CW data format description
Michael C. Nolan
2024 January 15

The delivered planetary radar Doppler spectra data format is a single CSV file which contains three csv subtables. These files are essentially a reformatted version of the internal "Radar Data Format" (RDF) used at the Arecibo and Goldstone planetary radar facilities. The first two tables contain metadata and the third table contains the actual radar Doppler spectra.

There is a basic Python3 script at the end of the bundle_description.txt file that can read in a set of these files. It is intended as an example to clarify the use of these files and is not a finished software package.

Most Doppler data are saved as power spectra for two orthogonal polarizations measured simultaneously. Occasionally only one polarization is obtained, either because of equipment issues or, rarely, because only one polarization was sampled. For this reason, these files each contain two spectra.

The data were processed using the procedure described in Magri et al. (2007). A radar survey of main-belt asteroids: Arecibo observations of 55 objects during 1999–2003. Icarus 186, 126-151. https://doi.org/10.1016/j.icarus.2006.08.018, except that for objects with tag nfreq = 0 or 1, that is, most near-Earth objects with short round-trip light-travel times, the "frequency hopping" described in that paper was not used. Instead, baselines are subtracted by excluding the (hand-noted) signal region from the polynomial fit described in that paper.

The first table contains "tags" that in principle may be different for each of the polarizations. These are saved in a four-column CSV table containing the name of the tag, the value for the first polarization, the value for the second polarization, and a short description with units if applicable. These tags are individually described later in this document. They are converted to ASCII from binary (integer or IEEE float) values. Every spectrum contains the same set of tags. Some of these values are not typically used, but are included as they could be meaningful in some older data.

The second table contains "extra tags" that are the same for each polarization. The user was free to add arbitrary "extra tags", so it is possible that some are undocumented, but the extra_tags generated by the normal processing software are documented here. These tags contain a name, a value, and a character indicating the data format from which they were converted (string: s integer: i, single-precision float: f, double-precision float: d), and a short description with units if applicable.

The third table contains the actual spectral data as a four-column CSV file. The columns are: frequency offset, polarization 1 power, polarization 2 power, <empty>. The spectra are computed using a Fast Fourier Transform, so the spacing between channels is exactly uniform,

with spacing in Hz = igw * ifft / 1000000, even if the tabulated frequencies suffer from finite precision.

Examined alone, the data file is a four-column CSV file begins with a row containing the string
`# Keywords,,,`
Followed by information used to create the PDS label and general useful text information, including the original RDF file it was created from, the version of the conversion script, and the date the CSV file was created. The next rows have the name of the entry in column 1, the value in column 2, and columns 3 and 4 are empty. The "observing system bookmark" is a shorthand that describes the transmit and receive stations that are fully spelled out in the label.

A typical entry follows:

```
Product Name,Phaethon CW 2018 Dec 15,,
Product Description,CW spectrum converted from RDF format,,
Product Version,1.0,,
Start Time,2017-12-15T23:36:07Z,,
Stop Time,2017-12-15T23:37:11Z,,
Target Name,Phaethon,,
Target Type,Asteroid,,
Editor List,"M. C. Nolan, Ed.",,
Product Processing Level,Calibrated,,
Science Search Facet,"Physical Properties",,
Product Wavelength Ranges,Microwave,,
Observing System Bookmark,AO TX;AO RX;AO RI,,
Original CW data file,/proj/radar/Phaethon/2017/dec15/cw/Phaethon.2017Dec15.p0156Hz.rdf,,
Software Version,20210411,,
Creation Date,2021-09-29T22:55:33Z,,
```

After those entries follows the string
`Tags,pol1tag,pol2tag,RDF Tags begin here,,`
Followed by the name of the tag and the value for each polarization, and short definition with units in four columns. The tag values are all presented as ASCII floating-point numbers, but many are in fact exact integers as noted below. The float values were stored as IEEE single-precision in the original data files. The definitions in the file are shortened from the description below.

The tags are:

rcsta	Integer	Receive start time in seconds from midnight (zero-based)
rcend	Integer	Receive end time in seconds from midnight (zero-based)
nffts	Integer	Number of FFTs summed to create spectrum
elev	Float	Pointing elevation of the telescope at receive weighted midtime in degrees from horizontal
azim	Float	Pointing azimuth of the telescope at receive weighted midtime in degrees

rttim	Float	Round-trip light time to the object at receive weighted midtime. This is a low-precision version of this value, and is intended for signal strength calculation only.
doppl	Float	Doppler offset (in Hz) in the original unprocessed spectrum, usually because the transmitted signal is purposely offset slightly from the nominal transmitter frequency. This will normally have been removed in processing.
phase	Float	Rotation phase at reference epoch (jd0), degrees. See the more complete Extra Tag for how this is used
itar	Integer	Target sequence number. Not often used.
irun	Integer	Run number within the original data file (1-based)
jgroup	integer	Group number that can be assigned in processing Almost never set.
ifft	Integer	Fourier transform length in initial processing.
igw	float	“Gate width” = original voltage sampling interval in microseconds. For spectra, usually an integer, though some systems do sample faster.
dfreq	float	Frequency spacing of spectrum. Should be exactly 1/(ifft * igw * 1E-6) Hz, and differences are due to rounding of dfreq.
tau	Float	Total integration time for this spectrum in seconds
rmsc	Float	Expected RMS of thermal background accounting for error propagation of mathematical operations.
rmsm	Float	RMS of thermal background measured.
xjcen	Integer	Channel in which 0 frequency lies, zero-based.
jsnr1	Integer	Channel number at start of signal region, zero-based
jsnr2	Integer	Channel number at end of signal region, zero-based
lljcp	Integer	Polarization: 1 = OC, 2 = SC
trpwr	Float	Transmitter power (kiloWatts) at the feed horn.
posfr	Integer	1 if frequency axis is increasing, -1 if it is decreasing.
tsys	Float	Receiver system temperature (Kelvin) at receive weighted midtime
gain	Float	Round-trip antenna gain (product of transmitter and receiver antenna gain) * 1E-12 at receive weighted midtime
obs	Integer	DSN number of receive station
sdev	Float	Conversion from spectrum power in a channel to km^2.
cross	Float	Cross section for spectrum (km^2). Set to 0 if uncalibrated
crerr	Float	Cross-section uncertainty (km^2). Set to 0 if uncalibrated
kpts	Integer	Number of frequency points used in original processing
nfreq	Integer	Number of frequency hops if used in processing
frstep	Float	Frequency step if used in processing
color	Integer	Hop number if individual hops are saved (zero-based)
freq1	Float	First frequency offset for hop processing

Jsnr1 and jsnr2 are set by the user to describe the region in the spectrum where the signal is believed to be. Xjcen is the channel that contains the ephemeris frequency of the target: If the ephemeris was incorrect or preliminary, it may not be the center of the target spectrum.

The frequency spacing is exactly 1000000/(ifft * igw) Hz and zero frequency is in the xjcen bin. The tabulated frequencies can have errors due to finite calculation precision.

After those entries follows the string
`ExtraTags,RDF Extra Tags begin here,,`
Followed by the name of the tag, the value, a letter indicating whether the value should be considered a string, an integer, or a floating-point value, and a short description with units in []. The values were stored as ASCII text in the original data file, and so numbers are in principle arbitrary precision, but typically should be considered IEEE double-precision.

The user can add arbitrary Extra Tags. The following set are generated by the standard reduction software.

file_date	ISO Date string for date the original processed data file was written.	s
diameter	If set, Diameter to use in Radar Albedo calculation. Rarely set correctly, usually 1.0. [km]	f
period	If set, rotation period for Doppler calculations. Rarely set correctly, usually 1.0. [h]	f
lambda	Radar wavelength [m]	f
xmit_sta	Normally Arecibo or Goldstone	s
phase0	If set, rotation phase of the asteroid at epoch jd0 Used to compute rotation phase. This is a user-defined quantity, and so may not match any particular standard coordinate system. [deg]	d
jd0	Julian Day Number of the reference epoch for rotation phase. [d]	d
jdstart	Julian Day Number for start of receive. To be interpreted as the nearest integer UTC second. This is the reference time for all astrometry. [d]	d
jdmean	Julian Day Number for receive weighted mid-time [d]	d
calmean	Receive mid-time as YYYY mmm dd hh:mm:ss UTC	s
jdend	Julian Day Number for end of receive [d]	d
distmin	Closest range to target during run [au]	f
distmean	Range to target at midpoint of run [au]	f
distmax	Largest range to target during run [au]	f
ramin	Smallest RA of target during run [h]	f
ramean	RA of target at midpoint of run [h]	f
ramax	Largest RA of target during run [h]	f

decmin	Smallest Declination of target during run [deg]	f
decmean	Declination of target at midpoint of run [deg]	f
decmax	Largest Declination of target during run [deg]	f
infile	Name and path of original processed file in rdf format.	s
xmit_poln	Polarization transmitted in engineering convention. Choices are LCP, RCP, H, and V	s
timezone	Time zone for dates and times (always UTC).	s
tzcorr	Timezone correction used in processing [h]	f
perror	Pointing error due to Arecibo platform tilt and focus offset [arcsec]	f
focusoff	Focus offset at Arecibo [m]	f
badcal	1 if Focus error > lambda/4 or pointing error > 40 arcsec, 0 otherwise	i

After that comes the string

```
# Column Definitions,,,
```

Followed by definitions of the columns in the data that follows: Frequency with respect to the observing ephemeris, Polarization 1 normalized to unit standard deviation and zero mean in baseline, Polarization 2 normalized to unit standard deviation and zero mean in baseline

Then the string

```
# Data,,,
```

And then the data, arranged as frequency, polarization 1, polarization 2, which are the Opposite-Circular and Same-Circular polarizations, respectively unless otherwise noted in the Tags and documentation for that experiment.