

УСПЕХИ ФИЗИЧЕСКИХ НАУКНОВЫЕ ПРИБОРЫ И МЕТОДЫ ИЗМЕРЕНИЙ

СИСТЕМЫ ДОКУМЕНТАЦИИ МОЛЕКУЛЯРНЫХ СПЕКТРОВ

В. В. Налимов, Б. С. Непорент

Интенсивное развитие молекулярной спектроскопии как плодотворного метода исследования строения молекул и межмолекулярных взаимодействий, а также как средства анализа потребовало создания удобной и доступной системы учета накапливаемых в быстро возрастающем количестве данных, относящихся к спектрам различных веществ. Сложность этой задачи становится очевидной, если учесть, что в настоящее время уже описано около миллиона органических соединений, для десятков тысяч которых уже получены в лабораториях и опубликованы спектры поглощения и, наконец, что число такого рода исследований прогрессивно возрастает. При этом следует иметь в виду, что практически нельзя найти двух вполне одинаковых спектров поглощения, принадлежащих различным, даже близким по строению веществам.

Мы более подробно рассмотрим вопрос о систематизации или, как принято его называть, о документации спектров поглощения в инфракрасной области как более сложный по сравнению с задачами документации спектров комбинационного рассеяния или электронных. Инфракрасные спектры по сравнению с другими физическими характеристиками являются, по-видимому, наиболее полным источником информации о физических и химических свойствах молекул. Даже при нынешнем состоянии молекулярной спектроскопии, не позволяющем еще полностью использовать всю возможную информацию, инфракрасный спектр может дать больше сведений о составе и состоянии веществ, чем, например, показатель преломления, температура плавления, дипольный момент и другие физические характеристики молекул, включая электронные спектры и спектры комбинационного рассеяния.

Особые свойства инфракрасного спектра обусловлены тем, что даже в ограниченной области 2—15 μ он представляет почти полный набор основных колебательных частот молекулы, многие из которых являются характеристическими, т. е. появляются только в случае наличия в структуре молекулы определенных группировок атомов. Частоты определяются при этом массами принимающих участие в колебании групп атомов, силами связей между ними, а также характером их относительного движения — формой колебаний. Взаимные влияния атомных группировок в молекуле проявляются в спектре главным образом в виде изменений интенсивностей, иногда весьма существенных, и приводят лишь к относительно небольшим смещениям характеристических полос. В результате спектральная кривая в целом, содержащая в области основных частот десятки полос различной интенсивности, представляет источник весьма богатой информации о строении и свойствах молекул, которая до сих пор исполь-

зуется не полностью; пока в молекулярной спектроскопии проведена только идентификация частот основных характеристических полос. Вопросы же об интенсивностях полос и о проявлениях взаимных влияний атомов и групп, несмотря на всю свою важность для современной химии, в настоящее время находятся еще в стадии разработки.

На рис. 1 приведен в качестве примера инфракрасный спектр поглощения сульфониламида; там же дана его структурная формула. Сама по себе структурная формула представляет по существу краткую запись, с помощью которой современная химия описывает не только структурные элементы молекул, но, задавая их взаимное расположение, определяет также влияние этих элементов друг на друга и происходящие отсюда тон-

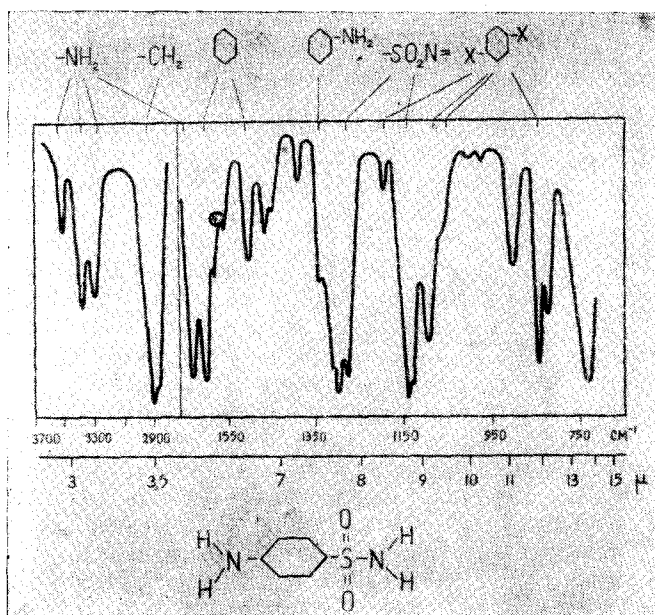


Рис. 1. Инфракрасный спектр поглощения сульфониламида, по Беллами¹. Вверху показано отнесение полос к колебаниям групп атомов, внизу — структурная формула вещества.

кие и вместе с тем весьма существенные детали свойств молекулы, а следовательно, и ее поведение как химического объекта. В соответствии с этим инфракрасный спектр поглощения содержит ряд полос, свидетельствующих о наличии в молекуле определенных структурных групп, а также полос, связанных с определенным взаимным расположением и взаимодействиями одних групп с другими, причем соотношения интенсивностей тех или иных полос отражают характер и силу этих взаимодействий.

Задача создания системы документации спектров состоит в том, чтобы сопоставить каждому из многих тысяч веществ свой спектр и обеспечить возможность выбора из этого множества: а) групп веществ, обладающих набором заданных спектральных признаков (сочетание полос известного положения и интенсивности), или б) групп спектров, отвечающих веществам, характеризующимся набором заданных элементов строения молекул (определенное сочетание атомов и атомных групп). Чем полнее определен набор признаков, по которым производится выбор, тем, естественно, меньше число спектров (или веществ) будет содержать выбираемая группа. Неко-

торому количеству признаков будет соответствовать единственный спектр (или вещество).

Существование такой идеальной системы документации позволило бы решать ряд существенных задач, далеко выходящих за пределы идентификации вещества по спектру или, с другой стороны, за пределы предсказания спектра по структурной формуле вещества. Если материал доброкачествен и если обеспечена возможность отбора желаемых данных (спектров или веществ) по любому количеству существенных для каждого случая признаков, т. е. если система документации имеет достаточное количество входов, то она может позволить, например, выяснить характер влияния тех или иных заместителей на свойства заданной полосы в разных соединениях, сопоставить спектральные свойства соединений с их реакционной способностью по отношению к тем или иным реакциям и т. п. В зависимости от сложности задачи должна быть обеспечена, естественно, большая или меньшая скорость обработки материала. Не говоря уже об аналитических применениях, без преувеличения можно считать, что осуществление надлежащей системы документации является в настоящее время одним из важных условий дальнейшего развития научных исследований, многие из которых, по-видимому, практически невозможны без обработки большого количества данных. Не следует, однако, недооценивать трудности создания системы документации, удовлетворяющей рассмотренным условиям. В сущности, каждая деталь спектральной кривой содержит ту или иную информацию о свойствах молекулы. Создание кода, обеспечивающего возможность дискретного представления спектров без потери информации, которое позволило бы выбрать необходимый спектр из множества других по нескольким полосам или сочетаниям полос или по другим признакам такого рода, является весьма сложной и трудно разрешимой задачей, к которой мы еще вернемся ниже. Аналогично этому, большие сложности представляет также кодирование структурной формулы, которое позволило бы найти соответствующую молекулу по одному или группе признаков, передаваемых этой формулой (количество тех или иных атомов, наличие и взаимное расположение определенных группировок, характер групп и связей между ними и т. д.).

Первым и весьма грубым приближением, позволяющим ориентироваться в спектрах в смысле определения важнейших структурных элементов молекул по наличию в их спектрах тех или иных характеристических полос, являются таблицы типа приведенной на рис. 2, предложенные в разное время разными авторами (см., например¹) для разных групп соединений. Такого рода таблицы, имея только два входа (свойства, представленные по горизонтали и по вертикали), не могут, при всей своей ценности для первоначальной ориентировки, рассматриваться как удовлетворяющие требованиям, предъявляемым к документации инфракрасных спектров, поскольку они не передают ни деталей спектров, ни всего многообразия структурных элементов молекул. Учет внутримолекулярных взаимодействий с помощью дополнительных таблиц для колебаний отдельных связей в разных молекулах (см., например, таблицу рис. 3) привел бы к слишком сложной системе. Подчеркнем особо, что в рассмотренных таблицах практически не учитываются интенсивности рассматриваемых полос.

Следующим шагом является составление атласов спектров, в которых представлены коллекции спектральных кривых для тех или иных групп соединений. Такие атласы, содержащие сотни спектров, были собраны отдельными авторами и выпущены в виде статей и книг (см., например¹).

Быстрое увеличение количества измеренных спектров потребовало введения новых форм их документации. Такой формой являются, напри-

мер, спектры API (American Petroleum Institute), коллекционируемые в виде отдельных карт (рис. 4), содержащих спектр и некоторые сведения об исследованном веществе. Фирмой Садтлер (США) был выпущен специальный «Атлас стандартных спектров», который непрерывно пополняется по мере накопления материала. Атлас издается в двух сериях: по одной таблице на картах размером 17×46 см или по три таблицы на страницах размером $21,5 \times 28$ см. Образец таблицы второго типа приведен на рис. 5. Кроме спектральной кривой, изображенной в виде зависимости пропускания образца (0—100%) от длины волны (2—15 μ), таблица содержит структурную и эмпирическую формулы вещества, его название, молекулярный вес, точку плавления, сведения об образце, использованном для получения спектра (агрегатное состояние, толщина слоя, растворитель и т. д.)

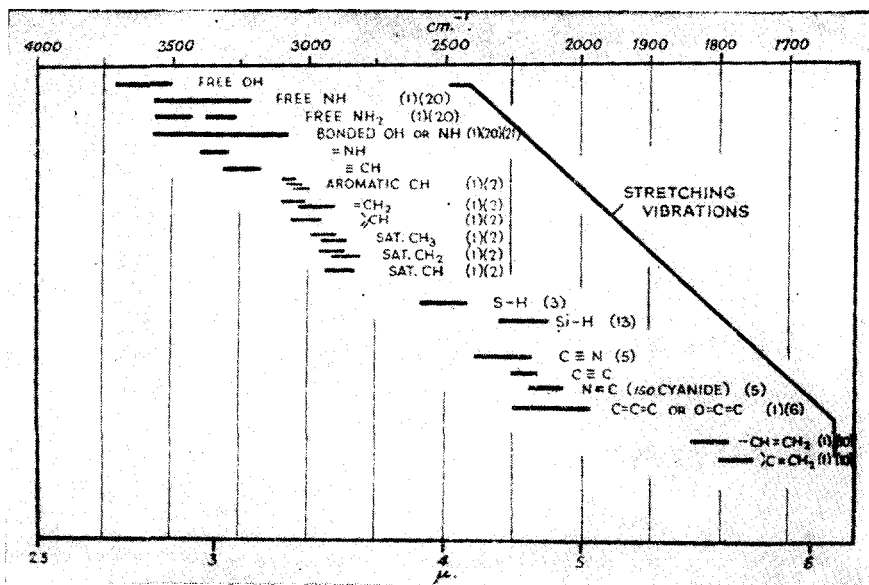


Рис. 2. Положение характеристических частот колебаний некоторых групп атомов, по Томсону¹.

и, наконец, литературную ссылку на источник, откуда заимствован спектр. Условия получения спектров поглощения специально не оговариваются. Обычно для нанесения на карту используются лучшие спектрограммы из полученных в данной работе. Часто используются спектры, полученные для 30% растворов в 0,1-мм кювете, или спектры чистых веществ в 0,03-мм кювете.

Представление фактического материала в виде спектральных кривых имеет значительное преимущество наглядности и возможности передачи многих тонких деталей спектра. Таким образом, каждая таблица атласа представляет, вместе со структурной формулой вещества, предельную достигаемую в условиях стандартного исследования информацию о спектре. Атлас Садтлера содержал уже в 1956 г. свыше 10 000 спектров и для того, чтобы обеспечить возможность и удобство пользования им, т. е. выбора спектров заданных веществ или веществ по заданным спектрам, фирма разработала ряд указателей: I—по порядковым номерам карточек, которые присваиваются им произвольно по мере поступления спектров, II—IV—по химическим характеристикам молекул—алфавитный по всему атласу, алфавитный по классам соединений, по эмпирическим формулам

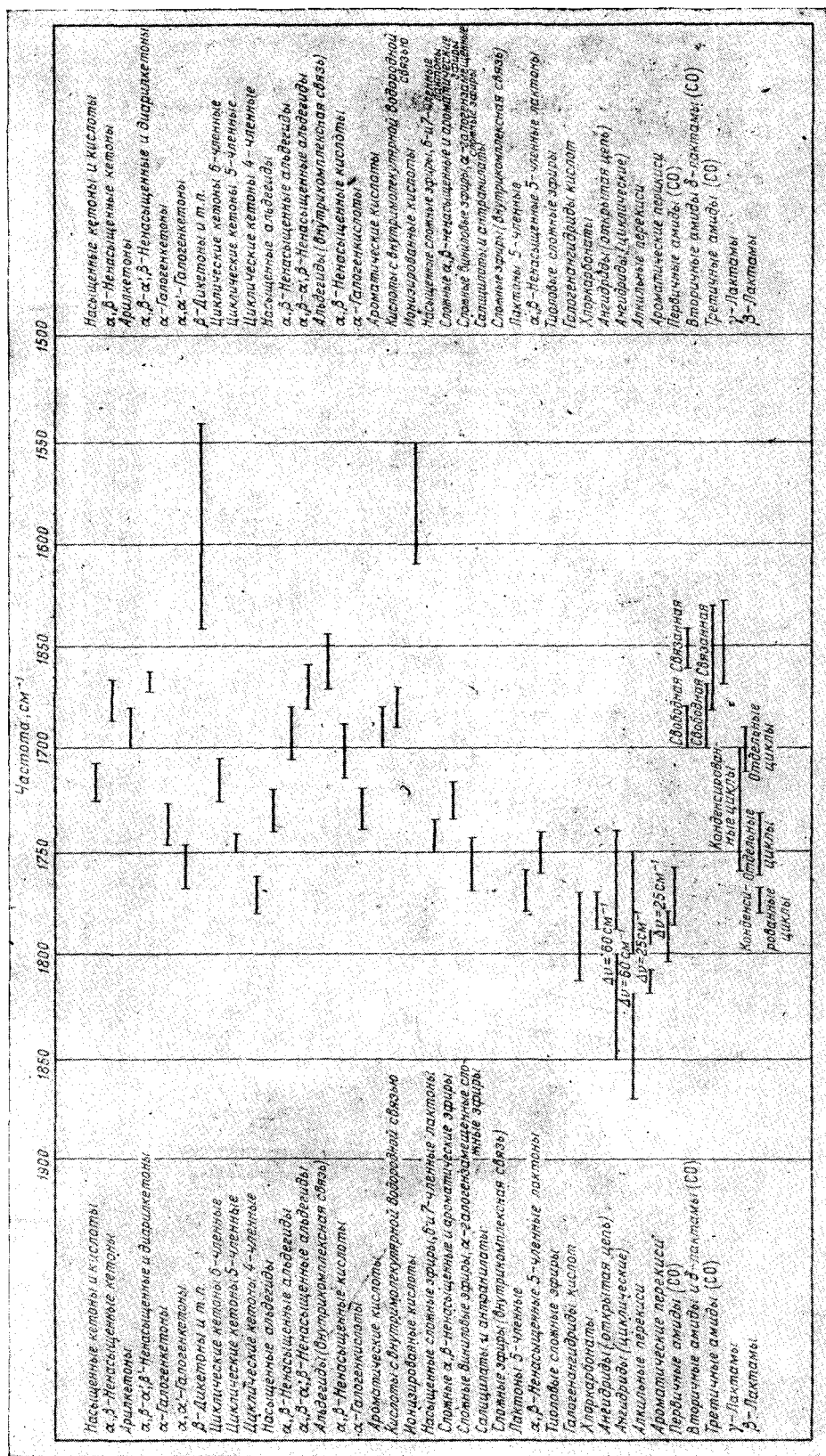


Рис. 3. Частоты колебаний карбональной группы в разных соединениях, по Беллами.

и по группам веществ, V—по спектральным характеристикам молекул—по главным полосам поглощения («спекфайндер»). В указателях II—IV использована система химической классификации, соответствующая принятой в Chemical Abstracts. В спектральном указателе V использован следующий прием: в каждом интервале в 1 микрон в области 2—15 μ в спектре выбирается наиболее интенсивная полоса, положение которой определяется с точностью до 0,1 μ , т. е. соответствующая длина волны выражается числом, имеющим одну цифру после запятой. Совокупность полученных цифр представляет набор соответствующих характеристических чисел. Эти числа (приведенные, например, на рис. 5 внизу) располагаются

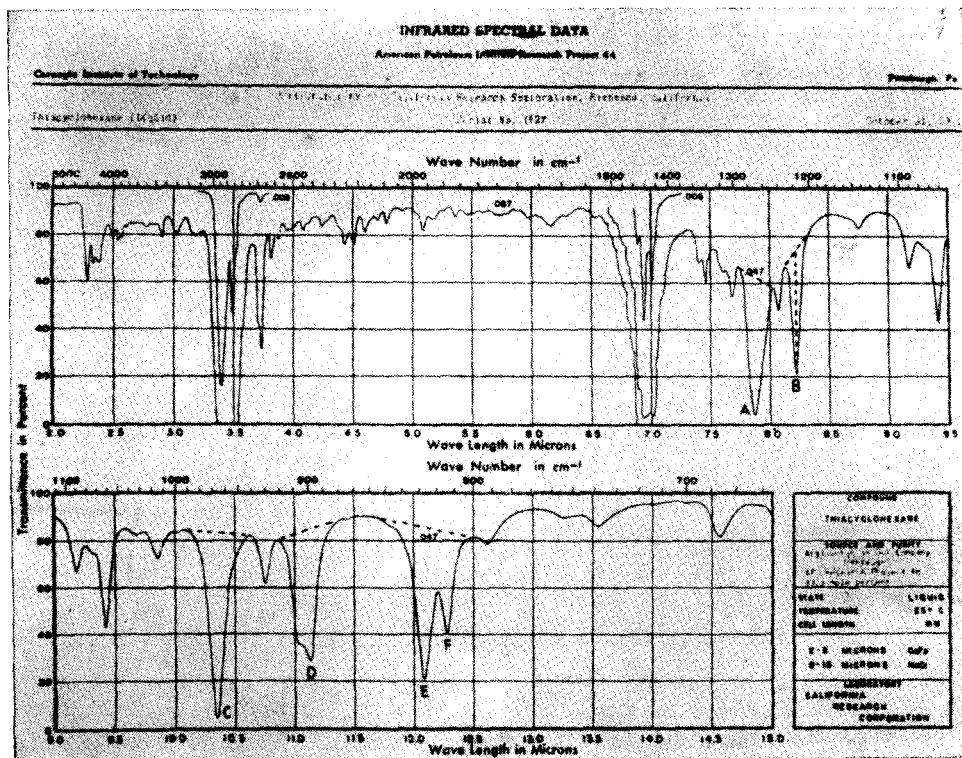


Рис. 4. Карта API с инфракрасным спектром тиациклогексана.

в указателе (см. рис. 6) в строку, в правой части которой указан номер спектра в атласе. Целочисленные значения длин волн максимумов, которые подразумеваются, показаны для удобства в верхней части каждой страницы указателя. Спектры располагаются в указателе в последовательности, отвечающей характеристическому числу (длине волны λ_{\max}) самой интенсивной полосы. Внутри групп с совпадающими значениями λ_{\max} спектры располагаются по возрастающим значениям чисел в последовательных столбцах. На рис. 6 отмечены строки, отвечающие спектру рис. 5. Этих строк две, поскольку в одном из интервалов (9—10 μ) две полосы имеют одинаковую интенсивность.

Атлас Садтлера с системой его указателей позволяет с относительно небольшой затратой времени найти спектр вещества по его структурной или эмпирической формуле, или, наоборот, определить вещество по его спектру, т. е. разрешать задачи, возникающие в повседневной

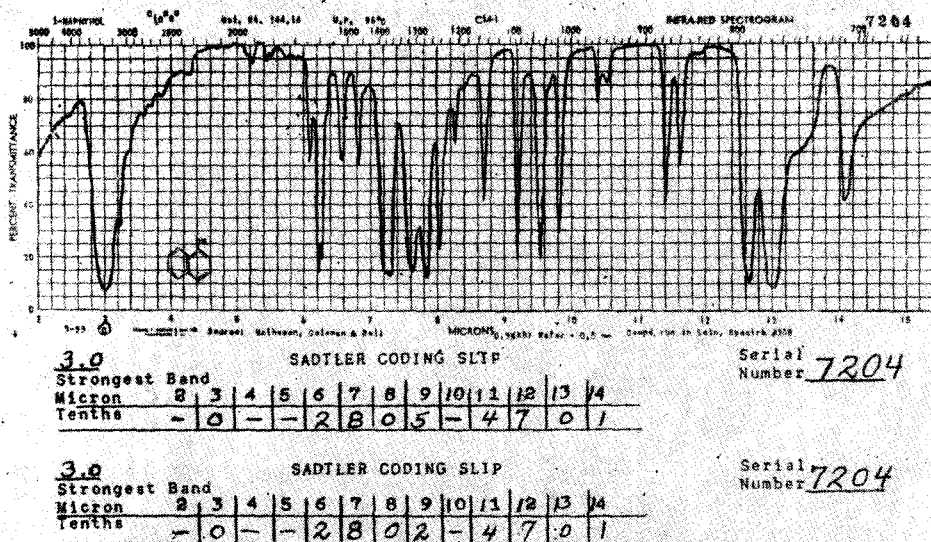


Рис. 5. Один из «стандартных спектров» Садтлера. Размер спектрограммы в меньшей серии — 85×185 мм. Над верхним краем таблицы указаны название вещества, эмпирическая формула, молекулярный вес, точка плавления (кипения), № спектрограммы. Под нижним краем — источник (журнал, лаборатория, авторы), характеристика образца (раствор, жидкость, пленка, таблетка КВг) и его толщина, ссылки на другие спектрограммы. Две таблицы в нижней части рисунка — наборы характеристических чисел для «спекфайндера».

0	2	0	8	2	6	3	0	30	379						
0	2	1	6	2		2	4	30	151						
0	2	2	8		0	6	4	1	5979						
MICRONS															
2	3	4	5	6	7	8	9	10	11	12	13	14	Strongest Band	Serial Number	STRONGEST BAND 3.0
0				2	6	0	7		6		6	3	30	7376	
0				2	7	1	1		3	3	1		30	4731	
0				2	8		4		2				30	8673	
0				2	8	0	2		4	7	0	1	30	7204	
0				2	8	0	5		4	7	0	1	30	7204	
0				2	8	1				8	3		30	4763	
0				2	9	2	6	5	4	2	7		30	8667	
0				2	9	3	6	0		4	5		30	8392	
0				3	0	0	2					5	30	9172	
0				3	0	0	1			1	6	6	30	9232	
0				3	0	3	1			1	6	6	30	9232	
0				3	2	8	4			3			30	6630	
0				3	3	2			9				30	994	
0				3	6	2	3		8	6			30	3671	
0				3	7	3	4	3		4			30	152	
0				3	7	3	9	3		4			30	152	
0				3	9	9	6	0	6	5	8		30	1065	
0				4	7	2	3		6	2	0	1	30	9050	
0				5		4					6	4	30	8574	
0				6	1	7	9	9	3	3	4		30	124	
0				6	4	0	1		8		5		30	268	
0				6	9	0	8		0	4	4		30	8447	
0				6	9	1		5	9	7		0	30	9049	
0				6	9	3	6	0		4	5		30	8392	
0				7		7		8	2	9	2	3	30	2206	

Рис. 6. Образец части страницы «спекфайндера», содержащий спектр рис. 5.

аналитической практике. Разрешение более сложных задач, требующих обработки более широкого материала, становится при этом несколько затруднительным, требуя значительной затраты времени, поскольку рассматриваемая система имеет ограниченное число входов и подбор данных, не предусмотренных этими входами (например, спектров всех соединений, имеющих в своем составе азо-, нитрильную или другую группировку атомов), по существу представляет трудоемкую работу по построению соответствующего входа в атлас.

Необходимое для разрешения сложных задач многомерное представление материала может быть осуществлено с помощью перфокарт². Каждую заполненную перфокарту можно рассматривать как точку многомерного пространства, свернутую на плоскость: каждая перфорация на плоскости карты—это координата такой многомерной точки, причем эти координаты могут принимать только два значения 0 и 1 (дырка пробита или

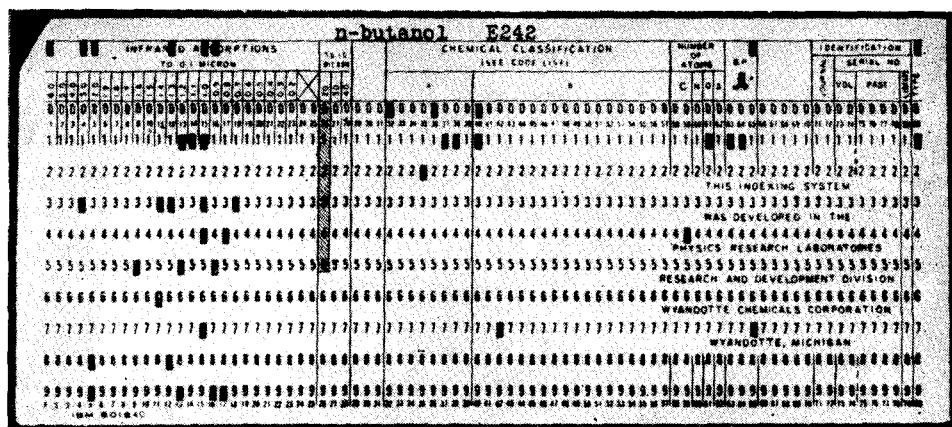


Рис. 7. Карта с внутренней перфорацией для машинной сортировки (IBM, Hollerith), заполненная по коду ASTM—Виандотт. Размер карты 83×188 мм.

не пробита). Стопку перфокарт можно рассматривать как множество точек многомерного пространства или как таблицу в многомерном пространстве с множеством входов.

Применение перфокарт в спектроскопии имеет уже свою историю. В 1948 г. в США для документации в области спектроскопии существовало 18 различных систем перфокарт и в связи с этим был создан комитет для выбора системы спектроскопических перфокарт³. В настоящее время в спектроскопии применяются две основные системы: перфокарты с внутренней перфорацией и перфокарты с перфорацией по краям. Карты заполняются как опубликованным материалам, так и по частным сообщениям отдельных лабораторий.

Перфокарты с внутренней перфорацией (такого же типа, как используются в СССР для машинной обработки большого цифрового материала и механизации счетных операций) применены в США для шифровки спектров, помещенных в атласе Садтлера⁴. Эти карты называют в литературе картами IBM или Hollerith, а также ASTM—Wyandotte (IBM, Hollerith—название фирм, выпускающих машины для обработки карт, ASTM—Американское общество по испытанию материалов, разработавшее совместно с фирмой Виандотт систему кодирования спектральных данных). На рис. 7 приведена фотография такой перфокарты, на которой имеется 80 вертикальных колонок и 10 горизонтальных рядов, переименованных

числами от 0 до 9; кроме того, перфорироваться могут также еще два верхних нумерованных ряда (в дальнейшем они будут обозначаться через x и y).

Спектры поглощения и данные, характеризующие физические и химические свойства молекулы, наносятся на карту (перфорируются) с помощью упомянутого кода. Колонки в левой части карточки служат для зашифровывания спектров поглощения в шкале обратных сантиметров (если нанесена перфорация в ряду y) или в шкале длин волн (ряд y не перфорирован). Если пользоваться шкалой длин волн, то первые 25 колонок служат для записи спектра поглощения от 2 до 25 μ с разрешением в 0,1 μ . Номер колонки обозначает число целых микрон, номер ряда — десятые доли микрона. Например, если в колонке № 12 пробита цифра 5, это означает 12,5 μ . Колонки №№ 26—28 служат для записи спектра в области 25—45 μ с разрешением в 1 μ . Например, если в колонке 27 пробита цифра 3, то это соответствует длине волны 33 μ . Для записи спектров в шкале обратных сантиметров используются те же колонки: в интервале 2000—4500 см^{-1} запись производится с разрешением в 100 см^{-1} , а в интервале 200—2000 см^{-1} с разрешением 10 см^{-1} . На перфокарту наносятся все максимумы в спектре, поглощение в которых на 20% или больше отличается от фона. Легко понять, что шифровка спектра не является исчерпывающей, хотя она и значительно более подробна, чем принятая в указателе Садтлера. Достаточно отметить, что интенсивности полос, хотя бы и относительные, никак не отражаются в карте. Этот недостаток, впрочем, свойствен всем имеющимся системам кодирования данных по инфракрасным спектрам, в которых приходится допускать значительную потерю информации вследствие недостаточной информационной емкости перфокарт.

Колонки 29—31 оставлены резервными, на колонках 32—62 подробно шифруются данные, характеризующие химический состав молекулы и ее строение. В колонках 32—57 зашифрованы сведения о классе соединений, к которому относится молекула, и о наличии в ней тех или иных структурных элементов. Примеры кодов для нескольких колонок приведены в таблице I. Колонки 58—62 несут сведения о содержании в молекуле атомов C, N, O и S, а 63—65 — сведения о температурах кипения и плавления. Колонки 66—70 оставлены для потребностей отдельных лабораторий, а 71—79 содержат сведения о литературных источниках, из которых заимствован спектр (журнал, том, страница и т. д.). Следует заметить, что химический код отличается полнотой и позволяет отбирать спектры или устанавливать корреляционные связи между зашифрованными на карте характеристиками спектров и структурными элементами строения молекул при большом богатстве выбора последних. Спектральный код, как указывалось, не столь полон, и свойственная картам ASTM некоторая неполнота информации о спектрах может существенно уменьшить скорость обработки материала, если для решения задачи окажется необходимым получить сведения о структуре спектров, обращаясь для каждого вещества к атласу Садтлера.

Карты ASTM сортируются машинами со скоростью 25—650 штук в минуту. После многократного прохождения через сортировальную машину эти карты изнашиваются, что не является существенным недостатком, так как они легко могут быть скопированы на чистые бланки с помощью специальной копировальной машинки. По данным 1956 г. было выпущено 12 000 этих перфокарт.

В связи с вопросом о неполноте спектрального кода следует указать, что в⁵ предложен способ нанесения на перфокарты IBM с внутренней перфорацией инфракрасного спектра поглощения с учетом всех его

Таблица I

Код ASTM—Wyandotte для нескольких колонок

Ряд	Колонка 32 элементы	Колонка 34 классы соеди- нений	Колонка 40 группы	Колонка 44 азот	Колонка 48 азот—кислород
0	O	Насыщенные алифатические	Метил	$-C(=NH)N<$	$>NC(=O)O-$
1	N	Ненасыщенные алифатические	Этил	$>NC(=NH)N<$	$>NC(=O)N<$
2	S	Насыщенные моноалициклич.	<i>n</i> -	$-C\equiv N$	$-C(=O)N<$
3	F	Ненасыщенные моноалициклич.	Изо-	$-N\equiv C$	$-N=C=O$
4	Cl	Насыщенные полиалициклич.	<i>n</i> -	$-NH_2$	$-OC\equiv N$
5	Br, J	Ненасыщенные полиалициклич.	Изо-	$>NH$	$>N-NO_2$
6	P, As, Sb, Bi	Бензолы, ароматические	Вторичн.	$\rightarrow N$	$>N-N=O$
7	Si, Ge, Sn, Pb	Ароматические полибензолные	Третичн.	$=NH$	$-N=N(=O)-$
8	Se, Te	Ароматические с конденсир. ядрами	<i>n</i> -амил	$>N-N<$	$-ONO_2$
9	B, Al	Кольца из 3—4 атомов	Фенил	$-N=N-$	$-ONO$
x	Другие	Кольца из 5—6 атомов	—	другие гетероциклические	Другие гетероциклические
y		Остальные	—		

деталей, отвечающих разрешающей способности прибора. При этом в каждой колонке записывается (в двоичной системе) интенсивность спектра в одной точке спектрограммы. Для записи всего спектра используется до 100 карт (!) (8000 точек), что дает возможность отразить тонкую структуру спектра. Перфокарты заполняются автоматически при съемке спектрограмм. Пользуясь таким дискретным представлением спектров, можно с помощью счетных машин выполнять разнообразные операции: производить сравнение спектров; из спектра, снятого для неизвестной смеси, выделить спектры известных соединений; пользуясь методами теории информации, улучшить отношение сигнала к шуму, используя счетную машину как фильтр, и т. п.

Следует упомянуть еще о системе использования карт с внутренней перфорацией для анализа смесей веществ по инфракрасным спектрам с помощью счетных машин⁶. При этом в качестве характеристических признаков вещества используются не только определенные полосы, но также и факт отсутствия полос в определенных участках спектра. Карты

заполняются по специальному, сравнительно простому коду, который является развитием предложенного ранее одним из авторов работы ⁶.

Иным путем разрешен вопрос об увеличении информационной емкости в системах карт с внешней перфорацией. Эти карты, отличаясь более медленной ручной сортировкой и относительной бедностью кода, несут, тем не менее, большую информацию, поскольку, кроме зашифрованных на перфорациях сведений о самой грубой структуре спектра, на картах непосредственно изображена спектральная кривая и приведен ряд других сведений.

Из многочисленных систем карт с внешней перфорацией мы рассмотрим подробнее, как наиболее типичную, систему DMS (Documentation of Molecular Spectroscopy) ^{7,8}, разработанную спектроскопистами Англии (Комитет под председательством Г. Томсона) совместно со спектроскопистами ФРГ (группы, объединенные Г. Кайзером) на базе опыта использования перфокарт других систем. Карты DMS состоят из двух серий: литературной, несущей информацию об источниках получения спектров, и спектральной, карты которой несут информацию о спектрах и веществах.

На рис. 8 и 9 приведены фотографии двух сторон спектральной перфокарты DMS. На лицевой стороне напечатано название соединения, указаны его основные физические свойства, дана структурная формула, приведено библиографическое описание источника, из которого заимствован спектр поглощения, и указаны условия, при которых он получен. На оборотной стороне карточки приведен спектр поглощения в области 200—4000 см^{-1} , представленный стандартным образом с линейной шкалой для обратных сантиметров и изменением масштаба при 2000 см^{-1} . Там же приведены табулированные данные для основных полос поглощения.

По краям карты имеется 203 отверстия. Для зашифрования информации эти отверстия могут быть прорезаны тремя различными способами: 1) вырезается лунка, захватывающая одну внешнюю дырку, 2) лунка захватывает обе дырки—внешнюю и внутреннюю и выходит на наружный край карты, 3) оба отверстия соединяются внутренней лункой, не выходящей на наружный край карты. Такая система дает возможность использовать каждую пару отверстий для зашифрования трех различных сведений. Сортировка карт производится вручную, при помощи спиц: карточки с вырезанными лунками извлекаются из общей стопки.

Нижний и правый края карты используются для кодирования 8—10 главных полос поглощения. Около отверстий отмечена длина волн в μ и частоты в см^{-1} . В области от 250—1800 см^{-1} отметки нанесены через 50 см^{-1} , в области 1800—2300 см^{-1} через 200 см^{-1} и далее до 3700 см^{-1} через 100 см^{-1} . Выбор полос поглощения для кодирования связан в этом случае с известным произволом—среди сильных полос могут встретиться не типичные для данного класса соединений. Поэтому положение полос на перфокарте только намечено—прорези лунок делают потребители карт, которые, исходя из своих потребностей, могут изменить систему нанесения прорезей.

Верхний и левый края карточки используются для кодирования информации, связанной со строением молекулы, которая рассматривается как состоящая из определенного количества структурных единиц, представляющих интерес с точки зрения химии или спектроскопии. На карте кодируется: число атомов углерода, строение основного скелета молекулы (число ненасыщенных углеродных связей, число колец—конденсированных и неконденсированных, наличие посторонних атомов в кольце), типы групп заместителей и их положение. Если структура соединения неизвестна, то кодируется по алфавиту название соединения. Подробное описание кода дается в специальном руководстве, приложенном

к перфокартам. Некоторое представление о системе кодирования можно получить по приведенным фотографиям перфокарты.

На рис. 10 и 11 даны фотографии лицевой и тыльной стороны библиографической перфокарты DMS. На них заносится библиографическая информация как для статей, использованных при составлении спектральных перфокарт, так и для статей общеметодического характера. Библиографические и спектральные перфокарты связаны перекрестными ссылками. На лицевой стороне карты указаны библиографические данные и приведен краткий реферат статьи. По краям карты имеются перфорированные отверстия, при помощи которых кодируются следующие данные: 1) цель исследования (например, анализ чистоты материала), 2) область спектра, в которой проводилась работа, 3) примененная аппаратура, 4) агрегатное состояние вещества, 5) год выхода журнала (указываются две цифры*), 6) три первые буквы фамилии первого автора и первая буква его имени**).

Кроме заполненных перфокарт, выпускаются также карты-бланки со структурным кодом, которые могут быть использованы для кодирования материала, имеющего частный интерес для тех или иных лабораторий. Все издание выпускается (по 2000 карт в год) на двух языках—английском и немецком. Литературные карты—желтого цвета, спектральные для инфракрасных спектров—розовые, спектральные для комбинационного рассеяния—белые. Предполагается также начать выпуск карт по электронным спектрам.

Из других типов карт с внешней перфорацией следует отметить прежде всего систему, разработанную в США уже упоминавшимся Комитетом³, созданным Национальным исследовательским советом (NRC) совместно с национальным Бюро стандартов (NBS) под председательством В. Броде, секретарь Е. Крейтц. Эта система, называемая картами NRC-NBS, явилась обобщением опыта американских спектроскопистов и послужила, по-видимому, также основой при разработке карт DMS. Сведения NRC-NBS о системе карт NRC-NBS изложены в обзорных статьях^{9, 10} и в статьях Крейтца¹¹. Система NRC-NBS состоит из двух серий—карты веществ (желтые) и библиографические (белые). На тыльной стороне карт веществ изображена спектральная кривая поглощения, в которой, в отличие от карт DMS, принят равномерный масштаб длин волн. Обе системы—европейская и американская—существенно различаются также и по способу представления и кодирования всего остального материала, о чем можно судить, сравнив фотографии этих карт на рис. 12—13 и 8—9, а также соответственно на рис. 14—15 и 10—11. Детали способа кодирования на картах NBS видны на фотографиях карт и едва ли есть необходимость рассматривать их здесь подробно, учитывая, что основные принципы кодирования сохранены в системе DMS.

Отметим еще карты с внешней перфорацией компании Шелл (Shell Development Co), в которых принята своеобразная система кодирования материала, разработанная при участии Р. Браттайна, Д. Танклиффа и др. Эти карты имеют сравнительно большой размер: 203×280 мм (размер карт DMS: 146×210 мм; карт NBC: 165×190 мм). Свободное поле в середине оставлено для изображения спектра. Отверстия по краям карты служат для кодирования материала по двум системам: 40 пар отверстий

*) Используется система, при которой числа от 0 до 9 кодируются двумя числами ряда: 7, 4, 2, 1, 0, такими, чтобы сумма их была равна кодируемому числу. Например, 54 обозначается так: в группе десятков пробиваются отверстия с номерами 4 и 1, в группе единиц пробиваются отверстия с номерами 4 и 0.

**) Каждая буква шифруется по двойному коду двумя числами. Это дает возможность при помощи 8 чисел кодировать 28 букв. Код указан на перфокарте.

Ethyl n-propionate				62				STRUCTURAL FORMULA			
(Ethyl n-propanoate)				$C_5H_{10}O_2$				$CH_3CH_2C(=O)OCH_2CH_3$			
E. K. Co.		Liquid		2-16 μ							
B. P. 97-98°											
REFERENCE NUMBER	DATE	SAMPLE DIMENSIONS	PRESS. CONT. SOLICIT	RANGE	DATE	FROM GRATING	TEMP.	REMARKS			
API-1229	L			2-15 μ	1951	NaCl	25°C	Spectrum			
8759-0454	L			500-1750	1945			Assignment chart			
3722-0481	G			1761	1948	CaF ₂		C=O freq.			
2515-0291	L	40mm 1 mm		7-2.5 μ	1929	Quartz		C=O study			
2909-0514	L	0.1, 0.5, 1 mm 0.5mm CCl ₄		1200-1900	1951	NaCl		Effect of C=O on C-H intensities			
8175-1261	L	1.5mm 0.5 1 μ		1-2.5 μ	1926	NaCl	Room	Spectrum			
7051-0511	Sol	14mm CCl ₄		2800-3000	1951	LiF		C-H stretch wt. spectrum			
0834-0291					1929			C-H freq.; absorpt. coeff.			
5530-0521	L	0.1mm		15-25 μ	1952	KBr		Transparent solvent except 17-18 μ			
NATIONAL RESEARCH COUNCIL - NATIONAL BUREAU OF STANDARDS COMPOUND CARD											

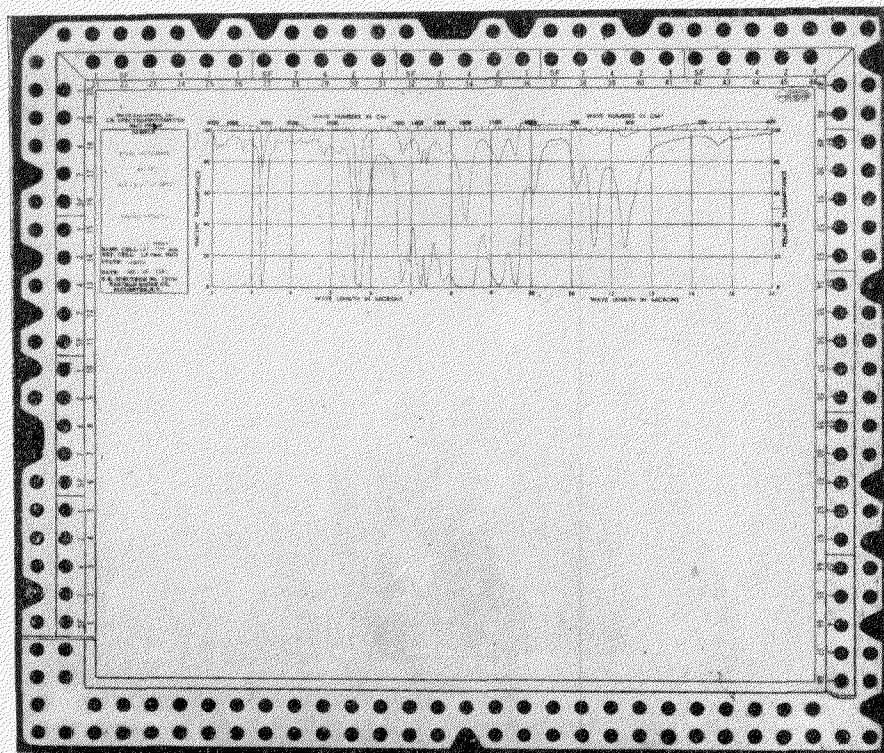


Рис. 12, 13. Спектральная карта NBS. Лицевая и обратная стороны.
Размер карты 165×191 мм.

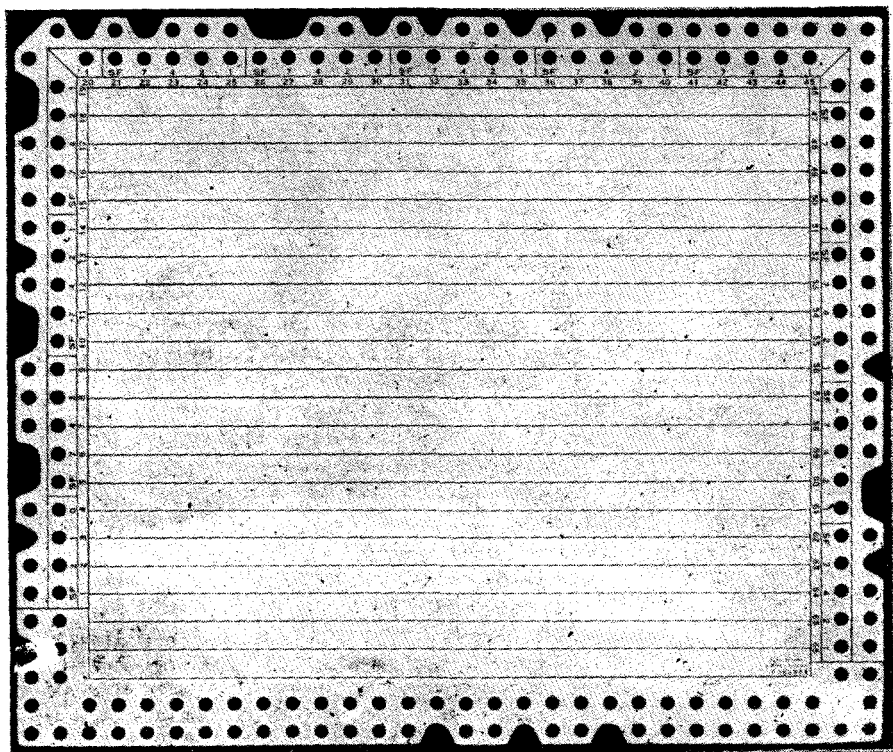
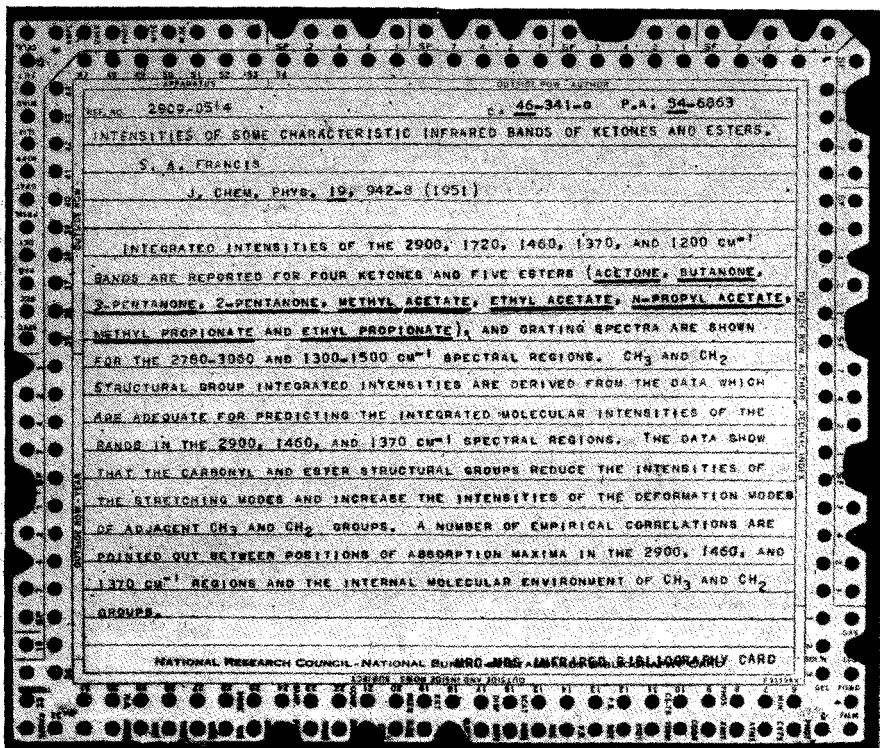


Рис. 14, 15. Литературная карта NBS. Лицевая и оборотная стороны.
Размер карты 165×191 мм.

на верхней стороне карты используются для шифровки по прямому коду (аналогичному коду NBS и DMS) самых основных сведений о типе исследованного спектра (инфракрасный, электронный, комбинационного рассеяния), области спектра, а также об основных структурных элементах молекулы. Остальные 100 пар отверстий, расположенные по краям и в нижней части карты, использованы для шифровки по двойному коду, в котором одно сведение кодируется на двух перфорациях. Такой код обладает значительно большей емкостью: на 200 отверстиях можно зашифровать

$$\frac{100!}{98! 2!} = 4950$$

сведений вместо 200 по прямому коду. Большая емкость информации позволила с достаточной подробностью зашифровать не только сведения о структурной формуле молекулы, как это сделано в других картах, но дать и более подробную шкалу длин волн для шифровки положения полос. Подробность этой шкалы не снимает, однако, затруднений, связанных с выбором полос для шифровки и невозможностью передачи сведений об их интенсивностях. Двойной код, обладая большой емкостью, имеет недостатки меньшей наглядности и несколько большей сложности отбора карт. (Применение такого кода для нескольких пар перфораций, как, например, для шифровки фамилии автора в системе DMS (рис. 11), позволяет при помощи треугольных сеток сделать его наглядным).

Рассмотрим кратко также вопрос о документации других молекулярных спектров. Документация спектров комбинационного рассеяния, учитывая меньшее их богатство полосами, не представляла бы дополнительных сложностей, если бы, как и для инфракрасных спектров, ограничиться шифровкой положений полос. При желании же обеспечить более полное кодирование сведений о спектрах комбинационного рассеяния возникает необходимость учета степени депolarизации и ширины линий, широко используемых в практике. Количественные измерения интенсивностей в спектрах комбинационного рассеяния распространены пока менее, чем в инфракрасных, но этот вопрос и для спектров комбинационного рассеяния потребует своего разрешения.

Говоря о существующих материалах по комбинационным спектрам, следует прежде всего указать на «Атлас спектров комбинационного рассеяния углеводородов»¹², составленный Г. С. Ландсбергом, П. А. Бажулиным и М. М. Суцинским и содержащий примерно для 300 веществ исчерпывающие сведения о спектрах и методике их измерений. Мы не располагаем данными или образцами карт DMS или NBS для комбинационного рассеяния, хотя по нашим сведениям эти карты выпускаются или предполагаются к выпуску.

Вопрос о документации электронных спектров поглощения характеризуется существенными отличительными чертами. Если отвлечься от колебательной структуры, которой обладает лишь небольшая часть спектров, то для каждого вещества в этом случае речь идет лишь о нескольких максимумах, отвечающих ближайшим электронным переходам, расположенным в доступной оптической области. Это обстоятельство позволяет, несмотря на ограниченную информационную емкость перфокарт, зашифровать на них также интенсивности полос, а не только их положения. Последнее было бы бессмысленно, учитывая кажущуюся малую характерность сплошных электронных спектров, которая, несмотря на легкость экспериментальных исследований, задержала надлежащее их использование в науке и технике.

Документация электронных молекулярных спектров в виде справочников и атласов уже предпринималась. Укажем здесь на известные справочники Ландольта и на Атлас спектров ароматических соединений Р. Фриделя и М. Оркина¹³. В последнее время появились предложения

о создании системы документации с помощью карт с перфорацией по краям. Мы кратко рассмотрим этим предложения главным образом в отношении шифровки спектров. Шифровка химических и физико-химических данных во всех системах—уже рассмотренных выше и рассматриваемых—значительно более полна и совершенна, чем шифровка спектров, поэтому ее различия не являются существенными, хотя для электронных и колебательных спектров существенны разные элементы строения молекул и, кроме того, для электронных спектров более важно знание условий измерений (агрегатное состояние изучаемого вещества, растворитель, температура и т. п.).

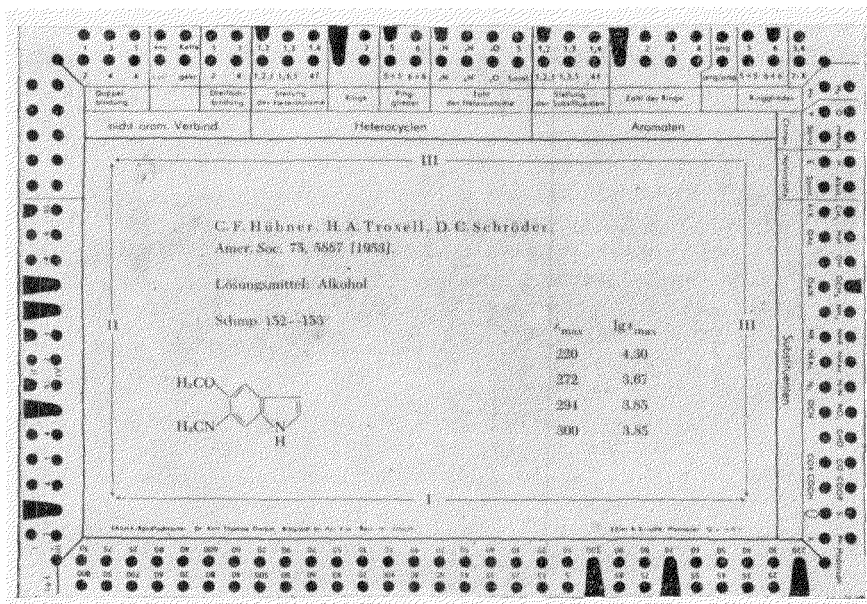


Рис. 16. Проект Штайдле¹⁴ карты для электронных спектров поглощения.

В статье В. Штайдле¹⁴ предложено использовать, аналогично системе DMS, нижнюю и левую части карты для шифровки спектра. В нижней части (рис. 16) на двух рядах отверстий по простому коду шифруются положения максимумов полос, отвечающих электронным переходам в области 200—850 $m\mu$: через 5 $m\mu$ до 350 $m\mu$, через 10 $m\mu$ до 500 $m\mu$, через 20 $m\mu$ до 700 $m\mu$ и т. д. Слева четыре группы по 7 отверстий служат для шифровки интенсивностей максимумов четырех полос поглощения в логарифмической шкале. При этом могут быть зашифрованы значения логарифма коэффициента поглощения $\lg \epsilon$ от 2,0 ($\epsilon=100$) до 4,85 ($\epsilon=70000$), причем весь интервал разбит на 15 ступеней, в соответствии с данными таблицы II. В каждой группе из 7 отверстий первые три служат для шифровки первой цифры значения $\lg \epsilon$ (2, 3 или 4), на следующих четырех отверстиях шифруются десятичные доли значения $\lg \epsilon$ (соответственно 0, 1; 0, 4; 0, 6; 0,85). Способ шифровки химических данных ясен из рисунка.

К. Кендалл¹⁵ предлагает несколько иную систему шифровки, используя один боковой ряд (22 пары отверстий) для указания длин волн максимумов полос: первая пара—короче 210 $m\mu$, затем через 10 $m\mu$ до 300 $m\mu$, через 20 $m\mu$ до 400 $m\mu$, через 50 $m\mu$ до 700 $m\mu$, последняя пара—длиннее 700 $m\mu$. Полосы, в которых $\epsilon < 1000$, шифруются вырезом мелкой лунки, захватывающей внешнее отверстие, при $\epsilon > 1000$ —вырезом глубокой лунки

Таблица II

lg ϵ	ϵ	Тип соединения
2,0	100	насыщенные и ненасыщенные кетоны
2,1	126	
2,4	250	
2,6	400	
2,85	700	
3,0	1000	бензол
3,1	1260	
3,4	2500	
3,6	4000	алкилбензолы длинноволновые полосы конденсированных ароматических соединений
3,85	7000	
4,0	10000	фенол, алкилфенолы
4,1	12600	
4,4	25000	производные пиридина
4,6	40000	
4,85	70000	полизамещен. произв. бензола, ароматические карбоновые кислоты, крезолы
		кетоны, альдегиды с одной ненасыщ. связью
		вторые полосы конденсиров. аромат. соединений
		производные бензола: анилин, стирол, ацетофенон и т. д.
		полиены, красители, конденсир. аромат. соед.

через оба отверстия. Химического кода мы не рассматриваем, отметим только, что система включает также и кодирование сведений о задаче исследования, в котором получен спектр. Тот же автор предлагает для карт с внутренней перфорацией системы IBM код для шифровки на первых двенадцати колонках сведений о спектре, включая интенсивности со ступенями $\Delta \lg \epsilon = 1$.

В обеих рассмотренных системах свободные поля в середине карт с двух сторон используются, очевидно, для нанесения спектральных кривых, структурных формул и других данных. Мы не имеем никаких сведений о выпуске или применении описанных карт.

Сравнение рассмотренных систем документации (речь идет в первую очередь об инфракрасных спектрах поглощения) весьма затруднительно, поскольку каждая из них обладает своими преимуществами и недостатками, а также, главным образом, потому, что решение вопроса о целесообразности той или иной системы не может быть независимым от задач, для решения которых привлекается документация. Поэтому мы ограничимся кратким сопоставлением характеристик этих систем.

Атлас Садтлера весьма удобен для установления связей между спектрами и строением отдельных соединений. Имея ограниченное число входов, этот атлас может уступать системам с внешней перфорацией (NBS, DMS), если необходимо достаточно быстро устанавливать такие связи для групп соединений и групп спектров. При необходимости использования широкого экспериментального материала для определения корреляционных связей между спектрами и структурными характеристиками молекул наиболее пригодны, очевидно, карты ASTM с машинной сортировкой. Если, однако, исходить из чисто практических соображений, то следует учесть, что для некоторых лабораторий, особенно небольших, именно вследствие трудности получения машин, процесс ручной сортировки карт NBS и DMS может иметь преимущества перед машинной сортировкой карт ASTM. Более того, некоторые задачи, даже для групп соединений и спектров, могут быть разрешены и с помощью атласа с ограниченным числом входов, вообще без применения перфокарт.

По полноте информации карты ASTM уступают атласу Садтлера, а также картам DMS и NBS, в которых кодированная информация

эффективно сочетается с обычной. Поэтому в тех случаях, когда необходимо установление корреляционных связей с незашифрованной на картах ASTM информацией, необходимость применения атласа Садтлера снижает скорость обработки этих карт. В таких случаях язык механической информации не может быть универсальным, если не сделать систему слишком сложной, как, например, упомянутая выше система ⁵.

Следует полагать, что возможным рациональным решением вопроса о документации молекулярных спектров было бы использование относительно простой системы для широкого употребления и одновременно применение быстродействующей электронной машины для установления сложных корреляционных связей, необходимых для решения научных и практических задач, выдвигаемых отдельными лабораториями. Информационная емкость запоминающих устройств такой машины должна быть при этом достаточной для помещения в ней всей практически важной информации, содержащейся в спектрах.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. R. B. Barnes, U. Liddel, V. Z. Williams, *Ind. Eng. Chem.* **15**, 659 (1943). H. W. Thompson, *J. Chem. Soc.* 328 (1948). N. B. Colthup, *J. Opt. Soc. Amer.* **40**, 397 (1950). Л. Беллами, *Инфракрасные спектры молекул*, ИИ, Москва, 1957, (New York, 1954). Н. М. Рандалл и др., *Infra-Red Determination of Organic Structures*, New York, 1949.
 2. R. Casey, I. Perry, *Punched Cards*, New York, 1951.
 3. *Anal. Chem.* **23**, 15A (1951) (October).
 4. L. Kuentzel, *Anal. Chem.* **23**, 1413 (1951).
 5. W. King, H. Blanton, J. Frawley, *J. Opt. Soc. Amer.* **44**, 397 (1954).
 6. D. Baker, N. Wright, A. Opler, *Anal. Chem.* **25**, 1457 (1953).
 7. H. W. Thompson, *J. Chem. Soc.* 4501 (1955).
 8. G. Bergmann, G. Kresze, *Angew. Chemie* **22**, 685 (1955).
 9. R. Mecke, E. Schmid, *Angew. Chemie* **65**, 253 (1953).
 10. A. Martin, *Industrial Chemist* 464 (1956).
 11. C. Creitz, *Appl. Spectroscopy* **7**, 40 (1953).
 12. Г. С. Ландсберг, П. А. Бажулин, М. М. Сущинский, *Атлас спектров комбинационного рассеяния углеводородов*, Изд. АН СССР, Москва, 1956.
 13. R. Friedel, M. Orchin, *Ultraviolet Spectra of Aromatic Compounds*, New York, 1951.
 14. W. Steidle, *Zs. Naturforsch* **10**, 284 (1955).
 15. C. Kendall, *Appl. Spectroscopy* **8**, 158 (1955).
-