International Telecommunication Union

ITU-R

Radiocommunication Sector of ITU

Recommendation ITU-R SM.2117-0
(09/2018)

Data format definition for exchanging stored I/Q data for the purpose of spectrum monitoring

SM Series
Spectrum management

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Annex 1 of Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	**Spectrum management**
SNG	Satellite news gathering
TF	Time signals and frequency standards emissions
V	Vocabulary and related subjects

Note*: This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2018

RECOMMENDATION ITU-R SM.2117-0

Data format definition for exchanging stored I/Q data for the purpose of spectrum monitoring

(2018)

Scope

This Recommendation defines a harmonized file format for exchanging stored In-phase and Quadrature components (I/Q) data.

Keywords

I/Q data, I/Q modulation, joint working, data exchange

Abbreviations/Glossary

HDF5	Hierarchical data format version 5
I/Q	In-phase and quadrature components
LPF	Low pass filter
TDOA	Time difference of arrival
UTF-8	Unicode transformation format 8-bit

The ITU Radiocommunication Assembly,

considering

a) that in-phase and quadrature components (I/Q) data files have a widespread and manifold use in many RF applications;

b) that I/Q data files may be device and manufacturer specific;

c) that the exchange of I/Q data files can be facilitated by a harmonized, easily accessible file format;

d) that stored I/Q data files offer a high degree of flexibility and may support many applications in the work of monitoring services;

e) that stored I/Q data files enable a reproducible reconstruction of a signal with either a signal generator or an analyser;

f) the usefulness of I/Q data files is increased by a device and manufacturer independent file format,

recommends

1 that devices and software applications supporting the use of I/Q data files offer a suitable conversion method from the original I/Q data format to the common format defined in Annex 1;

2 that the format defined in Annex 1 should be used to exchange I/Q data for RF monitoring applications.

Annex 1

Data format definition for exchanging stored I/Q data for the purpose of spectrum monitoring

1 General aspects

This annex provides the definition of a data format for exchanging stored I/Q data for radio monitoring purposes.

The data format provides:

– an easily understandable and implementable file format for I/Q data;

– a minimum set of characteristics to describe the data contained in the file;

– a data description suited for the exchange, analysis, and generation of signals on a spectrum analyser or for replay on a signal generator.

2 Description of I/Q data

2.1 Basics of I/Q data

The term I/Q data as used in the scope of this Recommendation describes a discrete time series of a complex valued baseband signal. With additional information like amplitude scaling and RF carrier frequency, a corresponding RF signal can be described.

In general, I/Q data describes a complex baseband signal $b(t)$ which can be transformed to or can be derived from a corresponding real valued RF signal $x(t)$. The "in-phase" component or real part of $b(t)$ is called $i(t)$. The "out-of-phase" component or imaginary part of $b(t)$ is called $q(t)$.

Thus, the corresponding RF signal $x(t)$ can be expressed as:

$$x(t) = i(t) \cdot \sqrt{2} \cdot \cos(2\pi f_0 t) - q(t) \cdot \sqrt{2} \cdot \sin(2\pi f_0 t) \quad (1)$$

It is evident that equation (1) can be used directly in a generator to synthesize the corresponding RF signal using $i(t)$ and $q(t)$ centred at a carrier frequency f_0.

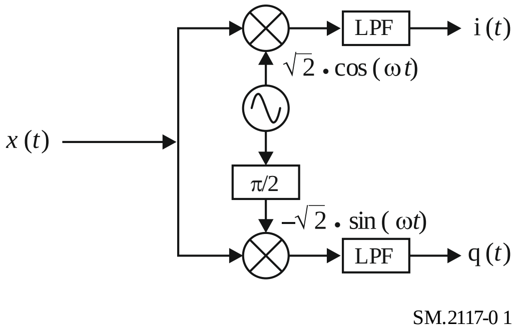
To obtain the corresponding complex baseband signal from a RF signal centred at a carrier frequency f_0, equation (2) would be implemented in an ideal analyser. The symbol 'j' indicates the imaginary component and is equal to the square root of minus one. The symbol '$*$' indicates the convolution operation. The impulse response $h(\tau)$ defines a low pass filter (LPF) with a gain of one inside its pass band. The filtering with a LPF is necessary to attenuate signal components centred at twice the carrier frequency. The passband of the LPF must be selected wide enough to cover the bandwidth of the corresponding RF signal.

$$b(t) = i(t) + \mathrm{j} \cdot q(t) = x(t) \cdot \sqrt{2} \cdot e^{-\mathrm{j} \cdot 2\pi f_0 t} * h(\tau) \quad (2)$$

Usually, f_0 is chosen in such way that f_0 is equal to or greater than half the bandwidth of the corresponding RF signal. If a smaller value is chosen for f_0 aliasing is not avoidable. Equation (2) is sketched in Fig. 1. Baseband I/Q data $i(t) + jq(t)$ are obtained from an RF signal $x(t)$ by the use of a quadrature detector (Fig. 1), where the real part $i(t)$ and the imaginary part $q(t)$ are expressed by the in-phase component and the quadrature component, respectively.

FIGURE 1

Quadrature detector by I/Q data



Currently, many signal generators and analysers work with sampled data. Thus, $i(t)$ and $q(t)$ are represented by a sequence of time discrete numbers with a sampling frequency f_s. The sample time interval $T = 1/f_s$ is constant. At every discrete point in time $t = nT$ the baseband signal is sampled and digitized in an analyser. In a generator, the sampled and digitized baseband signals are used to reconstruct an analogue continuous time RF signal. The sampled and digitized set of $i(t)$ and $q(t)$ consisting of N value pairs thus becomes $I(n)$ and $Q(n)$, with $n = \{0,1,2, \ldots N - 1\}$ and $t = nT$. The values are represented as dimensionless values. The actual RF signal can be retained by the multiplication with a so-called scaling factor (see § 4).

2.2 Considerations regarding sampling frequency and filter bandwidth

The maximum bandwidth that can be represented by the sampled I/Q data is equal to or smaller than the sampling frequency if aliasing is avoided. Therefore the actual RF bandwidth f_a represented by the I/Q data is often smaller than the sampling frequency. This Recommendation defines f_a as the equivalent noise bandwidth of the band limiting filter of an analyser. In generators working with sampled I/Q data often a corresponding band limiting filter exist in order to attenuate aliased components. The equivalent noise bandwidth of a band limiting filter of a generator should be equal to or greater than that of the analyser, used to record the I/Q data, in order to avoid a potential degradation of a regenerated RF signal.

3 File format

This section describes the format of a file containing data sets of continuous recordings of complex baseband signals and the associated additional information needed for a thorough description of the data sets.

The file format is based on the Hierarchical Data Format, version 5 (HDF5) file format (see https://support.hdfgroup.org/HDF5/doc/Specs.html [1]).

The raw data sets $I(n)$ and $Q(n)$ are stored in HDF5 data sets. The associated additional information (meta-data) is stored in a set of attached HDF5 attributes. An HDF5 data set, together with its attached HDF5 attributes, is stored in the group structure of an HDF5 file. The meta-data together with the raw data can completely describe the stored I/Q data, and thus, the original RF signal. An HDF5 data set built using the rules of this Recommendation is self-explanatory. Thus it can be interpreted even

[1] Copyright and license statements can be found in the COPYING file at the top level of the source code for open source software (see https://support.hdfgroup.org/products/licenses.html and https://support.hdfgroup.org/ftp/HDF5/releases/COPYING). The SZIP compression library is not part of HDF5 and thus also not part of this Recommendation.

if this Recommendation is not available. I/Q data sets created according to this Recommendation and other HDF5 objects can be stored in the group structure of the HDF5 file format (for example, groups, other data sets and named data types). The simplest HDF5 file containing ITU-R compliant I/Q data consists of only one I/Q data set in the root group of the file.

The file name extension is '.h5' to indicate an HDF5 file. An example for a complete filename is 'signal.h5'.

3.1 HDF5 attributes

Mandatory meta-data required to interpret the raw I/Q data is stored in a set of HDF5 attributes, which is attached to the HDF5 data set that contains the raw I/Q data. These mandatory attributes are listed and defined in Table 1.

TABLE 1

Mandatory attributes (data types expressed as HDF5 data types)

Name	Data type of value	Definition and comments
ITU-R data set class	variable length string	Fixed string with the content: 'I/Q'.
ITU-R Recommendation	variable length string	Fixed string: "Rec. ITU-R SM.2117-0"
RF carrier frequency (Hz)	H5T_IEEE_F64LE	The carrier frequency of the RF signal (f_0). Shall be positive. Use a value of zero if RF carrier frequency is unknown or not important.
Sampling frequency (Hz)	H5T_IEEE_F64LE	The sampling frequency of the I/Q data (f_s). Shall be greater than zero.
Data set type interpretation	variable length string	Fixed string with the content: 'Integer types, used to store I/Q data, are interpreted as fix point numbers with the radix point right to the most significant bit.'.
Data set unit	variable length string	The unit of the real world I/Q data. Only the following strings are valid: '', 'V', 'V/m', 'A/m' Use an empty string '' if the real world unit is not important.
Data set scaling factor	H5T_IEEE_F32LE	Multiply the normalized I/Q data stored in the data set with this factor to obtain I/Q values in their real world unit.

The attributes listed and defined in Table 2 are optional and may also be attached to the data set. Attributes which are not known should not be used.

TABLE 2

Optional attributes (data types expressed as HDF5 data types)

Name	Data type of value	Definition and comments
Comment	variable length string	Any useful comment regarding the I/Q data set.
Device	variable length string	Describes the device used to create the I/Q data set.
Filter bandwidth (Hz)	H5T_IEEE_F64LE	The equivalent noise bandwidth of the band limiting filter of the analyser (f_a). Valid range: 0 $<= f_a <= f_s$
Timestamp coarse (s)	H5T_STD_U32LE	UTC in POSIX time format at the occurrence of the first I/Q sample.
Timestamp fine (ns)	H5T_STD_U32LE	UTC sub seconds at the occurrence of the first I/Q sample.
Geolocation latitude (degree)	H5T_IEEE_F64LE	Latitude based on WGS 84 Valid range: −180 to 180
Geolocation longitude (degree)	H5T_IEEE_F64LE	Longitude based on WGS 84 Valid range: −90 to 90
Geolocation altitude (m)	H5T_IEEE_F32LE	Altitude above mean sea level Valid range: >= −10e3
Geolocation separation (m)	H5T_IEEE_F32LE	Geoid Separation: difference between WGS 84 ellipsoid and mean sea level
Speed over ground magnitude (m/s)	H5T_IEEE_F32LE	Valid range: >=0
Speed over ground azimuth (degree)	H5T_IEEE_F32LE	Valid range: 0 to 360, east = 90, true north = 0
Orientation azimuth (degree)	H5T_IEEE_F32LE	Valid range: 0 to 360, east = 90, true north = 0
Orientation elevation (degree)	H5T_IEEE_F32LE	Valid range: −90 to 90, up = 90
Orientation skew (degree)	H5T_IEEE_F32LE	Valid range: −180 to 180, right = 90
Magnetic declination (degree)	H5T_IEEE_F32LE	The difference between the reading of a magnetic compass and true north. The reported orientation azimuth must already be corrected by this value if a magnetic compass was used for orientation. The magnetic declination should not be reported if the orientation azimuth was not obtained by a magnetic compass.
Unsynced timestamp flag	H5T_STD_U8LE	If >0: The Timestamp might not be synchronized to a Reference clock.
Invalid flag	H5T_STD_U8LE	If >0: At least one sample in the stored I/Q data might be invalid.
PLL unlocked	H5T_STD_U8LE	If >0: At least one PLL might not have been locked and might affect at least one sample in the stored I/Q data.
AGC flag	H5T_STD_U8LE	If >0: At least one sample in the stored I/Q data might be affected by gain variations due to an active AGC.

TABLE 2 (*end*)

Name	Data type of value	Definition and comments
Detected signal flag	H5T_STD_U8LE	If > 0: A signal was detected during the time span defined by the stored I/Q data
Spectral inversion flag	H5T_STD_U8LE	If > 0: The stored I/Q data represent an inversed spectrum.
Over range flag	H5T_STD_U8LE	If > 0: At least one sample in the stored I/Q data might be affected by a detected over range condition.
Lost sample flag	H5T_STD_U8LE	If > 0: At least one sample in the data might have been skipped during recording of the stored I/Q data.
Attenuator (dB)	H5T_IEEE_F32LE	The attenuator setting. It is not needed to scale the I/Q values but its knowledge can be useful to estimate the measurement range of the analyser.
Antenna factor (1/m)	H5T_IEEE_F32LE	Antenna factor at the RF carrier frequency. It relates to the antenna used to obtain the stored I/Q data. It can be used to convert to other units if the antenna factor is constant within the filter bandwidth of the stored I/Q data.
Reference point	Variable length string	The reference point for the timestamp and scaling factor. It can be either "Antenna output port" or "Receiver input port".
Receiver input impedance (Ohm)	H5T_IEEE_F32LE	Nominal port impedance of the receiver in Ohms. If this attribute is omitted, 50 Ohms can be assumed.

All HDF5 attributes defined in this document are scalars and thus their dataspaces have only one dimension with a fixed size of one.

Optional attributes should be attached only if they are known, valid and applicable.

Attributes not defined in this document are allowed, but their names must be different from those defined in this Recommendation. In order to achieve this also for future revision of this Recommendation it is mandatory to use the string 'User' as a prefix for such attribute names. ITU will not define any attribute names in future revisions of this Recommendation which will begin with the string 'User'.

All attributes shall be attached in the same order as they are listed in Tables 1 and 2. The mandatory attributes in Table 1 are attached first. User defined attributes are attached after the optional attributes of Table 2. Optional attributes which are not used are skipped.

The variable length strings use the UTF-8 character set and are null terminated.

The geolocation and orientation attributes should have been fixed at a point in time close to the occurrence of the first I/Q sample.

The scaling factor and timestamp attributes should be valid at the reference point of interest. This reference point can be defined by the optional attribute 'Reference point', see Table 2. If no reference point is defined, the receiver input port is assumed. Thus any known and not negligible attenuation and group delay of devices between the analyser's input and the reference point should be compensated. Also the geolocation and orientation attributes should describe the position and orientation of the reference object. The hints in this paragraph are particularly important when for example lengthy cables are used in time difference of arrival (TDOA) measurement setups.

3.2 Data set

The dimensionless I/Q data are stored in a one dimensional HDF5 data set.

The name of the data set may be any valid HDF5 name.

The HDF5 file format is comprised of a compound data type which contains a complex number for each I/Q channel and an optional HDF5 bit field. The elements of the data set are stored in a one dimensional array.

If the bit field is used (see Table 3), it contains all the flags which are available as attributes. The flag values in the attributes are valid for the complete data set and are calculated by a logical "OR" function between all flag values associated with each sample of the data set. In the bit field, these sample-based flags can be made available for a more detailed analysis.

HDF5 compound data types are similar to structures as they appear in the programming language C. They must be defined before writing data of this compound data type and their definition is stored within the HDF5 file. For each member of an HDF5 compound data type a name and data type must be defined.

The member names used for an ITU-R I/Q data set shall be "Channel_*XYZ*" and "BitField". The part "*XYZ*" of the name string stands for a string, describing the specific channel. The describing strings must be different if more than one channel is used. A member with the name "BitField" is optional and, when used, has to be the last member.

The data type of "Channel_*XYZ*" members is also an HDF5 compound data type with the name "Real" for the first member and "Imag" for the second member and stores a complex I/Q value. Both members are of the same basic HDF5 data type which can be selected between:

– H5T_STD_I16LE

– H5T_STD_I32LE

– H5T_IEEE_F32LE

Note that the two HDF5 integer types are signed two's complement integers. They are interpreted here as fixed point numbers with the radix point right to the most significant bit. Also note that these fixed point numbers can only cover the range from -1 to $1-2^{-15}$ for the 16-bit integer type and -1 to $1-2^{-31}$ for the 32-bit integer type. An integer value of 1000 of type H5T_STD_I16LE is interpreted as $1000 / 2^{15}$. An integer value of 1000 of type H5T_STD_I32LE is interpreted as $1000 / 2^{31}$.

The data type of the "BitField" member is the basic HDF5 data type H5T_STD_B16LE. In Table 3 the bit positions of the flags are defined. The definition of these flags is exactly the same as the definition of the corresponding flags in Table 2. Note that the bit position zero is the position of the least significant bit. A flag which is true has a bit value of one. Flags that do appear as attributes are assumed to be not valid and the corresponding bit values shall be set to zero.

TABLE 3

Bitfield contents

Flag	Bit position	Comment
Unsynced_Timestamp	15	
Invalid	14	
PLL_Unlocked	13	
AGC	12	
Detected_Signal	11	
Spectral_Inversion	10	
Over_Range	9	
Lost_Sample	8	Is true for the first unlost sample after the loss

All valid HDF5 storage properties can be used for the data set. To store in chunks could be useful if small portions in big data sets need to be accessed quickly.

The following four examples show table column captions, representing valid ITU-R I/Q data sets:

– Channel_1.Real, Channel_1.Imag
– Channel_one.Real, Channel_one.Imag, BitField
– Channel_X.Real, Channel_X.Imag, Channel_Y.Real, Channel_Y.Imag
– Channel_1.Real, Channel_1.Imag, Channel_2.Real, Channel_2.Imag, BitField

3.3 Variable attributes

In some cases it is not sufficient to apply a single set of attributes to an entire I/Q recording. Some of the attributes might change over time.

For example, it might be necessary to apply different scaling factors because a gain setting has been changed. Another example is that the sample rate and time stamps might not be synchronized in some receivers and thus it might be necessary to use different time stamps for different sectors of the I/Q data.

Problems like this can be easily solved by using separate data sets for the different sectors of the I/Q data. Whenever a relevant attribute changes, the old data set is closed and a new one is created with its actual attributes attached to it.

The following conventions may be useful for the handling of multisector I/Q data:

– All data set names of the same multisector I/Q data should begin with a constant string like “Multisector_IQ_” followed by a suffix string.
– The suffix string should represent a number between zero and 9 999 999 999 and is formated as “0000000000” to “9999999999”.
– The number in the suffix is incremented by one from the actual to the next data set.
– The data set name of the first sector should have the suffix “0000000000”.
– All data sets of the same multisector I/Q data are stored in the same group and no other data sets or groups should be a contained in this group.

4 Interpretation and usage of the scaling factor

The scaling factor is the factor with which the dimensionless I/Q data stored in the data set must be multiplied to obtain their values in the unit indicated in the meta-data.

Example:

The stored dimensionless I/Q value pair is $I(n) = -0.6$ and $Q(n) = 0.8$. The unit is V and the scaling factor $sf = 0.005$. This results in:

$$i(nT) = I(n) \cdot sf \cdot \mathrm{V} = \text{-}0.003\ \mathrm{V} \quad (3)$$

$$q(nT) = Q(n) \cdot sf \cdot \mathrm{V} = 0.004\ \mathrm{V} \quad (4)$$

Thus the voltage magnitude of the corresponding RF signal at the corresponding moment in time is 0.005 V. This is equivalent to −46.02 dBV, 73.98 dBµV or −33.01 dBm over 50 Ω.