Astronomy & Astrophysics manuscript no. 3818
(DOI: will be inserted by hand later)
October 7, 2005

Representations of spectral coordinates in FITS

E. W. Greisen[1], M. R. Calabretta[2], F. G. Valdes[3], and S. L. Allen[4]

[1] National Radio Astronomy Observatory, PO Box O, Socorro, NM 87801-0387, USA
[2] Australia Telescope National Facility, PO Box 76, Epping, NSW 1710, Australia
[3] National Optical Astronomy Observatories, PO Box 26732, Tucson, AZ 85719, USA
[4] UCO/Lick Observatory, University of California, Santa Cruz, CA 95064, USA

Revised last 30 September 2005, 05 May 2005, / Received / Accepted

Abstract. Greisen & Calabretta (2002) describe a generalized method for specifying the coordinates of FITS data samples. Following that general method, Calabretta & Greisen (2002) describe detailed conventions for defining celestial coordinates as they are projected onto a two-dimensional plane. The present paper extends the discussion to the spectral coordinates of wavelength, frequency, and velocity. World coordinate functions are defined for spectral axes sampled linearly in wavelength, frequency, or velocity, linearly in the logarithm of wavelength or frequency, as projected by ideal dispersing elements, and as specified by a lookup table.

Key words. Methods: data analysis – techniques: image processing – techniques: radial velocities – techniques: spectroscopic – astronomical data bases: miscellaneous

1. Introduction

The present paper is the third in a series of papers that establishes methods by which information about the physical coordinates of FITS data may be transferred along with the binary image, random groups, and table data. In "Paper I" Greisen & Calabretta (2002) describe a revised method for transferring coordinate information in the FITS header and outline some rules governing the values assigned to the new standard header keywords. In "Paper II" Calabretta & Greisen (2002) specify the conventions necessary to define celestial coordinates in a two-dimensional projection of the sky. This paper defines the parameters and conventions needed to specify spectral information including frequency, wavelength, and velocity. In addition to these basic conversions, a world coordinate function is defined to describe ideal optical dispersers of several common types. Several concepts that were suggested by spectroscopic issues are generalized to apply to all types of coordinate axes. These are a generalized description of non-linear algorithms, the `-LOG` and `-TAB` non-linear algorithms, and an axis-naming keyword. In "Paper IV", Calabretta et al. (2005) specify methods to describe the distortions inherent in the image coordinate systems of real astronomical data.

Paper I describes the computation of the world or physical coordinates as a multi-step process. The vector of pixel offsets from the reference point is multiplied by a linear transformation matrix and then scaled to physical units. Mathematically, this is given by

$$x_i = s_i \; q_i = s_i \sum_{j=1}^{N} m_{ij} \; \left(p_j - r_j\right), \tag{1}$$

where p_j are pixel coordinates, r_j are pixel coordinates of the reference point given by `CRPIX`j, m_{ij} is a linear transformation matrix given either by `PC`i_j or `CD`i_j, N is the dimensionality of the WCS representation given by `NAXIS` or `WCSAXES`, and s_i is a scaling given either by `CDELT`i or by 1.0.

The final step in the computation is the conversion of these linear relative coordinates into the actual physical coordinates. The conventions to be applied to spectral axes, i.e. to frequency, wavelength, and velocity axes, are described in this paper.

2. Coordinate definition

At this stage in the discussion, we consider the spectral world coordinate at the reference point on all other axes. This covers the relatively simple case in which the spectral world coordinate is not dependent on the other world coordinates. Methods to describe deviations from this assumption due to the choice of spectral reference system are discussed in Sect. 7, while instrumental distortions are discussed briefly in Sect. 9 and, in more detail, in Paper IV.

Spectral coordinates are commonly given in units of frequency, wavelength, velocity, and other parameters proportional to these three. The coordinate types discussed here are then frequency, wavelength, and "apparent radial velocity" de-

Send offprint requests to: E. W. Greisen,
e-mail: `egreisen@nrao.edu`

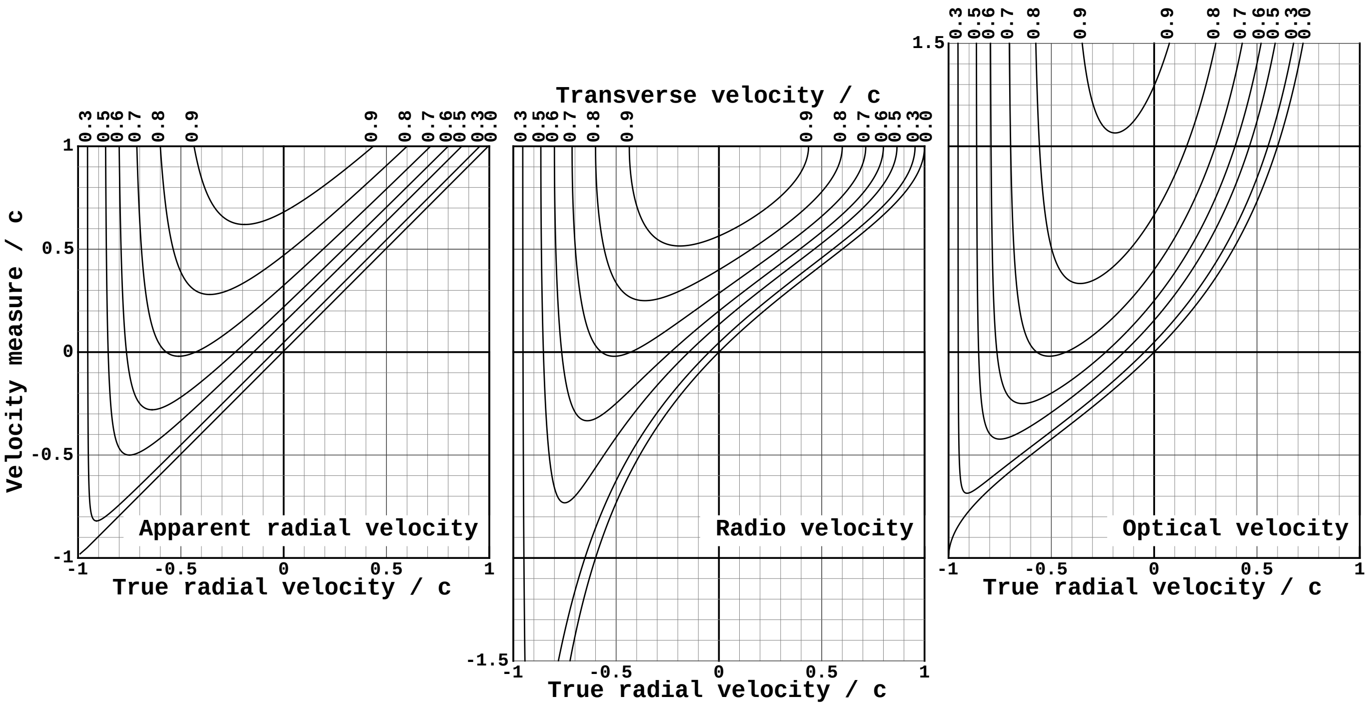


Fig. 1. Conventional velocities as a function of true radial velocity for selected values of the transverse velocity. The apparent radial velocity v/c is plotted in the left panel, the radio velocity V/c is plotted in the center panel, and the optical velocity Z/c is plotted in the right panel. Note that, at a transverse velocity of $0.9c$, the red shift z always exceeds 1. Note that each family of curves intersects the axis of zero velocity measure at the same values of v_t because the redshift is zero for these combinations of v_t and v_r. Thus, regardless of the value of v_t, for any given redshift the velocity measures agree on whether the object *appears* to be receding or approaching. However, the object may appear to be receding at high speed even when it is actually approaching at high speed, though the converse is never true. At transverse velocities of $v_t/c > \sqrt{2}/2(\approx 0.71)$ all of the velocity measures are positive (receding) regardless of whether the object is actually receding or approaching. Furthermore, for v_r/c below -0.5, the faster the object approaches, the smaller the transverse velocity required to make it appear to be receding!

noted by the symbols ν, λ, and v. There are also three conventional velocities frequently used in astronomy. These are the so-called "radio" velocity, "optical" velocity, and redshift, denoted here by V, Z, and z and given by $V = c(\nu_0 - \nu)/\nu_0$, $Z = c(\lambda - \lambda_0)/\lambda_0$ and $z = Z/c$. The velocities are defined so that an object receding from the observer has a positive velocity. We assume throughout that each image has at most one spectral coordinate axis; a similar limit on celestial coordinates was assumed in Paper II.

As discussed by Lindegren and Dravins (2003), the apparent radial velocity to be described here is itself a conventional velocity. The shift of a spectral line from its rest frequency is caused by a variety of effects, not just the Doppler shift due to a radial motion. The time dilation of an object moving with respect to the observer causes a spectral shift, even if that motion is entirely transverse. Gravitational fields at the radiating object and along the line of sight to the observer will shift the observed frequency. Furthermore, the apparent spectroscopic velocity may be produced at a peculiar point in the object, e.g. upwelling convective cells, rather than at the center of mass of the object. It may also be affected by absorption in intervening material and in particular by the cosmic redshift. We will use the term "apparent radial velocity" here simply to refer to that pseudo-physical velocity described by the equations presented here. The apparent radial velocity then is a sum of all spectroscopic effects presented as if they were solely a radial velocity. While this may be sufficiently accurate for many uses, observers wishing to achieve very high accuracies should consult Lindegren and Dravins, and references therein, closely.

The effect of a transverse velocity deserves a little more discussion. As given by Lang (1974),

$$\lambda = \lambda_0 \frac{c + v_r}{\sqrt{c^2 - v_r^2 - v_t^2}}, \tag{2}$$

where v_r is the true radial velocity, v_t is the true transverse velocity, c is the speed of light, and miscellaneous effects such as gravitational redshifts are assumed negligible. Even at ordinary astronomical velocities, this has a measurable effect; a transverse velocity of 100 km s^{-1} produces an apparent velocity of 15 m s^{-1}. The effects of transverse velocities such as might be found in astronomical jet sources on the apparent radial, radio, and optical conventional velocities are illustrated in Fig. 1. When the transverse velocity is zero, radio velocity ranges from $-\infty$ to $+c$, optical velocity from $-c$ to $+\infty$, and apparent radial velocity is v_r when $v_t = 0$. Because of all of the other processes that shift the observed frequency, it is inappropriate to propose keywords to describe true velocities at this time. Additional discussion of this matter is deferred to Appendix A.

For the purposes of this section and the next, we consider those spectral axes at a single celestial coordinate that are

Table 1. Spectral coordinate type codes, (characters 1–4 of `CTYPE`*ka*).

Code	Name	Symbol	Associate variable	Default units
`FREQ`	Frequency	ν	ν	Hz
`ENER`	Energy	E	ν	J
`WAVN`	Wavenumber	κ	ν	m^{-1}
`VRAD`	Radio velocity	V	ν	$\mathrm{m\ s}^{-1}$
`WAVE`	Vacuum wavelength	λ	λ	m
`VOPT`	Optical velocity	Z	λ	$\mathrm{m\ s}^{-1}$
`ZOPT`	Redshift	z	λ	–
`AWAV`	Air wavelength	λ_a	λ_a	m
`VELO`	Apparent radial velocity	v	v	$\mathrm{m\ s}^{-1}$
`BETA`	Beta factor (v/c)	β	v	–

linearly sampled in wavelength, frequency, or apparent radial velocity. The radio and optical "velocities" are directly proportional to frequency and wavelength, respectively, and thus do not constitute additional cases for the type of sampling. Frequency and wavelength axes may also be regularly sampled in their logarithm. Wavelengths are sometimes given "in air" rather than in vacuum and denoted here by λ_a. This non-linear distinction is discussed in Sect. 4. Frequency may also be described in energy (= $h\nu$) units and in Kaysers ("wave number" = $1/\lambda$) units.

Following the convention of Papers I and II, the first four characters of `CTYPE`*ka* specify[1] the coordinate type, the fifth character is `'-'` and the next three characters specify a predefined algorithm for computing the world coordinates from intermediate physical coordinates. When k is the spectral axis, the first four characters shall be one of the codes shown in Table 1. The mathematical symbols we will use for the codes and their default units are also listed. Units are specified with the `CUNIT`*ka* keyword. Standard values for unit strings are defined in Paper I. The IAU-standard prefixes for scaling the units are described in Paper I and should be used with all coordinate types, except that the dimensionless ones are not scaled.

The non-linear algorithm in use is specified by the final 3 characters of `CTYPE`*ka*. In spectral codes, the first of the three characters specifies the physical parameter type in which the data are regularly sampled, the second character is `2`, and the third character specifies the physical parameter type in which the coordinate is expressed. Non-linear algorithm codes, the last four to be introduced later, are summarized in Table 2. When the algorithm in use is linear, the last four characters of `CTYPE`*ka* shall be blank. The original FITS paper by Wells et al. (1981) contained a number of suggestions for the values of `CTYPE`*i*. Those suggestions are to be superseded by the conventions of this paper and of Papers I and II.

The relationships between the basic physical quantities ν, λ, and v are shown in Table 3. The symbols λ_0 and ν_0 are the rest wavelength and frequency, respectively, of the spectral line used to associate velocity with observed wavelength and frequency. The relationships between the derived quantities and the basic quantities with which they are associated are shown in Table 4. Derivatives are included in both tables since they will be needed in the coordinate computation.

[1] While Papers I and II use the generic intermediate axis number i, we use here the axis number k as the spectral intermediate axis number. The single-character version code a was introduced in Paper I. It has values blank and `A` through `Z`.

Table 2. Non-linear algorithm codes, (characters 6–8 of `CTYPE`*ka*).

Code	Regularly sampled in	Expressed as
`F2W`	Frequency	Wavelength
`F2V`	Frequency	Apparent radial velocity
`F2A`	Frequency	Air wavelength
`W2F`	Wavelength	Frequency
`W2V`	Wavelength	Apparent radial velocity
`W2A`	Wavelength	Air wavelength
`V2F`	Apparent radial velocity	Frequency
`V2W`	Apparent radial velocity	Wavelength
`V2A`	Apparent radial velocity	Air wavelength
`A2F`	Air wavelength	Frequency
`A2W`	Air wavelength	Wavelength
`A2V`	Air wavelength	Apparent radial velocity
`LOG`	logarithm	Any 4-letter coordinate type
`GRI`	detector	Any from Table 1
`GRA`	detector	Any from Table 1
`TAB`	not regular	Any 4-letter coordinate type

3. Basic coordinate computation

This section describes the computation of coordinates that are sampled linearly or logarithmically in a coordinate of similar type, and of coordinates that are sampled linearly in a coordinate of different type. The intermediate world coordinate for spectral axis k, given by Eq. (1), will be denoted, for convenience, by

$$w \equiv x_k. \tag{3}$$

The final world coordinate will be denoted by S. At the reference point, it has the value S_r, which is defined by the keyword `CRVAL`*ka*.

3.1. Linear coordinates

Linear coordinates are represented in `CTYPE`*ka* by characters 1-4 containing any code in column one of Table 1 and characters 5-8 blank. The world coordinate value for spectral axis k is computed with the above definitions and Eq. (1) as

$$S = S_\mathrm{r} + w. \tag{4}$$

As a general rule, non-linear coordinate systems will be constructed so that they satisfy Eq. (4) to first order at the reference point.

Table 3. Basic spectral transformation equations.

$\nu = \frac{c}{\lambda}$	(6)	$\frac{d\nu}{d\lambda} = -\frac{c}{\lambda^2}$	(7)
$\nu = \nu_0 \frac{c-v}{\sqrt{c^2-v^2}}$	(8)	$\frac{d\nu}{dv} = -\frac{c\nu_0}{(c+v)\sqrt{c^2-v^2}}$	(9)
$\lambda = \frac{c}{\nu}$	(10)	$\frac{d\lambda}{d\nu} = -\frac{c}{\nu^2}$	(11)
$\lambda = \lambda_0 \frac{c+v}{\sqrt{c^2-v^2}}$	(12)	$\frac{d\lambda}{dv} = \frac{c\lambda_0}{(c-v)\sqrt{c^2-v^2}}$	(13)
$v = c\frac{\nu_0^2-\nu^2}{\nu_0^2+\nu^2}$	(14)	$\frac{dv}{d\nu} = -\frac{4c\nu\nu_0^2}{(\nu^2+\nu_0^2)^2}$	(15)
$v = c\frac{\lambda^2-\lambda_0^2}{\lambda^2+\lambda_0^2}$	(16)	$\frac{dv}{d\lambda} = \frac{4c\lambda\lambda_0^2}{(\lambda^2+\lambda_0^2)^2}$	(17)

Table 4. Extended spectral transformation equations.

$\nu = \nu_0(1-\frac{V}{c})$	(18)	$\frac{d\nu}{dV} = -\frac{\nu_0}{c}$	(19)
$V = c\frac{\nu_0-\nu}{\nu_0}$	(20)	$\frac{dV}{d\nu} = -\frac{c}{\nu_0}$	(21)
$\nu = E/h$	(22)	$\frac{d\nu}{dE} = 1/h$	(23)
$E = h\nu$	(24)	$\frac{dE}{d\nu} = h$	(25)
$\nu = c\kappa$	(26)	$\frac{d\nu}{d\kappa} = c$	(27)
$\kappa = \nu/c$	(28)	$\frac{d\kappa}{d\nu} = 1/c$	(29)
$\lambda = \lambda_0(1+\frac{Z}{c})$	(30)	$\frac{d\lambda}{dZ} = \frac{\lambda_0}{c}$	(31)
$Z = c\frac{\lambda-\lambda_0}{\lambda_0}$	(32)	$\frac{dZ}{d\lambda} = \frac{c}{\lambda_0}$	(33)
$\lambda = \lambda_0(1+z)$	(34)	$\frac{d\lambda}{dz} = \lambda_0$	(35)
$z = \frac{\lambda-\lambda_0}{\lambda_0}$	(36)	$\frac{dz}{d\lambda} = \frac{1}{\lambda_0}$	(37)
$v = c\beta$	(38)	$\frac{dv}{d\beta} = c$	(39)
$\beta = v/c$	(40)	$\frac{d\beta}{dv} = 1/c$	(41)

3.2. Logarithmic coordinates

Data are often sampled in logarithmic increments in one or more coordinates. For example, spectra are sometimes sampled in logarithmic increments of wavelength or frequency. With this type of sampling, the source motion, expressed as V, Z, or z, shifts the "pixel" coordinates of spectral features by the same amount over the whole image. This allows relative velocities between spectra to be determined using cross-correlation methods in the pixel arrays.

For logarithmic axes, the last four characters shall be `'-LOG'`. While there are only three logarithmic coordinates commonly used in spectroscopy, namely `FREQ-LOG`, `WAVE-LOG` and `AWAV-LOG`, it would be unwise to forbid any coordinate type with the `-LOG` non-linear algorithm. Many such combinations may have little physical meaning or be intractable mathematically, but these are simply reasons to be cautious when using `-LOG`. This algorithm is evaluated simply

$$S = S_r e^{w/S_r}. \tag{5}$$

This form of the logarithm satisfies the requirement that $\frac{dS}{dw}\big|_r = 1$ so that Eq. (4) is satisfied to first order at points near the reference point. Thus $S \approx S_r + w$. This form of the logarithmic algorithm also has the desirable attribute that the units of the coordinate (S), reference coordinate, and scale are the same and are of a simple form. The units for `CRVAL`*ka*, `CDELT`*ka*, and `CD`*k_ja* are specified by the `CUNIT`*ka* keyword as defined in Paper I. These quantities are in normal, non-logarithmic units such as `'Hz'` for `FREQ-LOG` and `'m'` for `WAVE-LOG`. The prefixes and alternate units described in Paper I may be used, such as `'MHz'` and `'Angstrom'`.

Logarithmic quantities are frequently expressed as log base 10 rather than as natural logarithms. For such data, the `CDELT`*ka* and `CD`*k_ja* will need to compensate by including a factor of ln(10).

3.3. Coordinate axis names

The generality of this algorithm, and the `-TAB` algorithm to be introduced below, suggest the need for a more general description of the coordinate than may be indicated in the first four letters in the value of `CTYPE`*ka*. We hereby reserve the keyword

`CNAME`*ia* (character-valued)

to provide a description of a particular coordinate in up to 68 characters. Its default value will be all blank. For binary table vectors, the keyword will be *i*`CNA`*na*, while for pixel lists it will be `TCNA`*na*, where *i* is the intermediate world coordinate axis number and *n* is the table column number to which the keyword applies. This keyword may be used with standard axis types such as `GLON`/`GLAT` or `FREQ`, but will be of greatest value with non-standard axis types such as, hypothetically, `CTYPE`*i* = `'TFPY-LOG'` meant to indicate "log of Y position in telescope focal plane," which would be recorded in the `CNAME`*ia* card for that axis. This keyword provides a name for an axis in a particular WCS, while the `WCSNAME`*a* keyword names the particular WCS as a whole.

3.4. Non-linear spectral coordinates

We now consider the case where an axis is linearly sampled in spectral variable X, but is to be expressed in terms of variable S.

3.4.1. Spectral algorithm codes

Given the large number of spectral variables in Table 1 it is clear that there are very many pairwise combinations; in general, the relationship between any pair may be non-linear. However, each of the spectral variables in Table 1 is linearly related to one or other of ν, λ, λ_a, or υ. Thus we may restrict X to one of these four basic variables.

Even with this restriction on X there are still many possible combinations of X and S. In order to reduce the number still further we introduce an intermediate variable, denoted by P, that is the basic variable, ν, λ, λ_a, or υ, with which S is associated via a linear relation. This associate variable is listed for each spectral variable in column 4 of Table 1. Thus the sequence of transformations may be summarized as

$$p_j \rightarrow x_k\,(\equiv w) \rightarrow X \rightsquigarrow P \rightarrow S$$

where the non-linear transformation, indicated by the wiggly arrow, is between X and P.

Table 3 lists the equations, $X = X(P)$, and their inverses, $P = P(X)$, linking the basic spectral types ν, λ, and υ (discussion of λ_a is deferred to Sect. 4). These equations are generally non-linear. Likewise, Table 4 lists the linear relations, $S = S(P)$, and their inverses, $P = P(S)$, linking each spectral variable with its associate. When S is one of the basic types, ν, λ, λ_a, or υ, then $P \equiv S$ and $S(P)$ is identity.

Thus the functional relationship between S and X is specified via intermediate variable P as $S(X) = S(P(X))$ with inverse $X(S) = X(P(S))$. Since $S(P)$ is linear, P and X must always differ, otherwise $P(X)$ would be identity with the result that $S(X)$ would be linear, contrary to the assumption of a non-linear axis.

Non-linear spectral coordinate codes are constructed on this basis by combining the spectral coordinate type code for S from Table 1 with the non-linear spectral algorithm code in Table 2. The first letter of the algorithm code defines X as frequency (`F`), wavelength (`W`), air wavelength (`A`), or apparent radial velocity (`V`), and the third letter likewise defines P. For example, in `ZOPT-F2W`, X is `Frequency`, P is `Wavelength`, and the non-linear conversion between the two (`F2W`) is defined in Table 3, Eq. (10). The desired spectral coordinate S is redshift (`ZOPT`), and this is related to associate variable P, the `Wavelength`, by the linear equation given in Table 4, Eq. (36).

It is apparent from the foregoing that it is possible to construct invalid spectral `CTYPE` coordinate codes. For example, `ZOPT-F2V` is unrecognized since z is associated with λ, not υ. That is not to say that this code could not be interpreted in principle — after all, equations linking z and υ could have been included in Table 4 — simply that it is not recognized in practice. The associate variable, P, was introduced in the first place to reduce the possible number of combinations and adding new ones like this would defeat that purpose. This is particularly relevant to the software implementation; `ZOPT-F2V` would tell software to chain its $\nu - \upsilon$ function with a $\upsilon - z$ function but, in general, the latter function will not have been defined.

3.4.2. Spectral algorithm chain

Consider now the computation of spectral coordinate S for an axis that is linearly sampled in X. The statement that an axis is linearly sampled in X simply means that

$$X = X_r + w\,\frac{dX}{dw}, \tag{42}$$

where dX/dw is constant. This constant is determined by imposing the requirement that

$$\left.\frac{dS}{dw}\right|_r = 1, \tag{43}$$

so that Eq. (4) is satisfied to first order at points near the reference point:

$$S \approx S_r + w. \tag{44}$$

Thus

$$\frac{dX}{dw} = \left.\frac{dP}{dS}\right|_r \Big/ \left.\frac{dP}{dX}\right|_r , \tag{45}$$

which gives dX/dw as a function of X_r, dP/dS being constant and dP/dX a function of X. Given that the functions $S = S(P)$ and $P = P(X)$ are known, as are their inverses $X = X(P)$ and $P = P(S)$, then the equation for S as a function of w may be obtained from Eqs. (42) and (45)

$$S(w) = S\left(P\left(X(P(S_r)) + w\,\left.\frac{dP}{dS}\right|_r \Big/ \left.\frac{dP}{dX}\right|_r \right)\right), \tag{46}$$

where S_r is given by `CRVAL`*ka*.

Equation (46) suggests a three-step algorithm chain:

1. Compute X at w using Eq. (42); $X_r = X(P(S_r))$ and dX/dw are constants that need be computed once only, the latter obtained from Eq. (45) as a function of X_r.
2. Compute P from X using the appropriate equation from Table 3.
3. Compute S from P using the appropriate equation from Table 4.

The inverse computation by which the intermediate coordinate w is computed for a given value of S is effected by traversing the algorithm chain in the reverse direction. The inverse equations required are all listed in Tables 3 and 4. The inverse of Eq. (42) is, trivially,

$$w = (X - X_r) \Big/ \frac{dX}{dw}. \tag{47}$$

3.4.3. Example non-linear calculation

As an example of a non-linear coordinate computation, consider a `ZOPT-F2W` axis. `F` in the `F2W` code indicates that the axis is linearly sampled in frequency, but it is to be expressed in terms of redshift as indicated by the spectral coordinate type of `ZOPT`. Table 1 indicates that redshift is associated with wavelength, hence the `W` in `F2W` is correct. Of the X, P, and S variables in the preceding section, the axis is linear in frequency so X is ν, the associated variable is wavelength, so P is λ, and the required variable is redshift, so S is z.

Table 5. Sample non-linear coordinate combinations, where w, the intermediate coordinate of Eq. (3), has a different meaning for each of these equations It has units given by the spectral coordinate type code as listed in Table 1.

`FREQ-W2F`	$\nu = \left(\frac{\nu_\mathrm{r}}{\nu_\mathrm{r} - w}\right)\nu_\mathrm{r}$	(48)						
`VELO-W2V`	$v = c\,\frac{A_\lambda^2 - 1}{A_\lambda^2 + 1}$	(49)	where	$A_\lambda \equiv \frac{c^2 - v_\mathrm{r}^2 + cw}{(c - v_\mathrm{r})\sqrt{c^2 - v_\mathrm{r}^2}}$	(50)			
`VRAD-W2F`	$V = \frac{V_\mathrm{r}(c - V_\mathrm{r}) + cw}{c - V_\mathrm{r} + w}$	(51)						
`WAVE-F2W`	$\lambda = \left(\frac{\lambda_\mathrm{r}}{\lambda_\mathrm{r} - w}\right)\lambda_\mathrm{r}$	(52)						
`VELO-F2V`	$v = c\,\frac{A_\nu^2 - 1}{A_\nu^2 + 1}$	(53)	where	$A_\nu \equiv \frac{c^2 - v_\mathrm{r}^2 - cw}{(c + v_\mathrm{r})\sqrt{c^2 - v_\mathrm{r}^2}}$	(54)			
`VOPT-F2W`	$Z = \frac{Z_\mathrm{r}(c + Z_\mathrm{r}) + cw}{c + Z_\mathrm{r} - w}$	(55)						
`ZOPT-F2W`	$z = \frac{z_\mathrm{r}(1 + z_\mathrm{r}) + w}{1 + z_\mathrm{r} - w}$	(56)						
`WAVE-V2W`	$\lambda = \lambda_0\sqrt{\frac{1 + B_\lambda}{1 - B_\lambda}}$	(57)	where	$B_\lambda \equiv \frac{\lambda_\mathrm{r}^4 - \lambda_0^4 + 4\lambda_0^2\lambda_\mathrm{r}w}{(\lambda_0^2 + \lambda_\mathrm{r}^2)^2}$	(58)			
`FREQ-V2F`	$\nu = \nu_0\sqrt{\frac{1 - B_\nu}{1 + B_\nu}}$	(59)	where	$B_\nu \equiv \frac{\nu_0^4 - \nu_\mathrm{r}^4 - 4\nu_0^2\nu_\mathrm{r}w}{(\nu_0^2 + \nu_\mathrm{r}^2)^2}$	(60)			
`VOPT-V2W`	$Z = c\left(\sqrt{\frac{1 + D_Z}{1 - D_Z}} - 1\right)$	(61)	where	$D_Z \equiv \frac{Y_Z^4 - 1 + 4Y_Z w/c}{(1 + Y_Z^2)^2}$	(62)	with	$Y_Z \equiv 1 + Z_\mathrm{r}/c$	(63)

Since `CRVAL`ka for the `ZOPT-F2W` axis would be recorded as a redshift, z_r, this must first be converted to frequency by applying Eqs. (34),

$$\lambda_\mathrm{r} = \lambda_0(1 + z_\mathrm{r})$$

and (6),

$$\nu_\mathrm{r} = \frac{c}{\lambda_\mathrm{r}} = \frac{c}{\lambda_0(1 + z_\mathrm{r})}.$$

This provides X_r for Eq. (42). Then $\mathrm{d}\nu/\mathrm{d}w$ (i.e. $\mathrm{d}X/\mathrm{d}w$) is obtained from Eqs. (35) & (11) evaluated at the reference point:

$$\frac{\mathrm{d}\nu}{\mathrm{d}w} = \left.\frac{\mathrm{d}\lambda}{\mathrm{d}z}\right|_\mathrm{r} \Big/ \left.\frac{\mathrm{d}\lambda}{\mathrm{d}\nu}\right|_\mathrm{r} = \lambda_0 \Big/ \frac{-c}{\nu_\mathrm{r}^2} = -\frac{\nu_\mathrm{r}^2}{\nu_0}.$$

Redshift may now be computed for any given value of w. The first step is to compute ν (i.e. X) at w using Eq. (42):

$$\nu = \nu_\mathrm{r} + w\,\frac{\mathrm{d}\nu}{\mathrm{d}w}.$$

In this instance w will be a redshift. Then λ (i.e. P) may be computed from ν using Eq. (10) from Table 3:

$$\lambda = c/\nu.$$

The third and final step is to compute z (i.e. S) from λ using Eq. (36) from Table 4:

$$z = \frac{\lambda - \lambda_0}{\lambda_0}.$$

It is also possible to combine the three steps into a single equation, although there are many pair-wise combinations of spectral variables and for some the combined equations may be quite cumbersome. For the sake of illustration, consider the above equations for the `ZOPT-F2W` axis. Eq. (42) becomes

$$\nu = \frac{c}{\lambda_0(1 + z_\mathrm{r})} - \frac{cw}{\lambda_0(1 + z_\mathrm{r})^2}.$$

Substituting this in the equations for λ and z and simplifying we obtain

$$z = \frac{z_\mathrm{r}(1 + z_\mathrm{r}) + w}{1 + z_\mathrm{r} + w}.$$

A representative sample of these direct translation equations is shown in Table 5. They are useful in showing the nature of the coordinate non-linearity.

One thing to notice about the equations of Table 5 is that the meaning of w differs for each. For example, Eq. (55) may not be obtained from Eq. (56) simply by multiplying both sides by c: in Eq. (56) w is a redshift (`ZOPT`), whereas in Eq. (55) it is an optical velocity (`VOPT`).

3.4.4. Coordinate parameters

Aside from `CRVAL`ka, the coordinate computations of Sect. 3.4 require one extra parameter when evaluating the `F2V`, `V2F`, `W2V`, `V2W`, `A2V`, and `V2A` non-linear algorithms, namely the rest frequency or wavelength of the spectral-feature of interest. These are fundamental physical parameters so, rather than use the `PV`*i*_*ma* parameters defined in Paper I, which would tend to disguise them, the special floating-valued keywords

`RESTFRQ`*a* (floating-valued),

and

`RESTWAV`*a* (floating-valued),

are hereby reserved for the purpose. They are represented by symbols ν_0 and λ_0, respectively. Their units are `'Hz'` and `'m'` respectively, fixed to save having additional keywords to define them. `RESTWAV`*a* is for wavelengths in vacuum only. One or the other of these keywords must be included for the above-mentioned algorithm codes; usually `RESTFRQ`*a* would be given for `F2V` and `V2F`, and `RESTWAV`*a* for the others. FITS writers should always write one or other of these when it is meaningful to do so, even for algorithm codes such as `F2W` or `W2A` that do not require them.

Keyword `RESTFREQ` has been used in previous FITS files and should be recognized as equivalent to `RESTFRQ`.

4. Air wavelengths

The wavelengths so far discussed are measured in vacuum. However, dispersion coordinates for UV, optical, and IR spectra at $\lambda > 200$ nm are commonly given as wavelengths in air. The relative difference between the two varies between 0.028% and 0.032% across this range. To identify wavelengths measured in dry air at standard temperature and pressure rather than in vacuo, we introduce the coordinate type `AWAV` for which the reference value and increment must be expressed accordingly.

The two measures of wavelength are related by

$$\lambda = n(\lambda_a)\lambda_a, \tag{64}$$

where λ is the wavelength in vacuum, λ_a is the wavelength in air, and $n(\lambda_a)$ is the index of refraction of dry air at standard temperature and pressure. This varies non-linearly with wavelength. The standard relation given by Cox (2000) is mathematically intractable and somewhat dated. The International Union of Geodesy and Geophysics (1999) adopted the rather simpler formula[2]

$$n(\lambda_a) = 1 + 10^{-6}\left(287.6155 + \frac{1.62887}{\lambda_a^2} + \frac{0.01360}{\lambda_a^4}\right), \tag{65}$$

where λ_a is the wavelength in micrometers. This formula suffices for conversions from air to vacuum when no more than 0.25 parts per million accuracy is required. The derivative, which may be required in Eq. (45), is

$$\frac{d\lambda}{d\lambda_a} = 1 + 10^{-6}\left(287.6155 - \frac{1.62887}{\lambda_a^2} - \frac{0.04080}{\lambda_a^4}\right). \tag{66}$$

While the inversion of Eq. (65) is algebraically intractable, the vacuum wavelength may be used to evaluate Eq. (65) since it differs so little from the air wavelength. The resulting error in the index of refraction amounts to $1 : 10^9$, and this is less than than the accuracy of the empirical formula. Thus,

$$\lambda_a = \lambda / n(\lambda). \tag{67}$$

As usual, an axis that is linearly sampled and expressed in air wavelengths is described with a `CTYPE`*ka* of `'AWAV'` and evaluated with Eq. (4).

Algorithm codes for the non-linear conversions are `A2F`, `A2W`, `A2V`, and their complements, `F2A`, `W2A`, and `V2A` as listed in Table 2. Use of the three-step procedure described in Sect. 3.4 would require λ_a as a function of each of the other spectral variables, together with their inverses and derivatives. Thus it is much simpler to handle air wavelengths as a separate, extra step in the algorithm chain. For example, to compute world coordinates for `VRAD-A2V`, the value of `CRVAL`*ka* would first have to be converted from radio velocity to air wavelengths via Eqs. (18), (10), and (67), and $d\lambda_a/dw$ for Eq. (45) would be obtained from Eqs. (19), (7), and (66). Then, the value of λ_a computed for w via Eq. (42) would be transformed to vacuum wavelengths via Eq. (64), and then to radio velocity via Eqs. (6) and (20).

5. Dispersed spectra

One common form of spectral data is produced by imaging the light from a *disperser*. The wavelengths of the light at some position in the image are related to the position and wavelength of the light in the field of view illuminating the disperser; the relation is defined by the physics of the disperser. In general the light received at a pixel in the detector is a superposition of different wavelengths from different points in the field of view. However, if the field of view is limited spatially, usually by aperture masks and fiber optics, each pixel receives light from only a small range of wavelengths. It is then possible to define a world coordinate function relating pixel position and effective wavelength. This is the basis of many astronomical spectrographs.

In the following section we use the physical relation applicable to the dispersers commonly used in astronomical spectrographs to define a world coordinate function for computing spectral coordinates. The relation applies to the simple, though common, case of single dispersers. More complex spectrographs with multiple dispersers, such as those using multiple passes through prisms, are not described by the methods of this section. The equations developed below also assume that the radiation enters perpendicular to the face of the prism, a condition not met by some widely used spectrometers. Alternatives to using this ideal world coordinate function, based on the physics of simple dispersers, are the table lookup described in Sect. 6 and empirical function fits provided by the the distortion function mechanism described in Paper IV.

We require that the dispersion occurs along just one direction on the detector and that the intermediate coordinates are computed so that only one world coordinate axis corresponds to wavelength. Paper IV describes how distortions and effects of the aperture shape can be removed to satisfy this requirement. The distortion correction is also used to remove aberra-

[2] The quoted formulæ apply only at normal optical wavelengths. In the UV and IR spectral domains, atmospheric absorption lines cause refractivity to be a strong, weather dependent function of frequency. See, for example, Mathar (2004).

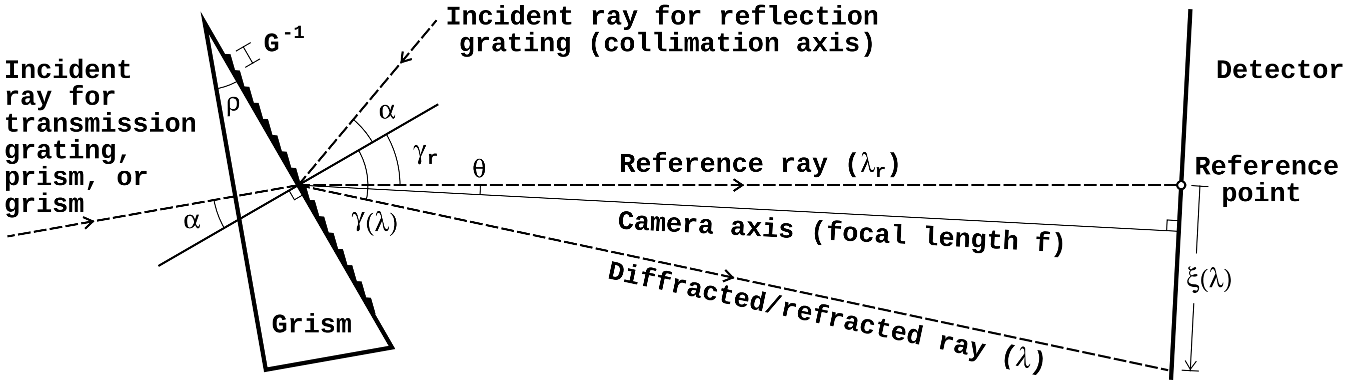


Fig. 2. Geometry of gratings, prisms, and grisms. This simplified representation omits the collimation and focusing optics. Dashed lines mark ray paths in the plane of the figure – the "dispersion plane". The normal to the grating/exit prism face and the normal to the detector plane are each projected onto the dispersion plane, and angles α, γ, and θ are measured with respect to these projected normals. Usually the incident ray for a prism or grism is perpendicular to the entry face so that α is equal to the prism angle, ρ. Angle γ is wavelength-dependent, and consequently so is the offset ξ in the dispersion direction on the detector. The intermediate spectral world coordinate, w, is proportional to ξ. Reference wavelength λ_r follows the reference ray defined by γ_r and illuminates the reference point at $w = \xi = 0$. The normal to the detector plane is shown tilted by angle θ from the reference ray though typically this angle is zero. The grating spacing G^{-1} is indicated.

tions causing departures from the ideal physical behavior of the dispersers assumed here.

The dispersers we consider are *gratings*, *prisms*, and *grisms*, which are a combination of a prism with a grating within or on the surface of the prism. Gratings may be reflecting or transmissive and may be fabricated with surface relief rulings, holographic surface relief patterns, and holographic volume phase patterns. In the discussion we use the terms *grooves*, *lines*, and *ruling* to refer to the periodic diffraction structures that produce the interference. By combining the two physical equations for interference and refraction a single world coordinate function covers all these cases with appropriate choice of parameters.

5.1. The grism coordinate function

This section defines a world coordinate function using the basic laws of refraction and interference. It is beyond the scope of this discussion to give the full background for the laws and concepts underlying dispersive spectroscopy. Of the many texts on the subject, a standard reference is *Astronomical Optics* by Schroeder (1999).

Spectroscopic dispersers are based on interference and refraction effects which are naturally described in terms of wavelength. Moreover, they are most often used in the regime where wavelength is commonly the spectral coordinate. For these reasons the grism function is defined in terms of wavelength. The function assumes that all of the dispersion occurs at one place. This is why combinations of dispersers, other than a single grism, are not described by this function. This assumption means that for prisms and grisms the function is rigorously correct only when the light enters perpendicular to the face of the prism and the grating is at the exit face. However, even when these conditions are not exactly met the function may still be a good approximation with slight adjustments of the parameters.

5.1.1. The grism equation

The basic physical relationship between the wavelength, λ, the angle of incidence of the light, α, and the diffracted/refracted angle, γ, is given by a combination of the grating interference equation and Snell's law of refraction: [3]

$$\frac{Gm\lambda}{\cos\epsilon} = n(\lambda)\sin\alpha + \sin\gamma, \tag{68}$$

where G is the grating ruling density, m is the interference order, and $n(\lambda)$ is the wavelength-dependent index of refraction of the prism material. For a pure prism the order m is zero and for a reflection or transmission grating $n(\lambda)$ is the unit function. The plane containing α and γ is the *dispersion plane*, illustrated in Fig. 2, and the angles are measured relative to the projection in the dispersion plane of the normal to the grating or exit prism face. Usually the normal does lie in the dispersion plane, but small values of ϵ, the angle between the normal and the dispersion plane, are sometimes used for instrumental design reasons.

By convention, the angles of incidence and diffraction/refraction are measured *from* the normal *to* the beam and α is always measured positive in the anticlockwise direction. The sign changes on either side of the normal and this determines the sign of γ. Thus, for the reflection grating in Fig. 2, $\alpha > 0$ and $\gamma < 0$, and vice versa for a transmission grating, prism, or grism.

Reflection grating geometry is sometimes defined by the angle ϕ measured from the camera axis to the collimator axis and the tilt t of the bisector relative to the grating normal. If ϕ and t obey the same sign convention as α, then $\alpha = \phi/2 + t$.

In spectrographs with prisms or grisms the requirement that the incident light be normal to the prism entry face means that α is equal to the prism angle (ρ in Fig. 2). Even for oblique entry, identifying α with the prism angle is often a good initial approximation.

[3] Also known as Descartes' Law of Sines; see for example, `http://wikipedia.org/wiki/Snells_law`

The requirement that the diffraction and refraction occur at one point means that α is fixed and independent of wavelength. The variation of γ with wavelength then defines the relation between wavelength and position ξ on the detector. The reference wavelength λ_r is the wavelength at the reference point corresponding to the zero point of ξ and the intermediate spectral world coordinate w.

The prism's dispersive power derives from the variation of its index of refraction with wavelength. While this variation depends on the material and is generally non-linear, it is sufficient to approximate it by a first-order Taylor expansion about the reference wavelength, λ_r:

$$n(\lambda) \approx n_r + n'_r(\lambda - \lambda_r), \tag{69}$$

where we have written n_r as a shorthand for $n(\lambda_r)$, and likewise n'_r for $\mathrm{d}n/\mathrm{d}\lambda|_r$. Combining Eqs. (68) and (69) yields the grism equation

$$\lambda = \frac{(n_r - n'_r\lambda_r)\sin\alpha + \sin\gamma}{Gm/\cos\epsilon - n'_r\sin\alpha}, \tag{70}$$

where the denominator must not be zero, though Gm, n'_r, and α may be zero in different types of spectrographs.

In order to define a world coordinate function we need λ as a function of the intermediate world coordinate, w, which is proportional to ξ. Since Eq. (70) gives λ as a function of γ it remains to determine the relationship between γ and ξ. First note that Eq. (70) evaluated at $\lambda = \lambda_r$ provides

$$\gamma_r = \sin^{-1}(Gm\lambda_r/\cos\epsilon - n_r\sin\alpha), \tag{71}$$

the exit angle of the reference ray for which $w = \xi = 0$, as in Fig. 2.

Fig. 2 shows angle θ, which is measured from the reference ray to the camera axis using the same sign convention as γ, i.e. if γ is clockwise-positive as for a grism then so is θ. Normally θ would be zero, but it is included here to provide a small correction. Depending on the sign convention for γ, simple geometry for a flat detector gives

$$\gamma = \gamma_r + \theta + \tan^{-1}(\pm\xi/f - \tan\theta), \tag{72}$$

where f is the effective focal length of the camera. The plus sign is taken when the sign convention for γ is such that ξ increases as γ increases, and the minus sign otherwise. Section 5.1.3 discusses the resolution of this potential sign ambiguity.

5.1.2. GRI coordinate axes

In keeping with the preceding sections we wish to define general grating, prism, and grism world coordinate representations such as `WAVE-GRI`, `FREQ-GRI`, etc.

Bearing in mind that the grism equation, Eq. (70), is expressed in terms of wavelength, then given a `FREQ-GRI` axis, for example, it would be straightforward to convert the reference frequency, ν_r, given by `CRVAL`k, from frequency to wavelength via $\lambda = c/\nu$. However, it would not be valid to convert the intermediate spectral world coordinate, w, from frequency to wavelength like this because it is not a true frequency. While it may be a close approximation near the reference point it deviates at points away from it.

Thus, interpretation of an axis such as `FREQ-GRI` necessarily involves a procedure similar to that of Sect. 3.4, and to apply that methodology we need a parameter, the *grism parameter*, Γ, that is a known function of the spectral variables and for which the axis is linearly sampled; this will substitute for X in Eq. (42). Since ξ is proportional to w

$$\xi = \sigma w, \tag{73}$$

where σ is a constant. Combining this with Eq. (72) we have

$$\tan(\gamma - \gamma_r - \theta) = -\tan\theta \pm w\sigma/f. \tag{74}$$

Thus the grism parameter may be identified as

$$\Gamma = \tan(\gamma - \gamma_r - \theta), \tag{75}$$

and this satisfies Eq. (42) at the reference point:

$$\Gamma_r = -\tan\theta. \tag{76}$$

The grism parameter has a simple geometrical interpretation in Fig. 2; it is the offset on the detector from the point of intersection of the camera axis measured in units of the effective focal length, f.

Following Sect. 3.4 (with X replaced by Γ), we have

$$\Gamma = \Gamma_r + w\frac{\mathrm{d}\Gamma}{\mathrm{d}w} \tag{77}$$

where $d\Gamma/dw$ is constant:

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}w} = \left.\frac{\mathrm{d}\Gamma}{\mathrm{d}P}\right|_r \left.\frac{\mathrm{d}P}{\mathrm{d}S}\right|_r \left.\frac{\mathrm{d}S}{\mathrm{d}w}\right|_r. \tag{78}$$

As before, S is the spectral variable in which the axis, and hence w, is expressed, and P is the basic variable, ν, λ, λ_a, or υ, with which S is most closely associated. Recognized spectral variables and their associates are listed in Table 1. When S is one of the basic types, ν, λ, λ_a, or υ, as is often the case, then $P \equiv S$ whence $\mathrm{d}P/\mathrm{d}S|_r = 1$.

As previously, we require that `CDELT`i or `CD`i`_`j be set so that

$$\left.\frac{\mathrm{d}S}{\mathrm{d}w}\right|_r = 1 \tag{79}$$

(exactly how this is done is addressed in Sect. 5.1.3). Since we only have Γ directly in terms of γ we may use

$$\left.\frac{\mathrm{d}\Gamma}{\mathrm{d}P}\right|_r = \left.\frac{\mathrm{d}\Gamma}{\mathrm{d}\lambda}\right|_r \left.\frac{\mathrm{d}\lambda}{\mathrm{d}P}\right|_r, \tag{80}$$

where $\mathrm{d}\Gamma/\mathrm{d}\lambda|_r$ itself may be deduced from the derivatives of Eqs. (75) & (70):

$$\begin{aligned}\left.\frac{\mathrm{d}\Gamma}{\mathrm{d}\lambda}\right|_r &= \left.\frac{\mathrm{d}\Gamma}{\mathrm{d}\gamma}\right|_r \Big/ \left.\frac{\mathrm{d}\lambda}{\mathrm{d}\gamma}\right|_r \\ &= \frac{Gm/\cos\epsilon - n'_r\sin\alpha}{\cos\gamma_r\cos^2\theta}.\end{aligned} \tag{81}$$

Combining Eqs. (78) to (81) we have

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}w} = \frac{Gm/\cos\epsilon - n'_r\sin\alpha}{\cos\gamma_r\cos^2\theta} \left.\frac{\mathrm{d}\lambda}{\mathrm{d}P}\right|_r \left.\frac{\mathrm{d}P}{\mathrm{d}S}\right|_r, \tag{82}$$

Table 6. Grism parameters, their default values, and required units.

Keyword	Default	Units
PVk_0$a = G$	0	m^{-1}
PVk_1$a = m$	0	
PVk_2$a = \alpha$	0	deg
PVk_3$a = n_\mathrm{r}$	1	
PVk_4$a = n'_\mathrm{r}$	0	m^{-1}
PVk_5$a = \epsilon$	0	deg
PVk_6$a = \theta$	0	deg

where $\mathrm{d}\lambda/\mathrm{d}P|_\mathrm{r}$ and $\mathrm{d}P/\mathrm{d}S|_\mathrm{r}$ come from Tables 3 & 4 respectively. In the common case where S, and hence P, is wavelength then $\mathrm{d}\lambda/\mathrm{d}P|_\mathrm{r}$ and $\mathrm{d}P/\mathrm{d}S|_\mathrm{r}$ are both unity.

The foregoing provides everything needed to compute Γ from w, and this forms the first step of a five-step algorithm chain for computing S from w:

1. Compute Γ at w using Eq. (77). Γ_r and $\mathrm{d}\Gamma/\mathrm{d}w$ are constants that need be computed once only; Γ_r from Eq. (76) and $\mathrm{d}\Gamma/\mathrm{d}w$ from Eq. (82) using the appropriate equations from Tables 3 & 4.
2. Compute γ from Γ using the inverse of Eq. (75):

$$\gamma = \tan^{-1}(\Gamma) + \gamma_\mathrm{r} + \theta. \tag{83}$$

3. Compute the wavelength λ from γ via Eq. (70).
4. Compute P from λ using the appropriate equation from Table 3.
5. Compute S from P using the appropriate equation from Table 4.

If S, and hence P, is wavelength then the last two steps are not performed.

As usual, the inverse computation is effected by traversing the algorithm chain in the reverse direction. The inverse equations required for the backward steps have already been given except that for Eq. (70):

$$\gamma = \sin^{-1}(\lambda(Gm/\cos\epsilon - n'_\mathrm{r}\sin\alpha) - (n_r - n'_\mathrm{r}\lambda_\mathrm{r})\sin\alpha), \tag{84}$$

and also the inverse of Eq. (77) which is trivial.

Grism parameters required for these calculations are provided by the PVk_ma keywords defined in Table 6, with γ_r given by Eq. (71), and λ_r computed from the coordinate reference value, CRVALk, by the appropriate equations from Tables 3 & 4. Default values for missing parameters are also defined in the table. Generally only the first three parameters will appear for gratings and the first five for grisms.

5.1.3. Determination of the grism scale

A question raised above was how CDELTi or CDi_j may be set by a WCS composer so that $\mathrm{d}S/\mathrm{d}w|_\mathrm{r} = 1$.

From Eqs. (1) & (3)

$$w = s_k q_k \tag{85}$$

so

$$s_k = \frac{\mathrm{d}w}{\mathrm{d}q_k} = \left.\frac{\mathrm{d}w}{\mathrm{d}S}\right|_\mathrm{r} \left.\frac{\mathrm{d}S}{\mathrm{d}P}\right|_\mathrm{r} \left.\frac{\mathrm{d}P}{\mathrm{d}\lambda}\right|_\mathrm{r} \left.\frac{\mathrm{d}\lambda}{\mathrm{d}q_k}\right|_\mathrm{r}. \tag{86}$$

Now, $\mathrm{d}w/\mathrm{d}S|_\mathrm{r} = 1$ by design, $\mathrm{d}S/\mathrm{d}P|_\mathrm{r}$ and $\mathrm{d}P/\mathrm{d}\lambda|_\mathrm{r}$ come from Tables 4 & 3 respectively, and $\mathrm{d}\lambda/\mathrm{d}q_k|_\mathrm{r}$ is the wavelength dispersion (e.g. in nm/pixel) measured at the reference point and this is an instrumental parameter. Thus we have everything needed to compute s_k. When S is wavelength, then all but the last factor in Eq. (86) is unity, and $s_k = \mathrm{d}\lambda/\mathrm{d}q_k|_\mathrm{r}$.

It is the correct choice of sign for $\mathrm{d}\lambda/\mathrm{d}q_k|_\mathrm{r}$ that resolves the sign ambiguity in Eq. (74).

5.1.4. GRA: grisms in air

Thus far we have ignored the distinction between vacuum and air wavelengths in the discussion of grism world coordinates. In fact, the λ variable that appears in the grism equation may be either, and in general $n(\lambda)$ in Eq. (68) is the index of refraction of the prism material with respect to the surrounding medium, either air or vacuum.

If the dispersion takes place in vacuum, as it does, for example, in a spectrograph on a spacecraft, then the grism equation is correct for vacuum wavelengths; if in air, it is correct for air wavelengths and λ in Sects. 5.1.1, 5.1.2, and 5.1.3 should be replaced everywhere by λ_a. Then as discussed in Sect. 4, the computation of quantities such as $\mathrm{d}\lambda_\mathrm{a}/\mathrm{d}P|_\mathrm{r}$ in Eq. (82) is best handled by using the vacuum wavelength, λ, as an intermediary, effectively introducing an extra step into the algorithm chain.

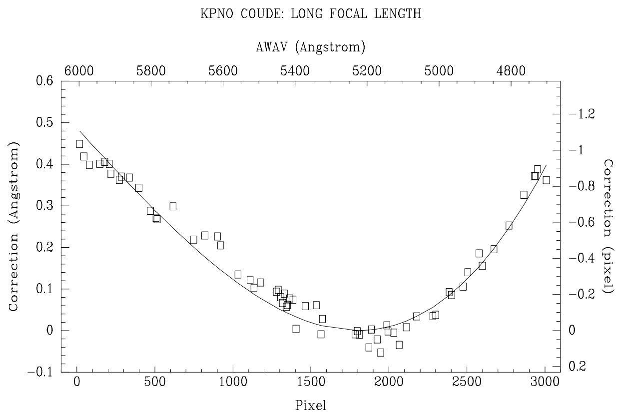
In order to distinguish between grisms operated in air and in vacuum we hereby reserve the use of GRI exclusively for vacuum operation and introduce GRA for operation in dry air at standard temperature and pressure.

Note that for real spectrographs operated in air at an observatory, the actual wavelength system that GRA describes is for air at the local conditions. The coordinate value and increment at the reference point are normally adjusted to the values at standard conditions during calibration. If the accuracy of the wavelength measurement requires it, any further correction between wavelength at local conditions and wavelength at standard conditions may be accomplished via the methods of Paper IV.

5.2. AWAV-GRA *examples*

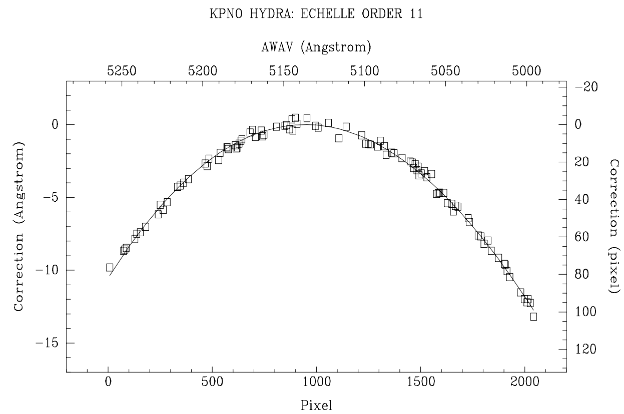
This section illustrates application of the GRA world coordinate function with three real-world examples of spectral images from Kitt Peak National Observatory (KPNO) spectrographs (Figs. 3–5). Each spectrum is of an arc calibration lamp that produces emission lines of known wavelength. The position of each line in the image along the spectral world coordinate axis is measured by centroiding on the spectral line profile and identified with the known rest wavelength to create a list of pixel positions and wavelengths.

Each example figure shows plots of the positions and wavelengths for a particular spectrograph. The pixel positions are plotted along the bottom axis. Rather than plot the wavelengths




```
CTYPE1  = 'AWAV-GRA' / Grating dispersion function
CUNIT1  = 'Angstrom' / Dispersion units
CRPIX1  =     1801.7 / [pixel] Reference pixel
CRVAL1  =     5225.2 / [Angstrom] Reference value
CDELT1  =    -0.4334 / [Angstrom/pixel] Dispersion
PV1_0   =     3.16E5 / [m^(-1)] Grating density
PV1_1   =          1 / Diffraction order
PV1_2   =       13.9 / [deg] Incident angle
```

Fig. 3. KPNO Coudé Feed spectrograph with a long focal length and a 3K CCD.




```
CTYPE1  = 'AWAV-GRA' / Grating dispersion function
CUNIT1  = 'Angstrom' / Dispersion units
CRPIX1  =      944.8 / [pixel] Reference pixel
CRVAL1  =     5136.8 / [Angstrom] Reference value
CDELT1  =    -0.1287 / [Angstrom/pixel] Dispersion
PV1_0   =     3.16E5 / [m^(-1)] Grating density
PV1_1   =         11 / Diffraction order
PV1_2   =       64.8 / [deg] Incident angle
```

Fig. 4. KPNO Hydra Fiber Bench Spectrograph using an echelle grating in 11th order with a 2K CCD.

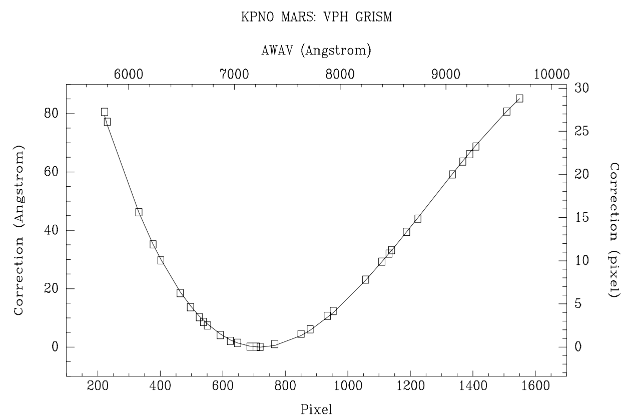
directly, where it would be difficult to see the shape of the curve, the difference or correction between the known (air) wavelengths and the simple linear world coordinate function `AWAV` ($\lambda_{\mathrm{a}} = S_{\mathrm{r}} + w$) are shown. The corrections are plotted along the left axes. In these examples the wavelength units are consistently Angstroms.

The top axis is labeled with simple `AWAV` linear coordinates. The right axis divides the wavelength correction by `CDELT1`, the linear dispersion at the reference pixel. This represents the shift, in pixels, of the wavelengths on the detector relative to where they would occur in a spectrograph with a linear dispersion relation.

The departure of the data points from zero indicate the magnitude of the world coordinate errors that would occur by using the simple linear `AWAV` world coordinate function. The solid lines in the figures are the difference between the wavelengths produced by the `AWAV-GRA` grating coordinate function and the linear `AWAV` coordinate function evaluated at the pixel positions in the image.

The usefulness of the ideal grating coordinate function is that the curves go through the measured data points with an appropriate choice of parameters. The parameters are essentially those known for the spectrograph and disperser with small adjustments to some of the parameters to produce the best fit to the calibration data. (These small adjustments may be viewed as corrections for the simplifications in the optical model.) In these examples the dispersion function represented by the grating world coordinate relation is as good as typically provided by empirical polynomial functions. Higher order effects due to aberrations are corrected by the distortion correction methods defined in Paper IV.

A close reading of the equations above will reveal that the seven grism parameters listed in Table 6 are not independent. We have chosen these parameters because of their physical meaning. However, the independent parameters are $Gm/\cos(\epsilon)$, $n_{\mathrm{r}}\sin(\alpha)$, $n'_{\mathrm{r}}\sin(\alpha)$, and θ. It is these combinations of parameters which must be used in fitting data.




```
CTYPE1  = 'AWAV-GRA' / Grating dispersion function
CUNIT1  = 'Angstrom' / Dispersion units
CRPIX1  =      719.8 / [pixel] Reference pixel
CRVAL1  =     7245.2 / [Angstrom] Reference value
CDELT1  =      2.956 / [Angstrom/pixel] Dispersion
PV1_0   =     4.50E5 / [m^(-1)] Grating density
PV1_1   =          1 / Diffraction order
PV1_2   =       27.0 / [deg] Incident angle
PV1_3   =      1.765 / Reference refraction
PV1_4   =  -1.077E6 / [m^(-1)] Refraction deriv
```

Fig. 5. KPNO MARS spectrograph with a 450 lines/mm volume phase holographic grism and a 2K CCD.

Each figure shows the arc line measurements, the coordinate function curve, and the relevant world coordinate keywords used to produce the curve. Not all of the WCS keywords are shown.

The first example shows the behavior of a long focal length spectrograph with a reflection grating used at a low angle of incidence. Because of the long focal length the deviation from linearity is relatively small though still clearly significant. The grating has a density of 316 lines/mm and is used in first order to produce a spectrum centered near 5225.2 Å with a dispersion of 0.43 Å/pixel.

The next example uses a 316 lines/mm echelle reflection grating, a grating designed for use at large angles from the grating normal, operated in a higher order. It is used to produce a spectrum centered near 5136.8 Å with a dispersion of 0.13 Å/pixel. The departure from linearity is in the opposite sense from the other examples because the echelle grating is used with the diffracted angle at a greater angle than the incidence angle. The higher order results in a fairly large departure from a linear WCS.

The last example illustrates use of a grism. However, this is not the more common grism with a ruled grating on the output face and the input face oriented normal to the incident beam. Instead, the grating is sandwiched in the middle of the prism. The prism is oriented with the beam entering and leaving the prism at equal angles to the faces resulting in a straight through configuration as with a standard grism.

Another unusual feature of this grism is that it uses a volume phase holographic (VPH) grating. While the intensity response is different from a surface relief grating (ruled or holographic) the dispersion behavior is the same.

The full optical equation is complex even in this symmetric configuration, but as shown in the figure using the `GRA` function, a very good description of the coordinates can be obtained.

The prism has a 27° angle with an index of refraction of 1.764 near the reference wavelength. The grating has an equivalent interference pattern of 450 lines/mm. Using these values and adjusting the derivative of the index of refraction produces Fig. 5.

6. Coordinates by table lookup

There are numerous instances in which a physical coordinate is well defined at each pixel along an image axis, but the relationship of the coordinate values between pixels cannot be described by a simple functional form. An obvious example of this would be a three-dimensional image consisting of a sequence of two-dimensional images of an astronomical object recorded at an arbitrary sequence of times determined in part by weather and time assignment committees. As another example, the calibration of some spectrographs, such as those employing a diode array detector, is represented best by a list of frequencies or wavelengths for each pixel on the spectral axis rather than some functional form.

6.1. *xxxx*-`TAB` *non-linear algorithm*

To support such representations for primary images and image extensions, we define a table-lookup form for the value of `CTYPE`*ia* as *xxxx*-`TAB`, where *xxxx* are four letters representing the type of coordinate, e.g. `TIME` or `FREQ`. As in Paper I, which established the "4–3" convention for `CTYPE`*i*, the coordinate *xxxx* is a "real" coordinate, such as `FREQ`, not an intermediate coordinate, such as `FREQ-F2W`, requiring an additional linear or non-linear algorithm in order to be converted into a physical coordinate.

6.1.1. -`TAB` indexing concepts

Consider first the case of a single (one-dimensional) coordinate axis. The -`TAB` algorithm uses a list of coordinate values, the *coordinate array*, to record coordinate values of the appropriate type for the coordinate axis. A second list of the same length, the *indexing vector*, may be used in addressing the coordinate array. The indexing vector provides one level of indirection which may be used to vary the sampling frequency of the coordinate array along the coordinate axis. The coordinate could then be sampled at smaller intervals over that portion of the range in which the instrument is more non-linear and sampled more coarsely over regions in which it is better behaved. The indexing algorithm, based on linear interpolation, is defined in Sect. 6.1.2. If the indexing vector is absent, it is taken to have values $1, 2, 3, \ldots, K$, where K is the number of values in the coordinate array. See Sect. 6.2.2 for an example of the use of the indexing vector.

The concept described above covers separable (one-dimensional) axes only. It may be extended simply to M non-separable axes so long as the indexing vectors for each of the M axes are separable. M coordinate values are required for each of the possible index positions. Therefore, the coordinates will be in a single $(1 + M)$-dimensional array. This *coordinate array* will have dimensions $(M, K_1, K_2, \ldots K_M)$, where K_m is the maximum value of the index on axis $m + 1$ of the coordinate array. For simplicity, degenerate axes are forbidden; therefore, $K_m > 1$. The indexing vectors for each of the M axes each contain a one-dimensional array of length K_m.

The data in the indexing vectors must be monotonically increasing or decreasing, although two adjacent index values in the vector may have the same value. However, it is not valid for an index value to appear more than twice in an index vector, nor for an index value to be repeated at the start or at the end of an index vector. Furthermore, repeated index values are allowed only in the case of one-dimensional separable axes. (See the following section for a discussion of the reasons for these limitations and how interpolation is done in such cases; see also Sect. 6.2.3 for an example of equal index values and further discussion of interpolation.) Application programs must not sort the index and coordinate arrays since this makes the relative order of the two equal index values indeterminate. The requirement for monotonic index values should eliminate any need for sorting. Note that it does not imply any ordering of the values in the coordinate array.

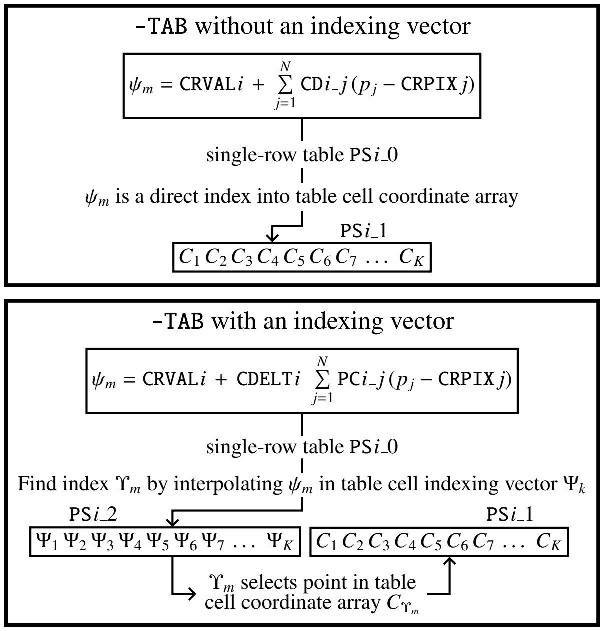


Fig. 6. -TAB logic flow with and without an indexing vector. The coordinate array subscript m associated with intermediate world coordinate axis i is specified with keyword PVi_3. In the case of an independent -TAB axis it would have value 1.
Note that ψ_m is computed either with the CDi_j form of the linear transformation matrix or the PCi_j plus CDELTi form.

6.1.2. Computing -TAB coordinate values

The indexing algorithm is illustrated schematically in Fig. 6 for the one-dimensional case. In the general case, to determine the M non-separable coordinate values C_m, one first determines the M indices, ψ_m for addressing the appropriate location in the table. If intermediate world coordinate axis i is associated with the m^{th} axis in the coordinate array, one evaluates

$$\psi_m = x_i + \mathtt{CRVAL}ia, \tag{87}$$

where x_i is computed following the prescriptions of Eq. (1). Using linear interpolation, if necessary, in the indexing vector for intermediate world coordinate axis i, one determines the location, Υ_m, corresponding to ψ_m. Then the coordinate value, C_m, of type specified by the first four characters of CTYPEia, is that at location $(m, \Upsilon_1, \Upsilon_2, \ldots \Upsilon_M)$ in the coordinate array, again using linear interpolation as needed.

In detail, the algorithm for computing Υ_m, and thence C_m, is as follows. Scan the indexing vector, $(\Psi_1, \Psi_2, \ldots)$, sequentially starting from the first element, Ψ_1, until a successive pair of index values is found that encompass ψ_m (i.e. such that $\Psi_k \leq \psi_m \leq \Psi_{k+1}$ for monotonically increasing index values or $\Psi_k \geq \psi_m \geq \Psi_{k+1}$ for monotonically decreasing index values for some k). Then, when $\Psi_k \neq \Psi_{k+1}$, interpolate linearly on the indices

$$\Upsilon_m = k + \frac{\psi_m - \Psi_k}{\Psi_{k+1} - \Psi_k}, \tag{88}$$

which yields an index into the coordinate array. However, if $\Psi_k = \Psi_{k+1} (= \psi_m)$ then the result is undefined.

In the case where an index value is equal to ψ_m, the algorithm above will find the interval $\Psi_k < \psi_m = \Psi_{k+1}$ for monotonically increasing index values or $\Psi_k > \psi_m = \Psi_{k+1}$ for monotonically decreasing index values, except when $\psi_m = \Psi_1$. Since two consecutive index values may be equal, the index following the matched index must be examined. If $\Psi_{k+2} = \Psi_{k+1} = \psi_m$ (or $\Psi_2 = \Psi_1 = \psi_m$), then Υ_m and C_m are undefined.

Linear interpolation via Eq. (88) applies for ψ_m in the range Ψ_1 to Ψ_K inclusive. Outside this range, for $K > 1$, linear extrapolation is allowed for values of ψ_m such that $\Psi_1 - (\Psi_2 - \Psi_1)/2 \leq \psi_m < \Psi_1$ or $\Psi_K < \psi_m \leq \Psi_K + (\Psi_K - \Psi_{K-1})/2$ for monotonic increasing index values, and for $\Psi_1 + (\Psi_1 - \Psi_2)/2 \geq \psi_m > \Psi_1$ or $\Psi_K > \psi_m \geq \Psi_K - (\Psi_{K-1} - \Psi_K)/2$ for monotonic decreasing index values. Extrapolation is also allowed for $K = 1$ with ψ_m in the range $\Psi_1 - 0.5 \leq \psi_m \leq \Psi_1 + 0.5$ (noting that Ψ_1 should be equal to 1 in this case) whence $\Upsilon_m = \psi_m$.

The value of Υ_m derived from ψ_m must lie in the range $0.5 \leq \Upsilon_m \leq K + 0.5$. These extrapolation limits permit assignment of coordinates to the regions of the pixels on the boundary of the array which are outside of the centers of the boundary pixels but within the conceptual "edge" of the boundary pixels. In the case of a single separable coordinate with $1 \leq k \leq \Upsilon_m < k+1 \leq K$, the coordinate value is given by

$$C_m = C_k + (\Upsilon_m - k)\,(C_{k+1} - C_k)\,. \tag{89}$$

For Υ_m such that $0.5 \leq \Upsilon_m < 1$ or $K < \Upsilon_m \leq K + 0.5$ linear extrapolation is permitted, with $C_m = C_1$ when $K = 1$.

Conceptually, to compute the change in coordinate value between Ψ_k and Ψ_{k+1}, in the case when $\Psi_{k+2} = \Psi_{k+1}$ and/or $\Psi_{k-1} = \Psi_k$, difference the two values of C_m obtained for $\psi_m = \Psi_{k+1} - \epsilon$ and $\psi_m = \Psi_k + \epsilon$ in the limit that ϵ goes to zero. In practice, the computation is straightforward and it is not necessary to take limits.

The inverse computation, in which one determines a ψ_m given a coordinate C_m, is relatively straightforward, at least in the case of a single separable coordinate. One scans the coordinate vector from the start looking for a pair of values that encompass C_m. Having found a pair, one must then check the indexing vector. If the two index values are unequal, then the correct pair has been found. Otherwise the search must be continued. When a correct coordinate pair has been found, the inverse of Eq. (89) is applied to determine Υ_m. The inverse of Eq. (88) then yields ψ_m.

It is understood that the general keywords CUNITma, CRDERma, and CSYERma apply to the output coordinate C_m rather than the -TAB coordinate keywords CRVALia et al. Similarly, if the table lookup determines celestial coordinates, the general keywords RADESYSa and EQUINOXa apply to the output celestial coordinates rather than the input -TAB coordinates.

6.1.3. -TAB implementation, parameters, and requirements

Standard FITS binary tables extensions (XTENSION = 'BINTABLE') will be used to hold the coordinate array in a single cell of a column of a one-row table. The length of this array is given in the FITS table header by the repeat count in the TFORMn keyword, where n is the number of the

Table 7. Parameter keywords used for the `-TAB` algorithm.

Keyword	Use	Default
PS*i*_0*a*	table extension name (EXTNAME)	—
PV*i*_1*a*	table version number (EXTVER)	1
PV*i*_2*a*	table level number (EXTLEVEL)	1
PS*i*_1*a*	column name for the coordinate array (TTYPEn_1)	—
PS*i*_2*a*	column name for the indexing vector (TTYPEn_2)	no index
PV*i*_3*a*	axis number (m) in array PS*i*_1*a*	1

column containing the coordinate array. The dimensions of the coordinate array will be given in the FITS table header by the keyword TDIM*n* set to '(M, K_1, K_2, ... K_M)', where n is the column containing the coordinate array. Note in particular that if $M = 1$ the coordinate array is formally *two-dimensional* though the first axis is degenerate. The repeat count in the TFORM*n* keyword value is the product of M and all the K_m. The indexing vectors for each of the M axes, if present, will occupy separate columns, each containing a one-dimensional array of length K_m.

The BINTABLE extension containing the coordinate array must be in the same FITS file as the data that reference it.

The parameters required by -TAB are the table extension name (EXTNAME), the table version number (EXTVER), the table level number (EXTLEVEL), the column name for the coordinate array (TTYPEn_1), the column name for the indexing vector (TTYPEn_2), and the axis number m associated with intermediate world coordinate axis i in the coordinate array. The keywords used for this purpose are PS*i*_0*a*, PV*i*_1*a*, PV*i*_2*a*, PS*i*_1*a*, PS*i*_2*a*, and PV*i*_3*a*, respectively. These are summarized in Table 7. For images, PS*i*_0*a* has no default; for tables a missing or blank PS*i*_0*a* is taken to be the current table (see below). PV*i*_1*a* and PV*i*_2*a* both have a default value of 1. PS*i*_1*a* never has a default; it must be present and be assigned a value actually occurring in table PS*i*_0*a*. If PS*i*_2*a* is missing or has a value of all blanks, the indexing vector is taken to be a list of integers from 1 to K_m and ψ_m becomes a direct index of axis PV*i*_3*a* in the array specified by PS*i*_1*a*. If PV*i*_2*a* is present and assigned a value, then that value must actually occur in table PS*i*_0*a*. Note that the values given to PS*i*_1*a* and PS*i*_2*a* are not case sensitive since the FITS Standard (Hanisch et al. 2001) states that "String comparisons with the values of TTYPE*n* keywords should not be case sensitive."

The use of -TAB for M related axes requires the header to specify the array PS*i*_1*a* to be the same for each of the M axes. The dimensions of this array must be given in the header as $(M, K_1, K_2, \ldots, K_M)$. These dimensions determine the maximum range of the array index (1 to K_m) for each axis. The indexing vectors for each of the M axes, if present, will occupy separate columns, each containing a one-dimensional array of length K_m. The M values of PV*i*_3*a* must account for all M axes. If any of these conditions are not met, the result is undefined.

If a FITS file contains multiple XTENSION HDUs (header-data units) with the specified EXTNAME, EXTLEVEL, and EXTVER, then the result of the WCS table lookup is undefined. If the specified FITS BINTABLE contains no column, or multiple columns, with the specified TTYPE*n*, then the result of the WCS table lookup is undefined. The specified FITS BINTABLE must contain only one row.

No units conversions are to be performed. CUNIT*i* must be the same as TUNIT*n* of the binary table, where n is the column number corresponding to PS*i*_1*a*.

We recommend strongly that the value chosen for EXTNAME always begin with the four letters WCS-. If this is done, generic programs will recognize the table as part of the WCS portion of the data and will be less likely to separate the table from the rest of the WCS.

The -TAB implementation is more complicated than most other WCS conventions because the coordinate system is not completely defined by keywords in a single FITS header. Software that supports -TAB must be able to gather all the necessary WCS parameters that are in general distributed over two FITS HDUs and in the body of the WCS extension table. The producers of FITS data products should consider the capabilities of the likely recipients of their files when deciding whether or not to use the -TAB convention, and in general should use it only in cases where other simpler WCS conventions are not adequate.

6.1.4. -TAB usage in tables

Binary table extensions containing array data columns may need a table-lookup function for coordinate values. Seemingly the most convenient form of -TAB would be one in which the -TAB array(s) are also table cells in the same row as the data array. However, a separate coordinate table would be more economical if it applied to data arrays in multiple rows. The -TAB keywords for such tables are *i*CTYP*na*, *i*S*n*_0*a*, *i*V*n*_1*a*, *i*V*n*_2*a*, *i*S*n*_1*a*, and *i*S*n*_2*a*, where n is the column number of the array of data and i is the intermediate world coordinate axis number. If *i*S*n*_0*a* is missing or blank, the present binary table is taken to be the coordinate lookup table. In that case, the coordinate array(s) are taken to be single-cell arrays in the same row as the data array and keywords *i*V*n*_1*a* and *i*V*n*_2*a* are ignored.

Strictly speaking, the -TAB representation is not required for binary or ASCII table extensions containing only one data value per cell because each of the coordinate values associated with the datum may be stored in separate columns. However, -TAB does provide a convenient method of solving one of the problems pertaining to such tables, namely identifying the data column. If the table contains a column of data and another one or more columns of coordinate values pertaining to those data, then one can define the coordinates of the data column as being of type -TAB. In this case, the critical keywords are *i*CTYP*na* to declare the coordinate axis type and *i*S*n*_1*a* to identify the corresponding coordinate column. Since the data value column contains only one value, the coordinate column should only contain one value. Then, the usual coordinate keywords (*j*CRPX*n* / *j*CRP*na*, *ij*PC*na*, *ij*CD*na*, *i*CDLT*n* / *i*CDE*na*) may be omitted since their defaults yield the desired result that output index pixel equals the input data pixel (both of which are 1 in the assumed case). Note too that tables of this type — one value per table cell — may be in either ASCII or binary table form.

6.2. -TAB examples

To illustrate the use of -TAB with two non-separable axes, let us consider a specific example of a four-dimensional image array. Assume that axes 1 and 3 are handled either linearly or by one of the non-linear single-axis cases (including -TAB). However, assume that axes 2 and 4 require table lookup in a mutually dependent fashion. Thus $i = 2$ for $m = 1$ and $i = 4$ for $m = 2$ and the required keywords are

```
PS2_0 = 'WCS-TAB'      PS4_0 = 'WCS-TAB'
PV2_1 = 1              PV4_1 = 1
PV2_2 = 1              PV4_2 = 1
PS2_1 = 'COORDS'       PS4_1 = 'COORDS'
PS2_2 = 'INDEX2'       PS4_2 = 'INDEX4'
PV2_3 = 1              PV4_3 = 2
```

where the first four keyword pairs must match exactly and the last two pairs must have different values.

A real example that could use this particular algorithm is represented by a spectral image of a portion of the sky taken with a single radio telescope. The observer commands the telescope to point at a regular grid of coordinates, but, due to wind loading and other pointing errors, the telescope achieves the commanded positions only approximately. The actual celestial coordinates observed are, however, accurately measured. These data could then be represented usefully as a two-dimensional array of spectra, but accurate celestial coordinates for each spectrum could be found only by a 2-dimensional table lookup.

Since there has been no generally agreed upon FITS format for spectral data with explicit wavelengths assigned to each pixel, data providers have resorted to defining their own formats. Examples from the Hubble Space Telescope and the Very Large Array are shown below, recast into the -TAB algorithm.

6.2.1. -TAB examples: HST data

Two types of Hubble Space Telescope data serve as examples of the -TAB algorithm. The two cases illustrate the evolution from simple images to the more powerful constructs provided by FITS extensions and binary tables. The purpose of discussing these formats is not to explain the HST formats, but to illustrate a couple of types of data that can be represented by the -TAB algorithm. Therefore, some details of these formats are ignored.

The early HST spectrographs, FOS and GHRS, adopted a format based only on simple FITS images. The basic concept is that the spectral flux values are given in one image and the vacuum wavelengths, in Angstroms, are given as the pixel values in another image. The two are associated by filenames. The filenames have the same basename, but different filename extensions for an exposure. To represent these spectra with a FITS WCS based on the -TAB method, while continuing to use an image representation for the spectral fluxes, one replaces the separate wavelength image with a table extension.

Table 8. WCS keywords in the spectral image for FOS and GHRS example.

CTYPE1	= 'WAVE-TAB'	/	Coordinate type
CUNIT1	= 'Angstrom'	/	Coordinate units
PS1_0	= 'WCS-TAB'	/	Coord table extension name
PS1_1	= 'WAVELENGTH'	/	Coord table column name

Table 9. WCS keywords in the STIS table for two of the spectral columns.

1CTYP4	= 'WAVE-TAB'	/	Coordinate type
1CUNI4	= 'Angstrom'	/	Coordinate units
1S4_1	= 'WAVELENGTH'	/	Coords for gross spectrum
1CTYP5	= 'WAVE-TAB'	/	Coordinate type
1CUNI5	= 'Angstrom'	/	Coordinate units
1S5_1	= 'WAVELENGTH'	/	Coords background spectrum

Table 8 shows the minimum WCS keywords required in the primary image header. The PS1_0 keyword value defines the coordinate table to be in a binary table extension with name 'WCS-TAB'. The coordinate table extension consists of one column, named 'WAVELENGTH', and one row. The array values and length correspond to the original image format.

Other WCS keywords defined in Paper I would also be included as needed. The intermediate coordinate transformation must produce indexing values equal to the pixel coordinate in the image. Since the defaults for these keywords are defined in Paper I to produce pixel coordinates, it is not strictly necessary to include these keywords.

The second generation HST spectrograph, STIS, adopted a binary table format very close to one of those defined for the -TAB algorithm. In the STIS one-dimensional extracted echelle format each exposure is stored in a separate binary table. The extracted data from an exposure consist of a number of echelle orders each of which is stored in a separate row. The wavelengths and various associated "spectra" are array elements in different columns. An associated spectrum is an array of a single type data quantity such as fluxes, errors, or data quality flags.

To convert this format to the -TAB representation only requires adding the 1CTYPn and 1Sn_1 keywords to specify the coordinate type and coordinate column name for each spectral column n. There are six types of spectral quantities for each order: gross, background, net, flux, error, and data quality flag. So there must be six pairs of keywords. Since the same wavelength column applies to each of the spectra, the keyword values are repeated six times, but with different column numbers in the keywords. Table 9 shows the WCS keywords for two of the spectral columns. The 1Sn_0 keywords are omitted, signifying that the coordinate array column is in the same table as the spectra.

As noted for the FOS and GHRS format, additional WCS keywords are included as needed and the intermediate coordi-

Table 10. Sample coordinate table. Each displayed column is actually a one-cell array of the (single-row) table.

iSn_2	iSn_1
1	ν_1
N_1	$\nu_1 + (N_1 - 1)\,\delta_1$
$N_1 + 1$	ν_2
$N_1 + N_2$	$\nu_2 + (N_2 - 1)\,\delta_2$
$N_1 + N_2 + 1$	ν_3
$N_1 + N_2 + N_3$	$\nu_3 + (N_3 - 1)\,\delta_3$
...	...
$\sum_1^{L-1} N_\ell + 1$	ν_L
$\sum_1^{L} N_\ell$	$\nu_L + (N_L - 1)\,\delta_L$

nate transformation values may be absent since this defines an indexing by pixel coordinate. In the STIS data, the units of the wavelength array are vacuum Angstroms so the 1CUNIn keyword would be required.

6.2.2. -TAB example: radio interferometry

Radio interferometry data are now best represented in binary tables with each row representing the visibility at one time on one antenna pair with a vector of data representing the complex values for all polarizations and frequencies observed. These telescopes often allow the user to observe N_ℓ regularly-sampled spectral channels starting from L arbitrarily chosen intermediate frequencies (IFs). Furthermore, this spectral data pattern may be used with one receiver and, a few minutes later, with other receivers at very different frequencies. Because of the great flexibility in the choice of the IFs and receivers, the frequencies present in the data may only be described by a table. Because that description is very repetitive, we choose to put it in a table separate from the visibility table.

We can construct a single column table of all $\sum_{\ell=1}^{L} N_\ell$ frequencies observed. However, by using a second column for an indexing vector, we can take advantage of the regular sampling of the N_ℓ spectral channels. If we set iCRVLn, iCDLTn, and iCRPXn (for binary table format) or CRVALi, CDELTi, and CRPIXi (for random groups and image formats) to one, then Table 10 may represent the actual frequencies, where ν_ℓ is the frequency of channel 1 in IF ℓ and δ_ℓ is the increment in frequency between the regularly sampled channels of IF ℓ. This example is shown graphically in Fig. 7.

6.2.3. -TAB examples: multiple exposures and sampling

In general, the FITS WCS papers do not consider the process by which the physical effects represented by the values in the FITS array have been measured. The measurable quantity may have been sampled over regions small enough to be considered as points in a widely separated grid of world coordinates. Alternatively, the measurable quantity may have been integrated across overlapping adjacent regions. The FITS WCS

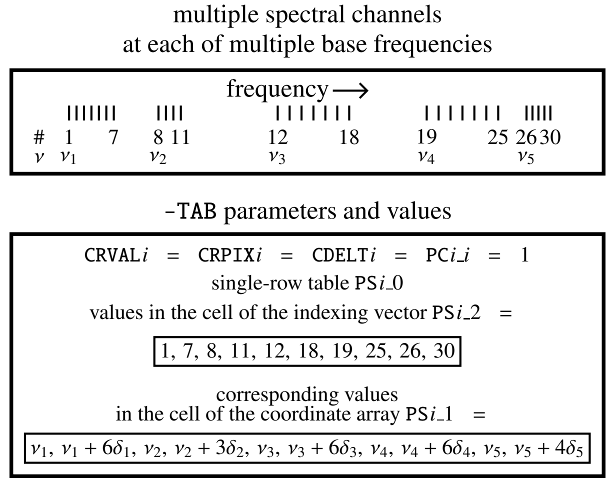


Fig. 7. Example taken from radio interferometry using -TAB with an indexing vector. The FITS keywords shown are suitable for the random groups format. The observation is made at a number of frequencies, with a number N_ℓ of regularly spaced δ_ℓ spectral channels beginning from each of a number of arbitrary base frequencies ν_ℓ. The use of an indexing vector reduces the number of values in the table from one array of 30 to two arrays of 10 each. In a real case the number of spectral channels would be significantly larger, making the space savings significant. In this example, pixel $p_j = 6$ produces $\psi_m = 6$. This lies at $\Upsilon_m = 1\frac{5}{6}$. The resulting coordinate is then $\frac{1}{6}\nu_1 + \frac{5}{6}(\nu_1 + 6\delta_1)$ which, as one would expect, equals $\nu_1 + 5\delta_1$. Note that this example involves only a single independent -TAB axis, so that PVi_3 $= m \equiv 1$.

papers do not indicate whether it is valid to presume that the values in adjacent array elements represent physically adjacent quantities, nor do they indicate whether it is valid to interpolate the WCS across the extent of an array element.

Some world coordinate axes are defined to have integral values with conventional meanings (e.g. COMPLEX, STOKES). Interpolating a world coordinate value across the extent of a FITS array element along the direction of an axis which is conventionally defined to be integral would clearly be inappropriate. Nevertheless, for other types of world coordinate axes in traditional FITS arrays, the adjacent elements in the NAXIS-dimensional array can often be presumed to represent adjacent locations in a measurement continuum. This adjacency is relevant, for example, to image display programs when the image is magnified so that one FITS array element extends across multiple display pixels. In such cases a FITS display program may attempt to interpolate WCS values across the extent of an array element. With CTYPEk values other than -TAB this is commonly done by linear interpolation between the presumably adjacent array elements. However for coordinate axes where the CTYPEk is described by the $xxxx$-TAB non-linear algorithm there can be no presumption of adjacency.

Notwithstanding these caveats, the use of -TAB does permit specification of changes in the coordinate value across the extent of a single array element. Consider a sequence of two-dimensional images of the sky which have been re-gridded such that the pixel array for each image shares the same 2-D WCS for celestial coordinates. A single 3-D FITS array can hold such a sequence of 2-D images. If the same instrument

Table 11. Top: keywords in main FITS header for `-TAB` example, bottom: table keywords and sample arrays for that example.

```
123456789 123456789 123456789 123456789 123456789 123456789 123456789 123456789
NAXIS   =                    3 / 2-d celestial coord, 1-d multiple exposure
NAXIS3  =                    4 / radio, infrared, optical, x-ray
WCSAXES =                    4 / multiple exposures in bandpass and time
WCSNAME = 'Multi-wavelength, multi-epoch'
CTYPE3  = 'WAVE-TAB'           / spectral axis by table look-up
CTYPE4  = 'TIME-TAB'           / temporal axis by table look-up
CNAME3  = 'Bandpass'
CNAME4  = 'Observation Date'
CUNIT3  = 'm'                  / wavelengths in meters
CUNIT4  = 'a'                  / observation dates in years
CRPIX3  =                  0.5 / spectral bin initial edge of initial pixel
CRPIX4  =                    1 / temporal bins are at begin/end time
CDELT3  =                    1 / spectral index array is one unit per bin
CDELT4  =                    1 / temporal index array is one unit per bin
CRVAL3  =                  0.5 / spectral reference is center of radio exposure
CRVAL4  =                    0 / temporal reference is start of radio exposure
PC3_3   =                    1 / 3rd pixel axis increments 3rd image coord axis
PC3_4   =                    0
PC4_3   =                    1 / 3rd pixel axis increments 4th image coord axis
PC4_4   =                    1 / 4th pixel axis is degenerate (1 point)
PS3_0   = 'WCS-table'
PS3_1   = 'WaveCoord'
PS3_2   = 'WaveIndex'
PS4_0   = 'WCS-table'
PS4_1   = 'TimeCoord'
PS4_2   = 'TimeIndex'
```

```
EXTNAME = 'WCS-table'
EXTVER  =                    1
EXTLEVEL=                    1

                        TUNIT2 = 'm'                                    TUNIT4 = 'a'
TFORM1 = '8E'           TFORM2 = '8D'           TFORM3 = '8E'           TFORM4 = '8D'
TTYPE1 = 'WaveIndex'    TTYPE2 = 'WaveCoord'    TTYPE3 = 'TimeIndex'    TTYPE4 = 'TimeCoord'
--------------------    --------------------    --------------------    --------------------
 (0.5,                   (0.21106114,            (0.0,                   (1997.84512,
   1.5,                    0.21076437,             1.0,                    1997.84631,
    1.5,                    2.0E-6,                 1.0,                    1993.28451,
     2.5,                    2.2E-6,                 2.0,                    1993.28456,
      2.5,                    500.0E-9,               2.0,                    2001.59234,
       3.5,                    650.0E-9,               3.0,                    2001.59239,
        3.5,                    1.24E-9,                3.0,                    2002.18265,
         4.5)                    2.48E-9)                4.0)                    2002.18301)
```

produced each of the images on different dates then the third dimension might be purely temporal. Alternatively, if different instruments (radio, IR, optical, X-ray) produced each of the images then the third dimension might be spectral. In practice, however, images in such a spectral 3-D image will typically also have different observation dates, so the image will actually span four dimensions of world coordinates.

Akin to the example of a long slit spectrum in Paper II, the WCS of this four-dimensional case can be represented by the 3-D array. This fourth dimension could be represented with `NAXIS = 4` and `NAXIS4 = 1`, or with `NAXIS = 3` and `WCSAXES = 4`. Whether the sequence of images is temporal or spectral, it is certain that the exposure duration or the spectral bandpass have a non-zero extent and the `-TAB` representation does provide the ability to communicate the start and end times of an observation, or the minimum and maximum wavelengths of an observation. For this example the WCS keywords are shown in Table 11, where the keywords relating to axes 1 and 2 are omitted because they are simply celestial coordinates as described by Paper II. In this example, note the clever use of the PC matrix to cause the fourth coordinate axis to depend on the third pixel axis, but not the degenerate fourth pixel axis. FITS writers can use details such as these to communicate how

Table 12. Recognized values for SPECSYS*a*, SSYSOBS*a*, and SSYSSRC*a*.

Value	Definition	Magnitude	ℓ	b	Reference
TOPOCENT	Topocentric	0.0 km s^{-1}	–	–	
GEOCENTR	Geocentric	0.5	–	–	
BARYCENT	Barycentric	30	–	–	Stumpff (1980)
HELIOCEN	Heliocentric	30	–	–	Stumpff (1980)
LSRK	Local standard of rest (kinematic)	20	† 56	23	Delhaye (1965), Gordon (1976)
LSRD	Local standard of rest (dynamic)	16.6	53	25	Delhaye (1965)
GALACTOC	Galactocentric	220	90	0	Kerr & Lynden-Bell (1986)
LOCALGRP	Local Group	300	90	0	de Vaucouleurs (1976)
CMBDIPOL	Cosmic microwave background dipole	368	263.85	48.25	Bennett et al. (2003)
SOURCE	Source rest frame	any	–	–	

† LSRK is normally quoted as right ascension 18 hours, declination 30 degrees (1900).

a display program might provide meaningful world coordinates at locations across an array element.[4]

In order to provide coordinate values across the entire array element, the coordinate values on the boundaries between the array elements are multiply defined. Such a result should be expected for array elements representing non-adjacent physical values. It is evident in this case that the order of the index arrays should not be permuted during the coordinate lookup. The lower portion of Table 11 shows a subset of the content of the table(s) referred to by PS3_0 and PS4_0. The four columns in the figure are meant to illustrate the association of table header keywords and table values. In the FITS file, the keywords would appear in the table header along with many other keywords while the four arrays of values would appear in four cells of the one-row binary table.

The indexing vectors in the table of Table 11 illustrate the reason to allow two values within the indexing vector to be identical. The rules for the linear interpolation within such vectors are clear. One must select the two index locations having values immediately surrounding ψ_m. Thus in the present example, if the time coordinate has $\psi_m = 1.1$, the third and fourth index and coordinate values are interpolated. The output time coordinate is

$$1993.28451 + \frac{1.1 - 1.0}{2.0 - 1.0}\,(1993.28456 - 1993.28451)\cdot$$

The result of the table look up is undefined if the ψ_m is equal to the repeated value in the indexing vector.

7. Coordinate reference frames

Frequencies, wavelengths, and apparent radial velocities are always referred to some selected standard of rest (reference frame), and while they are measured, of necessity, in the observer's rest frame, they may also be corrected to some other frame.[5] The velocity correction is computed from the vector dot product of the direction vector and the relative velocity vector of the two reference frames. In addition, the corrections from topocentric, the frame in which the observations are usually made, to geocentric and then to barycentric or heliocentric are dependent on the dot product with time-variable velocity vectors. As a consequence, the "velocity" with respect to the reference frame depends on direction. Differential effects over a field of view may be important in some high precision applications; for example, they may amount to 5 km s^{-1} for the Local Group correction over a field of one degree. In another example, observations of Galactic CO over two-degree fields separated by two degrees failed to align by two spectral channels (Mangum et al. 2000).

If a frequency, wavelength or apparent radial velocity associated with an image plane has been corrected to some other frame by transforming the CRVAL*ka* et al. values, then differential errors may arise at points away from the reference point. For example, it is normal in radio astronomy to observe a field while holding constant the velocity of the reference point with respect to the local standard of rest. See Sect. 10.2 for a more detailed discussion of the complications inherent in such data.

Nonetheless, each image plane shares a constant topocentric frequency (or apparent radial velocity). The velocity (*and* frequency and wavelength) with respect to the local standard of rest is then a function of celestial coordinate within the image. In order to denote this, we introduce two character-valued keywords. The first,

SPECSYS*a* (character-valued)

describes the reference frame in use for the spectral-axis coordinate(s). The second,

SSYSOBS*a* (character-valued)

describes the spectral reference frame that is constant over the range of the non-spectral world coordinates and has default value 'TOPOCENT'. The recognized values are given in

[4] The precise interpretation of CTYPE*k* keywords that describe TIME is deferred to a possible future paper on temporal coordinates in FITS.

[5] Proper relativistic formulæ should be used; see Gibbs & Tao (1997) or Rindler (1982). Better yet, observers should consider using standard software described by Lindegren and Dravins (2003) to perform the conversion from topocentric coordinate to the standardized Barycentric Celestial Reference System.

Table 12. In this table, the magnitude column gives the approximate magnitude of the velocity vector that defines the particular reference frame and the ℓ and b columns give the standard Galactic longitude and latitude of the reference frame, Blaauw et al (1960). This table includes the traditional heliocentric system which the authors believe should be deprecated. It has frequently been misused as an alias for barycentric and would, if high accuracy is desired, require the observing date in order to allow conversions between it and the other systems. The table also includes the WMAP microwave background dipole as a velocity system. Since, as the data are refined over time, this system will change, its use at present should be restricted to situations for which it is particularly appropriate and the parameters used should be well commented. The radio-astronomy example described above has `SPECSYS = 'LSRK'` and `SSYSOBS = 'TOPOCENT'` indicating that the spectral axis is expressed as a kinematic Local Standard of Rest velocity, but that each right ascension by declination image plane is at a constant topocentric spectral coordinate.

Many spectrometers suffer from a significant variation in their spectral coordinate as a function of celestial coordinate. Such variations are the primary subject of Paper IV, while the concepts of this work refer to images produced by, or corrected to, an ideal spectrometer. In general we recommend that the frame corrections will get folded into the instrumental ones and that the image produced after the Paper IV corrections will have identical `SSYSOBS` and `SPECSYS`. This is not required, but would certainly simplify matters.

It is not appropriate for this work to define the parameters of the recognized rest frames. The parameters needed to compute geocentric frequencies/velocities from topocentric are the sidereal time (or Earth rotation angle) and the observatory location. The observing date is needed to convert from geocentric to barycentric coordinates. The new keywords

`MJD-AVG` (floating-valued)

and

`DATE-AVG` (character-valued)

are reserved to give a representative time for the whole observation, suitable for calculating proper motion, apparent place, local apparent sidereal time, velocity with respect to the center of the Earth and barycenter, etc. The `DATE-AVG` keyword shall be expressed in the manner described by Hanisch et al. (2001) for the `DATE-OBS` keyword. If both keywords are present, they must agree to adequate accuracy or the result is undefined. If both keywords are absent, `DATE-OBS` may be used to determine the representative time, although `DATE-OBS` refers to the beginning of an observation and hence is not entirely suitable. For high-precision applications, a time system to apply to `MJD-AVG` must be defined in the manner proposed by Bunclark & Rots (1997). Note also that, for long integrations, no single time is adequate for computing proper motion, sidereal time, and the like.

Astronomers have traditionally quoted telescope positions using terrestrial longitude, latitude, and height, but, for computation of topocentric velocity, these values are immediately converted into a geocentric Cartesian triple. A precise conversion from longitude, latitude, and height, however, requires knowledge of the geodetic datum in which they are expressed. There are about 1000 geodetic datums in use. These are described by the United States Department of Defense (2000) in a printed document also available over the Internet. Some of the datums deviate from geocentricity by over a kilometer.

Rather than burden FITS with these details of geodetic tradition, three new keywords

`OBSGEO-X` (floating-valued)
`OBSGEO-Y` (floating-valued)
`OBSGEO-Z` (floating-valued)

are reserved to define a representative location for the observation in a standard terrestrial reference frame. The location values should be right-handed, geocentric, Cartesian coordinates valid at the epoch of `MJD-AVG` or `DATE-AVG`. Position errors of several kilometers should have negligible impact on the computation of the diurnal velocity correction, so for this purpose geodetic accuracy is not a requirement. However, it should be possible to provide coordinates with an accuracy of 1m or better based on a recent satellite-based geodetic reference frame and we recommend that FITS writers do so. Although it is beyond the scope of this paper to define particular terrestrial reference frames or tectonic models, only frames based on the ITRS (McCarthy & Petit 2004) are suitable for high-precision applications. Most post-1980 reference frames, including GPS, agree with recent versions of the ITRF to about 0.1 m. Web sites

http://earth-info.nga.mil/GandG/
http://www.ngs.noaa.gov/TOOLS/XYZ/xyz.html

provide general geodetic information and a solution based on the outputs of almost any hand-held GPS unit, respectively. These references, together with information about the origin of the traditional telescope position, should allow the determination of the keyword values defined in this paper.

It may be helpful for the FITS writer to provide information on the relative radial velocity between the observer and the selected standard of rest in the direction of the celestial reference coordinate. The keyword

`VELOSYS`*a* (floating-valued)

is hereby reserved for this purpose; its units shall be $\mathrm{m\,s^{-1}}$. `CUNIT`*ka* is not used since the WCS version *a* might not be expressed in velocity units.

The frame of rest defined with respect to the source (`SOURCE`) is useful for comparing internal apparent radial velocities in objects at different redshifts. This allows the FITS writer to apply all the cosmological and other model-dependent corrections, leaving a coordinate local to the object. For this frame of rest, it is necessary to define the velocity with respect to some other frame of rest. The keywords

`ZSOURCE`*a* (floating-valued)

and

`SSYSSRC`*a* (character-valued)

Table 13. New spectral keywords including forms for use in tables. All table keywords except the coordinate name have the same form for the `BINTABLE` vector image format and the pixel list format.

Primary Array	BINTABLE	Pixel list	Type	Sect.	Keyword Description
`CNAME`*ia*	*i*`CNA`*na*	`TCNA`*na*	character	3.3	Coordinate axis name
`RESTFRQ`*a*	`RFRQ`*na*		floating	3.4.4	Line rest frequency (Hz)
`RESTFREQ`	—		floating	3.4.4	Line rest frequency alternate for primary array only
`RESTWAV`*a*	`RWAV`*na*		floating	3.4.4	Line rest wavelength in vacuum (m)
`SPECSYS`*a*	`SPEC`*na*		character	7	Spectral reference frame (from Table 12)
`SSYSOBS`*a*	`SOBS`*na*		character	7	Spectral reference frame fixed during observation (from Table 12)
`MJD-AVG`	`MJDA`*n*		floating	7	Average date of observation (Julian date−2400000.5)
`DATE-AVG`	`DAVG`*n*		character	7	Average date/time of observation
`OBSGEO-X`	`OBSGX`*n*		floating	7	Observation X (m)
`OBSGEO-Y`	`OBSGY`*n*		floating	7	Observation Y (m)
`OBSGEO-Z`	`OBSGZ`*n*		floating	7	Observation Z (m)
`VELOSYS`*a*	`VSYS`*na*		floating	7	Radial velocity wrt standard of rest (m s^{-1})
`ZSOURCE`*a*	`ZSOU`*na*		floating	7	Redshift of source for `SOURCE` cases (unitless)
`SSYSSRC`*a*	`SSRC`*na*		character	7	Spectral reference frame for `SOURCE` cases (from Table 12)
`VELANGL`*a*	`VANG`*na*		floating	A	Angle of true velocity from tangent to line of sight

are hereby reserved; they should be used to document the adopted value for this systemic velocity of the source. `ZSOURCE`*a* is a unitless redshift; `SSYSSRC`*a* specifies the reference frame for this parameter and may have any value in Table 12 except `SOURCE`. The new floating-valued and character-valued keywords hereby reserved are listed[6] in Table 13 along with the other keywords defined in this paper.

8. Alternate FITS image representations: pixel list and vector column elements[7]

The use of coordinate keywords to describe a multi–dimensional vector in a single element of a FITS binary table (Cotton et al. 1995), and a tabulated list of pixel coordinates in a FITS ASCII (Harten et al. 1988) or binary table was discussed in Papers I and II. In this section, we extend those discussions to the keywords specific to spectral coordinates. This convention is considered an integral part of the full world coordinates convention.

8.1. Keyword naming convention

Table 13 lists the corresponding set of coordinate system keywords for use with each type of FITS image representation. The keywords defined in this paper and their allowed values are applicable to all three image storage formats. The following notes apply to the naming conventions used in Table 13:

– *a* is a 1-character coordinate version code and may be blank (primary) or any single uppercase character from `A` through `Z`.
– *n* is an integer table column number without any leading zeros (1–999).
– *i* is a one-digit integer axis number (1–9).

When using the `BINTABLE` vector image format, if the table only contains a single image column or if there are multiple image columns but they all have the same value for any of the keywords in Table 13 except `CNAME`*ia*, then the simpler form of the keyword name, as used for primary arrays, may be used. For example, if all the images in the table have the same location, then one may use one set of `OBSGEO-X`, `OBSGEO-Y`, and `OBSGEO-Z` keywords rather than multiple `OBSGX`*n*, `OBSGY`*n*, and `OBSGZ`*n* keywords. If both forms of these keywords appear, then column-specific values are applied to the specified columns and the non-specfic values are applied to all other columns. (Note, however, that the WCS keywords defined for tables in Papers I and II must always be specified using the more complex keyword name with the column number suffix and the axis number prefix.)

9. Spectral coordinate variation with other coordinates

There are instruments having spectral coordinates that are a function of other coordinates in the resulting image, e.g. celestial position. For example, optical astronomers use imaging Michelson interferometers and scanning Fabry-Perot instruments to produce three-dimensional images with two celes-

[6] The frame and velocity keywords in Table 13 apply to terrestrial observers only. Observations from spacecraft require additional description of their motions – either a position and velocity or an ephemeris. The required keywords and/or tables for this are beyond the scope of this paper.

[7] Suggested by contributions from William Pence, Arnold Rots, and Lorella Angelini of the NASA Goddard Space Flight Center, Greenbelt, MD 20771.

tial and one spectroscopic coordinate. Because the path length through such instruments increases off the optical axis of the telescope, the observed wavelength is a function of celestial coordinate in each plane of the image cube. In general, this is a subject reserved for Paper IV since such instruments almost always have lesser distortions that need to be parameterized and corrected along with the path length correction.

One particularly difficult case should be mentioned here. Objective prism instruments produce data in which the image of each object in the field has been spread out into a spectrum. Thus the output image is the convolution of the sky with both a spatial point-spread-function and a spectral dispersive point-spread function, both in the two dimensions of the output image. The methods used to specify coordinates described in the present paper and in Papers I and II are unable to handle such complex data. However, rather than ignore the world coordinates entirely, writers of such images should consider providing the celestial coordinates that would be correct for the image if it had been taken at a single representative wavelength. That wavelength would be specified in a third WCS axis. Alternate axis descriptions could then be used to specify the same information for other wavelengths. Multi-order echelle spectrographs also produce images in which multiple spectral orders overlap each other as well as the celestial axes and, hence, their coordinates cannot be described by the methods of Papers II and III.

Analysis of both objective prism and echelle spectrograph images can produce tables of spectra that are easily described by the present methods.

Table 14. Example image header.

```
123456789 123456789 123456789 123456789 123456789
SIMPLE  =                    T
BITPIX  =                  -32
NAXIS   =                    3
NAXIS1  =                 1024
NAXIS2  =                 1024
NAXIS3  =                   63
EXTEND  =                    T / Tables may follow
OBJECT  = '3C353'
TELESCOP= 'VLA'
DATE-OBS= '1998-09-29'         / Start obs
BUNIT   = 'Jy/beam'
EQUINOX =      2.000000000E+03
DATAMAX =      5.080558062E-01
DATAMIN =     -1.179690361E-01
WCSNAME = 'TopoFreq'
CTYPE1  = 'RA---SIN'
CRVAL1  =    2.60108333333E+02
CDELT1  =     -2.777777845E-04
CRPIX1  =                512.0
CUNIT1  = 'deg'
CTYPE2  = 'DEC--SIN'
CRVAL2  =   -9.75000000000E-01
CDELT2  =      2.777777845E-04
CRPIX2  =                513.0
CUNIT2  = 'deg'
CTYPE3  = 'FREQ'
CRVAL3  =    1.37835117405E+09
CDELT3  =      9.765625000E+04
CRPIX3  =                 32.0
CUNIT3  = 'Hz'
OBSGEO-X=     -1.601185365E+06 / [m]
OBSGEO-Y=     -5.041977547E+06 / [m]
OBSGEO-Z=      3.554875870E+06 / [m]
MJD-AVG =            51085.979
SPECSYS = 'TOPOCENT'
RESTFRQ =      1.420405752E+09 / [Hz]
RADESYS = 'FK5'

CNAME3Z = 'Barycentric optical velocity'
RESTWAVZ=          0.211061139 / [m]
CTYPE3Z = 'VOPT-F2W'
CRVAL3Z =         9.120000E+06
CDELT3Z =       -2.1882651E+04
CRPIX3Z =                 32.0
CUNIT3Z = 'm/s'
SPECSYSZ= 'BARYCENT'
SSYSOBSZ= 'TOPOCENT'
```

10. Example

10.1. Basic spectral-line header

Let us consider an example. The partial FITS header in Table 14 was produced for data from the Very Large Array. The FITS WCS standard states that all WCS keywords must be reproduced for the alternate descriptions, but those for non-spectral axes have been omitted for the sake of brevity and clarity. We have also updated the keywords to the standards of Papers I, II, and III. We will derive additional header keywords that could also have been written. These are shown in Table 15, again leaving out axes 1 and 2 of the alternate WCS versions.

The spectral-line cube is regularly sampled in frequency and the primary description is of a linear frequency axis (from the `CTYPE3` keyword value `'FREQ'`). `CRVAL3` and `CRPIX3` record that the 63-channel spectrum was centered with channel 32 at 1.37835117405 GHz, some 42 MHz from the rest frequency (`RESTFRQ` = 1.420405752 GHz) of neutral hydrogen (HI), the line under investigation. As indicated by the `SPECSYS` keyword, it is shown as if it had been observed at a constant topocentric frequency. For the moment, let us assume that the observations were made over a short interval and that this is correct.

As allowed by Paper I, we use alternate axis description letter codes in a mnemonic fashion. Using `Z` for optical, the first alternate axis description for the spectral axis reflects the observer's request that a Doppler shift be applied to the HI line frequency appropriate for a source with a barycentric optical-convention velocity of 9120 km s^{-1}. Combining Eqs. (30) and (6), we get

$$\nu = \nu_0/(1 + Z/c),$$

whence the barycentric HI rest frequency shifts to 1.37847122 GHz, some 120 kHz greater than the topocentric

Table 15. Additional keywords for example header.

```
123456789 123456789 123456789 123456789 123456789
VELOSYSZ=            26.108E+03 / [m/s]

CNAME3F = 'Barycentric frequency'
CTYPE3F = 'FREQ'
CRVAL3F =    1.37847121643E+09
CDELT3F =           9.764775E+04
CRPIX3F =                   32.0
CUNIT3F = 'Hz'
SPECSYSF= ‘BARYCENT'
SSYSOBSF= 'TOPOCENT'
VELOSYSF=            26.108E+03 / [m/s]

CNAME3W = 'Barycentric wavelength'
CTYPE3W = 'WAVE-F2W'
CRVAL3W =       0.217481841062
CDELT3W =       -1.5405916E-05
CRPIX3W =                   32.0
CUNIT3W = 'm'
SPECSYSW= ‘BARYCENT'
SSYSOBSW= 'TOPOCENT'
VELOSYSW=            26.108E+03 / [m/s]

RESTFRQR=      1.420405752E+09 / [Hz]
CNAME3R = 'Barycentric radio velocity'
CTYPE3R = 'VRAD'
CRVAL3R =    8.85075090419E+06
CDELT3R =       -2.0609645E+04
CRPIX3R =                   32.0
CUNIT3R = 'm/s'
SPECSYSR= 'BARYCENT'
SSYSOBSR= 'TOPOCENT'
VELOSYSR=            26.108E+03 / [m/s]

CNAME3V = 'Barycentric apparent radial velocity'
RESTFRQV=      1.420405752E+09 / [Hz]
CTYPE3V = 'VELO-F2V'
CRVAL3V =    8.98134229811E+06
CDELT3V =       -2.1217551E+04
CRPIX3V =                   32.0
CUNIT3V = 'm/s'
SPECSYSV= 'BARYCENT'
SSYSOBSV= 'TOPOCENT'
VELOSYSV=            26.108E+03 / [m/s]
```

frequency given by CRVAL3. The difference, of course, simply reflects the instantaneous relative velocity of the barycentric and topocentric reference frames at the time of observation. Using standard software (the STARLINK program rv), with the observatory position defined by OBSGEO-X, Y, Z, at the time indicated by MJD-AVG, and for the source at (J2000) right ascension and declination $17^{\mathrm{h}}20^{\mathrm{m}}26^{\mathrm{s}}$, $-00°58'30''$ indicated by CRVAL1 and CRVAL2, it may be verified that the topocentric correction amounts to $26.108\,\mathrm{km\,s^{-1}}$. Now this correction between reference frames is a true velocity and thus the relativistic formula, Eq. (8), should be used, whereupon the frequency recorded by CRVAL3 may be obtained. The frequency decreases as the observer moves away from the source position.

With CRVALZ set to $9.12 \times 10^6\ \mathrm{m\,s^{-1}}$ it remains to compute CDELT3Z. This is obtained by transforming CDELT3 from the topocentric to the barycentric frame with Eq. (8) expressed as

$$\Delta\nu = \Delta\nu_0 \frac{c - v}{\sqrt{c^2 - v^2}},$$

whence the frequency increment becomes 97.64775 kHz. Using Eq. (11) the corresponding wavelength increment is as recorded by CDELT3Z in Table 14. While SPECSYSZ is set to 'BARYCENT', SSYSOBSZ records the fact that the observation was actually made from the 'TOPOCENT' frame.

The correction for the observatory's motion due to the Earth's rotation and orbital motion with respect to the barycenter amounted to $26.108\,\mathrm{km\,s^{-1}}$ in the direction of the source. As illustrated in Table 15, this velocity shift can be represented in the FITS header with the new keyword VELOSYS.

The barycentric frequency may be used as another alternate axis description. The reference frequency and frequency increment for SPECSYSF = 'BARYCENT' have already been computed above. The relevant keywords have axis description F (for frequency) in Table 15.

The frequency description may be translated into a wavelength description simply using Eqs. (10) and (11), shown in Table 15 using axis description W (for wavelength).

The frequency description with respect to the BARYCENT may also be expressed simply as a radio velocity (also regularly sampled) using Eqs. (20) and (21). This is shown as axis description version R (for radio).

A apparent radial velocity description (version V for velocity) requires the use of Eq. (14) for the reference value and Eq. (15) for the increment.

10.2. Real-world complications

An assumption made throughout this paper, particularly in Sect. 7, is that data are provided with at least celestial coordinate information in addition to the spectral coordinates discussed here. The normal method of providing such data is with the full set of coordinate keywords including WCSAXES with a value of at least 3 (for two celestial and one spectral axis) even if only one celestial position was observed. For spectra in tables, it would be possible to have simple columns for the celestial coordinates, labeled as such. But, as discussed at the end of Sect. 6.1.4, this simple case may still use the full WCS nomenclature in order to associate the coordinate columns with the data column.

One important assumption made in Sect. 10.1 was that the topocentric velocity does not change significantly during the course of the observation. We now consider the consequences if this assumption is violated. In this we are interested in particular with three-dimensional "data cubes" containing multiple samples on each of two spatial axes and one spectral axis.

For some instruments, it may be possible to correct the data for each short exposure to a consistent WCS and then combine exposures to improve the sensitivity. For example, single-dish radio telescopes observe a single direction at a time, allowing the spectral axis in each observed spectrum to be scaled and resampled before it is added into a three-dimensional image. For this type of instrument it is a relatively simple matter to correct for topocentric motion.

However, other instruments observe multiple directions simultaneously and also routinely combine multiple exposures to form an image. While the error over a single exposure may be negligible, images made by aperture synthesis radio telescopes, for example, are often based on multiple exposures taken months apart. These exposures are usually made with different array configurations, which may change on a timescale of months, in order to sample the Fourier space adequately for imaging and deconvolution. Therefore, these data are affected not only by the rotation of the Earth but also by its motion around the Sun which usually is not negligible.

As a specific example, the actual observation on which Sect. 10.1 is based required an integration time of several hours. During this time the observing frequency was adjusted for the topocentric Doppler shift so that the barycentric frequency at the map center remained constant. This ensured that the spectral feature of interest, associated with a source near the map center, remained in the same spectral channel. However, this means that the primary spectral axis description, based as it is in the topocentric frame, is not strictly correct. It was used in this data as an aid to interferometric analysis, the resulting error being considered negligible, but strictly speaking a barycentric frequency reference frame, the `F` alternative WCS in the header, is the correct one.

Even so, the barycentric frame is only strictly correct at the map center since the topocentric Doppler correction varies across the field-of-view; the correction is computed as the dot product of the observatory's velocity vector and a unit direction vector which, of course, varies across the field. This gives rise to a differential error that is greatest when the velocity vector is orthogonal to the direction to the source and least when it is parallel or anti-parallel. In other words, the differential correction is greatest when the correction is zero, and least when the correction is greatest.

For diurnal rotation the worst case differential error occurs for a telescope at the equator observing a source on the local meridian. For an angular offset, γ, from the tracking center the maximum differential error is $v \sin\gamma$ which amounts to $8.7\,\mathrm{m\,s^{-1}}$ for $\gamma = 1°$. The worst case change in the differential error over the course of 12 hours occurs for an equatorial telescope observing a field at the north or south point on the horizon; it is $2v \sin\gamma$ or $17.5\,\mathrm{m\,s^{-1}}$ for $\gamma = 1°$. If data from exposures separated by several months are combined then the Earth's $30\,\mathrm{km\,s^{-1}}$ orbital velocity scales up the worst-case error to $1.0\,\mathrm{km\,s^{-1}}$ for $\gamma = 1°$. Note that, in the example of Sect. 10.1, the barycentric correction is close to the maximum value so the differential correction is nearly at its minimum.

The relevance of the `SSYSOBS`*a* keyword immediately becomes apparent in this context. It records the reference frame in which the frequency has no spatial variation; in other frames the reference frequency and frequency increment are only strictly correct at the reference point. While `SSYSOBS`*a* will most often indicate the topocentric frame, it is possible to regrid (resample) the spectral data by applying a position-dependent shift-and-scale so that `SSYSOBS`*a* has some other value.

Differential Doppler effects are often neglected for contemporary observations made with small angular fields or modest spectral resolution. However, instruments now under development will be capable of significantly greater sensitivity over wider fields-of-view and/or spectral resolution and hence will need to consider the effects discussed here.

11. Summary

The new keywords required for spectral coordinate systems are summarized in Table 13. `SPECSYS`*a*, `SSYSOBS`*a*, `SSYSSRC`*a*, and `ZSOURCE`*a* are used to specify velocity reference frames; `MJD-AVG`, `DATE-AVG`, `OBSGEO-X`, `OBSGEO-Y`, `OBSGEO-Z`, `VELOSYS`*a*, and `ZSOURC`*a* enable conversion between these standards of rest; and `RESTFRQ`*a*, `RESTFREQ`, and `RESTWAV`*a* define the spectral line for which velocities are measured. Variants of these keywords for use with tabular data are also defined.

These new keywords and their allowed values, along with new values for the standard header keywords formalized by Paper I, and the associated algorithms and methods introduced here allow an accurate description of spectral coordinates in FITS images. Wavelength, frequency, apparent radial velocity and oft-used quantities that are linearly related to them, such as redshift, energy, and radio velocity, may now be expressed as functions of pixel coordinate along an axis regularly sampled in wavelength, frequency, or apparent radial velocity.

A non-linear algorithm is also provided for spectral dispersers commonly used at optical wavelengths: gratings, prisms, and the combination of the two, grisms. Although developed for ideal dispersers, it was shown to apply quite well to real-world dispersers with suitable fine adjustment of the instrumental parameters.

The multi-dimensional `-TAB` table lookup algorithm developed for wavelength calibration will be useful for much more than just spectral axes. Several others of the methods introduced here are also completely general; logarithmic (`-LOG`) coordinates, and the `CNAME`*ia* keyword. The mathematical formalism described in Sect. 3.4 for constructing one-dimensional non-linear coordinate systems also has wide validity.

Appendix A: Relativistic space velocities

The discussion of Sect. 2, particularly Fig. 1, indicates the importance of transverse, as well as radial velocities, in spectroscopy at significant velocities. To be concrete, consider a relativistic jet emanating from a distant galaxy. In principle, the galaxy's systemic, cosmological redshift can be measured separately and used to correct the jet's observed redshift thereby providing its kinematic redshift (i.e. associated with a true velocity) in the reference frame of the galaxy.

Clearly knowledge of the velocity is fundamental in studying jet kinematics and dynamics. However, it can only be computed from the kinematic redshift if the jet's orientation angle is known. Note that the equations involving velocity in Table 3 are actually only correct if the transverse velocity is zero. For

Table A.1. Velocity equations using orientation and total velocity.

$$\nu = \nu_0 \frac{\sqrt{c^2 - v_s^2}}{c + v_s \sin\theta}, \tag{A.4}$$

$$\frac{d\nu}{dv_s} = -c\nu_0 \frac{v_s + c\sin\theta}{(c + v_s \sin\theta)^2 \sqrt{c^2 - v_s^2}}, \tag{A.5}$$

$$\lambda = \lambda_0 \frac{c + v_s \sin\theta}{\sqrt{c^2 - v_s^2}}, \tag{A.6}$$

$$\frac{d\lambda}{dv_s} = c\lambda_0 \frac{v_s + c\sin\theta}{(c^2 - v_s^2)^{3/2}}, \tag{A.7}$$

$$v_s = c\frac{-\chi^2 \sin\theta \pm \psi}{\chi^2 + \psi^2}, \tag{A.8}$$

where $$\chi = \frac{\nu}{\nu_0} = \frac{\lambda_0}{\lambda}, \tag{A.9}$$

and $$\psi = \sqrt{1 - \chi^2 \cos^2\theta}, \tag{A.10}$$

$$\frac{dv_s}{d\nu} = \frac{1}{\nu_0}\frac{dv_s}{d\chi}, \tag{A.11}$$

$$\frac{dv_s}{d\lambda} = \frac{-\lambda_0}{\lambda^2}\frac{dv_s}{d\chi}, \tag{A.12}$$

where $$\frac{dv_s}{d\chi} = -c\chi \frac{2\psi\sin\theta \pm (1 + \psi^2 \sin^2\theta)}{\psi(\chi^2 + \psi^2)^2}\,. \tag{A.13}$$

example, Eq. (12) is just a special case of Eq. (2). This is why the velocity is always referred to in the paper as being "apparent."

There are instances where the orientation angle may be inferred by geometry (e.g. by the observed tilt of an accretion disk) or by modelling. Defining the orientation angle as θ we have

$$v_r = v_s \sin\theta, \tag{A.1}$$

$$v_t = v_s \cos\theta, \tag{A.2}$$

$$v_s = \sqrt{v_r^2 + v_t^2}, \tag{A.3}$$

where v_r is the radial velocity, v_t is the transverse velocity, and $v_s \geq 0$ is the total space velocity. The orientation angle θ is defined so that $\theta = -90°$ is towards the observer, $\theta = 0$ is transverse, and $\theta = +90°$ is away from the observer. It is beyond the scope of this paper to describe how θ should be determined or what exactly it means for particular relativistic observers, other than that it resolves the kinematic velocity into radial and tangential components. With these definitions, the equations of Table 3 are given in Table A.1. Note that for $\chi > 1$ Eqs. (A.8) and (A.13) have two valid solutions for $\theta < 0$ but none otherwise. For $\chi \leq 1$ they have only one valid solution for any value of θ. Invalid solutions, i.e. those with $v_s < 0$, are the negative of the valid solution for $-\theta$.

There are other combinations that are of interest, including

$$\sin\theta = \frac{\nu_0\sqrt{c^2 - v_s^2}}{\nu v_s} - \frac{c}{v_s},$$

$$= \frac{\lambda\sqrt{c^2 - v_s^2}}{\lambda_0 v_s} - \frac{c}{v_s}\,.$$

It would perhaps be more instructive to express these sorts of equations in terms of the actual radial and transverse velocities. This is relatively straightforward, but the equations become even more messy and so we omit them here.

It has been proposed that we introduce a single keyword to express θ so that we may express velocities in terms of a true space velocity v_s rather than an apparent velocity v as done throughout this paper. However, astrophysical jets are thought to be conical in form, to interact along their boundaries with the external medium, and to change direction, even assuming helical shapes. Thus, there is no single angle θ that can normally be used to describe all directions in an image and one should probably not even assume that θ is independent of redshift in any one direction. For those cases in which a single angle may be appropriate, such as a spectrum of a spatially limited region in a relativistic jet, we reserve the keyword

`VELANGL`*a* (floating-valued)

with default value +90° to express the angle θ in degrees. All of this discussion has neglected non-velocity redshifts such as those due to the gravity of the black hole thought to be at the base of the astrophysical jet. Therefore, we choose to leave the full expression of internal velocities within a celestial object to those observers who are armed with both extensive observations and a detailed model. Apparent velocity is a suitable, semi-physical concept appropriate to a general FITS standard.

Acknowledgements. The authors wish to acknowledge constructive comments and contributions from the following: Wim Brouw (Australia Telescope National Facility), Bob Hanisch (Space Telescope Science Institute), Jonathan McDowell (Harvard University), William Pence (Goddard Space Flight Center), Arnold Rots (CfA Harvard University), William Thompson (Goddard Space Flight Center), Doug Tody (National Radio Astronomy Observatory), and Patrick Wallace (U.K. Starlink).

The `SLALIB` software library developed by Patrick Wallace was helpful in the preparation of Sect. 7 and should be of considerable help to users of conventions described in this work. It is available via `http://www.starlink.ac.uk/`.

The National Radio Astronomy Observatory is a facility of the (U.S.) National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

The Australia Telescope is funded by the Commonwealth of Australia for operation as a National Facility managed by CSIRO.

The National Optical Astronomy Observatory is a facility of the (U.S.) National Science Foundation operated under cooperative agreement by Associated Universities for Research in Astronomy, Inc.

UCO/Lick Observatory is operated by the University of California.

References

Bennett, C. L. Halpern, M., Hinshaw, G. et al 2003, ApJS, 148, 1.

Blaauw, A., Gum, C.S., Pawsey, J.L., & Westerhout, G. 1960, MNRAS, 121, 123

Bunclark, P., & Rots, A. 1997, Precise re-definition of DATE-OBS Keyword encompassing the millennium, `ftp://fits.cv.nrao.edu/fits/wcs/year2000.txt`

Cox, A. N. Ed. 2000, Allen's Astrophysical Quantities, AIP Press, Springer-Verlag, New York

Calabretta, M. R., & Greisen, E. W. 2002, A&A, 395, 1077, Paper II
Calabretta, M. R. et al. 2005, Representations of distortions in FITS world coordinate systems, in preparation, Paper IV
Cotton, W. D., Tody, D., & Pence, W. D. 1995, A&AS, 113, 159
de Vaucouleurs, G. 1976, Resolution of Commission 28 on Galaxies, Transactions IAU, 16B, 201
Delhaye, J. 1965, Stars and Stellar Systems, V, Blaauw, A., & Schmidt, M., ed., (University of Chicago Press), Chicago, Illinois, 61
Gibbs, P., & Tao, T. 1997, How do you add velocities in special relativity?, `http://math.ucr.edu/home/baez/physics/Relativity/SR/velocity.html`
Gordon, M. A., 1976, Methods of Experimental Physics, 12C, Meeks, M. L., Ed., (Academic Press), New York, New York, 277
Greisen, E. W., & Calabretta, M. R., 2002, A&A, 395, 1061, Paper I
Hanisch, R. J., Farris, A., Greisen, E. W., et al. 2001, A&A, 376, 359
Harten, R. H., Grosbøl, P., Greisen, E. W., & Wells, D. C. 1988, A&AS, 73, 365
International Union of Geodesy and Geophysics 1999, comptes rendus of the 22nd General Assembly, Birmingham, UK, 111 (`http://www.iugg.org/1999crendus.pdf`)
Kerr, F. J., & Lynden-Bell, D. 1986, MNRAS, 221, 1023
Lang, K. R. 1974, Astrophysical Formulæ(Berlin: Springer)
Lindegren, L., & Dravins, D. 2003, A&A, 401, 1185
Mangum, J. G., Emerson, D., & Greisen, E. W., 2000, *Imaging at Radio Through Submillimeter Wavelengths*, Mangum, J. G., Radford, S. J. E., ed., Astronomical Society of the Pacific Conference Series, 217, 179
Mathar, R. J. 2004, Applied Optics, 43, 928
McCarthy, D. D., & Petit, G. (eds.) 2004, IERS Conventions, IERS Technical Note No. 32, Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie, `http://www.iers.org/iers/publications/tn/tn32/`
Rindler, W. 1982, Introduction to Special Relativity, Oxford University Press, New York
Schroeder, D. J. 1999, Astronomical Optics, Academic Press, Inc., San Diego
Stumpff, P. 1980, A&AS, 41, 1
U. S. Department of Defense 2000, NIMA Technical Report TR8350.2. Available also at `http://earth-info.nima.mil/GandG/publications/tr8350.2/tr8350_2.html`
Wells, D. C., Greisen, E. W., & Harten, R. H. 1981, A&AS, 44, 363