

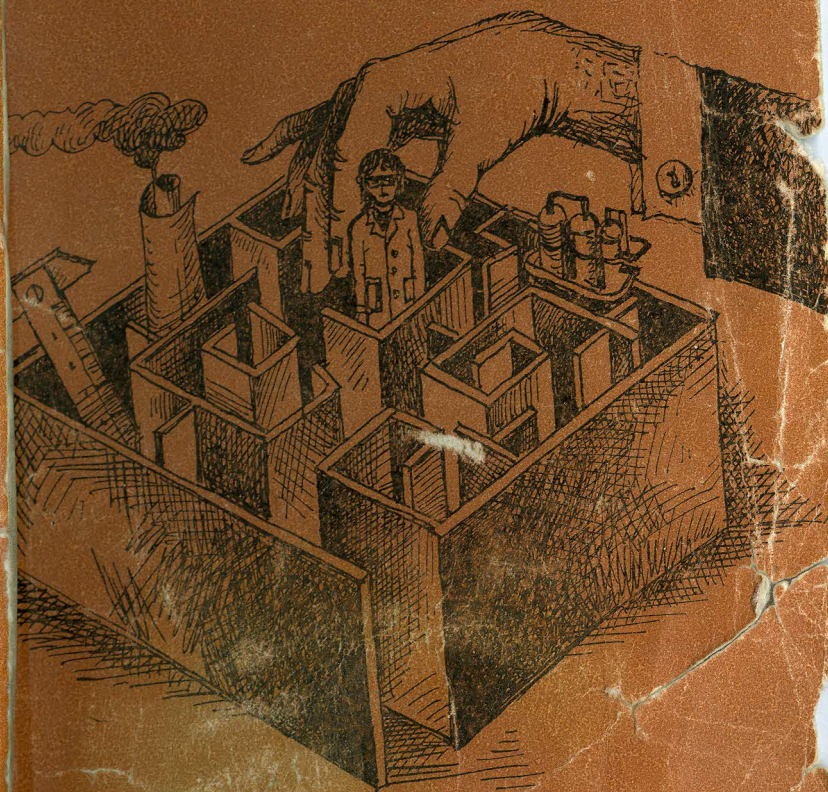
20 к.



«МЕТАЛЛУРГИЯ»



935
М. А. ШТРЕМЕЛЬ



ИНЖЕНЕР
В ЛАБОРАТОРИИ

V-1 669.01

Ш-935

М. А. ШТРЕМЕЛЬ

ИНЖЕНЕР В ЛАБОРАТОРИИ

(организация труда)

АБОНЕМЕНТ
УЧЕБНОЙ
ЛИТЕРАТУРЫ
ВЪЗ МАСХИ

Библиотека МИСиС



436138



МОСКВА • МЕТАЛЛУРГИЯ • 1983

~~142630000~~
12190000

3028
6200
3156
539
4800
1040
10044
13470
12233
16384

Уч. фонд
(и др.)

УДК 001.89

Рецензент докт. техн. наук С. З. Бокштейн

УДК.001.89

Инженер в лаборатории (организация труда). Штремель М. А.—
М.: Металлургия, 1983. 128 с.

Рассмотрена практика работы рядового инженера-исследователя в области металловедения и прикладной физики металлов. Проанализированы рациональные приемы работы, начиная от проработки технического задания, работы с литературой, выбора и создания методики исследования. Описано ведение первичной документации эксперимента, рациональный выбор методов обработки результатов, количественного анализа экспериментальных зависимостей и надежности выводов. Систематизированы существующие требования к содержанию и оформлению отчета о научно-исследовательской работе. Книга предназначена для широкого круга молодых инженеров черной и цветной металлургии, машиностроительной, химической, авиационной и других отраслей промышленности. Ил. 1. Табл. 1.
Библиогр. список: 56 назв.

Памяти
Георгия Павловича
БОГОВОГО,
учителя физики

ПРЕДИСЛОВИЕ

XXVI съезд КПСС поставил задачу «Поднять роль науки в создании материально-технической базы коммунизма». Эффективность научных исследований определяется прежде всего их организацией и квалификацией кадров, занятых в советской науке. Поэтому и в Советском Союзе, и за рубежом уделяется большое внимание организации исследований и разработок. В ряде недавно изданных монографий рассмотрено прогнозирование перспектив развития техники, планирование и руководство исследованиями, их эффективность в масштабах государства, отрасли, института. В обширной литературе 70-х годов по науковедению исследуется наука как социальное явление, как сфера занятости, как система потоков информации, изучается социология научного коллектива, психология научного творчества. Вместе с тем ни у нас, ни за рубежом нет почти никакой литературы по конкретной организации труда отдельного рядового инженера, занятого исследованием, хотя именно в качестве труда рядового работника — наибольший резерв роста эффективности исследований.

Исследованиями занимается все больше лабораторий заводов и институтов, но далеко не сразу и не везде утверждается высокая культура работы. Это ясно видно по разнице в уровне регистрируемых отчетов по НИР, поступающих в редакции, и даже многих публикуемых статей. Нередки работы, неэффективные по выбору метода, совершенно бездоказательные в изложении методики, в анализе воспроизводимости и надежности результатов, с дефектами логики в выводах.

Воспитать культуру труда, издавая наставления по его организации, вряд ли возможно. Большинство авторов хороших работ обходилось без них, но, несомненно, что-то было воспринято из опыта коллег. У давно работающего экспериментатора всегда есть система неписаных, но твердых правил работы. Для начинающего многие из этих правил выглядят мелочно придирчивыми, неоправданно жесткими и нудными. Однако, судя по унылому однообразию плохих работ, существует стандартный набор промахов. Учиться же на чужих ошибках

Ш 2601000000—010
040(01)—83 92—83

© Издательство «Металлургия», 1983

ках лучше, чем на собственных. Для коллективов новых и небольших лабораторий, где еще нет традиций, нет «внутренних стандартов» качества и организации работы, неизбежно — при всех его недостатках — заочное знакомство с полезным профессиональным опытом. Начинаящему инженеру-исследователю и адресовано это издание. В нем описаны некоторые правила и приемы работы, которые начинающему не всегда известны или не кажутся необходимыми, в то время как без их соблюдения результат не заслуживает доверия.

Поскольку по многим из затронутых вопросов почти не бывает публикаций, автор анализировал практику тех лабораторий заводов и институтов, где ему довелось работать или бывать, использовал собственный опыт рецензента научных отчетов и публикаций и руководителя небольших коллективов исследователей.

Все приведенные примеры относятся к деятельности инженера в лабораториях металловедения и физики металлов.

Автор будет считать свою цель достигнутой, если читатель найдет в книге хотя бы какую-то полезную для себя рекомендацию. Автор надеялся также привлечь внимание специалистов к научной организации и культуре труда такой массовой профессии, как инженер-исследователь в промышленности, и поэтому заранее признателен за возражения, вопросы и советы.

Автор благодарен С. З. Бокштейну и Л. З. Румшискому за ценные замечания, а также С. Н. Горбовой за помощь в работе над рукописью.

1. ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЯ

АНАЛИЗ ЗАДАНИЯ

Задание на исследование и разработку материала исходит в конечном счете от конструктора. Он формулирует перечень эксплуатационных и технологических характеристик (в виде цифр и пожеланий) — естественно, не полный и не всегда верный. Нередко предлагаемый материал, хотя и соответствует букве задания, но совершенно негоден по очевидным для конструктора причинам. Поэтому обстоятельный диалог с «лицом, принимающим решение» — конструктором — нужен в самом начале, до того, как задание станет директивным документом.

Как бы ни была сложна конструкция, материал разрабатывают для единственной детали. Необходимость разработки — предмет тщательного анализа. «Подобрать из старого или разработать новое» — первая дилемма для материалововеда. Для ответа нужно ясно видеть конструкцию, сроки, масштабы и перспективы производства. Является конструкция усовершенствованием или упрощением существующей, единственным прототипом нового класса изделий или лишь одним из ряда вариантов?

Каковы источники информации о недостатках в существующей конструкции: это надежная статистика отказов (с анализом повреждений в натуре или лишь в требованиях на запчасти) либо паника после одной аварии от непонятных причин? Понадобится тот же материал для следующего поколения изделий или все затраты на его разработку и освоение лягут на одну машину, а при просрочке погибнут? Определяет материал качество или технологичность всей конструкции, узла решающим образом или он обеспечивает второстепенные функции? Каковы конструктивные альтернативы: нельзя ли вообще обойтись без детали, для которой ищут материал? Каковы технологические альтернативы (если магний с вольфрамом трудно сварить, то почему трудно сделать разъемное соединение)?

Предполагается массовое, серийное или индивидуальное производство? Каково в перспективе годовое потребление материала, какой директивный срок начала выпу-

ска изделия и сколько лет его будут выпускать? Реально ли к этому сроку создать и освоить в производстве новый материал или процесс и оборудование для него либо надо подобрать старый и хорошо освоенный?

Если новый материал действительно необходим, уточним требования.

Каковы критерии эффективности конструкции и материала в ней? Что оптимизируется: абсолютное значение параметра (производительности, скорости и т. п.) или его относительное значение (на единицу веса, стоимости, трудоемкости)?

Каков критерий оптимального времени жизни конструкции? Это может быть: а) абсолютная долговечность каждой единицы (как у моста или нефтепровода); б) средняя «наработка на отказ» в партии изделий (как у электрической лампы); в) гарантированный ресурс (как у лопатки газовой турбины: задана малая вероятность отказа до срока для худшей детали в машине — безразлично, сколько проживет лучшая); г) цена ресурса (как у сверла или шины); д) гарантия однократного срабатывания (вероятность осечки для патрона).

Какой задан нормативный срок эксплуатации конструкции? Возможно ли профилактическое обслуживание, замена или ремонт детали? Какой расход запчастей приемлем? Есть ли кадры и дисциплина, гарантирующие обслуживание? Нормирована ли ремонтпригодность (во что обходится час простоя на ремонте и сколько нужно труда, чтобы разобрать до основания, заменить копеечную деталь, собрать все обратно и отрегулировать)? В любой ли момент неисправную машину можно остановить на ремонт (а если это самолет в воздухе)? Как часто нужны ревизии с разборкой (например, магистрального газопровода)? Какова надежность и стоимость диагностики: существуют ли объективные методы предсказания дальнейшей работоспособности (например, после дефектоскопии в процессе эксплуатации), не съедят ли проверки большую часть рабочего времени?

Из обоснованного таким образом ресурса детали нужно вывести требования к свойствам материала. Нужны подробности режимов эксплуатации. Как распределена во времени величина нагрузки и перерывы, какое число включений, каковы неизбежные перегрузки (например, при запуске)? Как меняются температура и

рабочая среда при эксплуатации? Какова агрессивность среды (механические или химические воздействия жидкости, твердых тел и частиц, газа и вакуума, адсорбция, облучение)? Какие требования добавляют транспортировка и хранение до эксплуатации и в промежутках (прочность прибора по дороге к потребителю важнее, чем при измерениях; зима простоя на улице портит комбайн не меньше, чем лето работы).

Как выглядят чертеж и расчетная схема? Какие нагрузки не учтены (монтажные, транспортные, термические, магнитные, вибрационные)? Какова сопрягаемая деталь и взаимодействие с ней? Какие допуски на размеры и на их стабильность при многолетней работе или хранении?

Какие нормы свойств материала записаны в чертеже? Все ли они необходимы? Какие свойства надо нормировать еще для гарантии работоспособности? Существуют ли стандартные методы контроля этих свойств или их также надо разработать?

Если материал и технология известны, то насколько «напряжены» требования к свойствам: велик ли будет брак от допускаемого рассеяния между плавками, партиями? Не понадобятся ли существенные усовершенствования в технологии выплавки, обработки, пооперационного и сдаточного контроля? Какой риск «выбросов» свойств допускает конструкция; нужна ли стопроцентная проверка всех деталей, какая и где: на входе, на выходе? Предусмотрены ли сдаточные испытания каждого узла, изделия и могут ли они отсеять брак материала или он вскрыется только в работе? Существует ли «запирающая операция», которая исключит попадание брака на сборку или к потребителю (при испытании гидросистемы высоким давлением все, что может лопнуть, — лопнет).

Пусть теперь нормативы для свойств и испытаний материала заданы. Но как доказать, что они действительно гарантируют работоспособность конструкции? Намечены ли натурные испытания разработанного узла до разрушения? С какой программой тензометрии, дефектоскопии и фрактографии после них? Поставят ли узел или машину на стендовые испытания? По какой программе нагрузок, на какой срок: до гарантийного ресурса, до первого отказа или до полного отказа всех узлов (с постепенной заменой выбывших)? Сколько эк-

земпляров? Организуют ли для этих экземпляров дополнительный контроль качества в производстве?

Планируют ли после стендовых полевые испытания конструкции (летные для самолета). Будут ли эти испытания многоступенчатыми (заводские — ведомственные — государственные)? До какой степени можно сравнивать варианты исполнения детали? Сколько вариантов можно поставить на испытания в каждой степени? Что внести в программу испытания узлов и деталей для проверки работоспособности материала?

Не понадобятся ли стендовые и полевые испытания специально для проверки данной детали (например, нового трака к трактору)? Будут ли полевые испытания детали пассивным экспериментом (сбором и сопоставлением поступающей информации о фактических характеристиках, отказах и условиях работы) или нужен активный эксперимент с заданным наперед распределением наработки по условиям работы? Какой ширины охват условий будет представительным (по характеру грунтов, по климатической зоне и времени года)?

Каково минимальное число параметров, позволяющее достаточно надежно задать эти условия? Какие крайние сочетания условий явно необходимо проверить? По какому принципу надо выбирать остальные сочетания условий из всех возможных? Чем ограничено число сочетаний сверху и снизу? Как следует классифицировать причины отказа при массовых испытаниях и в эксплуатации, например: заданный износ проушины трака (от трения по металлу), износ шпоры (трение по свободному абразиву), поломка (по какому месту?), раковины и другие макродефекты в изломе?

Существует ли классификатор отказов как документ, позволяет ли он безошибочно распознавать причину даже малоквалифицированному наблюдателю? Можно ли ограничиться отчетностью испытателей на бумаге? Какую долю поломок необходимо проверить в натуре на месте, какую отобрать для исследования в лаборатории? Нужно вернуть для исследования все испытанные детали, только некоторый процент их (по какому признаку отбор?) либо все детали с одним типом отказа (например, с литейными дефектами)?

Кто будет регистрировать условия работы и время наработки на отказ? Насколько это надежно? Существует ли сеть организаций, обязанных вести полевые ис-

пытания (например, машиноиспытательные станции МСХ СССР), или испытания проводит завод-изготовитель своими силами или по договоренности с эксплуатационниками? Не окажется ли все возложено на разработчика сплава?

Каковы реальные сроки испытаний (не только по «машино-часам», но и по загрузке испытателей, ожиданию нужных условий и т. п.)? Можно ли получить и использовать часть информации до завершения испытаний? Не понадобится ли программа испытаний разбить на «малую предварительную» и полную (куда первая входит как часть)?

Задав число деталей «на одну точку» и «матрицу вариантов», определим объем металла для испытаний. Достаточно одной плавки, садки или партий будет несколько (а при малостабильной технологии не понадобится ли предварительная выборка из многих плавков)? Как уравнивать разнородную выборку из многих плавков: каждую из них равномерно на все варианты приемами математического планирования (и маркировать при этом каждую деталь) или рассеять случайной выборкой (рандомизацией)? Или же технология производства и сбыта такова, что одна партия неизбежно попадает целиком в одни условия, а другая — в другие? Тогда какой проверкой можно доказать идентичность партий? Нужно ли оставлять себе от каждой партии детали — свидетели? Можно ли отбирать «сходные» партии до отгрузки или параллельно с полевыми испытаниями?

Не все эти вопросы начинающий инженер имеет возможность решить самостоятельно, но он обязан их обдумать и добиться однозначных и обоснованных ответов на свои предложения. «Ничей вопрос», оставшийся висеть между инстанциями, нередко, всплывая в конце, делает бесполезной всю работу.

Решение принципиальных «внешних» вопросов, связывающих исследователя с конструктором, технологом, эксплуатационником, экономистом, определяет постановку «внутренних» задач — о цели исследований в лаборатории (в том числе и «не прикладных» работ). Отсутствие технологии, неприемлемые сроки или затраты для ее перестройки, неизбежный побочный отрицательный эффект при избранном принципе обеспечения свойств могут и вообще свести на нет замысел материалововеда — до всяких экспериментов.

Когда натурное испытание просто и достоверно, а изменения в технологии несложны, новое иногда появляется и вообще без исследования (попробовали — получилось). Однако если проверка сделанного предложения требует существенной затраты сил, то и анализ его надо делать по всем правилам — прежде «опробования». «Метод тыка» никогда не решает коренные проблемы (если нужда всеобщая, то очень мало вероятно, чтобы никто не пробовал столь простое средство раньше).

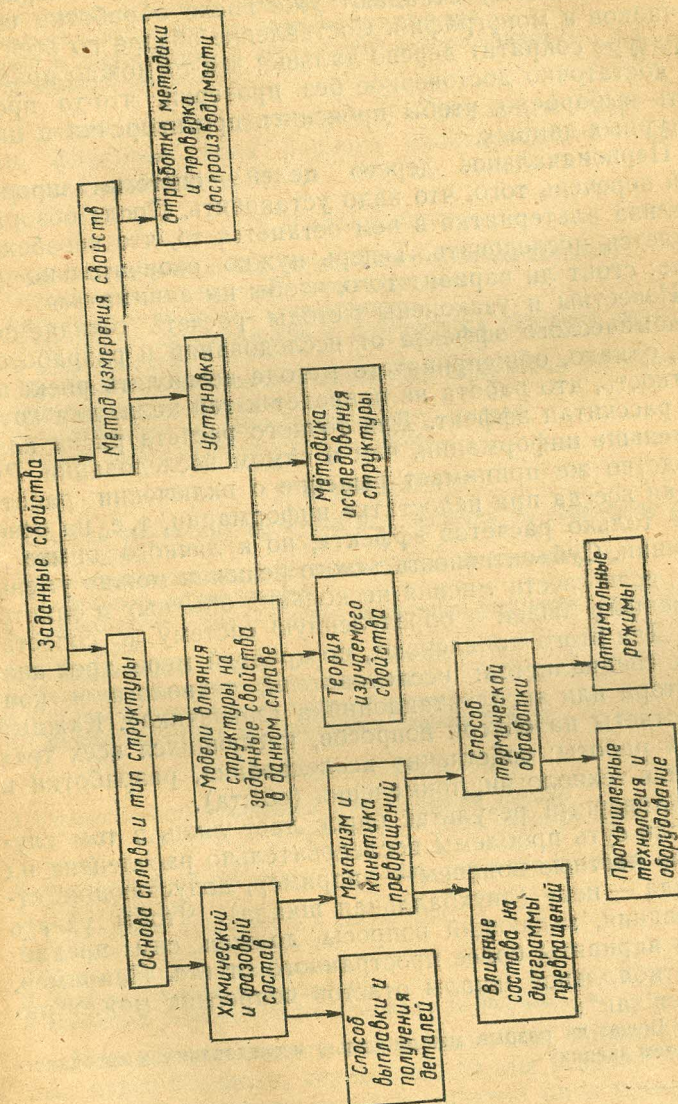
Рекомендательный библиографический список

1. Консон А. С. Экономика приборостроения. М.: Высшая школа, 1980, 576 с.
2. Martino Дж. Технологическое прогнозирование: Пер. с англ. М.: Прогресс, 1977. 592 с.
3. Диксон Дж. Проектирование систем: изобретательство, анализ и принятие решений: Пер. с англ. М.: Мир, 1969. 440 с.
4. Шеннон Р. Имитационное моделирование систем — искусство и наука: Пер. с англ. М.: Мир, 1978. 420 с.
5. Элементы теории испытаний и контроля технических систем / Под ред. Р. М. Юсупова. Л.: Энергия, 1978. 192 с.
6. Алтшуллер Г. С. Творчество как точная наука. Теория решения изобретательских задач. М.: Советское радио, 1979. 176 с.

ДЕРЕВО ЦЕЛЕЙ

Пусть установлено, что для достижения цели изменением конструкции или технологии затраты и сроки так велики, что, по-видимому, оправдано создание нового или усовершенствование имеющегося материала. Тогда надо выбрать путь, оценить затраты и сроки разработки. Конечные требования к материалу уже сформулированы в техническом задании. Для достижения этой цели надо решить серию частных задач, каждая из которых расчленяется на подзадачи. Чтобы обозреть весь объем и взаимообусловленность работ, можно составить дерево целей (см. рисунок) — граф соподчинения частных задач. Рисунок будет полезным, если сделать его сначала шире реального объема исследования. Проще нарисовать для полноты картины что-то лишнее (чтобы попутно вспомнить и все необходимое), чем потом, уже в ходе работы, пытаться дополнить объем и финансирование и отодвинуть сроки.

Имея перед глазами дерево целей, делают первую прикидку плана — перебор альтернатив (сплав с нужным комплексом свойств может быть медным или тита-



Начало дерева целей при разработке сплава

новым, литым или кованым, стареющим или упрочненным наклепом). Часть вариантов (и некоторые вершины дерева целиком) отпадают уже после обработки справочников и монографий. Составление обзора по всей литературе сократит дерево дальше: что-то можно принять за достаточно достоверное без проверки, что-то проверить выборочно, чтобы прояснить противоречия в литературных данных.

Первоначальное дерево целей — систематизированный перечень того, что надо установить. После обзора и анализа альтернатив в нем останется то, что неизбежно придется исследовать. Теперь нужно окончательно решить: стоит ли вариант того, чтобы им заниматься.

Известны и узаконены методы расчета ожидаемого экономического эффекта от исследований и разработок. Нет, однако, общепринятого метода вычислять риск: вероятность, что работа не приведет к той цели, для которой рассчитан эффект. Для точного расчета риска надо не меньше информации, чем в самом исследовании. Руководство же принимает решение о включении работы в план всегда при недостатке информации, т. е. на основе не только расчетов эффекта, но и личного опыта и интуиции. Субъективность такого решения можно уменьшить, если свести мнения нескольких опытных и заинтересованных людей в определенную систему и сопоставить. Для этого привлекают, по крайней мере, трех знающих специалистов: исследователя, технолога и конструктора или эксплуатационника (заказчика). Каждый дает ответы на список вопросов, касающихся всех трех стадий работы: собственно исследования, разработки и освоения технологии, применения (сбыта).

Достоверный результат опрос дает лишь в том случае, если суть проблемы последовательно расчленена на простые частные вопросы — например, допускающие ответы «да — нет» (двухбалльная шкала). Чтобы узнать обоснования, задающий вопросы должен сам предложить и варианты более пространного ответа. Например, для пятибалльной шкалы ответов вопросник может начинаться так*:

Вопрос. Велик ли разрыв между целью исследования и сегодняшним уровнем знания?

* По Флитвуду, 1977 г.

Ответ	Балл
Для решения недостает фундаментальных знаний	-2
Существенно недостает технологических знаний	-1
Небольшая недостача технологических знаний	0
Близко к достигнутому в предыдущих исследованиях	+1
В пределах достигнутого	+2

Главное в перечне вопросов — возможно шире охватить факторы риска разного происхождения — например, по такой схеме:

1. В исследовании

1. Величина разрыва между целью и сегодняшним знанием.
2. Ограничения по затратам на исследование.
3. Ограничения по срокам исследования.
4. Зависимость от смежников в изготовлении образцов, испытаниях и оборудовании.
5. Предсказуемость объема и детального содержания экспериментов.
6. Обеспеченность кадрами.
7. Наличие крупных специалистов — консультантов.
8. Личная заинтересованность участников.

2. В производстве

1. Степень улучшения продукта или процесса по сравнению с существующим.
2. Совместимость нового с существующим производством.
3. Зависимость от смежников.
4. Соответствие лабораторных испытаний условиям цеха или эксплуатации.
5. Определенность задачи для технологов.
6. Наличие специалистов в технологии.
7. Связи с производством и потребителем.
8. Соответствие цели долгосрочным планам фирмы.

3. В сбыте и эксплуатации

1. Совместимость технологии с окружающей средой.
2. Лимит ресурсов (материалов, энергии, оборудования, трудовых затрат).
3. Возможный объем сбыта — его гарантии и перспективы.
4. Конкуренция.
5. Социальная потребность в результате.
6. Четкость технической цели.
7. Компетентность заказчика.
8. Заинтересованность заказчика.
9. Важность экономического эффекта для фирмы.

Можно предложить много способов оценивать каждый ответ баллами, переводить баллы в вероятность неудачи, вычислять по цепочке событий риск прекращения работы на каждой стадии и возможные от этого убытки и, сопоставляя с эффектом, принимать решение: делать

или не делать. Несомненно, однако, что все перечисленные факторы действительно важны и должны быть учтены наперед, а мнение технолога и заказчика здесь так же важно, как и исследователя. Поэтому стоит отнять еще час времени у этих специалистов, чтобы получить конкретную оценку ситуации в виде списка ответов. Без понимания риска нет смысла в любых расчетах ожидаемого экономического эффекта исследования. Если риск приемлем и вариант принят, по дереву целей составляют план.

ПЛАН, ОБЪЕМ, СРОКИ

План, что и говорить, был превосходный: простой и ясный, лучше не придумаешь. Недостаток у него был только один: было совершенно неизвестно, как привести его в исполнение.

Л. Кэрролл. «Алиса в стране чудес»

План, написанный за два часа, иногда можно уточнить весь год. Поэтому вряд ли стоит сожалеть о времени, потраченном на составление подробного плана.

Даже для единолично выполняемой «внеплановой» поисковой работы пишут план. Только выписав по порядку все действия, можно убедиться, что работа имеет конец, и он достижим за разумное время.

Для утверждения плановых работ обычно требуется два или три документа: программа, календарный план — в виде списка работ и сроков или сетевого графика (иногда это часть программы) и методический план. Программа — перечисление основных этапов работы с указанием результата, метода, сроков и исполнителей. Это, по существу, такая роспись дерева целей «в столбик», где формулировки четки и однозначны также и для финансистов:

2.4.1. Дилатометрическое и магнитометрическое исследование кинетики изотермического превращения в сплавах № 1—5, 9—13 при температурах —60; —40; —20; 0° после закалки от 700, 720, 740 °С.

01.02.85—15.03.85

Лаборатория № 16

Программе предшествует более трудоемкий основополагающий документ — методический план. В нем конкретно и исчерпывающе полно перечислены сплавы, их

составы, варианты выплавки и обработки, схемы разрезки и отбора слитков, прутков, проб, образцов, порядок их нумерации. Указаны все изучаемые сочетания вариантов состава, способов обработки, методов испытания и структурного исследования, число образцов на точку для каждого варианта условий и вида испытания. Приведены расчет и таблицы числа образцов каждого типа для всех вариантов вместе и разбивки их по вариантам, подсчитано количество металла, число плавов, партий и общий объем всех испытаний. Как пункт методического плана формулируется и техническое задание на разработку установки или методики.

Обычные пункты плана:

«...1.3.2. Сплавы выплавляют в 3-т дуговой печи из шихтовых материалов (табл. ...), раскисляют в ковше алюминием (0,5 кг/т) и разливают в 200-кг слитки типа III (по нормам завода).

...1.3.4. Один слиток от каждого куста после разрезки по схеме рис. ... маркируют по табл. ... для исследования в литом состоянии. Остальные слитки после обрезки 15 % сверху и 5 % от донной части прокатывают при 1200—950° на квадрат 60×60 мм и разрезают на прутки длиной 2 м. По 10 прутков от головной и от донной части каждого слитка клеят по табл. ...

... 3.2.1. Испытания на растяжение проводят при +20 °С по ГОСТ 1497—73 на 5-кратных образцах диаметром 5 мм типа II, вырезанных вдоль и поперек направления прокатки по схеме рис. ... по 4 образца на точку. Маркировку образцов (табл. ...) наносят по чертежу (рис. ...) клеем на прибыльную часть до разрезки и переносят электрографом на головку образца перед обрезкой прибыли.

...3.2.2. Макроскопический анализ излома и измерения его геометрии делают на всех образцах на растяжение и на ударную вязкость. Один образец на точку после испытания на растяжение медленно покрывают капон-лаком для микроскопического исследования (обе половинки излома).

... 4.1.1. Раскрытие трещины измеряют приспособлением из двух скоб с тензодатчиком и электронных мостов. Обеспечивается синхронная запись перемещений в двух сечениях образца в диапазоне 0—0,5 мм с точностью не хуже 5 мкм и быстродействием не менее 1 с при температурах —70÷+20 °С.

... 5.1.1. От одной половинки разрушенного образца на электроискровом станке отрезают столбик высотой 3 мм для исследования излома в сканирующем электронном микроскопе и две шайбы толщиной 3 мм для шлифа и фольги. После исследования в сканирующем микроскопе от излома снимают реплику для извлечения частиц. Вторая половинка после магнитометрии служит резервным образцом для фрактографии.

... 6.2.2. Реплики с извлечением для электронно-микроскопического исследования натрия алюминия делают на шлифах от головок продольных разрывных образцов после закалки от 800 °С без отпуска (по 3 образца от головной части одного слитка на каждый вариант раскисления). Общая площадь просматриваемых реплик 5·10³ мкм² на один вариант. На этих же шлифах затем делают металлографический анализ неметаллических включений (п. 7.3.3.)»

Первый вариант методического плана содержит «по максимуму запросов» астрономическое число образцов, но зато он включает все интересные варианты и их перекрестную проверку. Прикинув по порядку величины достаточное количество металла, темпы изготовления образцов и испытаний, начнем сокращать план. Этот путь надежнее, чем «приспособление к обстоятельствам» с самого начала (сначала «как надо», а потом выявить «что здесь главное»).

Логика сокращения обычно такова.

Во-первых, лучше меньше, да лучше: лучше немного надежных точек с достаточным для статистических выводов числом образцов на точку, чем много «проб» по одному образцу, когда не поймешь, чему же верить. Во-вторых, совсем не обязательно делать все испытания для всех вариантов. Можно сразу видеть, что по некоторому аргументу отличия в данном свойстве будут заведомо невелики (или малоинтересны). Соответствующее исследование вдоль этого ряда исключим, сохранив контрольные точки (с двух краев либо на пересечении с рядом по другой переменной). В-третьих, обычно лучше основательно потратиться на два главных — трудных, но перспективных — варианта, чем сделать за их счет восемь попыток попроще, но без глубокого замысла. Наконец, чтобы обеспечить сопоставимость по всем параметрам с базовым вариантом («до улучшений»), на него тратят столько же, что и на самый многообещающий вариант. Поэтому окончательная таблица вариантов испытаний не квадрат, а причудливое кружево, для каждого метода — со своим узором, но их наложение всегда укажет общие опорные точки, где имеются все данные, и от них идет отсчет во всех рядах.

Можно сократить объем при той же надежности, применяя математическое планирование эксперимента, но все его приемы исходят из существования заданной наперед области поиска, количественно сформулированной цели поиска, однородного массива образцов, стандартизированной техники измерений, простых гипотез о виде искомых зависимостей. Поэтому оно полезно лишь как второй этап планирования — когда определена цель и области поиска для каждого типа экспериментов (но именно от их выбора больше всего зависит объем работы).

Нельзя запланировать открытие, но вся работа для него планируется (в том числе и по срокам): ресурсы и время всегда жестко ограничены. Кроме директивных сроков, существует и естественный срок, после которого работу прекращают за ненадобностью. Зная прогноз возможных причин неудачи, минимизируют риск, избавляясь по возможности от сомнительной методики и трудноорганизуемых промышленных экспериментов.

Многие технологические процессы и приемы отомрут раньше, чем их успеет постигнуть теория. Поэтому рискованно углубляться в экспериментальное исследование механизма явлений для решения срочной «разовой» практической задачи.

Основные затраты на эксперимент определяются по методическому плану: количество материала (с 2—3-кратным запасом на разрезку и другие потери), число плавок и экспериментов на оборудовании цеха, количество образцов каждого типа. Все это взвешивают не только по стоимости, но и, главное, расставляют в сеть по срокам, исходя из пропускной способности соответствующих цехов (с учетом приоритета других работ).

Сравнительно легко планировать по аналогии с предыдущим время, загрузку работников, оборудования для простых работ — стандартных испытаний и типовых измерений на серийном оборудовании (хотя и здесь потери от переналадки, ремонтов и загрузки узких мест соседями всегда превышают собственно машинное и подготовительно-заключительное время испытаний).

Наименее определенный по срокам и обычно наибольший по объему раздел — создание новой методики (или ее приспособление к новым условиям). Подобрать реактив для травления нового сплава можно за полдня, но иногда не хватает и недели. Фольги для электронной микроскопии иногда удается сделать за три дня, но иногда они не получаются и за месяц. Если измерительную установку можно собрать за месяц, то над стабильностью и воспроизводимостью можно биться полгода. Там, где нельзя вычислить срок надежнее, в план закладывают время ближе к верхнему пределу. Здесь важно также возможно подробнее расчленивать и спланировать наперед сам процесс отладки, но главное: ограничить выбор метода с учетом отведенных сроков.

Сложная методика оправдана для важной, рассчитанной на работу с «дальним прицелом» задачи.

Предварительный анализ затрат и сроков нередко вынуждает ограничить размах и цель исследования (много хуже, если это приходится делать по ходу работы).

Когда время и последовательность всех работ определены, составляют сетевой график. При большом числе исполнителей его оптимизируют на ЭВМ. Небольшой группе оптимизация на ЭВМ редко помогает (работники все те же, а неопределенность в загрузке оборудования у смежников большая). Однако сам сетевой график полезен и тут: на нем в каждый момент видно не только все, что осталось сделать, но и ближайшая точка наибольшего риска, о которой пора подумать.

2. ОБЗОР ЛИТЕРАТУРЫ

ЗАДАЧА

Так-то так, да дальше-то как?
Пословица

Обзор литературы — обязательная часть всякого отчета об исследовании. Он «должен полно и систематизированно излагать состояние вопроса..., позволять объективно оценивать научно-технический уровень работы, правильно выбирать пути и средства достижения поставленной цели и оценивать как эффективность этих средств, так и работы в целом. Предметом анализа в обзоре должны быть новые идеи и проблемы, возможные подходы к решению этих проблем, результаты предыдущих исследований..., данные экономического характера, возможные пути решения задач... Противоречивые сведения, содержащиеся в различных исходных документах, должны быть проанализированы и оценены с особой тщательностью» (ГОСТ 19600—74 «Отчет о научно-исследовательской работе. Общие требования и правила оформления»).

По обзору можно также судить об уровне знаний и общей культуре исследователя (очень плохой признак — обзор, составленный после работы, чтобы сдать отчет). Труд окупается, если обзор помог уяснить необходимость, цель и метод исследования — в начале пути, — и понять смысл и место собственных результатов в общем ряду — в конце.

Из анализа литературы должно быть видно, что в этом узком вопросе известно вполне достоверно, что сомнительно, спорно; какие задачи в поставленной технической проблеме первоочередные, ключевые; где и как стоит искать их решения. Хороший обзор — исчерпывающе полный, критический и обобщающий — это труд не менее сложный, чем оригинальное исследование. Нельзя научиться писать хорошие обзоры по инструкции — нужен собственный опыт, но можно попытаться показать, какой минимум труда надо вложить в критический обзор, как рационально организовать этот труд, каких элементарных грубых ошибок, повторяющихся у каждого новичка, надо стараться избежать.

Затраты времени на обзор складываются примерно так:

- 1) выписки из справочников, чтение и конспектирование основных монографий 3—5 %;
- 2) составление рабочего плана обзора 1—2 %;
- 3) поиск периодики (и составление картотеки) 5—8 %;
- 4) чтение и конспектирование периодики 30—40 %;
- 5) отбор материала из конспектов, его сопоставление и анализ 20—30 %;
- 6) написание обзора 10—20 %;
- 7) правка текста 10—15 %;
- 8) переписка и изготовление рисунков 5—6 %.

Заметьте, что конец конспектирования всех первоисточников — это только половина работы.

Исследование всегда имеет узкую конкретную цель. В заключении обзора обоснованы выбор цели и метода. Весь обзор должен подготовить это решение. Отсюда следует его план и отбор материала. В обзоре рассматривают только такие узкие вопросы, которые могут прямо влиять на решение задачи, но зато настолько полно, чтобы охватить практически всю современную литературу по этому вопросу и планировать исследование без риска «открыть Америку». Требование полноты защищает от растекания вширь. Бесплезно начинать «от Адама», с пересказа прописных истин из учебников — увлекшись этим нехитрым делом, до сути задачи можно так и не дойти. Мало пользы и в хронологии работ со всеми заблуждениями и дискуссиями. Важно отразить современное состояние нашей конкретной задачи. Здесь

нужно сформулировать немного (4—7) основных вопросов. Отбирая их, надо предусмотреть, без чего нельзя обойтись, чтобы показать необходимость нашей работы и обсудить затем ее результаты (все данные для их сопоставления с литературой и совместного анализа заготавливаются в обзоре).

ПОИСК ИСТОЧНИКОВ

- Знаешь ли, где то — не знаю что?
- Знаю.
- Укажи, сделай милость.

Русская сказка

Приступая к совсем новому для себя вопросу, надо взять из наиболее известных монографий и справочников (в том числе из справочника библиографических справочников [1]) основные сведения о нем и ссылки на основные статьи. Затем, просмотрев в систематическом каталоге технической библиотеки соответствующую рубрику (например, «Металловедение» — пол-ящика карточек), можно найти одну-две малоизвестные книги или брошюры по узкой теме работы (это вполне оправдывает затраченное время). Библиография в книге обрывается за 2—3 года до выхода в свет. Все, что было после, надо искать в периодических изданиях. Прежде всего просматривают реферативный журнал (РЖ «Физика» или «Металлургия») за оставшиеся годы. К нему есть алфавитный предметный указатель (отдельный том за каждый год). Составлять такой указатель сложно даже с помощью ЭВМ, поэтому, если ищешь сведения, например, о мартенситном превращении в хромомарганцевой стали, надо не поленившись посмотреть «сталь хромомарганцевая», «хром в стали», «марганец в стали», «мартенсит», «превращение мартенситное» и еще все мыслимые сочетания и перестановки. Если известен автор, то его работы можно отыскать по годовому авторскому указателю к РЖ. Реферат в РЖ появляется через 5—7 мес. после выхода статьи. Еще год проходит, пока он попадет в предметный указатель. Поэтому, кроме РЖ, надо просмотреть и основные журналы за последние год-два.

Когда разрабатывают новый сплав, технологический процесс, установку, сведения о патентах или их отсут-

ствия — обязательная часть технико-экономического обоснования. Формулы всех изобретений, зарегистрированных в СССР, помещают в бюллетене «Открытия. Изобретения. Промышленные образцы. Товарные знаки», а в США, Японии, ФРГ, Англии и Франции — в бюллетене ЦНИИ патентной информации «Изобретения за рубежом» (с 1972 г.). Для поиска нужно иметь представление о системах классификации изобретений в этих странах [2]. Полные сведения о патентах по данной теме, включая и иностранные, надо искать во Всесоюзной патентно-технической библиотеке и отраслевых и региональных патентных фондах [2, 3].

Во всех журналах статьи снабжены рефератами. Уяснив по реферату, аннотации, рисункам, выводам, что работа полезна, ее заносят на карточку с тем, чтобы вернуться к ней позже. Картотека много удобнее, чем тетрадь для ссылок, потому что карточки легко сортировать, выбирать, сравнивать. На карточке обязательно пишут фамилии и инициалы всех авторов, название журнала, год, том, номер, страницу. Некоторая избыточность информации здесь полезна (зная инициалы, легче искать работы того же автора в РЖ; зная и том и год, можно найти работу, даже если в одной из цифр описки. Не стоит ограничивать список первыми тремя авторами (как это заведено у библиографов): иногда чтобы судить, одной ли школы эта работа, какого она плана (и качества), надо видеть главу коллектива (или единственного известного мне автора), а он мог оказаться по алфавиту пятым...

Нет единого правила, что писать в карточке. Для личной картотеки выгоднее писать не название работы, а суть интересующего нас результата и метод. Так, по сравнению с заглавием

«Влияние зернограницных сеграгаций фосфора и микроструктуры на хрупкое разрушение стали 35Г2 после закалки и среднего отпуски»

почти такой же длины запись:

«35Г2, отпуск 300 °С, излом зернограницный. От 0,03 % до 0,1 % Р порог хладноломкости T_{50} поднимается на 150—300 °С. $T_{50} \sim -1/\sqrt{D_{\text{зерна}}}$ В изломе Р (Оже-спектр). Сегрегация при 950—1050 °С (по канавкам травления)»

содержит гораздо больше информации.

Рукописный обзор к отчету должен отражать все самые свежие работы — вплоть до вышедших за 1—2 ме-

сяца до защиты. Поэтому, пока пишется обзор и делается работа, картотека должна пополняться непрерывно. Выбрав круг чтения, просматривают хотя бы несколько журналов постоянно, из номера в номер, и заносят на карточки наиболее ценное. Никакие автоматизированные поисковые системы не заменят личного архива, где материал регистрируется лишь после оценки по надежности, полезности, перспективе. (Беда больших картотек общего пользования в том, что они сильно засорены малополезными и даже просто ошибочными работами с актуально звучащими заголовками и оптимистическими аннотациями).

Закончив просмотр, сортируют карточки и планируют чтение. Читают «с конца к началу»: в первую очередь — обзоры и наиболее свежие работы. В начале их обычно говорится о предыдущих работах. Отсюда можно понять, что стоит прочитать. Если сгруппировать карточки по пунктам плана обзора и читать родственные работы в один день, легче их сопоставлять. В прочитанном есть ссылка на другие интересные работы, о которых не знал раньше. Они составят второй круг чтения. Можно надеяться, что основная литература изучена, если «круг ссылок» замкнулся: в новых статьях уже не встречаются новые источники.

Если книга есть в одной из библиотек страны, ее можно получить по межбиблиотечному абонементу. Если упомянутой в РЖ работы нет ни в одной библиотеке, можно заказать ее микрофильм в ВИНТИ.

Иметь собственную копию статьи удобно (на ней можно делать пометки), но это не избавляет от необходимости конспектировать.

Рекомендательный библиографический список

1. Иениш Е. В. Ретроспективный поиск технической литературы. М.: Книга, 1980. 230 с. (Библиографические издания по металлургии и металловедению, с. 117—122).
2. Морозов А. Н. Патентная документация в библиотеках. Учебник. М.: Книга, 1979. 256 с.
3. Кичкин И. И., Скорняков Э. П. Патентные исследования при курсовом и дипломном проектировании в высших учебных заведениях. М.: Высшая школа, 1979. 112 с.

СОДЕРЖАНИЕ КОНСПЕКТА

Все-таки в науке есть что-то захватывающее. Дело пустяковое — количество фактов, а берешь колоссальный дивиденд в виде умозаключений.

Марк Твен. «Жизнь на Миссисипи»

Цель конспектирования — не только записать, но и предварительно обработать и проанализировать основные факты и аргументы автора, привести их к такому виду, чтобы можно было сравнивать с другими экспериментами и с теорией. Обычная ошибка — конспект выводов без анализа их надежности. «Необходимо критическое восприятие того, что сделано, причем критическое восприятие не только фактов — факты в науке почти всегда достоверны, — а трактовки фактов, она может быть чрезвычайно субъективна, и критическое отношение к этой трактовке необходимо для возникновения новой идеи исследования»*.

В публикациях экспериментов есть три основных дефекта, которые могут поставить под сомнение весь результат:

а) несоответствие обещанной точности и фактически доступной при данных средствах измерения (не учтены некоторые источники ошибки; в результате, например, автор утверждает, что результаты укладываются в параболу, а рассеяние точек такое, что можно провести и прямую);

б) не учтены побочные факторы (примеси, помехи), которые влияют на явление так же сильно, как и основной;

в) дефекты логики в выводах. Например, получив зависимость степени превращения y от времени t и коэффициента диффузии D в виде $y = k\sqrt{Dt}$, автор считает это подтверждением своей модели распада твердого раствора, где также $y \sim \sqrt{Dt}$. На деле же такую зависимость следует ожидать для всякого процесса объемной диффузии просто по соображениям размерности, а выбирать модель здесь можно только по величине k .

Подмечая подобного рода ошибки, сразу делаем в конспекте пометки (в квадратных скобках, чтобы не приписать их потом автору).

* Зильбер Л. Стратегия научного поиска. Природа, 1969, № 10, с. 48.

Чтобы иметь возможность критически сравнивать работы и отсеивать верное, все обоснование выводов фактами должно сохраниться в конспекте. Надо записать не вывод, а его логическую схему со своими замечаниями:

Автор: с ростом температуры закалки должны растворяться включения, поэтому меньше центров роста и старение замедляется.

[Но количество включений не мерили, а зерно выросло в 5 раз — не было ли зарождение зернограничным? — проверить по структуре?]

Автор: обработка холодом снижает пластичность [но рассеяние $\pm 5\%$ — значимого различия нет?].

Многие недоразумения и сомнения всплывут при сравнении одной работы с другими. Конспект полезен только, если по нему можно сличать работы, не перечитывая их целиком снова. Для этого должны быть сохранены все подробности методики.

Если в конспекте по работе [А] записано: «В сплаве после закалки выделений карбидов не видно», а по работе [Б]: «в закаленном состоянии есть цементит», то при сопоставлении обзора останется только разводить руками: «данные [А] и [Б] о наличии карбидов взаимно противоречивы». Стоило же записать так:

[Б]» Fe—0,01 % C; 0,005 % N; 0,10 % Si; 0,02 % Mn, образец $0,1 \times 10 \times 10$ мм, закалка из вакуума 0,1 Па, 700 °C, 5 мин, вода, зерно 40 мкм (средняя хорда). На фольге $\times 10\,000$ (УЭМВ—100 К) цементит [По снимку $L \sim 100$ нм, $B \sim 10$ нм, 4 рефлекса Fe_3C по микрофракции]

[А]» Fe—0,02 % C; 0,15 % Si; 0,5 % Mn, образец $1 \times 5 \times 20$ мм, закалка: 1 Па, 700 °C, 20 мин, вода, зерно 30 мкм (средний диаметр). Травление 5 %-ной HNO_3 в спирте, МИМ-8, $\times 1000$: никаких частиц (только зерно)»,

Тогда никакого противоречия нет, и в обзоре будет записано:

«При 0,01 % C после резкой закалки от 700 °C (образец толщиной 0,1 мм, охлаждение в воде) успевают выпасть карбиды не длиннее 0,1 мкм [Б]; карбидов крупнее 0,5 мкм нет и при меньшей скорости охлаждения (образец толщиной 1 мм) — при оптическом увеличении $\times 1000$ травление в 5 %-ной HNO_3 цементита не выявило [А]»

Все подробности методики, рассыпанные по тексту статьи, надо собрать и компактно изложить в начале конспекта (даже если сам автор не придает им значения, они-то иногда и решают все).

Все интересные для нас графики схематически зарисовываем (даже если есть фотокопии) — это облегчает последующую работу с конспектом. На рисунке обяза-

тельно пометить масштабы по осям или значения абсциссы и ординаты в характерных точках (нанести в масштабе по клеточкам точки начала, конца, экстремумов, а остальное провести на глаз). В тексте к графику отметить: «[13 точек, рассеяние до 10 кДж/моль]» — если потом придется ломать голову, какая кривая правильнее — в работе [А] или [Б], то важно вспомнить, что в [А] она проведена по 25 точкам, а в [Б] — по трем и одна выпала.

Графики, которые понадобятся для окончательного текста обзора, придется свести на кальку.

По снимкам полезно отмечать характер и размеры основных элементов микроструктуры, даже если в тексте о них не говорится: [зерно 300 мкм (крупнее кадра), включения 1—3 мкм, преимущественно у границ]. Это необходимо для качественного сопоставления структур из разных работ по конспектам.

ИЗУЧЕНИЕ ТЕОРИИ

Метод постулирования имеет много преимуществ, совпадающих с теми, какие имеет воровство по сравнению с честным трудом.

Бертран Расселл

Экспериментаторы часто жалуются, что им трудно читать теоретические работы. Следует помнить, что, и не владея соответствующим математическим аппаратом, можно критически оценивать эффективность теории, ее полезность в приложении к тем явлениям, которые мы изучаем. Если не ставится цель использовать аппарат теории или опровергнуть ее, нет смысла следить за выкладками и преобразованиями — иногда за выражением «после подстановки (10) и (8) в (5) получаем (11)» кроется неделя вычислений. Важно проверить прежде всего постановку задачи: какие факторы учтены, а какими пренебрегают с самого начала (и, стало быть, нет надежд увидеть в конце, что они дают). О неучтенных факторах часто умалчивают — важно их вспомнить и записать на видном месте, ибо типичный брак в работе теоретика — тонкое исследование эффектов второго порядка в пренебрежении одним из первых.

Далее надо выудить из текста все допущения, сделанные по ходу вычислений. Например, выражение «разложив z и y в ряд по t и ограничившись первыми двумя

членами, получаем решение в виде...» может означать, что автор только записал уравнения и, не сумев их решить, возвращается к тривиальной задаче, где ответ можно было увидеть и вообще без вычислений. Физический смысл всех попутных допущений надо продумывать сразу и пометать в конспекте. Надо выяснить, сколько же автор оставил себе свободных параметров, которые, надо определять из того же самого эксперимента? Если их 2—3, то можно объяснить все, что угодно, при любых формулах. Польза есть от такой теории, где параметры находят из независимых экспериментов. «Истинная теория явления отличается от простого пересказывания фактов учеными словами именно тем, что из нее следует гораздо больше, нежели дают сами факты, на которых она основывается»*.

Если окончательный результат для нас полезен, следует прикинуть, правильный ли он. Прежде всего — проверить переход к очевидным предельным случаям («если основной параметр $s \rightarrow 0$, то никакого эффекта не будет — верно ли?»; «если $a = 0$, все должно сводиться к ранее известному решению (A) — сведется ли?»; «если $c \gg k$, а $b = h$, то по смыслу следует ожидать $x \equiv y$ — так ли?»). Далее надо выяснить: при ограничениях, сделанных в постановке задачи и по ходу вычислений, для какой области значений основных переменных и параметров получено решение? Встречаются ли такие сочетания параметров (одновременно $a \gg 1$, $p \sim 1$, $c \ll 1$) в природе? (Т.е. описывает ли теория хоть какое-нибудь реальное явление?) Где она применима? Надо проверить и порядок величины результата: «во всяком случае $10^6 < p < 10^{10}$ см⁻², а $b \sim 10^{-8}$ см. Тогда формула дает всегда $e < 10^{-7}$, а в эксперименте измеряется $e \geq 10^{-6}$. Очевидно, теория относится совсем не к той области, где мы наблюдаем явление».

Далеко не всякий результат теории удастся количественно сопоставить с экспериментом — не хватает констант или надежности самого эксперимента. Однако не так легко и утверждать, что «эксперименты [5, 7] качественно подтверждают теорию [8]». Бывает, что $x(t)$ в эксперименте падает примерно по гиперболе, как и в теории, но если ожидалось падение на 10 %, то измеренные

* Ландау Л. Д., Румер Ю. Б. Что такое теория относительно-сти? М.: Советская Россия, 1959.

значения x в том же интервале t падают на порядок. Очевидно, кроме случайного сходства графиков, между теорией и экспериментом нет ничего общего. Не имея прямых данных, иногда за подтверждение теории выдают косвенные: «вычисленная сила блокировки дислокаций растет с концентрацией; в согласии с этим измеренный предел текучести тоже растет» — но где показано, что эти величины должны быть одного порядка и связаны линейно?

Итак, постановка задачи, промежуточные допущения и ограничения, область существования решения и способ сравнения с экспериментом — вот что должно войти в конспект теории для дальнейшего анализа и использования (разумеется, должна быть основная исходная и конечная формула и обозначения).

ТЕХНИКА КОНСПЕКТИРОВАНИЯ

- Когда вы начали?
- Я думаю, это было четырнадцатого марта, — ответил он.
- Пятнадцатого, — сказал Мартовский Заяц.
- Шестнадцатого, — добавил Соня.
- Запишите, — сказал Король.
- И судьи со рвением записали все три даты, затем сложили их вместе и перевели в шиллинги и пенса.

Л. Кэрролл. «Алиса в стране чудес»

Прежде всего, в отличие от просмотра для картотеки, статью читают целиком, до конца. Изучая для экономики отдельные куски, можно потерять где-нибудь в примечании такую подробность методики, которая ставит крест на всей работе. Только монографии можно читать выборочно — отдельными главами и параграфами (выяснив общие позиции автора по введению и заключению).

При чтении делают пометки, что выписать. Выписывать сразу по мере чтения неэкономно (трудно сгруппировать вместе всю методику, трудно выделить важное, не зная, чем дело кончится, одни и те же мысли повторяются: сначала намеком, потом более стройно). Прочитай и разметив текст, не откладывая, пока вся статья свежа в памяти, пишут конспект.

В толстой тетради, где все конспекты пронумерованы, начинают с «выходных данных» статьи: инициалы и фамилии всех авторов, полное название статьи (на язы-

ке оригинала), журнал, год, номер, страницы (от — до). Для книги — еще издательство, город, номера страниц и их общее число.

Выгоднее писать конспект полными словами, приемлемы только сокращенные обозначения терминов, если их одинаково используешь во всех конспектах («ЧД — частичная дислокация, ОР — обратная решетка...»). При чтении конспекта важно быстро схватывать суть дела, не отвлекаясь на расшифровку иероглифов. Использование конспекта — более напряженная работа, чем его составление, поэтому сокращение слов — дорогая «экономия на спичках».

Пропорция между объемом статьи и конспекта зависит, конечно, от сложности содержания, количества разнородных фактов, но в среднем, при некотором опыте, удается сжимать материал в пропорции $\sim 1:10$. Если интересен только какой-то частный результат или только методика, можно конспектировать не всю статью, но прочесть ее нужно целиком.

Обязательно надо записать словами смысл всех буквенных обозначений. То, что само собой разумеется при чтении, со временем забудется. Если в статьях на одну и ту же тему встречаются разные буквенные обозначения, полезно сразу записывать все формулы в одних и тех же обозначениях, привычных для себя, чтобы не отвлекаться, когда дело дойдет до сравнения формул из разных работ по существу.

Так же и все не метрические единицы в тексте и на графиках надо переводить сразу при конспектировании.

Все интересные литературные ссылки обозначим в тексте номером, а в конце сделаем выборку библиографии.

ГРУППИРОВКА И АНАЛИЗ МАТЕРИАЛА

Для чего толь многие опыты учинены в физике и химии? Для чего толь великих мужей были труды и жизни опасные испытания? Для того ли только, чтобы, собрав великое множество разных вещей и материй в беспорядочную кучу, глядеть и удивляться их множеству, не размышляя о их расположении и приведении в порядок?

Ломоносов

Безнадежное дело — писать обзор «по наитию», вспоминая разные факты из прочитанного. Бессмысленно пересказывать материал от статьи к статье. При этом не достигается главная цель обзора — критическое со-

поставление и анализ всех материалов вместе. Между конспектами и текстом обзора есть еще одна необходимая стадия промежуточной обработки материала. Прежде всего нужно перегруппировать материал конспектов в соответствии с планом обзора (данные одной статьи попадают в разные пункты плана, а в каждый пункт набираются факты из многих статей).

Полезно заготовить большие листы бумаги и на каждом надписать один заголовок из плана обзора. Разобьем лист еще и на подзаголовки. Затем, внимательно читая подряд все конспекты, отметим на этих листах каждый полезный нам факт в своей рубрике (номер конспекта и суть наблюдения). Иногда можно дать и мини-графики с числами в характерной точке. Так:

«2. Сверхструктуры внедрения на основе о. ц. к. металлов.

а) Теория

б) Наблюдавшиеся типы сверхструктур:

[А] Nb_6O — тип (а) [В].

[Б] NbO_2 — типа Ta_4O (микродифракция).

[Г] Nb_2O — о. ц. к., $c/a = 1,42$ (рентг.)

[Д] NbO_x — $a/c = 1,025$ (рентг.) — выбор осей?

в) область существования по c , T .

г) кинетика возникновения

Не надо путать факты и мнения — мнения заносят сюда отдельно.

Когда все конспекты обработаны, начинается главная работа — сличение и анализ фактов.

Для группы родственных фактов в каждой рубрике сводки надо выяснить и пометить: совпадают; прямо не сравнимы, так как относятся к разным интервалам условий, но не противоречат одни другим; противоречивы.

Для сравнения удобно однотипные кривые из разных работ свести на один сводный график, условия разных экспериментов — в одну сводную таблицу (ничего, если кое-где в ней будут прочерки оттого, что один автор, например, указал скорость охлаждения, а другой нет; пустые клетки могут как раз навести на мысль о том, почему в конце концов так разнились результаты).

Иногда для сопоставления надо перестроить некоторые графики в новые координаты (серии изохрон — в изотермы и т. п.), изменения одних величин в другие (например, снижение периода решетки и электросопротивления — в изменение концентрации раствора), используя градуировочные зависимости из других работ.

Можно построить из всех зависимостей (например, механических свойств от режимов обработки) сводные графики, например предел усталости как функция предела прочности. Лучше, если нужные для этого данные были в таблицах, но если нет — можно (очень аккуратно) снять их циркулем с графиков и построить таблицу самому. (Конечно, с графиков можно брать только точки, действительно соответствующие реальным образцам, а не любые точки с гладко проведенных кривых)

Всюду, где есть теория, надо сравнить ее с экспериментом, собрав, если надо, из других источников необходимые константы.

Когда все возможности сравнения исчерпаны, пишем страничку текста в обзор. Сжимая текст, надо сохранять конкретность фактов — пищу для размышлений читателю и себе, когда дойдет дело до обсуждения собственных результатов. Например:

«В большинстве работ [1—5] для сплавов с 30—33 % Cr при температуре 480—520 °C и выдержках от 50 до 1000 ч (табл. ...) является одна и та же кинетика превращения — рис. ... Нейтронографические данные о порядке (по интенсивности линии $\left\{ \begin{smallmatrix} 2 & 2 \\ 3 & 3 \end{smallmatrix} \right\} 0$)

[3] и рентгеновские (по ширине расщепления линии {220} [2, 1]) дают в сходных условиях (500 °C, 100 ч и 510 °C, 80 ч) примерно одну и ту же степень порядка ($s=0,6 \pm 0,7$). Выпадают из общей зависимости результаты для сплава с 30 % Cr [6] — в нем после 100 ч при 500 °C упорядочения не заметили (чувствительности рентгеновского метода $\pm 0,01$ нм [2] соответствует $s < 0,2$ [1]). Возможно, это связано с разной резкостью закалки (охлаждение с ампулой на воздухе [6], в воде или масле [1—5], табл. ...).

Автор обзора должен проявлять свое собственное отношение к фактам. Нельзя принимать на веру и пересказывать все подряд. «За что купил, за то и продаю» — довод не для науки. Ссылаясь на какой-либо факт или вывод без оговорок, мы тем самым признаем их надежность. Сомнения следует высказать. Если факты из разных работ противоречивы, а мы, упоминая те и другие, отдаем одним предпочтение, следует объяснить, почему.

Если мы знаем о каких-то работах, но не рассматриваем их из-за ненадежности результата, надо об этом сказать. Например:

«В ранних работах [А, Б] отмечали уменьшение количества карбида после наклепа. Однако эти результаты получили на репликах от шлифов, и наклеп мог влиять не только на количество карбидов, но и на их травимость».

Критическое отношение к делу должно выражаться не в обличении предшественников, а в обоснованном отборе того, что у них верно, и в конкретных предложениях: как поставить эксперимент (или выполнить расчеты, оценки), чтобы снять сомнения и вскрыть причину противоречий.

Обзор мнений не нужен — после обзора и анализа фактов достаточно выделить, чем можно объяснить всю совокупность фактов, какие из объяснений нельзя принять как слишком частные (они описывают лишь одну сторону явления), какие бесполезны как слишком общие (таким же способом можно объяснить не только это, но и многое другое — не видно, откуда появится специфика нашего случая). Например:

«Все факты согласуются со схемой превращения [3, 4], где скорость лимитирует зарождение на дислокациях. Представление о решающей роли закалочных вакансий [6] не подтверждается, потому что данные [6] не противоречат и модели стока на дислокации, а по электронно-микроскопическим наблюдениям [3, 7] частицы росли только от субграниц. Объяснить все зернограницным распадом [1] нельзя — тогда не видно, почему превращение ускоряется с наклепом».

Не всегда легко, но нужно докопаться, на каких, собственно, фактах основано общепринятое объяснение явления, повторяющееся из работы в работу. Иногда, дойдя по цепочке ссылок до первоисточника, видишь, что там это была всего лишь догадка, и уверенность в ней появилась впоследствии лишь от многих перепевов, но не от новых экспериментальных доказательств.

Везде нужно проводить четкую грань между объяснением (одним из многих возможных, которые пока не противоречат фактам) и доказательством (единственным объяснением, исключаяющим все альтернативы). Никогда не надо упускать случай придумать, какой эксперимент помог бы сделать выбор между несколькими из правдоподобных объяснений.

В обзоре теорий надо подчеркнуть все различия в исходных допущениях (например, сведя их в таблицу) и в качественных результатах — по вошедшим в ответ параметрам, виду зависимостей, порядку величин и лишь в самом конце — по численному результату, если он интересен сам по себе. Совпадение какого-то числа («ответа») очень редко может подтвердить теорию — важно общее согласие со всей совокупностью измеренных зависимостей.

Обзор теорий заканчивается сравнением условий и областей применимости каждой из моделей и указанием области реального эксперимента, где не годится ни одна из них (отсюда видно, куда должна развиваться теория).

Заключительный раздел обзора отвечает на вопрос, что в этой области известно твердо (в чем сходятся все наблюдения)? Чем это объясняется? Однозначно ли? На какие вопросы ответ противоречив? На какие вопросы вообще нет осмысленного ответа? Какие здесь есть догадки? Какие надо решить главные, ключевые вопросы на пути к решению всей проблемы? Каким методом можно надеяться получить на них однозначный ответ? Какова наша цель? Почему мы выбираем именно этот объект и метод исследования? Что мы хотим в конце концов узнать? «Формулировка пожеланий является первым шагом к их осуществлению»*.

3. МЕТОДИКА ЭКСПЕРИМЕНТА

ЦЕЛЬ И ПРИНЦИПИАЛЬНАЯ СХЕМА ИЗМЕРЕНИЙ

У меня, как говорится, очень развит талант к наблюдению, но только когда уже поздно и когда неприятность уже произошла.

Гашек. «Похождения бравого солдата Швейка»

Как можно большую часть измерений всегда стремятся выполнить на действующих стандартных установках, для которых наперед известны реальные оценки погрешностей и воспроизводимости (что, конечно, не освобождает нас от постоянного их контроля).

Проектирование новой установки, комплектация ее приборами, узлами и деталями, изготовление в цехе, отладка (и переделки) не только занимают много времени, но и вносят большую неопределенность в общие сроки исследования: во-первых, материальное обеспечение постройки не сводится к финансированию и зависит не только от нас; во-вторых, всегда есть риск, что качество работы установки окажется неприемлемо низким. Поэтому постройку затевают лишь при острой необходимости

сти: когда опробовать идею новой технологии нигде; для измерения новых свойств, наблюдения новых явлений; измерений или наблюдений в особых условиях, недоступных на существующих установках; в случае если нужна установка на порядок точнее или производительнее существующих; когда она одним прямым измерением заменяет сразу много косвенных. При этом «комбайн» для одновременного измерения многих свойств обычно оказывается «равномерно хуже», чем его специализированные прототипы.

Если необходимость постройки очевидна, то совместно прорабатывают принципиальную схему эксперимента и установку или приспособления для него.

Первый этап работы — четко установить цель эксперимента. «Общее доброе правило: приступая к задаче, попытаться ответить на вопрос: «Что мы собираемся делать с ответом?» — Активность и воображение могут дать многое для всего исследования уже на этой стадии, в том время как лень может помешать возникновению какого бы то ни было понимания, расходуя многие часы... для получения очевидных числовых результатов»*.

Всякий физический эксперимент сводится в конце концов к некоторым измерениям. Но полученные в результате цифры редко являются окончательной целью исследования. В большинстве случаев «цель расчетов — не числа, а понимание»*. Измерение нужно, чтобы доказать некоторый факт, отобрать главные факторы, влияющие на явление, найти закон, управляющий процессом. Поэтому прежде всего следует продумать: если измерения дадут ожидаемый результат, достаточно ли его для доказательства факта, для однозначного установления взаимосвязи явлений? Не допускает ли этот результат еще и несколько других истолкований? Какими экспериментами и оценками эти альтернативные истолкования можно отвергнуть? Не было ли все результатом не искомого эффекта, а недочетов методики?

Таким образом, планирование эксперимента начинается с обсуждения еще не полученного результата. Нужно четко представить, какие однозначные выводы удастся сделать, если будет получен ожидаемый результат. Нередко после такого анализа приходится не только пе-

* Хемминг Г. В. Численные методы для научных работников и инженеров. М.: Наука, 1972. 400 с.

* Миддлтон Д. Очерки теории связи. М.: Советское радио, 1966.

рассмотреть средства, но и заново сформулировать задачу исследования. Так составляют программу одного эксперимента (перечень той совокупности измерений, которая в случае успеха даст решение задачи).

Всякий, даже чисто качественный, эксперимент требует предварительного количественного анализа. Надо оценить: какого порядка величины ожидаемый эффект; какими средствами измерения можно его обнаружить (случайная погрешность измерений должна быть хотя бы на порядок меньше ожидаемого эффекта). И для чисто качественного обнаружения эффекта следует указать, какова вероятность получить те же результаты из-за случайных флуктуаций (помех). «Даже сильная вера должна быть подкреплена статистическим критерием»*.

Анализ начинается с перечисления побочных явлений, накладывающихся на исследуемый эффект. Составление перечня возможных источников ошибки — важнейшая стадия подготовки эксперимента. Все источники существенных погрешностей редко предвидит даже экспериментатор с большим опытом. Некоторые помехи обнаруживают и устраняют в процессе отладки схемы и самих измерений. Однако горький опыт учит, что тот, кто ставит эксперимент без тщательного анализа возможных ошибок, станет жертвой собственной самонадеянности. Ошибки всплывают не все сразу — одна за другой — и, если схема не продумана в деталях заранее, их можно устранять годами — до тех пор, пока некто более осмотрительный не достигнет той же цели своим путем.

Далее оценивают величину побочных эффектов. Если они сравнимы с основным, надо менять схему эксперимента, чтобы устранить или скомпенсировать помеху. Качество компенсаций также надо оценить; как правило, исключение одних погрешностей тут же вносит новые; необходимо знать, насколько они меньше старых. Вот примеры самых очевидных источников погрешности, важных почти в каждой схеме. При измерении температуры — неравномерность температуры по рабочему объему установки, по длине и сечению образца; непостоянство температуры холодных спаев термпары, а в радиационной пирометрии — изменение коэффициентов

* Феллер В. Введение в теорию вероятностей и ее приложения. Т. 2. М.: Наука, 1967, с. 101.

черноты металла и прозрачности окон со временем. При термостатировании: тепловая инерция печи и инерция датчиков, определяющие амплитуду автоколебаний температуры при автоматическом регулировании. При измерениях длины: тепловое расширение образца и элементов схемы от колебаний температуры окружающей среды, упругие деформации и люфты передаточных элементов. При измерениях силы: несоосность приложения нагрузки, силы и моменты трения в призмах, направляющих, шарикоподшипниках. При измерении электрических величин: нестабильность температуры элементов схемы, т.э.д.с. и гальванические э.д.с. в контактах, дрейф напряжения и тока питания схемы, утечки через изоляцию, наводки.

Предвидеть как можно больше источников ошибок и грубо оценить их с помощью простых физических моделей гораздо важнее, чем дать скрупулезный подсчет одной-двух ошибок (при более широком обзоре они могут и пропасть на фоне более мощных источников помех).

В предварительных расчетах неизбежно принимают на веру некоторые исходные допущения. Если от них существенно зависит точность и надежность результата, следует тут же продумать, как проверить справедливость этих допущений, когда установка будет готова.

Программу самого эксперимента всегда корректируют в ходе работы. Однако первую серию основных экспериментов надо спланировать заранее — при этом легко обнаружить недостатки разработанной схемы. Важно проверить, стабильны ли (в пределах заданной точности) результаты измерений, когда искомый эффект заведомо отсутствует («холостые» измерения). Надо решить:

а) если уже здесь окажутся нестабильные результаты, как можно, меняя условия измерений, найти причину нестабильности;

б) сколько одинаковых образцов понадобится, чтобы оценить воспроизводимость «рабочих» результатов и сравнить ее с априорной оценкой суммарной погрешности? Какую наибольшую ошибку воспроизводимости можно считать приемлемой?

в) как, изменяя условия эксперимента, искать причину плохой воспроизводимости? Что нужно для этого добавить в конструкции?

Проектирование установки требует обстоятельного знакомства с литературой. Прежде всего следует про*

смотреть самые капитальные монографии об исследуемом явлении и руководства по технике эксперимента. Затем — описания методики и ссылки к ним в своих конспектах для обзора. Наконец — разделы реферативных журналов и предметные указатели журналов по технике эксперимента за последние 2—3 года. Как правило, удастся найти ссылки на сходные работы и установки. Если даже заранее видно, что это не совсем то, до указанных в ссылках первоисточников следует добраться и внимательно прочесть. Затраты времени на литературу всегда окупаются. Даже если работы окончились неудачей, в этих статьях можно найти упоминания о таких осложнениях и помехах, которые нам пока и в голову не приходили. Если основной вопрос и не решен, у предшественников можно найти неожиданно простое и изящное решение частных задач. После внимательного ознакомления с работами предшественников нередко приходится расстаться и с основной идеей эксперимента — по этому пути уже кто-то ходил и встретил явно непреодолимые при сегодняшней технике трудности. Зато из других работ вытекают совсем другие пути решения, о которых мы не догадывались.

ПРОЕКТ УСТАНОВКИ

Это случилось после того, как взорвалась паровая машина, которую мы с Мишкой делали из консервной банки.

Носов. «Веселая семейка»

Пусть первая — труднейшая — часть проекта установки кончена. Обоснована идея эксперимента, его принципиальная схема. Оценками по порядку величины показано, что ожидаемый эффект можно наблюдать. Предусмотрены многие факторы, которые могут исказить результаты; указан способ их нейтрализации и необходимые нормы точности для измерительных средств и конструкции установки.

Надо приступать к рабочим чертежам. Как правило, на этой стадии вновь возникает ряд принципиальных вопросов; обнаруживаются, оцениваются и устраняются новые источники погрешности. Поэтому общий вид установки проработать во всех деталях обязан сам экспериментатор, даже когда в лаборатории есть профессиональный конструктор. Конструктор может принести

пользу в дальнейшем, упрощая конструкцию для удобства и надежности изготовления. Но никто не может лучше самого экспериментатора продумать все подробности будущей работы на установке и предвидеть возможные осложнения.

Первый чертеж установки делают сразу в масштабе, в трех проекциях — удобнее всего карандашом на миллиметровке. Начинают с изображения образца и места, куда должна вписаться установка (например, при размещении схемы в готовом термостате следует сразу нанести контуры внутреннего пространства термостата, имеющийся канал для ввода проводов и т.п.). Затем размещают основной активный элемент (измерительный датчик, узел передачи усилий, перемещений и т.п.). Когда действующие элементы готовы, изображают узлы их крепления, герметизации, несущий каркас и т.п. Затем чертят электрическую схему питания, вакуумную схему и т.п. (принципиальную схему электрических измерений или автоматического регулирования, конечно, прорабатывают еще при постановке задачи).

Способ изготовления каждой детали продумывают при изготовлении рабочего чертежа. Немыслимые кривоколенные глухие отверстия и микронные допуски появляются в проектах обычно не в силу необходимости. Если подумать, тот же результат дадут более простые и менее точно выполненные детали. Вместо одной высокоточной детали бывает выгоднее сделать из более простых деталей узел с точно регулируемым размером. Например, выдержать расстояние между двумя отверстиями в одной детали с точностью до 0,01 мм можно на прецизионном координатно-расточном станке. Однако можно перенести одно из отверстий в подвижную втулку и закрепить ее установочными винтами. С помощью дифференциального винта или клина втулку можно будет установить так же точно, хотя точность изготовления каждой из деталей много ниже.

Рискованно проектировать установку, где весь успех зависит от особо тонкой работы слесаря-лекальщика или сборщика. Когда есть сомнения, можно ли сделать такую деталь, надо советоваться непосредственно со стеклодувом, лекальщиком, токарем, наперед продумать допустимые варианты изменений. Полезно и просмотреть соответствующие пособия, чтобы не задавать глупых вопросов. Часто труднейшими для цеха оказываются вто-

ростепенные детали, совершенно не важные для работы конструкции. Поэтому в разговоре с исполнителем важно уяснить, что именно невыполнимо и почему. Со своей стороны, важно подчеркнуть, что безусловно обязательно по самому принципу действия установки, а что можно менять. Тогда тут же можно решить, как обойти трудности.

Что касается сборки, юстировки, регулировки, настройки механической, оптической, электронной схемы, с которой я один собираюсь работать, то здесь, несомненно, лучший критерий: «сумею ли сделать эту тонкую работу я сам?»

Надо рассчитывать на однократное изготовление установки в небольшом механическом цехе. Поэтому следует избегать применения отливок, поковок, штамповок, ограничиваться простейшей технологией сварки и пайки. Вместе с тем обычно широко используют готовые узлы серийных приборов (редукторы, лентопротяжные и пишущие механизмы, блоки питания, усилители, датчики), доступные как запчасти или при разборке устаревших и сломанных приборов.

По ходу проектирования надо делать заметки о способах сборки и юстировки. Для сборки, отладки и проверки схемы по частям могут понадобиться дополнительные разъемы, отверстия, базовые поверхности, монтажные крепежные детали, переключатели, выводы от промежуточных точек схемы. Их надо предусмотреть сразу. Окончив рабочий чертеж, надо шаг за шагом продумать всю последовательность операций сборки и наладки — как правило, при этом всплывает еще несколько недочетов. Устранить их на бумаге неизмеримо легче и быстрее, чем в металле. Общий вид установки должен содержать такую проработку всех деталей, чтобы не оставалось никаких сомнений при их детализовке. На общем виде наносят основные габаритные размеры и размеры, необходимые для присоединения других приборов и узлов, не входящих в задание. Там, где от этого зависит точность эксперимента или работоспособность конструкции, должен быть указан допуск на размеры.

Детализовку выгодно сделать с чернового (уже грязного) рабочего чертежа, а затем по детализовочным чертежам с проставленными размерами «собрать» окончательный общий вид установки. Для этого, не обращаясь к памяти и к черновику, надо пунктуально состав-

лять конструкцию только по размерам, указанным в детализовке. При этом обязательно обнаружатся неувязки в сопрягающихся размерах, «наложения» одной детали на другую, невозможность соединить или вставить одно в другое; иногда и не хватит некоторых деталей. Такая «контрольная сборка» на бумаге длится несколько больше, чем простое копирование общего вида нацисто, но экономит много труда рабочего и конструктора при выяснении и исправлении неувязок в цехе. Обязательна полная спецификация с указанием материала и (если надо) термообработки. Для всех стандартных комплектующих деталей и узлов (подшипники, реле, моторы, готовые приборы) следует давать обозначение (марку) и основную характеристику. Даже если этого не требует руководство, все расчеты и соображения к проекту стоит свести в пояснительную записку. Ее разделы:

1. Постановка задачи: логические соображения, показывающие, какие выводы можно будет сделать из планируемых наблюдений.

2. Количественная оценка ожидаемого эффекта, выбор средства измерения. Качественный анализ источников погрешности. Их количественные оценки. Выбор методов устранения помех и их количественные оценки.

3. Конструктивные расчеты (кинематические соотношения, потребные мощности, проверка на силовые, тепловые, электрические нагрузки и т. п.).

4. Указания по сборке и наладке схемы, если они не ясны из чертежа.

5. Перечень холостых и поверочных экспериментов, демонстрирующих работоспособность схемы и ее точность.

6. Программа первой серии основных экспериментов. Обоснование выбора числа образцов «на точку», числа точек, их размещения.

7. Список использованной литературы (включая все источники, откуда взяты формулы и значения констант).

Последовательность размышлений и оценок на стадии разработки простого приспособления иллюстрируется ниже. Чтобы не мешать пониманию большим числом известных формул (интересных только для данного конкретного случая), сами расчеты здесь не приведены.

Ссылки на публикации по существу задачи не приведены, но обозначены (Л), чтобы показать — что имен-

но пришлось разыскивать в справочниках и периодике. Поиски соответствующих формул и констант составляют, хотя и трудоемкую, но не главную часть работы. Гораздо важнее решить, что именно надо рассчитывать, чтобы, предвидя реальный ход эксперимента, количественно сформулировать конкретные требования к его схеме и методике, необходимые и достаточные для достижения поставленной цели.

Задача эксперимента: доказать, что фазовый наклеп возможен при высокотемпературном (выше $0,5 T_{пл}$) полиморфном превращении. Объемный эффект полиморфного превращения должен вызывать пластическую деформацию и сопутствующие ей изменения дислокационной структуры при любой температуре превращения. Однако при высокой температуре возврат или рекристаллизация снимут следы наклепа раньше, чем их удастся обнаружить.

Чтобы выяснить, есть ли наклеп сразу же после превращения (длящегося доли секунды) [Л], нужно достаточно резко «заморозить» полученную структуру и исследовать ее при низкой температуре. Для наблюдения выбрано $\alpha \rightarrow \gamma$ -превращение при нагреве железа. Было известно, что увеличение скорости нагрева до 1000 К/с меняет дислокационную структуру мартенсита при 20°C [Л]. Но эти изменения можно связать и не прямо с фазовым наклепом аустенита, а с изменениями механизма последующего $\gamma \rightarrow \alpha$ -превращения при охлаждении. Факт фазового наклепа будет доказан, если удастся обнаружить достоверные различия между дислокационным строением аустенита сразу после нагрева со скоростью 1000 К/с и после выдержки для снятия фазового наклепа.

Исследовать структуру надо быстро: при высокой температуре наклеп быстро снимается, при низкой — распадается аустенит. Сами значения количественных характеристик не нужны — достаточно доказать достоверное различие в какой-либо из качественных характеристик субзеренной структуры. Из них наиболее доступно для измерения физическое уширение линий рентгенограмм поликристалла или (при идентичной геометрии съемки) просто ширина линий. Отсюда идея опыта: сравнить ширину дифракционной линии аустенита сразу после быстрого нагрева до температуры T_1 (выше точки A_3) и после выдержки, достаточной для снятия фазового наклепа.

Ожидаемое уширение линии неизвестно. Фазовый наклеп в мартенситном превращении приводит к уширению линии в несколько раз [Л], а умеренная пластическая деформация — на 20–50 % [Л]. Среднеквадратичная ошибка воспроизводимости при многократной записи сильных линий составляет 8–10 % [Л], поэтому есть надежда обнаружить фазовый наклеп по изменению ширины линий аустенита. Для съемки нужна сильная линия аустенита, свободная от наложения линий феррита (иначе наблюдаемая ширина начнет меняться вследствие распада еще до того, как мы обнаружим его по появлению линии феррита). Такая линия $(311)_\gamma$. Если придется проверять, не появился ли за время съемки при температуре T_2 феррит, нужно снимать фон на месте сильнейшей линии α -фазы, свободной от наложения γ . Это линия $(211)_\alpha$. Цель съемки — возможно более точное измерение ширины довольно сильных линий. Чтобы контролировать идентичность геометрии съемки, желательно регистри-

ровать также линию α -фазы до нагрева (отметим для программы). Это гораздо легче на дифрактометре, чем при фоторегистрации.

Теперь сформулируем требования к сплаву. Прямых измерений скорости снятия фазового наклепа не было. О порядке величины ее можно судить по скорости рекристаллизации пластически деформированного аустенита: при умеренном легировании она завершается за доли минуты при 900°C [Л]. Значит, «свежий» аустенит при температуре T_1 (около точки A_3) надо исследовать за несколько секунд. Можно снизить точку A_3 сильным легированием никелем и марганцем, но это нецелесообразно: появится много остаточного аустенита в исходном состоянии и соответственно снизится наблюдаемый эффект наклепа при нагреве. Значит, надо либо снять рентгенограмму за несколько секунд, либо быстро переохладить аустенит и снимать его в нижнем интервале устойчивости (над мартенситной точкой).

Для скоростной съемки рентгенограммы на пленку надо использовать мощную импульсную рентгеновскую трубку, но с ней трудно получить идентичную геометрию съемки разных образцов (в отличие от обычных установок камеру нельзя заранее «выставить по пучку»). Нельзя и получить усреднение по многим зернам, так как вращающийся образец за время импульса освещается практически в одной точке. Линия умеренной ширины отличается от фона в интервале менее 2°, и если не требовать точного воспроизведения профиля, то при скорости счетчика 2–4 град/мин ее можно записать на серийном дифрактометре за 0,5–1 мин. Поэтому принимаем вариант с дифрактометрической съемкой переохлажденного аустенита.

Устойчивость аустенита при выбранной по диаграмме распада [Л] температуре T_2 проверим по времени появления линии $(211)_\alpha$ феррита при изотермической выдержке.

Снимаем линию $(211)_\alpha$ при нагреве феррита ниже A_1 , убедимся, что сама по себе выдержка при T_2 не меняет профиль линии и не снижает сильно ее интенсивность из-за непредвиденных деформаций или смещений конструкции и образца или из-за окисления поверхности. Не имея данных о температурной зависимости скорости возврата и рекристаллизации, можно лишь утверждать, что они идут при охлаждении не быстрее, чем при T_1 . Тогда время охлаждения до температуры съемки T_2 должно быть не более 5–10 с. Для съемки желательно, чтобы аустенит не распадался при T_2 возможно дольше. Но не стоит добиваться этого, вводя углерод: неоднородное распределение углерода при растворении карбидов может дать свое уширение линий, не связанное с фазовым наклепом. По имеющимся диаграммам изотермического распада из числа безуглеродистых сплавов с полным превращением $\alpha \rightleftharpoons \gamma$ наибольшую устойчивость имеют сплавы с 5–10 % Cr [Л]. Выбран сплав с 8 % Cr; аустенит в нем начинает распадаться через 12–15 мин при $T_2 = 510^\circ\text{C}$ [Л]. Изотермическая диаграмма его превращения показывает, что принятая по условиям рекристаллизации скорость достаточно, чтобы предотвратить распад в процессе охлаждения, но «запас» невелик, а С-образные диаграммы могут зависеть и от индивидуальных особенностей плавок. Поэтому запишем замечание: проверить, не появляется ли в процессе выдержки при T_2 α -фаза (иначе возможен фазовый наклеп от превращения при выдержке).

Выберем среду нагрева. На воздухе окисление образца во время съемки будет постепенно ослаблять интенсивность, а для вакуу-

ма или защитной атмосферы придется герметизировать камеру и, значит, пропускать рентгеновский пучок через два окна. Вполне надежны бериллиевые окна толщиной 0,8 мм (как в рентгеновских трубках [Л]). Более тонкие могут оказаться негерметичными. Чтобы получить возможно более широкие линии, нужно мягкое излучение хрома, кобальта, железа. По линейному коэффициенту поглощения [Л] находим, что два бериллиевых окна ослабят K_{α} -излучение почти в 4 раза. В то же время по кривым окисления сплавов Fe—Cr на воздухе [Л] находим, что привес за 1 мин при 900°C составляет около 0,1 г/м², а за 30 мин при 500°C—много меньше этого. Если весь окисел представляет собой Fe₂O₃, то по количеству кислорода и плотности окисла [Л] находим, что пленка окисла — не толще 0,1 мкм, тогда как дифракционную картину дает слой толщиной порядка 10 мкм [Л]. Поэтому можно пренебречь влиянием окислы и делать съемку на воздухе.

Чем больше интенсивность, тем меньше относительная погрешность от пуассоновых флуктуаций пучка. Чем меньше инструментальная ширина, тем легче обнаружить физическое уширение. Поэтому подберем возможно более узкие щели, а ток и напряжение трубки зададим, так, чтобы интенсивность в максимуме линии была чуть ниже предела регистрации (500 имп/с).

Съемка «по точкам» всегда точнее, чем прямая запись профиля линии. Однако за имеющиеся в нашем распоряжении 10—20 мин надо снять линию несколько раз и убедиться, что она не менялась за это время. Для анализа профиля линии нужно не менее 15—20 точек, и тогда останется всего по 10—20 с на точку—большая часть времени будет затрачена на холостое перемещение счетчика от точки к точке. Поэтому решаем пользоваться непрерывной записью. Ширину линий при скорости счетчика 2 град/мин за 10 мин можно записать до 8—10 раз. Надо проверить, не снимается ли фазовый наклеп в процессе съемки при T_2 —нет ли систематического изменения ширины V в каждой серии линий (311), снятой с одного образца.

Нагревать образец со скоростью 1000 К/с удобнее всего прямым пропусканием тока. Для этого образец должен представлять собой ленту постоянного сечения. Верхняя граница толщины ленты определяется допустимой скоростью свободного охлаждения от T_1 до T_2 на воздухе (по диаграмме изотермического распада надо не менее 30—50 К/с). Время охлаждения оценим, зная теплопроводность металла и коэффициент теплоотдачи в спокойном воздухе [Л]. Достаточной оказывается толщина образца около 0,5 мм (более тонкая лента обычно менее однородна по толщине).

Ширина образца должна быть не менее 5—7 мм, чтобы получать достаточную интенсивность отражения. Чтобы профиль линии не искажался из-за малого числа освещаемых зерен, обычно образец вращают; тогда за время менее постоянной времени интегратора (2 с) под пучок попадает площадь около 1 см². Но чтобы вращать ленточный образец, нужен щеточный токоподвод; это громоздко и ненадежно. Выгоднее сообщить ленте возвратно-поступательное движение в собственной плоскости. Тогда нужна амплитуда около 1 см и период не более 2 с. Для этого используем микромотор на 60 об/мин. Вставив в шкив штифт, получим кривошип, а для качания образца поставим шатун.

Из диаграммы изотермического распада видно, что область достаточной устойчивости аустенита над мартенситной точкой имеет

ширину около 50 К, так что колебания температуры около T_2 не должны превышать ± 25 К. С некоторым запасом назовем допуск ± 15 К. Из них ± 10 К составит допуск на регулирование T_2 , а ± 5 К—на перепад температуры по площади съемки.

При температуре T_2 в средней точке образца на длине 1 см перепад ΔT не должен превышать 10 К. Отсюда, предполагая (с запасом), что массивные клеммы имеют температуру 20°C, из известного решения задачи теплопроводности неизолированного стержня, нагреваемого током [Л], находим стационарное распределение температуры и минимальную допустимую длину ленты между клеммами 60 мм. Добавив по 1—1,5 см на зажим, получим длину образца (более длинный образец трудно разместить в габаритах установки). Отметим также, что при удалении от средней точки образца перепад ΔT растет, поэтому обязательно устанавливать образец так, чтобы в среднем положении движущегося образца его середина находилась как раз под пучком.

Размеры образца установлены. Теперь нужно выбрать схему нагрева, измерения и регулирования температуры. Из расчета стационарного теплообмена [Л] выберем (с запасом) такой максимальный ток, чтобы задавать образцу любую температуру до 1500°C. Отсюда по удельному сопротивлению при наибольшей температуре [Л] найдем напряжение на образце. По допустимой плотности тока [Л] 3 А/см² выбираем сечение медного кабеля, по его сопротивлению и длине (ориентировочно 6—8 м)—падение напряжения на нем, необходимое напряжение (до 20 В) и мощность источника питания (до 2 кВт). Выбираем для нагрева переменный ток 50 Гц—при больших токах (до 1000 А) его проще регулировать и переключать, чем постоянный. Мощность и напряжение выходного трансформатора теперь известны, но выпускаемые трансформаторы с плавной регулировкой не рассчитаны на нужное нам сравнительно низкое напряжение и большие токи. Поэтому перебрав имеющиеся в наличии трансформаторы, выбираем более сложную схему: секционированный трансформатор, понижающий с 220 В на 7—56 В на выходе; его вход питается от автотрансформатора с плавной регулировкой (РНО).

При $T_2=500$ °C радиационные (инфракрасные) пирометры ненадежны, так как коэффициент черноты образца зависит от характера окислы. Более надежны привариваемые термопары. Их диаметр должен быть во всяком случае на порядок меньше толщины образца, иначе они будут измерять только температуру холодного пятна, получающегося в месте приварки. Более точные оценки погрешности измерения от подстуживания из-за теплоотвода в термопару в данном случае можно не делать, поскольку важно не абсолютное значение температуры съемки T_2 , а только ее стабильность (устойчивость аустенита при T_2 все равно придется проверить). По той же причине мы не контролируем температуру холодных спаев термопары. Ажурные нити приваренной термопары должны быть возможно короче (4—6 см), чтобы не запутать их или не оборвать. Поэтому надо дать им достаточно массивные медные клеммы, чтобы холодный спай не прогревался излучением образца и температура T_2 не «сползала» (достаточно будет проверить клеммы на ощупь по окончании выдержки при T_2 —разницу между 20 и 30°C чувствует рука).

Температуру наибольшей устойчивости аустенита T_2 найдем, построив диаграмму изотермического распада рентгеновским методом,

Выберем щель у счетчика настолько широкой, чтобы она захватывала всю линию. Установим неподвижный счетчик на линии (211) α и будем записывать ее появление и усиление за время изотермической выдержки при разных T_2 .

Выберем регистрирующий и регулирующий прибор. Самопишущий электронный потенциометр имеет класс 0,5. Прямая ошибка регистрации составит 5 К, если вся шкала около 1000 К. Это приемлемо. При скорости нагрева 1000 К/с перо должно пробегать всю шкалу за 1 с. Поэтому пригоден только потенциометр с временем пробега каретки 1 с. При таком быстроедействии показания потенциометра все равно сильно отстают от истинной температуры: лишь через 5 с после нагрева со скоростью 1000 К/с отставание станет менее 0,5 % [Л]. Поэтому температуру отключения при нагреве придется подбирать «с упреждением», чтобы компенсировать инерцию прибора. Конечную температуру нагрева желательно задать как можно ниже (чтобы дольше сохранялся фазовый наклеп), но все-таки на 30–50° выше точки A_3 , что гарантирует превращение при любом срабатывании отключения.

Теперь рассмотрим способ поддержания температуры изотермы T_2 . Примем размах этих колебаний ± 10 К. Такой размах можно попытаться обеспечить контактором, без применения следящей схемы. При скорости охлаждения 50 К/с образец будет остывать на $\Delta T = -10$ К за 0,2 с. При однопозиционном регулировании (включено—выключено) контактор будет работать непрерывно и может давать сбой. Чтобы контакторы срабатывали реже, образец должен остывать медленнее. Придется ввести двухпозиционное регулирование—питать силовой трансформатор попеременно с двух отводов регулируемого автотрансформатора, на которых подобрать напряжение несколько больше и несколько меньше, чем надо для поддержания изотермы T_2 .

Если потенциометр нужного быстроедействия не имеет регулирующего контактного устройства, придется сделать контакт, управляемый стрелкой потенциометра. Размыкать им непосредственно катушку контактора нельзя: из-за большой индуктивности катушки контакты станут искрить, что ведет к ложным срабатываниям. Нужно промежуточное реле, с бесконтактным включением,—например, от фотосопротивления, заслоняемого стрелкой потенциометра (целесообразно поискать готовую схему включения силовой цепи от светового сигнала).

Осталось рассмотреть механическую часть—узел крепления образца с токоподводами и термодарным вводом. Главное требование, определяющее точность съемки: образец должен лежать в осевой плоскости гониометра. Смещение из этой плоскости на 0,1 мм смещает линию на 4' [Л]. Хотя знать период решетки нам и не нужно, вряд ли можно допустить смещение линии больше чем на 10–15', в противном случае «хвосты» могут выйти за пределы выбранного интервала съемки. Потребуем (с запасом), чтобы смещение из осевой плоскости не было более 0,1 мм.

Осевая плоскость задана в пространстве положением опорных бобышек столика гониометра. Устанавливаемое на гониометр приспособление должно иметь хорошо шлифованную базовую плоскость, которая ложится на эти бобышки. Чтобы поверхность образца легла точно в базовой плоскости, контакты, подводящие к нему ток, должны быть параллельны этой плоскости. Толщина образца меняется редко, поэтому вместо приспособления для плавного совме-

щения образца с базовой плоскостью можно просто зажать под контактом пачку пластинок (фольги), чтобы регулировать уровень. Надо решить, как проверять совпадение плоскостей. Зазор между плоскостями 0,05–0,10 мм хорошо виден «на свет», поэтому достаточно предусмотреть доступ к базовой плоскости, чтобы, прикладывая в разных направлениях металлический угольник, проверять совмещение плоскостей.

Базовую плиту придется жестко скрепить со столиком гониометра. Но образец должен иметь возвратно-поступательное движение. Значит, контакты двигаются по плите, и при этом поперечные колебания их не превышают 0,1 мм. В направляющих скольжения сильное трение. Сделаем направляющие качения—треугольные в сечении пазы, заполненные шариками. Чтобы шарики не сбегались в один конец паза, поставим сепаратор из тонкой латунной ленты. Важно, чтобы зазор в направляющих был много меньше 0,1 мм, иначе каретка будет качаться в поперечном направлении. Если делать пазы непосредственно в базовой плите, придется шлифовать их с точностью до 0,01 мм. Это хотя и возможно, но трудно. Проще шлифовать направляющие планки отдельно и прикрепить их к плите, предусмотрев регулировку (например, сделать широкие отверстия под крепеж и перед затяжкой винтов подгонять зазор легким постукиванием).

Образец придется нагревать. Если расстояние между его зажимами жестко фиксировано, ленточный образец из-за теплового расширения вспучится и выйдет из базовой плоскости. По температурному коэффициенту расширения [Л] можно найти увеличение длины 0,1 мм, а по длине образующейся при нагреве дуги—стрелу прогиба 0,4 мм, что неприемлемо велико. Значит, один из контактов должен при нагреве смещаться по отношению к другому. Оба контакта уже поставлены на шариковые опоры, но вряд ли они сами будут легко расходиться—тонкий образец скорее погнется, чем раздвинет довольно массивные контакты. Придется ставить пружину, раздвигающую контакты при нагреве. Но слишком сильная пружина вызовет ползучесть образца. Пластическая деформация менее 1 % не повлияет заметно на уширение линий [Л]. Исходя из данных о температурной зависимости ползучести железа [Л], можно оценить, при каких напряжениях деформация за время нагрева будет менее 1 %. По напряжению и сечению образца найдем допустимое усилие пружины и ее размеры. Этот расчет даст лишь порядок величины и надо предусмотреть контроль реально получающегося остаточного удлинения после съемки каждого образца. Оно не должно превышать 1 %.

Натяжение пружины раздвигает свободно стоящие захваты, и образец останется плоским при нагреве и охлаждении. При съемке кривошипно-шатунный механизм задает возвратно-поступательное движение образца. Если связать шатун только с одним из захватов, то при ходе «на сжатие» образец согнется. Значит, по достижении температуры съемки захваты надо скреплять быстроедействующим зажимом: проще всего—заклинивать штифтом проходящую через оба зажима зубчатую рейку (винт).

Требования к механической схеме приспособления сформулированы. Осталось продумать, вполне ли сочетаются электрическая и механическая схемы между собой и с основной (рентгеновской) установкой.

Силовой токоподвод—тяжелый и жесткий кабель. Поэтому не

стоит крепить кабели прямо к подвижной каретке, лучше подать их на неподвижные клеммы на базовой плите, а для соединения с зажимами образца использовать голый многожильный гибкий провод или пакет медной фольги.

И клеммы, и зажимы следует, конечно, изолировать от корпуса (слюдяными и текстолитовыми прокладками), чтобы цепь нагрева не замкнулась на корпус. Качество изоляции надо проверить мегомметром после сборки узла и еще раз — перед пуском установки.

Гониометр есть, по существу, прецизионный редуктор со строго выверенной осью столика. Нагрузка на столик должна быть как можно меньшей, а установка на столик нашего приспособления не должна заметно увеличить вращающий момент привода гониометра. Поэтому все приспособление надо не ставить на столик, а подвесить над ним на штанге. Штангу повесим на упорном подшипнике, а подшипник — на кронштейне. Большой П-образный кронштейн прикрепим не к гониометру, а прямо к корпусу рентгеновской установки. Сверху по оси вращения спустим и кабели — так они меньше мешают вращению.

Чтобы совмещать оси подвески и гониометра, надо оставить две степени свободы для горизонтального движения подшипника по кронштейну. Совместив оси, можно скрепить подвеску с гониометром (предусмотреть — чем). Возникающую от плохой соосности перегрузку привода проще всего проверить, вращая столик от руки.

Кронштейн не должен мешать вращению счетчика в любом положении, а кабели при этом не должны заслонять падающий и отраженный пучок.

Термопара приварена к движущемуся образцу. Чтобы не оборвать тонкие нити, надо, чтобы и термопарные клеммы двигались вместе с образцом — поэтому прикрепим планку с клеммами к каретке. От нее к потенциометру будет идти тонкий кабель с заземленной экранирующей оплеткой (прокладывать его придется параллельно с силовым кабелем, и без заземления экрана будут сильные наводки).

Схема эксперимента и конструкция приспособления в принципе ясны — составим программу первой серии экспериментов. Для этого надо выбрать возможно более экономно переменные в условиях эксперимента (факторы), их значения (уровни), их сочетания в каждом опыте (точки) и число повторных опытов «на точку».

В основном эксперименте достаточно ограничиться одной температурой T_1 , двумя крайними скоростями (наибольшая ограничивается схемой нагрева, наименьшая — искажением и ослаблением линии от окисления) и менять только время выдержки при T_1 . Начнем с крайних случаев: а) скоростной нагрев без выдержки при T_1 ; б) то же, но с наибольшей выдержкой (1—2 мин); в) медленный нагрев.

По всем записям линии (311), при T_2 (их можно получить около 10 на одном образце) найдем среднюю ширину линии $\bar{B}$ и ее дисперсию s_B^2 . Оценим значимость различия средних по режимам (а) и (б). Сделав 3—4 образца по каждому режиму, вычислим средние $\bar{B}$ и s_B^2 по каждой группе образцов. Если разброс величины $\bar{B}$ внутри группы статистически значимо превышает разброс отсчетов $\bar{B}$ для одного образца, придется подключить к термопаре шлейфовый осциллограф и на серии образцов проверить, нет ли большого раз-

броса скоростей нагрева. Если различие между образцами внутри группы статистически незначимое, а различие между режимами (а) и (б) — значимое, основной результат достигнут.

Чтобы оценить время снятия фазового наклепа, сравним режимы (а) и (б), а после этого назначим и другие выдержки: короче или длиннее (б) — в зависимости от того, отличается ли (б) от (а).

Если же режимы (а) и (а) при данном числе образцов и записей оказались неразличимы, то, зная дисперсию B , рассчитаем, сколько снять образцов, чтобы с достоверностью не ниже 90—95 % обнаружить изменение величины B на 5—10 %. Если при меньшем уширении считать, что фазового наклепа нет, надо оценить, ниже какой величины плотность дислокаций (или хотя бы степень деформации). Во всяком эксперименте нужно заранее знать, что можно будет утверждать, если искомый эффект окажется ниже границы обнаружения.

«СТАНДАРТНЫЕ» ИЗМЕРЕНИЯ

— Сама меряет, — сказал молодой человек, передавая астролябию покупателю, — было бы что мерять.

Ильф и Петров. «Двенадцать стульев»

Большинство лабораторий (металлографические, рентгеновские и т. д.) ныне организовано по принципу общности методики, а не задач. Это обеспечивает квалифицированное обслуживание оборудования, его полную загрузку, гарантирует определенный профессиональный уровень измерений. Обратная сторона медали: появляются «собственно исследовательские» группы, где люди знают цель и раздают заказы на измерения, но плохо владеют методикой. В чисто промышленном эксперименте это неизбежное зло: собственно опыты делает цех, результаты выдают лаборатории, задача автора идеи — заставить всех работать, ничего не перепутать и выяснить, что из этого вышло. Успех такого исследователя-организатора («диспетчера») возможен, если он использует только стандартную, хорошо отлаженную в данной лаборатории технику эксперимента (не ставит задач: «измерь мне это — придумай, как»), владеет всей основной методикой (сам в свое время делал), сам наблюдает и регистрирует все основные эксперименты и, наконец, имеет разветвленную программу контрольных опытов для перекрестной проверки воспроизводимости условий в цехе и результатов в лаборатории.

Даже при определении состава мы доверяем химикам в стандартных анализах (подбрасывая все же под дру-

гими номерами дубль-образцы для ответственных точек), но вынуждены вникать в существо их методики, когда содержание примеси на нижнем пределе измерения или анализ на нее делают редко. Что же касается структуры и свойств, то здесь всякое отступление от стандарта требует совместной работы ведущего эксперимент и «владельца техники» еще при постановке задачи и разработке методики. Иначе технологические результаты будут ненадежны, а научные и вовсе бессмысленны.

Некоторые массовые измерения делают автоматизированные установки: многоканальные спектрофотометры, электронные и рентгеновские микроанализаторы, мессбауэровские спектрометры, анализаторы изображений и т. д. Встроенная в них ЭВМ по вложенной фирмой программе обрабатывает отсчеты, вносит поправки на условия измерений и выводит на печать или экран сразу конечный результат: число, график и даже оценку ошибки. Пробившись на такую установку, экспериментатор вовсе не избавлен от анализа ошибок. Напротив, приходится довольно долго выяснять, какие именно поправки и ошибки учитывает фирменная программа, что в ней не учли составители, а что — мы сами, поставив не предусмотренную фирмой задачу. Указываемая в паспорте погрешность анализа обычно ниже фактической и не включает наиболее сильных — систематических ошибок (дрейфа режимов и калибровки, влияния приготовления и ввода пробы). Надежный контроль воспроизводимости и стабильности результатов в автоматизированных установках требует постоянного внимания и даже изобретательности экспериментатора (не одного только оператора прибора).

Так, например, в полностью автоматизированном анализаторе микроструктуры от личного решения оператора зависит выбор степени контраста изображения, выбор степени черноты, по которой частица отличается от матрицы, способа подавления дефектов шлифа, выбор схемы промеров, диапазона и шага разрядов для классификации размеров, величины полей зрения, их числа и схемы размещения и, наконец, возможность изменения всего этого от места к месту и от шлифа к шлифу. Подгонка «наилучших» условий для каждой картинки делает результаты совершенно несравнимыми. При абсолютном же постоянстве условий для половин образцов результат будет заведомо испорчен.

Гибкая тактика заключается в тщательном подборе неизменных условий, пригодных для одной серии образцов, и выборе эталонного образца «для привязки» при смене условий измерения. Но для этого надо понимать и установку (что может изменять режимы ее работы?), и суть различия между образцами, и разницу между главным и второстепенным в них (что можно принести в жертву в данном эксперименте). Поэтому осмысленные результаты дает лишь работа вдвоем с оператором.

МАТЕМАТИЧЕСКОЕ ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТА

Если существует стандартный метод измерения исковой величины y с постоянной воспроизводимостью, известен набор влияющих на нее факторов $x_1, x_2, \dots, x_n$ (вектор аргумента $\vec{x}$) и определена интересующая нас область существования аргумента, возможно определение оптимального числа, условий, последовательности и группировки однотипных единичных экспериментов методами математической статистики. Критерий оптимальности — наименьшее число измерений для описания зависимости $y(\vec{x})$ с заданной надежностью (или максимум надежности при данном объеме измерений). Объем определяется либо прямо числом или временем измерений, либо через некоторую *функцию трудности* — зависимость затрат на измерения от координаты $\vec{x}$ или величины y при неизменной точности.

Планирование эксперимента наиболее эффективно, когда факторов x_i много (задача многомерная), а вид зависимости $y(\vec{x})$ — простой (или допускает безболезненное упрощение). Такая ситуация обычна для технологических экспериментов. Полная задача: определить вид зависимости $y(\vec{x})$. Частные задачи: отсеять факторы x_i , влияние которых не превышает ошибки эксперимента, или найти точку экстремума $y(\vec{x})$.

В полной задаче *поверхность отклика* $y(\vec{x})$ обычно ищут в виде полинома. Например, в трехмерном пространстве $\{\vec{x}\}$

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_1 x_2 + b_5 x_1 x_3 + b_6 x_2 x_3 + b_7 x_1 x_2 x_3 + b_8 x_1^2 + \dots \quad (3.1)$$

отклик складывается из линейного влияния каждого аргумента x_i , их «перекрестных» парных воздействий (x_i, x_l) , тройного взаимодействия x_1, x_2, x_3 , затем — квадратичных эффектов (x_i^2) и т. д.

Если известно, что для описания $y(\vec{x})$ достаточно в полиноме (3.1) оставить n данных (обычно — «первых») членов, то измерение y в $N=n$ любых точках x_k в принципе достаточно, чтобы из N линейных относительно b_j уравнений (3.1) найти все коэффициенты b_j , т. е. описать всю зависимость. Однако желательно эти N точек отсчета выбрать так, чтобы удовлетворить и дополнительные требования. Во-первых, требование ортогональности: ошибки найденных b_j должны быть не связаны (статистически независимы) — если все b_j будут отклоняться от своих истинных значений согласованно, то суммарная ошибка поверхности отклика $y(\vec{x})$ может быть катастрофической. Во-вторых, хорошо иметь «лишние» точки ($N > n$), чтобы по их отклонению от поверхности отклика судить о приемлемости уравнения (3.1) («адекватности модели»). В-третьих, план должен быть оптимальным (ошибка окончательного результата при данном N наименьшей). Поскольку ищут сразу несколько величин b_j , возможны разные критерии оптимальности, например, минимум наибольшей ошибки предсказания y во всей области (G — оптимальность), минимум наибольшей из ошибок коэффициентов b_j (E -оптимальность), минимум суммы дисперсий всех коэффициентов $\sum_{j=1}^n s_j^2$ (A -оптимальность) или мини-

мум произведения всех дисперсий b_j : $\prod_{j=1}^n s_j^2$; (D — оптимальность, означающая максимум информации). Обычно добиваются D -оптимального размещения наименьшего возможного числа точек.

Если область наблюдений по всем координатам ограничена плоскостями $A_i \leq x_i \leq B_i$, то преобразованием $x' = [2x - (A + B)] / (B - A)$ она отображается в единичный куб $|x'| \leq 1$. Если каждому из m аргументов (факторов) x_i задавать по k значений (уровней), то перебор всех их сочетаний — *полный факторный план* — содержит $N = k^m$ точек наблюдения. При двух уровнях ($x_i = \pm 1$) и трех факторах $N = 2^3 = 8$ точек (в восьми вершинах ку-

ба $\langle 111 \rangle$). Но если ищут линейную зависимость $y(x_1, x_2, x_3)$, то в ней всего четыре коэффициента b_j , и достаточно четырех точек отсчета, т. е. можно оставить от плана 2^3 его половину — *дробную реплику*.

Долгое время считали очевидным, что влияние фактора x_1 (коэффициент b_1) надо искать, меняя только x_1 , при постоянных остальных факторах: $b_1 = \frac{1}{2} [y(1, 0, 0) - y(-1, 0, 0)]$. Если изменять факторы x_i поодиночке, каждое b_j будет найдено из двух отсчетов y . Если дисперсия каждого отсчета s_y^2 , то дисперсия каждого коэффициента b_j $s_b^2 = s_y^2/2$. Существует, однако, более экономный (D -оптимальный) способ: надо измерить y в четырех вершинах куба, образующих правильный тетраэдр, и определить сразу три b_j (при m факторах — надо соответственно мерить y в вершинах правильного m -мерного многогранника с $m+1$ вершиной — *симплекса*). Тогда дисперсия $s_b^2 = s_y^2/(m+1)$, т. е. при равном числе измерений получается выигрыш в точности в $\sqrt{(m+1)/2}$ раз потому, что для определения каждого b_j использованы, например, при $m=3$ все четыре отсчета y вместо двух. Так:

$$b_1 = \frac{-y(-1, -1, 1) + y(1, -1, -1) - y(-1, 1, -1) + y(1, 1, 1)}{4},$$

а точки отсчета расставлены по диагоналям куба. В m -мерном кубе диагональ в $\sqrt{m}$ раз длиннее ребра; соответственно точнее находится и «наклон» $y(x_j)$, т. е. коэффициент b_j .

Когда число точек N равно числу n искоемых коэффициентов b_j , «план насыщен». Чтобы судить, насколько хорошо найденные коэффициенты описывают реальную зависимость $y(\vec{x})$, нужен *ненасыщенный план* ($N > n$). Так, добавив к симплексу еще одну точку в центре куба $[000]$, мы прямо измерим величину $y(0, 0, 0) = b_0$ и можем сравнить ее с b_0 , вычисленным по всем остальным y . Лишние степени свободы ($N - n > 0$) позволяют проверить, отражает ли уравнение (3.1) при выбранном n действительный вид поверхности $y(\vec{x})$.

Если уравнение «не адекватно», в него надо включить дальнейшие члены — перекрестные или квадратичные, добавив точки отсчета. Чтобы менять план по ходу из-

мерений, надо заранее иметь такой *композиционный план*, который будет *D*-оптимальным и с «малым», и с «большим» набором точек. Существуют таблицы *D*-оптимальных планов на разное число факторов и уровней, в том числе и композиционных [3].

Иногда сама поверхность отклика $y(x)$ и не имеет физического смысла, а задача эксперимента — выяснить, какие из факторов x_j влияют на величину y (т. е. какие $b_j \neq 0$). При этом все или часть факторов могут быть качественными (да и нет — 0 и 1) — например, $x_1 = 1$ — с обработкой синтетическим шлаком, а $x_1 = 0$ — без обработки. *D*-оптимальный план с дробными репликами эффективен и в этих случаях, где цель — не поиск уравнения, а лишь выделение значимых факторов по величине невязки — «остаточной дисперсии» (дисперсионный анализ [2]).

Бывают факторы — помехи: их влияние надо не изучать, а исключить (например, разницу в шихтовом материале, если одной его партии заведомо не хватит на все плавки). Тогда цель плана — так расставить факторы помехи по вариантам эксперимента, чтобы неоднородность равномерно перешла в случайную ошибку, не искажая вида искомой зависимости $y(x)$. Так, если зависимость y от двух переменных x_1, x_2 при трех уровнях каждой осложнена использованием разнородного исходного материала (варианты *A, B, C*), то систематическую ошибку устраняет *ортогональный план* в виде *латинского квадрата*:

Латинский

$x_1 =$	1	2	3
$x_2 = 1$	A	B	C
2	B	C	A
3	C	A	B

Греко-латинский

$x_1 \rightarrow$			
$x_2 \downarrow$	A α	B γ	C β
	B β	C α	A γ
	C γ	A β	B α

Так, если надо выяснить влияние и формы слитка ($x_1 = 1, 2, 3$) и смазки изложницы ($x_2 = 1, 2, 3$), а из одной плавки выходит всего три слитка, то латинский квадрат задает такое сочетание x_1 и x_2 для каждой из плавков (*A, B, C*), чтобы влияние этих факторов не искажалось,

например, разницей в температуре разливки или колебаниях состава между плавками. Аналогично размещение двух факторов помех (*A, B, C* и α, β, γ) задает греко-латинский квадрат. Существуют таблицы подобных ортогональных квадратов (кубов и гиперкубов) на разное число факторов.

Мало известны возможности планирования пассивного эксперимента — использования данных производственного контроля для выделения главных факторов и поиска оптимума. Это большие массивы зачастую не вполне надежных данных. Трудность их обработки в том, что точки $\vec{x}$ размещены сильно неравномерно («кучками», «полосами»). Эта корреляция аргументов «перекашивает» и функцию отклика $y(\vec{x})$. Можно, изучив сначала закономерность в размещении аргументов $\vec{x}$, выделить *главные компоненты*: такие линейные комбинации $z_k = \sum_j \alpha_{jk} x_j$, что между ними статистической зависимости нет. Задача представления отклика в виде $y(z_k)$ — *факторный анализ*. Однако нередко z_k определяются настолько плохо, что даже линейная зависимость $y(z_k)$ сильно размыта. Другой путь — ввести *обучающий массив*: отобрать только точки (или их группы), правильно расположенные «по плану», а затем, построив по ним $y(x)$, убедиться, что множество «прочих» точек не ухудшает зависимости.

Особенность эксперимента с непрерывным производственным процессом в цехе (например, на агрегате непрерывного отжига), в том, что условия $\vec{x}$ разрешено менять лишь в узких пределах (чтобы не шел брак), а отклик y зависит не только от $\vec{x}$ и времени t , но и от производных dx_j/dt . Тогда задача заключается в том, чтобы, задавая малые изменения $\vec{x}$, определить вид уравнения $y(\vec{x}, d\vec{x}/dt)$, описывающего процесс. Для такой «задачи идентификации» существуют «*D*-оптимальные тестирующие сигналы»: малые изменения $\vec{x}$ задают в виде серии коротких возмущений аргумента (например, $x = 1, 1, -1, -1, 1, -1, 1, 1, -1$), которым ставят в соответствие отсчеты отклика $y(x)$ в другие моменты.

Каким бы планом ни пользоваться, эксперимент от-

ражает реальную картину лишь тогда, когда в число изучаемых аргументов включены все существенные факторы. «Подозреваемых» факторов обычно больше, чем действующих, а объем эксперимента и при оптимальном плане растет с числом факторов m пропорционально 2^m . Чем больше факторов, тем менее заметно влияние каждого из них, и тем труднее реализовать даже дробную реплику от 2^m . При $m \geq 10$ все факторы нельзя учесть, но можно — в случае удачи — отсеять неважные из них за $N < m$ экспериментов методом *случайного баланса* [1—3]. Для этого, выделив $k < N$ факторов в качестве значимых, все остальные складывают в один: $x = x_{h+1} + \dots + x_{k+2} + \dots + x_m$. Если незначимые факторы угаданы правильно, то их сложение ни на чем не отразится, зависимость $y(x_1, x_2, \dots, x_k)$ окажется адекватной, и все факторы от x_{h+1} до x_m можно вместе отбросить. Если же среди них попал хотя бы один существенный, нужно начать сначала, выбрав другой «хвост». Тупой перебор сочетаний x_j в «хвосте» не дешевле полного факторного эксперимента 2^m — успех здесь зависит от интуиции.

Сходные идеи перебора факторов реализует «метод группового учета аргументов». В нем план сверхнасыщенный ($N < m$), проверить адекватность модели по рассеянию точек относительно поверхности отклика нельзя, и критерием истинности считают совпадение моделей $y(\vec{x})$, полученных для одной и для другой половины экспериментальных данных (критерий минимума смещения моделей между «обучающей» и «проверочной» выборкой). Такие методы облегчают интуитивный поиск модели, но они в принципе не могут гарантировать единственность найденного решения.

Иногда нужно знать не вид отклика $y(\vec{x})$, а только положение или величину его экстремума — оптимум состава, значение свойств в оптимуме и т. п. [2, 6]. Если $y(\vec{x})$ линейно во всей области существования $\vec{x}$, то оптимум лежит на краю этой области (и всегда находится методом линейного программирования).

Чтобы найти оптимум внутри области по уравнению $y(\vec{x})$, в нем часто не хватит даже и членов с x_j^2 — реальная поверхность отклика обычно сложнее, чем гладкая поверхность второго порядка. Поэтому оптимум находят, не определяя саму функцию $y(\vec{x})$ — методом крутого вос-

хождения [1, 2]. Для этого достаточно в одной точке найти градиент $y(\vec{x})$ и двигаться вдоль него «в гору» — до максимума, время от времени уточняя направление градиента. Выигрыш в том, что градиент укажет уже линейное приближение: чтобы определить $\text{grad } y$ в одной точке $\vec{x}$ в принципе достаточно при m факторах измерить y в $m+1$ точках — вершинах симплекса с центром $\vec{x}$. Кроме того, соседний симплекс, куда указывает градиент, имеет с данным всего одну общую точку, где и надо добавить одно измерение y , чтобы сделать следующий шаг. Только вблизи максимума, когда градиент станет незначим, надо усложнить план и найти поверхность второго порядка вблизи экстремума.

При бесспорной идее успех эксперимента по математическому плану определяется, прежде всего, пониманием существа задачи и метода экспериментатором. Во-первых, нужен физически обоснованный вид аргументов. Так, если в описание кинетики температура входит линейно (или лучше как $x_1 = 1/T$), то время не может входить линейно — аргументом должно быть $x_2 = \lg t$, ибо для термически активируемых процессов $t \sim e^{Q/kT}$. Если положить $x_2 = t$, то «оптимальный» план при малом шаге x_2 охватит малый отрезок процесса, а при большом — пропустит наиболее интересную «середину». В обоих случаях согласие с наблюдениями будет хорошее, а предсказание — никуда не годное (не там стояли точки).

Во-вторых, даже для переменных одной размерности важен выбор масштаба: если x_1 — размер детали в мм, а x_2 — шероховатость ее поверхности (тоже в мм, но на три порядка меньше), то при любом описании $y(x_1, x_2)$ градиент будет направлен строго вдоль x_1 . Чтобы сделать наклоны $y(x_j)$ в обе стороны сравнимыми, нужно ввести переменные $x_j = x_j/M_j$. Масштабы M_j задают равными ошибке определения x_j или диапазону изменения x_j .

В-третьих, размер симплекса для определения градиента ($\pm \Delta x_j$) надо сделать достаточно большим, чтобы y изменялось значимо (иначе градиент неопределенный), и достаточно малым, чтобы поверхность $y(\vec{x})$ на площадке $\pm \Delta x_j$ была сначала примерно линейной.

В-четвертых, надо правильно выбрать длину шага

вдоль градиента* (коротким шагом долго идти до максимума, длинным легко перескочить через него).

Наконец, всякое описание отклика достаточно гладкой поверхностью равносильно допущению, что не будет неожиданностей: разрывов от смены механизма процесса, острых пиков, гребней и т. п. Процедура направлена на сглаживание, а не поиск скачков. Поэтому, решая простые задачи, она может запутывать сложные. Математическое планирование эксперимента — полезный вспомогательный инструмент, а не самодовлеющий аппарат.

Нельзя забывать также, что все методы оптимального планирования уязвимы при потере точек (от брака при обработке, промахов в испытаниях и т. п.). Даже при ненасыщенном плане от пропуска наблюдений появляется взаимосвязь ошибок в искомых b_i , а учет потерянного как нового неизвестного ухудшает надежность выводов.

Математическое планирование облегчает поиск экстремума одного свойства y . В технике же обычно ищут компромисс — возможно лучший уровень сразу нескольких переменных y_k (например, и прочности, и пластичности). Задачу можно свести к предыдущей, вводя одну функцию желательности $Y(y_k)$ [6]. Синтез одного критерия Y из нескольких y_k — задача квалитметрии, и главное здесь в том, как по здравому смыслу задать вид функции $Y(y_k)$, чтобы ее наибольшее значение действительно соответствовало наилучшему для нас сочетанию свойств y_k . Естественно использовать объединение критериев в безразмерные комбинации (например, физические критерии подобия или просто отношения двух величин одной размерности: предела прочности к пределу текучести). Если есть сильная корреляция между двумя y_k , следует оставить один из них.

При пассивном эксперименте, когда все измерения сделаны заранее, можно и не вводить «синтетический» критерий, а вместо этого по поверхности отклика нанести на плоскости x горизонтали $y_k = \text{const}$, очерчивающие область «хороших» и «приемлемых» результатов. Пересечение областей для y_k оставит, быть может, столь узкую область, «приятную во всех отношениях», что далее не понадобится никакая оптимизация (компромисс

* Существуют многие усовершенствования метода с переменным шагом, корректировкой направления, ограничением шага пространством сильнейших переменных (меньшей размерности).

естественный). Даже если компромисса нет, его поиск приводит к четкой постановке дальнейшей задачи: «нужен способ поднять y_4 в области x_1 , или снизить y_3 в области x_2 , все другие подходы еще труднее» (требуют улучшения сразу нескольких параметров).

Другой класс задач планирования — оптимизация физического эксперимента. Цель такого эксперимента — доказать существование зависимости именно данного вида (против четко сформулированной альтернативы), измерить параметры такой зависимости, обнаружить различие параметров либо согласие их с теорией. Принципиальное отличие от предыдущего в том, что вид функции отклика $y(x)$ задан наперед. Иногда задана также и функция трудности — зависимость «цены измерений» (трудоемкости, стоимости, времени) от значений аргумента x (например, измерять интенсивность рассеяния излучения тем труднее, чем ближе к первичному пучку). Оптимальный план [8] указывает координаты точек измерения и распределение на них ресурса (доли цены), чтобы получить наилучший при данном ресурсе результат (и указать объем измерений для заданной надежности результата). Поэтому аппарат оптимизации физического эксперимента с единообразной методикой, исходящий из принципа максимума правдоподобия, отличен от планирования факторного эксперимента.

В любом случае математическое планирование — не первооснова, а лишь составная часть плана, требующая нешаблонного выбора приемов для каждой задачи. Жалоба «от математического планирования мало пользы» часто означает, что для данного объекта выбор метода планирования противоречил здравому смыслу, и это можно было видеть еще до эксперимента.

Рекомендательный библиографический список

1. Налимов В. В. Теория эксперимента. М.: Наука, 1971. 208 с.
2. Рохваргер А. Е., Шевяков Л. Ю. Математическое планирование научно-технических исследований. М.: Наука, 1975. 440 с.
3. Ахназарова С. Л., Кафаров В. В. Оптимизация эксперимента в химии и химической технологии. М.: Высшая школа, 1978. 320 с.
4. Адлер Ю. П., Маркова Е. В., Грановский Ю. В. Планирование эксперимента при поиске оптимальных условий: Программированное введение в планирование эксперимента. М.: Наука, 1971. 284 с.

5. Новик Ф. С., Арсев Я. Б. Оптимизация процессов технологии металлов методами планирования экспериментов. М.: Машиностроение — София: Техника, 1980. 304 с.
6. Вознесенский В. А., Ковальчук А. Ф. Принятие решений по статистическим моделям. М.: Статистика, 1978. 192 с.
7. Применение математических методов для исследования многокомпонентных систем / Зедгинидзе М. Г., Новик Ф. С., Чемлева Т. А., Преображенская Г. Б. М.: Металлургия, 1974. 176 с.
8. Федоров В. В. Теория оптимального эксперимента (Планирование регрессионных экспериментов). М.: Наука, 1971. 312 с.

ЭВМ В ЛАБОРАТОРИИ

Не только сам эксперимент, но и одни лишь образцы для него настолько дороже счета, что всякие результаты следует обрабатывать наилучшим образом, не смущаясь объемом вычислений. Поэтому экспериментатор должен владеть употребляемым в его учреждении алгоритмическим языком хотя бы настолько, чтобы написать и суметь врезать в готовую программу короткую вставку — приспособить ее для своей задачи. Имея доступ на машину, всякий быстро выучивается этому по кратчайшим самоучителям: Фортрана [1], Алгола [2], Бэйсика [3]. Дальнейшее по мере необходимости дается практикой (например, соединение стандартных и собственных программ и хранилищ данных с помощью операционных систем больших ЭВМ [4]).

Счет на ЭВМ доступен всем. Особой же подготовки требует «не счетное» применение ЭВМ для эксперимента — в системах автоматизации, поиска и распознавания. Прямой ввод в ЭВМ данных измерения — задача техники связи (нужны преобразователи аналог — код, каналы связи, их согласование и т. п.). Заниматься этим самим есть смысл лишь в том случае, когда считывание (например, с лент самописцев) и «набивка» для ввода в ЭВМ оказываются во много раз дольше эксперимента.

Следующая ступень автоматизации — управление самим экспериментом «в реальном масштабе времени»: изменение режимов работы прибора, точек наблюдения и схемы обработки по наперед заданной жесткой программе (как в автоматических дифрактометрах) или автоматически в зависимости от получающихся результатов (как в микрозондовых анализаторах высокого класса). Все действительно эффективные установки такого рода (или их первые прототипы) созданы не заводами

приборостроения, а самими экспериментаторами. Только они могут решать, какие операции действительно важно автоматизировать. Постройка и ввод установки такого уровня — дело не одного года, и оно оправдано лишь там, где производительность увеличивается на порядки: либо автоматизация позволяет получить и переработать такой массив информации, какой вручную никогда и не пытались получить, либо ЭВМ решительно улучшает точность эксперимента за счет непрерывной корректировки режима работы, что самому экспериментатору было ранее недоступно по быстрдействию и по объему функций.

Другая сфера «не счетного» применения ЭВМ экспериментатором: упорядочение больших массивов экспериментальных данных (в том числе чужих) для их качественного анализа и обобщений. Одномерная задача упорядочения по некоторому трудно измеримому или принципиально качественному признаку (например, по степени перехода от пластинчатого перлита к зернистому в серии шлифов или по шиферности изломов) — *ранжировка экспертных оценок* [7, 8, 10]. Если, по крайней мере, для некоторых пар объектов наблюдатель может указать отношения неравенства $a_1 > a_3$ (первый синее третьего: a_1 ближе к синему, а a_3 — к зеленому), то *метод парных сравнений* [8] по таким парам осуществляет ранжировку всех наблюдений и количественную оценку признака (баллом от 0 до 1) с учетом неопределенности некоторых отношений («не видно разницы»), противоречий (у одного наблюдателя оказалось $a_4 > a_5$ и $a_5 > a_7$, но $a_4 < a_7$) и расхождений между разными наблюдателями («экспертами»). Метод парных сравнений обобщается и на многомерные задачи упорядочения по многим косвенным признакам одного свойства.

Выделение групп родственных объектов (например, сплавов) по совокупности многих признаков — задача *кластерного анализа* [11, 12]. В нем главный и не формализуемый шаг — задать некоторым образом расстояние между любыми двумя объектами. Например, если «да» и «нет» кодировать как 0 и 1, то при n признаках один объект изображается n -мерным вектором $\vec{y} = [01100010]$, и родство можно мерить как евклидово расстояние ρ между точками $\vec{y}_1$ и $\vec{y}_2$ в n -мерном пространстве (предостережения о важности выбора масшта-

бов остаются в силе). Второй неформальный шаг — определить признак «родства»: например, одному кластеру (гнезду) принадлежат такие точки, что расстояние между любыми из них менее ρ_0 , или хотя бы до одной — менее ρ_0 и т. п. После этого дело техники (ЭВМ) радионально перебрать все паросочетания и выделить гнезда — кластеры. В зависимости от выбора формулы расстояния, критерия и величины ρ_0 получаются разные разбиения на кластеры. Задача экспериментатора — угадать, в каком варианте результат имеет наилучшую физическую интерпретацию. Как и все статистические методы, кластерный анализ — человеко-машинная процедура, где необходимы также знания, привнесенные извне («априорная информация»).

Когда классы объектов заданы, распознавание принадлежности предмета (минерала, соединения) к одному из классов — задача *таксономии*. Ее решению предшествует обучение — построение по признакам множества известных объектов *дискриминантной функции*, которая должна для данного класса попадать в определенный интервал значений.

Простая частная задача таксономии — разбиение пространства непрерывных количественных признаков на области, содержащие точки одного класса (например, построение изотермических диаграмм превращения — в том числе многомерных, включая зависимость от степени деформации, состава и т. п.) по небольшому числу точек каждого класса («до», «во время», «после превращения»). Идея заключается в выборе некоторых естественных спрямляющих координат, где все точки каждого класса размещены компактно. Тогда строят выпуклый многогранник, который содержит все точки данного класса и ни одной «чужой». Плоскость, разделяющая два многогранника и наиболее удаленная от обоих, — наилучшая граница, чтобы предсказать состояние сплава при любых условиях. Двумерную диаграмму быстрее нарисовать на бумаге, чем ввести ее точки в ЭВМ. Однако при многих переменных только ЭВМ может проводить n -мерные границы областей, учитывая положение всех точек сразу (а не только на данном разрезе). Выигрыш здесь в большей надежности результата при меньшем числе точек — ускоряется не обработка, а сам эксперимент. Сходным способом строят многофазные или многомерные диаграммы состояния по умеренному

числу точек. Здесь надо учитывать топологические ограничения, разрешающие соседство лишь определенных полей (например, по правилу фаз).

Кластерный анализ и таксономия — частные случаи *распознавания образов* [11, 13]; группировки серии изображений (снимков, карт и т. п.) на сходные и несходные (между собой или с эталоном). В кластерном анализе «изображение» проще тем, что это набор изолированных точек «равной яркости».

В распознавании образов, помимо численных «поточечных», бывают эффективны и физические интегральные процедуры: оптическое сравнение Фурье — спектров изображений, их голограмм, вычитание изображений и т. п. Например, интегральное (без промеров линий) голографическое опознавание сложного линейчатого спектра (дебаеграммы смеси веществ) оптическим сравнением с серией эталонов [15].

Еще одна область применения ЭВМ, где личное участие экспериментатора нужно на всех стадиях разработки, — информационно-поисковые системы (ИПС). Даже простая задача: поиск патентов о сплаве — требует для автоматизации четкого структурирования. Чтобы сгруппировать для поиска признаки состава и свойств, нужно хорошо представлять реальные взаимосвязи свойств и возможную частоту разных сочетаний признаков. Широкое профессиональное знание объекта поисков здесь важнее, чем программирование. Тем более это относится к сложным ИПС, например, для опознавания веществ (в том числе в смесях) по их спектрам, рентгенограммам, построения структур по спектрам [16] и, наконец, предсказания существования соединений с нужной структурой или свойствами по совокупности данных о родственных веществах.

Активно используя ЭВМ, экспериментатор независимо от его исходной подготовки начинает ориентироваться и в своем разделе прикладной математики. Усвоив, что всякое описание зависимостей от многих переменных приводит к операциям с матрицами [18], он вскоре обнаружит, что если многие связи выпадают, то в матрице нулей много больше, чем других чисел, и тогда есть своя, более экономная техника. Всякая задача отделения «инструментальных функций» от «физических зависимостей» — обратная задача математической физики, для которой есть не только удачные алгоритмы [24], но и прин-

ципы регуляризации — введения априорной информации о зависимости, без которой не может быть однозначного решения [23]. Находя экстремум функции нескольких переменных, надо знать, что если он не единственный, то нет универсального алгоритма, и выбор способа зависит от вида функций [21].

Выигрыш от ЭВМ — не в ускорении счета, а в том, что с ней экспериментатор решает задачи, иными способами неразрешимые.

Рекомендательный библиографический список

1. Салтыков А. И., Семашко Г. Л. Программирование для всех. М.: Наука, 1980. 160 с.
2. Брудно А. Л. Алгол. 2-е изд. М.: Наука, 1972. 80 с.
3. Кетков Ю. Л. Программирование на Бэйсике. М.: Статистика, 1978. 158 с.
4. Землянский А. В., Перенц М. Г. Основы операционной системы ЕС ЭВМ. М.: Советское радио, 1980. 144 с.
5. Абрамов С. А. Экономическое обоснование автоматической обработки информации. М.: Статистика, 1974. 183 с.
6. Евтеев Ю. И., Юрухин Б. Н. Цифровые системы обработки сканируемой информации. М.: Энергия, 1975. 136 с.
7. Дэвид Г. Порядковые статистики. Пер. с англ. М.: Наука, 1979. 335 с.
8. Дэвид Г. Метод парных сравнений. Пер. с англ. М.: Статистика, 1978. 144 с.
9. Лоули Д., Максвелл А. Факторный анализ как статистический метод. Пер. с англ. М.: Мир, 1967. 144 с.
10. Миркин Б. Г. Анализ качественных признаков и структур. М.: Статистика, 1980. 320 с.
11. Елисеева И. И., Рукавишников В. О. Группировка, корреляция, распознавание образов. (Статистические методы классификации и измерения связей). М.: Статистика, 1977. 144 с.
12. Диоран Б., Оделл П. Кластерный анализ. Пер. с англ. М.: Статистика, 1977. 128 с.
13. Джурс П., Айзенауэр Т. Распознавание образов в химии. Пер. с англ. М.: Мир, 1977. 230 с.
14. Ярославский Л. П. Введение в цифровую обработку изображений. Советское радио, 1979. 312 с.
15. Василенко Г. И. Голотографическое опознавание образов. М.: Советское радио, 1977. 328 с.
16. Эляшберг М. Е., Грибов Л. А., Серов В. В. Молекулярный спектральный анализ и ЭВМ. М.: Наука, 1980. 308 с.
17. Египко В. М. Организация и проектирование систем автоматизации научно-технических экспериментов. Киев: Наукова думка, 1978. 232 с.
18. Демидович Б. П., Марон Н. А. Основы вычислительной математики. 4-е изд. М.: Наука, 1970. 664 с.
19. Хемминг Г. В. Численные методы для научных работников и инженеров. Пер. с англ. М.: Наука, 1972. 400 с.

20. Моисеев Н. Н., Иванюков Ю. П., Столярова Е. Н. Методы оптимизации. М.: Наука, 1978. 352 с.
21. Стронгин Р. Г. Численные методы в многоэкстремальных задачах: Информационно-статистические алгоритмы. М.: Наука, 1978. 240 с.
22. Поттер Д. Вычислительные методы в физике: Пер. с англ. М.: Мир, 1975. 392 с.
23. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. 2-е изд. М.: Наука, 1979. 288 с.
24. Вергань А. Ф., Сизиков В. С. Методы решения интегральных уравнений с программами для ЭВМ: Справочное пособие. Киев: Наукова думка, 1978. 292 с.
25. Сегерлинд Л. Применение метода конечных элементов: Пер. с англ. М.: Мир, 1979. 392 с.

4. ОРГАНИЗАЦИЯ И АНАЛИЗ ЭКСПЕРИМЕНТА

ПЕРВИЧНЫЕ ДОКУМЕНТЫ

— Вы заняты?...

— Да, разные дела на память в книгу вносим. Забудется, того гляди.

Грибоедов. «Горе от ума»

Любое исследование заканчивается отчетом. Общие требования к отчетам и правила их оформления в СССР едины — их установил ГОСТ 19600—74 «Отчет о научно-исследовательской работе». ГОСТ гласит: «Отчет о НИР является основным документом, в котором излагаются исчерпывающие сведения о выполненной работе». Этим единым документом автор отчитывается перед руководством предприятия и ведомства в полезности своих результатов для производства, перед планово-финансовыми органами — в целесообразности расходов, перед учеными — в научной ценности результатов.

В отчете приведены первичные результаты эксперимента, дан анализ их достоверности и значимости, выделены новые, ранее не известные связи и зависимости, из их числа отобраны фактические данные, полезные для техники, и принципиально новые факты и выводы, являющиеся вкладом в науку, обоснованы рекомендации для практики и указаны перспективы дальнейших исследований. Прочую документацию по работе: предполагаемые технологические инструкции, технические условия на новые сплавы, проектные задания на оборудование для новых процессов, заявки на изобретения, статьи — пишут на основе отчетов, содержащих весь необходимый материал.

Отчет об экспериментальном исследовании составляют на основании первичных документов, которые хранятся в лаборатории и в любое время могут быть проверены. Главный первичный документ — рабочие журналы всех участников работы. Кроме них, в лаборатории годами хранят все образцы, негативы, рентгенограммы, осциллограммы, ленты с самописцев и цифропечатающих устройств, диаграммы с испытательных машин, плавочные сертификаты, протоколы анализов и сделанных «на стороне» испытаний. Хранят также все черновики вычислений, перфоленты или карты массивов данных и программ для ввода в ЭВМ, распечатки программ и выведенных массивов. Все эти материалы (в день получения!) зарегистрированы и описаны в рабочем журнале, на каждом из них поставлены даты и номер, отсылающие к соответствующей странице журнала.

Правила и сроки хранения в разных учреждениях различны, но общий принцип: материал следует сохранять, пока можно ожидать каких-либо возражений и сомнений в его надежности — после публикации, при использовании результатов в производстве, при выполнении следующих исследований в развитие этого. Чтобы спать спокойно, большинство экспериментаторов хранят свои рабочие журналы пожизненно, а образцы — по меньшей мере, несколько лет. (В сохранившихся поныне записных книжках Фарадея последний опыт № 16041).

Рабочий журнал — толстая тетрадь, где на первом листе указаны автор, лаборатория, учреждение (адрес и телефон — а вдруг...?). Все записи сделаны во время эксперимента. Записей без даты нет. Не должно быть никаких листочков-черновиков. Записи «по памяти» за прошлые дни недопустимы — это уже не протокол, а легенда. Науку нельзя строить на воспоминаниях отдельных личностей.

В журнале, как и вообще в документах, ничего не вычеркивают и не стирают.

Первичные данные не исправляют никогда. Если есть сомнения, все перемеряют и записывают заново. Найдя, что именно было неправильно (обнаружен дефект в схеме, брак образцов и т. п.), делают об этом запись сегодняшним днем:

8.X. В схеме фотоусилителя, собранной 26.IX, корпус не заземлен; в повторных измерениях эталона разброс до 10 % от наводки 60 Гц... (далее описание измерений, доказывающих причину ошибки)... Все измерения с 28.IX по 7.X неправильные.

Затем все старые данные обводят цветным карандашом и надписывают: «Измерения неправильные — см. запись 8.X.»

Ошибку в расчетах при обработке результатов зачеркивают так, чтобы старые цифры были видны (как знать — в чью пользу будет вторая или третья проверка?). Только при очень большом объеме обработки (на ЭВМ) массивы первичных данных переписывают из журнала на отдельные листы (группируя их так, как удобно для ввода). Хотя и эти листы, и распечатка с ЭВМ хранятся, таблицу окончательных данных переписывают в рабочий журнал, чтобы собрать все вместе так, как удобно для дальнейшего анализа.

Самая непоправимая и непростительная ошибка — перепутать образцы. Поэтому журнал начинают с описания материала и образцов.

17.X.88 получено в НИИАВВ от Э. Ю. Яковлева 0,8 кг сплава ЭП000 (плавка 7198) в виде проволоки диам. 2,0 мм. В мотке 4 отдельных куска проволоки: № 1—3 совсем светлые, № 4 — местами с пятнами нестравленной окалины. Длина 9,0; 9,1; 7,0; 8,2 м. Диаметры: № 1—2,00—2,02; № 4—1,97—1,99 (по измерениям в 10 точках по всей длине, крест-накрест). На мотке фанерная бирка «пл. 7198». На каждый кусок нанизал жестяные бирки с клеймом 1, 2, 3, 4. Получен сертификат № 981/14 от 7.V.88. Анализ плавочный (пробы от слитка), сделан в лаборатории № 11 НИИАВВ. Состав по сертификату:...

По данным Э. Ю. обработка: закалка в бунтах 1100 °C, 1 ч, вода, травление в расплаве щелочей, затем холодное волочение — всего 4 заковки, последнее обжатие 50 % — материал наклепан. Технология плавки описана в отчете НИИАВВ № 217. Термообработка и волочение — по технологии завода Ц для жаропрочных сплавов. Других сведений у Э. Ю. нет.

Здесь нет ни одной лишней подробности. Все дальнейшие превращения зависят от исходного состояния. О нем пока известно немного, но сразу узнали, где найти остальное. Состав пробы от определенного места слитка может существенно отличаться от состава проволоки из-за ликвации или выгорания некоторых элементов при переделах. (Поэтому, получив плавочный анализ, стоит повторить его на окончательном продукте, чтобы не перепроверять задним числом на уже испытанных образцах). На одну серию сопоставляемых экспериментов надежнее взять образцы от одного куска: отпадут сомнения, что попал образец другой плавки. Тогда надо знать, сколько в каком куске метров (сколько выйдет образ-

* Все протоколы анализов и испытаний на стороне подшиваются и, кроме того, вписываются в журнал.

Металлу нужны броские надписи (что, чье, дата, «не трогать!»).

18.X. От головы и хвоста каждого мотка проволоки пл. 7198 отрезано по образцу $L=50$ мм для количественного спектрального анализа на Cr , Al , Ti , Mo , W , Nb в лаборатории № 9. Мотки № 1—4 помечены соответственно 1—4 надпилами надфилем. Сданы как 8 отдельных образцов: Б1, пл Б4».

Чтобы перепроверить, где какой моток, достаточно сосчитать по журналу, от какого сколько отрезано. Информация должна быть избыточна, иначе ошибки неисправимы.

Система нумерации, маркировки и хранения образцов продумана до их изготовления. Она единая для всей работы (если в июне и в декабре сделаны два образца под одним номером—как различать их через год?). Надо предвидеть, будет ли маркировка безвредной и живой: клейма могут помешать испытаниям, надписи — обгореть при термообработке. Для ленточных образцов часто применяют систему надразов («№ 1 — один угол сточен, № 2 — два смежных угла с торца, № 3 — два по диагонали, ..., № 11 — два угла по диагонали и один надпил на торце...»). Вся система маркировки сначала распилана в журнале, а уж затем нанесена на образцы — иначе путаница неизбежна. При вырезке образцов из листа надо зарисовать, как все они размещались по его длине, ширине, направлению, чтобы сохранить возможность проверить неоднородность и анизотропию материала.

66

Технологию получения исходного металла записывают возможно подробнее по данным поставщика. Чаще всего она нигде полностью не описана, и, кроме поиска документов, приходится наблюдать самому и расспрашивать исполнителей (сначала составить полную схему описания технологии, чтобы ничего не упустить, затем заполнять ее по вопросам). В схему обычно входят: шихта (марка и состав компонентов по ГОСТу или ТУ), способ выплавки (печь, материал тигля, шлаковый режим, атмосфера или вакуум, вес плавки и слитков), способ раскисления и разливки, режимы гомогенизации, режим и способ нагрева под ковку, прокатку (и охлаждения после них), последовательность обжиганий, промежуточные отжиги, смазка, травление, термообработка продукции — «состояние поставки».

5*

ница в распределении наклепа по сечению при одинаковой суммарной деформации. Чтобы не развивать гипотезы на атомном уровне, когда все дело в том, каким концом образец лежал под нагревателем печи, а каким — у холодной заслонки, надо своими глазами видеть от начала до конца все подготовительные операции и записывать все подробности немедленно (для чего полезно продумать наперед, что может оказаться впоследствии важным?).

Если может влиять коррозия на воздухе или превращения при $+20^{\circ}\text{C}$, записывают историю хранения:

16.1. Начало прокатки 12³⁰. 13⁴⁵ — прокатаны все ленты. 14³⁰ — ленты вывешены на улицу (-10°C)... 25.1. Температура хранения с 16.1 не выше -5°C ; 9²⁰ — лента внесена в помещение ($+22^{\circ}\text{C}$)... 14²⁵ — после резки и маркировки коробка с образцами поставлена в морозильную камеру холодильника (по термометру -7°C).

Отсюда можно оценить эквивалентное (приведенное по известной энергии активации, например, к 20°C) время старения от прокатки до испытания образца.

Никому нельзя передоверять отбор проб для анализа (если их режет станочник, то в присутствии экспериментатора). При безупречной работе химиков-аналитиков один и тот же кусок фольги может оказаться из стали 03X18H9 у одного исследователя и из 17X18H9 — у другого, потому что первый промыл ленту ацетоном, затем спиртом и, не прикасаясь к вымытому концу, настриг пробу в конверт из кальки, а второй мыл бензином, резал в руках и завернул в тетрадный лист. В радужных пятнах, отпечатках пальцев и махрах от бумаги достаточно углерода, чтобы перекрыть весь эффект от чистейшего феррохрома, плавки и разливки под вакуумом.

Все, что можно записать до эксперимента (его цель, условия, перечень и состояние образцов, результаты проверки схемы), записывают перед началом работы.

Начнем с записей об оборудовании:

17.XI. Измерения электросопротивления. Компенсационная схема с образцовой катушкой. Потенциометр ППТН-1, заводской № 0714, по паспорту пределы 0,1 В — 0,1 мкВ и основная погрешность 0,015 % $+0,04$ мкВ. Гальванометр М25/3 № 8795, $0,45 \cdot 10^{-3}$ мкВ/(мм·м), от зеркала до шкалы 1,4 м. Образцовая катушка Р310 0,01 Ом, № 901, класс 0,02 при $+15 \div +30^{\circ}\text{C}$. Нормальный элемент 3-го класса № 518. Питание цепи потенциометра — 5 батарей «Девиз» последовательно (номинал 1,28 В, 525 А·ч, ток не более 0,6 А), цепи образца — 3 батареи «Девиз» последовательно.

Далее зарисована схема измерения, записаны конкретные условия:

Предел 0,1 мкВ (деление в последней декаде); ток цепи образца 50 мА, температура термостата $+19^{\circ}\text{C}$, каретка для образцов с латунными ножами, база по штангенциркулю 74,7 мм (74,75 у левого, 74,65 у правого крепления ножа). Потенциальные отводы образца выведены на клеммы «1», а эталона — на клеммы «0» декадного переключателя,

Затем наперед подробно записан план действий при измерении:

Цепь рабочего тока включена на весь день. Установка тока по нормальному элементу два раза в день (перед началом и после перерыва). Измерения: эталон — образец — эталон при двух направлениях тока. Время отсчета записываем с округлением до 1 мин в момент перекидки тока. Истинное напряжение на эталоне в момент измерения образца вычислим линейной интерполяцией по времени, температуру — перед каждым циклом измерения.

После этого идут контрольные эксперименты, проверяющие правильность выбранных режимов, чувствительность, воспроизводимость:

Образец № 17. Рабочий ток 100 мА. Уравновешен после выдержки под током 3 мин и выключен, через 2 мин включен — разбаланс от охлаждения 3 мм шкалы гальванометра (1 мм = 0,2 мкВ на декаде). Поставлен ток образца 50 мА. При той же проверке разбаланс за 2 мин около 0,5 мм. Для дальнейшего принят ток 50 мА.

Далее идет таблица измерений на сегодня: № образца, колонки для всех отсчетов (с обозначением величин и единиц измерений), колонки для обработки всех результатов (в том числе и для контрольных подсчетов).

По ходу измерений записывают все изменения условий в течение дня:

14³³ — не сработало реле термостата, температура упала до $24,3^{\circ}\text{C}$; 16⁰² — в реле подрегулирован зазор, измерения начаты с № 43 снова. Ток цепи образца 48 мА, в остальном без изменений.

При многодневных однотипных измерениях каждый день начинают с записи всех изменений на сегодня:

14.III. Измерения электросопротивления. Схема 5.III. Рабочий ток не устанавливается. 3 батареи заменены на новые, соединения пропапаны. Ток 60 мА, в остальном без изменений.

В этой бюрократии, к сожалению, нет ничего лишнего: при анализе результатов много раз будут возникать сомнения в причинах рассеяния, систематических сдвигов отсчетов, выбросов. Исключить из-под подозрения влияние изменений в схемах и условиях измерения задним числом можно только последовательным хронологическим сопоставлением: на какие дни приходится сомни-

тельные результаты, совпадает ли их начало и конец с каким-либо из изменений в условиях работы?

Пока схему монтируют и отлаживают, в журнал вносят все переделки (иногда, чтобы понять, какой же вариант в конце концов работает, нужно либо «поднять архив», либо распаять всю схему). Здесь же все снятые характеристики блоков в каждом варианте с комментариями: почему они не работают.

Большинство приборов подлежит обязательной периодической государственной поверке. Поверитель ставит клеймо или выдает свидетельство, где указан срок следующей поверки. Измерения на просроченных приборах незаконны, поэтому все даты поверки есть в журнале. Клеймо госповерителя не гарантирует, что машину не сломают до срока — ее следует поверять самому перед началом большой серии измерений или при подозрении в повреждении и вносить в журнал полный протокол поверки.

Самопишущие или цифropечатающие приборы не освобождают нас от необходимости записей. Наоборот, их приходится делать дважды. В журнале записано.

11.X. Дифрактометр ДРОН-0,5 трубка БСВ6, Fe, гониометр ГУР-4, приставка для съемки с вращением. Юстировка З.Х. Нуль счетчика $+0^{\circ}03'$, нуль столика $+0^{\circ}14'$, далее все углы по записи с экрана ГУР, без исправления на истинное положение нулей. Оправка — кольцо с пластином, параллельность образца выверяли на плите с индикатором до 0,01 мм. Счетчик СРС-4, настройка З.Х. Щели 0,25; 0,25; 0,1 мм. Вертикальные 4 мм. Фильтр Мп—фольга 20 мкм на второй щели. Скорость счетчика 0,5 град/мин, ленты 20 мм/мин. На трубке 16 кВ (3-я ступень), 10 мА. Диапазон 0—200 имп/с, постоянная интегратора IV (8 с). Линия (200) α , интервал $2\theta=83-88^{\circ}$. Образец № 14: с одного установка снят 3 раза по 2 прохода с переменной направлением вращения счетчика. № 15, 19, 20 — то же, на той же ленте.

На ленте перед началом записи — те же сведения о режиме и дата, затем номер образца, стрелки $\theta \rightarrow$ (направление роста), цифры отметки углов. Без этой двойной записи невозможно разобраться с ворохом лент (особенно при пересъемках: номера образцов повторяются, режимы разные — какие где?).

При одновременной записи на нескольких лентах (например, тока, напряжения и температуры для температурной зависимости сопротивления) важно нанести надежные синхронизирующие отметки — например, закорачивая все входы самописцев нажатием одной кнопки до и после каждой записи.

Особенно важны подробные записи о простых «ручных» процедурах, почти не связанных с измерениями:

14.XII. Закалка образцов № 85—108. Печь МП2 № 4. Потенциометр ЭПД12 № 1291, термopара ХА диам. 2 мм (поверка З.Х — см.). Для закалки бак с 10 л 10 % NaCl (см. 11.XII), без охлаждения, на половине глубины сетка.

13⁵⁵ температура 940°C , по диаграмме колебания $\pm 10^{\circ}\text{C}$ с периодом 6 мин.

13⁵⁸ — загружены образцы № 85—89, на подставку, под термopару, в один ряд, без зазоров. Температура упала до 910°C .

14⁰² — восстановилась до 950°C .

14⁰⁴ — образцы сравнились по цвету с печью.

14¹⁷ — закалка, шум кипения 6 с.

Здесь опять-таки ничего лишнего. Можно получить совсем другую структуру закалки при тех же показаниях стационарной термopары, если вместо подставки класть образцы на под (на 30° горячее). Если в один день закалочный бак был с сеткой и образцы омывались со всех сторон, а в другой они долго кипели прямо на дне, то никакими тонкими структурными методами не найти причину загадочных различий в свойствах, когда нет достоверных первичных записей, чтобы сосчитать скорость охлаждения до 100°C для каждой серии образцов.

Чтобы не отвлекаться записями, когда следишь за образцами, можно вписывать цифры в заготовленный текст. Но не забудьте записать все непредвиденные отклонения от шаблона и промахи!

При работе с микроскопом указывают объектив, окуляр, общее увеличение, способ освещения. После общей характеристики структуры записывают видимые качественные различия между образцами и между разными полями зрения в одном образце. Чтобы сопоставить структуры качественно, важно иметь количественные характеристики:

на 10 полях при $\times 200$ карбиды от 1 до 3 мкм, от 2 до 12 штук в поле (поле $9 \times 12 \text{ см}^2 = 27 \cdot 10^4 \text{ мкм}^2$).

Фотографируя, кроме номеров негативов, образца и увеличения, отмечают, почему выбрано место съемки: если сняты три «типичных» поля и два диковинных включения, то через полгода без журнала не вспомнишь, как часто было то и другое. Высохшие негативы сразу подписывают (автор, дата, номер), а закончив работу, составляют их общий список для архива лаборатории.

Кроме протокола, полезны соображения о причинах неудач, о том, что проверить, как переделать, что учесть при обработке.

В общем журнале уместен и конкретный план действий на некоторый срок: перечень измерений и прочих дел (предвидимая цепь событий для контроля), текущие задания руководителя, результаты обсуждений и т. д.

ПЕРВИЧНАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ

Запомните, экспериментатор, в отличие от теоретика, ошибается только один раз, и потом ему уже не верят.

Л. А. Арцимович

То, что измерено, — это еще не то, что мы хотим знать. Все измерения обрабатывают. Первичная обработка — вычисление искомых функций от набора измеренных величин для данного образца в данном состоянии (например, удельного электросопротивления — по отсчетам напряжения, тока и размерам образца), усреднение и описание рассеяния результатов повторных измерений «в одной точке».

Соблюдаются обычные правила округления; недоразумения неизбежны, когда не пишут «для краткости» нули (21,0 и 21,00 — это совсем разные числа). В ряде случаев округление регламентирует ГОСТ: так, предел текучести, упругости округляют до 5 МПа при $100 \leq \sigma \leq 500$ МПа, до 1 МПа — ниже и до 10 МПа — выше этого интервала σ . Для относительного удлинения и сужения правило ГОСТ 1497—73 другое: до 0,5 % округляют значения $10 \% \leq \delta \leq 25 \%$ (ниже — до 0,1 %, выше — до 1 %). Округление среднего из измерений на нескольких образцах ГОСТ не регламентирует, но ясно, что гарантировать следующий знак могут не менее 10 образцов на точку.

Все расчеты должны быть доступны для перепроверки, все промежуточные результаты — в таблицах в журнале, без черновиков. Лучше вести первичные записи чернилами, а обработку карандашом, чтобы не спутать измеренное (раз и навсегда) с расчетом, который еще не поздно перепроверить.

Не слишком длинные вычисления (в несколько действий) делаем, как обычно, с одним лишним знаком (стало быть, на линейке получим всего двузначный окончательный результат). Для большей точности нужны клавишные машины. В простых первичных расчетах ЭВМ помогает редко: выбрать из журнала и набить ис-

ходный массив на перфоленту — почти так же долго, как просчитать на клавишной машине.

При сотне первичных измерений даже для простых вычислений полезно обзавестись графиком, например, диагональ отпечатка (в делениях именно этого окуляра) — число микротвердости. Если первичные вычисления сложны, стоит потратить время на специальную номограмму (например, толщина образца — прогиб — относительная деформация при сильном нелинейном изгибе ленты, где в расчет входят эллиптические интегралы). Даже на квадратном метре миллиметровки ошибка графика или номограммы будет не менее 0,1 % от диапазона изменения. Поэтому номограммы приемлемы там, где допустимо округление до 1 %. Графические методы быстрее численных, когда надо найти много корней трансцендентного уравнения (на плоскости), неявные, многозначные обратные функции. Точность графического метода определена ошибкой проведения кривой по точкам.

Графическое интегрирование и дифференцирование почти всегда хуже численных. Полярный планиметр измеряет площадь в 100 см^2 с погрешностью 0,5 % (и эта ошибка обратно пропорциональна площади). Нормаль и касательную к кривой с помощью зеркала можно провести с ошибкой в $1\text{--}2^\circ$. Тогда наименьшая — для наклона в 45° — ошибка в производной составит 3—6 %, но ошибка проведения кривой по точкам (по лекалу или от руки) обычно гораздо больше, и для нее нет объективной оценки погрешности. Поскольку графические методы не учитывают первичные ошибки кривой (в каждой точке измерений), они создают иллюзию точности, и многократное применение графической обработки часто приводит к катастрофическим результатам. Так, если найденные производные входят как первичные данные в новую задачу (например, по глубине слоев и времени нашли коэффициент диффузии $D(T)$, по его температурной зависимости — энергию активации Q , по ее зависимости от состава — термодинамические константы и т. д.), то графическую обработку можно продолжать как угодно далеко (поскольку никакой информации о предыдущих погрешностях на новом графике нет). Только численное решение с оценкой дисперсий параметров напомним, что информация исчерпана и пора остановиться — с учетом наследования ошибок все новые параметры значимо не отличаются от нуля. При графичес-

кой обработке автор пробивается к цели, не взирая на подобные мелочи — наметанному глазу нетрудно слегка повернуть прямую, куда следует, или дорисовать пик там, где он явно должен быть. Когда нет никаких вычислительных средств, графическая обработка проще, но это, видимо, тот самый случай, когда «простота хуже воровства». Графическая обработка полезна лишь для предварительных прикидок, а «чистовая» обработка — почти всегда численная. При комбинированной обработке (например, графическое отделение фона «на глаз» с последующим вычислением параметров линии на ЭВМ) чаще всего тоже получается недопустимая диспропорция точности (произвол в выборе фона делает бессмысленным дальнейший точный счет).

Особый случай — обработка непрерывных записей (лент самописцев, осциллограмм и т. д.). Здесь встречаются два типичных недоразумения. Во-первых, нельзя считать, что случайная ошибка записи определяется классом прибора. Когда сигнал меняется от нуля до U_0 мгновенно, показания прибора приближаются к истинному U_0 по закону $U = U_0(1 - e^{-t/\tau})$. Постоянная времени τ — мера инерционности системы — определяется самым тихоходным звеном. Иногда это инерция самописца, но чаще — датчиков и преобразователя (в частности, при счете квантов τ — постоянная времени интегратора). Прибор класса 0,5 с постоянной времени $\tau = 1$ с регистрирует постоянный сигнал с погрешностью 0,5 % только через 5 с. При динамических измерениях он сглаживает резкие скачки как частотный фильтр с передаточной характеристикой $K(\omega) = 1/(1 + \omega^2\tau^2)$. Только после сопоставления статической погрешности и инерционности всех звеньев системы непрерывной записи можно судить о точности и выбирать моменты отсчетов «постоянных» величин.

Во-вторых, точки промеров для обработки нельзя брать с ленты сколь угодно часто. По теореме Котельникова, вся полезная информация о непрерывном сигнале сохраняется при дискретных отсчетах с шагом 2τ (интервал корреляции). Более частое считывание не повышает точности (любые промежуточные значения однозначно заданы этими опорными отсчетами)*.

* Отсюда же исходят при статистических оценках: непрерывная запись за время T равноценна $n = T/2\tau$ независимым отсчетам.

Обработав данные по каждому образцу, находим средние «на точку» (для группы одинаковых образцов или для многократных измерений на одном образце в идентичных условиях). Если искомая величина — нелинейная функция измеряемой, то не все равно, когда усреднять. Так, твердость по Виккерсу HV зависит от диагонали отпечатка d как d^{-2} и, конечно, среднее HV по n измерениям не совпадает с фиктивным числом твердости, найденным по средней диагонали $\bar{d}$. Усреднять надо окончательную, имеющую физический смысл величину (твердость), а не промежуточные отсчеты.

Легко увидеть положение среднего на графике, где уже нанесены точки для каждого образца. При правильно выбранном масштабе усредняемые отсчеты различаются на миллиметровку на небольшое число целых клеточек. Нанеся на глаз черту для ожидаемого среднего x' суммируем (со знаком) отклонения от нее всех точек $\Delta x_i = x_i - x'$ прямо в клеточках (т. е. в однозначных целых числах). Действительное среднее будет

$$\bar{x} = x' + \sum_{i=1}^n \Delta x_i / n.$$

Почти в каждой задаче можно предложить несколько способов оценки искомого x . Выбирая между «быстрой» и «хорошей» оценкой, сравнивают их качество по ряду признаков: состоятельности, достаточности, эффективности, несмещенности, устойчивости, линейности. *Состоя-*

тельны такие оценки $\hat{x}$, где с ростом числа измерений n дисперсия оценки $\sigma_{\hat{x}}^2$ уменьшается. *Достаточны* оценки,

использующие всю содержащуюся в отчетах x_i информацию об x . *Эффективны* оценки, например, 0,5 означает, что при данном n дисперсия $\sigma_{\hat{x}}^2$ вдвое больше, чем

в наилучшем возможном методе (асимптотическая эффективность соответствует пределу отношения этих дисперсий при $n \rightarrow \infty$). Оценка $\hat{x}$ — *несмещенная*, если при $n \rightarrow \infty$ она сходится к «истинному значению» — математическому ожиданию x_0 . Оценка *устойчива*, если ее надежность мало страдает при отклонениях вида распределения $p(x)$ от предполагаемого. *Линейная оценка* — линейная функция наблюдаемых x_i .

Состоятельны почти все разумные оценки*. Всегда желательны оценки достаточные и с эффективностью 1 (чтобы обойтись наименьшим числом измерений). Так, оценка $\bar{x}$ по среднему значению $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ — достаточная и эффективная. Ее дисперсия убывает как $1/n$. Оценка же из размаха $\hat{x} = (x_{\max} - x_{\min})/2$ — не достаточная (не использует все остальные отсчеты x_i) и нулевой асимптотической эффективности: ее дисперсия убывает как $1/n$. Можно получить оценку лучше эффективной, если, кроме наблюдений x_i , ввести еще и априорную информацию. Например, если x — масса, то заведомо $x > 0$ (распределение x имеет обрыв), и тогда существует *сверхэффективная оценка* $\hat{x}$, где дисперсия убывает как $1/n^2$. Несмещенность — не абсолютное достоинство. Если в результате смещения $\lim_{n \rightarrow \infty} (\hat{x} - x_0) \neq 0$ можно уменьшить дисперсию оценки $\sigma_{\hat{x}}^2$, то может быть выгоднее добиваться не минимума $\sigma_{\hat{x}}^2$ при $\Delta = 0$, а минимума суммы $\Delta^2 + \sigma_{\hat{x}}^2$ — квадратов систематического смещения и случайной ошибки (*гребневой анализ, ридж-анализ*). Устойчивость важна для критериев надежности: большинство из них выведено только для нормальных распределений, а доказать нормальность трудно. Линейность оценки говорит лишь о легкости счета.

СИСТЕМАТИЧЕСКИЕ ОШИБКИ

Можно поправить, да будет хуже.

Пословица

Обработав первичные данные, нужно оценить их надежность, чтобы решить, какие из них заслуживают дальнейшего анализа, а какие бесполезны из-за больших ошибок эксперимента.

* Несостоятельны оценки спектральной плотности («спектр мощности»). Чем большее время T мы наблюдаем сигнал $x(t)$, тем шире и наблюдаемый спектр (больше интервал частот $f \geq 1/T$). Поэтому за счет T нельзя улучшить оценку мощности на одной данной частоте. Состоятельные оценки возможны только для сглаженного спектра.

Деление ошибок на систематические и случайные довольно условно. Упоминать о неизвестной постоянной систематической ошибке есть смысл, когда, например, все данные получены сравнением с некоторым «эталоном», и значение этой опорной величины a точно неизвестно, но постоянство ее гарантировано (т. е. наперед известно, что любые возможные колебания a за все время измерений много меньше случайной ошибки самих измерений δx). Тогда можно утверждать, что относительные изменения x содержат только случайную ошибку δx , а в абсолютные значения $x + a$ входит, кроме того, неизвестная систематическая ошибка эталона — постоянное отклонение его от номинала a .

Если мы знаем о постоянной систематической ошибке и можем ее измерить, то надо просто исключить ее из всех результатов соответствующей поправкой. Если же мы не уверены в ее постоянстве (не регистрировали соответствующие параметры), то поправку внести нельзя, и эта ошибка — такая же случайная, как и все остальные. Так, когда опорной точкой отсчета температуры служит точка кипения жидкого азота, надо учитывать ее колебания с атмосферным давлением (до $\pm 0,7^\circ\text{C}$). Если же мы не записываем при каждом измерении давление, то ввести поправку нельзя, и эта систематическая ошибка переходит в число случайных (колебания отсчетов в связи с погодой). Отсюда, в частности, следует рецепт: чтобы повысить точность эксперимента, надо в случайной ошибке найти основной систематический фактор и исключить его изменением условий измерения, компенсацией в схеме либо поправкой после опыта.

Измерение всех величин, входящих в расчет поправки, вносит в нее случайные ошибки. Поэтому когда вводят поправку на систематическую ошибку, суммарная случайная ошибка самой поправки переходит в ошибку «исправленной» величины. Поправок не может быть много (поправки вычитаются, но ошибки складываются!). Если одна и та же ненадежная поправка внесена во многие независимые измерения, они перестают быть статистически независимыми. Одинаковая (или рассчитанная из одних и тех же измерений) поправка создает корреляцию между величинами и потому изменяет закон сложения ошибок: если для независимых x и y разность $z = x - y$ имеет дисперсию $s_z^2 = s_x^2 + s_y^2$, то при коэффициенте корреляции между ними ρ она будет $s_z^2 = s_x^2 +$

$+2\rho s_x s_y + s_y^2$ (и в худшем случае вместо квадратов складываются первые степени ошибки: $s_z = s_x + s_y$). Дальше всюду придется учитывать, что матрица ошибок стала недиагональной. Поэтому полезны лишь поправки, которые существенно больше собственной ошибки (иначе «исправление» только увеличивает суммарную ошибку, ничего не уточняя). Так, если на дифрактометрической линии участок «фона» короче «линии», то процедура отделения фона портит больше, чем все ошибки измерения самой линии.

Для выявления систематической ошибки часто полезно распределение большого числа отсчетов (50—100) изобразить гистограммой. Иногда она имеет четкий физический смысл (например, два максимума для диаметра зерна — свидетельство разноразмерности, и тогда мало смысла в среднем диаметре). Гистограммы полезны и там, где их смысл наперед не очевиден: двугорбое распределение твердости в большой партии исходных отожженных образцов укажет на какую-то неоднородность исходной структуры (например, недогрев части образцов в середине садки).

Если неоднородность исходного состояния не удалось устранить, меняя способ подготовки образцов, у нас два пути: либо отобрать для всех экспериментов только однородные исходные образцы (под одним, по смыслу лучшим, пиком гистограммы), выбросив остальные, либо рандомизировать условия эксперимента: разбив образцы по гистограмме на однородные группы, подбирать на каждый вариант дальнейшей обработки одинаковый набор образцов из всех групп; при этом неоднородность исходного состояния равномерно переводится в случайную ошибку среднего результата, и мы застрахованы от искажения зависимостей из-за неблагоприятной случайной группировки исходных состояний.

Хуже всего, если отбраковкой по исходному состоянию придет нужда заниматься уже после основного эксперимента, а потому, не закончив всю обработку одних измерений, не начинают следующих опытов. Если все-таки исходная неоднородность обнаружена лишь задним числом, приходится систематически проверять, где она повлияла далее. Можно, например, построить сводную для всех обработок зависимость «исходная твердость образца — отклонение его предела текучести (после окончательной обработки) от среднего по группе». Вытяну-

тость такого «облака рассеяния» — признак корреляции. Коэффициент корреляции покажет значимость влияния исходной неоднородности на результаты. Быструю оценку здесь дает ранговый коэффициент корреляции по Спирмену R [1]: если в таблице $y(x)$ дважды перенумеровать все n пар отсчетов (i — по возрастающим x , j — по возрастающим y), то остается только сложить квадраты разности номеров каждого отсчета:

$$R = 1 - 6 \sum (i - j)^2 / n(n-1)(n+1);$$

Коэффициент $R=1$ при однозначной монотонной зависимости y от x . Если зависимость полностью отсутствует, R имеет при больших n нормальное распределение с $\bar{R}=0$ и дисперсией $1/(n-1)$. Если измеренное R значительно отличается от нуля, то y зависит от x . Обычный коэффициент линейной корреляции ρ при нормально распределенных x и y связан с коэффициентом Спирмена R при $n \gg 1$ соотношением $\rho \approx 2 \sin(\pi R/6)$.

Другая трудно учитываемая систематическая ошибка — дрейф условий эксперимента во времени («плывет» напряжение питания, температура среды и т. д.). Дрейф обнаруживают либо длинной серией холостых измерений одного и того же образца (той же «точки»), либо периодическим повторением измерений одних и тех же точек (если надо измерить 20 разных образцов по 4 раза каждый, то лучше измерить сначала все 20 по одному разу, потом — по второму и т. д.).

Монотонный дрейф (неуклонный рост или спад отсчетов) при холостых измерениях «в одной точке» обнаруживается (на фоне случайных ошибок) простой обработкой графика $x(t)$: обозначим каждый отсчет знаком плюс, если он больше и знаком минус — если он меньше предыдущего. Если критерий знаков [1] обнаружит значимое превышение числа плюсов над минусами, x растет. Если измеряли попеременно k разных образцов, то плюсом отметим превышение x_{ik} над предыдущим отсчетом того же образца.

Чаще дрейф — не монотонный: условия колеблются. Даже при питании от сети через электронные стабилизаторы напряжение в схеме реагирует на приближение обедненного перерыва и конца смены; все приборы в лаборатории регистрируют включение печи или пресса у соседей. При таких осцилляциях условий в ряду

++++-----++----- серии плюсов удлиняются

или укорачиваются, и по их числу критерий серий (скачков [1]) выявит периодические флуктуации.

Дрейф не так легко обнаружить, но еще гораздо труднее устранить или учесть поправками. Если вычислять поправку на дрейф по одному «контрольному» отсчету из серии, случайная ошибка ее определения будет больше самой поправки. Линейный дрейф можно измерить, используя весь массив контрольных отсчетов для уравнения регрессии $x(t)$; $[\Delta x_k(t)]$ — если образцы разные].

Для поправки на осцилляции режимов понадобится много больше «холостых» измерений, чем «рабочих». В таких случаях ограничиваются размещением повторных отсчетов для каждого образца в порядке, заданном таблицей случайных чисел (рандомизацией). При этом неизвестная систематическая ошибка дрейфа перейдет в четко измеряемую случайную ошибку. Когда рандомизировать отсчеты трудно (например, в дифрактометрии), делают хотя бы многократное сканирование с возможно большей скоростью, чтобы суммарная линия не искажалась дрейфом.

Если метод измерения таков, что ошибка прямого измерения суммы не намного больше, чем каждого слагаемого, а порядок развертки по координате t (например, по длине спектра) можно менять, то от дрейфа можно избавиться, измеряя в каждый момент суммы отсчетов во многих точках t_i , группируемых попеременно по мультиплексной схеме [2] (спектроскопия по Адамару).

Рекомендательный библиографический список

1. Ван дер Варден Б. Л. Математическая статистика: Пер. с англ. М.: Физматгиз, 1963. 436 с.
2. Сороко Л. М. Мультиплексные системы измерений в физике. М.: Атомиздат, 1980. 120 с.

СЛУЧАЙНЫЕ ОШИБКИ И НАДЕЖНОСТЬ ПЕРВИЧНЫХ ДАННЫХ

Ныне ученые люди... мало взирают на родившиеся в одной голове вымыслы и пустые речи, но больше утверждаются на достоверном искусстве... Мысленные рассуждения произведены бывают из надежных и много раз повторенных опытов.

Ломоносов

Многократные измерения одной величины x в идентичных условиях «в одной точке» позволяют оценить

случайную ошибку воспроизводимости. Она отражает рассеяние результатов от нерегистрируемых условий эксперимента: состояния установки, колебаний ее рабочих режимов, индивидуальных различий между номинально одинаковыми образцами.

Понятие «величина случайной ошибки δx » не имеет смысла, пока не задана надежность — *доверительная вероятность* p того, что истинное значение x_0 лежит в указанных пределах $\bar{x} \pm \delta x$ (*доверительных границах, доверительном интервале*) около среднего $\bar{x}$. (Чем большая требуется надежность, тем шире эти пределы при данном объеме измерений.) *Доверительный уровень* (риск) $\alpha = 1 - p$. Это вероятность того, что истинное x_0 лежит вне указанного поля ошибки.

Оценку случайной ошибки воспроизводимости «в точке» по размаху — разности крайних отсчетов $x_{\max} - x_{\min}$ практически не применяют из-за низкой эффективности (с ростом числа отсчетов размах растет). Существует универсальная мера рассеяния отсчетов при неизменных условиях измерения. При неограниченном повторении измерений величины x на одном образце или на множестве идентичных образцов в одной точке «генеральная совокупность» отсчетов сходится к некоторому «теоретическому распределению» $P(x)$ с дисперсией σ_x^2 . Наша цель — получить оценку этой дисперсии s_x^2 по тем измерениям, какие есть (по выборке). Найти s_x^2 можно и по размаху $|x_{\max} - x_{\min}|$, но лучше — прямым счетом в виде

$$s_x^2 = \sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1). \quad (4.1)$$

Эта оценка — эффективная (с наименьшей ошибкой из всех возможных при данной выборке). С ростом числа отсчетов n величина s_x^2 сходится к «генеральной дисперсии», «дисперсии распределения»: $s_x^2 \rightarrow \sigma_x^2$ при $n \rightarrow \infty$.

Сама по себе величина s_x не отражает повышения точности значения $\bar{x}$ с ростом числа отсчетов n . Польза многократных измерений в том, что среднее $\bar{x}$ имеет дисперсию $\sigma_{\bar{x}}^2 = \sigma_x^2 / n$; так что рассеяние среднего $\bar{x}$ в $\sqrt{n}$ раз меньше, чем рассеяние самих отсчетов x .

Поэтому для характеристики надежности среднего $\bar{x}$, кроме оценки дисперсии $s_{\bar{x}}^2$, нужно указать число отсче-

тов n (число образцов — для воспроизводимости между образцами) либо оценку дисперсии среднего

$$s_x^2 = s_x^2 / n. \quad (4.2)$$

Очевидная разница между двумя оценками: дисперсии s_x^2 самой физической величины x и дисперсии s_x^2 оценки этой величины $\bar{x}$ по n наблюдениям — источник множества недоразумений (особенно если автор не оговорил, что он называл «среднеквадратичным отклонением» — s_x или $s_{\bar{x}}$). Если x — масса, то качество весов характеризует «среднеквадратичное отклонение x » — величина s_x (и она не зависит от того, сколько раз взвешен образец). Если же речь идет о массе данного образца, то после n взвешиваний мы ее знаем с ошибкой $s_{\bar{x}}$ («среднеквадратичное отклонение среднего $\bar{x}$ »), которая, конечно, убывает* как $1/\sqrt{n}$.

Если распределение $P(x)$ нормальное, то с вероятностью p истинное значение x_0 лежит в пределах

$$\bar{x} - t(p, n) s_{\bar{x}} < x_0 < \bar{x} + t(p, n) s_{\bar{x}}, \quad (4.3)$$

где величина $t(p, k)$ имеет распределение Стьюдента (границы двусторонние, $k = n - 1$ — число степеней свободы).

Хотя в формулу (4.3) входит дисперсия среднего $s_{\bar{x}}^2$, поле ошибки все еще зависит, хотя и слабо, от числа отсчетов n :

n	3	5	10	∞
t для $p=0,90$	2,9	2,1	1,8	1,6
t для $p=0,70$	1,4	1,2	1,1	1,0

Для надежности $p=0,9$ нужно указать интервал $\bar{x} \pm 2s_{\bar{x}}$ при пяти образцах на точку и $\bar{x} \pm 3s_{\bar{x}}$ — при трех. Не реже, чем в одном случае из трех, истинное x_0 лежит за пределами поля $\bar{x} \pm s_{\bar{x}}$.

* Уточнения результата не будет, когда ошибка метода s_x меньше «ошибки квантования» прибора — погрешности от неизбежного округления результата до цены последнего деления. Если $n=100$ измерений линейкой (с округлением до 0,5 мм) дали одинаковый отсчет «306 мм», это не значит, что длина составляет $(306,0 \pm 0,05)$ мм, а только гарантирует отсутствие грубых ошибок (промахов) в 1 мм и более.

Неравенство двух величин x и y при равенстве числа отсчетов $n_1 = n_2$ и дисперсий $s_x^2 = s_y^2$ проверяют по отличию от нуля разности $z = \bar{x} - \bar{y}$ при $s_z^2 = 2s_x^2$ и $k = 2n_1 - 2$ степенях свободы.

Поэтому общеизвестная формула « x и y неотличимы, если поля $\bar{x} \pm s_{\bar{x}}$ и $\bar{y} \pm s_{\bar{y}}$ перекрываются или касаются» соответствует (по t -критерию) надежности 0,88 при трех образцах на точку, 0,90 — при пяти и 0,92 — при $n \rightarrow \infty$ (еще раз подчеркнем, что s_x^2 — дисперсия среднего $\bar{x}$, а не распределения x !). Именно поэтому на графиках чаще всего наносят «поле ошибки» $\pm s_{\bar{x}}$.

Результаты связаны в работе в логическую цепь, и при m звеньях надежность окончательного вывода убывает как p^m . Поэтому вряд ли можно допускать надежность отдельного результата ниже, чем $p=0,9$ (что приведет к $p=0,53$ — «ни да, ни нет» при $m=7$ звеньях последовательных преобразований и умозаключений). Вот почему считают разумным делать не менее 3—5 отсчетов на точку.

Есть критерии для выбора решения, но нет критерия для выбора критерия. Приемлемый уровень риска (1— p) задает экспериментатор. Это неформализованное личное решение. Однако представление всех результатов с численной характеристикой их надежности — не украшательство: автор указывает, где и какой произвол им допущен. Без этого результаты разных работ будут несравнимы.

Когда нужно найти среднее из неравноточных измерений x_i одной и той же величины x , полученных в разных условиях и имеющих неравные дисперсии s_i^2 , определяют взвешенное среднее

$$\bar{x} = s_x^2 \sum_i x_i^2 / s_i^2. \quad (4.4)$$

Вес отсчета i — величина s_i^{-2} , обратная его дисперсии. Для нормировки весов на единицу их умножили в формуле (4.3) на дисперсию среднего

$$s_x^2 = \left(\sum_i s_i^{-2} \right)^{-1} \quad (4.5)$$

(чем больше погрешность x_i , тем с меньшим весом этот отсчет входит в среднее).

Когда по той же формуле (4.4) находят среднее $\bar{x}$ по нескольким группам отсчетов со средними $\bar{x}_i$, то s_i^2 — дисперсии средних $\bar{x}_i$.

Если в разных точках i измерены разные значения величины x_i и оценки их дисперсии s_i^2 различаются мало, иногда для характеристики ошибки всех x_i используют одну, объединенную оценку* дисперсии величины x . Для первичного анализа эксперимента лучше показать все s_i^2 , как они есть. Нередко и из физической модели ясно, что дисперсия зависит от координат точки (например, дисперсия межплоскостного расстояния d от брэгговского угла θ) или от величины отсчетов x_i (например, глубины диффузионного слоя, если неровности фронта усиливаются по мере роста). Если даже ожидали однородность дисперсий, то неплохо продемонстрировать видимое различие s_i^2 — как намек на подозреваемый промах, который не удалось исключить на законном основании. В ряде случаев как раз изменения дисперсии несут основную информацию: «двугорбое» распределение микротвердости с очень большой дисперсией — признак двухфазной структуры, рассеяние ударной вязкости растет с началом интервала хладноломкости.

Многое можно понять, построив распределение данных цехового контроля: его асимметрия и выбросы — следы округлений, записей задним числом, отбрасывания или «перепроверки» плохих результатов.

Когда в серии отсчетов есть выбросы (одиночные точки далеко от среднего), обычно подозревают некоторые грубые промахи (ошибки экспериментатора в обращении с прибором, в записях, брак образцов и т. п.). Возникает сильное желание выбросить этот промах из вычислений среднего и ошибки, однако, исключая из совокупности один отсчет как невероятный (промах), мы сразу много теряем в надежности. Так, например, если рассеяние отсчетов x_i симметрично относительно x_0 , x_1 — наименьший, а x_n — наибольший из n сделанных отсчетов, то вероятность наблюдать n отсчетов $x_i < x_0$ и ни одного $x_i > x_0$ составляет $q = (1/2)^n$. Другими словами, это вероятность того, что $x_0 > x_n$ (истинное x_0 больше,

чем самый большой из сделанных отсчетов). Ту же вероятность имеет событие $x_0 < x_1$ (истинное x_0 меньше, чем самый малый из n сделанных отсчетов). Тогда вероятность, что истинное x_0 лежит между измеренными ($x_1 < x_0 < x_n$), будет

$$p = 1 - 2q = 1 - (1/2)^{n-1}. \quad (4.6)$$

Вероятность, что в упорядоченном ряду отсчетов $x_1 < x_2 < \dots < x_n$ ровно r отсчетов x_i лягут справа от истинного x_0 и $n-r$ — слева (т. е. $x_{n-r} < x_0 < x_{n-r+1}$), составит $q(r) = (1/2)^n C_n^r$. Тогда $x_0 > x_k$ с вероятностью $p_k =$

$$= \sum_{r=0}^k q(r), \text{ и точно так же } x_0 < x_l \text{ с вероятностью } p_l = \sum_{r=l}^n q(r).$$

Отсюда вероятность, что x_0 лежит в интервале $x_l < x_0 < x_k$, составляет

$$p(k, l) = 1 - p_k - p_l = 1 - (1/2)^n \left(\sum_{r=0}^k C_n^r + \sum_{r=0}^l C_n^r \right).$$

Положив $l=0$, $k=1$, находим вероятность того, что x_0 лежит в пределах от x_1 до x_{n-1} [поле ошибки $\pm (x_{n-1} - x_1)/2$], а x_n является единственным выбросом

$$p = 1 - (1/2)^n (n+2). \quad (4.7)$$

По сравнению с уравнением (4.6) риск вырос в $(n/2+1)$ раз. Так, например, когда нет других сведений о рассеянии, по пяти образцам на точку положение x_0 между x_1 и x_5 определяется с надежностью $p=0,93$. Если же крайний из пяти отсчетов исключить как промах, то достижимая надежность сразу падает до $p=0,78$. Еще хуже — исключить без дополнительных оснований несколько промахов ($k \neq 0$, $l \neq 0$).

Приведенная оценка надежности — непараметрическая, верная при любом (симметричном) распределении величины x . Если из теории или других измерений достоверно известен вид распределения x и некоторые другие его параметры (математическое ожидание x_0 , дисперсия σ_x^2), то эта информация уменьшает риск при исключении промаха. Однако ситуация, когда x_0 и σ_x^2 за-

* Классическая статистика требует предварительно проверить однородность этих дисперсий по критериям Бартлетта или Кохрена. Однако результат проверки настолько чувствителен к нарушениям нормальности распределения, что практически бесполезен.

ранее точно известны, — практически нереальная (зачем тогда измерения?). В лучшем случае есть доказательство нормальности распределения. Тогда предполагают, что все образцы (или отсчеты), кроме подозреваемого, принадлежат к одной генеральной совокупности («нормальных» образцов), а подозреваемый x_n — к другой («испорченных» образцов — с другим средним и дисперсией). Исключив один этот отсчет, находим оценки среднего $\bar{x}$ и дисперсии s_x^2 по оставшимся и вероятность, что исключенный отсчет принадлежал к общей генеральной совокупности. При небольшом числе образцов на точку эта процедура ненамного лучше непараметрической оценки (4.7), поскольку $\bar{x}$ и s_x определены из тех же измерений. Исключить промах одними лишь методами статистики удастся в редких случаях — при сотнях отсчетов с распределением известного вида, когда в общем известна физическая причина промахов (например, при счете числа квантов много работавший счетчик Гейгера изредка «захлебывается» из-за плохого гашения разряда — получается огромный, в несколько раз выше среднего N выброс, когда распределение «нормальных» N — заведомо пуассоновое).

Никакой отсчет нельзя исключать без четкого обоснования. Для подозреваемого промаха следует прежде всего поискать арифметическую ошибку в расчете. Проверить все вспомогательные измерения (например, входящих в расчет размеров образца). Осмотреть самый образец (намагниченность или пластичность могли выпасть из-за закалочной трещины, твердость — из-за того, что два укола ближе один к другому, чем разрешено ГОСТом и т.п.). Если можно, следует измерить много раз повторно тот же образец. Когда причина промаха достоверно обнаружена, его исключают из всех подсчетов с подробной записью причин в журнале. Если же это не удалось, придется довольствоваться точками с большой дисперсией, лежащими не там, — таково качество работы.

Каково бы ни было обоснование, исключение из результатов некоторого промаха — событие чрезвычайное. Упоминания обо всех исключенных промахах — их числе, причинах и обосновании — должны сохраниться в окончательном описании методики. Заметное число промахов — признак плохого качества образцов или ненадежной техники эксперимента.

Особый случай представляет работа с малым числом дефицитных образцов или быстротечными превращениями, когда нельзя накопить достаточную статистику отсчетов. Статистический анализ надежности результата здесь тем более обязателен, и для этого надо специально спланировать эксперимент: если нельзя повторять отсчеты во всех точках, их надо сделать многократно в избранных точках, наиболее важных для результата, а если в них нельзя, то на концах интервала. В тех редких случаях, когда повторение уникального эксперимента вообще невозможно, для оценки воспроизводимости ставят серию холостых экспериментов. Так, измерить теплоту превращения для единственного образца с редкой структурой можно один раз, но для оценки воспроизводимости надо измерить серию других, одинаковых образцов с близкой температурой и теплотой превращения.

Снимая серию рентгенограмм с одного образца в процессе нагрева, нельзя вернуться и снять одну и ту же точку дважды, а повторные образцы вряд ли удастся снять на той же стадии превращения. Тогда колебания интенсивности и ширины линий «эталона», где ничего не происходит, покажут, измеряли мы в образце степень превращения и дисперсность частиц или же в основном дрейф режима трубки и счетчика, окисление поверхности, сдвиг линий от колебаний температуры и коробление образца. Эталон должен быть, конечно, возможно ближе по форме, размеру, составу к образцу, чтобы уравнивать окисление, теплоотдачу и т.п.

Если невозможны дифференциальные измерения (непосредственное измерение разности, например, э.д.с. совместно нагреваемого образца и эталона), то выгодно измерять образец и эталон попеременно (если только их смена не расстраивает установку) и свойства эталона интерполировать к моменту измерения образца.

Надежность вариантов решения для событий, неповторимых в принципе, например в экономике, сопоставляют по неопределенности исхода (энтропийные оценки). Для эксперимента в физике и технике проверка воспроизводимости всегда возможна и необходима. Если нельзя повторить уникальный эксперимент, оценивают дисперсию результата как функцию дисперсий всех сделанных промеров, которые находят порознь из холостых измерений (но не из априорной погрешности приборов по их паспортным данным!). Чем сложнее опыт, тем боль-

ше проверок на воспроизводимость делают для каждого участвующего в нем прибора и приема измерения. При этом надо прежде всего искать возможности прямо оценить воспроизводимость окончательного результата, ибо нет гарантий от побочных источников ошибки, не учтенных проверкой «по элементам».

Кроме классических определений объемной доли, удельной поверхности и пространственных размеров частиц, методы статистической металлографии [4, 5] позволяют решать и многие более тонкие задачи: по наблюдениям на одном шлифе определять пространственную форму частиц (отличать иглы от пластин), измерять сплошность системы зерен одной фазы (или пор) в пространстве, преобладающую ориентировку частиц, равномерность их распределения и ее анизотропию (строчечность), анизотропию пространственных ориентировок границ зерен и частиц, следов скольжения и трещин; истинные (пространственные) углы между ними и их кристаллографию; степень неравномерности системы зерен по их огранке (числу вершин и статистике ребер в сечении) или по пространственной кривизне границ, текстуры — по ориентировкам фигур травления, мартенситный сдвиг — по рельефу шлифа, извилистость и размах колебаний трещин. Все эти методы используют анализ гистограмм, а их надежность определяется прежде всего объемом измерений.

Автоматические считывающие устройства (типа «Квантимета» или микроденситометра) позволяют быстро набрать массив, достаточный для подробной гистограммы, но это не освобождает нас от статистического анализа воспроизводимости. При большом числе промеров N на первое место выступают ошибки выборки от неоднородности металла и ошибки распознавания границ, частиц и т. п., зависящие от настройки установки и качества шлифа. Первые обнаруживают и оценивают, сравнивая гистограммы от многих полей зрения и от многих номинально одинаковых образцов, вторые — повторяя съемку одного и того же «эталонного» образца с каждой новой партией. Различие между структурами стоит обсуждать, если оно много больше, чем для «эталона» в разные дни съемки.

Основная трудность количественного анализа микроструктур — не в математических методах и даже не в объеме необходимых измерений, а в надежном опознава-

нии элементов структуры (особенно в электронной микроскопии, где анализ делают только со снимков и вернуться к кадру нельзя). Нечеткие детали изображения (частиц, дислокаций, границ) наблюдатель распознает по интуиции. Чтобы оценить субъективные ошибки, при отработке методики количественно сопоставляют надежность опознавания объектов несколькими наблюдателями, понимающими существо задачи.

Предварительно просмотрев серию негативов и попытавшись промерить один, они совместно (и обязательно письменно) формулируют общие для всех кадров правила опознавания: когда надо исключить часть снимка как нечеткую (и как это отметить), что служит признаком объекта (что, например, считать мелкой частицей, а что — дефектом шлифа?), как проводить границы фаз (посередине зоны размытого контраста, по какому краю?), чего нельзя «домысливать» (например, замыкать пунктирные границы, которые, по здравому смыслу, обрываться не могут). После этого каждый из трех наблюдателей независимо очерчивает искомые объекты (например, субграницы) для одних и тех же негативов (на бумаге под фотоувеличителем).

Сложив два рисунка разных наблюдателей, обводим совпадающие контуры одним цветом, несовпадающие — другим, «дорисовки» — третьим, а затем измеряем контуры (или площади) каждого сорта (методом секущих или точечным). Измерения микроструктур следует продолжать, если хотя бы на 90 % объекты у разных наблюдателей совпадают. В противном случае у нас просто непригодные для счета снимки (или не внушающие доверия наблюдатели?). Когда объективность наблюдателя проверена таким образом, он может продолжать все промеры один, время от времени повторяя измерения на одном и том же негативе для проверки воспроизводимости.

Дисперсия от расхождения между наблюдателями входит в ошибку определения абсолютной величины структурной характеристики. Если достаточен сам факт ее изменения, учитывается только ошибка воспроизводимости одного наблюдателя.

Ошибку воспроизводимости для направлений, углов, расстояний на электронограммах также нужно определять многократными промерами (прочерчивая, например, один и тот же угол на разных листках подложенной бумаги).

Допущение об однородности распределения частиц в образце можно проверить, если, положив их распределение на плоскости пуассоновым, сравнить наблюдаемую дисперсию числа частиц в кадре с ожидаемой. Если дисперсия значимо больше, чем пуассонова, это говорит о неслучайном размещении частиц, а если меньше — обычно о субъективном отборе «типичных» полей для съемки.

Когда одно поле зрения (один снимок) содержит много однородных элементов (частиц, фасеток поверхности), можно получать «обобщенный портрет» частицы, не измеряя отдельных частиц: Фурье-преобразованием изображения в оптической системе [6] или как голограмму. Характерные размеры, анизотропию формы или ориентировки частиц измеряют сразу в среднем как детали Фурье-спектра. Из-за конструктивных трудностей в таких измерениях обычно участвует фотография, и тогда для измерения годятся лишь такие детали спектра, которые не пострадают от неконтролируемого нелинейного преобразования сила света — почернение пленки.

Рекомендательный библиографический список

1. Кенуй М. Г. Быстрые статистические вычисления. М.: Статистика, 1979. 70 с.
2. Васманов В. В. Приборы для математической обработки и построения кривых. М.: Машиностроение, 1973. 184 с.
3. Иоаницкий Г. Р., Литинская Л. П., Шихматова В. Л. Автоматический анализ микрообъектов. М.: Энергия, 1967. 224 с.
4. Салтыков С. А. Стереометрическая металлография (Стереология металлических материалов) М.: Металлургия, 1976. 272 с.
5. Чернявский К. С. Стереология в металловедении. М.: Металлургия, 1977. 280 с.
6. Иоаницкий Г. Р., Куниский А. С. Исследование микроструктуры объектов методами когерентной оптики. 2-е изд. М.: Энергия, 1981. 168 с.
7. Катис Г. П. Объемное и квазиобъемное представление информации. М.: Энергия, 1975. 368 с.

ГРАФИКИ, ГИСТОГРАММЫ, СНИМКИ

Балаганов с уважением посмотрел на треугольник. Дозоды Паниковского показались ему не особенно убедительными, но в треугольнике чувствовалась такая правдивая безнадежность, что Балаганов поколебался.

Ильф и Петров. «Золотой теленок»

Для предварительного обсуждения надо построить не только ту зависимость, которую ищем, но и все зависимости, которые могли бы иметь смысл, не только $x(t)$ и

$y(t)$, но также и $(x-y)=f(t)$ и $x(y)$ при одинаковом t или $x(y, t)$ — в трех измерениях. Так, измерив для изотерм превращения долю новой фазы f и средний радиус частиц при нескольких температурах, кроме собственно изотерм $f(t)$ и $r(t)$ при $T=\text{const}$, построим изохроны $f(T)$ и $r(T)$ для разных моментов $t=\text{const}$ (чтобы видеть разницу в скорости на разных стадиях) и те же зависимости для условного числа частиц $N=f/r^3$ и сводные по всем температурам и временам зависимости $N(f)$, $f(r)$ (чтобы рассмотреть связи между степенью превращения и размером частиц). Пусть многие из этих графиков окажутся неинтересными, но построить их гораздо быстрее, чем из общих соображений предугадать, в каком будет больше полезной информации.

Человек гораздо легче опознает закономерность в рисунках, чем в таблицах. Стоит строить график, даже если не видно наперед, почему здесь может быть зависимость. Хуже упустить закономерность на том лишь основании, что мы ее не ожидали. Бывают полезны графики не только искомым, но и первичным или промежуточным величин — по ним, например, можно обнаружить скачки или закономерные колебания условий измерения, которые из таблиц не улавливаются. Так, сведенная на один лист зависимость от календарной даты электросопротивления отожженного эталона, измеряемого с каждой серией образцов, может обнаружить его старение или скачки от неполадок в схеме. Первичные графики нужны и перед обработкой на ЭВМ: прежде чем пускаться в расчеты, надо видеть, стоит ли этот материал обрабатывать или рассеяние и выбросы лишают его всякого смысла.

Если вместо сводного графика получилось «облако точек», это не значит, что он бесполезный. По форме облака можно обнаружить косвенные связи и скрытые зависимости величин. Если на сводном по всем обработкам графике «твердость—относительное сужение образцов» точки распались на два облака, то не указывает ли это на два возможных механизма разрушения при одинаковой твердости? Если образцы с разной температурой отпуска наносили разным цветом, то по цвету пятен сразу видно интервал отпускной хрупкости (по изотермам его могли и не увидеть из-за рассеяния). Если вдруг выявилась бессмысленная корреляция между диаметром отожженной проволоки и твердостью, правильно ли выбрана нагрузка твердомера?

Построение любого (и чернового!) графика начинают с рационального выбора масштаба, нанесения цифр и размерностей на осях и подписи (какая зависимость, из какой серии измерений, когда измерено и построено).

Сначала наносят деления (не меньше трех на каждой оси — с числами) и надписи, только потом — точки. Нельзя путать место запятой: « $R, 10^2 \text{ Ом}$ » означает: «одна единица на шкале = 100 Ом», но « $R.10^2, \text{ Ом}$ » — «одна единица на шкале = 0,01 Ом».

Характерное время разных процессов различается чаще всего по порядку величины. Чтобы они были обозримы вместе, шкалу времени обычно делают логарифмической, при этом нуль шкалы $\lg t$ — это 1 с, 1 мин и т. п., так что состояние при $t=0$ условно наносят левее оси ординат и соединяют с остальным графиком пунктиром. При логарифмическом масштабе вдоль оси обычно идут два ряда цифр:

0	1	2	3	4	$\lg t, \text{ с}$
1	10	10^2	10^3	10^4	

(иногда и три — с минутами и часами).

Две шкалы (линейную $1/T, \text{ К}^{-1}$ и нелинейную $T, \text{ К}$ или $^{\circ}\text{С}$) дают для обратной температуры.

«Точки» никогда не наносят точками — это кружок, треугольник и т. п. — иначе их не разглядишь, когда проведут кривые. Масштаб графиков должен быть таким, чтобы «точка» рисунка была в несколько раз меньше ошибки измерения. Пропорций масштабов по двум осям обычно такие, чтобы график (или их серия в одних координатах) заполнял стандартный лист, потому что глаз лучше всего различает наклоны около 45° .

Для сопоставимых зависимостей на разных графиках принимают один масштаб и одно начало координат, чтобы сравнивать не только «на глаз», но и «на свет». Оси проводят всегда по жирным линиям миллиметровки. 100 кг могут соответствовать 20 или 50 клеточкам миллиметровки, но если взять 30 или 70 клеточек, то построение и чтение превратятся в устный счет с дробями.

На первичные графики наносят, как правило, все точки — для каждого отдельного образца, отсчета — так виднее выбросы и другие странности. Лишь после этого отмечают средние (быстрее всего — считать прямо по клеточкам).

На чистовые графики для обсуждения при 3—5 и более образцах на точку обычно наносят только средние значения, а в качестве интервала рассеяния — чаще всего поле $\pm s_{\bar{x}}$ (меру рассеяния среднего $\bar{x}$, а не самих отсчетов x_i , что оговаривается и в тексте). В сводных же таблицах лучше приводить дисперсию самого x и число образцов на точку — иначе трудно проверить однородность рассеяния x в точках с разным числом образцов. Иногда вместо $\pm s_{\bar{x}}$ наносят доверительный интервал $\pm \delta x$, но это дополнительное преобразование редко необходимо, поскольку δx и $s_{\bar{x}}$ довольно жестко связаны формулой: $\delta x = t(p, n) s_{\bar{x}}$.

Если одинаково существенны ошибки и аргумента и функции, наносят точку и «крест» ошибок или прямоугольник вместо точки.

Нужно продумать, что будут означать вторичные графики отношений или разностей. Так, если два близких профиля линий вычесть один из другого, предварительно нормировав на единицу площадь под линией, мы выделим в чистом виде сдвиги в координатах и интенсивности и далее можем анализировать только изменения профиля.

Не злоупотребляйте «изменениями в процентах». Даже если измеряли как раз отношение искомой величины к ее исходному значению $x(t)/x(0)$, лучше обсуждать $x(t)$, измерив $x(0)$ независимо, ибо чаще всего физический смысл имеют изменения абсолютных величин: когда коэрцитивная сила выросла с 40 до 50 Э в одном случае и с 45 до 55 Э — в другом, какой смысл утверждать, что «относительный прирост снизился с 25 до 22 %»?

Первичные гистограммы строят обычно для абсолютных частот (числа отсчетов n_i в i -том разряде). Для анализа удобнее относительные частоты («частости») $p_i = n_i/N$, где $N = \sum n_i$, но переходя к p_i , надо не забыть указать суммарное N для оценок надежности.

При равномерной разбивке измеряемой величины x на разряды наименьшее среднеквадратичное отклонение найденного распределения $p(x)$ от истинного достигается * при числе разрядов $k \approx \sqrt[3]{N}$ и убывает с числом

* Ченцов Н. Н. Статистические решающие правила и оптимальные выводы. М.: Наука, 1972. 520 с.

всех отсчетов N как $1/\sqrt[3]{N}$. Таким образом, для $N=100$ отсчетов оптимальное число разрядов $k=5$, а при $N \leq 27$ $k \leq 3$, и вряд ли вообще стоит строить гистограмму. Чтобы имела смысл более подробная гистограмма с 10 разрядами, надо не менее 1000 отсчетов. Для m -мерных гистограмм* оптимальное $k \approx N^{m/(m+2)}$, а уклонение убывает как $N^{1/(m+2)}$. Для двумерной гистограммы с $k=5 \times 5=25$ группами надо набрать не менее 625 отсчетов, а чтобы снизить погрешность вдвое, объем измерений следует увеличить в 16 раз.

Когда x меняется более, чем на порядок (например, диаметр зерна), равномерная разбивка по x не годится, и строят (равномерную же) гистограмму для $\lg x$. Если же объединять «для гладкости» соседние разряды, то из одного набора данных можно получать гистограммы любого вида.

Чтобы сравнить гистограммы (естественно, при одинаковой разбивке x на разряды) нужно указать некоторые поля рассеяния $\pm s_{p_i}$. Если бы n_i не были связаны (условием $\sum n_i = N$), то дисперсия их была бы пуассоновой: $s_{n_i}^2 = n_i$, а дисперсия частоты $s_{p_i}^2 = p_i/N$. Это соотношение иногда используют для грубых оценок. Обычно о теоретическом распределении $p(x)$ наперед ничего не известно, и задача заключается в сравнении двух гистограмм. Если первая гистограмма содержит n'_i отсчетов в i -том разряде, а всего в k разрядах $N' =$

$= \sum_{i=1}^k n'_i$ отсчетов, а вторая — n''_i и N'' , соответственно, эти

гистограммы можно с риском (уровнем значимости) α считать различными, когда критерий

$$\chi^2 = \frac{1}{q(1-q)} \left[\left(\sum_{i=1}^k n'_i q_i \right) - qN' \right]$$

превышает заданный уровень $\chi^2(\alpha)$ при $k-1$ степенях свободы. Здесь $q_i = n_i/(n'_i + n''_i)$ — «вес» гистограммы в разряде i , $q = N'/(N' + N'')$ — ее «общий вес». Для выборки равного объема $N' = N''$ правило: «гистограммы достоверно отличимы, если в каждом разряде поля ошибки $p_i \pm s_{p_i}$ не перекрываются» равносильно условию $\chi^2 \geq 2k$, что дает надежность $p \approx 0,95$ при $k=5$ разрядах и

$p \approx 0,98$ — при 10 разрядах гистограммы. Если же, наоборот, любое значение p'_i попадает в поле $p_i \pm s_{p_i}$, то $\chi^2 \leq k/2$ и нельзя утверждать, что гистограммы различны (риск ошибки 0,73 и 0,88 соответственно).

Убедившись таким образом в однородности двух серий измерений (скажем, размеров зерен на двух шлифах), гистограммы можно объединить (конечно, складывая n_i , а не усредняя p_i).

Все черновые графики подколоты или вклеены в журнал. В отчете графики приводят в таких координатах, где зависимости наиболее четко видны. Надо показать все первичные зависимости от «естественных» аргументов: структуры и свойств — от температуры и времени превращения и т. п.

Для анализа вида зависимости удобны такие координаты, где она ожидается линейной (и кривизна сразу покажет, где есть отклонения от модели). Если приведена некоторая зависимость $z(x)$ для одного варианта условий (температуры, состава), то для сопоставлений надо показать $z(x)$ и для остальных вариантов, даже если они «не смотрятся». Все эти зависимости можно совмещать на одном листе, если точки четко изображены разными значками и не совпадают, а кривые пересекаются редко. Если вариантов много, лучше дать несколько графиков рядом на одном листе (конечно, в одном масштабе и с одним положением нуля).

Из дополнительных графиков, перестроенных из первичных, приводят только те, что используют в обсуждении.

Сколько бы ни было экспериментальных точек, на окончательном графике показывают все — даже 256 или 512 точек спектра ядерного гамма-резонанса. График без точек — бессмыслица.

В тексте должно быть четко сказано, какая именно величина нанесена в качестве поля ошибки на каждом из графиков.

Когда точек много, а дисперсии почти одинаковы (например, для дифрактометрических линий), можно показать поле ошибки лишь в избранных, критических точках графика (включая точки с наибольшей и наименьшей ошибкой).

Не усредненные значения наносят, когда данных мало (1—2 образца на точку) или когда приходится обобщать выборы.

Точки отражают экспериментальные факты, а кривые на графиках — мнение экспериментатора об этих фактах. Это мнение не должно противоречить результатам анализа ошибок. Если через все поля можно провести прямую, то и должно проводить прямую, хотя и ожидали совсем другую зависимость (рассеяние результатов таково, что никаких сведений о кривизне в экспериментах не содержится). Все кривые на графике — простейшие из возможных в пределах полей ошибок (чтобы показать максимум, необходимо статистически значимое превышение точки над соседними). Поэтому не нужно соединять точки ломаными. Если есть всего 2—3 точки, лучше вместо графика $x(t)$ дать ряд столбиков с аргументом $t = \dots$ под каждым, а поля ошибки изобразить штриховкой.

Зависимость от качественных факторов показывают только сериями столбиков, но не графиком: нельзя соединить ни кривой, ни ломаной точки, изображающие свойства после охлаждения в воде, масле и на воздухе, если не измерены средние скорости охлаждения.

«Хотя различие в величине z для t_1 и t_2 не выходит за пределы ошибки, на всех трех графиках $z_2 > z_1$; очевидна тенденция к повышению z » — это философствование о тенденциях при отсутствии фактов. Если автору кажется, что здесь что-то должно быть, он волен повторить измерения столько раз, чтобы различия стали статистически значимыми. Если кажется, что слабое различие повторяется в разных сериях экспериментов, надо показать что объединенная по всем сериям оценка разности ($z_2 - z_1$) значимо больше нуля*. Пока это не сделано, нет оснований рисовать и обсуждать подъем кривой на графиках.

Изображение многомерных зависимостей требует нескольких проб. Если точки $z(x, y)$ размещены вдоль плоских разрезов ($x = \text{const}$ или $x/y = \text{const}$), то сначала проводят эти сечения в виде пачки кривых на плоскости и по ним строят в аксонометрии поверхность $z(x, y)$, обрисованную сеткой линии $x = \text{const}$ и $y = \text{const}$ (как, например, на диаграммах рекристаллизации). Хуже, если

* Если разность средних не значима из-за рассеяния между образцами, но известны разности $\Delta_i = z_2 - z_1$ для каждого образца, непараметрический критерий знаков может обнаружить факт $z > 0$, не используя оценки дисперсии z_1 и z_2 .

серии точек не попадают на плоскости, параллельные оси z . Иногда удается и аксонометрия с частоклом высот, восстановленных из каждой точки на плоскости (x, y), но чаще приходится показывать «вид сверху»: нанести все точки на плоскость (x, y) и написав число z около каждой из них, выделить разные уровни z размером, цветом или чернотой точек, а затем проводить систему горизонталей $z = \text{const}$ между точками, указывая, как у топографов, высоту уровня в разрывах горизонталей. Так, например, изображают поверхность солидуса тройной системы в концентрационном треугольнике. Иногда легче сначала найти точки $z = \text{const}$ на некоторых (не обязательно плоских) разрезах, а уж по ним нанести горизонталей $z = \text{const}$ на карту. Так отображают текстуру на полюсных фигурах. Высоту полезно оттенять штриховкой или цветом. При любом способе изображения обязательно показывают все первичные точки.

Кроме графиков с точками, в отчете приводят скопированные на кальку первичные непрерывные записи — дилатометрические кривые, диаграммы деформации, дифрактометрические линии и пр. Если обсуждаются лишь некоторые точки этих кривых (температура превращения, предел текучести), то приводят только 2—3 рисунка в описании методики. На каждом из них пачка кривых для одинаковых условий (чтобы показать рассеяние), а разные пачки демонстрируют крайние (по условиям и свойствам) случаи, чтобы был виден весь диапазон различий. Все кривые, конечно, приведены к одному началу и масштабу.

Если же анализируется самый вид кривых (например, профили линий), то надо приводить их все, по порядку сопоставления, стопкой одну над другой при едином масштабе и общем нуле абсцисс.

В небольшом отчете помещают все имеющиеся доброкачественные фотографии структур, электронограммы, рентгенограммы. При сотнях снимков приходится отбирать «типичные». Это можно сделать только после полного (письменного) обсуждения всех без исключения снимков. Все рассуждения о сходстве структур (в разных сплавах или состояниях) и об их систематических различиях должны быть подкреплены соответствующими сериями кадров. Приводят и комментируют также все заслуживающие обсуждения (или непонятные автору) аномалии структуры и снимки с типичными артефак-

тами (морщинами реплик, выкрошенными включениями и т. п.), мешающими пониманию структур.

Под каждым снимком указывают увеличение: « $\times 1000$ » (или « $\times 400 \times 2$ », если оно складывается из увеличения микроскопа и фотоувеличения с негатива). Когда приводят снимки с разным увеличением или обсуждают размеры элементов структуры, на сами снимки наносят масштабный штрих (черту с надписью «1 мкм») — одной и той же истинной длины (в мкм) на всех сравниваемых снимках.

Одну-две обсуждаемые детали структуры можно указать на снимке стрелкой с буквой (тушью), но если их больше, лучше обвести и обозначить эти детали на кальке, наклеенной рядом или, еще лучше, за один край точно над снимком. Так же указывают кристаллографические направления на электронно-микроскопических снимках.

Снимок прорисовывают на кальку полностью, когда нужно противопоставить две группы трудно различимых объектов (границы и субграницы, похожие частицы разных фаз).

Так же на кальке индицируют электронограммы и рентгенограммы; над дебаеграммами рисуют штрих-рентгенограмму.

Под любым рисунком обязательно подпись — настолько исчерпывающая, что можно ознакомиться с отчетом по рисункам, не читая. В каждой подрисункной подписи перечислены все обозначения.

КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ЭКСПЕРИМЕНТАЛЬНЫХ ЗАВИСИМОСТЕЙ

Даже качественный эксперимент требует количественного анализа, чтобы установить надежность результата (достоверность факта). Апостериорный количественный анализ эксперимента имеет следующие цели (в ряду по возрастающей сложности):

- указать доверительный интервал для некоторой прямо измеренной величины;
- оценить достоверность ее изменений;
- доказать существование некоторой зависимости (корреляционной связи) между измеренными величинами;
- аппроксимировать эти зависимости для последую-

щей экстраполяции, дифференцирования и интегрирования, поиска экстремума;

д) показать согласие наблюдений с некоторым физическим законом [«теоретической зависимостью» $y(x)$];

е) определить входящие в этот закон константы и доверительные интервалы для них (параметризовать гипотезу).

Соответствующим задачам математической статистики посвящены обстоятельные руководства. Мы рассмотрим лишь «способ употребления» (выбор статистического метода) и интерпретацию результатов статистических вычислений.

Критерии для проверки простых гипотез и условия их применения сведены в таблицу.

СТАТИСТИЧЕСКИЕ КРИТЕРИИ ДЛЯ ПРОВЕРКИ ГИПОТЕЗ

Дано	Гипотеза	Критерий
НР; a НР; неизвестные $\sigma_x^2 = \sigma_y^2$	$x_0 = a$ $x_0 = y_0$	t -Стюдента
—	$x_0 = y_0$	Вилкоксона
—	$x_0 = y_0$	Серий
—	$x_0 = y_0$	Знаков
НР; σ^2	x и y независимы	R -Спирмена
НР	$\sigma_x^2 = \sigma^2$	χ^2 -Пирсона
	$\sigma_x^2 = \sigma_y^2$	v^2 -Фишера (F)

Здесь x_0 — математическое ожидание наблюдаемой величины x , σ_x^2 — ее дисперсия, НР — нормальное распределение x с плотностью вероятности

$$p(x) = (\sigma \sqrt{2\pi})^{-1} \exp[-(x - x_0)^2 / 2\sigma^2];$$

«дано» означает априорную информацию. Все отсчеты в совокупности $\{x\}$ предполагаются статистически независимыми. Один и тот же критерий может быть односторонним, если альтернативная гипотеза выражена одним неравенством (например, возможно либо $x_0 > a$, либо $x_0 = a$, но заведомо невозможно $x_0 < a$), или двусторонним, если как альтернатива к $x_0 = a$ равно приемлемы оба неравенства: $x_0 < a$ или $x_0 > a$.

Как видно, для наиболее употребительных критериев (v^2 -Фишера и его частные случаи: критерии t -Стюдента и χ^2) нужно нормальное распределение. Критерий Стюдента устойчив при умеренных отклонениях от нормального распределения, но теряет надежность, когда отсчеты зависимы; χ^2 — критерий не устойчив к нарушениям нормальности.

Для проверки различия средних t критерий Стюдента легко заменить непараметрическим критерием Вилкоксона, который и без допущения о нормальности и равенстве дисперсий имеет почти ту же мощность (чтобы отвергнуть неправильную гипотезу по критерию Вилкоксона, нужно набрать 22 отсчета в таких условиях, когда по Стюденту достаточно 21 отсчета [11]). Широкое пространство t -критерия объясняется тем, что x и s_x^2 все равно нужно вычислять как первичные величины для дальнейшей обработки. Для анализа дисперсий по v^2 - и χ^2 -критериям нормальность необходима.

Нормальное распределение в природе встречается очень часто — всюду, где x рассеяно в результате множества независимых возмущений примерно равной силы. В нормальное вырождаются при большом числе измерений также биномиальное распределение (для многократных независимых выборок из конечного числа дискретных элементов) и пуассоново распределение (для числа независимых случайных событий из однородного потока за определенное время). Есть, однако, и заведомо не нормальные физические величины (например, в однородной структуре нормально распределен не размер зерна, а его логарифм). Когда x сравнимо с σ_x , никак нельзя полагать нормальным распределение величины, которая в принципе не бывает отрицательной (например, концентрация или объемная доля) — иначе допускается большая вероятность, что $x < 0$.

У экспериментатора три пути обоснования нормальности распределения:

- а) из физической модели явления (например, число сосчитанных квантов N имеет пуассоново распределение, которое при $N \rightarrow \infty$ должно переходить в нормальное);
- б) по совокупности предшествующих измерений по той же методике (например, по анализу журнала испытаний за несколько лет);
- в) по собственным данным.

Если нормальность гарантирована физикой явления или статистикой многих предшественников, ее можно не проверять по статистическим критериям, но строить распределение $p(x)$ все равно полезно, чтобы вовремя заметить грубые нарушения методики: выбросы, раздвоение $p(x)$ от неоднородности и т. п. Для проверки нормальности в одной точке (с одним x) по χ^2 -критерию нужно не менее 50 отсчетов, и часто легче получить такой массив в специально поставленных холостых экспериментах, чем из основных измерений. Если использовать много разных точек (с разными и средними $\bar{x}_i$ и дисперсиями σ_i^2), то в принципе достаточно трех отсчетов на точку (две степени свободы расходуются на определение среднего и дисперсии), но необходимо около тысячи точек.

Нормальность доказана лишь для того интервала, где были наблюдения x ; для более широкого интервала по ним можно оценить интеграл $\int_{-\infty}^{\infty} p(x) dx$, но не вид

$p(x)$ — распределение выбросов больше $x_{\max}$ предсказать таким образом нельзя.

Доказательство нормальности — не самоцель, а обоснование применимости статистических критериев. Из всех распределений с равной дисперсией нормальное приносит наименьшую информацию о величине x , и потому допущение нормальности не занижает оценку ошибки (4.3), но может ее завесить. Если оценки уже сделаны, то доказательство нормальности лишь дает уверенность, что при наших измерениях результат и не мог быть точнее.

Когда совокупность измерений обнаруживает факт существования некоторой зависимости $y(x)$, это, как правило, еще не окончательный результат. Его надо анализировать дальше. Для технологических целей (чтобы отобрать важнейшие факторы, нащупать область оптимума) достаточно аппроксимировать $y(x)$ некоторым выражением $Y(\alpha_j, x)$, вид которого выбран лишь по соображениям удобства дальнейших вычислений, а параметры α_j не имеют какого-либо физического смысла.

Если речь идет только об аппроксимации, то удобнее всего полиномы $Y = P(x)$. Их степень выбирают по виду $y(x)$: если это явно не прямая, начнем с $n=2$, если есть и максимум, и минимум — с $n=3$. Вместо полиномов

высокой степени для сложных кривых и поверхностей выгоднее применять кусочно-параболическую аппроксимацию (метод сплайнов [7]), задавая условия гладкости на стыках парабол. Когда x меняется на несколько порядков, можно ввести в полином аргумент $z = \lg x$ (может быть функцию $u = \lg y$), но нет пользы от «подбора эмпирических формул» более сложного вида, если для них нет ясного физического обоснования.

Цель физического эксперимента — подтвердить существование зависимости определенного вида (выбрать вид $Y(\alpha_j, x)$ из нескольких допускаемых теорией альтернатив) или только найти физические константы α_j для известного закона (или то и другое вместе).

Если вид закона $Y(x)$ в теории не известен, то прежде всего надо сформулировать гипотезу с наименьшим числом m параметров α_j так, чтобы по смыслу задачи все искомые α_j были взаимно независимы (попытка раздельного определения двух параметров, в природе сильно связанных, дает плохо определенный ответ при любом способе вычислений). Наилучшее в этом смысле представление* — когда искомые α_j есть коэффициенты разложения $Y(x)$ по системе ортогональных функций $P_j(x)$:

$$Y(x) = \sum_{j=1}^m \alpha_j P_j(x). \quad (4.8)$$

Если, измерив $y(x)$ в n точках $(x_1, x_2, \dots, x_n)$, решить линейную относительно α_j систему n уравнений (4.8), то найденные параметры α_j окажутся статистически независимы лишь в том случае, когда для функций $P_j(x)$ выполнено условие ортогональности с весом (при неравной точности y_i), которое на дискретном множестве n точек отсчета x_i имеет вид

$$\sum_{i=1}^n s_i^{-2} P_l(x_i) P_k(x_i) = 0 \quad (l \neq k). \quad (4.9)$$

* Другое, вычислительное преимущество представления (4.8): хотя искомые зависимости $Y(x)$ нелинейные, неизвестные α_j связаны с наблюдаемыми y_i линейно, так что решение всегда единственно, а дисперсии s_j^2 всех параметров α_j определяются из той же матрицы системы. Это достоинство практически очень важно: если $Y(\alpha_j)$ нелинейно, придется найти все множество решений системы нелинейных уравнений (и универсального алгоритма здесь нет), а затем среди них искать единственное физическое решение.

Важно, что для данного набора точек наблюдения $\{x\}$ набор функций $P_j(x)$ не зависит от величины ожидаемых отсчетов $y(x_i)$ в этих точках. Но он зависит от ошибок этих отсчетов: формула (4.9) содержит *веса точек* $w_i = 1/s_i^2$ (чем хуже измеряется y в точке x_j , тем меньше учитывается ее вклад).

В принципе всегда возможно для любого набора точек $\{x\}$ и их весов w_i найти любое число ортогональных полиномов $P_j(x)$, определив их коэффициенты численно. Если есть возможность выбора, точки наблюдения x_i обычно расставляют с равным шагом и задают в них равные дисперсии s_i^2 (регулируя, если надо, число или длительность отсчетов). Выигрыш в том, что на множестве равноотстоящих точек равного веса ортогональным разложением служит такой же тригонометрический ряд Фурье, как и для непрерывной функции на отрезке (разница лишь в том, что значения y на концах $x_{\max}$ и $x_{\min}$ надо вводить в расчет α_j с весом $1/2$).

Сегодня любая численная обработка данных много быстрее и дешевле самих измерений. Поэтому всегда нужно обрабатывать эксперименты наилучшим методом и до полного извлечения всей информации. Лучше всего (без потерь) используют информацию методы, основанные на принципе максимума правдоподобия: «наилучшим описанием явления будет то, которое дает наибольшую вероятность получить в результате измерений именно те значения, которые и были фактически получены» [2]. Методы, основанные на этом принципе, дают для искомых параметров α_j оценки, которые достаточны (используют всю содержащуюся в эксперименте информацию об α_j) и эффективны (имеют наименьшую дисперсию из всех возможных при любых способах решения) [2].

Если отсчеты в каждой точке y_i имеют нормальное распределение, то принцип максимума правдоподобия приводит к методу наименьших квадратов (МНК) с весом: для данной гипотезы $Y(\alpha_j, x)$ наилучшие α_j те, которые дадут минимум квадратичной формы

$$M = \sum_{i=1}^n w_i [y_i - Y(\alpha_j, x_i)]^2. \quad (4.10)$$

Вес $w_i = \sigma^2/s_i^2$ учитывает неравноточность отсчетов.

Сумма весов нормирована на единицу $\left(\sum_{i=1}^n w_i = 1\right)$ умножением на коэффициент $\sigma^2 = \left(\sum_{i=1}^n s_i^{-2}\right)^{-1}$, характе-

ризующий дисперсии первичных данных в совокупности. Рассеяние же всех наблюдений относительно найденной линии $Y(x)$ описывает оценка дисперсии линии регрессии [мера уклонения экспериментальных точек от найденной кривой $Y(\alpha_j, x)$]:

$$s_{II}^2 = M_{\min}/(n-m). \quad (4.11)$$

Если отклонения найденной кривой $Y(x)$ от экспериментальных точек обусловлены лишь их случайным рассеянием, то s_{II}^2 и σ^2 должны совпадать. Если же отношение $v^2 = s_{II}^2/\sigma^2$ значительно отличается от единицы (по v^2 -критерию Фишера*), то найденная зависимость $Y(\alpha_j, x)$ неприемлема — гипотеза не согласуется с фактом, и надо менять вид предполагаемой зависимости $Y(x)$.

Когда гипотеза $Y(\alpha_j)$ линейная, оценки дисперсии всех искомых параметров α_j находят просто умножением обратной матрицы системы на дисперсию (4.11). При этом если первичные отсчеты y_i равноточные, то и все α_j статистически независимы (что важно для дальнейших вычислений, где α_j послужат первичными данными).

Неудобство нелинейных задач МНК — отсутствие простых выражений для дисперсии искомых параметров α_j . Хотя в принципе их можно определить из соотношений типа

$$s_{\alpha_j}^2 \approx \sum_{i=1}^n (\partial \alpha_j / \partial y_i)^2 s_i^2,$$

даже сами производные редко можно выписать в явном виде. Поэтому в нелинейных задачах часто удобнее искать $s_{\alpha_j}^2$ непосредственно по изменению M : имея решение $\{\alpha_j^0\}$ варьируем поодиночке α_j , пока отношение $M(\alpha_j^0 + \Delta \alpha_j)/M(\alpha_j^0)$ не превысит заданного по v^2 -кrite-

* Для числителя число степеней свободы $f_{II} = n - m$, для знаменателя же $f_I = \sum_{i=1}^n f_i$ зависит от происхождения оценок f_i .

рию уровня. Соответствующее $(\Delta \alpha_j)^2$ — оценка дисперсии α_j . При этом обязательно искать границы для отклонения $\Delta \alpha_j$ обоих знаков — сильная асимметрия обычно указывает, что приемлемое решение не единственно (рядом с ним есть ненайденный алгоритмом другой «подводный минимум»).

Теоретическую зависимость можно уточнять, вводя в нее дополнительные члены, но с каждым из них входит новый неизвестный коэффициент α_j , а общее число коэффициентов m ограничено: $m < n$. При $m > n$ неизвестных будет больше, чем независимых уравнений, число которых n равно числу экспериментальных точек. При $m = n$ кривая $Y(x)$ проходит точно через все точки y_i — нельзя ни проверить согласие решения с экспериментом (по v^2 -критерию), ни оценить ошибки α_j .

Среди всех m есть одно — оптимальное — число параметров v , при котором оценка дисперсии регрессии s_{II}^2 по формуле (4.11) наименьшая (принцип Гаусса). Дальнейшее уточнение вида зависимости $Y(x)$ увеличивает общую неопределенность решения и отдаляет найденную зависимость от истинной. Когда $m > v$, для некоторых α_j ошибка $s_j > |\alpha_j|$. Поэтому важно вовремя остановиться: вычисление лишних (незначимых) параметров α_j не только бессмысленно, но и вредно, так как ухудшает и надежные α_j , и общее согласие решения с экспериментом (так, вычисляя более v гармоник профиля линии, мы его портим, поскольку уточнения статистически незначимы, а вносимая с ними случайная ошибка синтезированного профиля растет). Не во всякой задаче сразу виден «естественный» порядок номеров j , но его надо задать, прежде чем приступать к «упорядоченной минимизации риска» [4].

Вообще говоря, восстановление любой зависимости $y(x)$ по ее наблюдениям в отдельных точках x_i — задача некорректная: всегда существует бесконечное множество сильно различающихся решений, не противоречащих наблюдениям. Так, при любой расстановке точек невозможно доказать, что между смежными точками x_1 и x_2 не было высоких узких пиков. Для разумного однозначного решения надо ограничить класс функции, на котором его ищут. Такую регуляризацию [4] можно задать чисто математически (например, потребовать достаточной гладкости решений), либо обосновать ее из некоторых физических соображений. В любом случае экспери-

ментатор вносит априорную информацию, не вытекающую из наблюдений, и важно четко указать, где эти допущения. Так, отбрасывая незначимые гармоники по условию $|\alpha_j| < Ks_j$, мы вводим в качестве параметра регуляризации K , т. е. можем выбирать надежность ($K=1$ означает, что риск сохранить член с $\alpha_j=0$ составляет 0,32).

Если некоторый параметр α_k оказался незначимым, надо упростить гипотезу $Y(x)$, положив $\alpha_k=0$, и новый результат будет точнее. Все, что удалось аппроксимировать прямой, не пытайтесь описать зависимостью с большим числом параметров m (даже если на то есть физические соображения). Так, если по v^2 -критерию верно, что глубина диффузионного слоя y линейно зависит от времени t , не надо загонять эту зависимость в ожидаемую параболическую $y \sim \sqrt{t}$ (вычислить параметры параболы можно, но один из них или оба не будут значимо отличаться от нуля).

Метод наименьших квадратов применим и к многомерным задачам, когда Y функция нескольких независимых аргументов $\{x\}$ в каждой экспериментальной точке или, что то же самое, функция векторного аргумента x . Новая трудность здесь в том, что возможна корреляция аргументов x'_i и x''_i в первичном массиве (что легче всего увидеть на графике). Так, может оказаться, что температура и время выдержки сильно связаны: для низких температур были только большие выдержки, для высоких — только маленькие. Связанные аргументы делают решение неопределенным; если существует корреляционная связь $x' \approx 2x''$, то $y = x' + 5x''$; и $y = 10x' - 13x''$ — одинаково хорошие решения. Система уравнений регрессии плохо обусловлена и поэтому неразрешима*. Исправить дело можно только соответствующим выбором экспериментальных точек (до измерений — методами планирования эксперимента).

Выбор вида $Y(x)$ всегда ограничен принципом Гаусса и желанием получить возможно менее связанные α_j . Но если для аппроксимации главное — простота выражений (легкость вычислений), то для параметризации

* Иногда сильно связанные аргументы целесообразно заменить их независимыми линейными комбинациями — главными компонентами, которые находят приведением матрицы ковариаций аргументов к диагональному виду [1].

гипотез параметрам α_j нужен четкий физический смысл. Поэтому часты физические гипотезы, где $Y(\alpha_j)$ нелинейно. Некоторые из них преобразуются в линейные простой заменой переменных. Так, для температурной зависимости коэффициента диффузии $D(T)$ естественные координаты $y = \lg D$, $x = 1/T$. Вид распределения y_j при этом меняется, и дисперсия $s_{y_i}^2 = s_{D_i}^2 / D_i^2$ будет сильно зависеть от D_i . Сходный случай: определение периода решетки a по нескольким линиям. Координаты центров тяжести θ_i определены (в угловой мере) почти с равной точностью, но дисперсия периода a будет сильно зависеть от θ (наиболее точно a по задним линиям), так что для экстраполяции к $\theta=90^\circ$ нужна регрессия с весом.

Иногда известно, что в разных областях x физический закон будет разным, но граница этих областей наперед неизвестна (например, диаграмма усталости $\sigma_{-1}(\lg N)$ изображается двумя прямыми, но точку перелома надо определить из эксперимента). В этом случае мы несколько раз проводим методом наименьших квадратов два отрезка, систематически смещая их границу, пока не найдем по формуле (4.10) минимум M (на действительной границе области).

Наиболее трудны случаи, когда физически обоснованная зависимость $Y(\alpha_j)$ не линеаризуется ни функциональным преобразованием, ни разложением в ряд. Так, когда для анализа профиля линий преобразованием свертки отделяется линия «эталона», центр тяжести «физического профиля» x_k входит в его описание через

$$y(x) = \sum_{j=1}^v a_j \cos j(x - x_k), \text{ и нет разложения в ряд по } x_k,$$

пригодного для всех j . Сам метод наименьших квадратов применим и для таких задач, но уравнения становятся нелинейными, для них нет универсального алгоритма и неизвестно даже число решений. Физическое решение одно — в абсолютном минимуме M . Чтобы найти его, нужно найти все решения. Хотя известен алгоритм поиска абсолютного минимума M в m -мерном пространстве с доказанной сходимостью, нет даже оценок скорости сходимости. Поэтому практически очень важно иметь удачную программу МНК — решения нелинейных систем, которая бы автоматически перебирала минимумы M , не «застывая» в них. Такая программа не универсальна, а создается «под задачу».

С точки зрения статистики одинаково хороши все минимумы M , приемлемые по χ^2 -критерию. Если таких решений опять несколько, значит, точности эксперимента недостаточно для решения физической задачи, и тогда нет смысла выбирать из них «лучшее по физическим соображениям»: если мы знали наперед, как это сделать, соответствующие ограничения на область существования или соотношения между параметрами надо было заложить в алгоритм. Тогда и программа стала бы быстрее — одно лишь упорядочение параметров (если заранее известно, что $\alpha_1 < \alpha_2 < \dots < \alpha_m$) сужает m -мерное пространство поиска в 2^m раз. Если же наперед этого сказать нельзя, то нечего гадать и по набору решений: надо увеличить точность измерений, чтобы абсолютный минимум отличался от следующего ближайшего к нему по χ^2 -критерию.

Всякое найденное решение надо, конечно, показать на графике, синтезировав кривую по ее параметрам. Этим мы, во-первых, проверяем правильность всех вычислений (если ошибка была еще при вводе данных, ее не обнаружат никакие статистические критерии), а, во-вторых, уясняем себе физический смысл решения (иногда неожиданный: предполагали два максимума, а оказалось, что лучшее соответствие с экспериментом дает один широкий пик, покрывший их вместе).

Применение метода наименьших квадратов исходит из допущения, что ошибки аргументов x_i влияют на результат много меньше, чем ошибки их функций y_i . Это обычно верно, когда x — время, температура, угол рассеяния. Если измерить аргумент ничуть не легче, чем функцию, то принцип максимума правдоподобия приводит не к методу наименьших квадратов, а к уравнениям конфлюэнтного анализа [2], где на равных правах учтены дисперсии $s_{x_i}^2$ и $s_{y_i}^2$.

Когда распределения y_i не нормальные, но с конечной дисперсией, оценка МНК все равно дает наименьшую дисперсию α_j из всех возможных несмещенных оценок, если зависимость $Y(\alpha_j)$ — линейная [2].

Еще одно (неявное) допущение МНК: дисперсия σ_i^2 в данной точке постоянна. Если, например, на обычные помехи изредка накладываются случайные выбросы, то сама дисперсия переменна: σ_i^2 случайная величина, и эксперимент может находить лишь ее «среднее значение».

Если она распределена нормально (наихудший случай — минимум информации о σ_i^2), то тогда измеряемое y_i имеет распределение Лапласа

$$p(y) = \frac{\sqrt{\pi}}{4\sigma} e^{-\frac{\sqrt{\pi}}{2\sigma}|y - \bar{y}|}$$

В таком случае максимуму правдоподобия соответствует наименьшая сумма не квадратов, а модулей отклонений [3]:

$$M = \sum_i |y_i - Y(\alpha_j, x_i)|.$$

Метод наименьших модулей (МНМ) — наиболее осторожный при риске выбросов. Если нужды в нем не было (дисперсии в действительности постоянные), то МНМ завysит дисперсию s_j^2 параметров α_j в $\pi/2$ раз против МНК. Неправильное применение МНК к задаче с «плавающими» дисперсиями завysит s_j^2 вдвое [3]. Истина в любом случае — между этими крайними вариантами (МНК и МНМ).

Особые приемы регрессии нужны для «цензурированных данных» [7], когда для части точек известно не значение y_i , а лишь неравенство $y_i > t_i$ (например, если y — долговечность и есть образцы, выдержавшие время t_i без разрушения).

В наиболее общем случае количественный анализ эксперимента сводится к следующему ряду действий:

а) оценка рассеяния первичных данных, отбор значимых изменений; оценка сложности разрешимых гипотез по объему информации;

б) формулировка гипотез $Y(x)$ — альтернативных физических моделей и выбор (по максимуму правдоподобия) модели, не противоречащей совокупности измерений;

в) определение параметров модели α_j , их дисперсий и анализ их зависимости от прочих условий опыта, т. е. вновь операции (а) и (б) над результатами анализа и так далее, пока не иссякнет информация.

Статистические критерии асимметричны: не все, что принято, имеет физический смысл, но то, что отвергнуто, не заслуживает обсуждения. Статистика — фильтр, и только после него стоит обсуждать существо наблюдаемых явлений.

Рекомендательный библиографический список

1. Девис Дж. Статистика и анализ геологических данных: Пер. с англ. М.: Мир, 1977. 576 с.
2. Клепиков Н. П., Соколов С. Н. Анализ и планирование экспериментов методом максимума правдоподобия. М.: Наука, 1964. 188 с.
3. Мудров В. И., Кушко В. Л. Методы обработки измерений (квазиправдоподобные оценки). М.: Советское радио, 1976. 192 с.
4. Вапник В. Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. 448 с.
5. Харкевич А. А. Борьба с помехами. М.: Наука, 1965. 276 с.
6. Новицкий П. В. Основы информационной теории измерительных устройств. Л.: Энергия, 1968. 248 с.
7. Себер Дж. Линейный регрессионный анализ: Пер. с англ. М.: Мир, 1980. 456 с.
8. Сборник научных программ на фортране. Руководство для программиста. Вып. 1. Статистика. М.: Статистика, 1974. 316 с.
9. Жовинский А. Н., Жовинский В. Н. Инженерный экспресс-анализ случайных процессов. М.: Энергия, 1979. 112 с.
10. Бард И. Нелинейное оценивание параметров: Пер. с англ. М.: Статистика, 1979. 352 с.
11. Ван дер Варден Б. Л. Математическая статистика: Пер. с англ. М.: Физматгиз, 1963. 436 с.

ОРГАНИЗАЦИЯ И ОПИСАНИЕ ВЫЧИСЛЕНИЙ

В любой задаче физики и техники математические методы всегда эффективны, если только они правильно выбраны. Далеко не все решается на бумаге, но теоретический анализ полезен уже и тем, что указывает, чего не надо мерить.

Не навязывая советов теоретикам, рассмотрим элементарную тактику действий. Стадии решения задачи: формулировка, аналитическое исследование, счет, интерпретация. Вопреки распространенному мнению, если инженер не использует математику в своих повседневных задачах, то чаще всего не из-за недостатка математических знаний, а потому, что не умеет поставить задачу.

Постановка задачи — это перечисление исходных условий и формулировка цели: какую найденную зависимость (между чем и чем) можно будет назвать полным решением? Какая главная величина («выход») нас интересует? Одна ли она? Каков полный набор аргументов («вход»), влияющих на искомый выход? Все ли они измеримы? Какие из них нельзя измерить, но

можно поддерживать неизменными? Какие совсем не управляемы? Не означает ли это, что при наших возможностях не стоит и искать зависимость? Можно ли описывать влияние входных переменных поодиночке?

Какие связи уже известны? Сколько осталось неизвестных связей? Нельзя ли сократить число переменных и связей по соображениям размерности и подобия? Целесообразно ли при оставшихся m переменных искать всеохватывающую физическую модель? Если модель все равно не удастся проверить систематическими измерениями во всем m -мерном пространстве переменных, может быть лучше искать «по возможности» некоторую простую аппроксимацию для наблюдаемой области, чтобы решить практически важную частную задачу (например, найти экстремум)?

Только после такого качественного анализа стоит писать какие-то формулы. Начинаящий закапывается в вычисления, не уяснив, на какой вопрос ищется ответ. По прошествии изрядного времени оказывается, что найден ответ не от той задачи или задача неразрешима, потому что плохо сформулирована.

Рациональная последовательность действий: в письменном виде формулируем цель вычислений; перечисляем все переменные и ожидаемые области их существования; начинаем вести полный список обозначений, в том числе и «общепринятых» (много бедствий бывает от того, что в разных местах длинных вычислений одна и та же буква означает немножко разные величины); выписываем все наперед принятые («очевидные») допущения; рисуем блок-схему решения: какие зависимости хотим получить и где их затем использовать. Только после этого начинается собственно решение.

На листах со сквозной нумерацией пишем подробно все преобразования, выделяем основные формулы, делаем ссылки на источники формул, констант и (что очень важно для чтения и размышлений впоследствии) записываем словами частные цели и результаты и комментируем действия.

Одинаковые индексы, точки, штрихи должны всюду означать одно и то же, и в них есть своя мнемоническая система: когда четвертый момент некоторой величины обозначен M_4 , центрированный m_4 , а нормированный μ_4 , делать преобразования, проверять и понимать их гораздо легче, чем если они обозначены M_4 , α''' и $x^{(4)}$.

Компактные обозначения иногда и сами по себе дают ответ на запутанный вопрос. Запись $F \times g = F \times g$ вместо

$$\begin{aligned} & \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(x, y, z) g(x_0 - x, y_0 - y, z_0 - z) dx dy dz = \\ & = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y, z) F(x_0 - x, y_0 - y, z_0 - z) dx dy dz \end{aligned}$$

делает формулу самоочевидной. Надо вводить одну букву для повторяющегося сложного выражения (r_{12} вместо $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$), но сразу записать эту замену на видном месте, не надеясь на память.

Проверки — неотъемлемая часть процесса решения. Лучше вовремя остановиться, чем переделывать все, поэтому нужны систематические проверки по ходу решения, а не в конце. Непрерывно контролируется размерность результата, однородность имеющих физический смысл многочленов. Проверка должна быть намного быстрее решения, поэтому по ходу вычислений надо примечать выгодные для контроля предельные переходы, численные подстановки, все случаи, где ответ ясен и из других соображений. В окончательном результате проверяем все возможные предельные переходы, порядок величины результата при крайних комбинациях параметров. Как бы ни были красивы и наглядны окончательные формулы, их надо проверить, построив численные зависимости для нескольких наборов параметров, покрывающих всю область существования решения (при этом можно убедиться, что потерян множитель 10^{12} , а после всех подстановок получено $\sqrt{-2}$ кДж/моль или $e^{5\text{кг}}$).

Для численного исследования (собственно счета) тоже нужно сначала составить список обозначений (с единицами измерения), таблицу констант, сводную таблицу рабочих формул, блок-схему их использования, заготовить таблицы результатов (в том числе промежуточных) и установить правила контроля. Даже для функции $y = ax^2 + bx + c$ таблица y при 10 значениях x и всего пяти возможных a , b и c будет с четырьмя входами и 1250 значениями результата. Поэтому стоит заранее продумать, что быстрее, писать ли на одном ли-

сте в столбцах и колонках $y(a, x)$ при фиксированных b, c или $y(a, b)$ при $c, x = \text{const}$? Не надо ли заготовить вспомогательные листы $ax^2(a, x)$; $bx(b, x)$, чтобы потом лишь складывать цифры из разных колонок?

Легче заранее проверить 5—10 контрольных сумм по строкам $\sum_i a_i x_i^2 = a_i \sum_i x_i^2$ и столбцам $\sum_i a_i x_i^2 = x_i^2 \sum_i a_i$,

чем в конце искать ошибку среди тысячи результатов. Чтобы выявить арифметические ошибки, нужно нанести окончательные результаты на графики.

Подготовка к счету на ЭВМ в принципе не отличается от «ручного» счета, но требует еще большей педантичности (чем быстроходней машина, тем тяжелее разобратся в последствиях собственного недомыслия). Даже имея отлаженную программу, нужно записать словами обозначения всех величин и массивов, выписать все рабочие формулы, проверить объем памяти для массивов, обязательно сосчитать контрольные суммы для ввода, выписать порядок вывода на печать, чтобы легче ориентироваться в результатах.

Не начинайте обрабатывать большой массив данных, не окинув взглядом его на графике: одна запятая не на месте убьет впустую все машинное время. Прикиньте на глаз и порядок величины контрольных сумм массива. Ошибки ввода и сбой «на один час счета» у машин не реже, чем у человека. Поэтому выведенные с ЭВМ данные тоже проверяют «по здравому смыслу» (без пересчета) — по знаку, порядку величины, гладкости графиков. В подозрительных случаях проверяют входной массив — его распечаткой, а программу — контрольным примером, а затем уж и повторяют счет.

Описание результатов вычислений в окончательном тексте отчета начинается с изложения цели, постановки задачи, перечня допущений, идеи решения, блок-схемы вычислений. Только после этого идут исходные формулы и преобразования. Все без исключения буквы, значки, индексы описаны словами там, где они впервые вводятся (или в отдельной таблице перед текстом). Численные значения всех констант даны со ссылкой на источник (лучше — в одном месте текста).

Не надо приводить в отчете все преобразования, но нужно указать все переходы, подстановки и привести основные промежуточные формулы, не оставляя за кад-

ром и «маленькие хитрости» вычислений. Вычисления надо показать так, чтобы их можно было проверить не только в принципе, но и в точности тем способом, как у автора (что хорошо и для собственного спокойствия).

Надо приводить в отчете и некоторые промежуточные таблицы, важные для проверки или сами по себе (например, вычисленные значения интегралов, которые могут понадобиться для аналогичных расчетов в будущем).

Исследуя окончательные формулы и зависимости, показывают их предельные переходы, область существования (границы применимости решения), очерчивают диапазон изменения переменных, перечисляют возможные варианты поведения решения в зависимости от констант. Графики приводят не для отдельных сочетаний констант, но для некоторой систематической их совокупности.

Формулы — исходные, конечные и все, о каких приходится говорить, имеют по ГОСТу сквозную нумерацию в пределах главы: например, 3.11 — формула 11 из главы 3. ГОСТ 7.32—81 требует, чтобы сразу под формулой были описаны все буквенные обозначения, каждая буква с новой строчки, в том порядке, как в формуле (это относится, естественно, к исходным формулам, где вводятся обозначения, а не ко всем промежуточным преобразованиям).

Читателю гораздо удобнее, когда величины в тексте называют и словами, а не только буквами:

вместо С изменением $V(r)$ будет $\gamma_{\min}$ при $AB=1$	стоит написать При изменениях потенциала взаимодействия $V(r)$ плотность γ минимальна, когда силовые константы A и B удовлетворяют соотношению $AB=1$.
---	---

При машинном счете нужно привести описание алгоритма, рабочие формулы, блок-схему программы, перечень использованных стандартных подпрограмм, саму программу (распечатку со списком всех обозначений), отладочную задачу и ее решение. Нужно указать машину, язык, транслятор, что и в каком порядке вводится, время счета одного варианта (или диапазон затрат в зависимости от объема массива и значений параметров), гарантированную наибольшую погрешность счета (например, при интегрировании). Если эта по-

грешность не указана в стандартной библиотечной программе, придется сделать ее оценку.

Для вновь разработанных программ указывают также объем памяти, возможности дополнительного вывода промежуточной информации или частичной блокировки вывода, проверку некоторых контрольных соотношений, заложенную в программу (например, $\sum_k a_k^2 =$

$= \sum_i u_i^2$ для Фурье-анализа), вывод контрольных точек

(например, невязок после подстановки решения в исходную систему уравнений). Надо указать, какова реакция программы на ошибки ввода, недостаточность данных, предусмотрена ли печать признаков ошибки, переполнения.

Описание новой программы заканчивают перечнем ее преимуществ и недостатков по сравнению с известными и указанием условий, где ее полезно применять, а также области, где она работает плохо.

Для публикации полагается указывать также, где и в каком виде можно получить программу.

5. ОТЧЕТ ОБ ИССЛЕДОВАНИИ МАТЕРИАЛ И МЕТОДИКА

Высоко поднял, да снизу не подпер.

Пословица

Квалифицированное описание методики — самая трудная часть отчета. Оно должно быть настолько подробным, чтобы любой другой исследователь, поставив эксперимент по одному этому описанию (не додумывая подробностей), получил бы те же (в пределах указанных ошибок) результаты. Если у других специалистов этот результат не воспроизводится, виноват автор публикации: его данные либо ненадежны, либо бестолково изложены. В систематическом описании подробностей с упоминанием всех отклонений и промахов — главная трудность.

Приводятся все сведения о металлургической истории и составе сплава, группировке образцов по плавкам, слиткам, пруткам. Даны чертежи образцов (кроме тех,

что есть в ГОСТах), указан способ их изготовления, диапазон рассеяния ответственных рабочих размеров между образцами и в пределах образца (разнотолщинность по длине и ширине и т. п.).

После этого описана обработка давлением и термическая обработка, входившие в программу работы, порядок использования образцов и затем — все методы измерений. Для каждого из них перечислено оборудование и режимы его работы (по всем параметрам), число и порядок измерений, способ обработки результатов, характеристики рассеяния и оценки ошибки воспроизводимости.

Для стандартных установок указывают использованный диапазон измерений и все паспортные данные о чувствительности, погрешности, стабильности в этом диапазоне, описывают приемы и периодичность юстировки, поверки и контроля работы. Для установки из стандартных блоков описывают блок-схему, принцип действия, характеристики и режим работы блоков, последовательность настройки, работы, контроля. Для нестандартных электронных блоков приводят еще и принципиальную схему со всеми номиналами и краткие указания по настройке. Конструкцию нестандартных механических устройств показывают в одной-двух проекциях или в аксонометрии.

Должны быть описаны и все «не уставные», не очевидные постороннему меры техники безопасности и безопасности техники.

Методику удобно описывать в два приема: сначала заготавливаем «шаблон» с пропусками для цифр и некоторых фактических сведений, беспокоясь только о том, чтобы описание было полным, а анализ достоверности — всесторонним. На этой стадии не надо оглядываться, есть ли у нас все данные. Когда такой бланк составлен «как должно», заполняем его «как было»: внимательно выбираем сведения из рабочего журнала, перечисляя и все отклонения от идеальной схемы. При этом первый вариант полного текста о методике стоит написать после первых серий опытов: пытаюсь доказать посторонним надежность своих измерений, замечаешь, каких проверок не хватает.

Для каждого метода должно быть четко показано: в чем конкретно гарантии надежности результата? Предвидя, с какой стороны результат можно оспаривать, нуж-

но (в компактном неявном виде) наперед заложить в описании методики ответы на основные возражения и придирки (какие делали бы мы сами, будь это чужая работа со странным результатом).

Если эффект оказался статистически незначимым, надо указать, при какой величине он был бы достоверно обнаружен этими измерениями.

При использовании в схеме эксперимента быстродействующих счетных устройств (например, многоканального анализатора при автоматической количественной металлографии или мессбауэровской спектроскопии) априорно оцениваемые (пуассоновы) ошибки счета обычно малы. Основную ошибку создает неконтролируемый сдвиг условий регистрации (порогов дискриминации, скорости развертки и т. п.). Поэтому нужно приводить оценки нестабильности настройки установки из независимого эксперимента (скажем, по воспроизводимости «эталонов» во все дни измерений).

Сравнивая в предварительном опыте всего два уникальных образца, мы тоже должны привести оценки надежности — по воспроизводимости повторных измерений на других образцах (или холостых измерений без образца).

Надо четко показать, какие именно ошибки компенсированы применением эталона, а какие остались. Например, почему изменения эталона во времени можно аппроксимировать гладкой кривой — откуда видно, что период колебаний режимов работы установки существенно больше интервала между отсчетами (если колебания режима быстрые, нормировка на эталон ничего не даст).

Описывая методику, нужно внимательно перечитать рабочий журнал: что в нем осталось ценного? Неудачи не принято расписывать в подробностях и публиковать, но все размышления о промахах и ошибках, о разном в методике и неоднородности материала нужно вставить — в поучение последователям — короткими замечаниями по ходу изложения. Только так накапливается и передается коллективный опыт экспериментаторов.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Обсуждение начинается с краткой формулировки цели и программы основных экспериментов, логики их связей, схемы использования образцов (все это в 2—3 аб-

зацах). Попутно для удобства читателя еще раз перечисляют методы и объем измерений.

Целью работы было сопоставление структурных факторов, определяющих упрочнение сплава Ni_2Cr и поиск режимов обработки, обеспечивающих наивысший предел упругости. Изучена кинетика превращений при отпуске закаленного сплава от 5 мин до 1000 ч при 450—550 °С. Для оценки степени дальнего порядка измеряли удельное электросопротивление при +20 и —196 °С, а также интенсивность рентгеновских линий $\left(\frac{2}{3} \frac{2}{3} 0\right)$ и (220), а для оценки размера доменов

ширину этих линий. Измерено сопротивление малым пластическим деформациям (0,002—0,03 %) при продольном изгибе и микротвердость H_{200} . Дислокационную и доменную структуру наблюдали в двух состояниях сплава (после закалки и при максимальном упрочнении) до деформации и после растяжения на 2 %.

Состав сплава приведен в табл. ..., режимы обработки в табл. ... и... Электросопротивление, предел упругости и микротвердость измерены на одних и тех же ленточных образцах размером $0,3 \times 5 \times 100$ мм (по трем образцам на точку); после измерений от них же отрезаны образцы для дифрактометрической съемки. Образцы для электронной микроскопии толщиной 0,1 мм получены прокаткой из того же листа и закалены вместе с остальными образцами (из той же ампулы), но отпускались отдельно.

Здесь же изложены мотивы выбора методики и условий эксперимента, если они не очевидны, и ожидаемый, эффект:

Полный порядок при 600 °С устанавливается за 600—800 ч [А, Б], так что при выдержках до 1000 ч максимум упрочнения должен быть обнаружен.

По сужению сверхструктурной линии можно зафиксировать начало коалесценции доменов и проверить, является ли размер доменов ведущим фактором упрочнения в этих сплавах.

Даже если рабочая гипотеза не подтвердилась, или не все задуманное оказалось технически выполнимым, надо упомянуть о первоначальной идее и причинах, по которым она не реализована, — без этого трудно понять логику действий (упоминания о причинах неудач — иногда самое ценное в работе).

После вступления идет систематическое описание первичных результатов с обсуждением значимости различий. По каждому графику указывают, что изображено, какие различия значимы, каков вид зависимости:

На рис. показана кинетика изменения периода решетки a сплава В при отжигах в интервале 660—750 °, 5 мин—24 ч. Приведены средние значения a по трем образцам на точку и среднеквадратичные отклонения среднего a , вычисленные для каждой точки; исходное значение a (после закалки от 1240 °С) нанесено слева от оси ординат. При 600 и 650 °С нет значимых изменений a по сравне-

нию с исходным состоянием; через 24 ч при 700 °С период решетки меньше исходного (с достоверностью 0,90), между остальными точками различие не значимо. Только при 750 °С удается проследить кинетику уменьшения периода решетки в интервале 1—24 ч. Скорость обеднения твердого раствора во всем этом интервале нарастает; замедления распада вплоть до 24 ч еще не видно.

Судя по концентрационной зависимости периода решетки (рис...), наибольшее $\Delta a = (0,10 \pm 0,02)$ пм (за 24 ч при 750 °С) соответствует удалению из твердого раствора $(1,5 \pm 0,3)$ % Nb тогда как по диаграмме состояния (рис. ...) пересыщение после закалки может достигать 5 % (ат.) Nb. Очевидно, наши наблюдения охватывают только первую треть всего процесса распада.

В этой части обсуждения обычны две крайности:

а) автору скучно перебирать факты, поэтому он сразу делает вывод, не доискиваясь «противоречий в мелочах»;

б) автор объясняет каждый изгиб графика, не утруждая себя анализом их достоверности и доказательством объяснений.

До выводов тут еще далеко. После перечисления группы родственных фактов идет анализ их взаимосвязей:

Остаточное электросопротивление и период решетки меняются с температурой отжига качественно одинаково (рис. ...). Из-за большей чувствительности метода обеднение твердого раствора заметно по электросопротивлению уже через 4 ч при 660 °С. Известная концентрационная зависимость остаточного сопротивления [Л] позволяет независимо оценить степень распада: наибольшее падение $\Delta \rho = (2,0 \pm 0,4)$ мкОм·см (через 24 ч при 750 °С) по рис. ... соответствует $\Delta c = (1,0 \pm 0,2)$ % (ат.). Это не противоречит найденному по периоду решетки $\Delta c = (1,5 \pm 0,3)$ % (ат.).

Если измерения разными методами согласуются качественно — по виду зависимостей и порядку величин, исследуем полностью количественное соответствие (здесь может быть и новая информация):

«На рис. ... изменения периода решетки a показаны в зависимости от изменений электросопротивления при —196 °С $\Delta \rho$ (те и другие по отношению к исходному состоянию после закалки; измерения a и ρ на одних и тех же образцах, каждая точка соответствует одному образцу). Как видно, для температур отжига 700—750 °С есть общая линейная зависимость $\Delta a = k \Delta \rho$. При низких же температурах (600 и 650 °С) есть значимый подъем электросопротивления ($\Delta \rho = 1,0$ мкОм·см) без соответствующего изменения периода решетки, хотя по рис. ... при таком $\Delta \rho$ следовало ожидать $\Delta a = 0,05$ пм, что при нашей методике должно надежно обнаружиться. Очевидно, значительный подъем электросопротивления (до 1 ч при 600 °С и 15 мин при 650 °С) не связан с изменением концентрации твердого раствора. Аналогичное явление в системах Al—Cu, Al—Zn объясняют аномалиями рассеяния на зонах Гинье—Престона [Л].»

Примерно в той же последовательности описывают и фотографии (структуры). После «общего вида» (типа структуры) отмечают форму и размеры (от и до) основных элементов, особенности их взаимного расположения, ориентировочно — количество. Сначала описывают то общее, что есть в серии структур, а затем — их различия.

Следует быть очень осторожным в описании качественных различий структур, четко указывая, где гарантия, что это не просто различия между выбранными полями:

В исходном состоянии на 30 полях зрения (на 3 образцах) площадью по 0,1 мм² каждое не найдено ни одного участка цементитной сетки; после отжига отдельные участки сетки (в 2—3 зерна длиной) замечены на 7 из 30 полей на трех образцах.

Для более сильных утверждений нужна количественная металлография с достаточной статистикой.

Доказать, что чего-то нет, гораздо труднее, чем что оно есть: наличие видно по единственному снимку или рентгенограмме, а об отсутствии можно говорить только после анализа возможностей метода (границ обнаружимости при данном объеме измерений).

Все внутренние противоречия в данных сразу анализируют. Нужно обосновать и принять решение, является ли неожиданное расхождение между сериями экспериментов следствием различий в их условиях и методике? Если серии экспериментов по методической причине оказались принципиально несравнимыми, то следует исключить из дальнейшего анализа одну из них как неправильно выполненную или она имеет самостоятельную ценность? Можно ли считать противоречие самостоятельным экспериментальным фактом? Не осталась ли причина противоречия невыясненной? Этот, наихудший случай ставит под сомнение не только данную пару, но и все остальные, «благополучные» зависимости, полученные по той же методике. Тогда надо четко указать масштабы потерь:

Если, в худшем случае, была неконтролируемая утечка водорода из образцов до испытания, то все зависимости от концентрации водорода сомнительны, но зависимости от температуры испытания (от 0° и ниже — рис. ...) остаются достоверными, потому что каждая из них получена из серий образцов с одинаковой исходной концентрацией и временем хранения.

Показав, что искомые зависимости существуют (рассматриваемые изменения статистически значимы), каче-

ственно однородны (для одного процесса — одного вида), изменяются с условиями эксперимента закономерно и внутренне не противоречивы (по всем экспериментальным данным, полученным разными методами), можно переходить к сопоставлению фактов с литературой. Здесь обычно меньше возможностей для сравнения (различаются составы, режимы, методика), но вид закономерностей, диапазоны изменений, порядки величин можно сопоставлять систематически. Если даже в литературе нет достаточного набора количественных данных, надо поискать некоторые сильные неравенства для оценки:

Изменение концентрации углерода от 10⁻² до 10⁻⁴ % (по массе) не изменило температурную зависимость предела текучести. Однако это не отвергает гипотезу о твердорастворном упрочнении [Л], поскольку граница растворимости проходит около 10⁻⁶ % (по массе) [Л], и все наши сплавы — двухфазные, с примерно равной концентрацией углерода в твердом растворе.

После всех сопоставлений должно быть четко сказано, какие зависимости в общем качественно соответствуют известным? Какие факты противоречат известным в литературе? Какие из достоверных данных — новые (ранее неизвестные).

Итак, вырисовывается общая схема изложения результатов:

1. Описание отдельной зависимости, выявление в ней значимых эффектов.
2. Сопоставление однотипных зависимостей при различии некоторого параметра; анализ качественных изменений вида зависимости и величины эффекта при изменении этого параметра.
3. Сопоставление взаимосвязей зависимостей, полученных разными методами, анализ внутренней непротиворечивости.
4. Сравнение с литературой — выявление качественных и количественных соответствий, противоречий и отделение действительно нового материала.
5. Собственно обсуждение — сопоставление с теорией, выдвижение гипотез о природе явлений и причинах вновь обнаруженных зависимостей; анализ альтернатив и отбор гипотез.

Процесс обсуждения есть последовательная многоступенчатая свертка информации: закончив анализ первичных фактов, получаем более компактную группу

обобщений «первого уровня», их анализ дает меньшее число следующих обобщений. С вопросом «Ну и что?» мы переходим от одного уровня к следующему.

Каркас обсуждения задан исходной постановкой задачи. «Научные комментарии» к тому, что известно до нас, должны укладываться в 1—2 предложения. Цель обсуждения — установить, что же дал новый факт для разрешения конечной задачи и чего еще не достает во всей цепочке фактов?

Объяснения могут быть «частные» — для разрешения противоречий в наблюдениях или же «основные» — для понимания новых фактов, чтобы решить главную задачу.

При некотором уровне общего образования не составляет труда придумать правдоподобное объяснение отдельному факту (и не одно). Первая трудность здесь в том, чтобы объяснение не противоречило ни одному из остальных фактов (и у нас, и до нас). Чем больше разнородных фактов, затрагиваемых нашей гипотезой, извлечено из памяти и нанизано рядом, тем ближе к истине выжившая гипотеза.

Однако главная работа — доказательство единственности — еще впереди. Нужно выдвинуть достаточно содержательные альтернативы и выводить из них следствия, пока они не предскажут достоверно различимые исходы. Мы увидели дискриминирующий эксперимент — ставился ли он, возможен ли при доступной нам технике? Очень трудно, но необходимо формулировать серию гипотез так, чтобы каждая пара альтернатив была не только взаимоисключающей, но и полной. Для этого надо уметь множество исходов разбивать «последовательно на два». Аргумент «Из трех гипотез две явно отпали, значит, правильна третья, а иначе это никак нельзя объяснить!» — лишен логики. Работа заключается именно в том, чтобы найти: какая система проверок оставит единственно приемлемую модель явления?

Эффектно отвергнуть гипотезу, найдя для нее абсурдное следствие, противоречащее достоверным фактам. Поэтому надо продумать основные возражения по интерпретации результатов и кратко ответить на них при обсуждении, не дожидаясь, пока возразят другие.

Перечитывая обсуждение, проверим: дан ли в нем ответ на все вопросы, поставленные в начале работы?

Можно ли при нашем материале дополнить ответы? Или пока не поздно, лучше сузить вопросы?

Последняя трудная задача автора — самому дать анализ эффективности работы и перспектив.

Научный эффект: какая задача окончательно решена? Где лежит главный путь к решению проблемы в целом? Какие главные вопросы достаточно решить для достижения цели? Что надо для этого изменить в методике и тактике исследования? Какие вопросы возникли в результате нашей работы? Лежат ли они на магистральном пути к решению проблемы или сбоку? Не окзалось ли у нас ключа к решению других задач?

Технический эффект: в чем и насколько созданные установки, материалы или способы их получения лучше существующих? Что даст их применение для конструкции (или для технологического процесса) в целом? Как должны использоваться результаты в производстве? Чего для этого не хватает в экспериментальных данных, в технологии, оборудовании и в организации производства? Какие изменения следует внести в производство сейчас, какие — в дальнейшем, и при каких условиях?

Экономический эффект: как сопоставить в совокупности изменения расхода материалов, эксплуатационных затрат и капиталовложений, выигрыш от повышения качества металла, долговечности и надежности машины, упрощения технологии и повышения производительности? Какие затраты и время нужны для реализации наших предложений? За какой срок эти затраты окупятся? При каких условиях они целесообразны?

Ни одна работа не дает ответа на все эти вопросы, но уж совсем плохо, если она не отвечает ни на один.

ВЫВОДЫ И ПРЕДЛОЖЕНИЯ

Обычно каждый из выводов говорит о достижении одной из основных подцелей, а последний — о главном результате работы (этот «вывод из выводов» частенько не успевают додумать). Главным итогом может быть получение данных о процессах, явлениях, закономерностях; разработка новых методов исследования, методики измерений, алгоритмов, программ; получение качественных и количественных характеристик объектов и явлений; получение новых материалов, испытание, внедрение в производство. Если результат работы отрицательный,

ГОСТ 7.32—81 требует указать это в выводах. Автор обязан наметить в выводах также пути и цели дальнейшей работы либо мотивировать нецелесообразность ее продолжения.

Число выводов, независимо от объема самой работы, обычно от трех до десяти. Если их много, то, скорее всего, в работе не было главной цели.

В несколько пунктов выводов (по 2—4 строчки в каждом) должны уложиться:

а) вновь созданные методы исследования и их основные преимущества перед известными ранее;

б) достоверные новые результаты — главные факты и обобщения;

в) найденное решение технической задачи и главные технологические рекомендации;

г) степень продвижения к решению более общей проблемы и рекомендуемый путь.

Один вывод может быть описательным, перечисляющим все основные изученные зависимости:

1. Изучены изменения диаграммы растяжения (от 0,002 % до разрушения) и твердости закаленных сплавов никеля с 10, 15 и 20 % (ат.) Мо, вызванные отпуском при 500—760 °С (от 15 мин до 200 ч). При всех температурах сопротивление пластической деформации в процессе отпуска монотонно растет, а пластичность падает.

Конкретный и доказательный вывод демонстрирует и объем измерений, и надежность фактов:

Плохо:
Сопоставлением механических свойств после различных режимов отпуска и большого числа рентгенограмм однозначно доказано, что сопротивление малой пластической деформации определяется в первую очередь размерами доменов порядка.

Лучше:
Предел упругости $\sigma_{0,01}$ для всех режимов отпуска (от 15 мин до 12 ч при 400—500 °С) меняется обратно пропорционально размеру доменов L , определенному по ширине сверхструктурной линии (100) (при 17 образцах коэффициент корреляции 0,91).

Выводы читают все — редактируйте их тщательно.

ПУБЛИКАЦИЯ

Служебная обязанность автора — представить отчет о законченной работе во Всесоюзный научно-технический информационный центр, а также подать все возможные заявки на изобретения по материалам отчета (о чем следует упомянуть в тексте.) Рефераты отчетов

публикуют в сборниках ВНИИЦ (ознакомиться с любым отчетом можно, заказав во ВНИИЦ микрофильм).

Принципиально новый материал, представляющий общий научный интерес, публикуют в журналах. По сравнению с отчетом статья много короче (в разных журналах предел: 6—15 страниц машинописи). В ней нельзя приводить одни и те же данные и в таблицах, и на графиках. Описание методики обычно укладывается в 2—3 абзаца, опускаются промежуточные преобразования, но все требования к полноте описания фактов и доказательности при этом сохраняются. Поэтому писать статью по законченному отчету — это кропотливый литературный труд с многократной проверкой и правкой.

В библиографии к этой статье нельзя ссылаться на работу, которую сам не читал, а видел лишь ссылку на нее в другой статье, в РЖ, в картотеке. Причина простая: нельзя анализировать и критиковать работу, не прочитав ее до конца. Уровень доверия к «собственно-ручно обработанным» источникам и к «пересылкам» совсем разный, и нельзя ставить их на одну доску. Если приходится упоминать непрочитанную работу, то надо назвать в тексте ее автора, а ссылку дать на тот источник, откуда о ней узнал, например:

17. Мотт Н., 1936. Цитируется по: Займан Дж. «Принципы теории твердого тела», М.: Мир, 1966, с. 406.

С 1967 г. требования к авторам научных и технических статей регламентированы международными нормами ЮНЕСКО*. Важнейшие положения следующие:

«Главной целью научной статьи является ясное, недвусмысленное и четкое изложение научных идей и информации...

...3. Оригинальная научная статья... написана таким образом, что квалифицированный научный работник имеет возможность на основе полученной информации воспроизвести эксперимент и получить результаты той же точности, проверить наблюдения, расчеты, теоретические выводы автора и, таким образом, оценить результаты его работы.

...6. Авторы должны четко указывать, что в статье является их собственным вкладом, а что принадлежит другим ученым. Они должны четко определить границы

* Рекомендации к публикации научных статей и рефератов, УФН, 1970, т. 100 (2), с. 349—353.

своей работы: возможные ошибки в приведенных ими расчетах, причины ошибок, степень обоснованности выводов. Им не следует слишком переоценивать точность своей работы, общность выводов и возможности для применения результатов.

...8. Материал необходимо излагать лаконично, простым языком, избегая жаргонных слов... Если возникает необходимость во введении новых терминов, они должны четко определяться.

9. ... Неаккуратная подготовка рукописей отнимает время у редакторов, рецензентов и (если статья напечатана) у читателей, создает впечатление научной необоснованности работы.

... 11. Если требуется отметить вклад, который внесла работа в развитие научных идей, то в статье следует привести ясные библиографические ссылки на предшествующие работы по этому же вопросу».

Там же сформулированы требования к реферату к статье:

«2. ... Публикация реферата преследует следующие три цели:

а) помочь ученым, работающим над теми же вопросами, по которым написан реферат, уточнить, имеет ли им смысл читать всю работу полностью;

б) сообщить читателям, для которых статья не представляет специального интереса, как можно больше сведений с тем, чтобы у них не возникло необходимости полностью читать работу;

в) упростить и ускорить работу реферативных журналов».

Редколлегия и ее рецензенты — серьезный фильтр. Его проходит работа, безусловно грамотная по методике, новая по замыслу и результату. Только такие работы и есть наука. Только на таком уровне и имеет смысл работать с самого начала.

СОДЕРЖАНИЕ

	Стр.
Предисловие	3
1. Постановка задачи исследования	5
Анализ задания	5
Дерево целей	10
План, объем, сроки	14
2. Обзор литературы	18
Задача	18
Поиск источников	23
Содержание конспекта	25
Изучение теории	27
Техника конспектирования	28
Группировка и анализ материала	32
3. Методика эксперимента	32
Цель и принципиальная схема измерений	32
Проект установки	36
«Стандартные» измерения	47
Математическое планирование эксперимента	49
ЭВМ в лаборатории	58
4. Организация и анализ эксперимента	63
Первичные документы	63
Первичная обработка результатов	72
Систематические ошибки	76
Случайные ошибки и надежность первичных данных	80
Графики, гистограммы, снимки	90
Количественный анализ экспериментальных зависимостей	98
Организация и описание вычислений	110
5. Отчет об исследовании	115
Материал и методика	115
Результаты и обсуждение	117
Выводы и предложения	123
Публикация	124