inspec

CHEMICAL AND NUMERICAL INDEXING: IMPLEMENTATION ON

STN[SM] INTERNATIONAL

June 1995

Introduction

When searching in online databases for specific numerical data and for chemicals by formulae, obtaining good results has always been difficult; at best, the search strategy is complex, and at worst, the search is infeasible. With the encouragement and assistance of a wide range of users from all over the world, INSPEC developed two special indexing systems designed to overcome many of the problems. These are:

- **Chemical Indexing** for searching for all kinds of inorganic chemical compounds and material systems.
- **Numerical Data Indexing** for limiting searches to a specific value or range of values of a particular physical quantity.

Both appear in relevant records added to the INSPEC Database since the beginning of 1987.

This guide describes how these indexing systems appear and are searchable on STN.

Further Information

This booklet is designed to be read in conjunction with INSPEC's *Numerical Data Indexing Thesaurus*; both the *Thesaurus* and this booklet contain extracts from the *INSPEC User Manual.*

These and further information on INSPEC products and services are available from:

All areas excluding the Americas

INSPEC Marketing Department
Michael Faraday House
Six Hills Way
Stevenage
Herts SG1 2AY
UNITED KINGDOM

Tel:	National	01438 313311
	International	+44 1438 313311
Fax:	National	01438 742840
	International	+44 1438 742840

E-mail: inspec@iee.org.uk

Web: http://www.iee.org.uk/publish/inspec/

North, Central & South America

INSPEC Inc.
379 Thornall Street
Edison, NJ 08837-2225
USA

Tel: +1 (732) 321 5575

Fax: (732) 321 5702

E-mail: inspec@inspecinc.com

1.0 Chemical Indexing

INSPEC's Chemical Indexing has been included in appropriate records added to the Database from the start of 1987. It is a controlled indexing system for inorganic substances and material systems. If you want to search for organic compounds (there are a substantial number of relevant references in the Database), you should use the compound name and search the supplementary terms (uncontrolled indexing). To search for inorganic substances in papers added to the Database before 1987, you should also use the supplementary terms, remembering that the formula, rather than the name of the substance, will have been indexed.

A list of chemical elements and their symbols, and a list of major chemical groups, are given in Sections 1.2 and 1.4.

The chemical indexing is designed to deal with a number of problems which arise in searching for chemical substances in the identifier field. These include:

- substances which occur in solid-state and semiconductor physics, such as gallium aluminium arsenide, may be represented in a very large number of ways, including for example:

  GaAlAs *or* AlGaAs *or* $Ga_xAl_{1-x}As$ *or* $(GaAs)_{0.5}(AlAs)_{0.5}$

  meaning that it is impracticable to enter all the search terms necessary to achieve comprehensive retrieval. The chemical indexing solves this by allowing you to search for any system containing Ga, Al and As.

- some chemical symbols and formulae have the same spellings as common English words, for example gallium phosphide (GaP - gap). This problem is solved because only chemical information appears in the Chemical Indexing.

Furthermore, you have a range of search options, because INSPEC applies Chemical Indexing to each substance at three levels. For example, you could retrieve references to sulphuric acid, H_2SO_4, by using one of a number of strategies, including:

- H_2SO_4 itself, because the complete substance is indexed as a whole
- any substance containing an SO_4 group (or alternatively, an H_2 group), because the components are indexed
- any substance containing H, S and O, because the individual elements which make up the substance are indexed

To enable you to distinguish between references to, for example, silicon (Si) as an element and as part of a more complex system, each Chemical Index Term has an associated role indicator. There are seven role indicators:

el	element	ss	system with 3 or more components
dop	dopant	int	interface system
bin	binary system	sur	surface or substrate
		ads	adsorbate or sorbate

On STN this is displayed against the label 'CHI'. The index terms for each material or system are separated by semi-colons. The index terms for a single material are separated by commas, and within each index term, the chemical symbols and associated roles are separated by spaces.

Therefore, a reference to elementary silicon would be indexed:

```
CHI Si el
```

whilst a reference to silicon dioxide, SiO_2, would be indexed:

```
CHI SiO2 bin, Si bin, O2 bin, O bin
```

If a given system, be it compound or alloy, contains two elements, then it is indexed as ‘binary’, even if one of the elements is only a very minor component such as a dopant.

Thus you have the option to search for SiO_2 either as a complete formula or by its constituents, Si and O (or O_2). This is discussed further in Section 1.1.7.

1.1 Examples of searching

1.1.1 Elements in any form

On STN the chemical indexing field is searched with the suffix ‘/CHI’ or ‘/MAI’. To retrieve all references to an element, search for that element without specifying the role.

```
=> S SI/CHI ; D TI,CHI
L1       47212 SI/CHI

L1   ANSWER 1 OF 47212
TI  Tunable integrated active filters at microwave frequencies
    using gallium arsenide MMIC technology.
CHI Si3N4 int, Si3 int, N4 int, Si int, N int, Si3N4 bin, Si3 bin,
    N4 bin, Si bin, N bin; GaAs int, As int, Ga int, GaAs bin, As
    bin, Ga bin
```

1.1.2 Elements as elements

To search for a chemical index term including its associated role, no contextual operator is required. References to an element as a simple substance may be distinguished from those to that element in a system by searching for that element with the role ‘EL’. For example:

```
=> S SI EL/CHI ; D TI,CHI
L2       28306 SI EL/CHI

L2   ANSWER 1 OF 28306
TI  Using silicon and gallium arsenide technologies for new
    supercomputer design.
CHI GaAs int, As int, Ga int, GaAs bin, As bin, Ga bin; Si int, Si
    el
```

1.1.3 Elements as dopants

To retrieve references to an element as a dopant, search for that element with the role ‘DOP’. For example:

```
=> S P DOP/CHI ; D TI, CHI
L3        1362 P DOP/CHI

L3   ANSWER 1 OF 1362
TI  Correlations between optical, electrical, and structural
    properties. of in-situ phosphorus-doped hydrogenated
    microcrystalline silicon-effects of rapid thermal annealing on
    material properties.
CHI Si:H,P ss, Si ss, H ss, P ss, Si el, H el, P el, H dop, P dop
```

1.1.4 Doped systems

To retrieve a complete doped system, search for the material with a colon separating the host and dopant. For example, for boron-doped silicon:

```
=> S SI:B BIN/CHI ; D TI, CHI
L4         904 SI:B BIN/CHI

L4   ANSWER 1 OF 904
TI  Effects of heat-treatments on electrical properties of
    boron-doped silicon crystals.
CHI Si:B bin, Si bin, B bin, Si el, B el, B dop
```

1.1.5 Elements in compounds and alloys

To retrieve references to an element as a component of a compound or alloy, search for that element with the role 'BIN' (for compounds or alloys with 2 elements) or with the role 'SS' (for compounds or alloys with 3 or more elements). For example, for simple chlorides:

```
=> S CL BIN/CHI ; D TI, CHI
L5        5263 CL BIN/CHI

L5   ANSWER 1 OF 5263
TI  Substrate temperature dependence of the structural properties
    of glow discharge produced a-Ge:H.
CHI Ge:H bin, Ge bin, H bin, Ge el, H el, H dop; Al sur, Al el; Be
    sur, Be el; NaCl sur, Cl sur, Na sur, NaCl bin, Cl bin, Na
    bin; C sur, C el; Si sur, Si el
```

Or for multi-component nickel systems:

```
=> S NI SS/CHI ; D TI, CHI
L6       10140 NI SS/CHI

L6   ANSWER 1 OF 10140
TI  Differential hysteresigraph using the transverse magneto-optic
    Kerr effect applied to thin magnetic films.
CHI YCo sur, Co sur, Y sur, YCo ss, Co ss, Y ss; LaCo sur, Co sur,
    La sur, LaCo ss, Co ss, La ss; YNi sur, Ni sur, Y sur, YNi ss,
    Ni ss, Y ss; LaNi sur, La sur, Ni sur, LaNi ss, La ss, Ni ss
```

1.1.6 Groups of compounds

To retrieve references to a multi-element group of compounds, search for that group with the role 'SS'. For example, for niobates:

```
=> S NBO3 SS/CHI ; D TI, CHI
L7         2267 NBO3 SS/CHI

L7   ANSWER 1 OF 2267
TI  A lithium niobate bandpass modulator for optical subcarrier
    transmission systems.
CHI LiNbO3 ss, NbO3 ss, Li ss, Nb ss, O3 ss, O ss
```

A list of chemical groups which are indexed as groups by INSPEC is given in Section 1.4.

1.1.7 Specific compounds or alloys

Where the chemical formula is certain and well established, you can search for the whole system. For example, for sulphuric acid:

```
=> S H2SO4/CHI ; D TI, CHI
L8          551 H2SO4/CHI

L8   ANSWER 1 OF 551
TI  The composition, structure and properties of anodic oxide films
    formed on aluminum in a 13 M sulfuric acid solution.
CHI Al sur, Al el; Al-Al2O2.29SO4H2O int, Al2O2.29SO4H2O int, O2.29
    int, Al2 int, SO4 int, Al int, H2 int, O4 int, H int, O int, S
    int, Al2O2.29SO4H2O ss, O2.29 ss, Al2 ss, SO4 ss, Al ss, H2
    ss, O4 ss, H ss, O ss, S ss, Al el; Al-Al2O2.81SO4H2O int,
    Al2O2.81SO4H2O int, O2.81 int, Al2 int, SO4 int, Al int, H2
    int, O4 int, H int, O int, S int, Al2O2.81SO4H2O ss, O2.81 ss,
    Al2 ss, SO4 ss, Al ss, H2 ss, O4 ss, H ss, O ss, S ss, Al el;
    H2SO4 ss, SO4 ss, H2 ss, O4 ss, H ss, O ss, S ss
```

In many cases, however, it will be necessary to search for several components of a system. One example of this would be retrieving all of the various formulae for gallium aluminium arsenide which were referred to in Section 1.0. Since it is not feasible to search directly for all of the possible formulae, the search must be for Ga, Al and As occurring together in the same chemical system. The three search terms would be GA SS, AL SS and AS SS. However, these terms would also appear in a reference which discussed two separate chemical systems, indium gallium arsenide and aluminium indium phosphide:

```
CHI InGaAs ss, As ss, Ga ss, In ss; AlInP ss, Al ss, In ss, P ss
```

Note that in this case there are two separate chemical index entries (subfields), each consisting of a number of terms. Therefore, to avoid retrieving such references, ensure that all your search terms occur in the same subfield (i.e. in the same chemical system) by linking them with the subfield operator, (S). Thus your search for references to gallium aluminium arsenide is:

```
=> S (GA SS(S)AL SS(S)AS SS)/CHI
```

This single strategy will retrieve all references to the various possible formulae for gallium aluminium arsenide, such as those listed in Section 1.0.

1.1.8 Interface components

To retrieve references to an interface component, search for that component with the role ‘INT’. For example, for all interfaces involving silicon or its compounds or alloys:

```
=> S SI INT/CHI ; D TI, CHI
L9       16225 SI INT/CHI

L9   ANSWER 1 OF 16225
TI  Tunable integrated active filters at microwave frequencies
    using gallium arsenide MMIC technology.
CHI Si3N4 int, Si3 int, N4 int, Si int, N int, Si3N4 bin, Si3 bin,
    N4 bin, Si bin, N bin; GaAs int, As int, Ga int, GaAs bin, As
    bin, Ga bin
```

1.1.9 Specific interfaces

To retrieve references to a specific interface, search for all the components of that interface with the role ‘INT’ and if necessary with the roles ‘EL’, ‘BIN’ and ‘SS’ as appropriate. For example, for interfaces involving elemental gold and elemental germanium:

```
=> S (AU EL(S)AU INT(S)GE EL(S)GE INT)/CHI ; D TI, CHI
          3885 AU EL/CHI
          1880 AU INT/CHI
          3505 GE EL/CHI
          1705 GE INT/CHI
L10         68 (AU EL(S)AU INT(S)GE EL(S)GE INT)/CHI

L10  ANSWER 1 OF 68
TI  Ultra-thin film superconductors of artificially synthesized
    Au/Ge layers.
CHI Au-Ge int, Au int, Ge int, Au el, Ge el
```

For some interface systems this approach would produce very long search statements. For example, for all representations of GaAlAs-InP interfaces:

```
=> S (GA SS(S)AL SS(S)AS SS(S)IN BIN(S)P BIN(S)GA INT(S)AL INT(S)
AS INT(S)IN INT(S)P INT)/CHI
```

These can be shortened by searching for the component parts and specifying an interface simply by ensuring that the indexing also contains the role ‘INT’.

```
=> S (GA SS(S)AL SS(S)AS SS(S)IN BIN(S)P BIN(S)INT)/CHI ; D TI, CHI
         15975 GA SS/CHI
         21673 AL SS/CHI
         14687 AS SS/CHI
          6461 IN BIN/CHI
          6768 P BIN/CHI
         45798 INT/CHI
L11        101 (GA SS(S)AL SS(S)AS SS(S)IN BIN(S)P BIN(S)INT)/CHI
```

```
L11  ANSWER 1 OF 101
TI  A highly dispersive wavelength division demultiplexer in
    InGaAlAs-InP for 1.5 mu m operation.
CHI InGaAlAs-InP int, InGaAlAs int, InP int, Al int, As int, Ga
    int, In int, P int, InGaAlAs ss, Al ss, As ss, Ga ss, In ss,
    InP bin, In bin, P bin
```

1.1.10 Surfaces/substrates and adsorbates

These may be searched for in a similar way to those used for interfaces with the role ‘SUR’ or ‘ADS’ replacing ‘INT’ as appropriate. For example, for GaAs surfaces:

```
=> S (GA BIN(S)AS BIN(S)GA SUR)/CHI ; D TI, CHI
         20944 GA BIN/CHI
         21023 AS BIN/CHI
          3444 GA SUR/CHI
L12       2930 (GA BIN(S)AS BIN(S)GA SUR)/CHI

L12  ANSWER 1 OF 2930
TI  MBE-growth of InAs and GaSb epitaxial layers on GaAs
    substrates.
CHI InAs bin, As bin, In bin; GaSb bin, Ga bin, Sb bin; GaAs sur,
    As sur, Ga sur, GaAs bin, As bin, Ga bin
```

For example, for hydrogen as an adsorbate:

```
=> S H ADS/CHI ; D TI, CHI
L13       1471 H ADS/CHI

L13  ANSWER 1 OF 1471
TI  Coadsorption of Cs and hydrogen on W(110) studied by metastable
    impact electron spectroscopy.
CHI H ads, H el ; Cs ads, Cs el
```

1.1.11 Number of elements in a material system (/ELC)

STN's element count feature allows you to restrict a chemical substance search according to the number of unique components in the complete compound or system (i.e. in the subfield). It does not represent a count of the number of atoms. The element count may be searched for with the field suffix ‘/ELC’, normally in conjunction with other chemical indexing search statements. Since one record may contain indexing for several substances, use the link operator, (L), to ensure that the element count and chemical indexing terms refer to the same substance. For example, for quaternary GaAlAsP compounds:

```
=> S (GA SS(S)AL SS(S)AS SS(S)P SS)/CHI(L)4/ELC ; D TI, CHI
         16055 GA SS/CHI
         21819 AL SS/CHI
         14736 AS SS/CHI
          7999 P SS/CHI
         38573 4/ELC
L14         14 (GA SS(S)AL SS(S)AS SS(S)P SS)/CHI(L)4/ELC
```

```
L14  ANSWER 1 OF 14
TI   Alloy disorder effects on the electronic properties of III-V
     quaternary semiconductor alloys.
CHI  AlGaAsP ss, Al ss, As ss, Ga ss, P ss; AlGaAsSb ss, Al ss, As
     ss, Ga ss, Sb ss; InGaAsP ss, As ss, Ga ss, In ss, P ss;
     InGaAsSb ss, As ss, Ga ss, In ss, Sb ss
```

This retrieves only the quaternary systems. However, a search without the /ELC term retrieves unwanted records of systems with five or more compounds:

```
=> S (GA SS(S)AL SS(S)AS SS(S)P SS)/CHI ; D TI, CHI
         15975 GA SS/CHI
         21673 AL SS/CHI
         14687 AS SS/CHI
          7949 P SS/CHI
L15         44 (GA SS(S)AL SS(S)AS SS(S)P SS)/CHI

L15  ANSWER 1 OF 44
TI   Thermodynamic analysis of the MOVPE and MBE growth of pentanary
     III-V alloy semiconductors.
CHI  AlGaInAsP ss, Al ss, As ss, Ga ss, In ss, P ss
```

1.1.12 Periodic group designation of elements (/PG)

In loading the INSPEC file, STN has added the possibility of searching for any element from a given periodic group. This is done by using the notation shown in Tables 1.2 and 1.3 with the field suffix '/PG' and the (L) operator, often in conjunction with other chemical indexing search statements. For example, for III-V binary systems:

```
=> S BIN/CHI(L)A3/PG(L)A5/PG ; D TI, CHI
        118199 BIN/CHI
         67741 A3/PG
         61710 A5/PG
L16      25862 BIN/CHI(L)A3/PG(L)A5/PG

L16  ANSWER 1 OF 25862
TI   Tunable integrated active filters at microwave frequencies
     using gallium arsenide MMIC technology.
CHI  Si3N4 int, Si3 int, N4 int, Si int, N int, Si3N4 bin, Si3 bin,
     N4 bin, Si bin, N bin; GaAs int, As int, Ga int, GaAs bin, As
     bin, Ga bin
```

1.2 Table of chemical elements with symbols and periodic groups

Element	Sy	Group	Element	Sy	Group	Element	Sy	Group
actinium	Ac	B3/ACTN	hafnium	Hf	B4	praseodymium	Pr	LNTH
aluminium	Al	A3	hahnium	Ha	B5	promethium	Pm	LNTH
americium	Am	ACTN	helium	He	A8	protactinium	Pa	ACTN
antimony	Sb	A5	holmium	Ho	LNTH	radium	Ra	A2
argon	Ar	A8	hydrogen	H	A1	radon	Rn	A8
arsenic	As	A5	*see also*			rhenium	Re	B7
astatine	At	A7	*deuterium,*			rhodium	Rh	B8
barium	Ba	A2	*tritium*			rubidium	Rb	A1
berkelium	Bk	ACTN	indium	In	A3	ruthenium	Ru	B8
beryllium	Be	A2	iodine	I	A7	rutherfordium	Rf	B4
bismuth	Bi	A5	iridium	Ir	B8	samarium	Sm	LNTH
boron	B	A3	iron	Fe	B8	scandium	Sc	B3
bromine	Br	A7	krypton	Kr	A8	selenium	Se	A6
cadmium	Cd	B2	lanthanum	La	B3/LNTH	silicon	Si	A4
caesium	Cs	A1	lawrencium	Lr	ACTN	silver	Ag	B1
calcium	Ca	A2	lead	Pb	A4	sodium	Na	A1
californium	Cf	ACTN	lithium	Li	A1	strontium	Sr	A2
carbon	C	A4	lutetium	Lu	LNTH	sulphur	S	A6
cerium	Ce	LNTH	magnesium	Mg	A2	tantalum	Ta	B5
chlorine	Cl	A7	manganese	Mn	B7	technetium	Tc	B7
chromium	Cr	B6	mendelevium	Md	ACTN	tellurium	Te	A6
cobalt	Co	B8	mercury	Hg	B2	terbium	Tb	LNTH
copper	Cu	B1	molybdenum	Mo	B6	thallium	Tl	A3
curium	Cm	ACTN	neodymium	Nd	LNTH	thorium	Th	ACTN
deuterium	D	A1	neon	Ne	A8	thulium	Tm	LNTH
see also			neptunium	Np	ACTN	tin	Sn	A4
hydrogen			nickel	Ni	B8	titanium	Ti	B4
dysprosium	Dy	LNTH	niobium	Nb	B5	tritium	T	A1
einsteinium	Es	ACTN	nitrogen	N	A5	*see also*		
erbium	Er	LNTH	nobelium	No	ACTN	*hydrogen*		
europium	Eu	LNTH	osmium	Os	B8	tungsten	W	B6
fermium	Fm	ACTN	oxygen	O	A6	uranium	U	ACTN
fluorine	F	A7	palladium	Pd	B8	vanadium	V	B5
francium	Fr	A1	phosphorus	P	A5	xenon	Xe	A8
gadolinium	Gd	LNTH	platinum	Pt	B8	ytterbium	Yb	LNTH
gallium	Ga	A3	plutonium	Pu	ACTN	yttrium	Y	B3
germanium	Ge	A4	polonium	Po	A6	zinc	Zn	B2
gold	Au	B1	potassium	K	A1	zirconium	Zr	B4

1.3 Periodic Table of the Elements showing atomic no. and group

Group A1	A2	B3	B4	B5	B6	B7	B8	B8	B8	B1	B2	A3	A4	A5	A6	A7	Group A8
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57 * La	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89 ** Ac	104 Rf	105 Ha													

* Other Lanthanides	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
** Other Actinides	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr

Additional grouping designations:

LNTH	Lanthanides, i.e. elements 57 to 71
ACTN	Actinides, i.e. elements 89 to 103
T1	First-row transition metals, i.e. elements 21 to 29
T2	Second-row transition metals, i.e. elements 39 to 47
T3	Third-row transition metals, i.e. elements 57 and 72 to 79

1.4 Table of chemical groups and their formulae

Chemical Group	**Formula**	**Chemical Group**	**Formula**
alumina	Al2O3	iodate	IO3
aluminium garnet	Al5O12	iron garnet	Fe5O12
arsenate	AsO4	*see also ferrite*	
arsenate	As2O3	magnesate	MgO3
bismuthate	Bi2O3	manganate	MnO4
borate	BO4	molybdate	MoO4
borate	B2O3	niobate	NbO3
borate	B3O6	niobate	Nb2O5
borate (per-)	BO3	niobate	Nb2O7
bromate	BrO3	nitrate	NO3
carbonate	CO3	nitrite	NO2
carbonyl	CO	phosphate (ortho-)	PO4
chlorate	ClO3	phosphate	P2O5
chromate	Cr2O3	phosphate (pyro-)	P2O7
chromate	CrO4	phosphite	PO3
chromate (di-)	Cr2O7	phosphite	P4O12
chromate (per-)	CrO3	selenate	SeO4
cyanide	CN	selenite	SeO3
ferrite	Fe2O3	silica	SiO2
ferrite	Fe2O4	silicate	SiO4
ferrite	Fe3O4	sulphate	SO4
ferrite (ortho-)	FeO3	sulphate (thio-)	S2O3
see also iron garnet		sulphite	SO3
gallium garnet	Ga5O12	tantalate	TaO3
garnet (aluminium)	Al5O12	titanate	TiO3
garnet (gallium)	Ga5O12	tungstate	WO3
garnet (iron)	Fe5O12	tungstate	WO4
germanate	GeO2	vanadate	V2O5
germanate	GeO3	vanadate (meta-)	VO3
hydroxyl	OH	vanadate (ortho-)	VO4
hydroxyl (deuterated)	OD	zirconate	ZrO3

2.0 Numerical Data Indexing

INSPEC has applied Numerical Data Indexing to relevant documents added to the Database from the beginning of 1987. Indexed data is selected by INSPEC's subject specialists as being important in the particular document and useful for retrieval. This data is structured in a standard format. INSPEC designed and introduced the Numerical Data Indexing in order to cope with specific problems which arose in searching for this type of information by previously-available methods.

As an example of these problems, consider searching for all references to power stations generating between 20 and 30 MW. Papers would be likely to be of interest whether they discussed a single value in the range (e.g. 23 MW) or a range of values, such as 25 to 35 MW, which might significantly, but not necessarily wholly, overlap the specified range. Furthermore, 1 MW is equivalent to 1000 kW or 1000000 W. Finally, the abbreviation MW may be written more fully as MWatt or megawatt.

Therefore, there is an unlimited number of ways in which authors might specify values which should be retrieved by the search, including for example:

29.2 MWatt	27500 KW
25 to 30 megawatts	20 to 35 MW
27000000 Watts	2.1×10^7 W

The situation is further complicated by the fact that some quantities may be measured in any of a number of units. For example, temperature may be measured in degrees Centigrade or degrees Fahrenheit as alternatives to the SI unit, Kelvin. Therefore, in addition to the above considerations, in searching the basic index for references mentioning 100 degrees Centigrade, you would also have to include the equivalent values 373 Kelvin and 212 degrees Fahrenheit as alternative search statements. These variations in the format of numerical values make it impossible to achieve comprehensive retrieval of numerical information by searching the basic index.

To solve these problems, INSPEC introduced a standardised indexing format for numerical data. On STN, numerical data is displayed against the label PHP (physical properties), for example:

PHP frequency 5.0E+01 Hz	(for a frequency of 50 Hz)
PHP wavelength 1.06E-05 m	(for a wavelength of 0.0000106 metres)
PHP temperature 2.73E+02 to 3.73E+02 K	(for 0^o to 100^o Centigrade)

More generally, each Index Term takes the following format:

Quantity Value [to Value] Unit

where:

'**Quantity**' is a physical quantity, e.g. 'frequency' or 'temperature', and has only one unit.

'**Value**' is usually expressed in floating point format, e.g. 1.8E+04 for 18000, 9.5E-01 for 0.95. Floating point format is explained in Section 2.7.

'**Unit**' is usually the SI unit, e.g. 'Hertz', 'Kelvin'. A unit can apply to more than one quantity, e.g. the unit 'metre' applies to the quantities 'depth', 'distance' and 'size'. However, each quantity is always measured in the same unit, e.g. 'Kelvin' for 'temperature'.

The Quantities and Units are given in the *Numerical Data Indexing Thesaurus*, which contains approximately 50 preferred quantities, and is reviewed each year and revised as necessary. STN has added the numerical thesaurus to the online file, in the /PHP index (see Section 2.4).

You can make the best use of numerical data indexing by using it to refine a search already carried out on other criteria, e.g. communications at over 9600 bit/s, or atmospheric effects occurring at an altitude of between 100 and 150 km. Numerical data index terms are only applied when the data appear in the title or abstract of the paper and when they are likely to be of use for retrieval. This will most probably be when they relate to:

> Relevant and essential operating characteristics or features of instruments, systems and devices, e.g. frequency range for a signal generator.
>
> Relevant and important criteria relating to effects, phenomena and processes. This is much more likely to be experimental or operating conditions than measured values or observations. For example, if a reference describes an experiment to measure variations in pressure with altitude, the altitude range would be more likely to be indexed than the pressures.

2.1 Numerical searching using the thesaurus

The first step in a numerical search is to look in the numerical data indexing thesaurus for the required physical quantity. If an entry is found it will either be a preferred quantity or point to a preferred quantity. The preferred quantity is the one used for searching. Either the printed or the online numerical data indexing thesaurus may be used for this purpose.

The next step is to determine the preferred unit for the quantity. This will be the unit in which the data appear in the database. If the quantity being searched for is expressed in units other than the preferred unit, then a conversion may be required. If this is the case, then the conversion formula is likely to be given in the thesaurus.

Sometimes the quantity being searched for is expressed in units listed in the thesaurus but with the addition of a standard prefix or its abbreviation such as k (for ‘kilo’ meaning times 1,000). In these cases the value or values should be converted into the base unit by applying the multiplication factor prior to using the data either directly in a search or in a conversion specified in the thesaurus. The possible prefixes and their associated multiplying factors are given in the thesaurus.

For example:

a. A search is required for a power level of greater than 500MW (perhaps a power station).

The thesaurus is searched for ‘power’ and the following preferred entry is found:

> **power : W** (watts)

Thus the preferred unit is watts, but the search data is in megawatts. The thesaurus gives a conversion factor of 1,000,000 and so the 500MW is converted to 500,000,000 watts and the search is then for power greater than 5.0E+08 W.

b. A search is required for documents containing a numerical data index entry of 100 kilocalories of heat.

The thesaurus is searched for ‘heat’ and the following entry found:

> heat *use* energy

The preferred entry for ‘energy’ is:

> **energy : J** (Joules)

Thus the preferred unit is Joules and the search data are in kilocalories. One of the energy entries is:

cal *use* J [J=cal × 4.1868]

There is no record containing kilocalories. Thus the first conversion required is from kilocalories to calories. From Table 1in the thesaurus, the factor for this is 1,000 and so the searched value becomes 100,000 calories. The data is then converted to Joules using the conversion given in the thesaurus and becomes 418,680 Joules. This data would appear in the numerical data indexing field as:

energy 4.1868E+05 J

It is often advisable to search for a range of values to avoid missing data with a different level of precision and to avoid any problems of rounding errors. Thus, here, the search might be for an energy of 4.18E+05 to 4.19E+05 J or for a wider range depending on the particular requirements.

2.2 Numerical searching on STN

On STN numerical values are searchable in quantity-specific fields (e.g. '/FRE' for frequency or '/W' for wavelength). The field suffices are listed in Section 2.6; alternatively, they may be displayed by using the online numerical data indexing thesaurus (see Section 2.4).

Examples of various types of search are given below. However, the following specific points illustrated in the examples may be noted:

Although all values are displayed in a standardised exponential form, they can be entered in a variety of formats, e.g. any of the following are valid for entering 8.0E-02: 8.0E-02, 8E-2, 0.08 or .08. The '+' sign need not be used for specifying positive exponents. However, if it is used, then the string containing it must be contained within quotes. e.g. '1.5E+2' for 150.

STN can handle values in the range 5.4E-79 to 7.2E+75. Values outside this range are converted to 5.4E-79 or 7.2E+75 as appropriate for searching. It is suggested that, for all practical purposes, users regard the operating range as 1.0E-70 to 9.9E+70 as this is easier to remember.

2.3 Examples of searching

2.3.1 Exact value searching

Exact values can be retrieved by searching for the value in any format followed by the quantity specific field suffix. For example, consider a search for infra-red detectors for 0.0000106 m wavelength. Note that, since the wavelength specified is in the infra-red range, it is not necessary to search for the term infra-red.

```
=> S DETECTOR# AND 1.06E-5/W ; D TI, PHP
         66598 DETECTOR?
          1954 1.06E-5 M /W
L1         212 DETECTOR# AND 1.06E-5 M /W

L1   ANSWER 1 OF 212
TI  (Hg,Zn)Te photon detectors of 10.6 mu m laser radiation with
    optical resonance cavity.
PHP wavelength 1.06E-05 m
```

2.3.2 Range searching

Ranges of numerical values can be searched by specifying the values for the start and end of the range separated by ‘-’. These are then followed by the quantity specific field suffix. For example, to search for thin film properties at 0.00000106 to 0.00000107 m wavelength:

```
=> S THIN FILM# AND 1.06E-6-1.07E-6/W ; D TI, PHP
        126416 THIN
        148394 FILM#
         78647 THIN FILM#
                 (THIN(W)FILM#)
          2087 1.06E-6 M - 1.07E-6 M /W
L2          13 THIN FILM# AND 1.06E-6-1.07E-6/W

L2   ANSWER 1 OF 136
TI  Near IR optical properties of sputtered InN films.
PHP wavelength 7.0E-07 to 2.7E-06 m
```

Overflows may occur if the range searched for is wide. If this happens, search the range as several narrower ranges with an OR operator.

2.3.3 ‘Greater than’ searching

Searching for quantities greater than a particular value may be achieved by searching for the quantity specific field abbreviation followed by the ‘greater than’ sign, i.e. ‘>’ and then the particular value. ‘Greater than or equal to may be achieved by using ‘>=’ instead of ‘>’.

e.g. Search for all aspects of turbogenerators operating at or above 1GW (1,000,000,000 W).

```
=> S TURBOGENERATOR# AND POW>=1E9 ; D TI, PHP
          4512 TURBOGENERATOR#
           715 POW>=1E9 W
L3          25 L5 AND L6

L3   ANSWER 1 OF 25
TI  Generator loss of field: experience and studies for AEP’s
    Rockport Plant.
PHP power 1.3E+09 W
```

2.3.4 ‘Less than’ searching

Searching for quantities less than a particular value may be achieved by searching for the quantity specific field abbreviation followed by the ‘less than’ sign, i.e. ‘<’ and then the particular value. ‘Less than or equal to’ may be achieved by using ‘<=’ instead of ‘<’.

c.f. ‘Greater than’ searching.

2.4 Online numerical data indexing thesaurus

In addition, the physical properties thesaurus (the numerical data indexing thesaurus) is available online for searching. This consists of the fields listed in Section 2.6 and lead-ins. The fields can be identified in the normal way with the /PHP delimiter. For a full list type EXPAND A/PHP and continue with more E commands.

As an example, consider a search for a range of altitudes. First EXPAND to see the list of numerical quantities, and then expand again to see the thesaurus:

```
=> E ALTITUDE/PHP
E#   FREQUENCY    AT     TERM
—    ————-    —     ——
E1           0     2     A/PHP
E2        2783     2     AGE/PHP
E3        2411     1 —> ALTITUDE/PHP
E4           0     2     AMP/PHP
E5         674     2     APPARENT POWER/PHP
E6           0     3     ASTRONOMICAL UNIT/PHP
E7           0     2     ATM/PHP
E8           0     2     ATMOSPHERE/PHP
E9           0     3     AU/PHP
E10       1730     1     BANDWIDTH/PHP
E11          0     2     BAR/PHP
E12          0     2     BECQUEREL/PHP

=> E E3+ALL/PHP
E1        2411   —>  ALTITUDE/PHP
                  FQS  /ALT
                  INSPEC UNIT M (METRE)
                  CGS UNIT CM
                  ENG UNIT FT
                  FPS UNIT FT
                  MKS UNIT M
                  SI UNIT M
                  STN UNIT M
                  OTHER UNITS MILE, YARD
                  DEF  MEASURED FROM SURFACE (LIQUID OR SOLID) FOR
                       EARTH AND ALL PLANETARY BODIES. MEASURED
                       FROM PHOTOSPHERE FOR SUN.
*********  END  *********
```

Alternatively, the thesaurus display may be reached in one step with the command:

```
=> E ALTITUDE+ALL/PHP
```

STN allows you to search in any of a number of unit systems. For the following search, the foot-pounds-seconds (fps) system would be the easiest, so set all quantities to fps units before searching for the required altitude range in feet. Thus, to find out about winds at 20-25,000 feet:

```
=> SET UNIT ALL=FPS

SET COMMAND COMPLETED

=> S 2E4-2.5E4/ALT AND WIND#
           351 2E4 FT - 2.5E4 FT /ALT
         35066 WIND#
L4          80 2E4 FT - 2.5E4 FT /ALT AND WIND#

=> D TI,PHP

L4   ANSWER 1 OF 80
TI  Enhanced frequency spectra of winds at the mesoscale based on
    radar profiler observations.
PHP time 6.0E+02 to 8.6E+04 s; time 1.4E+04 to 6.5E+04 s; altitude
    5.8E+03 to 8.8E+03 m; altitude 1.25E+04 to 1.9E+04 m
```

For more information, type HELP SET UNITS at the prompt sign.

2.5 Tolerance

With the tolerance specification feature, you may cause a single-value numeric search specification to be expanded into a numerical range search.

The 'default tolerance' for every numerical field is zero. This means that a single-value numerical search specification will not be expanded into a numerical range search.

You may specify the tolerance for a numerical field by entering 'SET TOLERANCE' at an arrow prompt (=>). For information, enter 'HELP TOLERANCE' and 'HELP SET TOLERANCE' at arrow prompts.

More than one field code may be specified at a time, but all field codes must be valid for the current file. Note that an equals sign (=) is required with no intervening spaces. Note also that tolerances may be specified either as an absolute value or as a percentage (%), and may be up to 15 digits.

For example, the following command sets the tolerance for altitude to 5% and that for time to 10 seconds:

```
=> SET TOLERANCE ALT=5% TIM=10
```

2.6 Quantities and suffices on STN

Property	Unit	Search Code
Age	yr (Year)	/AGE
Altitude	m (Metre)	/ALT
Bandwidth	Hz (Hertz)	/BAW
Bit Rate	bit/s (Bit per Second)	/BIR
Byte Rate	Byte/s (Byte per Second)	/BYR
Capacitance	F (Farad)	/CAP
Computer Execution Rate	IPS (Instruct. per Second)	/COE
Conductance	S (Siemens)	/CON
Computer Speed	FLOPS (Floating Point Operation per Second)	/COS
Current	A (Amp)	/CUR
Depth	m (Metre)	/DEP
Distance	m (Metre)	/DIS
Electric Conductivity	S/m (Siemens per Metre)	/ECND
Electron Volt Energy	eV (Electron Volt)	/EEV
Efficiency	percent	/EFF
Energy	J (Joule)	/ENE
Electrical Resistivity	ohmm (Ohm Metre)	/EREST (or /REE)
Frequency	Hz (Hertz)	/FRE
Galactic Distance	pc (Parsec)	/GAD
Gain	dB (Decibel)	/GAI
Geocentric Distance	m (Metre)	/GED
Heliocentric Distance	AU (Astronomical Unit)	/HED
Loss	dB (Decibel)	/LOS
Mass	kg (Kilogram)	/M
Memory Size	Byte	/MES
Magnetic Flux Density	T (Tesla)	/MFD (or /B)
Noise Figure	dB (Decibel)	/NOF
Picture Size	pixel (Picture Element)	/PIS
Apparent Power	VA (Volt-Amps)	/POA
Reactive Power	VAr (Volt-Amp (reactive))	/POR
Power	W (Watt)	/POW
Pressure	Pa (Pascal)	/PRES (or /P)
Printer Speed	cps (Character per Second)	/PRSP (or /PRS)
Radioactivity	Bq (Becquerel)	/RAD
Radiation Absorbed Dose	Gy (Gray)	/RADA
Radiation Dose Equivalent	Sv (Sievert)	/RADE
Radiation Exposure	C/kg (Coulomb per Kilogram)	/RAE
Resistance	ohm	/RES
Storage Capacity	bit	/SCA
Size	m (Metre)	/SIZ
Stellar Mass	Msol (Solar Mass)	/STM
Temperature	K (Kelvin)	/TEMP (or /T)
Time	s (Second)	/TIM
Velocity	m/s (Metre per Second)	/VEL (or /V)
Voltage	V (Volt)	/VOLT
Word Length	bit	/WOL
Wavelength	m (Metre)	/WVL (or /W)

2.7 Floating point format

The practice of expressing numbers in floating point format avoids the problems of managing large numbers of zeros in such numbers as 1952000000000 or 0.000000000753. To convert a number into floating point format, use the following guidelines.

Consider, for example, 100. This is $1\times 10 \times 10$, which is 1×10^2, or 1E2; this is the same as 1.0E+02.

Similarly, 150 is 1.5×10^2, which is 1.5E2, or equivalently 1.5E+02.

Conversely, 0.015 is $1.5 \div 10^2$, which is 1.5E-2, or equivalently 1.5E-02.

It follows that the initial examples above are expressed as the floating point numbers 1.952E+12 and 7.53E-10 respectively.

Numbers are always expressed with a single digit to the left of the decimal point. When you display a record, you will see that INSPEC also uses at least one figure to the right of the point, as well as plus signs and two-digit exponents (e.g. 1.0E+02 rather than 1E2). However, for searching on STN, as well as in general use, the shorter form is also accepted. In addition, STN accepts conventional numbers (e.g. 100) as an alternative to the floating point format.