



Standard Practices for Identification of Material by Infrared Absorption Spectroscopy, Using the ASTM Coded Band and Chemical Classification Index¹

This standard is issued under the fixed designation E 204; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 These practices describe a data system generated from 1955 through 1974. It is in world-wide use as the largest publicly available data base. It is recognized that it does not represent the optimum way to generate a new data base with the most modern computerized equipment.

1.2 These practices describe procedures for identification of individual chemical substances using infrared absorption spectroscopy and band indexes of spectral data. Use of absorption spectroscopy for qualitative analysis has been described by many (1-8),² but the rapid matching of the spectrogram of a sample with a spectral data in the literature by use of a band index system designed for machine sorting was contributed by Kuentzel (9). It is on Kuentzel's system that the ASTM indexes of absorption spectral data are based.

1.3 Use of these practices requires, in addition to a recording spectrometer and access to published reference spectra, the encoded data and suitable data handling equipment.³

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 ASTM Standards:

E 168 Practices for General Techniques of Infrared Quantitative Analysis⁴

E 932 Practice for Describing and Measuring Performance of Dispersive Infrared Spectrometers⁴

E 1252 Practice for General Techniques for Qualitative Infrared Analysis⁴

3. Summary of Practices

3.1 A representative sample of the material to be analyzed is separated into its individual components, if required, and each component is introduced into a suitable sample cell or matrix, mainly according to its physical state. The spectrum is recorded over a characterizing range. The choice of spectral range and instrument is dictated by a general consideration of the chemical nature of the sample (3-5). A note is made of the spectral positions of prominent absorption bands and, optionally, of known chemical and physical properties of the material. The qualitative chemical composition of the material may then be identified by searching the coded data file for compounds having matching characteristics. Details on searching procedures are available elsewhere.⁵ Details of the code are in the following sections.

4. Apparatus

4.1 *Infrared Spectrophotometer*—A spectrophotometer with capabilities equivalent to an instrument with a rock salt prism operated under parameters compatible with Analytical Spectra (8, 10) and with wavelength accuracy to 0.05 μm by comparison with the indene spectrum in Practice E 932.

4.2 Laboratory procedures for obtaining spectra are described in Refs (3-5) and in Practices E 168, and E 1252.

4.3 *Data-Handling Equipment*—It is possible to convert data on the ASTM magnetic tape to IBM cards, and to use sorters or collators to manipulate the data. However, the file is large and it is more efficient, and with good software, more effective, to use computers. These may be either dedicated or time-shared. Thus, the minimum equipment requirement is a

¹ These practices are under the jurisdiction of ASTM Committee E-13 on Molecular Spectroscopy and are the direct responsibility of Subcommittee E13.03 on Infrared Spectroscopy.

Current edition approved Dec. 10, 1998. Published August 1999. Originally published as E 204-62T. Discontinued 1998. Reinstated December 1998.

² The boldface numbers in parentheses refer to the list of references at the end of these practices.

³ The *ASTM Infrared Spectral Index, AMD 33* and its supplements may be purchased in the form of magnetic tapes, from Sadtler Research Labs., Inc., 3316 Spring Garden St., Philadelphia, PA 19104.

⁴ *Annual Book of ASTM Standards*, Vol 03.06.

⁵ Publicly available systems are as follows: *IRGO*, Chemir Labs., 761 W. Kirkham, St. Louis, MO 63122; *SPIR* (Canada only), National Research Council, 100 Sussex Dr., Ottawa, Ontario, Canada K1A 0R6.

computer, a program, and the coded data (and either batch processing facilities) or a teletypewriter or terminal with modem for accessing these resources⁵ for interactive searches.

5. Index

5.1 The index data on approximately 145 000 spectra are available on magnetic tape. The main absorption bands of each spectrum are coded to the nearest 0.1 μm .

5.2 In addition to the code for spectral data of chemical substances, there are codes for chemical-structure classification, empirical formula, melting or boiling point, and serial number reference. Other codes include data on sample state, wavelength intervals of strongest bands, and no-data areas. For a given substance, the coded spectral data are almost invariably unique as is the pattern for coded chemical structure and physical properties. Variables may be searched in any desired combination to locate a standard spectrum similar to that of a sample of unknown composition, to correlate type of structure with absorption band positions, to locate spectra of compounds having given structural features in common, and in other ways that are too numerous to include here.

5.3 Spectral and chemical data from the user's own laboratory may be coded in a compatible system from details given in subsequent sections.

5.4 Molecular formula-name tabulations comprise complementary data systems for use in conjunction with the spectral band codes and chemical classification tapes. These carry the molecular formulas, chemical names, and reference serial numbers for the compounds included in the indexes described in 5.1 and 5.2. The tapes are commercially available and the indexes have been published in book form as alphabetical, numerical, and molecular formula indexes (**11,12,13**). These books enable one to determine the name of the compound involved from a knowledge of the serial number of a spectrogram or to locate a published standard spectrogram for a compound when the name is known. The serial-number listing permits one to obtain the names of possible solutions to analytical problems from spectra serial numbers produced by search operations even though complete files of standard spectra (as listed in Table 1) are not at hand. Often the name of the compound together with other available information will suffice; however, it is desirable to have as many standard spectra as feasible on hand for detailed study and comparison, because positive identification depends upon matching the unknown spectrum with one from published material or one obtained from a *bona fide* sample of the compound. The molecular formula and alphabetical indexes are useful for accessing band data for a suspected answer to an unknown.

6. General

6.1 The system described below is designed to handle the spectral absorption data obtained in the spectral range from 2 to 16 μm , and the system provides for a band-position coding resolution of 0.1 μm .

6.2 The original coding was on an IBM card format. The numerical values therefore correspond to columns and rows. See Fig. 1.

6.3 Columns 1 through 15 are used for coding absorption band positions.

TABLE 1 Catalogs of Spectrograms Covered by ASTM Punched Cards Indexing Infrared Absorption Data

A	API Research Project 44 ^A
B	User's own file of spectrograms ^B
C	Sadtler catalog of spectrograms ^C
D	NRC-NBS file of spectrograms ^D
E	Literature
F	Documentation of Molecular Spectroscopy ^E
G	Coblentz Society Spectrograms ^F
H	Chemical Manufacturer's Association (CMA) ^G
J	Infrared Data Committee of Japan ^H
K	Aldrich Library of Infrared Spectra ^I , 1970 Edition

^A American Petroleum Institute, Research Project 44, *Infrared and Ultraviolet Spectral Data*, Texas Agricultural and Mechanical College, College Station, TX, 1943 to date. Loose-leaf.

^B Users are encouraged to submit spectrograms (or the pure compound in some cases) to one of the other organizations listed. It is unlikely that any individual laboratory can code its spectral data and punch cards at the cost of the ASTM cards (about one cent each).

^C *Standard and Commercial Spectra*, Sadtler Research Laboratories, 3316 Spring Garden St., Philadelphia, Pa. 19104. Loose-leaf. The Sadtler organization also offers a "Spec-Finder" book method of matching spectrograms with those in its catalog.

^D *National Research Council-NBS Committee on Spectral Absorption Data*, National Bureau of Standards, Washington, D. C. 20025. Card file.

^E *The DMS System*, Butterworth Scientific Publications, London WC2. Distributed in U. S. by Butterworth, Inc., 7235 Wisconsin Ave., Washington, D. C. 20014.

^F *Coblentz Society Spectra*, sold by Sadtler Research Laboratories, 3316 Spring Garden St., Philadelphia, Pa. 19104 and The Coblentz Society, Inc., P.O. Box 9952, Kirkwood, MO 63122.

^G Chemical Manufacturer's Association (CMA) 1825 Connecticut Ave., N. W., Washington, D. C. Loose-leaf. Spectra are no longer available from CMA.

^H Infrared Data Committee of Japan, Sanyo Shuppan Doeki Co., Inc., Hoyu Bldg., 8, 2-chrome, Takaracho, Chuo-ku, Tokyo, Japan. Card file distributed in U. S. by Preston Technical Abstracts Co., 1718 Sherman Ave., Evanston, Ill.

^I The Aldrich Library of Infrared Spectra, Aldrich Chemical Co., 940 N. St. Paul St., Milwaukee, Wis. 53233.

6.4 The chemical classification code is in columns 32 through 57, and columns 58 through 62 provide for coding the number of C, N, O, and S atoms in the compound under consideration. A melting or boiling point is coded in 63 to 65. The rest of the card provides space for the private use of individual laboratories and the identification of the source of the coded data. The codes concerned with each of these areas are discussed separately.

CODING OF INFRARED ABSORPTION BANDS (COLUMNS 1 THROUGH 25)

7. Codes for Absorption Band Positions

7.1 *Columns 1 to 15 of "A" Cards* (Note)—Coding is done in terms of wavelength in micrometres. From columns 1 through 15, the column number is taken as the whole number value of the absorption band, and the fractional part is rounded to the nearest 0.1 μm (values ending in five hundredths are considered as next higher tenths) and the number corresponding to the 0.1 μm value is added to the number of the column. Thus a band at 7.38 μm is coded to correspond to position 4 in column 7, for a value of 7.4. The coding resolution of 0.1 μm has been found to be adequate for searching and correlating published spectra.

NOTE 1—"A" is the designation for rock salt region infrared data (see 18.4).

7.2 *Columns 1 to 25 of "G" for Far-Infrared*—The coding of far-infrared absorption bands is done in terms of wavelength in micrometres. The whole number value of the band position

INFRARED ABSORPTIONS 2-16 MICRONS TO 0.1 MICRON										CHEMICAL CLASSIFICATION (SEE CODE CHART)										NUMBER OF ATOMS			- + D P + + M P		IDENTIFICATION																																																						
										A B										C N O S					SERIAL NO. VOL. PAGE																																																						
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0																														
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1																				
WYANDOTTE-ASTM																																																																															
PUNCHED CARD INDEX TO																																																																															
INFRARED SPECTRAL ABSORPTION DATA																																																																															
DISTRIBUTED BY																																																																															
AMERICAN SOCIETY FOR TESTING MATERIALS																																																																															
1916 RACE STREET																																																																															
PHILADELPHIA 3, PENNSYLVANIA																																																																															
AND IN COOPERATION WITH																																																																															
COMMITTEE E-13																																																																															
1955 BY THE AMERICAN SOCIETY FOR TESTING MATERIALS, PHILADELPHIA, PA.																																																																															

FIG. 1 Infrared Spectral Data Card

is obtained by adding 10 to the column number and the nearest tenth of a micrometre is represented by the decimal value to the nearest tenth. Thus, a band at 18.57 μm is coded as 8.6.

7.3 To indicate the range of data covered by the spectrogram, an "x" code is coded for each column that codes a spectral range where no data are available. This is to distinguish such regions from those in the spectrogram that have been examined and found to contain no bands of sufficient intensity to code, or to mark those regions where the spectral data are obscured by strong solvent bands. Additionally, a "y" code is added to each column that indexes a very strong band. The coding of such strong bands is limited to a very few, usually about three, which may be expected to persist in the spectrum of a considerably diluted sample of the material. Use of such codes may be made in the analysis of mixtures where individual components may be present in relatively low concentrations so that only the strongest bands are readily detectable.

8. Criteria for the Selection of Bands to be Coded

8.1 Experience has shown that it is not desirable to code all of the bands of most spectra. Major and medium strength bands are coded to identify the compounds uniquely. However, coding of too many weak bands minimizes the effectiveness of negative searching, which is valuable for mixtures. Therefore, the selection of which bands to code and which to omit requires some judgment; and because of the nature of published spectrograms, the judging can be guided only by rather flexible rules. Several factors enter into the determination of the strength of an absorption band, and what may be a good set of factors for the production of an excellent spectrogram from one material is not necessarily a good set to provide a spectrogram from another material. Moreover, the quality of published spectra varies widely and any system of coding absorption bands must allow for making the best possible use of all such data.

8.2 As a general rule, bands selected to be coded have an absorbance ratio with the strongest band in the spectrogram of 1:10 or more. This means that when the strongest band has between 1 and 5 % transmittance, bands are coded which have 70 % or less transmittance as measured from a reasonably

adjacent background (not necessarily at 100 % transmittance); or if the strongest band is between 5 and 20 % transmittance, bands are coded which have 80 % or less transmittance as measured from a reasonably adjacent background. Thus, to be coded, a band stands out from its adjacent background, at least on one side, by 20 to 30 % transmittance on the chart. Therefore, "shoulders" and weak bands on the sides of strong bands are not coded. Likewise, bands whose percent transmittance may be as low as 60 to 50 as read from the chart, but which extend from backgrounds having transmittance values of 80 to 70 %, are not coded. Some examples are provided in Fig. 2.

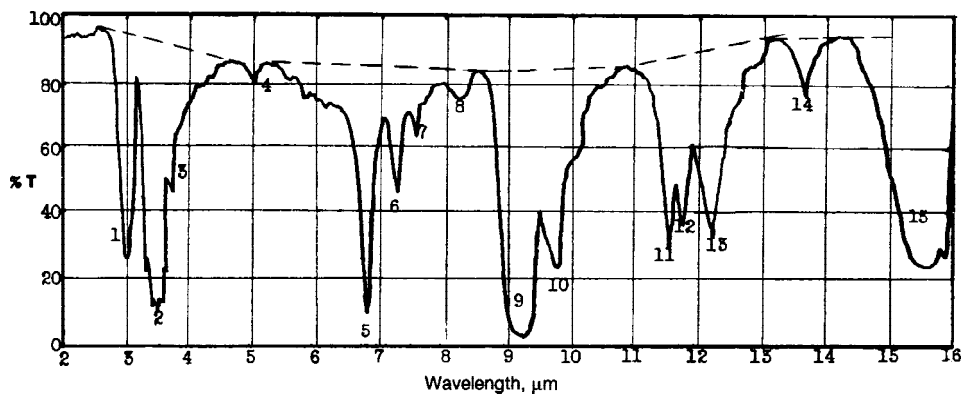
8.3 Searching absorption band data is much the same as coding the bands. First, the spectrogram of the unknown material should have its strongest bands between 1 and 20 % transmittance since it is to be compared with data coded on that basis. Then one proceeds by two different methods depending upon whether the unknown is a single component or is a mixture of two or more components in roughly equivalent amounts. In the former case, positive searching on the bands is in order, while the latter case requires that negative inputs be included in the search request. Each method is discussed briefly in Sections 9 and 10.

8.4 The optimum combination of searching techniques depends upon the computer algorithm used. Instructions specific for each program should be followed.⁵

9. Positive Searching for Individual Spectra

9.1 In this method, the search data are selected with the expectation that all or most of the bands in the unknown spectrogram are caused by a single compound. Search programs vary, but it is desirable that they include provisions for weighting the bands by their importance. This weighting may be systematic, as by the strength of the bands, or it may be on the basis of bands that the spectroscopist recognizes as important for other reasons.

9.2 If one cannot be certain to 0.1 μm of the location of the band, then searching should be carried out over as broad an interval as required to make certain the band is included in the search. Thus, a particular band on the original standard spectrogram may have been measured to be 13.46 μm. When it



NOTE 1—The above hypothetical spectrogram is included to assist in describing the application of rules prescribing which bands to code and which to omit from the index card. It will be noted that band No. 9 is the strongest and has a transmittance value between 1 and 5 %; therefore all bands having a transmittance of 70 % or less as measured against a reasonably adjacent background should be coded. The dotted lines indicate what is meant by such an adjacent background. The distance by which coded bands must project from such a background is equal to one and a half units of the vertical scale. Applying this rule, one can code without question the following bands: No. 1, 2, 5, 6, 9, 10, 11, 12, 13, and 15, and furthermore, bands 2, 9 and 15 should receive the “y” overpunch code also. Band No. 5, while rather strong, is not expected to persist in considerably diluted samples of this material.

It will be noted that band No. 3 was not included. This is a case of a rather weak band on the side of a strong one which has no value in sorting and need not clutter up the card. Therefore it was omitted. On the other hand, band No. 10 was included as it is prominent enough to be used in sorting operations. Also, No. 12, which does not fit the coding criteria when measured from its immediately adjacent background, is included in those bands coded because it obviously is one of three rather strong bands which are close enough together to overlap appreciably. An ill-defined shoulder on the side of band No. 10 is ignored as is the fine structure in the No. 2 band. Band No. 14 represents a possible borderline case that should not be coded. A good rule is “when in doubt, leave it out.” The spectrogram is typical of many that appear published in the literature and serves to illustrate why a coding resolution of 0.1 μm is entirely adequate.

FIG. 2 Example of Infrared Curve

was coded the position was 13.5, or number 5 in column 13. In an unknown spectrogram, this same band might be read as 13.44, or if a longer cell path was used the band may have spread to the extent that one cannot be certain whether the minimum is 13.4 or 13.5 μm. In such a case, the search program should accept 13.4 or 13.5 μm, or both, and thus not miss the desired compound.

10. Negative Searching for the Analysis of Mixtures


10.1 When the unknown infrared spectrogram represents a mixture of two or more compounds in appreciable amounts, positive searching on the absorption band positions must

follow a procedure that considers the many possible combinations of bands that may characterize an individual constituent, since it is not known which bands are produced by each component. Such an approach may be carried out directly on the spectrogram of a mixture. However, considerable time may be saved if the bands are subjected to “negative” searching to eliminate all of the spectra that do have bands in regions of the spectrum where the unknown spectrogram does not have bands, since none of these could possibly be a component of the mixture. Positive searching of the reduced spectral file which results from the negative searching is more selective than searching the total file.

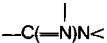
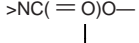
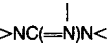
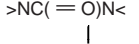
TABLE 2 Chemical Classification Code Chart Part A

Row	Column 32 Elements	Column 34 Structure	Row	Column 36 Code Units	Column 38 Miscellaneous	Row
0	O	acyclic	0	12 or more	solid	0
1	N	alicyclic	1	1	liquid	1
2	S	aromatic	2	2	gas	2
3	F	heterocyclic	3	3	organo-metallic	3
4	Cl	fused alicyclic	4	4	isotopic	4
5	Br, I	fused aromatic	5	5	indeterminate	5
6	P, Bi	fused heterocyclic	6	6	solution	6
7	As, Sb		7	7	polymer	7
8	Si, Ge	3 or 4-member ring	8	8	chelate	8
9	Sn, Pb	5-member ring	9	9	hydrate	9
X	B, Al	6-member ring	X	10	salt plate	X
Y	other	7-or more member ring	Y	11		Y

Row	Column 33 Unsaturation	Column 35 Rings—Chains	Row	Column 37 Substitutions	Column 39 Miscellaneous	Row
0	ring	rings	0	[mono]	cis	0
1	1	1	1	1 [1, 2]	trans	1
2	2	2	2	2 [1, 3]	spiro	2
3	3	3	3	3 [1, 4]	dextrorotary	3

Row	Column 33 Unsaturation	Column 35 Rings—Chains	Row	Column 37 Substitutions	Column 39 Miscellaneous	Row
4	4	4	4	4 [1, 2, 3]	levorotary	4
5	5	5	5	5 [1, 2, 4]	symmetrical	5
6	6	6	6	6 [1, 3, 5]	unsymmetrical	6
7	7	7	7	7 [1, 2, 3, 4]	vicinal	7
8	8	8	8	8 [1, 2, 4, 5]	salt	8
9	9	9	9	9 [1, 2, 3, 5]	inorganic ester	9
X		10	X	10 [penta]		X
Y	—C=C—	11 or more	Y	[hexa]	inorganic	Y

Part B

Row	Column 40 C—H	Column 42 O	Column 44 N	Row	Column 46 S	Column 48 N—O	Row
0	—CH ₃ methyl	—C(=O)OH		0	—C(=S)S—		0
1	—C ₂ H ₅ ethyl	—C(=O)O—		1	—C(=S)H		1
2	—C ₃ H ₇ <i>n</i> -propyl	—C(=O)H	—CN	2	>C=S	—C(=N)O—	2
3	—C ₃ H ₇ isopropyl	>C=O	—NC	3	—SC(=S)S—	—NCO	3
4	—C ₄ H ₉ <i>n</i> -butyl	—OC(=O)O—	—NH ₂	4	—SH	—OCN	4
5	—C ₄ H ₉ isobutyl	C(OR) ₄	>NH	5	—S—	>NNO ₂	5
6	—C ₄ H ₉ <i>sec</i> -butyl	—C(OR) ₃	>N—	6	—SS—	>NNO	6
7	—C ₄ H ₉ <i>ter</i> butyl	—OH	>C=N—	7	>S*	—NN(=O)—	7
8	—C ₅ H ₁₁ <i>n</i> -pentyl	—O—	= NN =	8	= S = S	—ONO ₂	8
9	—C ₆ H ₅ phenyl	—OO—	>NN<	9		—ON=O	9
X	other	other	= N=N	X	other	other	X
Y	heterocyclic	heterocyclic	heterocyclic	Y	heterocyclic	heterocyclic	Y

Row	Column 41 C—H	Column 43 O	Column 45 N	Row	Column 47 S	Column 49 N—O	Row
0	—CH=CH ₂ vinyl	>O*	—N=N—N<	0		—NO ₂	0
1	>C=CH ₂ <i>ter</i> methylene	—O ₃ —	—N≡N*	1		—NO	1
2	= CHCH ₃ ethylidene	>C(—O—) ₂	>N*	2		= NO—	2
3	—CH ₂ CH=CH ₂ allyl		—NH ₄	3		>NO—	3
4	—CH=CHCH ₃ propenyl		>NCN	4		>N=O	4
5	—C(CH ₃)=CH ₂ isopropenyl		—N ₃	5			5
6	= C(CH ₃) ₂ isopropylidene			6			6
7	—C≡CH ethynyl			7			7
8	—C≡CCH ₃ 1-propynyl			8			8
9	—CH ₂ C≡CH 2-propynyl			9			9
X	conjugated	conjugated	conjugated	X	conjugated	conjugated	X
Y				Y			Y

Row	Column 50 N—	Column 52 O—S	Row	Column 54 N—O—S	Column 56 Elements	Row
0	>NC(=S)N<	—OC(=S)S—	0	>NC(=S)O—	Se, Te, Po	0
1	—SC(=N)N<	—SC(=O)S—	1	>NC(=O)S—	Ga, In, Tl	1
2	>NC(=S)S—	—OC(=S)O—	2	—C(=O)NS—	Zn, Cd, Hg	2
3	—SC(=N)S—	—OC(=O)S—	3	—C(=S)NO—		
4	—C(=S)N<	—C(=S)O—	4	—OC(=N)S—		
5	—C(=N)S—	—C(=O)S—	5	—S(=O)N<	Cu, Ag, Au	3
6	>NSN<	—S(O ₂)O—	6	>NS(O ₂)N<	Fe, Co, Ni, Mn	4
7	= NS—	S _x O ₆	7	>NS(O ₂)O—	Cr, Mo, W, U	5
8	> NS—	(x = 2–6)	8		V, Cb, Ta, Pa	6
9	—N = S	>SO ₂	9	>NS(=O)O—	Ti, Zr, Hf, Th	7
X	>S = N—	>S = O	X	—NSO	Se, Y, La, Ac	8
Y	other	—OSO—	Y	other	Ru, Rh, Pd, Os, Ir, Pt	9
		other			rare earths	X

Y

heterocyclic

heterocyclic

Y

heterocyclic

heterocyclic

Y

Row	Column 51 N—S	Column 53 O—S	Row	Column 55 N—O—S	Column 57 Elements	Row
0		—OS(O ₂)O—	0		Li	0
1		—OS(=O)O—	1		Na	1
2			2		K	2
3			3		Rb, Cs	3
4			4		Be	4
5			5		Mg	5
6			6		Ca	6
7			7		Sr, Ba	7
8			8			8
9			9			9
X	conjugated	conjugated	X	conjugated	conjugated	X
Y			Y		other	Y

ORGANIC CHEMICAL CLASSIFICATION CODE (COLUMNS 32 THROUGH 57)

11. General

11.1 The chemical classification code for organic compounds is designed to present the chemical structure of such compounds in a convenient form for use in the preparation and use of ASTM indexes of absorption spectral data. The accompanying chart (Table 2) relates the code positions on the card in terms of column and row numbers to the coded items of structural features used to characterize compounds. Reference to codes are made by giving the column number followed by the row designations. For example, 32-0,2,y indicates positions 0,2, and “y” in column 32. The overpunch positions are referred to as “x” and “y” to avoid digital confusion, but they correspond to the 11 and 12 positions. While every effort was made to keep the codes simple and unambiguous, the complexity of some structures that must be coded requires that a few rules be provided in the interest of uniformity. These rules, together with notes on the interpretation of them, follow in a column-by-column discussion of the chart. At the end will be found general instructions for the application of the codes together with a number of examples (Table 3).

11.2 *Part A*—This section of the chart is concerned with providing a means of coding elements commonly found in organic materials; the number, size, and kind of gross structural features; the number and locations of code units or substituent groups of atoms upon these gross features; and a number of general descriptive terms that may be applied to the materials encountered.

11.2.1 *Column 32*—This column provides for the coding of the identity of elements commonly found in organic compounds. Carbon and hydrogen are not coded directly, but hydrocarbons are indicated when there are no value 32 codes. The code is the designated value for each different element regardless of the number of such elements in the compound. The coding of less common elements is provided for in columns 56 and 57. Whenever any of these elements or any not listed in the chart are coded, a code of “y” or “other” should be made in column 32.

11.2.2 *Column 33*—This column codes the type and location of unsaturated carbon-to-carbon bonds. In every case, except for aromatic unsaturation, the presence of such unsaturation is coded as to type (that is, double bond or triple bond, or both), by 33-x or 33-y, or both. Numbers in this column are

used to indicate the location of these unsaturated bonds subject to the following rules:

11.2.2.1 If the unsaturation is located in a ring, then a code of 33-0 is required. When this is lacking, it is understood that unsaturation in a chain is being coded.

11.2.2.2 Unsaturation at positions requiring numbers higher than nine, Greek letters, or primed numbers are not coded.

11.2.2.3 The use of the position codes is restricted to compounds containing a single chain, a single ring, or a fused ring system where the Geneva System for chains and the Patterson Ring Index for cyclic compounds can be applied without ambiguity.

11.2.2.4 Unsaturation in benzene rings, fused or otherwise, or in alicyclic rings as a result of fusion with aromatic rings is not coded here.

11.2.2.5 Where both cyclic and chain systems are present in a single compound and unsaturation is present in only one or the other, it is to be coded as to location.

11.2.2.6 Where both cyclic and chain systems are present in a single compound and both contain unsaturation, the position code is applied to the largest ring or fused ring system.

NOTE 2—Space for coding unsaturated hydrocarbon groups and conjugated unsaturation is provided for in column 41.

11.2.3 *Column 34*—This column is used to code the major structural features of a compound and is largely concerned with the type and size of rings. The use of these codes in describing a molecular structure is governed by the following rules:

11.2.3.1 An “acyclic” code is used whenever there are one or more carbon atoms which are not part of a ring. Thus, methane, benzaldehyde, toluene, ethyl benzene, and benzoic acid require “acyclic” codes, but phenol, aniline, and phenyl hydrazine would not.

11.2.3.2 Each individual type of ring present in a single molecule is coded by a single code. Each member of a fused system is coded separately if different types are involved. All rings other than aromatic or heterocyclic are considered alicyclic and only benzene rings are coded “aromatic.”

11.2.3.3 No portion of any ring, except that involved in fusion, is coded more than once. Thus, multiple ring systems formed by bridging are individually coded but the enveloping ring is not.

11.2.3.4 The size of aromatic rings is not coded.

NOTE 3—Spiro compounds are coded in column 39 as well as in column 34.

TABLE 3 Examples of the Types of Compounds Coded by the Code Units in the Chemical Classification Code Chart

NOTE 1—Following are examples of the types of compounds which the various code units in the chart may index. It is to be understood that these examples do not restrict the use of the code units in the indexing of other types of compounds in which they may appear.

42-0 acids	48-4 cyanates
42-1 esters, salts, lactones, anhydrides	48-5 nitro amines
42-2 aldehydes	48-6 nitroso amines
42-3 ketones	48-7 azoxy compounds
42-4 carbonates	48-8 nitrates
42-5 ortho carbonates	48-9 nitrites
42-6 ortho carboxylates	
42-7 alcohols, phenols	49-0 nitro compounds
42-8 ethers, oxy compounds	49-1 nitroso compounds
42-9 peroxides	49-2 isonitroso compounds, oximes
	49-3 amine oxides
43-0 oxonium compounds	
43-1 ozonides	50-0 thiourido compounds
43-2 acetals	50-1 thiocarbamyl compounds
	50-2 thioamides, thiomides
44-0 amidines	50-3
44-1 guanidines	50-4 thiocyano compounds
44-2 nitrilo or cyano compounds	50-5 isothiocyano compounds
44-3 isonitrilo compounds	50-6 diamino sulfides
44-4 primary amines	50-7 sulfimes, sulfenamides
44-5 secondary amines	50-8 sulfamino and sulfinyl compounds
	50-9 sulfilimines
44-6 tertiary amines	
44-7 imines	52-0 dithiocarbonates
44-8 hydrazones, hydrazines	52-1 thiocarbonates
44-9 azo or diazo compounds	52-2 thiolic, thionic compounds, carbothioates
	52-3 sulfonates
45-0 triazenes	52-4 sulfimates
45-1 diazonium compounds	52-5 thiosulfimates
45-2 quaternary ammonium compounds	
45-3 ammonium compounds	52-6 thionates
45-4 cyanamides	52-7 sulfones
45-5 triazo compounds, azides	52-8 sulfoxy compounds, sulfinyls
	52-9 sulfenates
46-0 thionothiolic compounds, carbodithioates	
46-1 thioaldehydes	53-0 sulfates
46-2 thiones, thioketones	53-1 sulfites
46-3 trithio carbonates	
46-4 thiols	54-0 thiocarbamates
46-5 sulfides	54-1 carboxamido sulfides
46-6 disulfides, polysulfides	54-2
46-7 sulfonium compounds	54-3 sulfenamides
46-8 perthio compounds	54-4 sulfamides
	54-5 sulfamates
48-0 carbamyl compounds, carbamates	54-6 sulfonyl amines, sulfonamides
48-1 ureido compounds	54-7 amino sulfimates
48-2 amides, imidic compounds, lactams	54-8 sulfinyl amines
48-3 isocyanates	54-9

11.2.4 *Column 35*—This column provides for coding the length of carbon chains or the number of rings in a compound. Use of the following rules will ensure uniformity of coding:

11.2.4.1 If there is only one ring or if there are no rings in the compound, the length of the longest, normal carbon-to-carbon chain is coded. One carbon atom is considered a “chain,” but carbon atoms in the rings are not to be counted as part of such chains.

11.2.4.2 If there are two or more rings in the compound and aromatic rings are involved, both the total number of rings and

the number of benzenoid rings are coded into the column along with a code at 35-0 to indicate that rings are being coded. Each ring in a fused ring, spiro, or bridged system is counted separately (11.2.2.3 applies here).

NOTE 4—Compounds involving one ring only are not coded at 35-0.

11.2.5 *Column 36*—Column 36 codes the total number and the number of different kinds of “code units” as identified by Part B of the chart and the rules stated below. Both numbers should be coded into this column when there is a difference between the total number and the number of different kinds. The following rules assist in arriving at the proper totals:

11.2.5.1 Consider all code designations specified by Part B of the chart except codes for “heterocyclic” and “conjugated” groups and those in columns 40 and 41.

11.2.5.2 Consider all atoms other than C, H, N, O, and S. Each such element counts as a code unit.

11.2.5.3 For the total number of code units, count each unit as many times as it appears in the structure.

11.2.5.4 For the number of different kinds of code units, count each type once.

NOTE 5—The number of different kinds of code units should equal the sum of the code assignments made under Part B of the chart, except as noted in 11.2.5.1, plus the number of different kinds of atoms other than C, H, N, O, and S that are included in columns 32, 56, and 57.

11.2.6 *Column 37*—This column provides for locating the positions of substituent groups of “code units” in a limited number of cases. It is intended that this column provide a means of differentiating molecular isomers and is not rigorously applied in coding all compounds. One should not attempt to code all substitutions in compounds where there is ambiguity as to just what is substituted on what. The following rules apply:

11.2.6.1 Substitution positions requiring numbers higher than 10 or the use of Greek letters or primed numbers are not to be coded here.

11.2.6.2 Except as provided in 11.2.6.4, use of the code is restricted to indicating substitution positions on a single carbon chain, a single ring, or a fused ring system where application of the Geneva System for chains and the Patterson Ring Index for cyclic compounds can be made without ambiguity.

11.2.6.3 In monocyclic compounds which also have acyclic or chain systems, code the location of substitutions on the ring.

11.2.6.4 In polybenzenoid compounds not involving fusion, code designations within the brackets (on the chart) are used to indicate the degree and location of substitution on the several rings using the lowest numbering arrangement.

11.2.6.5 The locations of heteroatoms in heterocyclic rings are not to be made with this code.

11.2.7 *Columns 38 and 39*—These columns provide for coding miscellaneous facts about the compounds. For the most part they are self-explanatory, but the following interpretations should be made:

11.2.7.1 Punches 38-0, 1, 2, and 6 are used to indicate both the physical state of the compound at the time it is analyzed in the spectrometer and the physical state of the compound at room conditions. Thus, codes of 38-0,6 indicate that the material is normally a solid but that it was analyzed in the

spectrometer as a solution, whereas a code of 38-0 only would indicate that the material is a solid and the spectrum was determined in the solid state. A code at 38-x indicates that a salt plate has been employed in sample preparation. A code of 38-5 is used if the structure is unknown, but not if the state is unknown. Codes of 39-5,6,7 are not to be applied to coding trisubstituted benzene or other cyclic compounds but rather to describe the arrangement of heteroatoms in heterocyclic rings such as the triazines where substitutions play no part in determining the use of the terms. Such application is not limited to rings containing one kind of heteroatom and the code may be used for both five- and six-member rings.

11.2.7.2 39-y indicates that the compound involved is inorganic and the special inorganic structure codes apply. Thus, different codes must be designated to search for inorganic compounds and organic compounds (see 14.1).

11.3 *Part B*—This section of the chart is primarily concerned with providing a means for coding groups of atoms that are commonly considered as substituent groups or reactive groups in molecules as revealed by detailed organic structural formulas. It is desirable that no particular name be associated with these structural units lest the name tend to limit the use of the code. Therefore, they are referred to as “code units” and, with the exception of the units involving carbon and hydrogen only, about which there can be little question, the units are illustrated by structural arrangements of atoms as they are commonly represented in structural formulas. An accompanying list provides examples of the type of structural groups that the unit codes may index, but the use of such codes is not to be limited by any exemplary name supplied. The sole criterion for the use of the unit codes is that the precise arrangements of atoms as depicted on the chart are present in structural formulas being sorted for or being coded for punching cards. Except for the code units involving C and H alone, no unit contains more than one carbon atom.

11.3.1 *Columns 40 and 41*—Code units involving carbon and hydrogen only are depicted here. Only the smaller of many possible such groups are listed and the coding of such structural units is confined to those illustrated. The following rules apply:

11.3.1.1 Code each unit that is observed in the structural formula in question, regardless of the simplicity or complexity of such structure.

11.3.1.2 Use the largest unit code that will characterize a group and do not code smaller parts of such a unit. Thus, if a $-\text{C}_2\text{H}_5$ group is present, code a 40-1 but do not code the $-\text{CH}_3$ group which forms a part of the larger group. A code of 40-8 is used for any straight chain longer than pentyl.

11.3.1.3 Under 41-x, code all conjugated double-bond systems in rings or chains, or both, that involve carbon only, except purely benzene ring conjugation. Do code conjugated carbon-carbon systems involving a benzene ring if there is at least one carbon-carbon double bond outside the ring or if there are two or more benzene rings forming such a system.

11.3.2 *Columns 42 through 55*—These columns provide for coding unit groups or code units involving oxygen, nitrogen, or sulfur with or without a single carbon atom. They are arranged in columns depending on whether the above elements are

involved singly, in pairs, or altogether. The following rules assist in the application of the code:

11.3.2.1 Code each unit that is observed in the structural formula, whether it be part of a ring or not.

11.3.2.2 Use the largest unit code that will characterize a group and do not code small parts of such a unit. Thus, if the unit group present is $>\text{NC}(=\text{O})\text{O}-$, characterize it by a code of 48-0 and do not code 42-3, 42-8, or 44-6. These smaller units will always be understood to be parts of the larger unit. Also, if the bonds of a code unit are satisfied with H or C atoms, the smaller units thus formed are not to be coded. Thus, $\text{H}_2\text{NC}(=\text{O})-\text{OH}$ is coded only as 48-0, without codes of 44-4 and 42-7, and $>\text{C}=\text{N}-\text{NH}_2$ requires only a 44-8 code.

11.3.2.3 Code larger groups than appear in the chart, or those involving two or more carbon atoms, by using the least number of units containing the largest number of heteroatoms. Strict application of this rule, regardless of one’s feelings for the chemistry or naming of compounds, is essential. In many cases this rule will necessitate the coding of an atom or two in each of two code units. Thus, for example, in $\text{CH}_2=\text{NNHC}(=\text{S})\text{NHNH}_2$ one can observe the following code units: $>\text{C}=\text{N}-$, $=\text{NN}<$, $>\text{NC}(=\text{S})\text{N}<$, $>\text{NN}<$, $>\text{C}=\text{S}$, $>\text{NH}_2$ and $-\text{NH}_2$. The problem is to include all of these structural arrangements in as few units as possible. One begins by selecting a central carbon atom and observing the greatest number of heteroatoms that are attached to it. In the above example this process yields the code unit $>\text{NC}(=\text{S})\text{N}<$ or 50-0. All that remains are the $>\text{NN}<$ groups which require a code of 44-8. Any other possible choice of code units such as 50-2, 44-7,8 or 46-2, 44-7,8 would not involve the largest possible units and obviously any further breakdown would not involve the least number of units. Note that in assigning a code unit to a given structural group, only the internal arrangement of atoms in the code unit must be rigidly matched. Whether the external bonds are attached to one or two or more atoms is of no consequence. Thus, a 44-8 codes all of these types of compounds: $\text{R}=\text{NN}=\text{R}$, $\text{R}=\text{NNH}_2$, and RHNNH_2 .

11.3.2.4 Conjugated double-bond systems involving the elements listed at the head of each column on the chart are coded with an “x” as indicated. A conjugated double-bond system consists of the complete series of alternate double and single bonds which may stretch through one or more benzene rings and involve two or more heteroatoms. One should code each separate system with the appropriate “x” which identifies the elements involved but, as with heterocyclic rings below, code only the system that involves all heteroatoms. This includes double-bond arrangements that are conjugated by virtue of attachment to aromatic rings.

11.3.2.5 The presence of heterocyclic rings involving elements listed at the head of the columns in the chart are coded by “y” in the appropriate columns. If two or more different heteroatoms are involved in a single ring, code only the unit that involves all of such atoms. Thus, a heterocyclic ring involving both oxygen and sulfur is coded as 52-y, and codes of 42-y and 46-y are not used. If a heterocyclic ring involves an element other than O, N, or S, either alone or with O, N, or S, then a single code of 56-y is appropriate.

11.3.2.6 Organic salts, including amine salts, etc., are to be coded in the un-ionized form and a code at 39-8 assigned to indicate that a salt is involved. Thus, diethylamine hydrochloride would be coded from the structural formula $(C_2H_5)_2NH \cdot HCl$ rather than $(C_2H_5)_2NH_2^+Cl^-$, and the code for the nitrogen group would be 44-5. Organometallic codes are used only if there is a metal-to-carbon bond involved and compounds such as metal salts of organic acids are not included. Also, only chelate compounds involving metal elements are to be coded with the 38-8 code.

11.3.3 *Columns 56 and 57*—These columns provide for coding the less common elements of organic chemistry. Such elements as are listed are coded in the appropriate position. Elements not listed either in columns 56 and 57 or in column 32 are coded at 57-y. Whenever elements are coded in columns 56 and 57, a code at 32-y is also required. However, if the only elements involved in the compound being coded are those in column 32, then there is no code at 32-y or in columns 56 and 57. Compounds containing elements other than C, N, O, and S that are involved in heterocyclic rings or conjugate double bond systems are coded in column 56 or 57, or both.

APPLICATION OF THE CHEMICAL CLASSIFICATION CODE CHART TO THE CHARACTERIZATION OF ORGANIC COMPOUNDS

12. General

12.1 The philosophy behind the development and use of the Chemical Classification Code Chart attempts to divorce the complexities of the names and chemistry of compounds from the codes and coding operations by which such compounds may be characterized. It is not intended that such characterization be unique for each different molecule since the purpose of the code is merely to provide a means of segregating compounds into related groups. Coding is based upon a detailed structural formula and a recognition of “code units” which make up the formula. These code units in many cases are the same as familiar reactive groups or radicals that enter into the chemistry and naming of organic compounds, but such names and chemistry as may be associated with the code unit must not restrict the use of the code wherever applicable under the rules previously presented. Thus, a code of 42-3 should be applied to

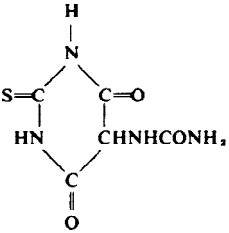
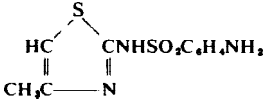
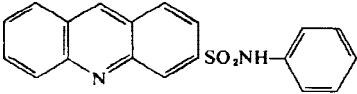
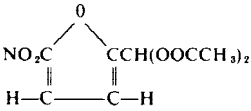
any compound that exhibits the $>C=O$ unit, common to ketones and frequently called the “keto” group, regardless of the name or the chemistry of the compound provided the unit does not form a part of another unit in the chart.

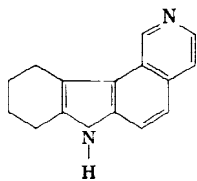
12.2 Codes in Part A apply to all compounds and it is convenient to start applying the codes in this section first, as far as possible. Thus, the initial examination of a structural formula should reveal the identity of all atoms other than carbon and hydrogen and the proper notations for columns 32, 56, and 57 can be made. Structural arrangements described in column 34 codes follow conveniently, with details provided in columns 33 and 35 being assigned next. Assignment of codes in columns 36 and 37 must await completion of the Part B coding. This leaves columns 38 and 39 of Part A which code miscellaneous and general items.

12.3 Codes in Part B require a careful examination of the structural formula being coded for the presence of the “code units” that are drawn in detail on the chart. Care should be taken to ensure that any unit accepted for coding is the largest or contains the greatest number of atoms. Such units separated by one or more carbon atoms should be coded separately. However, if two or more units are directly associated in the structure, they should be coded by using the smallest number of the largest units, even though this involves using some connecting heteroatoms twice. The association of two or more of the code units usually is made through atoms of N, O, or S so such atoms are involved in multiple use. Once the Part B code unit assignments have been made, the total number of such groups plus all other elements is indicated in column 36. Likewise, the number of different kinds of units and atoms can be determined and coded into the same column. It should be remembered that the “heterocyclic” and “conjugated” code designations and the code units in columns 40 and 41, the hydrocarbon units, are not counted as units for the column 36 totals.

12.4 Application of the chemical classification code chart can best be facilitated by a study of examples (Table 4). For sake of brevity, the codes are indicated by giving the column number first, followed by a dash and the row designations separated by commas. Thus, 32-0,2,4 means that codes of 0, 2, and 4 in column 32 are assigned to indicate the presence of oxygen, sulfur, and chlorine in the compound.

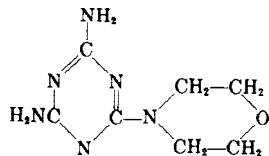
TABLE 4 Examples of Chemical Classification^A

1. Parachlorophenol	$\text{ClC}_6\text{H}_4\text{OH}$	32-0,4 34-2 36-2	37-1,4 38-0,6 42-7
2. Isoprene	$\begin{array}{c} \text{CH}_3 \\ \\ \text{CH}_2=\text{C}-\text{CH}=\text{CH}_2 \end{array}$	33-1,3,x 34-0 35-4	37-2 38-1 41-0,5,x
3. 2-thio-Pseudouric acid		32-0,1,2 34-0,3,x 35-1 36-3,4	38-0 39-6 44-6 48-1,2 50-0
4. Sulfamethylthiazole		32-0,1,2 33-0,2,x 34-0,2,3,9 35-0,1,2 36-3 38-0 39-6	40-0 44-4 45-x 50-0,y 53-x 54-6
5. Potassium salt of gamma parachlorophenoxy crotylmercaptomethyl penicillin	$\begin{array}{c} \text{O} \\ \\ \text{CH}_3(\text{ClC}_6\text{H}_4\text{O})\text{C}=\text{CHCH}_2\text{SCH}_2\text{CNHCH}-\text{CH}-\text{S}-\text{C}(\text{CH}_3)_2 \\ \\ \text{O}=\text{C}-\text{N}-\text{CHCOOK} \end{array}$	32-0,1,2,4,y 33-2,x 34-0,2,6,8,9 35-0,1,3 36-6,8 38-0 39-6,8	40-0 42-1,8 44-y 46-5 48-2 50-y 57-2
6. 3-Acridinesulfonanilide		32-0,1,2 33-0,x 34-2,5,6,x 35-0,3,4 36-2 38-0	40-9 41-x 44-7,y 54-6 55-x
7. 3-Buten-2-ol, 1-bromo-3,4-dichloro-	$\begin{array}{c} \text{Cl} \\ \\ \text{ClCH}=\text{C}-\text{CHCH}_2\text{Br} \\ \\ \text{OH} \end{array}$	32-0,4,5 33-3,x 34-0 35-4	36-3,4 37-1,2,3,4 38-1 42-7
8. Diethyl fluorophosphate	$\begin{array}{c} \text{O} \\ \\ (\text{C}_2\text{H}_5\text{O})_2\text{P}-\text{F} \end{array}$	32-0,3,6 34-0 35-2 36-4,5	38-1 39-9 40-1 42-8,x
9. 5-Nitrofurfural diacetate		32-0,1 33-0,2,4,x 34-0,3,9 35-2 36-3,4 37-2,5	38-0 40-0 41-x 42-1,8,y 49-0,x
10. 5,6,7,8 Tetrahydro-9H-pyrido (4,3c) carbazole		32-1	



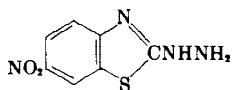
33-0,x
34-4,5,6,9,x
35-0,1,4
36-2
38-0
41-x
44-5,7,y
45-x

11. 2-(4-morpholinyl)-4,6-diamino-s-Triazine



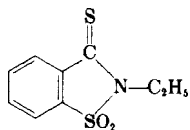
32-0,1
34-3,x
35-0,2
36-2,4
38-0
39-5
42-8
44-1,y
45-x
48-y

12. 2-Hydrazino-6-nitrobenzothiazole



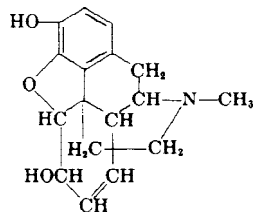
32-0,1,2
34-5,6,9
35-0,1,2
36-3
37-2,6
38-0
39-6
44-8
49-0,x
50-0,y

13. 2-Ethyl thiosaccharin



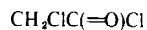
32-0,1,2
34-0,5,6,9
35-0,1,2
36-2
37-1,2,3
38-0
39-7
40-1
50-2,y
53-x
54-6

14. Morphine



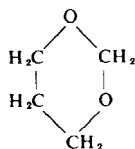
32-0,1
33-0,x
34-0,4,5,6,9,x
35-0,1,5
36-3,4
38-0
40-0
42-7,8,y
44-6,y

15. Chloroacetyl chloride



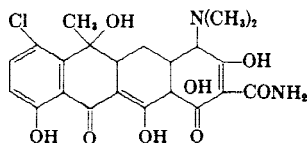
32-0,4
34-0
35-2
36-2,3
37-1,2
38-1
42-3

16. *m*-Dioxane



32-0
34-3,x
42-y
43-2

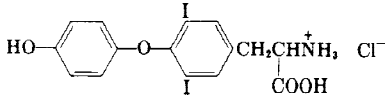
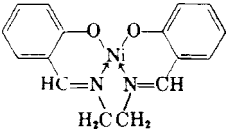
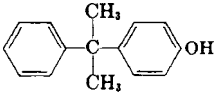
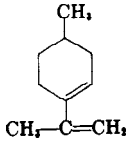

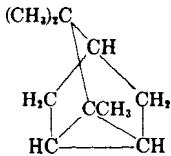
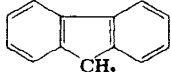
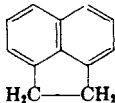
17. Aureomycin



32-0,1,4
33-0,x
34-0,4,5,x
35-0,1,4
36-5,x
38-0
40-0
42-3,7
43-x
44-6
48-2

18. *d,l*-3,5-Diiodothyronine-HCl

32-0,1,4,5
39-3,4,8

		34-0,2 35-0,2 36-6,7 37-3,9 38-0	42-0,7,8 44-4
19. <i>bis</i> -Salicylaldehyde-ethylenediamine nickel		32-0,1,y 34-5,6,9,x 35-0,2,5 36-3,5 38-0,8 39-7 42-8 44-7	45-x 56-4,y
20. <i>p</i> -Cumylphenol		32-0 34-0,2 35-0,2 36-1 37-0,3	38-0 40-0,9 42-7
21. Methyl ethyl ether	$\text{CH}_3\text{-O-C}_2\text{H}_5$	32-0 34-0 35-2 36-1	38-1 40-0,1 42-8
22. <i>n</i> -Butyl acetate	$\text{CH}_3\text{COO(CH}_2\text{)}_3\text{CH}_3$	32-0 34-0 35-4 36-1	38-1 40-0,4 42-1
23. <i>l</i> -Isopropenyl-4-methylcyclohexene-1		33-0,1,x 34-0,1,x 35-3 37-1,4 38-1 40-0 41-5,x	
24. <i>p</i> -Hydroxycyclohexanecarboxylic acid lactone		32-0 34-6,x 35-0,2 36-1 38-1 42-1,6	
25. Tricyclene		34-0,4,8,9 35-0,3 38-0 40-0	
26. Fluorene		34-4,5,9 35-0,2,3 38-0 41-x	
27. Acenaphthene		34-4,5,9 35-0,2,3 38-0 41-x	

^AThe use of the terms symmetrical, unsymmetrical, and vicinal in connection with indicating the arrangement of heteroatoms will be useful here. Thus, ortho is vicinal, meta is unsymmetrical, and para is symmetrical. Heteroatoms may be the same or different.

**CHEMICAL CLASSIFICATION CODE SEARCHING
INSTRUCTIONS**
13. General

13.1 Searching techniques for compounds having particular structural features follows much the same procedure as the application of the code to classifying compounds. One searches for the groups or structural features as coded in the chart and keeps in mind the rules and limitations that apply to each code. The most unusual structures known for the compound are usually included in a positive search so that the greatest selectivity is achieved. Conversely, structures selected to be input in the negative sense, that is, structures known not to be present, are selected to eliminate the greatest number of compounds and therefore are common structures. Negative searches for chemical classifications are generally recommended over positive searches because chemical coding for some materials is not available in the file.

**INORGANIC CHEMICAL CLASSIFICATION CODE
(COLUMNS 32 THROUGH 57)**
14. General

14.1 The large number of good infrared absorption spectra of inorganic compounds makes it desirable to have a chemical classification code for indexing such materials. Such a code has been used on punched ASTM cards indexing X-ray diffraction data for some time with considerable success. Therefore, this

code has been adapted for use in formatting the infrared data and all inorganic compounds are so coded. It should be pointed out that the coding for organic and inorganic compounds is identical except for the chemical classification sections. For making searches of chemical data it is necessary to distinguish between codes for organic and inorganic compounds. This is achieved through the “y” or 12 code for column 39. Thus, a search on 39,y isolates inorganic compounds.

14.2 The inorganic chemical classification code is organized into two sections, as is the organic compound code. Part A provides a means of identifying all elements that may be present in the compound. Part B codes inorganic radicals by name. Like the organic classification code, the inorganic code is not intended to provide unique coding for all compounds, but rather to provide a means for segregating classes of compounds.

14.3 *Part A*—Each element is given a single direct code position by assigning it a column and row position. The code list is arranged in both numerical and alphabetical order for the convenience of the user (see Table 5). Column and row are designated in the usual manner.

14.4 *Part B*—The inorganic radicals listed in Table 6 have been sufficient in kind to code all of the compounds in the ASTM catalog of X-ray diffraction powder data. The use of suffixes and prefixes to further qualify these radicals has not been attempted, so one will find that, for example, all phosphates, whether pyro-, ortho-, meta-, etc., will all be coded as 43.6.

TABLE 5 Elements Code

32-0 Actinium—Ac	34-0 Erbium—Er	36-0 Neptunium—Np	38-0 Sodium—Na
32-1 Aluminum—Al	34-1 Europium—Eu	36-1 Nickel—Ni	38-1 Strontium—Sr
32-2 Americium—Am	34-2 Fluorine—F	36-2 Nitrogen—N	38-2 Sulfur—S
32-3 Antimony—Sb	34-3 Francium—Fr	36-3 Osmium—Os	38-3 Tantalum—Ta
32-4 ^A Argon—A	34-4 Gadolinium—Gd	36-4 Oxygen—O	38-4 Technetium—Tn
32-5 Arsenic—As	34-5 Gallium—Ga	36-5 Palladium—Pd	38-5 Tellurium—Te
32-6 Astatine—At	34-6 Germanium—Ge	36-6 Phosphorus—P	38-6 Terbium—Tb
32-7 Barium—Ba	34-7 Gold—Au	36-7 Platinum—Pt	38-7 Thallium—Tl
32-8 Beryllium—Be	34-8 Hafnium—Hf	36-8 Plutonium—Pu	38-8 Thorium—Th
32-9 Bismuth—Bi	34-9 Holmium—H	36-9 Polonium—Po	38-9 Thulium—Tm
32-x Boron—B	34-x Hydrogen—H	36-x Potassium—K	38-x Tin—Sn
32-y Bromine—Br	34-y Indium—In	36-y Praseodymium—Pr	38-y Titanium—Ti
33-0 Cadmium—Cd	35-0 Iodine—I	37-0 Promethium—Pm	39-0 Tungsten—W
33-1 Calcium—Ca	35-1 Iridium—Ir	37-1 Proactinium—Pa	39-1 Uranium—U
33-2 Carbon—C	35-2 Iron—Fe	37-2 Radium—Ra	39-2 Vanadium—V
33-3 Cerium—Ce	35-3 Lanthanum—La	37-3 Rhenium—Re	39-3 Ytterbium—Yb
33-4 Cesium—Cs	35-4 Lead—Pb	37-4 Rhodium—Rh	39-4 Yttrium—Yt
33-5 Chlorine—Cl	35-5 Lithium—Li	37-5 Rubidium—Rb	39-5 Zinc—Zn
33-6 Chromium—Cr	35-6 Lutecium—Lu	37-6 Ruthenium—Ru	39-6 Zirconium—Zr
33-7 Cobalt—Co	35-7 Magnesium—Mg	37-7 Samarium—Sm	39-7
33-8 Columbium—Cb	35-8 Manganese—Mn	37-8 Scandium—Sc	39-8
33-9 Copper—Cu	35-9 Mercury—Hg	37-9 Selenium—Se	39-9
33-x Curium—Cm	35-x Molybdenum—Mo	37-x Silicon—Si	39-x
33-y Dysprosium—Dy	35-y Neodymium—Nd	37-y Silver—Ag	39-y Inorganic

^AAlso:

Helium—He
Krypton—Kr
Neon—Ne
Radon—Rn
Xenon—Xe

TABLE 6 Radicals

40-0 aluminate	41-x cyanide	43-7 phosphide	45-4 tungstate
40-1 ammonium	41-y ferrate	43-8 phosphite	45-5 uranate
40-2 antimonate		43-9 plumbate	45-6 vanadate
40-3 antimonite	42-0 ferrite	43-x plumbide	45-7 zincate
40-4 arsenate	42-1 fluoride	43-y rhenate	45-8 zirconate
40-5 arsenide	42-2 fulminate		45-9 zirconyl
40-6 arsenite	42-3 germanate	44-0 selenate	45-x platinate
40-7 bismuthate	42-4 hafniate	44-1 selenide	45-y platinite
40-8 borate	42-5 hexammine	44-2 selenite	
40-9 boride	42-6 hydride	44-3 silicate	46-0 chromite
40-x bromate	42-7 hydroxide	44-4 silicide	46-1 gallate
40-y bromide	42-8 iodate	44-5 stannate	46-2 palladite
	42-9 iodide	44-6 stannide	46-3
41-0 carbamate	42-x manganate	44-7 sulfate	46-4
41-1 carbide	42-y molybdate	44-8 sulfide	46-5
41-2 carbonate		44-9 sulfite	46-6
41-3 cerate	43-0 nitrate	44-x tantalate	46-7
41-4 chlorate	43-1 nitride	44-y telluride	46-8
41-5 chloride	43-2 nitrite		46-9
41-6 chlorite	43-3 osmate	45-0 tellurite	46-x
41-7 chromate	43-4 oxide	45-1 thionate	46-y
41-8 cyanamid	43-5 pentammine	45-2 titanate	
41-9 cyanate	43-6 phosphate	45-3 thorate	

14.4.1 A section of miscellaneous items is provided for the inorganic code and is intended to be used in the same manner as the corresponding miscellaneous codes, Column 38, of the organic code (see Table 7).

NUMBER OF ATOMS AND MELTING OR BOILING POINT CODES (COLUMNS 58 THROUGH 65)

15. Codes for Number of Atoms—Columns 58 to 62

15.1 These columns provide for recording the number of C, N, O, and S atoms in the compound being indexed. These values are coded directly into the appropriate columns (see Fig. 1). Provision for indicating a larger number of atoms than 9 in columns 60, 61, and 62 is achieved by use of positions “x” and “y.” A “y” overpunch adds 10 to the value punched into the column; an “x” adds 20; and a 0 adds 30 to the number. Thus a material containing 48 carbon atoms, 18 oxygen atoms, and 6 sulfur atoms would be coded as follows:

$$58-4; 59-8; 61-y, 8; 62-6 \quad (1)$$

NOTE 6—A “0” alone in columns 60, 61, or 62 means 30 atoms. Therefore one should never use the 0 position in any of these columns to indicate no atoms. All numbers over 38, except the number of carbon atoms, are coded at 39. Searching these data must take into consideration the double-column code. This follows the same procedure as searching for letters of the alphabet, which is discussed later. Also, the number of atoms in polymers is to be determined by the structure of the recurring unit if known. In copolymers, the recurring unit must have a 1:1 ratio of monomers regardless of the actual ratio in the polymer. If fractional atoms are indicated in monomer units, clear to smallest whole numbers. A38-7 implies that this is done where possible. Finally, double the formulas of hemisalts in arriving at the number of atoms for such compounds.

TABLE 7 Miscellaneous

56-0 Solid
56-1 Liquid
56-2 Gas
56-3 Solution
56-4 Salt plate
56-5 Hydrate
56-6 Isotopic

16. Codes for Melting or Boiling Point—Columns 63 to 65

16.1 These three columns provide for the coding of a melting or a boiling point. Melting points are indicated when the material being coded is a solid at 20°C, and boiling points at 760-mm Hg (101-kPa) pressure for all others. The numerical value of the constant is rounded to the nearest whole number and coded directly into the columns with the hundreds going into column 63, tens into column 64, and units into column 65.

16.1.1 The following code identifies the number in these columns:

16.1.1.1 “y” in column 65 indicates boiling point above 0°C,

16.1.1.2 “y” in column 64 indicates boiling point below 0°C,

16.1.1.3 “x” in column 65 indicates melting point above 0°C, and

16.1.1.4 “x” in column 64 indicates melting point below 0°C.

16.1.2 These codes aid in recognizing the nature of the constant. Normally the code of 64-x would never be used, since a material with a melting point below 0°C would be either a liquid or a gas at 20°C. However, if the only information available in such cases is the melting point, then it can be so coded. Searching operations on these data must first segregate the type of information being searched for thorough use of the x and y codes. Then the proper numerical values can be obtained. A code of 63-y is used for all corrected or reissued data.

RESERVED (COLUMNS 29, 30, AND 31 AND 66 THROUGH 70)

17. Space for Private Use—Columns 29 through 31 and 66 through 70

17.1 These columns are unassigned.

IDENTIFICATION CODES (COLUMNS 71 THROUGH 80)
18. General

18.1 Since the coding described previously is merely an indexing system which assists in locating the spectral data as it exists in published form, each document must bear a serial number or reference with which one may locate directly the original data from which it was prepared. This is provided for by columns 71 through 80. Provision is made for coding a journal reference or a serial number. It has been found to be impractical to use journal references for locating individual infrared spectrograms because of the practice of publishing a number of spectra on a single page so that the page number gives no indication of which curve is actually involved. Therefore, all spectra abstracted from journals by ASTM-sponsored groups are assigned a serial number so that each may be uniquely located and a numerical index of such serial numbers giving both the name of the compound and the journal reference is available.

18.2 Column 79 provides for indicating the particular catalog or collection of infrared spectra to which the serial number in columns 73 through 78 applies. See Table 1 for the significance of letters in this column.

18.3 Further subdivisions or special codes involving the individual collections of spectra as indicated by codes in column 79 are made possible through the use of columns 71 and 72. Thus a letter A in column 71 of the cards indexing DMS data as coded with an F in column 79, indicates that the data were obtained from inorganic compounds. For cards that have a letter C in column 79, the following letters in column 71 have the indicated significance for the Sadtler Commercial Spectra:

A— Agricultural Chemicals N— Inorganic

B— Polyols	P— Petroleum Chemicals
C— Surface-active agents	Q— Natural Resins and Gums
D— Monomers and Polymers	R— Pharmaceuticals
E— Plasticizers	S— Steroids
F— Perfumes and flavors	T— Textiles
G— Waxes and Derivatives	U— Food Additives
J— Elastomers and Rubbers	W— Attenuated Total Reflectance
K— Fibers	X— Pigments and Dyes
L— Solvents	Y— Rubber Chemicals
M— Intermediates	

Individual laboratories using the letter *B* in column 79 for their own data may make use of columns 71 and 72 in any manner. Whenever a journal reference is coded directly, there should be no code in column 79.

18.4 Column 80 provides for coding the type of coded data. The following assignments have been made:


Code	Type of Data
A—	Infrared absorption data
B—	X-ray diffraction powder data
C—	Ultraviolet absorption data
D—	Visible absorption data
E—	Mass spectral data
F—	Raman data
G—	Far-infrared data
H—	Near-infrared data

18.5 *Miscellaneous Coding Information*— The IBM code for letters involves two entries to describe each letter as follows:

A—y,1	J—x,1	S—0,2
B—y,2	K—x,2	T—0,3
C—y,3	L—x,3	U—0,4
D—y,4	M—x,4	V—0,5
E—y,5	N—x,5	W—0,6
F—y,6	O—x,6	X—0,7
G—y,7	P—x,7	Y—0,8
H—y,8	Q—x,8	Z—0,9
I—y,9	R—x,9	

REFERENCES

- (1) Barnes, R. B., Gore, R. C., Liddel, U., and Williams, V. Z., *Infrared Spectroscopy*, Reinhold Publishing Corp., New York, N.Y., 1944.
- (2) Clark, G. L., ed., *The Encyclopedia of Spectroscopy*, Reinhold Publishing Corp., New York, N.Y., 1960.
- (3) Smith, A. L., *Treatise on Analytical Chemistry*, edited by Kolthoff, I. M., and Elving, P., Part 1, Vol 6, Wiley Interscience, New York, N.Y., 1965.
- (4) Smith, A. L., *Applied Infrared Spectroscopy*, Wiley-Interscience, New York, N.Y., 1979.
- (5) Potts, W. J., Jr., *Chemical Infrared Spectroscopy*, Vol 1, Wiley-Interscience, New York, N.Y., 1962.
- (6) Bellamy, L. J., *The Infra-Red Spectra of Complex Molecules*, 3rd ed., John Wiley & Sons, New York, N.Y., 1975.
- (7) Colthup, N., Daly, L., and Wiberley, S., *Introduction to Infrared and Raman Spectroscopy*, 2nd ed., Academic Press, New York, N.Y., 1975.
- (8) Craver, C. D., ed., *The Coblenz Society Desk Book of Infrared Spectra*, 2nd Ed., 1982. Coblenz Society, P.O. Box 9952, Kirkwood, Mo. 63122, 1977.
- (9) Kuentzel, L. E., "New Codes for Hollerith-Type Punched Cards," *Analytical Chemistry*, ANCHA, Vol 23, 1951, pp. 1413–1418.
- (10) Coblenz Society Specifications, *Analytical Chemistry*, ANCHA, Vol 38, 1966, and Vol 47, 945A (1975).
- (11) *Serial Number List of Compound Names and References to Published Infrared Spectra*, ASTM AMD 32 and 32-S15, Am. Soc. Testing Mats., 1969 and 1974.
- (12) *Molecular Formula List of Compounds, Names, and References to Published Infrared Spectra*, ASTM AMD 31 and 31-S15, Am. Soc. Testing Mats., 1969 and 1974.
- (13) *Alphabetical List of Compounds, Names and References to Published Infrared Spectra*, ASTM AMD 34 and 34-S15, Am. Soc. Testing Mats., 1969 and 1974.

 **E 204 – 98 (2002)**

ASTM International takes no position respecting the validity of any patent rights asserted in connection with any item mentioned in this standard. Users of this standard are expressly advised that determination of the validity of any such patent rights, and the risk of infringement of such rights, are entirely their own responsibility.

This standard is subject to revision at any time by the responsible technical committee and must be reviewed every five years and if not revised, either reapproved or withdrawn. Your comments are invited either for revision of this standard or for additional standards and should be addressed to ASTM International Headquarters. Your comments will receive careful consideration at a meeting of the responsible technical committee, which you may attend. If you feel that your comments have not received a fair hearing you should make your views known to the ASTM Committee on Standards, at the address shown below.

This standard is copyrighted by ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959, United States. Individual reprints (single or multiple copies) of this standard may be obtained by contacting ASTM at the above address or at 610-832-9585 (phone), 610-832-9555 (fax), or service@astm.org (e-mail); or through the ASTM website (www.astm.org).