1^2 + 10^2} \approx 10.0\%$ . Если же снизить вдвое погрешность  $Y$  (до  $\delta_Y = 5\%$ ), получим для погрешности результата  $\delta_Q = \sqrt{2^2 + 5^2} \approx 5.4\%$ . При рассматриваемой связи измеряемой величины с аргументами величины  $X$  и  $Y$  рекомендуется измерять с примерно одинаковой относительной погрешностью.

Другая ситуация складывается, если величина  $Q$  связана с прямоизмеряемыми величинами соотношением  $Q = X + Y$  или  $Q = X - Y$ . Здесь рекомендации зависят от того, насколько сильно различаются значения  $X$  и  $Y$ .

Пусть  $Y \ll X$ . В этом случае измерение величины  $Y$  даже со значительно большей относительной погрешностью, чем величины  $X$ , незначительно снижает относительную погрешность результата. Например, при  $X = 10000 \pm 1$  и  $Y = 100 \pm 1$   $Q = 10100 \pm 2$  ( $\Delta Q = \sqrt{1^2 + 1^2} \approx 1.7$ ). При этом относительные погрешности таковы:  $\delta_X = 0.01\%$ ,  $\delta_Y = 1\%$ ,  $\delta_Q = 0.02\%$ . При измерении  $Y$  с погрешностью, в пять раз большей  $Y = 100 \pm 5$ , получим  $\delta_Q = 0,05\%$ .

В случае когда физическую величину вычисляют как разность двух близких по значению, независимо измеренных величин, погрешность результата катастрофически возрастает. Например, при  $X = 100 \pm 2$  и  $Y = 96 \pm 2$ , будем иметь  $Q = X - Y = 4 \pm 3$  ( $\Delta Q = \sqrt{2^2 + 2^2} \approx 2,83$ ). Хотя обе

прямоизмеряемые величины  $X$  и  $Y$  определены с относительной погрешностью 2 %, относительная погрешность результата  $\delta_Q = 75\%$ . Следовательно, необходимо найти другой метод измерения величины  $Q$ .

### 1.8. Представление окончательного результата

При симметричном доверительном интервале результат измерений записывают так:

$$X \pm \Delta, P_{\text{д}}, \quad (1.123)$$

где  $X$  — результат измерений,  $\Delta$  — доверительный интервал,  $P_{\text{д}}$  — доверительная вероятность.

Если вид функции распределения для составляющих погрешности результата измерений неизвестен, то результат, как правило, представляется в форме

$$X, S_X, n, \theta, \quad (1.124)$$

где  $S_X$  — оценка СКО случайной погрешности,  $n$  — число измерений,  $\theta$  — граница не исключенной систематической погрешности.

В форме (1.124) рекомендуется представлять и результаты работ, имеющих первостепенное значение. Это значительно облегчает сравнение результатов с данными экспериментов, проведенных другими исследователями. Зачастую в прецизионных экспериментах случайная погрешность значительно меньше систематической, которую может выявить последующий анализ методики измерения.

Предполагается, что в формулах (1.123), (1.124)  $X$  — среднее арифметическое. Применение какого-либо другого метода для оценки результата измерений должно быть указано и обосновано.

Доверительный интервал всегда выражается в тех же единицах, что и сама измеряемая величина. Например, правильной будет запись  $U = (5,372 \pm 0,004) \text{ В}$ , а запись:  $U = 5,372 \text{ В} \pm 4 \text{ мВ}$  неправильна.

Погрешность результата измерений физической величины определяет, сколько значащих цифр следует оставлять в результате измерений. Погрешность результата измерений, как следует из формулы (1.56), определяется с погрешностью около 20 %. Поэтому нужно руководствоваться следующими правилами.

1. Погрешность результата измерений указывается двумя значащими цифрами, если первая из них равна 1 или 2, и одной, если первая цифра 3 или больше. Погрешность всегда округляется в сторону увеличения ее по модулю.

2. Числовое значение результата измерений округляется до того же разряда, которым оканчивается округленное значение абсолютной погрешности. Цифры в разрядах числового значения результата, соответствующих разрядам погрешности, являются сомнительными. Цифры, находящиеся левее сомнительных, — верные, а правее — неверные. Неверные цифры в целых числах заменяются нулями, а в десятичных дробях отбрасываются. Если десятичная дробь оканчивается нулями, то они отбрасываются только до того разряда, которым оканчивается значение погрешности. Нельзя опускать значимые нули. Например, неправильно писать:  $25,7 \pm 0,03$ , нужно записать:  $25,70 \pm 0,03$ . Правильной записью чисел больших единицы является такая запись, при которой все цифры значащие. Последние — это все цифры, кроме нулей спереди, показывающих положение запятой, и нулей в конце числа, в том случае если они поставлены для обозначения разрядов вместо неизвестных цифр. Например, если число 440217 округляется до трех значащих цифр, а именно до 440000, то следует записать:  $440 \cdot 10^3$ , или  $44,0 \cdot 10^4$  или лучше всего —  $4,40 \cdot 10^5$ . Как было отмечено ранее, знаки в разрядах, соответствующих разрядам погрешности, сомнительные, а в более высоких разрядах — верные. Например, если для значения  $x = 6,9358$  абсолютная погрешность  $\Delta = 0,0025$ , то число  $x$  имеет три верных знака 6, 9, 3; остальные знаки сомнительные.

3. Округление производится лишь в окончательном результате, а все предварительные вычисления проводятся с одним-двумя лишними знаками. В частности, границы доверительных интервалов должны вычисляться по неокругленным значениям СКО. Это делается для того, чтобы сам процесс вычисления не вносил дополнительную погрешность.

### *Контрольные вопросы*

1. Охарактеризуйте термин «погрешность» применительно к измерениям.
2. Назовите виды погрешностей.
3. Как уменьшить методические погрешности?
4. Что такое функция распределения? Какие свойства она имеет?
5. Чем отличаются интегральная и дифференциальная функции распределения?
6. Нарисуйте графики интегральной функции распределения и распределения вероятностей для дискретной случайной величины.
7. Нарисуйте графики интегральной функции распределения и распределения плотности вероятностей для непрерывной случайной величины.
8. Чем отличаются начальные и центральные моменты случайных величин?
9. Что характеризует математическое ожидание случайной величины?
10. Что характеризуют дисперсия и среднее квадратичное отклонение случайной величины?
11. Опишите свойства нормального распределения. Какой смысл имеют параметры этого распределения?
12. Приведите примеры композиций распределений.
13. Что такое точечные оценки математического ожидания и дисперсии?
14. В чем различие равноточных и неравноточных измерений?
15. Что такое взвешенное среднее? Вес измерения?
16. Охарактеризуйте понятия доверительной вероятности и доверительного интервала.
17. Что такое промахи в измерениях? Назовите методы их исключения.
18. Определите погрешность косвенных измерений при линейной связи аргументов.
19. Определите погрешность косвенных измерений при нелинейной связи аргументов.
20. Назовите принципы суммирования погрешностей случайных и систематических.



## 2. АППРОКСИМАЦИЯ ФУНКЦИОНАЛЬНЫХ ЗАВИСИМОСТЕЙ ФИЗИЧЕСКИХ ВЕЛИЧИН

Одна из важнейших задач экспериментальной физики — нахождение функциональных зависимостей между физическими величинами. При этом обычно стараются поставить эксперимент так, чтобы изучить взаимозависимость только двух величин, поддерживая остальные влияющие физические величины постоянными. Таким образом, эксперимент сводится к наблюдению за тем, что происходит с интересующей нас величиной, если изменять некоторую другую физическую величину.

В эксперименте получают  $n$  значений  $y_i$  зависимой физической величины (называемой функцией отклика), измеренных при заданных экспериментатором значениях  $x_i$  влияющей на нее физической величины (фактора). Пара значений  $(x_i, y_i)$  — результат одного измерения,  $n$  — число измерений. Такие пары значений образуют выборку объема  $n$  из двумерной генеральной совокупности. В общем случае обе величины характеризуются ошибками измерения. Анализ зависимости начинают с графического представления в прямоугольной системе координат полученных пар наблюдений в виде экспериментальных точек с координатами  $(x_i, y_i)$ . Построение графиков является эффективным средством предварительного анализа экспериментальных зависимостей, поэтому графики рекомендуется строить обязательно. По экспериментальным точкам, как правило, можно провести гладкую кривую, которая приближенно описывает полученные результаты. Задача заключается в построении зависимости  $y = f(x)$  оптимальной формы в соответствии с выбранным критерием оптимальности. Из-за случайных погрешностей измерений график аппроксимирующей функции  $f(x)$  не должен обязательно проходить через все экспериментальные точки. Напротив, искомая функция должна сгладить случайные по-

грешности эксперимента, удовлетворив выбранному критерию оптимальности. Такую функцию называют сглаживающей или выравнивающей.

При известной теоретической формуле, описывающей связь двух указанных величин,  $y = f(a_1, a_2, \dots, a_k, x)$  по экспериментальным точкам находят параметры этой зависимости  $(a_1, a_2, \dots, a_k)$ . Даже, если формула адекватно описывает связь физических величин  $x$  и  $y$ , точки  $(x_i, y_i)$  отклоняются от теоретической кривой вследствие случайных погрешностей измерений. Это усложняет анализ зависимости и обуславливает погрешности в определении значений параметров  $a_i$ .

Если зависимость неизвестна, подбирают наиболее удобную эмпирическую формулу, описывающую основные особенности экспериментальной зависимости, и находят оптимальные параметры  $a_i$  для данной формулы. Задачу выбора эмпирической формулы невозможно формализовать, поскольку одну и ту же кривую на заданном интервале значений  $x$  можно с примерно одинаковой точностью описать различными аналитическими выражениями. При выборе аналитической зависимости обычно руководствуются тем, чтобы в дальнейшем ее было удобно использовать. Аппроксимирующая функция должна быть компактной и, по возможности, интерпретируемой. Последнее означает, что константам и выражениям, входящим в предлагаемую формулу, можно придать определенный физический смысл.

## **2.1. Простейшие виды аппроксимаций экспериментальных зависимостей**

Для подбора аппроксимирующей функции  $y = f(x)$  изучают график, построенный по точкам с координатами  $(x_i, y_i)$ . Плавная кривая графика должна проходить возможно ближе к экспериментальным точкам, и по обеим сторонам кривой должно находиться примерно одинаковое число точек. Если некоторые точки лежат далеко от кривой, то это часто свидетельствует о промахе при измерении. В таком случае излом кривой вблизи отдельной выпадающей точки недопустим. Поэтому, прежде чем проводить аппроксимацию, нужно из экспериментальных данных исключить промахи. Если на кривой зависимости ожидается максимум или минимум,

необходимо в области экстремума иметь достаточное количество экспериментальных точек для его четкого проявления.

Если говорить о чисто графическом анализе, то единственная зависимость, которую возможно провести на глаз по экспериментальным точкам, имеющим разброс вследствие случайных погрешностей измерений, — это прямая. Поэтому задача приближенного графического нахождения параметров аппроксимирующей зависимости легко решается только для линейной функции:

$$y = a + bx. \quad (2.1)$$

Значение  $a$  равно значению  $y$ , при котором проведенная прямая пересекает ось ординат,  $x = 0$ , а значение  $b$  находится как угловой коэффициент проведенной прямой. Угловой коэффициент прямой часто называют тангенсом угла наклона. Однако необходимо помнить, что численное значение тангенса нельзя определять по измеренному в градусах или радианах углу наклона прямой. Величина этого угла, а значит, и тангенса зависит от масштабов осей графика. Тогда как значение искомого углового коэффициента прямой во-первых, не должно зависеть от масштабов осей графика и во-вторых, он может иметь размерность, если размерности  $x$  и  $y$  различны.

Для определения  $b$  на проведенной прямой (рис. 2.1) выбирают две произвольные, удаленные друг от друга точки с координатами  $(\tilde{x}_1, \tilde{y}_1)$  и  $(\tilde{x}_2, \tilde{y}_2)$ . Здесь и далее знак «тильда» над значением переменной означает, что это рассчитанное по аппроксимирующей формуле, а не экспериментально полученное значение. Коэффициент  $b$  рассчитывается по формуле

$$b = \frac{\tilde{y}_2 - \tilde{y}_1}{\tilde{x}_2 - \tilde{x}_1} = \frac{\Delta\tilde{y}}{\Delta\tilde{x}}. \quad (2.2)$$

Если экспериментальная зависимость нелинейна, то для ее аппроксимации удобно подбирать такие функции, которые при замене переменных легко линеаризуются. Это эквивалентно использованию функционального масштаба по одной или обеим координатным осям графика. Два самых употребительных класса аппроксимирующих функций — степенные и показательные двухпараметрические зависимости.

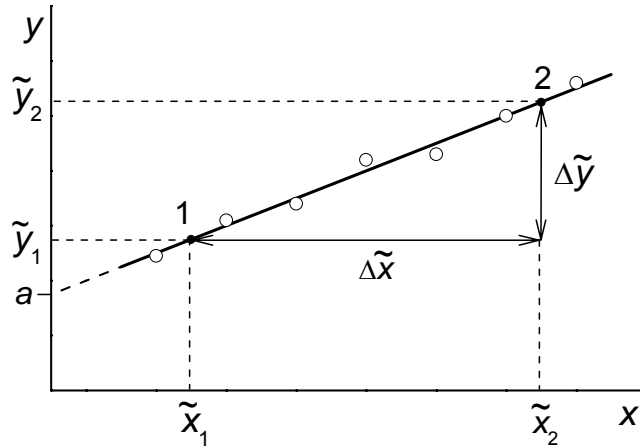


Рис. 2.1. К определению параметров линейной зависимости

Степенная функция (рис. 2.2, а)

$$y = ax^n \quad (2.3)$$

после логарифмирования

$$\lg y = \lg a + n \lg x \quad (2.4)$$

превращается в уравнение прямой, построенной в координатах  $(\lg x, \lg y)$  (рис. 2.2, б).

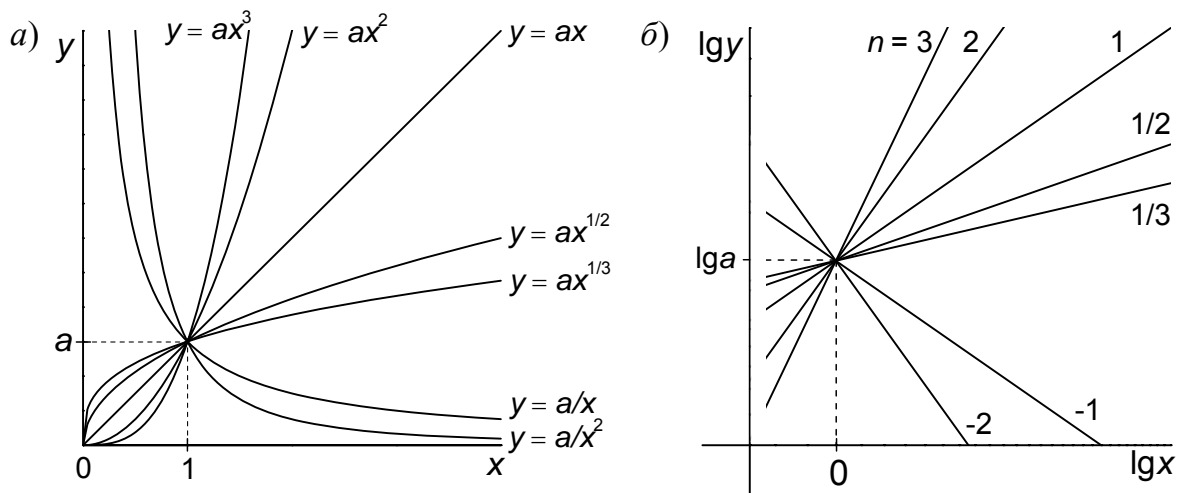


Рис. 2.2. Степенная функция  $y = ax^n$  при различных значениях  $n$  в линейном (а) и двойном логарифмическом (б) масштабе

Порядок нахождения параметров  $\lg a$  и  $n$  для такой линейной зависимости (рис. 2.3) аналогичен рассмотренному ранее (см. рис. 2.1):

$$n = \frac{\lg \tilde{y}_2 - \lg \tilde{y}_1}{\lg \tilde{x}_2 - \lg \tilde{x}_1} = \frac{\Delta \lg \tilde{y}}{\Delta \lg \tilde{x}}, \quad (2.5)$$

$$a = \frac{\tilde{y}_1}{\tilde{x}_1^n} \text{ или } a = \frac{\tilde{y}_2}{\tilde{x}_2^n}. \quad (2.6)$$

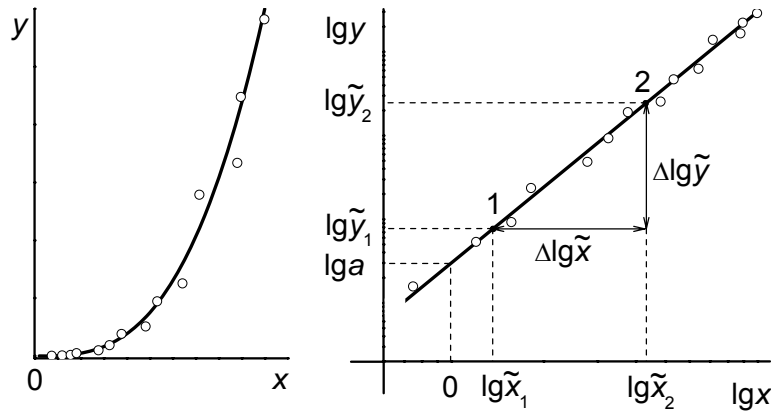


Рис. 2.3. Пример линейзации и графического определения параметров степенной функции  $y = ax^n$

Для определения параметра  $a$  удобно использовать соотношение  $\lg \tilde{y} = \lg a$  при  $\lg \tilde{x} = 0$ . Если значение  $\lg \tilde{x} = 0$  расположено вне интервала изменения  $x$  в эксперименте, то ордината  $\lg \tilde{y} = \lg a$  находится путем экстраполяции прямой до значения  $\lg \tilde{x} = 0$ .

Аналогично при аппроксимации результатов измерений показательной функцией (рис. 2.4, а) вида

$$y = ae^{bx} \quad (2.7)$$

для линейзации этой зависимости строят график в осях  $(x, \lg y)$ , то есть в полулогарифмическом масштабе (рис. 2.4, б):

$$\lg y = \lg a + b(\lg e)x, \quad (2.8)$$

где  $e$  — основание натуральных логарифмов  $e = 2.71828\dots$

При построении графиков в логарифмическом и полулогарифмическом масштабах по осям откладываются именно десятичные логарифмы соответствующих величин. Это дает возможность наносить на шкалу осей значения самой величины, кратные 10 и располагающиеся через равные интервалы. Но всегда следует помнить, что при расчетах параметров используются не сами величины, а их логарифмы.

Параметры показательной функции по линеаризованному графику определяются следующим образом. Значение  $\lg a$  равно значению  $\lg \tilde{y}$ , при котором построенная прямая пересекает ось ординат,  $x = 0$  (рис. 2.4, б). Из значений координат двух точек на прямой рассчитывают  $b$ :

$$b = \frac{\lg \tilde{y}_2 - \lg \tilde{y}_1}{\tilde{x}_2 - \tilde{x}_1} \cdot \frac{1}{\lg e} = \frac{\Delta \lg \tilde{y}}{\Delta \tilde{x}} \cdot \frac{1}{\lg e}. \quad (2.9)$$

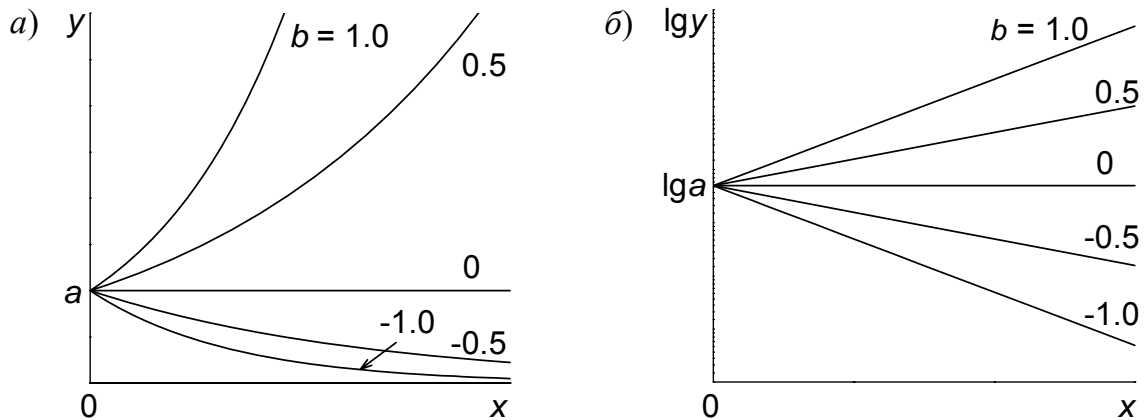


Рис. 2.4. Показательная функция  $y = ae^{bx}$  при различных значениях  $b$  в линейном (а) и полулогарифмическом (б) масштабах

Обычно удается линеаризовать двухпараметрические аппроксимации самого разного вида, например зависимость  $y = a + bx^2$  линеаризуется на графике  $y = f(x^2)$ , зависимость  $y = \frac{a}{x} + b$  — на графике  $y = f(1/x)$ . Зави-

симосьть  $y = \frac{x}{a + bx}$  можно привести к виду  $\frac{x}{y} = a + bx$  и, соответственно,

линеаризовать, откладывая по оси абсцисс  $x$ , а по оси ординат —  $x/y$ .

Функция  $y = ax^2 e^{-b/x}$  (температурная зависимость тока термоэлектронной эмиссии) представляет собой прямую в координатах  $\lg(y/x^2)$  и  $1/x$ .

Примеры линеаризующих преобразований для часто встречающихся двухпараметрических зависимостей представлены в табл. 2.1. После линеаризации вследствие замены переменных зависимости приобретают вид  $y' = a' + b'x'$ .

Таким образом, преобразование переменных может состоять в логарифмировании, возведении в степень или извлечении корней различной

степени. Спрявление графиков, описываемых неизвестными функциями, можно проводить, опираясь на рекомендации, представленные на рис. 2.5.

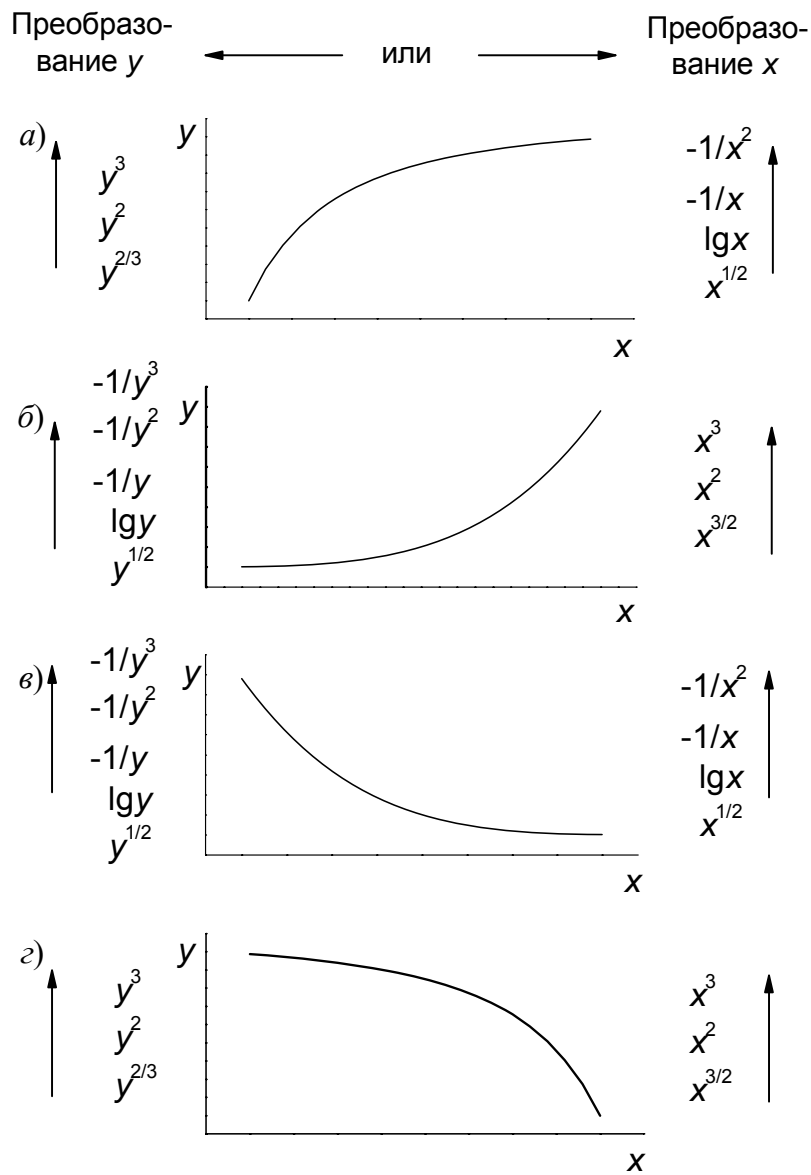


Рис. 2.5. Выбор линеаризующих преобразований переменных  $x$  или/и  $y$  в зависимости от формы кривой (эффективность спрявления кривых возрастает в направлении стрелок при преобразованиях, соответствующих столбцам слева и справа от графиков (Котельников, 1986))

Обычно выполняют преобразование масштаба вначале по одной из осей. По результатам и возможной физической интерпретации спрявление

с помощью преобразования  $x$  не то же самое, что с помощью преобразования  $y$ . Если попытка преобразования по одной из осей не дала хорошего результата, нужно попробовать преобразование по другой переменной. При необходимости улучшения результата целесообразно провести дополнительное преобразование и по второй переменной.

**Таблица 2.1**

**Линеаризующие преобразования для некоторых функций**

Функция	Замена переменных при линеаризации		Выражения для параметров линеаризованной зависимости	
	$y'$	$x'$	$a'$	$b'$
$y = a + b/x$	$y$	$1/x$	$a$	$b$
$y = 1/(a + bx)$	$1/y$	$x$	$a$	$b$
$y = x/(a + bx)$	$x/y$	$x$	$a$	$b$
$y = ab^x$	$\lg y$	$x$	$\lg a$	$\lg b$
$y = ae^{bx}$	$\ln y$	$x$	$\ln a$	$b$
$y = ae^{b/x}$	$\ln y$	$1/x$	$\ln a$	$b$
$y = 1/(a + be^{-x})$	$1/y$	$e^{-x}$	$a$	$b$
$y = ax^b$	$\lg y$	$\lg x$	$\lg a$	$b$
$y = a + b \lg x$	$y$	$\lg x$	$a$	$b$
$y = a/(b + x)$	$1/y$	$x$	$b/a$	$1/a$
$y = ax/(b + x)$	$1/y$	$1/x$	$1/a$	$b/a$
$y = ax^2/(b + x)$	$x^2/y$	$x$	$b/a$	$1/a$
	$x/y$	$1/x$	$a$	$b/a$
$y = a + bx^n$	$y$	$x^n$	$a$	$b$

Рассмотрим пример графической линеаризации зависимости, представленной на рис. 2.5, *a*. Это монотонно возрастающая функция, скорость роста которой уменьшается с увеличением значения аргумента (рис. 2.6, кривые *l*).



Кривая 4 (см. рис. 2.6, б) очень близка к прямой. Это свидетельствует о том, что аппроксимируемая зависимость близка к степенной  $y \sim x^{1/3}$ . При некотором опыте о близости зависимости к степенной можно догадаться и из небольшого отклонения кривой 2 от прямой (см. рис. 2.6, а) на всем ее протяжении. Если есть основания полагать, что зависимость степенная, дальнейший анализ нужно проводить, строя график зависимости в двойном логарифмическом масштабе (см. рис. 2.2, б).

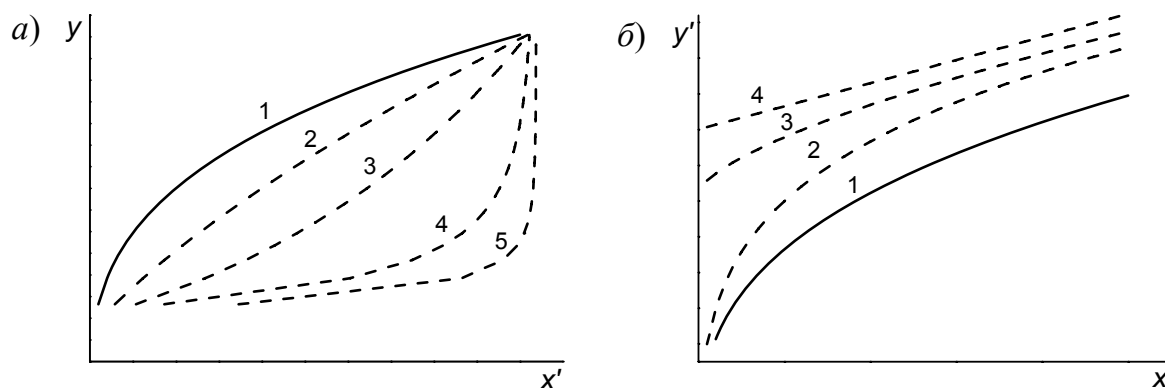


Рис. 2.6. Графическая линейзация зависимости  $I$  (см. рис. 2.5, а): а) — изменение масштаба по оси  $x'$ : 1 —  $x$ ; 2 —  $x^{1/2}$ ; 3 —  $\lg x$ ; 4 —  $(-1/x)$ ; 5 —  $(-1/x^2)$ . б) — изменение масштаба по оси  $y'$ : 1 —  $y$ ; 2 —  $y^{2/3}$ ; 3 —  $y^2$ ; 4 —  $y^3$

Если не удастся линейзовать всю зависимость, то выполняют спрямление отдельных участков кривой.

При анализе экспериментальных зависимостей, для описания которых необходимо использовать аппроксимирующие функции с более чем двумя параметрами, также возможно применение рассмотренных методов линейзации. Экспериментальную кривую можно рассматривать как состоящую из нескольких отрезков, каждый из которых аппроксимируется двух- или однопараметрической кривой. Например, для функции

$$y = a + bx + \frac{c}{x} \quad (2.10)$$

при достаточно больших значениях  $x$  последним слагаемым можно пренебречь. Зависимость становится линейной:

$$y \approx a + bx. \quad (2.11)$$

При малых значениях  $x$  можно пренебречь вторым слагаемым:

$$y \approx a + \frac{c}{x}. \quad (2.12)$$

Зависимость (2.12) легко линеаризуется в координатах  $\left(\frac{1}{x}, y\right)$ .

В физике часто встречаются зависимости, представляющие собой сумму двух и более экспонент с существенно различающимися по значению показателями (рис. 2.7) (например, вольтамперная характеристика полупроводникового  $p$ - $n$ -перехода, температурная зависимость концентрации носителей заряда в полупроводнике):

$$y = ae^{bx} + ce^{dx}, \quad d \gg b. \quad (2.13)$$

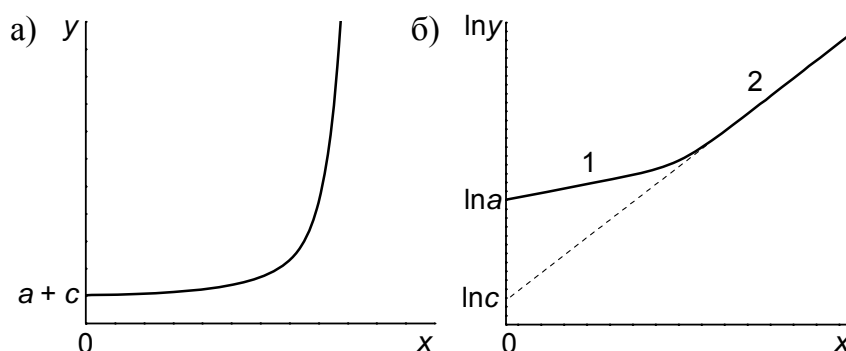


Рис. 2.7. Функция в виде суммы двух экспонент (2.13) до — (а) и после — (б) линеаризации (на участке 1  $y \approx ae^{bx}$ , на участке 2  $y \approx ce^{dx}$ )

Если диапазон изменения аргумента достаточно широк, на зависимости  $\ln y$  от  $x$  четко проявляются два линейных участка (рис. 2.7, б), по которым легко определить параметры  $a$ ,  $b$ ,  $c$  и  $d$ .

## 2.2. Определение параметров аппроксимирующей функции методом наименьших квадратов

При графическом анализе экспериментальных зависимостей был использован нестрогий критерий для построения сглаживающей кривой. Полагалось, что примерно одинаковое количество экспериментальных точек должно располагаться по обе стороны от сглаживающей кривой на примерно равных расстояниях от нее. При таком подходе невозможно утверждать, что найденные параметры зависимости оптимальны. Кроме того, не

удается рассчитать погрешности определения этих параметров. Для корректного построения аппроксимирующей зависимости необходимо задать строгий критерий выбора оптимальных параметров приближающей функции. Любой разумный критерий, очевидно, должен быть основан на минимизации параметра, учитывающего отклонения всей совокупности экспериментальных точек от аппроксимирующей зависимости.

При использовании аналитических методов расчет неизвестных параметров аппроксимирующей функции  $f(a_1, \dots, a_k, x)$  сводится к решению системы из  $n$  независимых уравнений ( $n$  равно числу проведенных измерений):

$$\begin{cases} f(a_1, a_2, \dots, a_k, x_1) = y_1, \\ f(a_1, a_2, \dots, a_k, x_2) = y_2, \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ f(a_1, a_2, \dots, a_k, x_n) = y_n. \end{cases} \quad (2.14)$$

Для определения  $k$  коэффициентов  $a_j$  необходимо не менее  $k$  независимых уравнений. Если  $n = k$ , решение системы (2.14) единственно, и для каждого коэффициента  $a_j$  можно найти определенное числовое значение. Однако при этом рассчитанные значения  $a_j$  будут отягощены неизвестными погрешностями, поскольку значения  $y_i$  измерены с погрешностью. Для нахождения погрешностей оценок  $a_j$  число уравнений в системе (2.14) должно превышать количество искомых величин.

При числе измерений  $n$  больше, чем число  $k$  искомых параметров ( $n > k$ ), число независимых уравнений системы избыточно. Если из этих уравнений в разных комбинациях составить несколько подсистем из  $k$  уравнений, каждая из них в отдельности даст свое решение. Но между собой эти решения несовместны. Это означает, что, подставив одно из полученных решений в другие подсистемы уравнений, мы их не удовлетворим. Поэтому уравнения системы типа (2.14) принято называть условными уравнениями. Решению каждой из подсистем будет соответствовать аппроксимирующая функция со своими параметрами. Если все их построить на графике, получим пучок аппроксимирующих кривых. Область, занимаемая этими кривыми, характеризует погрешности эксперимента. Если каким-либо образом произвести усреднение всех найденных кривых, ус-

редненная кривая будет гораздо точнее описывать исследованное явление. При этом параметры усредненной кривой будут определены тем точнее, чем больше сделано измерений. Так, для проведения прямой достаточно двух точно известных точек  $y_1(x_1)$  и  $y_2(x_2)$ . Однако при наличии погрешностей прямую необходимо проводить не через какие-то отдельные точки, а учитывая положение всех экспериментальных точек на графике.

Усреднение несовместных решений избыточной системы уравнений может быть произведено различными способами. Одним из наиболее широко используемых является метод наименьших квадратов (МНК), иначе называемый регрессионным анализом.

По методу наименьших квадратов наилучшие параметры аппроксимирующей кривой находят из условия минимума суммы квадратов всех отклонений по ординате  $\Delta y_i$  для экспериментальных точек  $(x_i, y_i)$  от расчетной кривой (принцип Лежандра) (рис. 2.8).

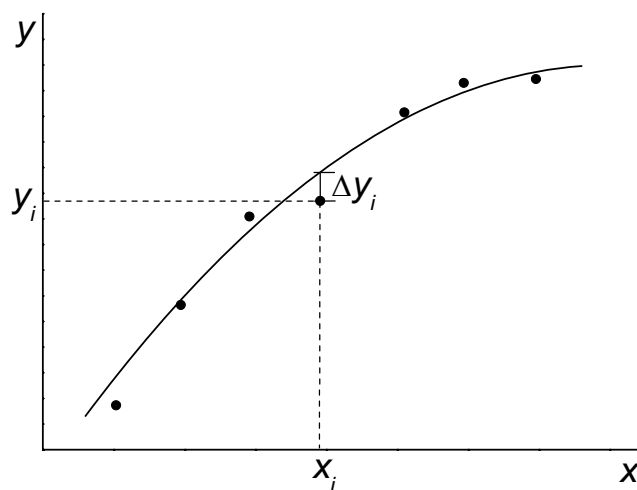


Рис. 2.8. Отклонения  $\Delta y_i$  экспериментальных точек  $(x_i, y_i)$  от аппроксимирующей кривой в методе наименьших квадратов

Теоретически показано, что МНК приводит к оценкам параметров, удовлетворяющим принципу максимального правдоподобия (то есть к наиболее вероятным оценкам), при следующих предположениях о свойствах результатов измерений  $(x_i, y_i)$ :

- 1) значения  $x_i$  известны точно, то есть погрешности значений  $x_i$  пренебрежимо малы по сравнению с погрешностями  $y_i$ ;
- 2) результаты измерений  $y_i$  содержат только случайные погрешности и

имеют одинаковые дисперсии  $\sigma_i^2$  (см. главу 1);

3) погрешности  $y_i$  распределены по нормальному закону (см. главу 1).

При этих условиях оценки параметров аппроксимирующей зависимости, получаемые по МНК, являются несмещенными и эффективными, то есть имеющими минимальные дисперсии. При выполнении только первых двух условий оценки МНК являются несмещенными и имеют минимальные дисперсии среди линейных оценок параметров. Если погрешности измерения  $x_i$  сравнимы с погрешностью  $y_i$ , то они при линейной зависимости  $y$  от  $x$  могут быть пересчитаны для дальнейшего анализа методом наименьших квадратов в погрешность  $y_i$ .

Обозначим  $\tilde{y}_i = f(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_k, x_i)$  рассчитанные по аппроксимирующей формуле значения функции отклика при оптимальных, но пока неизвестных нам значениях параметров  $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_k$ . Знак «тильда» ( $\tilde{\phantom{x}}$ ) над параметром, как обычно, означает, что это оценка параметра, то есть его приближенное значение. Между  $\tilde{y}_i$  и экспериментальными отсчетами  $y_i$  будут наблюдаться отклонения  $\Delta y_i$ , называемые остатками или остаточными погрешностями:

$$\Delta y_i = \tilde{y}_i - y_i. \quad (2.15)$$

МНК позволяет найти такие значения  $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_k$ , при которых

$$\sum_{i=1}^n [f(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_k, x_i) - y_i]^2 = \sum_{i=1}^n \Delta y_i^2 \rightarrow \min. \quad (2.16)$$

Для нахождения значений параметров  $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_k$ , минимизирующих сумму квадратов отклонений, необходимо вычислить частные производные от этой суммы по  $\tilde{a}_j$  и приравнять их нулю:

$$\frac{\partial \sum_{i=1}^n \Delta y_i^2}{\partial \tilde{a}_j} = 0; \quad j = 1, 2, \dots, k. \quad (2.17)$$

Причем частные производные второго порядка от суммы квадратов при этом должны быть положительными.

Систему (2.17) называют системой нормальных уравнений. Число уравнений системы равно числу неизвестных. Решая эту систему уравне-

ний для неизвестных  $\tilde{a}_j$ , находим наиболее достоверные по критерию (2.16) значения  $\tilde{a}_j$ .

Как пример применения МНК рассмотрим наиболее простую аппроксимацию — линейную зависимость между  $y$  и  $x$ . Уравнение искомой прямой представим в виде

$$y = a + bx. \quad (2.18)$$

Имеется  $n$  измеренных значений  $(x_i, y_i)$ . Ищем минимум выражения

$$\sum_{i=1}^n (\tilde{a} + \tilde{b}x_i - y_i)^2 = \min. \quad (2.19)$$

Приравняем частные производные по параметрам  $\tilde{a}$  и  $\tilde{b}$  нулю:

$$\begin{cases} \sum_{i=1}^n 2(\tilde{a} + \tilde{b}x_i - y_i) = 0, \\ \sum_{i=1}^n 2x_i(\tilde{a} + \tilde{b}x_i - y_i) = 0. \end{cases} \quad (2.20)$$

Получаем систему уравнений, линейных относительно параметров  $\tilde{a}$  и  $\tilde{b}$ :

$$\begin{cases} n\tilde{a} + \tilde{b}\sum_{i=1}^n x_i = \sum_{i=1}^n y_i, \\ \tilde{a}\sum_{i=1}^n x_i + \tilde{b}\sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i. \end{cases} \quad (2.21)$$

Систему (2.21) решаем, подсчитывая определители в соответствии с правилом Крамера

$$\tilde{a} = \frac{\Delta_{\tilde{a}}}{\Delta}, \quad \tilde{b} = \frac{\Delta_{\tilde{b}}}{\Delta}, \quad (2.22)$$

где  $\Delta$  — главный определитель системы,  $\Delta_{\tilde{a}}$  и  $\Delta_{\tilde{b}}$  — определители матриц, получаемых заменой столбца при соответствующем искомом параметре на столбец свободных членов из правой части системы (2.21).

Получим

$$\Delta = \begin{vmatrix} n & \sum x \\ \sum x & \sum x^2 \end{vmatrix} = n\sum x^2 - (\sum x)^2, \quad (2.23)$$

$$\Delta_{\tilde{a}} = \begin{vmatrix} \sum y & \sum x \\ \sum xy & \sum x^2 \end{vmatrix} = \sum y\sum x^2 - \sum xy\sum x, \quad (2.24)$$

$$\Delta_{\tilde{b}} = \begin{vmatrix} n & \sum y \\ \sum x & \sum xy \end{vmatrix} = n \sum xy - \sum x \sum y. \quad (2.25)$$

В формулах (2.23)–(2.25) для краткости у знака суммы  $\sum$  опущены индексы.

Таким образом,

$$\tilde{a} = \frac{\sum_{i=1}^n y_i \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i y_i \sum_{i=1}^n x_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}; \quad (2.26)$$

$$\tilde{b} = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}. \quad (2.27)$$

Коэффициент  $\tilde{b}$ , характеризующий наклон прямой, называют коэффициентом регрессии.

Отметим, что полученная прямая в соответствии с первым уравнением системы (2.21) проходит через точку  $(\bar{x}, \bar{y})$ , где  $\bar{x}, \bar{y}$  — средние значения  $x$  и  $y$  соответственно.

Оценка дисперсии отклонения  $y_i$  от прямой регрессии (2.18) вычисляется по формуле

$$\tilde{D}(Y) \equiv S^2 = \frac{\sum_{i=1}^n \Delta y_i^2}{n-2}. \quad (2.28)$$

Здесь оценка дисперсии записана как квадрат оценки среднего квадратичного отклонения  $S$  случайной величины  $Y$ .

Оценки дисперсий найденных значений параметров дают формулы:

$$S^2(\tilde{a}) = \frac{\sum_{i=1}^n x_i^2}{\Delta} S^2, \quad (2.29)$$

$$S^2(\tilde{b}) = \frac{n}{\Delta} S^2. \quad (2.30)$$

Интерпретация выражений (2.28)–(2.30) дана далее при обобщении ап-

проксимации МНК на случай трехпараметрической зависимости (см. (2.39) (2.40)).

Используя метод наименьших квадратов, можно найти наилучшие значения параметров  $a, b, \dots$  для любых *линейных по этим параметрам* аппроксимаций. Так, например, такие нелинейные по  $x$  аппроксимации, как

$$y = ax + bx^2 + cx^3 + \dots \quad (2.31)$$

или, в общем виде,

$$y = ag(x) + bh(x) + cm(x) + \dots, \quad (2.32)$$

где  $g(x), h(x), m(x)$  — известные функции от  $x$ , являются линейными аппроксимациями по параметрам  $a, b, c, \dots$ . Поскольку значения  $x_i$  нам известны, то любые функции от  $x$  (будь то  $x^2, x^3$  или  $g(x), h(x), m(x)$ ) при известном  $x_i$  превращаются просто в числовые коэффициенты при неизвестных  $a, b, c, \dots$

Используя МНК, можно рассчитать параметры и тех нелинейных по параметрам зависимостей, которые поддаются линеаризации. Один из методов преобразования нелинейных уравнений в линейные мы обсудили. Он состоит в такой замене переменных, которая приводит к условным уравнениям, линейным относительно параметров. При этом для расчета погрешностей оценок параметров исходных уравнений используют те же методы, что и при расчете погрешностей косвенных измерений (см. главу 1). Другой метод заключается в разложении нелинейной функции в ряд Тейлора с сохранением линейной части ряда. Но вообще говоря, решение системы условных уравнений нелинейного вида очень часто связано с непреодолимыми трудностями.

Чтобы было ясно, как решение системы линейных условных уравнений может быть представлено в общем виде, рассмотрим аппроксимацию результатов измерений функцией с тремя неизвестными параметрами  $a, b, c$ . Пусть система условных уравнений имеет вид

$$ag_i + bh_i + cm_i = y_i \quad (i = 1, 2, \dots, n; \quad n > 3), \quad (2.33)$$

где  $g_i = g(x_i), h_i = h(x_i), m_i = m(x_i)$ .

Соответствующая система нормальных уравнений для расчета оценок параметров  $\tilde{a}, \tilde{b}, \tilde{c}$ , получаемая из условий минимизации суммы квадратов



отклонений (2.17), будет такой:

$$\begin{cases} \tilde{a} \sum_{i=1}^n g_i^2 + \tilde{b} \sum_{i=1}^n g_i h_i + \tilde{c} \sum_{i=1}^n g_i m_i = \sum_{i=1}^n g_i y_i, \\ \tilde{a} \sum_{i=1}^n h_i g_i + \tilde{b} \sum_{i=1}^n h_i^2 + \tilde{c} \sum_{i=1}^n h_i m_i = \sum_{i=1}^n h_i y_i, \\ \tilde{a} \sum_{i=1}^n m_i g_i + \tilde{b} \sum_{i=1}^n m_i h_i + \tilde{c} \sum_{i=1}^n m_i^2 = \sum_{i=1}^n m_i y_i. \end{cases} \quad (2.34)$$

Далее для более компактной записи формул воспользуемся обозначениями Гаусса для сумм:  $\sum_{i=1}^n g_i^2 = [gg]$ ,  $\sum_{i=1}^n g_i h_i = [gh]$  и т.д. Очевидно, что

$$\sum_{i=1}^n g_i h_i = \sum_{i=1}^n h_i g_i, \text{ или } [gh] = [hg] \text{ и т.д.}$$

Нормальные уравнения приобретают вид

$$\begin{cases} [gg]\tilde{a} + [gh]\tilde{b} + [gm]\tilde{c} = [gy], \\ [hg]\tilde{a} + [hh]\tilde{b} + [hm]\tilde{c} = [hy], \\ [mg]\tilde{a} + [mh]\tilde{b} + [mm]\tilde{c} = [my]. \end{cases} \quad (2.35)$$

Используя правило Крамера, вычисляем оценки параметров через определители:

$$\tilde{a} = \frac{\Delta_{\tilde{a}}}{\Delta}, \quad \tilde{b} = \frac{\Delta_{\tilde{b}}}{\Delta}, \quad \tilde{c} = \frac{\Delta_{\tilde{c}}}{\Delta}, \quad (2.36)$$

где  $\Delta$  — главный определитель системы,

$$\Delta = \begin{vmatrix} [gg] & [gh] & [gm] \\ [hg] & [hh] & [hm] \\ [mg] & [mh] & [mm] \end{vmatrix}. \quad (2.37)$$

Определители  $\Delta_{\tilde{a}}$ ,  $\Delta_{\tilde{b}}$ ,  $\Delta_{\tilde{c}}$  получаем из главного определителя, заменяя столбец с коэффициентами при неизвестном  $\tilde{a}$ ,  $\tilde{b}$  или  $\tilde{c}$  соответственно столбцом со свободными членами из правой части системы уравнений (2.35).

Остаточные погрешности  $\Delta y_i$  вычисляем, подставляя в условные уравнения (2.33) оценки  $\tilde{a}$ ,  $\tilde{b}$ ,  $\tilde{c}$ :

$$\Delta y_i = \tilde{y}_i - y_i = (\tilde{a}g_i + \tilde{b}h_i + \tilde{c}m_i) - y_i \quad (2.38)$$

Оценку дисперсии остаточных погрешностей дает формула (обобщение формулы (2.28))

$$D(Y) \equiv S^2 = \frac{\sum_{i=1}^n \Delta y_i^2}{n - k}, \quad (2.39)$$

где  $n$  — число условных уравнений,  $k$  — число искомых параметров (в формуле (2.28)  $k = 2$ ).

Оценки дисперсий вычисленных значений параметров рассчитываются по формулам (обобщение формул (2.29), (2.30)):

$$S^2(\tilde{a}) = \frac{\Delta_{11}}{\Delta} S^2, \quad S^2(\tilde{b}) = \frac{\Delta_{22}}{\Delta} S^2, \quad S^2(\tilde{c}) = \frac{\Delta_{33}}{\Delta} S^2, \quad (2.40)$$

где  $\Delta_{11}$ ,  $\Delta_{22}$ ,  $\Delta_{33}$  — алгебраические дополнения элементов  $[gg]$ ,  $[hh]$ ,  $[mm]$  определителя, получаемые путем удаления из определителя  $\Delta$  столбца и строки, на пересечении которых находится данный элемент,

$$\Delta_{11} = \begin{vmatrix} [hh] & [hm] \\ [mh] & [mm] \end{vmatrix}, \quad \Delta_{22} = \begin{vmatrix} [gg] & [gm] \\ [mg] & [mm] \end{vmatrix}, \quad \Delta_{33} = \begin{vmatrix} [gg] & [gh] \\ [hg] & [hh] \end{vmatrix}.$$

Определив дисперсию оценок искомых параметров, строим доверительные интервалы для действительных значений параметров, для чего в соответствии с разделом 1.3 главы 1 нужно используем распределение Стьюдента. При этом число степеней свободы для распределения Стьюдента равно  $(n - k)$ , поскольку для определения параметров кривой регрессии необходимо выполнение  $k$  дополнительных условий. Из формул (2.40) следует, что точность определения параметров аппроксимирующей функции зависит от соотношения числа искомых параметров  $k$  и числа условных уравнений  $n$ . Чем больше число условных уравнений, равное числу проделанных пар измерений, по сравнению с числом искомых параметров, тем точнее полученный результат.

Как было отмечено ранее, если дисперсии для параметров линеаризованных зависимостей найдены методом наименьших квадратов, дисперсии параметров исходной нелинейной аппроксимации рассчитываются по ним теми же методами, что и дисперсии косвенных измерений (см. главу 1).

Проверкой согласия выбранной аппроксимирующей функции с экспериментальными данными может служить сравнение дисперсии остаточных погрешностей (2.39) с известной оценкой дисперсии отдельных измерений  $y_i$ . Значения указанных дисперсий должны быть близки.

Обсуждая МНК, мы рассматривали случай равноточных измерений величины  $y$ . В общем случае измерения  $y_i$  при разных  $x_i$  могут быть неравноточными. Если веса результатов  $w_i$  известны (например, оценены по результатам многократных измерений), то после умножения правой и левой частей условных уравнений (2.14) на  $\sqrt{w_i}$  получим систему равноточных условных уравнений

$$f(a_1, a_2, \dots, a_k, x_i) \sqrt{w_i} = y_i \sqrt{w_i}. \quad (2.41)$$

Действительно, дисперсия  $D$  значений правой части уравнений (2.41)

$$D(y_i \sqrt{w_i}) = \sigma_i^2 w_i = \sigma_i^2 \frac{\sigma_0^2}{\sigma_i^2} = \sigma_0^2 \quad (2.42)$$

одинакова для всех уравнений. Следовательно, для так преобразованных условных уравнений будут справедливы все формулы, полученные для равноточных измерений. Естественно, эти формулы будут включать в себя значения  $w_i$  в соответствии с выполненным преобразованием.

При использовании МНК следует быть внимательным, чтобы не получить совершенно неприемлемый результат. Как правило, такой результат есть следствие высокой чувствительности МНК к статистической неоднородности результатов измерений или к недостаточному числу экспериментальных точек в области быстрого изменения функции. Для оценки приемлемости результата аппроксимации рекомендуется найденное МНК решение представлять графически, при этом возможные грубые несоответствия легко выявляются.

Для построения зависимостей по экспериментальным данным помимо МНК, имеющего отмеченные ограничения в использовании, применяются и другие методы.

Когда значения аргументов  $x_i$  известны с погрешностями, сравнимыми с погрешностями  $y_i$ , при использовании МНК получают смещенные оценки, а главное, неправильные значения погрешностей параметров. Для об-

работки данных в этом случае используют весьма разнообразные конъюнктные методы.

Когда имеет место существенное отклонение распределений погрешностей  $y_i$  от предполагаемого нормального, неоднородность статистики  $y_i$  или промахи, МНК также приводит к результатам, имеющим большие погрешности, и не позволяет правильно оценить параметры аппроксимации. Например, при использовании МНК наличие даже одного-двух далеко отстоящих промахов может привести к абсурдному решению. В таком случае для нахождения параметров зависимости используют робастные методы, устойчивые к ошибкам.

### *Контрольные вопросы*

1. С какой целью нужно проводить аппроксимацию экспериментальных зависимостей?

2. Как следует проводить аппроксимирующую кривую на глаз?

3. Как определить параметры линейной зависимости?

4. Назовите основные правила графического спрямления нелинейных зависимостей.

5. Выполните графический анализ зависимостей  $y = a \exp(bx)$ ,

$$y = a \exp(-bx), y = a \exp(b/x), y = a + b \exp(nx), y = [a + b \exp(nx)]^{-1},$$

$$y = a + b \exp(-nx), y = ax^n \exp(-bx), y = a + bx^n, y = -\frac{a}{x^n} + \frac{b}{x^m} \text{ при } n < m.$$

Определите параметры  $a, b, n, m$  из проведенного графического анализа.

6. В чем заключаются достоинства метода наименьших квадратов при определении параметров аппроксимирующей функции?

7. В каких случаях аппроксимация экспериментальных результатов с помощью метода наименьших квадратов приводит к большим ошибкам?

### **3. ИСПОЛЬЗОВАНИЕ ПРОГРАММЫ MICROCAL ORIGIN ДЛЯ ГРАФИЧЕСКОГО ОТОБРАЖЕНИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ**

Microcal Origin — одна из лучших программ для графического представления экспериментальных данных. Origin позволяет просто и эффективно отображать числовые данные в виде двух- или трехмерных графиков самого разного типа (всего более 40 типов графиков). Все атрибуты графического представления научных данных можно легко изменять, что позволяет удовлетворять любым требованиям к оформлению графиков. Подготовленные в Origin графики пригодны как для профессиональной печати в научно-технических изданиях, так и для публикации в электронном виде. Origin содержит и достаточно мощные средства для анализа и математической обработки численных данных.

Возможности Origin можно практически неограниченно расширять, используя встроенные языки программирования LabTalk и Origin C. Последний появился в 7-й версии программы и является C++-подобным объектно-ориентированным языком. В отличие от интерпретируемого языка LabTalk язык Origin C — язык компилируемый, и вследствие этого написанные на нем программы обработки и представления данных в Origin работают значительно быстрее. Создатели программы рекомендуют использовать именно Origin C. Язык LabTalk сохранен для совместимости с более ранними версиями.

Далее будут рассмотрены основные возможности программы по графическому представлению и аппроксимации экспериментальных зависимостей на примере 7-й версии Origin. Однако все приведенные рекомендации в полной мере применимы и для версий 6 и 7.5, а в значительной их части и для пятой версии.

### 3.1. Главное окно Origin

Интерфейс Origin соответствует общепринятым стандартам Windows (рис. 3.1).

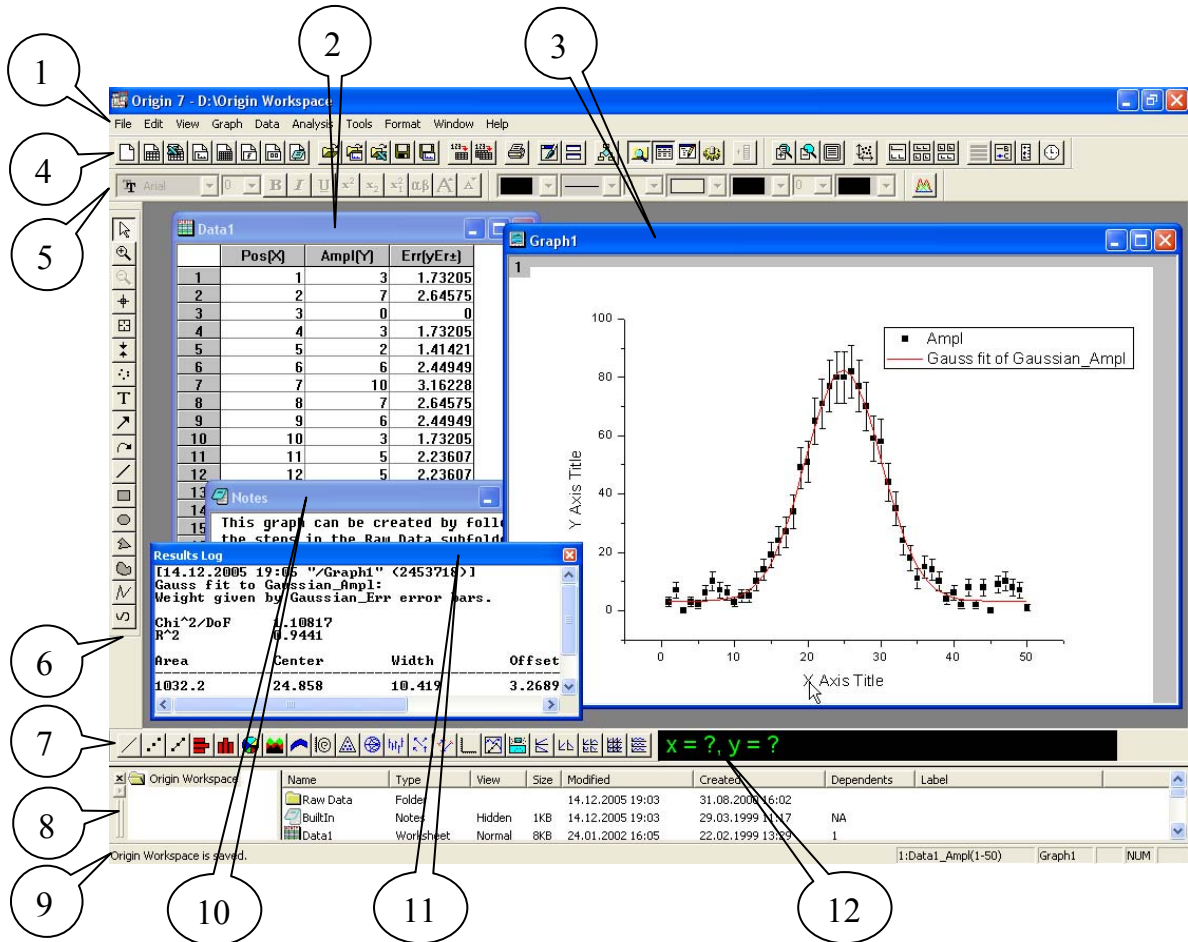


Рис. 3.1. Рабочее пространство Origin: 1 — строка главного меню; 2 — Worksheet Window (Рабочая таблица, или Окно данных); 3 — Graph Window (Графическое окно); 4–7 — панели инструментов: 4 — Standard (Стандартная) и Graph (График); 5 — Format (Форматирование (надписей)) и Style (Стиль (рисованных объектов)); 6 — Tools (Инструменты); 7 — 2D Graphs (Двухмерные графики); 8 — Project Explorer (Проводник проекта); 9 — строка состояния; 10 — Notes Window (Окно заметок); 11 — Results Log Window (Окно протоколирования результатов); 12 — Data Display (Дисплей данных)

Назначение многих элементов главного окна проекта Origin, таких как основные пункты меню **File** (Файл), **Edit** (Редактирование), **View** (Вид) и др., и соответствующих им кнопок на панелях инструментов, хорошо понятно. Кнопками на панелях инструментов продублировано большинство команд главного меню программы. При наведении указателя мыши на

кнопку появляется всплывающая короткая подсказка о назначении кнопки, а более подробная информация выводится в строке состояния в нижней части окна.



При настройках по умолчанию в главном окне программы присутствуют далеко не все панели инструментов. Чтобы отобразить необходимую панель инструментов, нужно воспользоваться пунктом **Toolbars...** (Панели инструментов) меню **View** (далее обращение к пункту меню будем записывать так: **View/Toolbars...**). В открывшемся диалоговом окне **Customize Toolbar** (Настройка панели инструментов) устанавливаем (отмечаем галочкой) флажок, соответствующий выбираемой панели, и щелкаем мышью по кнопке **ОК**. Новая панель появляется в строке инструментов. В главном окне можно разместить и отдельные кнопки из любой панели инструментов. Выбранную кнопку нужно просто перетащить мышью при нажатой левой кнопке в главное окно Origin из вкладки **Button Groups** (Группы кнопок) диалогового окна. Следует отметить, что в Origin часть команд доступна только через кнопки панелей инструментов и отсутствует в главном меню. Подробно с элементами интерфейса Origin мы будем знакомиться по мере детального обсуждения приемов работы с программой.

Основные типы окон в Origin — **Worksheet Window** (Окно данных) с именами по умолчанию Data1, 2 и т.д. и **Graph Window** (Графическое окно) с именами по умолчанию Graph1, 2 и т.д. Пункты главного меню и состав входящих в них команд меняются в зависимости от того, окно какого возможного в Origin типа активно в данный момент в главном окне программы. Например, при активном окне данных в главном меню появляется пункт **Statistics** (Статистика) с командами, позволяющими проводить статистический анализ данных в таблице, при активном графическом окне — команда **Data** (Данные) для операций с данными на графике.

Как во всякой современной программе с дружественным к пользователю интерфейсом, каждое действие, как правило, можно выполнить разными способами, выбирая для себя наиболее удобный. Мы будем использовать для выполнения необходимых команд в первую очередь кнопки панелей инструментов и контекстные меню, открывающиеся по щелчку правой кнопки мыши на соответствующих объектах в окнах Origin. Как правило открыть диалоговое окно для изменения свойств объекта можно

также, сделав двойной щелчок левой кнопкой мыши по этому объекту. Рекомендуем пользоваться этой широко реализованной в Origin возможностью.

### 3.2. Работа с таблицей Origin

При запуске программы в окне нового проекта Origin открывается окно с таблицей (**Worksheet Window**), в которую необходимо поместить данные для последующей обработки. Далее это окно будем называть просто таблица. При необходимости можно открыть дополнительные окна данных (несколько таблиц) щелчком по кнопке  (**New Worksheet**). В каждом столбце таблицы располагается набор данных, значения которых при построении графика будут откладываться по одной из осей. Данные в ячейки таблицы можно вносить либо с клавиатуры, либо импортируя их из файлов различных форматов, созданных другими программами. Возможен, в частности, импорт данных из текстовых файлов командой меню **File/Import/Single ASCII...** или щелчком по кнопке  (**Import ASCII**) для заполнения сразу нескольких столбцов таблицы. Колонки чисел в текстовом файле должны быть отделены друг от друга каким-либо разделителем. Origin автоматически распознает в качестве разделителя пробел, символ табуляции и запятую.

Большое количество наборов данных, которые удобно использовать при обучении построению графиков, представлено в папке C:\Program Files\OriginLab\OriginPro70\Samples\Data\. Рекомендуем воспользоваться наборами данных Expdecay.DAT, Fitpoly.DAT, Gaussian.DAT.

Числа в ячейки можно вводить либо в общепринятой форме, либо в так называемом научном формате с мантиссой и порядком (например, 518 как 5.18E2).

Для ввода данных в столбец таблицы могут использоваться и функциональные выражения. Для этого прежде всего выделяется столбец таблицы щелчком мыши по его заголовку. Из контекстного меню для выделенного столбца выбирается команда **Set Column Values...** (Установить значения в столбце). В текстовое поле в нижней части открывшегося одноименного диалогового окна (рис. 3.2) вводим выражение, определяющее



вносимые в ячейки столбца значения. В выражение могут входить встроенные в Origin функции и значения из столбцов таблицы. Функции и имена столбцов выбираются из соответствующих выпадающих списков. Вставка осуществляется щелчком мыши по кнопке **Add Function** (Добавить функцию) или **Add Column** (Добавить столбец). В выпадающем списке представлены далеко не все функции, встроенные в Origin. Всего встроенных математических функций почти 50. Среди них — специальные и многопараметрические функции. Все их можно применять для ввода в данном окне. Ознакомиться с ними можно в справке к языку LabTalk.

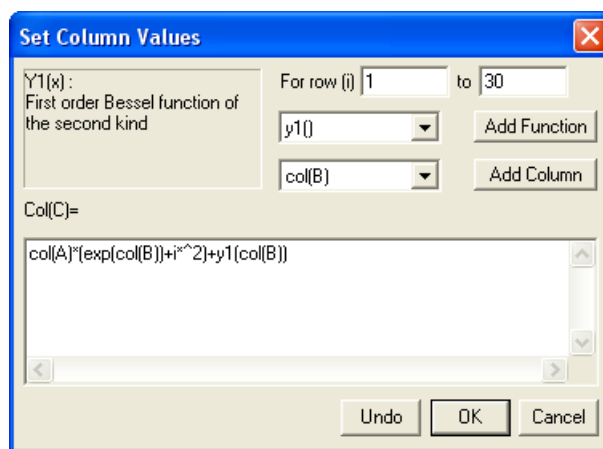


Рис. 3.2. Диалоговое окно для ввода в столбец таблицы значений, определяемых функциональной зависимостью

В функциональном выражении можно использовать переменную  $i$ , значение которой равно номеру строки таблицы. Используются общепринятые обозначения операций (+, −, \*, /, ^) и условия о приоритете операций. Приоритет операций можно изменить введением в выражение скобок.

По умолчанию колонкам таблицы присваиваются имена A(X), B(Y), C(Y), ... . Буквы в скобках указывают на роль данных в колонке: X — значения, откладываемые по оси X; Y — по оси Y. Еще один интересующий нас тип колонок — колонки со значениями погрешностей по соответствующим осям ( $x \text{Er} \pm$ ,  $y \text{Er} \pm$ ). Тип колонки можно изменить, вызвав контекстное меню щелчком правой кнопки мыши по заголовку столбца и выбрав пункт меню **Set As** (Задать в качестве).

Имя столбца можно изменить на более содержательное в диалоговом окне **Worksheet Column Format** (Формат колонки таблицы) (рис. 3.3). Вызывается окно двойным щелчком по заголовку таблицы или из контекстного меню столбца (пункт **Properties...** (Свойства)). В этом окне можно задать или изменить любые атрибуты колонки. В частности, в дополнение к имени колонки, которое может содержать не более 18 знаков и только латинские буквы и цифры без пробелов, можно ввести **Column Label** (Маркировка колонки). Маркировка колонки позволяет более развернуто характеризовать данные, представленные в колонке, поскольку может содержать до 63 знаков в одной строке, включая специальные и пробел, и иметь несколько строк. Ширину колонки можно изменить не только в данном диалоговом окне, но и просто перетаскивая границу между заголовками колонок при нажатой левой клавише мыши. При этом курсор должен иметь вид двунаправленной стрелки.

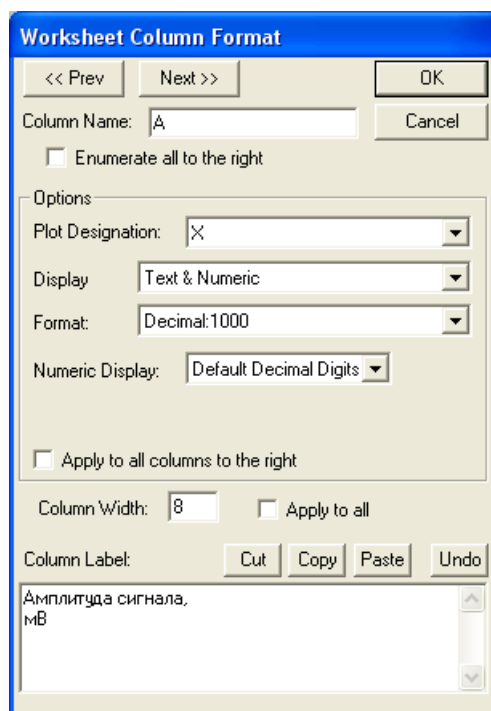


Рис. 3.3. Диалоговое окно для задания атрибутов колонки таблицы данных

В таблицу можно добавлять или удалять из нее ячейки, строки и столбцы, используя контекстное меню для предварительно выделенных элементов таблицы. Добавляемые в таблицу ячейки и строки вставляются

выше выделения, а столбцы — левее. Необходимое число столбцов можно добавить и в конец таблицы, воспользовавшись главным меню: **Column/Add New Columns...** (Колонка/Добавить новые колонки). Весь столбец или строка выделяются щелчком по заголовку столбца или номеру строки соответственно, а группа смежных столбцов, строк или ячеек — либо протаскиванием курсора мыши при нажатой левой клавише, либо щелчком мыши на первом выделяемом элементе, а затем на последнем при нажатой клавише **Shift**. Для выделения несмежных столбцов необходимо при щелчке мышью удерживать нажатой клавишу **Ctrl**.

Если в проекте несколько таблиц с данными, удобно дать им содержательные имена и снабдить надписями (**Label**), выбрав из контекстного меню по щелчку на заголовке окна таблицы пункт **Rename...** (Переименовать).


### 3.3. Построение графиков на основе данных таблицы

Объекты графического окна Origin имеют иерархическую структуру. Графическое окно включает в себе единственную графическую страницу (**Page**). Это белая область графического окна, включающая в себя графики и другие объекты. Именно она выводится на печать. Графическая страница может содержать до 50 слоев. Слой (**Layer**) содержит единственный набор координатных осей. В этих координатных осях могут быть построены зависимости по нескольким наборам данных. Помимо построенных зависимостей слой может содержать в себе и другие объекты, например текстовые или рисованные объекты. При щелчке мышью по одной из координатных осей слоя появляется рамка выделения слоя. Сдвигая маркеры этой рамки мышью при нажатой левой кнопке, можно менять размеры слоя. Для перемещения слоя, имеющего рамку выделения, достаточно щелкнуть левой кнопкой мыши по внутренней области слоя и перетащить его при нажатой кнопке (при этом маркер имеет вид креста со стрелками). Слои на графической странице могут быть взаимно разнесены, но могут и накладываться друг на друга. Кроме того, слои могут быть связаны друг с другом пространственно или математически таким образом, что изменения в положении, размерах или масштабах шкал родительских слоев будут влиять на положение, размер и масштабы шкал дочерних слоев. Поскольку

каждый слой является отдельным объектом Origin, возможна индивидуальная настройка свойств этого объекта, что в сочетании с комбинацией слоев обеспечивает практически неограниченные возможности для графического представления числовых данных.

Основные причины добавления нового слоя на графическую страницу:

- Необходимость представить график данных на одной странице в разных масштабах, например линейном и логарифмическом.
- Желание разместить на одной графической странице несколько графиков для разных данных.
- Необходимость сделать вставку в поле основного графика в виде отдельной небольшой координатной системы, на которой будет в бóльшем масштабе, как бы под микроскопом, представлена часть основного графика.

Для построения двухмерного графика в декартовой системе координат достаточно выделить в таблице столбец типа Y со значениями ординат графика (при этом может быть выделен и столбец типа X) и выбрать один из пунктов меню **Plot/Line** (График/Линия), **Plot/Scatter** (График/Точки) или **Plot/Line + Symbol** (График/Линия + Символы) в соответствии с желаемым видом графика. При выборе **Line** на графике будет представлена только линия, последовательно соединяющая точки данных  $(x_i, y_i)$ , при выборе **Scatter** — с помощью выбранного вида значков только точки данных, при выборе **Line + Symbol** — как точки, так и соединяющая их линия. Еще удобнее для построения графика воспользоваться соответствующими кнопками панели **2D Graphs** (Двухмерные графики) , располагающейся в нижней части окна Origin (см. рис. 3.1). Если в таблице представлены данные о погрешностях в столбце типа  $y \pm \text{Er}$ , для отображения этих данных на графике в общепринятой форме выделяют наряду со столбцом типа Y соответствующий ему столбец типа  $y \pm \text{Er}$ .

Если в таблице несколько столбцов типа Y, то для построения этих нескольких зависимостей, отражающих разные наборы данных на одном графике, достаточно выделить соответствующие столбцы и щелкнуть мышью по кнопке на панели **2D Graphs**, отвечающей выбранному типу графика. Построенные зависимости будут различаться цветом линий, цветом

и видом значков для точек данных. При этом все атрибуты зависимостей окажутся взаимосвязанными и будут изменяться согласованно при изменении характеристик зависимости, соответствующей левому из выделенных столбцов типа Y. Такую взаимосвязь можно снять в диалоговом окне **Plot Details** (Детали графика) на вкладке **Group** (Группа) (рис. 3.4).

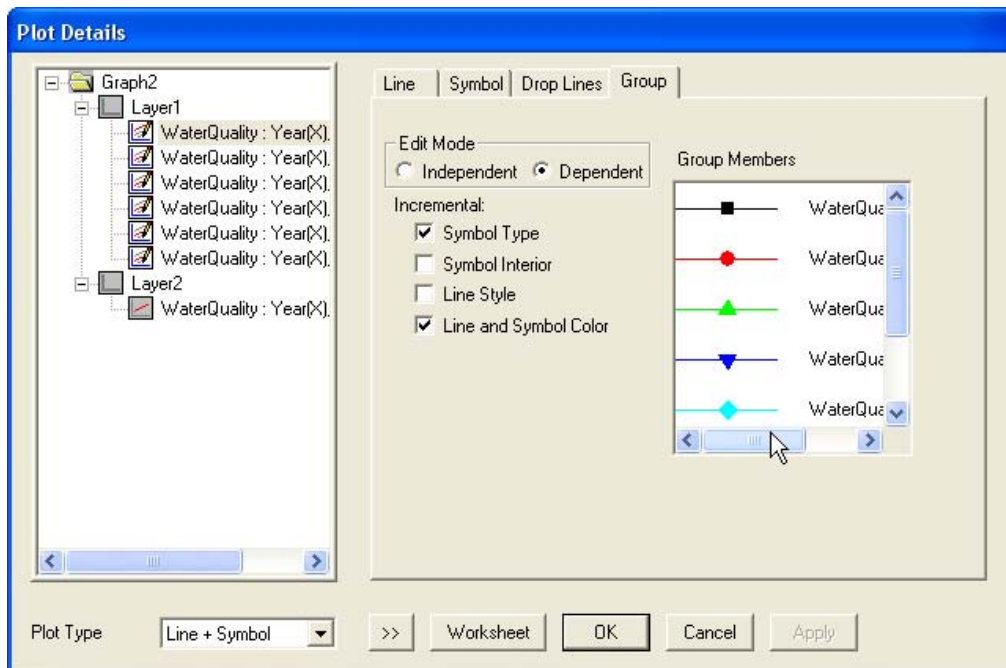



Рис. 3.4. Диалоговое окно Plot Details для управления деталями графического представления зависимостей

В таблице данных может быть более одного столбца типа X. При этом в качестве значений аргумента при построении графика автоматически используются значения в столбце типа X, ближайшем слева к выделенному столбцу типа Y.

Таким образом, при использовании описанных выше приемов построения графиков программа Origin создает новое графическое окно, один слой в этом окне и строит столько зависимостей в координатных осях созданного слоя, сколько столбцов типа Y выделено в таблице.

Добавить новые кривые на график к уже имеющимся кривым в координатных осях слоя можно, воспользовавшись при активном слое к команде главного меню **Graph/Add Plot to Layer/** (График/Добавить зависимость к слою) и выбрав тип графика. В открывшемся диалоговом окне

**Select Columns for Plotting** (Выбрать колонки для построения) выделяем имя колонки данных для построения новой кривой. Если таблиц в проекте несколько, то предварительно в выпадающем списке **Worksheet** (Таблица) диалогового окна выбираем необходимую. Затем переносим выделенное в левом поле диалогового окна имя столбца в одно из полей **Column** (Колонка) справа, щелкая мышью по соответствующей кнопке с двусторонней стрелкой. Выбор кнопки соответствует желаемой роли данных, представленных в колонке (аргумент это или функция). Для добавления выбранной зависимости к слою щелкаем по кнопкам **Add** (Добавить) и **OK**. Диалоговое окно **Select Columns for Plotting** (Выбрать колонки для построения) можно открыть и из контекстного меню, щелкнув правой кнопкой мыши по пиктограмме слоя (серый квадратик с цифрой, соответствующей номеру слоя в левом верхнем углу графического окна) и выбрав пункт меню **Plot Associations...** (Связи графика).


Еще проще добавить зависимость на график слоя методом символического перетаскивания мышью с нажатой левой кнопкой соответствующего набора данных из таблицы в поле графика. Предварительно должен быть выделен нужный столбец типа Y, а также выбран такое положение курсора мыши на выделении, при котором курсор приобретает вид .


Для построения зависимости (в том числе и методом перетаскивания) можно использовать только часть внесенных в таблицу числовых значений, размещенных в смежных ячейках. Для этого, прежде чем выполнять построение, выделяем необходимый диапазон ячеек (или диапазон строк, если используются все столбцы). Иногда удобно скрыть в таблице значения, размещенные в неиспользуемых для построения кривой ячейках. С этой целью в контекстном меню, появляющемся при щелчке правой кнопкой мыши по ячейке таблицы с минимальным выбранным для построения зависимости номером строки, необходимо выбрать команду **Set as Begin** (Установить как начальное), а затем на ячейке с максимальным номером **Set as End** (Установить как конечное). Вернуть скрытые значения на место можно, воспользовавшись главным меню: **Edit/Reset to Full Range** (Редактирование/Восстановить полный диапазон). При этом столбец со скрытыми данными должен быть предварительно выделен.

Управлять свойствами графической страницы, слоя, деталями графического представления зависимостей, а также добавлять к слою новые кривые удобно через контекстное меню, открывающееся при щелчке правой кнопкой мыши по пиктограмме слоя. Выбор пункта **Layer Properties...** (Свойства слоя) контекстного меню приводит к открытию уже знакомого нам диалогового окна **Plot Details** (Детали графика) (см. рис. 3.4) с набором вкладок для редактирования свойств графической страницы. В левой части диалогового окна располагается древообразное иерархическое меню **Graph/Layer/Plot** (График/Слой/Зависимость). При выделении разных элементов этого меню в правой части окна **Plot Details** меняется набор вкладок, позволяющих изменять свойства графических окон, слоев и кривых. Диалоговое окно **Plot Details** можно вызвать также из контекстного меню графического окна (щелчком правой кнопкой мыши в любом свободном месте графического окна и выбором пункта **Properties...** (Свойства)) или из другого контекстного меню, открывающегося при щелчке правой кнопки мыши по полю графика (пункт **Plot Details...**). При этом окно открывается с набором вкладок, соответствующих объекту.

Для размещения в графическом окне еще одного графика в новых координатных осях необходимо создать новый слой (при этом создаются и новые координатные оси), что можно сделать, обратившись к контекстному меню графического окна (пункты **New Layer (Axes)** (Новый слой (Оси)) и **Add and Arrange Layers...** (Добавить и упорядочить слои)). Однако большие возможности появляются при использовании диалогового окна **Layer** (рис. 3.5). Это окно открывается при обращении к команде главного меню **Tools/Layer** (Инструменты/Слой) при открытом графическом окне.

При создании нового слоя необходимо определить, будут ли его координатные оси независимы от координатных осей уже имеющегося слоя или будут связаны с ними. Это устанавливается щелчком по одной из кнопок вкладки **Add** диалогового окна **Layer** (см. рис. 3.5, а):

 — создается слой с обычным образом расположенными независимыми осями;

 — ось X создаваемого слоя привязана к осям слоя, активного в момент щелчка по кнопке, и располагается сверху, как изображено на кнопке. При этом масштаб оси X создаваемого слоя не связан с масштабом

оси X исходного слоя. Ось Y общая. Размеры и положение добавленного слоя изменяются в соответствии с изменением размеров и положения исходного слоя, но не наоборот;

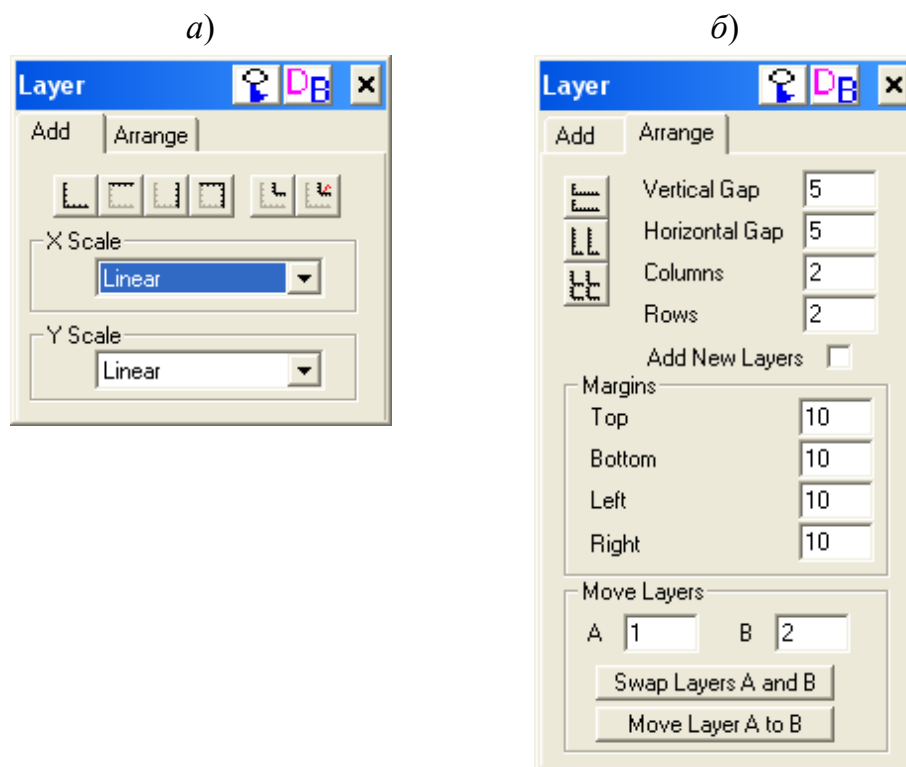








Рис. 3.5. Диалоговое окно Layer для добавления слоя в графическое окно: *a* — вкладка Add (Добавить); *б* — вкладка Arrange (расположить)

  — к осям слоя, активного в момент щелчка по кнопке, оказываются привязанными ось Y или обе оси X и Y соответственно;

 — создается слой с осями в четыре раза более короткими, чем оси слоя, к которому он привязан (так называемая врезка). Новый слой располагается в правом верхнем углу родительского слоя;

 — на создаваемой врезке воспроизводятся все зависимости родительского слоя, что позволяет легко построить на врезке выделенный фрагмент исходного графика в ином масштабе или представлении, например, в бóльшем масштабе, как бы под лупой. Для этого выбираем в панели инструментов **Tools** (Инструменты), располагающейся вдоль левого края главного окна Origin, инструмент **Zoom In** (Приблизить) . Затем на врезке, нажав левую кнопку мыши и удерживая ее, охватываем пунктирной рам-



кой необходимый участок графика. После отпускания кнопки мыши на врезке останется только этот участок, заполняющий весь график врезки. Вернуться к исходному масштабу можно, щелкнув по инструменту **Zoom Out** (Отодвинуть) .

На вкладке **Arrange** (Расположить) диалогового окна **Layer** (см. рис. 3.5, б) указывают, сколько будет слоев в графическом окне и как они будут расположены.

### 3.4. Форматирование графика

Для изменения свойств кривой на графике (цвет, толщина и тип линии, вид значков, обозначающих точки, по которым строится кривая, и т.д.) необходимо вызвать уже известное нам диалоговое окно **Plot Details** (см. рис. 3.4) и воспользоваться его вкладками, соответствующими редактируемой зависимости. Выбор редактируемой кривой, как было отмечено в предыдущем разделе, осуществляется в иерархическом меню, расположенном в левой части диалогового окна. Раскрыть сразу нужный набор вкладок окна можно, выполнив двойной щелчок мышью по линии графика. Полный набор вкладок окна для редактирования кривой — **Line** (Линия), **Symbol** (Значок), **Drop Lines** (Линии проекции), **Group** (Группа), **Error Bar** («Усы», или Планка погрешностей). Часть вкладок может отсутствовать, что определяется типом редактируемого графика. Предоставляемые для редактирования возможности понятны и не требуют дополнительных пояснений, особенно учитывая возможность визуально контролировать результаты проводимых на графике изменений, не закрывая диалоговое окно **Plot Details** и оставляя эти изменения обратимыми. Чтобы заданные изменения параметров отобразились на графике, достаточно щелкнуть по кнопке **Apply** (Применить). При этом окно **Plot Details** остается открытым. Лишь щелчок по кнопке **OK** закрывает диалоговое окно и делает внесенные изменения необратимыми.

Кроме кривых на графике необходимо иметь возможность форматировать оси координат, штрихи на осях (**Axis Ticks**), подписи к штрихам (**Axis Tick Labels**), координатную сетку (**Grid Lines**). Настройки параметров всех этих элементов доступны в диалоговом окне **X (Y) Axis - Layer**

(рис. 3.6). Вызвать его можно, воспользовавшись главным меню: **Format/Axis** (Форматирование/Ось) (или **Format/Axis Tick Label** (Форматирование/Подписи к штрихам)). Выбор редактируемой оси графика осуществляется в списке **Selection** (Выбор) в левой части данного диалогового окна. Это же окно можно открыть быстрее и сразу для выбранной оси, дважды щелкнув левой кнопкой мыши по оси графика или один раз по оси правой кнопкой мыши и выбрав в контекстном меню пункт **Properties...** (Свойства).

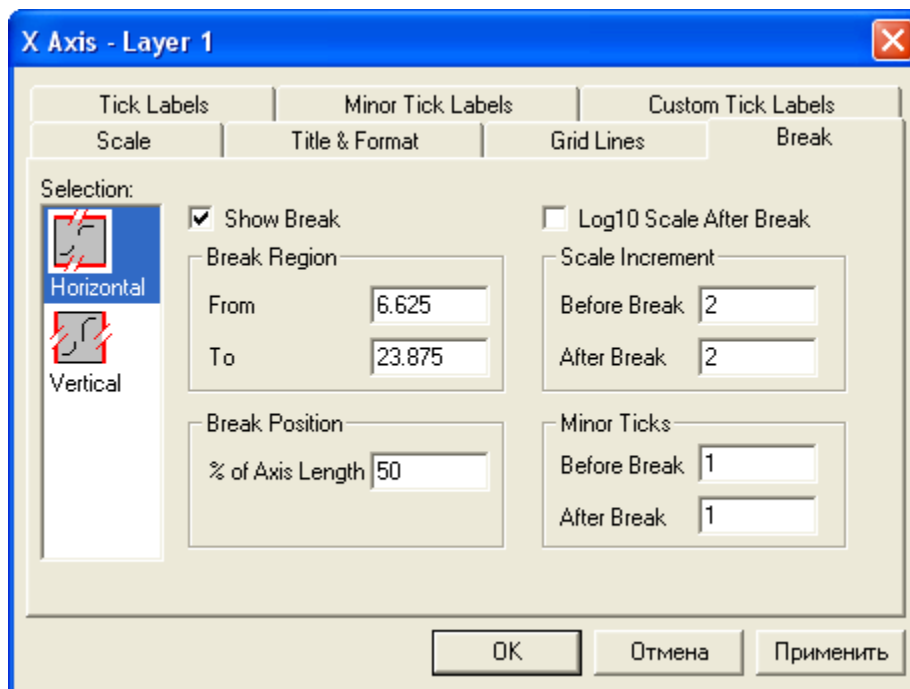


Рис. 3.6. Диалоговое окно форматирования графика, открытое на вкладке Break

В предоставляемых на семи вкладках данного диалогового окна возможностях редактирования свойств графика легко разобраться, поскольку задаваемые параметры сгруппированы по типу объектов, а все управляющие элементы вкладок снабжены подписями. Полное описание представленных элементов настройки можно быстро получить, обратившись к разделу **Help** (Справка) программы, соответствующему открытой вкладке. Для получения такой контекстной справки достаточно нажать клавишу **F1**.

Рассмотрим подробнее только вкладку **Break** (Разорвать) (см. рис. 3.6). Она дает возможность разорвать график по одной или обеим

осям неотображаемой областью значений аргумента или функции. Этой возможностью пользуются, например, если отображаемая на графике зависимость имеет особенности в двух удаленных друг от друга диапазонах изменения аргумента, а в промежуточной протяженной области изменяется слабо. Возможности вкладки **Break** становятся доступными, если галочкой пометить флажок **Show Break** (Показать разрыв). В списке **Selection** выбираем, для какой оси, горизонтальной или вертикальной, будем задавать параметры разрыва. В группе **Break Region** (Область разрыва) задаем диапазон значений аргумента или функции, попадающих в разрыв, в группе **Break Position** (Положение разрыва) — положение символов разрыва на оси в процентах от общей длины оси, в группе **Scale Increment** (Шаг шкалы) — шаг между основными штрихами на оси до и после разрыва. Отметив галочкой флажок **Log10 Scale After Break** (Десятичная логарифмическая шкала после разрыва), можно сделать масштаб оси после разрыва логарифмическим. Наконец, в группе **Minor Ticks** (Промежуточные штрихи) задаем, сколько дополнительных штрихов нужно проставить на оси между основными штрихами.

Подписи координатных осей (**X Axis Title** и **Y Axis Title** по умолчанию) можно редактировать непосредственно на графике. При двойном щелчке левой кнопкой на подписи в ее тексте появляется курсор в виде вертикальной черточки, и становится возможным редактирование. Изменение формата подписи осуществляется посредством кнопок панели **Format** (Форматирование), располагающейся во второй строке панели инструментов (см. рис. 3.1 и 3.7).

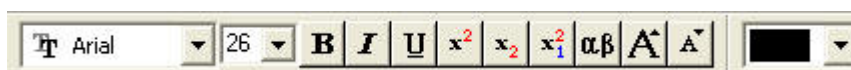


Рис. 3.7. Панель инструментов **Format** (Форматирование) для редактирования текстовых объектов

Те же и некоторые дополнительные возможности редактирования (например, поворот подписи на произвольный угол) дает диалоговое окно **Text Control** (Управление текстом). Открывается это окно при выборе пункта **Properties...** (Свойства) контекстного меню, появляющегося при щелчке правой кнопкой мыши на подписи.

Точно так же осуществляется редактирование любых надписей, помещаемых на график. Добавить надпись в выбранное место графического окна можно, либо воспользовавшись пунктом **Add Text...** (Добавить текст) контекстного меню, вызываемого в этом месте, либо щелкнув по кнопке **Text Tool** (Инструмент текст) **T** на панели инструментов **Tools** (Инструменты), а затем в месте будущего размещения надписи.

Формат подписей к штрихам на осях графика, в отличие от других надписей, задается на вкладке **Tick Labels** (Подписи штрихов) диалогового окна **X(Y) Axis - Layer** (X(Y) Ось слоя) (см. рис. 3.6). Окно открывается на вкладке **Tick Labels** при двойном щелчке на подписи к штрихам. Дополнительные операции форматирования подписей к штрихам доступны на вкладке **Custom Tick Labels** (Настройка подписей штрихов) того же окна.


В графическом окне помимо надписей можно разместить простые графические объекты: прямые, прямые со стрелкой, изогнутые кривые со стрелкой, прямоугольники, овалы, многоугольники, ломаные линии, а также линии произвольной формы. Для этого используется панель **Tools**, располагающаяся у левого края главного окна программы. Атрибуты графических объектов задаются в диалоговом окне **Object Properties** (Свойства объекта), которое открывается при выборе пункта **Properties...** контекстного меню на объекте. После выделения созданных объектов щелчком по ним мышью можно изменять размеры объектов, перетаскивая при нажатой кнопке мыши маркеры выделения (рис. 3.8, а).



Рис. 3.8. Маркеры редактирования объекта «изогнутая кривая со стрелкой», появляющиеся после одного (а), двух (б), трех (в) и четырех (г) щелчков мыши по объекту

Второй щелчок по объекту (именно второй, а не двойной) приводит к изменению вида маркеров выделения, и пользователь получает возможность вращать объект. Для некоторых объектов и последующие щелчки по ним приводят к изменению вида маркеров выделения (как, например, на рис. 3.8) и к появлению новых возможностей редактирования объекта.

### 3.5. Представление полученных результатов

Содержимое любого активного окна проекта Origin можно вывести на печать, нажав на кнопку **Print** (Печать)  панели инструментов. При этом печать начинается сразу с использованием текущих установок печати. Использование команды главного меню **File/Print...** (Файл/Печать) позволяет, прежде чем начать печать, проконтролировать и изменить установки печати в открывающемся диалоговом окне **Print** (Печать). Лишь после щелчка по кнопке **ОК** начинается вывод документа на принтер. Состав элементов диалогового окна зависит от типа активного окна (графическое окно, окно данных и т.д.).

При печати графического окна в выпадающем списке группы **Print Graph** (Печать графика) можно выбрать печать либо текущего окна, либо всех открытых графических окон, либо всех графических окон проекта. При этом содержимое каждого окна печатается на отдельном листе. Иногда удобно воспользоваться возможностью не выводить на печать часть элементов графического окна, не удаляя их, а просто скрыв. Для этого необходимо снять отметки у соответствующих элементов в меню **View/Show...** (Вид/Показать). Например, можно сделать невидимыми все слои, кроме активного, убрав галочку у пункта меню **View/Show/All Layers** (Вид/Показать/Все слои).

При распечатке таблицы данных в окне **Print** можно задать печать только выбранного диапазона строк и столбцов. При этом будут напечатаны также заголовки столбцов и номера строк.

Можно вывести на печать произвольный набор окон проекта разного типа, выделив соответствующие пиктограммы в левой части окна **Project Explorer** (Проводник проекта) Origin и выбрав в контекстном меню пункт **Print...** Окно Проводника располагается в нижней части главного окна Origin (см. рис. 3.1). Для выделения смежной группы окон в Проводнике нужно щелкнуть по первой пиктограмме группы, а затем по последней при нажатой клавише **Shift**. Для выделения несмежных пиктограмм необходимо удерживать нажатой клавишу **Ctrl**.

Origin дает возможность не только сохранить графическое окно в файле собственного формата, но и экспортировать его в файлы почти двух


десятков других графических форматов (**File/Export Page...** (Файл/Экспортировать страницу), как растровых, так и векторных. Для размещения графиков, созданных в Origin, в документах приложений Microsoft Office удобнее всего использовать векторные форматы Enhanced Metafile Format и Windows Metafile Format (расширения EMF и WMF соответственно), поддерживаемые этими приложениями. Редактировать рисунки таких форматов возможно средствами самих приложений Microsoft Office. Для публикации в печати рекомендуются форматы EPS, PDF, TIF, TGA. Причем более предпочтительны масштабируемые без потери деталей векторные форматы EPS, PDF. Для размещения графиков в Интернете используют экономные растровые форматы JPG, PNG, GIF.


Таблицы данных можно экспортировать в файлы ASCII форматов (**File/Export ASCII...** (Файл/Экспорт ASCII). Таблицы, экспортированные в текстовые форматы с расширениями DAT или TXT с разделителями между столбцами в виде знаков табуляции, корректно открываются в Microsoft Excel. В файл ASCII может быть экспортирована как вся таблица, так и только выделенная ее часть. В обоих случаях можно экспортировать в такие файлы и имена столбцов при установке флажка **Include Column Names** (Включить имена столбцов) в диалоговом окне, открываемом по команде записи в файл с заданным именем.

Origin предоставляет удобную возможность для подготовки презентаций из созданных в проекте графиков и таблиц. «Холст», на котором готовится презентация, — окно **Layout Page** (Страница компоновки) (рис. 3.9).

Для добавления в проект окна **Layout Page** обращаются к пункту меню **File/New** (Файл/Новый) и выбирают в открывшемся диалоговом окне строку **Layout** (Компоновка).

Для размещения в окне **Layout Page** графика в контекстном меню, открываемом при щелчке правой кнопкой мыши по белому полю окна, выбирают пункт **Add Graph...** . В появившемся окне **Select Graph Object** присутствует список всех графических окон проекта. После выделения од-

ного из них курсор приобретает вид . Щелчок таким курсором по полю окна **Layout Page** приводит к появлению на нем выбранного графика с размерами по умолчанию. Если, не отпуская кнопку мыши, протянуть этот курсор по полю окна, размер графика будет соответствовать возникшей

пунктирной рамке. Таким же образом помещаются в окно **Layout Page** таблицы **Origin** после выбора в контекстном меню команды **Add Worksheet...** (Добавить таблицу). При этом курсор приобретает вид .

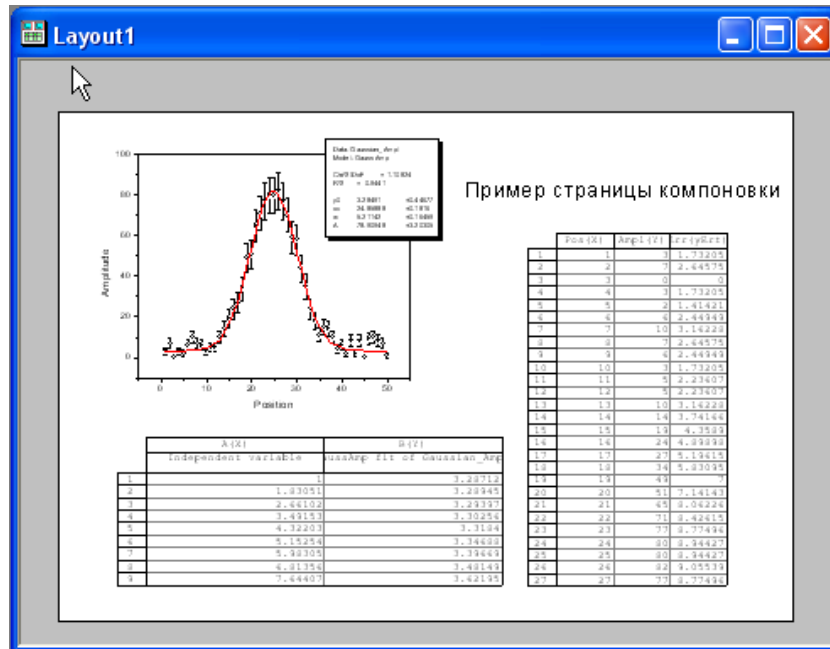


Рис. 3.9. Окно **Layout Page** с примером страницы, скомпонованной для презентации

Размеры и положение графиков и таблиц на странице компоновки изменяются стандартным образом после выделения соответствующего элемента. Кроме того, на страницу компоновки можно добавить текст, рисованные объекты **Origin** и поместить на нее рисунки, созданные в других программах.

Подготовленные таким образом листы презентации (см. рис. 3.9), как и рассмотренные ранее графики, можно сохранить в одном из доступных при экспорте из **Origin** графических форматов, обратившись к меню **File/Export Page...**

### **Контрольные вопросы**

1. Охарактеризуйте назначение элементов главного окна программы Origin.
2. Выполните ввод данных в таблицу, дайте или измените имена столбцов таблицы.

3. Как скрыть часть ячеек или строк в таблице? Как их восстановить?
4. Постройте график по данным таблицы, смените форму экспериментальных точек на графике.
5. Добавьте новые зависимости на график.
6. Сделайте вставку части основной кривой на графике в другом масштабе, в других координатных осях.
7. Отформатируйте график: измените вид точек, толщину и цвет линий, сделайте подписи к осям, введите разрыв оси графика.
8. Подготовьте страницу для презентации, включающую в себя таблицу, график, текст, символ организации.



#### 4. АППРОКСИМАЦИЯ ЗАВИСИМОСТЕЙ В MICROCAL ORIGIN

Любой график экспериментальной зависимости строится с учетом того, что экспериментальные точки получены с некоторой погрешностью. Плавная кривая (аппроксимирующая функция) не должна обязательно проходить через экспериментальные точки, а проводится так, чтобы определенным образом минимизировать совокупность отклонений экспериментальных точек от кривой.

Программа Microcal Origin позволяет осуществлять аппроксимацию экспериментальных данных функциями самого разного вида, как линейными по параметрам аппроксимации, так и нелинейными. Всего в Origin имеется 140 встроенных аппроксимирующих функций, объединенных в 12 групп. Наиболее часто используемые функции, кроме того, включены в группу **Origin Basic Functions** (Основные функции Origin). Также можно создавать пользовательские аппроксимирующие функции.

Аппроксимация линейными и нелинейными по параметрам функциями осуществляется в Origin с использованием разных алгоритмов. В обоих случаях для нахождения значений параметров аппроксимирующей функции минимизируется сумма квадратов отклонений экспериментальных точек от аппроксимирующей кривой, точнее, дисперсия остаточных погрешностей с учетом веса отдельных измерений, которая в руководстве к программе Origin называется  $\chi^2$  (хи-квадрат), а в уравнениях обозначена  $\chi^2$ :

$$\chi^2(a_1, a_2, \dots, a_k) = \frac{1}{N-k} \sum_{i=1}^N w_i (f(a_1, a_2, \dots, a_k, x_i) - y_i)^2 \rightarrow \min, \quad (4.1)$$

где  $y_i$  — измеренные значения величины  $Y$ ,  $w_i$  — вес  $i$ -го измерения,  $N$  — число измеренных пар значений  $(x_i, y_i)$ ,  $a_1, a_2, \dots, a_k$  — искомые параметры

аппроксимирующей функции,  $k$  — число неизвестных параметров,  $f(a_1, a_2, \dots, a_k, x)$  — аппроксимирующая функция.

При аппроксимации линейными по параметрам функциями для нахождения оптимальных параметров используется классический метод наименьших квадратов, то есть значения  $a_1, a_2, \dots, a_k$  получаются как результат решения системы линейных нормальных уравнений:



$$\frac{\partial \chi^2}{\partial a_j} = 0; \quad j = 1, 2, \dots, k. \quad (4.2)$$

Если параметры в аппроксимирующую функцию входят нелинейным образом, значения  $a_1, a_2, \dots, a_k$  подбираются с использованием алгоритма Левенберга–Марквардта (Levenberg–Marquardt), минимизирующего  $\chi^2$ . В этом случае перед запуском итерационной процедуры минимизации необходимо задать начальные значения искомых параметров. Следует отметить, что эффективность нелинейной аппроксимации может существенно зависеть от выбранных начальных значений параметров.


#### 4.1. Подготовительные операции

Прежде всего необходимо активировать график зависимости, которую нужно аппроксимировать. Это можно сделать, воспользовавшись пунктом главного меню **Data** (Данные). Список зависимостей, построенных на активном в данный момент слое, представлен в отдельном разделе, располагающемся ниже всех команд меню **Data**. Активная в данный момент зависимость помечена галочкой. Щелкнув мышью по имени аппроксимируемой зависимости, активируем ее, если она не была активной. Можно активировать выбранную зависимость, воспользовавшись контекстным меню. При этом не нужно разбираться, какая запись в меню **Data** соответствует выбранной кривой на графике. Просто щелкаем по аппроксимируемому графику правой кнопкой мыши и из контекстного меню выбираем пункт **Set as Active** (Сделать активным).

Можно аппроксимировать зависимость не во всем диапазоне представленных на графике значений аргумента, а только в заданном. Для выбора аппроксимируемого диапазона щелкаем по кнопке **Data Selector** (Се-

лектор данных)  на панели инструментов **Tools**. На обоих концах активного графика появляются маркеры данных в виде таких же двух стрелок, как на кнопке. Кроме того, становится активным инструмент той же панели **Data Display** (Отображение данных) . Для маркировки необходимого сегмента кривой щелкаем мышью по маркеру и, удерживая нажатой клавишу мыши, перетаскиваем маркер на один из концов задаваемого диапазона. При этом координаты соответствующей точки отображаются на черном фоне полоски **Data Display** (Дисплей данных), располагающейся по умолчанию в нижней части главного окна над **Project Explorer**. Для перемещения маркеров можно использовать также клавиши с правой ( $\rightarrow$ ) и левой ( $\leftarrow$ ) стрелками при нажатой клавише Ctrl. При каждом нажатии на клавиши  $\leftarrow\rightarrow$ ,  $\leftarrow\leftarrow$  маркер перемещается на следующую точку данных. Если удерживать нажатыми одновременно клавиши Ctrl и Shift, маркер будет сдвигаться сразу на пять точек. При использовании клавиш со стрелками численные данные в столбцах X должны быть отсортированы по возрастанию значений  $x_i$ . Чтобы выполнить сортировку данных во всех столбцах типа X по возрастанию, следует сделать активной соответствующую таблицу и дать команду **Analysis/Sort Worksheet/Ascending** (Анализ/Сортировка таблицы/По возрастанию). При этом данные в колонках типа Y перемещаются так, что соответствие  $(x_i, y_i)$  сохраняется.

В некоторых случаях, особенно для сложных зависимостей, удобно увеличить размеры красного крестика, выделяющего при маркировке активную точку данных на графике. Для ступенчатого циклического изменения размеров крестика достаточно последовательно нажимать клавишу пробела.




Чтобы зафиксировать выделенный для аппроксимации диапазон, выполняем одно из следующих действий: делаем двойной щелчок мышью внутри выделенного диапазона X, или нажимаем клавишу Enter или Esc, или щелкаем по кнопке **Pointer** (Указатель)  панели **Tools**. В результате маркеры по краям диапазона приобретают вид неперемещаемой двунаправленной стрелки. Для того чтобы скрыть на графике данные за пределами выделенного диапазона дайте команду главного меню **Data/Set Display Range** (Данные/Установить диапазон отображения). Ко-

манда **Data/Reset to Full Range** (Данные/Восстановить полный диапазон) снимает выделение. Те же команды доступны из контекстного меню, открывающегося при щелчке правой кнопкой мыши по кривой.


Можно исключить из анализа и отдельные экспериментальные точки. Это дает возможность проанализировать, как изменяются параметры аппроксимирующей функции при исключении точек, являющихся предположительно промахами. Исключение части точек из анализа в программе Origin называется маскировкой (masking) и осуществляется с помощью панели инструментов **Mask** (Маскировка) (рис. 4.1). Для отображения панели выбираем пункт меню **View/Toolbars** (Вид/Панели инструментов) и в списке панелей устанавливаем флажок у **Mask**.



Рис. 4.1. Панель инструментов Mask (Маскировка)

На появившейся после закрытия диалогового окна панели щелкаем по кнопке **Mask Point Toggle** (Переключатель маскировки точки) . Это активирует инструмент **Data Reader** (Считыватель данных), и курсор приобретает вид . Делаем двойной щелчок по маскируемой точке. Инструмент **Data Reader** закрывается, точка меняет цвет на красный и приобретает статус замаскированной. Чтобы демаскировать точку, снова щелкаем по кнопке , а затем делаем двойной щелчок по точке. Точка возвращается в число анализируемых.

Маскировку части значений для исключения их из анализа можно производить и в окне с таблицей данных. Чтобы отметить ячейку или группу ячеек в таблице как замаскированные, необходимо выделить их и щелкнуть по кнопке **Mask Range** (Замаскировать область значений) . Фон ячеек становится красным, и ячейки исключаются из дальнейшего анализа. Если по рассматриваемым данным уже построен график, то точки на графике, соответствующие замаскированным ячейкам, также окрашиваются в красный цвет. Если график строится заново по данным таблицы с замаскированными ячейками, соответствующие точки на графике будут отображены как замаскированные. Для того чтобы вернуть замаскированную ячейку

в число анализируемых данных, выделяем эту ячейку и нажимаем кнопку **Unmask Range** (Снять маскировку области) .

## 4.2. Аппроксимация линейными по параметрам функциями при обращении к меню **Analysis**

### *Аппроксимация прямой*

По команде **Analysis/Fit Linear** (Анализ/Аппроксимировать прямой) осуществляется аппроксимация активной зависимости с помощью прямой  $Y = A + BX$ , параметры которой  $A$  и  $B$  определяются методом наименьших квадратов. Если на графике представлены «усы» погрешностей, то при аппроксимации любыми зависимостями, в том числе и прямой, учитывается вес каждого результата. После расчета параметров создается новая (скрытая) таблица данных с именем по умолчанию **LinearFit1**, содержащая координаты  $(x_i, y_i)$  20-ти точек аппроксимирующей прямой. Построенная по этим точкам прямая, отображается в активном графическом окне. Параметры аппроксимации и результаты статистического анализа данных записываются в окно **Results Log** (Протоколирование результатов). Информационный блок, создаваемый в **Results Log**, начинается с ввода даты, времени и полного имени окна. Число в круглых скобках в конце первой строки — дата и время, представленные в формате Origin Julian Day Number. Затем записывается вид выполненного анализа и его результаты. В рассматриваемом случае в окне **Results Log** представлены: параметры прямой  $A$ ,  $B$  и их средние квадратичные отклонения;  $R$  — коэффициент корреляции между измеренными  $y_i$  и вычисленными по аппроксимирующей формуле  $(A + Bx_i)$  значениями  $y$ ;  $SD$  — стандартное отклонение аппроксимации;  $N$  — число экспериментальных точек, по которым строится аппроксимация;  $P$  — вероятность равенства коэффициента корреляции  $R$  нулю, то есть вероятность отсутствия корреляции между  $x_i$  и  $y_i$ . Коэффициент корреляции вычисляется по формуле

$$R = \sqrt{1 - \frac{\sum_{i=1}^N w_i [y_i - (A + Bx_i)]^2}{\sum_{i=1}^N w_i (y_i - \bar{y})^2}}, \quad w_i = \frac{1}{\sigma_i^2}, \quad (4.3)$$


где  $w_i$  — вес результата;  $\sigma_i$  — погрешность из соответствующей ячейки столбца типа  $yEr\pm$ .

Стандартное отклонение аппроксимации определяется выражением

$$SD = \sqrt{\frac{\sum_{i=1}^N w_i [y_i - (A + Bx_i)]^2}{N - 2}}. \quad (4.4)$$

Если значение погрешности не указано (отсутствует столбец типа  $yEr\pm$ ), то вес принимается равным единице ( $w_i = 1$ ).

Чем ближе  $R$  к единице и чем меньше  $SD$  и  $P$ , тем корректнее выполненная аппроксимация.

Когда окно **Results Log** открывается впервые, оно оказывается пристыкованным к нижнему краю главного окна программы. Его можно переместить в любое удобное место. Чтобы при перемещении избежать пристыковывания окна, необходимо удерживать нажатой клавишу **Ctrl**. Если окно результатов скрыто, для его появления в главном окне Origin достаточно щелкнуть по кнопке **Results Log**  панели инструментов.

### *Аппроксимация полиномом*

Для выполнения аппроксимации полиномом вида

$$Y = A + B_1X + B_2X^2 + B_3X^3 + \dots + B_kX^k$$

даем команду **Analysis/Fit Polynomial** (Анализ/Аппроксимировать полиномом). Открывается диалоговое окно **Polynomial Fit to Имя колонки данных** (рис. 4.2).

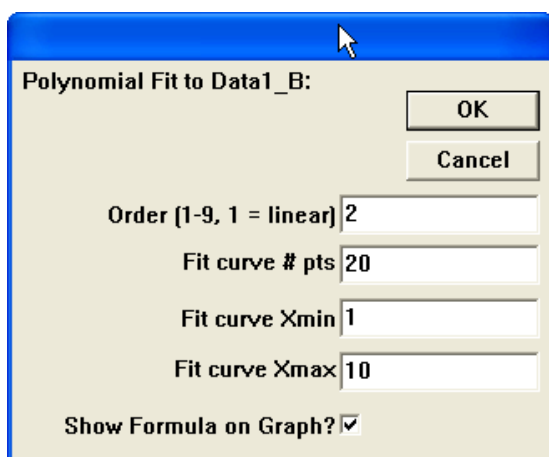


Рис. 4.2. Диалоговое окно, открываемое при аппроксимации полиномом через меню **Analysis**

В диалоговом окне указываем: порядок полинома (от 1 до 9); число точек, по которым будет строиться аппроксимирующая кривая; минимальное и максимальное значения  $X$  для аппроксимируемого интервала. Если установить флажок **Show Formula on Graph** (Показать формулу на графике), в поле графика будет выведено уравнение аппроксимирующей функции.

В окне **Results Log** выводятся: параметры  $A, B_1, B_2, \dots, B_k$  и их средние квадратичные отклонения; R-Square(COD) — квадрат коэффициента корреляции  $R$  (4.3) (называемый также коэффициентом детерминации (Coefficient of Determination));  $SD$  — стандартное отклонение аппроксимации (4.4);  $N$  — число точек данных, по которым строится аппроксимация;  $P$  — вероятность равенства  $R^2$  нулю.

#### 4.3. Аппроксимация линейными по параметрам функциями при обращении к меню **Tools**

Обращение к командам меню **Tools** (Инструменты) для аппроксимации данных с помощью прямой или полинома обеспечивает бóльшие возможности для управления процессом аппроксимации и позволяет вывести в окно **Results Log** более подробные статистические данные.

##### *Аппроксимация прямой*

Команда **Tools/Linear Fit** (Инструменты/Линейная аппроксимация) открывает инструмент **Linear Fit**, представляющий собой диалоговое окно, имеющее две вкладки: **Settings** (Установки) и **Operation** (Управление) (рис. 4.3). Понять назначение представленных на вкладках элементов помогают всплывающие подсказки, появляющиеся при задержке курсора мыши на поле ввода или другом элементе вкладок.

Рассмотрим вкладку **Settings** (Установки). В группе **Fit Curve Options** (Параметры аппроксимирующей кривой) вкладки **Settings** можно задать в поле **Points** (Точки) число точек, по которым строится аппроксимирующая кривая, в поле **Range Margin** (Протяженность края) — длину участков аппроксимирующей кривой, выходящих слева и справа за пределы диапазона заданных значений данных  $X$  (в процентах); в поле **Confidence** (Доверие)

— доверительную вероятность  $P$ , используемую при расчете погрешностей вычисляемых параметров (также в процентах).

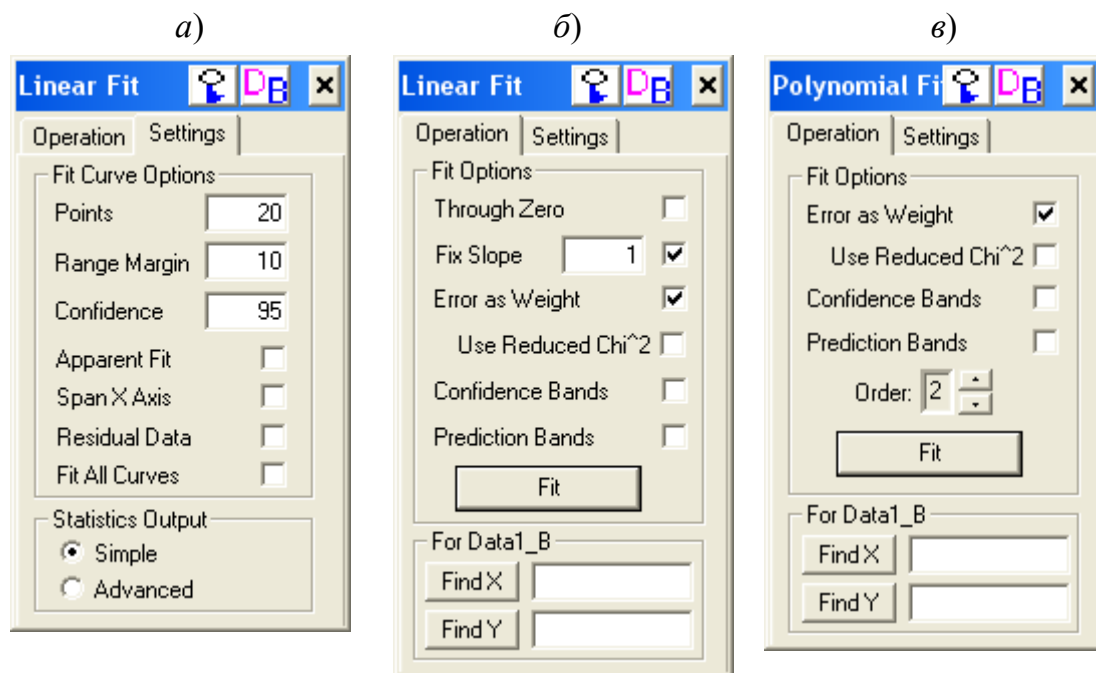


Рис. 4.3. Инструмент Linear Fit, открытый на вкладках: Settings (Установки) (а) и Operation (Управление) (б). Инструмент Polynomial Fit, открытый на вкладке Operation (Управление) (в)

Флажок **Apparent Fit** (Видимая аппроксимация) устанавливается в том случае, если для аппроксимации линейной зависимостью используются значения координат точек, преобразованные в соответствии с выбранным нелинейным масштабом осей из числа стандартных нелинейных масштабов для Origin, например, если выполняется аппроксимация прямой линией экспоненциально падающей зависимости, построенной в полулогарифмическом масштабе. Отметим также, что если флажок установлен и для данных указаны погрешности, Origin в качестве веса использует бóльшую по модулю ошибку.

Флажок **Span X Axis** (Диапазон оси X) устанавливается в том случае, если аппроксимация выполняется по части диапазона представленных на графике значений, а кривая аппроксимации должна быть экстраполирована на весь представленный диапазон.

Если установлен флажок **Residual Data** (Остатки), то после выполнения аппроксимации справа от колонки  $Y$  в таблице данных, по которым по-



строена зависимость, появляются две новые колонки: **Fit(Y)** — содержит значения  $Y$ , рассчитанные по аппроксимирующей формуле, **Residual(Y)** — содержит остатки, то есть  $\Delta Y = y_i - Y$ .

Если установлен флажок **Fit All Curves** (Аппроксимировать все кривые), то выполняется аппроксимация всех зависимостей, представленных в активном слое.

В группе **Statistics Output** (Вывод статистических данных) можно выбрать один из двух вариантов вывода сведений в окно **Results Log**. В варианте **Simple** (Простой) выводятся те же сведения, что и в случае использования для выполнения аппроксимации меню **Analysis**, а в варианте **Advanced** (Расширенный) — все статистические сведения о выполненной аппроксимации, включая доверительные интервалы по критерию Стьюдента (t-test) и таблицу результатов дисперсионного анализа (ANOVA Table, где ANOVA — сокращение от analysis of variance (дисперсионный анализ)).

На вкладке **Operation** (Управление) располагается группа флажков **Fit Options** (Параметры аппроксимации).

При установке флажка **Through Zero** (Через ноль) аппроксимирующая прямая будет проходить через начало координат.

При установке флажка **Fix Slope** (Фиксированный наклон) можно задать фиксированное значение углового коэффициента  $B$  прямой в связанном с флажком текстовом поле.

При установке флажка **Error as Weight** (Ошибка как вес) величина, обратная квадрату значения, представленного в колонке  $y_{\text{Err}\pm}$ , будет использоваться как вес результата. Для этого на графике должны быть построены «усы» погрешностей.

Когда установлен флажок **Error as Weight**, становится доступным флажок **Use Reduced Chi<sup>2</sup>** (Использовать редуцированную  $\chi^2$ ). Установка флажка **Use Reduced Chi<sup>2</sup>** влияет не на процесс подгонки, а только на выводимые значения погрешностей параметров. Когда данный флажок снят, погрешность параметров аппроксимации рассчитывается как стандартное отклонение. По умолчанию в качестве погрешности используется именно эта величина. При установленном флажке **Use Reduced Chi<sup>2</sup>** приводимая погрешность рассчитывается по формуле

$$\Delta a_i = \sqrt{\sigma(a_i) \sqrt{\frac{\chi^2(a_1, a_2, \dots, a_k)}{N - k}}}, \quad (4.5)$$

где  $\sigma(a_i)$  — стандартное отклонение для параметра аппроксимации  $a_i$ ;  $\chi^2(a_1, a_2, \dots, a_k)$  определяется формулой (4.1);  $(N - k)$  — число степеней свободы.

Для прямой, определяемой двумя параметрами,  $k = 2$ . Множитель, представленный корнем квадратным, не что иное, как значение  $SD$ , выводимое в окне **Results Log**.

При установленном флажке **Confidence Bands** (Доверительные полосы) на графике выводятся кривые, представляющие собой верхнюю и нижнюю доверительные границы для аппроксимирующей кривой. Они рассчитываются по формуле

$$\tilde{Y}(X) \pm t(P, \nu) s(\tilde{Y}(X)), \quad (4.6)$$

где  $\tilde{Y}(X)$  — значение  $Y$ , рассчитанное для заданного  $X$  по аппроксимирующей формуле;  $t(P, \nu)$  — коэффициент Стьюдента при доверительной вероятности  $P$  и числе степеней свободы,  $\nu = N - k$ ;

$$s^2(\tilde{Y}(X)) = S^2(\tilde{Y}(X)) \left[ \frac{1}{N} + \frac{X - \bar{X}}{\sum_{i=1}^N (X_i - \bar{X})^2} \right] \quad (4.7)$$

$(S^2(\tilde{Y}(X))) \equiv (SD)^2$  — оценка дисперсии  $\tilde{Y}(X)$  в соответствии с формулой (4.4).

Ширина доверительной полосы минимальна при  $X = \bar{X}$  и уширяется к краям интервала аппроксимации.

Если установлен флажок **Prediction Bands** (Прогнозируемая полоса), граничные кривые строятся по формуле

$$\tilde{Y}(X) \pm t(P, \nu) s(\text{pred}), \quad (4.8)$$

где

$$s(\text{pred}) = \sqrt{S^2(\tilde{Y}(X)) + s^2(\tilde{Y}(X))}. \quad (4.9)$$

Как следует из формул (4.7), (4.9), прогнозируемая полоса шире доверительной.

После задания всех параметров аппроксимации для построения аппроксимирующей кривой достаточно щелкнуть по кнопке **Fit** (Подгонка) вкладки **Operation** (Управление).

Можно выполнить аппроксимацию одновременно с построением графика, обратившись непосредственно к таблице данных. Выделяем в таблице столбец типа  $Y$  и, если необходимо, соответствующий ему столбец типа  $Y_{Err\pm}$ . Вызываем инструмент **Linear Fit**, задаем необходимые параметры аппроксимации и щелкаем по кнопке **Fit**. Будет построен график (линия аппроксимации), а параметры аппроксимации выведены в окно **Results Log**.

После выполнения аппроксимации становятся активными кнопки **Find X** (Найти  $X$ ) и **Find Y** (Найти  $Y$ ) на вкладке **Operation**. При вводе в ассоциированное текстовое поле значения  $X$  во втором текстовом поле после щелчка по кнопке **Find Y** выводится значение  $Y$ , рассчитанное по аппроксимирующей формуле (в наших обозначениях  $\tilde{Y}(X)$ ). Точно так же по заданному значению  $Y$  после щелчка по кнопке **Find X** выводится рассчитанное значение  $X$ .

### *Аппроксимация полиномом*

Для аппроксимации зависимости полиномом необходимо командой **Tools/ Polynomial Fit** открыть инструмент **Polynomial Fit**. Вкладка **Settings** этого инструмента аналогична соответствующей вкладке инструмента **Linear Fit** (см. рис. 4.2, *a*).

На вкладке **Operation** (рис. 4.2, *в*), помимо уже знакомых нам элементов, присутствует только счетчик, позволяющий выставить значение степени полинома от 1 до 9.

Все параметры аппроксимации, выводимые в окно **Results Log**, уже обсуждались при рассмотрении аппроксимации линейной зависимостью.

#### **4.4. Аппроксимация линейной по параметрам функцией произвольного вида**

Microcal Origin позволяет выполнить аппроксимацию экспериментальных данных зависимостью, представляющей собой линейную комбинацию заданных произвольных функций одной переменной:

$$y(x) = a + b_1 f_1(x) + b_2 f_2(x) + \dots + b_k f_k(x), \quad (4.10)$$

где  $a, b_1, b_2, \dots, b_k$  — постоянные, подлежащие определению.

Возможна также аппроксимация линейной функцией от нескольких независимых переменных  $x_1, x_2, \dots, x_k$ ;

$$y(x) = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k. \quad (4.11)$$

В Origin аппроксимация функцией (4.11) называется множественной регрессией (multiple regression). Для расчета параметров множественной регрессии создается особым образом организованная таблица данных. В первую колонку таблицы помещаются значения зависимой переменной  $y$ , а в последующие — значения независимых переменных  $x_1, x_2, \dots, x_k$ . Для выполнения аппроксимации выделяются колонки с независимыми переменными и выбирается команда меню **Statistics/Multiple Regression...** (Статистика/Множественная регрессия). Появляется информационное окно с сообщением о том, какие колонки представляют независимые переменные и какая — зависимую. Это диапазон выделенных колонок и первая колонка таблицы соответственно. После щелчка по кнопке **ОК** информационного окна выполняется расчет параметров аппроксимации, и результаты выводятся в окне **Results Log**. Выводятся значения параметров аппроксимирующей зависимости, их стандартные отклонения и другие результаты дисперсионного анализа данных. Значения параметра  $a$  и его стандартного отклонения представлены в строке **Y-Intercept** (отрезок, отсекаемый на оси  $Y$ ), значения коэффициентов  $b_i$  и их стандартных отклонений — в строках, имеющих имена соответствующих колонок таблицы.

Отметим, что даже для линейной по параметрам аппроксимации необходим графический контроль результатов. Чем больше в аппроксимирующей зависимости параметров и чем более разнородными функциями она образована, тем менее точным будет результат.

#### 4.5. Аппроксимация наиболее часто применяемыми нелинейными по параметрам функциями

Аппроксимацию проще всего выполнить нелинейными функциями, выделенными в качестве отдельных команд меню **Analysis**. Это наиболее

часто используемые на практике нелинейные аппроксимирующие функции.

При выборе **Analysis/Fit Exponential Decay/First Order** (Анализ/Аппроксимация экспоненциально затухающей функцией/Первый порядок) осуществляется аппроксимация функцией

$$y = y_0 + A_1 e^{-x/t_1}, \quad (4.12)$$

где  $y_0$  — сдвиг по оси  $Y$ ;  $A_1$  — амплитуда;  $t_1$  — постоянная затухания.

Как было отмечено ранее, при нелинейной аппроксимации перед началом итерационного процесса необходимо задать начальные значения искомых параметров. При аппроксимации функцией (4.12) через команду меню **Analysis** начальные значения параметров задаются программой автоматически, и пользователь не может вмешаться в этот процесс или зафиксировать какие-либо параметры.

После проведения аппроксимации в окне Results Log, помимо значений параметров аппроксимирующей функции и их стандартных отклонений, выводится значение  $\text{Chi}^2/\text{DoF}$ :

$$\text{Chi}^2/\text{DoF} \equiv \chi^2 / (N - k), \quad (4.13)$$

где  $\chi^2$  определяется (4.1);  $N$  — число точек на графике,  $k$  — число параметров зависимости (в нашем случае  $k = 3$ );  $(N - k)$  — число степеней свободы.

Выводится также значение  $R^2$  — квадрат коэффициента корреляции  $R$ , определяемого формулой (4.3).

**Analysis/Fit Exponential Decay/Second Order** (Анализ/Аппроксимация экспоненциально затухающей функцией/Второй порядок). Аппроксимация выполняется функцией

$$y = y_0 + A_1 e^{-x/t_1} + A_2 e^{-x/t_2}. \quad (4.14)$$

Автоматический подбор начальных значений сразу пяти параметров зависимости (4.14) не всегда возможен. В таком случае выводится сообщение о неудаче при подборе параметров. Это, скорее всего, свидетельствует о том, что для аппроксимации выбрана функция, не соответствующая характеру экспериментальной зависимости. Но если у вас есть основания считать, что данные все же описываются функцией вида (4.14), можно открыть сес-

сию нелинейной подгонки (см. раздел 4.7), чтобы попытаться подобрать параметры вручную. При этом процедура подгонки значительно облегчается, если известны хотя бы грубые предварительные оценки параметров.

**Analysis/Fit Exponential Decay/Third Order** (Анализ/Аппроксимация экспоненциально затухающей функцией/Третий порядок). Аппроксимация выполняется функцией

$$y = y_0 + A_1 e^{-x/t_1} + A_2 e^{-x/t_2} + A_3 e^{-x/t_3}. \quad (4.15)$$

Еще чаще, чем в предыдущем случае, может появиться сообщение о невозможности автоматического подбора начальных параметров. Параметры зависимостей (4.14) и (4.15) определяются методом итеративного приближения с приемлемой точностью только в том случае, если постоянные затухания процессов существенно (на порядки) различаются, а диапазон значений  $x$  достаточно широк.

**Analysis/Fit Exponential Growth** (Анализ/Аппроксимация экспоненциально растущей функцией). Аппроксимация выполняется функцией

$$y = y_0 + A_1 e^{x/t_1}, \quad (4.16)$$

где  $y_0$  — смещение по оси  $Y$ ;  $A_1$  — амплитуда;  $t_1$  — постоянная нарастания.

Аппроксимации функциями (4.12) – (4.16) можно проводить как при  $t > 0$ , так и при  $t < 0$ . Таким образом, с точки зрения использования для аппроксимации функции (4.12) и (4.16) эквивалентны. Однако, нужно иметь в виду, что при  $t > 0$  более корректно осуществляется автоматический подбор начальных значений параметров аппроксимации.

**Analysis/Fit Gaussian** (Анализ/Аппроксимация гауссианом). Аппроксимация осуществляется колоколообразной кривой (рис. 4.4, *a*), подобной кривой плотности нормального распределения:

$$y = y_0 + \frac{A}{w\sqrt{\frac{\pi}{2}}} e^{-\frac{2(x-x_0)^2}{w^2}}, \quad (4.17)$$

где  $y_0$  — смещение базовой линии;  $A$  — площадь под кривой, отсчитываемая от базовой линии;  $x_0$  — центр пика;  $w$  соответствует  $2\sigma$  в распределении Гаусса и равняется примерно 0,849 от ширины пика на половине его высоты.

Origin проводит аппроксимацию данных гауссианом и в том случае, если диапазон представленных значений не включает в себя максимум. Точность при этом, естественно, снижается.

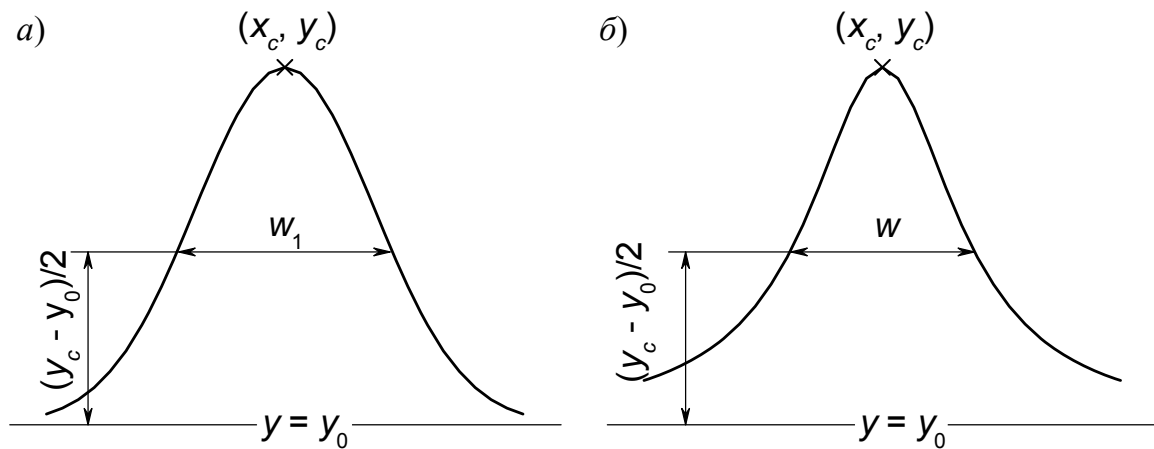


Рис. 4.4. Кривые Гаусса (а), Лоренца (б) и их параметры ( $x_c$  соответствует  $x_0$  в формулах (4.17) и (4.18));  $(x_c, y_c)$  — координаты максимума)

В окне **Results Log** выводятся значения:  $\text{Chi}^2/\text{DoF}$  (4.13),  $R^2$  (4.3), а также Area — площадь под кривой, равная параметру  $A$ ; Center — положение центра пика  $x_0$ ; Width (Ширина) — значение параметра  $w$ , при этом ширина на половине высоты пика, отсчитываемой от базовой линии  $y = y_0$ , равна  $w_1 = w\sqrt{\ln 4} \approx 1,177w$ ; Offset — смещение базовой линии  $y_0$ ; Height — высота пика, отсчитываемая от базовой линии  $(y_c - y_0)$ .

**Analysis/Fit Lorentzian** (Анализ/Аппроксимация лорентцианом). Аппроксимация осуществляется функцией

$$y = y_0 + \frac{2A}{\pi} \frac{w}{4(x - x_0)^2 + w^2} \quad (4.18)$$

Построенная по уравнению (4.18) кривая представлена на рис. 4.4, б. В окне **Results Log** выводятся те же значения, что и для гауссовой кривой.

Кривые Гаусса или Лоренца описывают, в частности, форму спектральных линий, вид которых определяется физической причиной их уширения.

С помощью Origin можно оценить параметры перекрывающихся пиков в результате аппроксимации кривой несколькими гауссианами или лоренцианами. Такая аппроксимация осуществляется по команде меню **Analysis/Fit Multi-peaks/** (Анализ/Аппроксимировать несколькими пиками).

Значительно более широкие возможности для подобного анализа дает **Peak Fitting Module** (Модуль аппроксимации пиков), который может быть установлен как расширение Origin.

**Analysis/Fit Sigmoidal** (Анализ/Аппроксимация сигмоидальной кривой). Аппроксимация осуществляется монотонно возрастающей или монотонно убывающей функцией либо с двумя горизонтальными асимптотами — при  $x \rightarrow -\infty$  и при  $x \rightarrow +\infty$  (функция Больцмана), либо с одной горизонтальной асимптотой — при  $x \rightarrow +\infty$  (логистическая функция).

Если для построения графика по оси  $x$  использована линейная шкала, то при аппроксимации по команде **Fit Sigmoidal** будет применена функция Больцмана

$$y = \frac{A_1 - A_2}{1 + e^{(x-x_0)/\Delta x}} + A_2, \quad (4.19)$$

где  $A_1$  — начальное значение  $y$ ,  $A_1 = y(-\infty)$ ;  $A_2$  — конечное значение  $y$ ,  $A_2 = y(+\infty)$ ;  $x_0$  — центр сигмоидальной кривой, определяемый как значение  $x$ , при котором  $y$  равен среднему между двумя предельными значениями ( $y(x_0) = (A_1 + A_2)/2$ );  $\Delta x$  — ширина перехода.

При использовании по оси  $x$  логарифмической шкалы при аппроксимации будет применена логистическая функция

$$y = \frac{A_1 - A_2}{1 + (x/x_0)^p} + A_2, \quad (4.20)$$

где  $A_1$  — начальное значение  $y$ ,  $A_1 = y(0)$ ;  $A_2$  — конечное значение  $y$ ,  $A_2 = y(+\infty)$ ;  $x_0$  — центр сигмоиды ( $y(x_0) = (A_1 + A_2)/2$ );  $p$  — показатель степени.

Чем больше  $p$ , тем резче переход.

Управление значениями параметров сигмоидальной кривой при аппроксимации осуществляется инструментом **Tools/Sigmoidal Fit** (рис. 4.5).

Часть элементов на вкладке **Settings** нам уже знакома. При выборе переключателя **Boltzmann** (Больцман) или **Dose Response** (Зависимость от дозы) аппроксимация зависимости при прологарифмированных по оси  $X$  значениях (например,  $10^7$  отложено как 7) будет выполняться соответственно функциями (4.19) или (4.20). Но в том случае, если по оси  $X$  исполь-



зован логарифмический тип шкалы, то независимо от положения переключателя аппроксимация осуществляется логистической функцией.

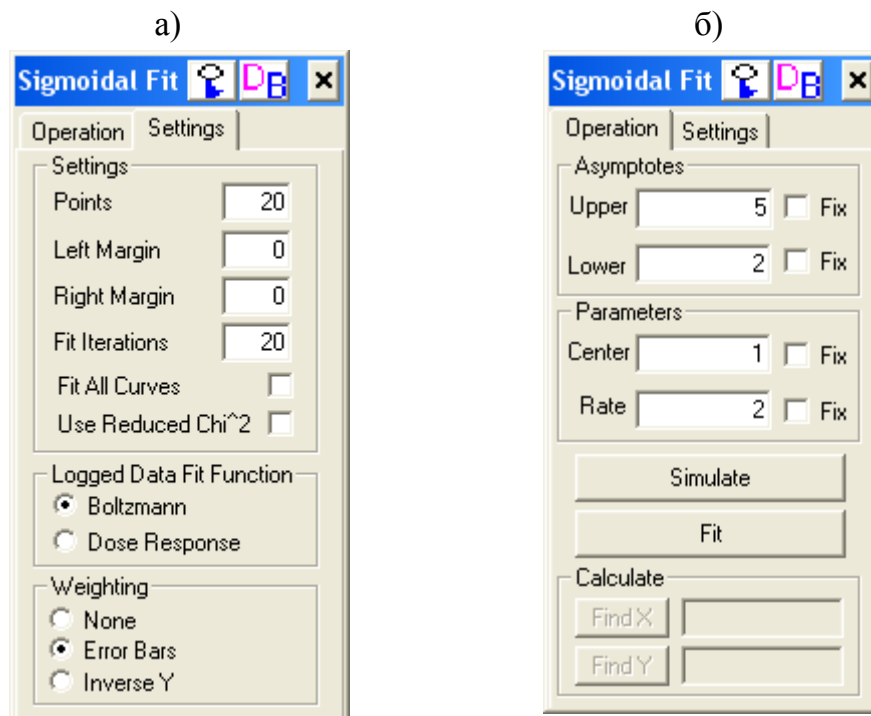


Рис. 4.5. Инструмент Sigmoidal Fit: *a* — вкладка Settings (Установки); *б* — вкладка Operation (Управление)

На вкладке **Operation** в текстовых полях **Upper/Lower** (Верхнее/Нижнее) задаем начальные значения  $A_2$  и  $A_1$  соответственно, в поле **Center** (Центр) —  $x_0$ , в поле **Rate** (Скорость) —  $\Delta x$  или  $p$ . Значение любого из этих параметров можно зафиксировать, установив флажок в соответствующем поле **Fix** (Зафиксировать). При этом отмеченные параметры сохраняют свое значение при выполнении итерационной процедуры подгонки. При щелчке по кнопке **Simulate** (Смоделировать) в активном графическом окне строится сигмоидальная кривая с параметрами, заданными на вкладке **Operation**. Это облегчает подбор начальных значений параметров

#### 4.6. Проверка статистической однородности двух групп данных

На практике часто одну и ту же величину измеряют в разных условиях, разными методами, получая несколько групп результатов наблюдений. К повышению точности результатов измерений может привести объедине-

ние этих групп данных для совместной обработки. Однако объединение результатов правомерно и приводит к повышению точности лишь при статистической однородности групп измерений, то есть в том случае, если представленные группы являются выборками из одной генеральной совокупности и различия данных не являются значимыми. Сопоставление групп данных — неременная процедура при их совместной обработке.

В Origin однородность двух групп данных устанавливается при помощи инструмента **Fit Comparison** (Сравнение аппроксимаций) (рис. 4.6), открывающегося по команде меню **Tools/Fit Comparison...**. Заключение об однородности или значимом различии групп данных дается на основе сравнения результатов аппроксимаций этих групп данных функцией одного и того же вида. Процедура сравнения состоит из нескольких последовательно выполняемых действий.

1. Из двух выпадающих списков в группе **Datasets** (Наборы данных) инструмента **Fit Comparison** выбираются сравниваемые колонки типа Y. В группе **Category and Function** (Категория и функция) — выбирается функция для аппроксимации выбранных данных. После щелчка по кнопке

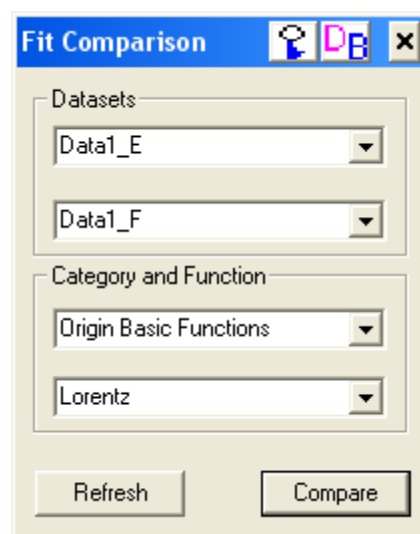


Рис. 4.6. Инструмент Fit Comparison (Сравнение аппроксимаций)

**Compare** (Сравнить) Origin выполняет по отдельности аппроксимацию каждой из двух выбранных групп данных указанной функцией. Затем осуществляется аппроксимация той же функцией всего набора данных, со-

ставленного объединением исходных групп данных. Для этих трех аппроксимаций рассчитываются:

$SSR$  — сумма квадратов отклонений от аппроксимирующей кривой:

$$SSR = \sum_{i=1}^N (f(a_1, a_2, \dots, a_k, x_i) - y_i)^2, \quad (4.21)$$

где  $f$  — аппроксимирующая функция,  $a_1, a_2, \dots, a_k$  — параметры аппроксимации,  $y_i$  — значения из таблицы данных,  $N$  — число значений в наборе данных;

$$DOF = N - k \quad (4.22)$$

( $DOF$  — число степеней свободы,  $k$  — число параметров аппроксимирующей функции).

Таким образом, получаются следующие величины:  $SSR1$ ,  $DOF1$ ,  $SSR2$ ,  $DOF2$  и  $SSRcombined$ ,  $DOFcombined$  — для первого набора данных, для второго и для объединенного набора соответственно.

2. Подсчитываются

$$SSRseparate = SSR1 + SSR2 \quad (4.23)$$

и

$$DOFseparate = DOF1 + DOF2. \quad (4.24)$$

3. На следующем шаге Origin рассчитывает значение  $F_{1,2}$  по формуле:

$$F_{1,2} = \frac{(SSRcombined - SSRseparate) / (DOFcombined - DOFseparate)}{SSRseparate / DOFseparate} \quad (4.25)$$

4. Вычисляется вероятность  $P$  того, что значение  $F_{1,2}$  для двух однородных групп данных может превысить рассчитанное. Случайная величина  $F_{1,2}$  имеет  $F$ -распределение плотности вероятностей, поэтому используется встроенная функция  $invf(value, m, n)$ , обратная интегральной функции  $F$ -распределения. Здесь  $value$  — значение  $F_{1,2}$ ,  $m, n$  — число степеней свободы для сравниваемых групп данных. Вероятность вычисляется так:

$$P = 1 - invf(F_{1,2}, (DOFcombined - DOFseparate), DOFseparate). \quad (4.26)$$

Если  $P > 0.05$ , в окне **Results Log** выводится сообщение о незначимости отличий двух групп данных с доверительной вероятностью 95 %. Заключение о значимости статистических отличий дается именно для самих сравниваемых групп данных, а не для параметров их аппроксимации.

#### 4.7. Расширенный контроль над процессом нелинейной аппроксимации

Возможность использования всех встроенных аппроксимирующих функций Origin при почти полном контроле над процессом аппроксимации обеспечивает сеанс нелинейной аппроксимации методом наименьших квадратов. Для открытия такого сеанса необходимо выбрать команду меню **Analysis/Non-linear Curve Fit/Advanced Fitting Tool...** (Аппроксимация нелинейной кривой/Расширенный инструмент аппроксимации) при активном графическом окне или окне данных. По команде меню открывается окно **NonLinear Curve Fitting** (Аппроксимация нелинейной кривой) в одном из двух представлений: базисном (рис. 4.7, *a*) или расширенном (рис. 4.7, *б*). Для перехода между этими типами представлений используются кнопки **More...** (Больше...) и **Basic Mode** (Базисная мода) в соответствующих окнах **NonLinear Curve Fitting**.

Базисная мода включает в себя пять диалоговых окон: **Select Function** (Выбор функции), **Define New Function** (Задание новой функции), **Edit Function** (Редактирование функции), **Select Dataset** (Выбор набора данных) и **Fitting Session** (Сеанс подгонки). Любое из этих окон открывается при щелчке по соответствующим кнопкам, располагающимся в активном диалоговом окне.

Рассмотрим аппроксимацию зависимости экспоненциальной функцией  $y = y_0 + A_1 \exp(-x/t_1)$  на примере проекта FITEXMP1.OPJ, размещенного в папке Tutorial программы Origin, при использовании режима базисной моды.

Открываем файл проекта и активизируем графическое окно с построенным в нем графиком. Выбираем команду меню **Analysis/Non-linear Curve Fit/Advanced Fitting Tool...**. Если при этом открывается окно **Select Function** в расширенном представлении (см. рис. 4.7, *б*), щелкаем по кнопке **Basic Mode** для перехода в режим базисной моды (см. рис. 4.7, *a*).

В списке функций в окне **Select Function** выбираем экспоненциальную функцию ExpDec1, щелкнув по ее имени. В этом окне представлены функции только той категории, которая использовалась в предыдущем сеансе аппроксимации. Для выбора другой категории функций следует временно перейти в расширенное представление окна подгонки, щелкнув по

кнопке **More...**. В окне просмотра, располагающемся правее списка функций, выводится либо уравнение функции, если выбран переключатель **Equation** (Уравнение), либо обобщенный вид кривой и характеристики некоторых ее особенностей (переключатель **Curve** (Кривая)).

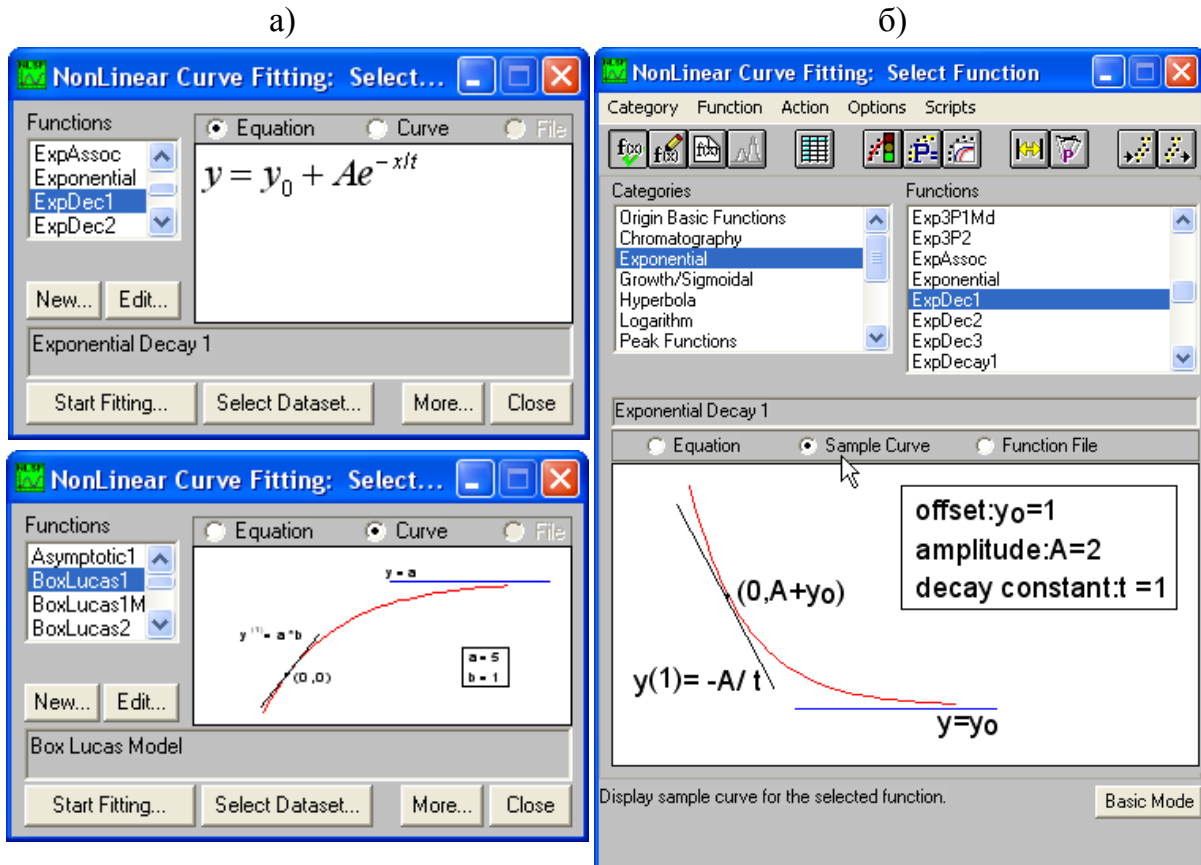


Рис. 4.7. Окно NonLinear Curve Fitting: Select Function (Аппроксимация нелинейной кривой: Выбор функции) в базисном (а) (верхнее окно — при включенном переключателе Equation (Уравнение), нижнее — при включенном переключателе Curve (Кривая)) и в расширенном (б) представлении

После выбора аппроксимирующей функции щелкаем по кнопке **Select Dataset...** (Выбор набора данных). Открывается одноименное окно для выбора колонки аппроксимируемых данных (рис. 4.8, а). В списке **Variables** (Переменные) выделяем строку **y Dep** — зависимая переменная  $y$ , а в списке **Available Datasets** (Доступные наборы данных), содержащем имена всех наборов данных проекта, — имя соответствующей колонки данных. После щелчка по кнопке **Assign** (Назначить) колонка данных закрепляется в качестве набора значений зависимой переменной. Аналогично за  $x$  **Indep**

закрепляется колонка значений независимой переменной  $x$ . В текстовых окнах справа и слева от надписи **<=Row<=** для зависимой переменной указывается диапазон строк, используемых для аппроксимации. Для независимой переменной можно задать диапазон либо строк для аппроксимации, либо значений переменной после щелчка по надписи **<=Row<=**, которая при переходе к независимой переменной превращается в кнопку **<=Row<= / <=X<=**. Подбор параметров аппроксимации будет осуществляться только по значениям переменных в указанном диапазоне, но при выводе на график кривая аппроксимации экстраполируется на весь диапазон значений  $x$ , представленных на графике. Для вывода аппроксимирующей кривой только в выбранном диапазоне аппроксимируемых данных нужно для выбора диапазона воспользоваться инструментом **Data Selector** (см. раздел 4.1). Используя окно **Fitting Step Size** (Величина шага подгонки), при выполнении аппроксимации можно пропустить часть точек. Например, если в окно введено число 3, для поиска параметров аппроксимации используется только каждая третья точка.

а)

б)

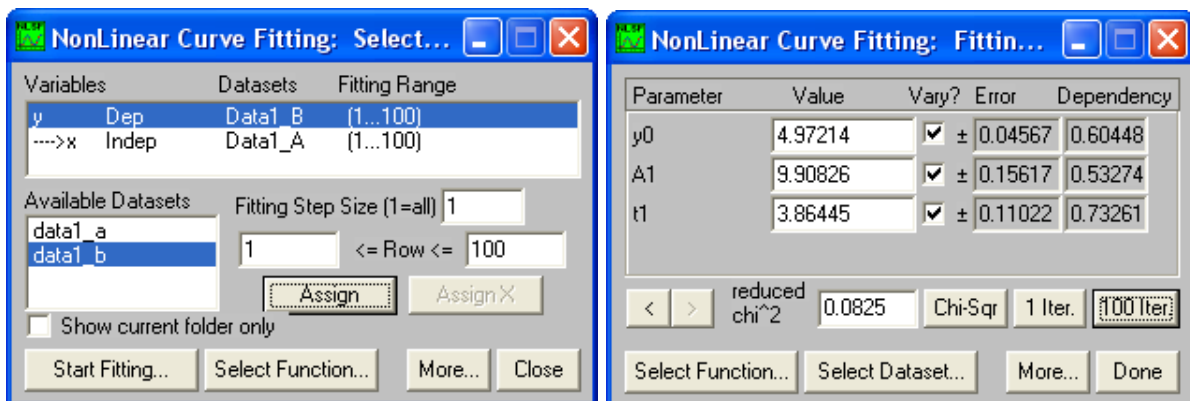


Рис. 4.8. Окна Select Dataset (Выбор набора данных) (а) и Fitting Session (Сессия подгонки) (б) в базисном представлении

Если после выбора аппроксимирующей функции сразу щелкнуть по кнопке **Start Fitting...** (Начать подгонку), появится информационное окно, предупреждающее о необходимости выбора аппроксимируемой колонки данных. После щелчка по кнопке **Dataset** этого окна открывается уже известное нам окно **Select Dataset**. Если в таблице данных предварительно была выделена нужная колонка  $Y$ , информационное окно примет другой

вид. В нем предлагается либо согласиться с аппроксимацией данных из выделенной колонки (кнопка **Active Dataset** (Активный набор данных)), либо выбрать другую колонку для аппроксимации, нажав на кнопку **Other Dataset** (Другой набор данных). В первом случае сразу открывается окно **Fitting Session** (Сессия подгонки), а во втором — окно **Select Dataset**.

Окно **Fitting Session** (рис. 4.8, б) включает в себя текстовые поля **Value** (Значение) со значениями параметров аппроксимирующей функции. Исходно в эти поля внесены начальные значения параметров, подобранные Origin автоматически. Значения параметров, с которых стартует итерационная процедура, могут быть в любое время изменены пользователем. Если снять флажок в ассоциированном поле **Vary?** (Изменять?), заданное значение параметра не будет меняться при итерациях. После завершения подгонки в поле **Error** (Ошибка) выводится значение стандартного отклонения для параметра, а в поле **Dependency** (Зависимость) — число, отражающее взаимозависимость параметров аппроксимирующей функции. Чем сильнее проявляется взаимозависимость параметров в процессе подгонки, тем ближе это число к единице. При избыточном числе параметров в функции между некоторыми из них имеет место жесткая связь, и соответствующие значения **Dependency** равны единице. В таком случае итерационный процесс не всегда сходится.

В поле **reduced chi<sup>2</sup>** выводится редуцированное значение хи-квадрат (4.13). При щелчке по кнопке **Chi-Sqr** выполняется расчет значения приведенного  $\chi^2$  при заданных в полях **Value** значениях параметров, и на графике строится кривая выбранной аппроксимирующей функции при этих параметрах. При каждом щелчке по кнопке **1 Iter** (Одна итерация) выполняется один итерационный цикл, выводятся новые значения параметров функции, величины  $\chi^2$ , и изменяется кривая на графике в соответствии с этими параметрами. Приближение аппроксимирующей кривой к графику исходной зависимости с каждой итерацией свидетельствует о том, что начальные значения параметров выбраны удачно. Для завершения итерационной сессии щелкаем по кнопке **100 Iter** (100 итераций), если необходимо, то несколько раз. Реальное число выполненных итерационных циклов для достижения заданной точности может быть и намного меньше 100 (табл. 4.1). По умолчанию точность определения редуцированного  $\chi^2$

равняется  $10^{-9}$ . После щелчка по кнопке **Done** (Готово) результирующая кривая аппроксимации выводится в графическое окно, а данные о параметрах аппроксимации — в окно **Results Log**.

На рис. 4.9 представлен результат аппроксимации экспоненциальной зависимостью данных из проекта FITEXMP1.OPJ.

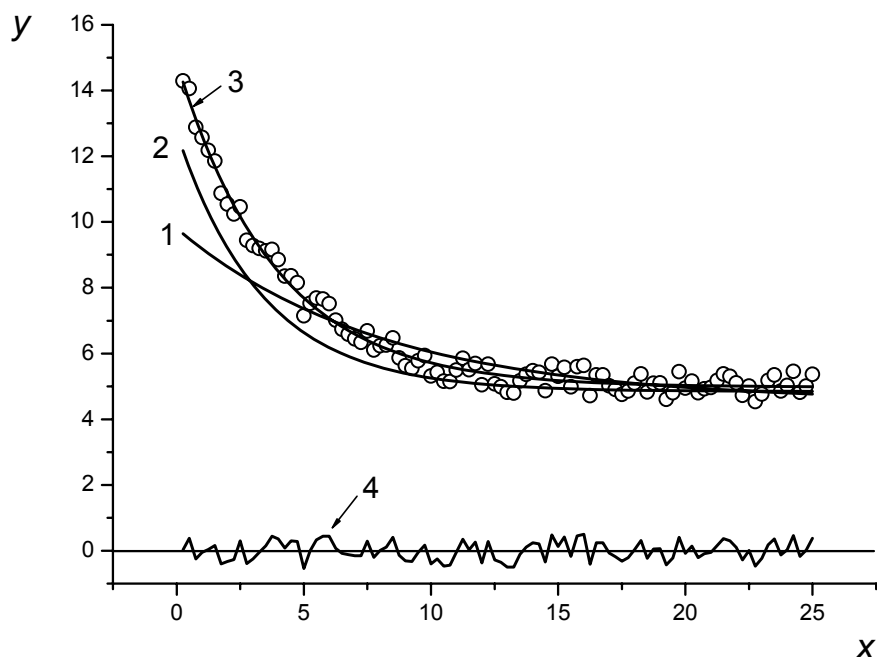


Рис. 4.9. Аппроксимирующая функция с автоматически подобранными начальными параметрами (1), после одного итерационного цикла (2), после шести циклов итераций ( $\Delta\chi^2 < 10^{-9}$ ) (3) и график остатков ( $\Delta y_i$ ) (4)


**Таблица 4.1**



Значения параметров аппроксимирующей экспоненциальной функции на разных этапах подгонки

Число итераций	Параметры функции $y = y_0 + A_1 \exp(-x/t_1)$ и их стандартные отклонения $\sigma$			$\chi^2/(n-3)$
	$y_0$	$A_1$	$t_1$	
0	$4.53320 \pm 0.04567$	$5.27587 \pm 0.15617$	$8.01678 \pm 0.11022$	1.25341
1	$4.84319 \pm 0.12586$	$7.88992 \pm 0.50372$	$3.37935 \pm 0.36816$	0.73592
6	$4.97214 \pm 0.04567$	$9.90826 \pm 0.15617$	$3.86445 \pm 0.11022$	0.08250



Сессию нелинейной аппроксимации можно начать, обратившись и непосредственно к таблице данных. Чтобы в последующем можно было пропустить шаг выбора аппроксимируемых данных в окне **Select Dataset**, нужно выделить колонку типа  $Y$  с аппроксимируемыми данными или, если будет использоваться только часть значений, — диапазон значений в данной колонке. При активной таблице данных выбираем команду меню **Analysis/Non-linear Curve Fit/Advanced Fitting Tool...**. Origin открывает графическое окно, строит в нем график зависимости и открывает сессию нелинейной аппроксимации.

При использовании расширенной моды (см. рис. 4.6, б) появляются дополнительные возможности управления процессом аппроксимации и выводом полученных результатов. Например, щелкнув по кнопке **100 Simplex Iter.** в окне расширенной моды **Fitting Session** (открывается при щелчке по кнопке  панели инструментов окна **Non-linear Curve Fitting**), можно выполнить аппроксимацию, используя алгоритм Симплекса, а не стандартный для Origin алгоритм Левенберга–Марквардта. В некоторых случаях алгоритм Левенберга–Марквардта не сходится, а с помощью алгоритма Симплекса удается найти параметры аппроксимации. Но следует помнить, что обычно алгоритм Левенберга–Марквардта работает эффективнее и требует при подгонке значительно меньшего числа итераций, чем алгоритм Симплекса.

Щелчок по кнопке **Param. Worksheet** (Таблица параметров) окна **Generate Results** (Генерировать результаты), которое открывается при щелчке по кнопке , приводит к созданию окна с таблицей, содержащей подробные сведения о результатах аппроксимации. После щелчков по кнопкам **Conf. Band** и **Pred. Band** на графике будут построены границы доверительной и прогнозируемой полос соответственно при указанных в текстовых полях вероятностях. Щелчок по кнопке **Residuals Plot** (График остатков), а затем по кнопке **Create Curve** (Создать кривую) в окне **Simulate Curves** (Моделировать кривые) () выводит в графическое окно график остатков (см. рис. 4.8).

Для реализации некоторых дополнительных возможностей при использовании расширенной моды необходимо знакомство со встроенными

языками Origin C или LabTalk. К расширенной моде приходится обращаться при решении специфических задач аппроксимации. В практике обработки результатов физического эксперимента, как правило, в ее использовании нет необходимости.

### ***Контрольные вопросы***

1. Для чего нужна аппроксимация экспериментальных зависимостей?
2. Как производится выбор функции для аппроксимации экспериментальных данных?
3. Выберите диапазон аппроксимируемых данных.
4. Каким образом исключаются сомнительные экспериментальные точки из общего набора данных?
5. Выполните аппроксимацию данных линейной функцией. Определите параметры полученной зависимости.
6. Выполните аппроксимацию данных полиномом. Как задать параметры аппроксимации полиномом?
7. Как выполняется аппроксимация данных нелинейными функциями?
8. Как выполняется итерационная процедура подгонки?
9. Отрадите погрешности на графике, постройте доверительную и прогнозируемую полосы.
10. Каким способом проверяют статистическую однородность двух групп данных?

## ЗАКЛЮЧЕНИЕ

В заключение сформулируем ряд выводов из изложенного материала и дадим некоторые практические советы начинающему экспериментатору.

При обработке результатов эксперимента возможны такие варианты соотношения эксперимента и теории:

1° Теоретическая зависимость известна и на основе экспериментальных данных определяются входящие в нее постоянные.

2° По экспериментальным данным подбирается формула, имеющая теоретическую интерпретацию.

3° Теория процесса не до конца ясна и подбираемая формула служит лишь для сглаживания или интерполяции результатов, а физического смысла пока не имеет. В этом случае аппроксимация обычно проводится полиномами по методу наименьших квадратов.

Обработку результатов измерений можно выполнить графически, аналитически и с помощью компьютерной программы. Компьютерная обработка результатов измерений значительно упрощает процедуру обработки, увеличивает возможности при обработке, снижает погрешность выполненных оценок.

Преимущества графической обработки:

- простота;
- наглядность (результат виден сразу);
- быстрая обработка почти любой закономерности.

Недостатки графической обработки:

- субъективность;
- зависимость от опыта экспериментатора;
- сложность оценки величины погрешности.

Аналитическая обработка уточняет результат и позволяет оценить погрешность более точно. Для графического определения параметров по измеренной зависимости ее спрямляют с помощью выбора системы координат ( $y-x$ ;  $\lg y-\lg x$ ;  $\lg y-1/x$  и т. д.). Строго говоря, прямую по полученным экспериментальным точкам нужно проводить по методу наименьших квадратов, однако эту вычислительную операцию при предварительной графической обработке обычно заменяют проведением прямой на глаз.

Опытный экспериментатор в состоянии провести такую прямую близко к вычисленной. При этом необходимо придерживаться следующих правил:

- 1) сумма положительных и отрицательных отклонений точек от прямой должна быть равна (зрительно) нулю;
- 2) не должно быть компенсации положительных отклонений, например, в начале прямой отрицательными в ее конце.

Оценка погрешностей эксперимента всегда должна быть объективной. «Вычисление» погрешности до третьего знака не имеет смысла. Даже очень тщательная математическая обработка не совсем качественных экспериментальных данных не даст хороших результатов и может привести к «открытию» сомнительных физических закономерностей. В таком случае улучшение условий эксперимента важнее самой изощренной обработки.

Погрешности экспериментальных данных могут быть систематическими и случайными. В первом случае погрешность остается постоянной или закономерно изменяется на протяжении всех измерений, во втором случае погрешность каждого измерения может оказаться как положительной, так и отрицательной.

Случайные погрешности устраняются путем статистической обработки результатов измерений.

Систематические погрешности в принципе опаснее случайных, если их не учитывать или не знать. Результат может выглядеть вполне надежным, хотя будет неверным. Систематические погрешности не вычисляются статистическими методами, их необходимо выявлять и устранять или учитывать. Общих рецептов нет.

Причины систематических погрешностей:

- 1) неточны измерительные приборы;
- 2) реальный эксперимент отличается от идеального (теоретического);
- 3) не совсем верна теория эффекта, что-то не учитывается.

Всегда необходимо уделять особое внимание тем измеряемым величинам, которые дают наибольший вклад в погрешность результата.

Ничто не может заменить практического опыта, хотя опыт без осмысливания — это медленный и трудный путь приобретения знаний.

Никто не застрахован от случайных ошибок.

Если эксперимент длится более одного—двух дней, то всегда следует обрабатывать результаты измерений в ходе эксперимента, поскольку может обнаружиться необходимость внесения изменений в измерения по ходу эксперимента.

Необходимо:

- сразу фиксировать все проделанное;
- записывать так, чтобы по истечении месяца, года и более было легко разобраться в записанном. Всегда ставить даты проведения измерений;
- все результаты записывать немедленно и без какой-либо обработки, даже без самых простых арифметических расчетов в уме. Не исправлять неверно записанный результат — лучше зачеркнуть и записать правильный;
- сохранять первичные данные измерений, так как может потребоваться иная интерпретация;
- оформлять результаты измерений в виде таблиц и по возможности сразу строить графики. Нанесение результатов на график позволяет контролировать ход эксперимента. Через экспериментальные точки следует проводить плавную кривую, а не соединять последовательно все точки. Схемы, таблицы и графики позволяют добиться большей ясности в проводимом эксперименте. Один рисунок лучше тысячи слов.

Итак, в путь. Счастливого плавания в эксперименте!

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

Богданов А.А. Визуализация данных в Microsoft Origin. М.:Альтекс-А, 2003. 104 с.

Грановский В.А., Сирая Т.Н. Методы обработки экспериментальных данных при измерениях. Л.: Энергоатомиздат. Ленингр. отд-ние, 1990. 288 с.

Зайдель, А.Н. Погрешности измерений физических величин. Л.: Наука, 1985. 112 с.

Котельников Р.Б. Анализ результатов наблюдений. М.: Энергоатомиздат, 1986. 144 с.

Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений. 2-е изд., перераб. и доп. Л.: Энергоатомиздат. Ленингр. отд-ние, 1991. 304 с.

Селиванов М.Н., Фридман А.Э., Кудряшова Ж.Ф. Качество измерений: Метрологическая справочная книга. Л.: Лениздат, 1987. 295 с.

Сергеев А.Г., Крохин В.В. Метрология: Учеб. пособие для вузов. М.: Логос, 2002. 408 с.

Тейлор Дж. Введение в теорию ошибок. Пер. с англ. М.: Мир, 1985. 272 с.

### Пример построения графика по экспериментальным данным и анализ полученной зависимости

При измерении вольт-амперной характеристики диода с  $p-n$ -переходом получены значения для электрического тока  $I$  через диод при разных значениях приложенного к диоду прямого напряжения  $V$ :

$V, В$	$I, А$	$V, В$	$I, А$	$V, В$	$I, А$	$V, В$	$I, А$
0	0	0.030	$2.19 \cdot 10^{-6}$	0.150	$3.29 \cdot 10^{-4}$	0.270	0.03420
0.005	$2.13 \cdot 10^{-7}$	0.040	$1.00 \cdot 10^{-5}$	0.175	$8.68 \cdot 10^{-4}$	0.280	0.05040
0.010	$4.72 \cdot 10^{-7}$	0.050	$5.91 \cdot 10^{-6}$	0.200	0.00228	0.285	0.04000
0.015	$7.86 \cdot 10^{-7}$	0.075	$1.72 \cdot 10^{-5}$	0.220	0.00495	0.290	0.07420
0.020	$1.17 \cdot 10^{-6}$	0.100	$4.68 \cdot 10^{-5}$	0.240	0.01070	0.295	0.09000
0.025	$1.63 \cdot 10^{-6}$	0.125	$1.25 \cdot 10^{-4}$	0.260	0.02320	0.300	0.10900

Для построения графика зависимости  $I$  от  $V$  в первую колонку таблицы Origin A(X), имеющую по умолчанию тип X и имя A, вводим значения  $V$ . Введенное в ячейку таблицы значение фиксируется при перемещении курсора в любую другую ячейку. Перемещение между ячейками осуществляется клавишами со стрелками, щелчком мышью на соответствующей ячейке или нажатием клавиши Enter. В последнем случае курсор перемещается в ячейку, располагающуюся ниже исходной. Во вторую колонку (типа Y) с именем B вводим значения  $I$ . Числа с порядком, например  $2.13 \cdot 10^{-7}$ , вводим в виде 2.13E-7. При этом в свойствах колонки должен быть указан тип вводимых данных **Text & Numeric** (Текст и число), который установлен по умолчанию, или **Numeric** (Число). Свойства колонки задаются в диалоговом окне **Worksheet Column Format** (Формат колонки таблицы), которое вызывается двойным щелчком по заголовку колонки.

Имея подготовленные таким образом исходные данные, уже можно построить график. Но прежде присвоим колонкам таблицы имена, соответствующие обозначениям измеряемых величин. Первой колонке — имя V, а второй I. Кроме того, выполним маркировку колонок: «V, В» и «I, А». Для этого вводим выбранные надписи в текстовые поля **Column Name** (Имя колонки) и **Column Label** (Маркировка колонки) соответственно диалого-

вого окна **Worksheet Column Format** выбранной колонки. Имя таблицы по умолчанию **Data1** заменим на **VAcurve**. Ввод нового имени производится в диалоговом окне **Rename** (Переименование), которое открывается из контекстного меню, появляющегося при щелчке правой кнопкой мыши по заголовку окна таблицы.

Для построения графика выделяем обе колонки щелчком мыши по заголовку одной из них с дальнейшим протаскиванием курсора мыши при нажатой левой кнопке на заголовок другой колонки. Даем команду **Plot/Line + Symbol** (График/Линия + Символ), выбрав соответствующий пункт главного или контекстного меню. Открывается графическое окно с построенным программой графиком — символы, соединенные прямыми линиями (согласно заданным по умолчанию в шаблоне Origin.otp параметрам). Изменим диапазон значений, откладываемых по осям графика. Двойным щелчком по оси *X* открываем диалоговое окно **X Axis - Layer 1** (Ось *X* — Слой 1) на вкладке **Scale** (Шкала) и в текстовом окне **From** (От) меняем значение  $-0.025$  на  $0$ . Выбираем в окне **Selection** (Выбор) вкладки пиктограмму **Vertical** (Вертикальная) и меняем начальное значение для оси *Y* также на  $0$ . Заменим прямоугольнички, которыми на графике обозначены экспериментальные точки, на другой символ. Двойным щелчком по кривой графика открываем диалоговое окно **Plot Details** (Детали графика) на вкладке **Symbol** (Символ) и в выпадающем наборе значков выбираем, например, закрашенный кружок.

Заменяем подписи у осей *X* и *Y* графика на «V, В» и «I, А» соответственно. Редактировать подписи можно либо прямо на месте, сделав двойной щелчок по прежней подписи, при этом в поле подписи появляется текстовый курсор, либо в окне **Text Control** (Управление текстом), выбрав в контекстном меню для подписи пункт **Properties...** (Свойства).

Теперь представим на врезке в основной график более детально начальную часть кривой, соответствующую малым токам. Для создания нового слоя, в котором будет размещен график-врезка, откроем диалоговое окно **Layer** (Слой) командой главного меню **Tools/Layer** (Инструменты/Слой). На вкладке **Add** (Добавить) этого окна щелкаем по крайней правой кнопке для создания второго слоя с размещенным в нем графиком таким же, как исходный, но меньшего размера. Выделяем его щелчком



мышью по одной из осей и перетаскиваем на свободную площадь исходного графика слева вверху. Немного увеличиваем размер графика-врезки перетаскивая угловой маркер выделения. Выставляем диапазон значений по оси  $X$  0–0.12, а по оси  $Y$  0– $5 \cdot 10^{-5}$ . Увеличиваем размер подписей значений у штрихов графика-врезки. Для этого щелкаем дважды по подписи и в открывшейся вкладке **Tick Labels** (Подписи штрихов) в поле **Point** (Пункт) задаем необходимый размер символов.

Чтобы сделать подписи штрихов оси  $Y$  более компактными, можно выбрать научный формат представления чисел. Выбор формата подписей осуществляется на вкладке **Tick Labels** окна **Y Axis - Layer 2**, открывающегося при двойном щелчке по оси  $Y$  графика-вкладки. Выбрав в выпадающем списке **Format** пункт **Scientific:1E3**, будем иметь, например, вместо 0.00002 подпись  $2 \cdot 10^{-5}$ . Еще удобнее оставить формат **Decimal:1000**, но в поле **Divide by Factor** (Разделить на коэффициент) ввести коэффициент  $1E-6$ . При этом подписи у штрихов будут соответствовать значениям величины тока, выраженным в микроамперах. Делаем соответствующие подписи у осей. Поскольку какие-либо подписи у них отсутствуют, необходимо ввести их в окошке **Title** вкладки **Title & Format**, а затем отредактировать для получения требуемого размера символов. Результат представлен на рис. П.1.

Выбранное нами представление данной зависимости не очень удачное. Как правило, если значения откладываемой по оси величины меняются на несколько порядков, следует выбирать по этой оси логарифмический масштаб, что легко осуществить в Origin. Снова строим график зависимости тока от приложенного напряжения и, дважды щелкнув по оси  $Y$  в выпадающем списке **Type** (Тип) вкладки **Scale** (Шкала), выбираем **Log10**. График будет перестроен в полулогарифмическом масштабе. Автоматически выбранный программой диапазон значений тока не включает в себя участок малых токов. Диапазон можно изменить, введя соответствующее значение в окошко **From** (От) вкладки **Scale**. Однако удобнее использовать команду главного меню **Graph/Rescale to Show All** (График/Изменить масштаб, чтобы показать все) для отображения всех данных из таблицы на графике, а затем при необходимости подправить границы диапазона (рис. П.2).

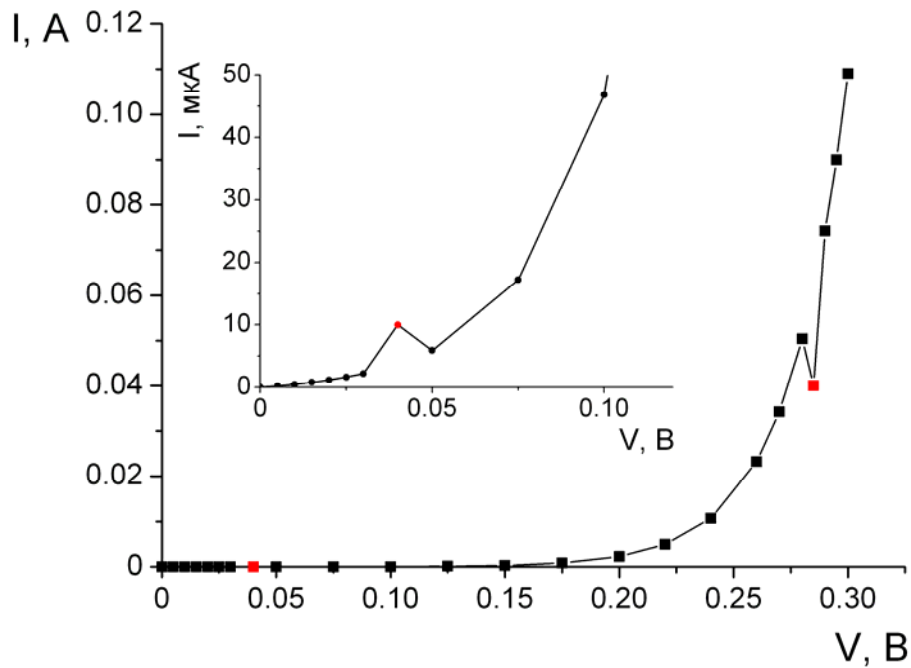


Рис. П.1. Вольт-амперная характеристика полупроводникового диода с врезкой для малых токов

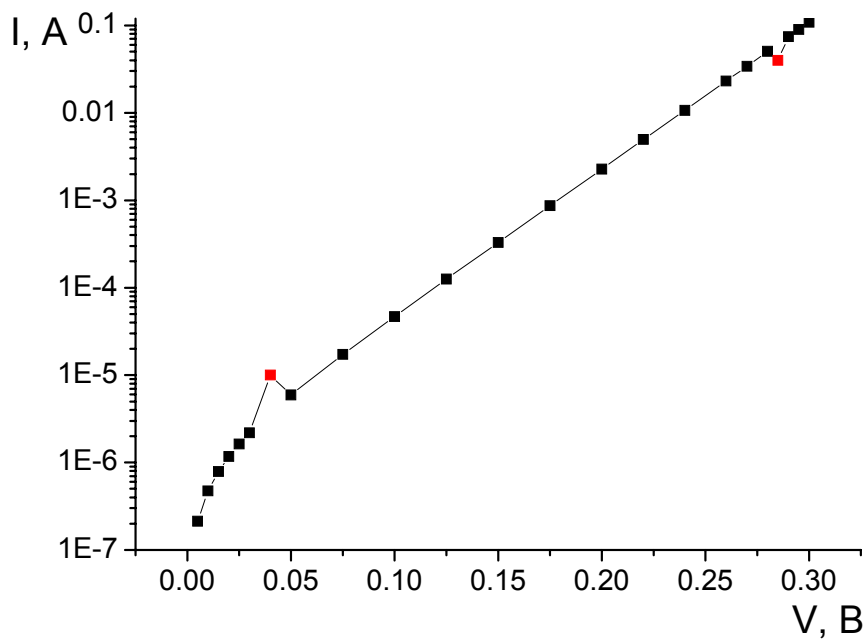


Рис. П.2. Вольт-амперная характеристика полупроводникового диода, построенная в полулогарифмическом масштабе

График, построенный в полулогарифмическом масштабе, показывает, что при напряжениях на диоде выше примерно 50 мВ зависимость тока от

напряжения примерно экспоненциальная, поскольку на этом участке график представляет собой прямую линию.

Подберем аппроксимацию для исследуемой зависимости. Известно, что статическая вольт-амперная характеристика идеального  $p$ - $n$ -перехода описывается формулой Шокли

$$I = I_s \left( e^{\frac{V}{V_T}} - 1 \right), \quad (\text{П.1})$$

где  $I_s$  — обратный ток насыщения диода,  $V_T$  — температурный потенциал.

Определим параметры зависимости (П.1) для рассматриваемых экспериментальных данных.



Строим график в линейном масштабе, на котором представлены только точки, отражающие экспериментальные данные. Для этого выделяем в таблице столбец **I(Y)** и выбираем команду **Plot/Scatter**.

Попробуем подобрать аппроксимирующую функцию вида (П.1) из стандартных аппроксимирующих функций Origin. Команда главного меню **Analysis/Non-linear Curve Fit/Advanced Fitting Tools...** (Анализ/Аппроксимация нелинейной кривой/Инструменты аппроксимации с расширенными возможностями) открывает диалоговое окно **NonLinear Curve Fitting: Select Function** (Аппроксимация нелинейной кривой: Выбор функции). Среди функций категории **Exponential** (Экспоненциальные) выбираем функцию **BoxLucas1** вида

$$y = a(1 - e^{-bx}). \quad (\text{П.2})$$

Функция (П.2) совпадает с зависимостью (П.1) при замене переменных  $I = y$ ,  $V = x$ ,  $I_s = -a$ ,  $1/V_T = -b$ . Для выполнения аппроксимации необходимо сообщить программе, в каких колонках таблицы данных располагаются значения независимой переменной и значения величины, функционально от нее зависящей ( $V$  и  $I$  соответственно). В меню диалогового окна выбираем команду **Action/Dataset** (Операция/Таблица данных). Вид окна изменяется. В верхнем поле выделена строка **y Dep**. В нашем случае зависимая переменная  $I$ , поэтому в поле **Available Datasets** (Доступные наборы данных) выбираем строку **VAcurve\_I** и щелкаем по кнопке **Assign** (Назначить). Сделанный нами выбор отражается в верхнем поле. Для определения параметров выбранной зависимости  $a$  и  $b$  даем команду меню **Action/Fit** (Операция/Подгонка). Содержимое диалогового окна меняется в

соответствии с заданным действием. В его полях **Value** (Значение) указаны начальные параметры аппроксимации, подобранные программой:  $a = 0.109$ ,  $b = 3.52441$ . Попробуем запустить итерационный процесс подгонки параметров зависимости (П.2). При щелчке по кнопке **1Iter.** (1 итерация) получаем на графике кривую, весьма далекую от экспериментальных точек. Обратите внимание, что значения в полях **Error** (Ошибка) значительно больше значений параметров  $a$  и  $b$  соответственно. Сто итераций, выполняемые при щелчке по кнопке **100 Iter.**, не улучшают аппроксимацию. Итерационный процесс не сходится. Это свидетельствует о том, что крайне неудачно подобраны начальные значения параметров зависимости  $a$  и  $b$ . Такой результат — следствие того, что алгоритм автоматического подбора начальных значений параметров построен в предположении, что  $a$  и  $b$  — положительные числа. В нашем случае  $a$  и  $b$  меньше нуля и  $a$  весьма мало. При необходимости параметр  $b$  легко оценить по наклону линейной части графика, построенного в полулогарифмическом масштабе. Но мы не будем этого делать, а в расчете на хорошую сходимость итерационного процесса введем в соответствующие поля более реалистичные, чем автоматически заданные, начальные значения  $a = 0$ ,  $b = -1$ . Действительно, уже после первого щелчка по кнопке **100 Iter.** аппроксимирующая кривая хорошо ложится на экспериментальные точки, а после второго щелчка перестают меняться значения параметров  $a$  и  $b$  и их погрешностей. Результат:  $a = -3.5017 \cdot 10^{-7} \pm 2.7527 \cdot 10^{-7}$ ,  $b = -42.12486 \pm 2.76912$  (конечно, в приведенных результатах много незначущих цифр, но это потому, что мы записали их в виде, выданном программой).

Качество аппроксимации можно улучшить, исключив из аппроксимируемых данных явные промахи, полученные для значений тока при  $V = 0.04$  В и  $V = 0.285$  В. Для исключения указанных точек из данных, по которым проводится аппроксимация, щелкаем по кнопке **Mask Point Toggle** (Переключатель маскировки точки) , курсор приобретает вид . Делаем двойной щелчок по маскируемой точке. Она окрашивается в красный цвет, что свидетельствует об ее исключении из анализа. Промах, соответствующий значению  $V = 0.285$  В, легко обнаружить и замаскировать на обрабатываемом графике (рис. П.3), а точку, соответствующую  $V = 0.04$  В,

гораздо удобнее замаскировать на врезке к графику на рис. П.1 или на рис. П.2. Обратите внимание на то, что выбранные точки оказываются помеченными, как замаскированные, на всех графиках, построенных по данным приведенным на с. 127. Соответствующие ячейки таблицы также помечены голубой подсветкой. Щелкаем по кнопке **100 Iter.**. Результат аппроксимации:  $a = -1.0085 \cdot 10^{-6} \pm 6.4363 \cdot 10^{-9}$ ,  $b = -38.63853 \pm 0.0224$ . Как видим, относительная погрешность результатов уменьшилась на много порядков.

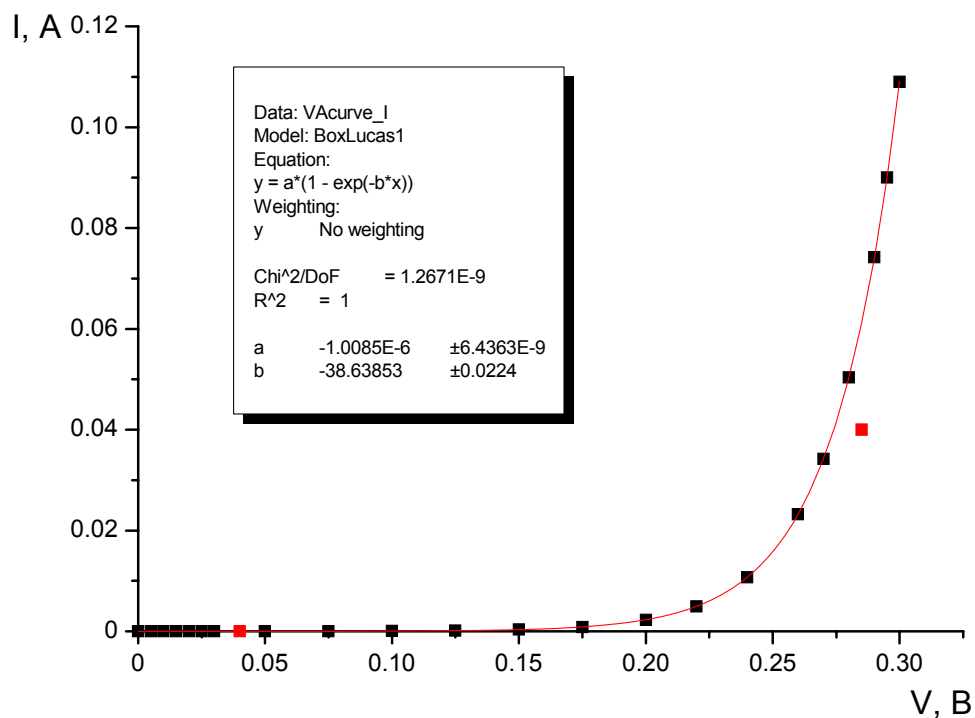


Рис. П.3. Аппроксимирующая кривая и ее параметры

При щелчке по кнопке **Done** (Готово) в поле графика выводится сводка данных по выполненной аппроксимации. Процесс аппроксимации закончен. Ее результаты представлены на рис. П.3.

Если в большом наборе аппроксимирующих функций Origin вы не обнаружите подходящую, можно задать собственную аппроксимирующую функцию. В качестве примера зададим функцию вида (П.1), используя расширенное представление моды нелинейной аппроксимации.

Открываем окно расширенной моды командой главного меню **Nonlinear Curve Fit/Advanced Fitting Tool** (Аппроксимация нелинейной кривой/Расширенный инструмент аппроксимации). Если откроется базовое

окно (см. рис. 4.7, а), щелкаем по кнопке **More...** . Создаем новую категорию функций, которую назовем **User**. Для этого выбираем в меню открывшегося окна **Nonlinear Curve Fitting** (Аппроксимация нелинейной кривой) (рис. П.4, а) команду **Category/New** (Категория/Новая) и вводим выбранное нами имя **User**. Для задания новой аппроксимирующей функ-

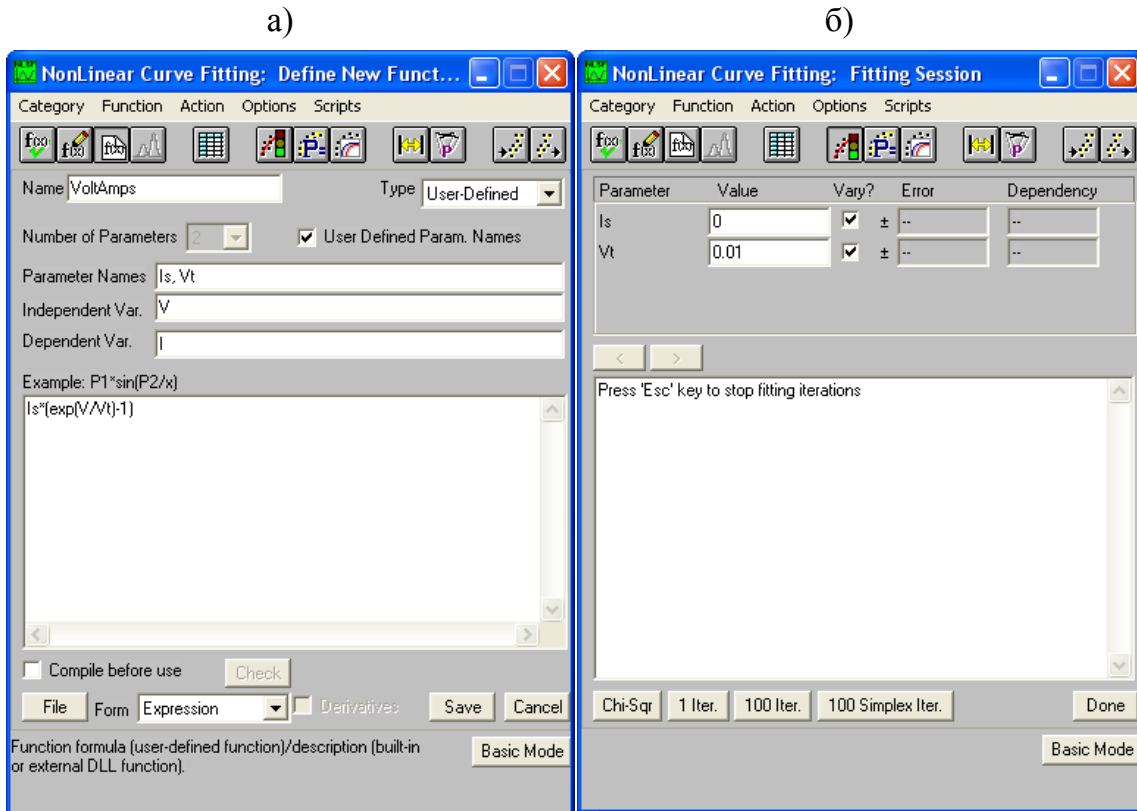




Рис. П.4. Окна Define New Function (Определение новой функции) (а) и Fitting Session (Сессия подгонки) (б) в расширенной моде

ции выбираем в меню окна **Nonlinear Curve Fitting** команду **Function/New** (Функция/Новая) или щелкаем по третьей слева кнопке панели инструментов окна  (**Define a new function** (Определение новой функции)). Открывается окно, представленное на рис. П.4, а. Вводим имя задаваемой функции, которую назовем **VoltAmps**, в поле **Name**. Аппроксимирующая функция (П.1) определяется двумя параметрами. Чтобы воспользоваться именами параметров по умолчанию **P1** и **P2**, достаточно в выпадающем списке **Number of Parameters** (Число параметров) выбрать значение 2. Присвоим собственные имена параметрам функции, для чего установим флажок **User Defined Param. Names** (Определенные пользователем имена

параметров). В поле **Parameter Names** (Имена параметров) вводим через запятую  $I_s, V_t$ , в поля **Independent Var.** (Независимая переменная) и **Dependent Var.** (Зависимая переменная) — соответственно  $V$  и  $I$ , в нижнее поле ввода — выражение  $I_s \cdot (\exp(V/V_t) - 1)$ . Щелкаем по кнопке **Save** (Сохранить) — новая пользовательская функция создана. В окне, открываемом после команды **Action/Dataset** (Операция/Набор данных) или щелчка по кнопке  (**Specify fitting dataset** (Укажите аппроксимируемые данные)), в качестве колонки независимых данных указываем  $V_{Acurve\_V}$ , а зависимых данных —  $V_{Acurve\_I}$ , как мы делали это ранее при аппроксимации встроенной функцией **BoxLucas1**.

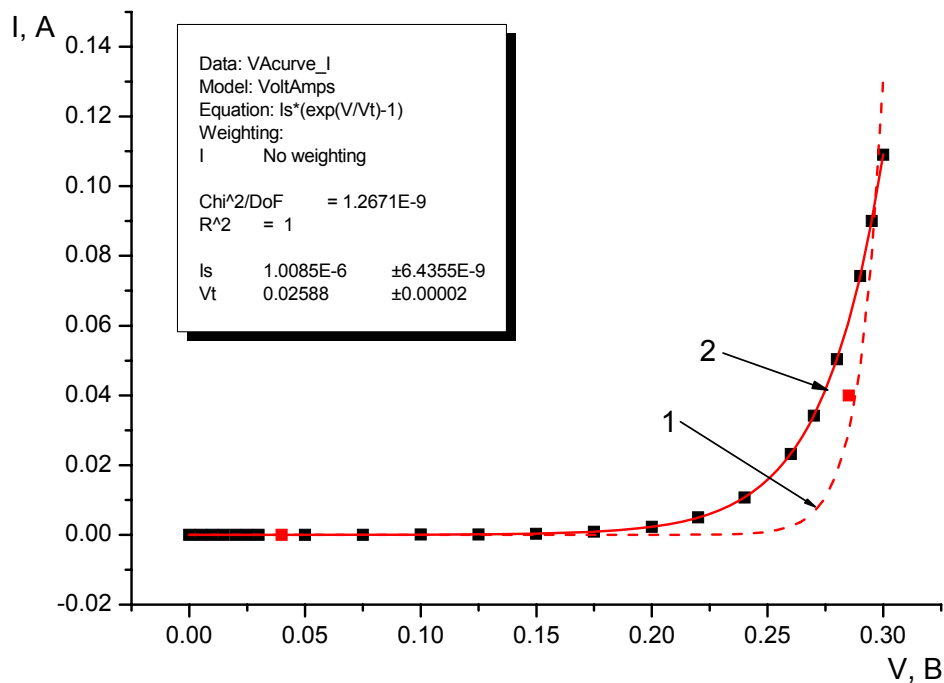



Рис. П.5. Результат аппроксимации данных пользовательской функцией после одной итерации (1) и после 200 итераций (2)

Команда **Action/Fit** (Операция/Подгонка) или щелчок по кнопке  (**Enter fitting session and perform nonlinear curve fitting** (Вход в сессию подгонки и выполнение нелинейной аппроксимации)) открывают окно, изображенное на рис. П.4, б. Начальные значения параметров  $I_s$  и  $V_t$  в полях ввода отсутствуют, поскольку мы не задали алгоритм вычисления начальных параметров. Это можно сделать с помощью скрипта, записанного на языке LabTalk или программы на языке Origin C. Вручную вводим

$I_s = 0$ ,  $V_t = 0.01$ . Щелкаем по кнопке **1 Iter.** (1 итерация), дважды по кнопке **100 Iter.** (то есть 200 итераций), а затем по кнопке **Done** (Готово). Результат и параметры аппроксимации представлены на рис. П.5.



## ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- А**
- Аппроксимация 57
- линейными по параметрам функциями 72, 101, 103, 107
  - нелинейными по параметрам функциями 60, 61, 64, 108
  - полиномом 72
  - прямой 59, 70, 103
- Арифметическое среднее 28
- В**
- Вероятность
- доверительная 34
  - плотность 14
- Вес результата измерения 32
- Взвешенное среднее 31
- Выборка 26
- Г**
- Гаусса распределение 19
- Гауссиан 110
- Генеральная совокупность 26
- Д**
- Дисперсия
- оценка 28
  - погрешности 20
  - случайной величины 18
- Доверительный интервал 33
- З**
- Значимости уровень 34
- И**
- Измерения
- косвенные 39
  - неравноточные 31
  - прямые 30
  - равноточные 30
- К**
- Квантиль 34
- Квантильный множитель 35
- Ковариация 40
- Композиция распределений 24
- Корреляции коэффициент 41
- Корреляционный момент 40
- Л**
- Лапласа функция 22
- Лоренца функция 111
- Лоренциан 111
- М**
- Максимального правдоподобия

принцип 27  
Математическое ожидание 17  
– оценка 28  
Момент распределения случайной  
величины 15  
– начальный 15  
– центральный 15  
Метод  
– моментов 27  
– наименьших квадратов 68

## Н

Наименьших квадратов метод 68  
Нормальное распределение 20  
Нормировки условие 14

## О

Отклонение  
– среднее квадратичное (СКО) 18  
– стандартное 18  
Оценка 26  
– дисперсии 28  
– математического ожидания 28  
– несмещенная 27  
– состоятельная 26  
– среднего квадратичного откло-  
нения 29  
– точечная статистическая 26  
– эффективная 27

## П

Погрешность 10  
– абсолютная 10  
– инструментальная 11  
– методическая 11  
– относительная 10  
– систематическая 11  
– случайная 12  
– суммирование 46  
– – арифметическое 48  
– – геометрическое 48  
Промах 37

## Р

Распределение 13  
– композиция 24  
– нормальное (гауссово) 20, 21  
– нормированное 22  
– одномодальное 43  
– равномерное (прямоугольное) 23  
– Стьюдента 35  
– трапецеидальное 24  
– треугольное 24

## С

Сигмоидальная кривая  
(сигмоида) 112  
Случайная величина 13  
– дискретная 15  
– непрерывная 14  
– нормированная 22

Среднее *17*  
– арифметическое *28*  
– взвешенное *31*  
– квадратичное отклонение *18*  
Статистическая однородность  
групп данных *31, 113*  
Стьюдента распределение *35*

## Ф

Функция  
– Больцмана *113*  
– Гаусса *19, 111*  
– Лапласа *23*  
– Лорентца *111*  
– распределения *14*  
– – дифференциальная (плотность  
вероятностей) *14*  
– – интегральная *14*

## Ц

Чебышева неравенство *36*

## Э

Экссесс распределения *30*

ПРОШИН Владимир Иванович  
СИДОРОВ Валерий Георгиевич

**МЕТРОЛОГИЯ, СТАНДАРТИЗАЦИЯ И СЕРТИФИКАЦИЯ:  
МЕТОДЫ ОБРАБОТКИ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ**

Учебное пособие

Редактор О.В. Махрова  
Технический редактор А.И. Колодянская  
Оригинал-макет подготовлен авторами

Директор Издательства Политехнического университета *А.В. Иванов*

Свод. темплан 2007 г.

Лицензия ЛР № 020593 от 07.08.97

Налоговая льгота — Общероссийский классификатор продукции  
ОК 005-93, т. 2; 95 3005 — учебная литература

---

Подписано в печать 01.09.07. Формат 60×84/16

Усл. печ. л. 8,75. Уч.-изд. л. 8,75. Тираж 100.

---

Санкт-Петербургский государственный политехнический университет.

Издательство Политехнического университета, член издательско-  
полиграфической ассоциации университетов России.

Адрес университета и издательства:

195251, Санкт-Петербург, Политехническая ул., 29.