The FITS format

L.Chiappetti

EWASS 2012 - Special Session 12
Rome 05 July 2012

The FITS format

what astronomers use it for ?

L.Chiappetti

on behalf of the IAU FITS Working Group

EWASS 2012 - Special Session 12
Rome 05 July 2012

The FITS format

what astronomers use it for ?

L.Chiappetti

INAF
IASF Milano

on behalf of the IAU FITS Working Group

thanks to: W.Pence (GSFC), A.Dobrzycki (ESO), R.Seaman (NOAO)

EWASS 2012 - Special Session 12
Rome 05 July 2012

The FITS format

what astronomers use it for ?
how can non-astronomers use it ?

L.Chiappetti

INAF
IASF Milano

on behalf of the IAU FITS Working Group

thanks to: W.Pence (GSFC), A.Dobrzycki (ESO), R.Seaman (NOAO)

EWASS 2012 - Special Session 12
Rome 05 July 2012

FITS in a nutshell – 2.a

- established in 1979 x

- endorsed by IAU in 1982
- not only images !
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.b

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982
- not only images !
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.c

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982
- not only images !
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.d

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982
- not only images !
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.e

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982 ▸
- not only images ! ▸
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.f

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982
- not only images !
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.g

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982 ▸
- not only images ! ▸
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

FITS in a nutshell – 2.h

- established in 1979 x

From the NOST "FITS basics and information (periodic posting)"

"FITS (Flexible Image Transport System) is a data format designed to provide a means for convenient exchange of astronomical data between installations whose standard internal formats and hardware differ."

- endorsed by IAU in 1982
- not only images !
- not only transport but archive and working format !
- somebody proposed renaming as *Flexible Image and Table Systems*
- since 2005 also MIME standard (RFC 4047)
- de facto standard for archiving and exchange of astronomical data and used as internal working format by many analysis packages

the International Astronomical Union – 3.a

The IAU is one of earliest *international scientific unions*

created in the framework of the *Conseil International de Recherches*

(London 1918; Paris 1918; established Bruxelles 1919)

since 1931 within ICSU: International Council of Scientific Unions

first IAU General Assembly took place in Rome in 1922

the International Astronomical Union – 3.b

courtesy of O.Straniero

the International Astronomical Union – 3.c

The IAU is one of earliest *international scientific unions*

created in the framework of the *Conseil International de Recherches*

(London 1918; Paris 1918; established Bruxelles 1919)

since 1931 within ICSU: International Council of Scientific Unions

first IAU General Assembly took place in Rome in 1922

. . . the XXVIII GA will take place in August 2012 in Bejing !

the IAU FITS Working Group – 3.d

The IAU FWG was established by the IAU in 1988 under

- Division 12 (Union-Wide Activities)
  - Commission 5 (Documentation and Astronomical Data)
    - FITS Working Group

Regional FITS Committees

(. . . to be re-arranged at Bejing GA)

the IAU FITS Working Group – 3.e

The IAU FWG was established by the IAU in 1988
Details on site http://fits.gsfc.nasa.gov/iaufwg/

- *maintains the standard*
- *reviews, approves and maintains future extensions, recommended practices and thesaurus of approve[d] FITS keywords*
- maintains list of registered extensions
- maintains registry of conventions
- with formal voting rules
- with the support of Regional FITS Committees and of the community at-large (FITSBITS exploder)

- may establish adhoc temporary task forces
- last (3.0) standard revision by dedicated tech panel (2006-2008)

the IAU FITS Working Group – 3.f

The IAU FWG was established by the IAU in 1988
Details on site http://fits.gsfc.nasa.gov/iaufwg/

- *maintains the standard*
- *reviews, approves and maintains future extensions, recommended practices and thesaurus of approve[d] FITS keywords*
- maintains list of registered extensions
- maintains registry of conventions
- with formal voting rules
- with the support of Regional FITS Committees and of the community at-large (FITSBITS exploder)

- may establish adhoc temporary task forces
- last (3.0) standard revision by dedicated tech panel (2006-2008)

the IAU FITS Working Group – 3.g

The IAU FWG was established by the IAU in 1988
Details on site http://fits.gsfc.nasa.gov/iaufwg/

- *maintains the standard*
- *reviews, approves and maintains future extensions, recommended practices and thesaurus of approve[d] FITS keywords*
- maintains list of registered extensions
- maintains registry of conventions
- with formal voting rules
- with the support of Regional FITS Committees and of the community at-large (FITSBITS exploder)

- may establish adhoc temporary task forces
- last (3.0) standard revision by dedicated tech panel (2006-2008)

the IAU FITS Working Group – 3.g

The IAU FWG was established by the IAU in 1988
Details on site http://fits.gsfc.nasa.gov/iaufwg/

- *maintains the standard*
- *reviews, approves and maintains future extensions, recommended practices and thesaurus of approve[d] FITS keywords*
- maintains list of registered extensions
- maintains registry of conventions
- with formal voting rules
- with the support of Regional FITS Committees and of the community at-large (FITSBITS exploder)

- may establish adhoc temporary task forces
- last (3.0) standard revision by dedicated tech panel (2006-2008)

FITS documentation – 4.a

FITS is "*completely defined down to bit level by refereed published papers unlike formats implicitly defined by the software which reads and writes them*"

Published papers - links are alive !!

http://fits.gsfc.nasa.gov/fits_standard.html	Current standard with latest updates (3.0): always check here !
A&A 524, A42 (2010)	Latest published standard (3.0) so far coincides with current
	absorbs and supersedes
A&AS 44, 363 (1981)	Original FITS paper (Wells et al. 1981)
A&AS, 44, 371 (1981)	Random groups (Greisen & Harten 1981)
A&AS, 73, 359 (1988)	Generalized extensions (Grosbol et al. 1988)
A&AS, 73, 365 (1988)	ASCII table extension (Harten et al. 1988)
A&AS, 105, 53 (1994)	IMAGE extension (Ponz et al. 1994)
A&AS, 113, 159 (1995)	BINary TABLE extension (Cotton et al. 1995)
	incorporated by reference in the standard
A&A, 395, 1061 (2002)	World Coordinate System concept (Greisen & Calabretta 2002, WCS I)
A&A, 395, 1077 (2002)	WCS sky projection (Calabretta & Greisen 2002, WCS II)
A&A, 446, 747 (2006)	WCS spectral (Greisen et al. 2006, WCS III)
MNRAS, 381, 865 (2007)	WCS HEALPIX projection (Calabretta & Roukema 2007)
more to come	WCS IV V in preparation
	not on journals
RFC 4047	FITS as MIME datatype (Allen & Wells 2005)
(text file at GSFC)	Floating point agreement
(text file at GSFC)	Y2K date agreement

FITS documentation – 4.b

Ουκ εστι βασιλικη οδος There is no Royal Way

Euclid's way of telling King Ptolemy: *Read The Fine Manual* !

Published papers - links are alive !!

http://fits.gsfc.nasa.gov/fits_standard.html	Current standard with latest updates (3.0): always check here !
A&A 524, A42 (2010)	Latest published standard (3.0)
	absorbs and supersedes
A&AS 44, 363 (1981)	Original FITS paper (Wells et al. 1981)
A&AS, 44, 371 (1981)	Random groups (Greisen & Harten 1981)
A&AS, 73, 359 (1988)	Generalized extensions (Grosbol et al. 1988)
A&AS, 73, 365 (1988)	ASCII table extension (Harten et al. 1988)
A&AS, 105, 53 (1994)	IMAGE extension (Ponz et al. 1994)
A&AS, 113, 159 (1995)	BINary TABLE extension (Cotton et al. 1995)
	incorporated by reference in the standard
A&A, 395, 1061 (2002)	World Coordinate System concept (Greisen & Calabretta 2002, WCS I)
A&A, 395, 1077 (2002)	WCS sky projection (Calabretta & Greisen 2002, WCS II)
A&A, 446, 747 (2006)	WCS spectral (Greisen et al. 2006, WCS III)
MNRAS, 381, 865 (2007)	WCS HEALPIX projection (Calabretta & Roukema 2007)
more to come	WCS IV V in preparation
	not on journals
RFC 4047	FITS as MIME datatype (Allen & Wells 2005)
(text file at GSFC)	Floating point agreement
(text file at GSFC)	Y2K date agreement

FITS documentation – 4.c

Ουκ εστι βασιλικη οδος There is no Royal Way

Euclid's way of telling King Ptolemy: *Read The Fine Manual* !

Network resources (also alive !)

http://fits.gsfc.nasa.gov/	NASA FITS support office
http://www.cv.nrao.edu/fits/	NRAO FITS archive
http://fits.gsfc.nasa.gov/fits_documentation.html	Reference to published papers
http://fits.gsfc.nasa.gov/fits_libraries.html	Software libraries
http://fits.gsfc.nasa.gov/fits_viewer.html	FITS viewers
http://fits.gsfc.nasa.gov/fits_wcs.html	WCS @ NASA
http://www.atnf.csiro.au/people/mcalabre/WCS/index.html	WCS @ ATNF
http://listmgr.cv.nrao.edu/mailman/listinfo/fitsbits	fitsbits mail exploder
http://fits.gsfc.nasa.gov/xtension.html	Registry of extensions
http://fits.gsfc.nasa.gov/fits_registry.html	Registry of conventions
http://fits.gsfc.nasa.gov/fits_conventions.html	Other conventions

Once FITS always FITS – 5.a

- Updates to the standard are strictly controlled by the IAU FWG.
- However the standard allows a wide degree of freedom in defining format and usage conventions for particular purposes.

- Astronomy comes from the past and shall preserve data (*not just documents*) for the future.
- E.g. IAU Task Force for the Preservation and Digitization of Photographic Plates
  http://www.lhobs.org/PDPP.html
- IVOA deals with gregorian proleptic calendar
- WCS Paper V draft quotes Herschel, 1851 and the bull *Inter Gravissimas* by Pope Gregory XIII

Once FITS always FITS – 5.a

- Updates to the standard are strictly controlled by the IAU FWG.
- However the standard allows a wide degree of freedom in defining format and usage conventions for particular purposes.

Once FITS forever FITS

"*Any structure that is a valid FITS structure* shall remain a valid FITS structure at all future times. Use of certain valid FITS structures may be deprecated [by the standard]"

- Astronomy comes from the past and shall preserve data (*not just documents*) for the future.
- E.g. IAU Task Force for the Preservation and Digitization of Photographic Plates
  http://www.lhobs.org/PDPP.html
- IVOA deals with gregorian proleptic calendar
- WCS Paper V draft quotes Herschel, 1851 and the bull *Inter Gravissimas* by Pope Gregory XIII

Once FITS always FITS – 5.c

- Updates to the standard are strictly controlled by the IAU FWG.
- However the standard allows a wide degree of freedom in defining format and usage conventions for particular purposes.

Olim FITS semper FITS

"*Any structure that is a valid FITS structure* shall remain a valid FITS structure at all future times. Use of certain valid FITS structures may be deprecated [by the standard]"

- Astronomy comes from the past and shall preserve data (*not just documents*) for the future.
- E.g. IAU Task Force for the Preservation and Digitization of Photographic Plates
  http://www.lhobs.org/PDPP.html
- IVOA deals with gregorian proleptic calendar
- WCS Paper V draft quotes Herschel, 1851 and the bull *Inter Gravissimas* by Pope Gregory XIII

Once FITS always FITS – 5.c

- Updates to the standard are strictly controlled by the IAU FWG.
- However the standard allows a wide degree of freedom in defining format and usage conventions for particular purposes.

Quondam FITS semper FITS

"*Any structure that is a valid FITS structure* shall remain a valid FITS structure at all future times. Use of certain valid FITS structures may be deprecated [by the standard]"

- Astronomy comes from the past and shall preserve data (*not just documents*) for the future.
- E.g. IAU Task Force for the Preservation and Digitization of Photographic Plates
  http://www.lhobs.org/PDPP.html
- IVOA deals with gregorian proleptic calendar
- WCS Paper V draft quotes Herschel, 1851 and the bull *Inter Gravissimas* by Pope Gregory XIII

Once FITS always FITS – 5.e

- Updates to the standard are strictly controlled by the IAU FWG.
- However the standard allows a wide degree of freedom in defining format and usage conventions for particular purposes.

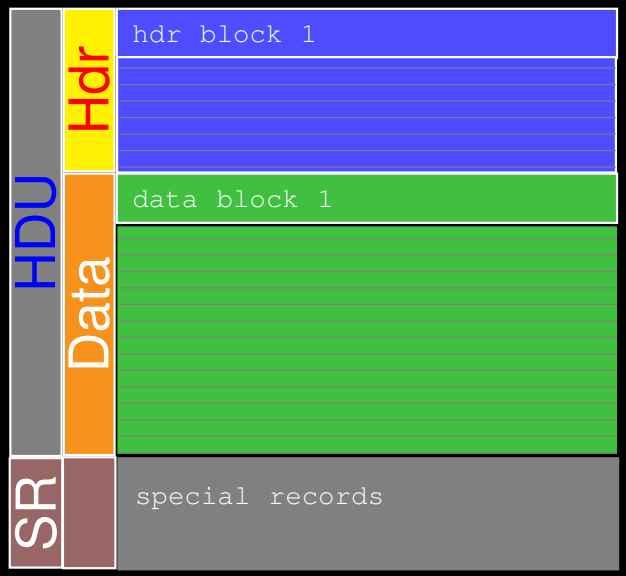
Semel FITS semper FITS

"*Any structure that is a valid FITS structure* shall remain a valid FITS structure at all future times. Use of certain valid FITS structures may be deprecated [by the standard]"

- Astronomy comes from the past and shall preserve data (*not just documents*) for the future.
- E.g. IAU Task Force for the Preservation and Digitization of Photographic Plates
  http://www.lhobs.org/PDPP.html
- IVOA deals with gregorian proleptic calendar
- WCS Paper V draft quotes Herschel, 1851 and the bull *Inter Gravissimas* by Pope Gregory XIII

SIMPLE□□=□T – 6.a

Basic FITS (1979-81)

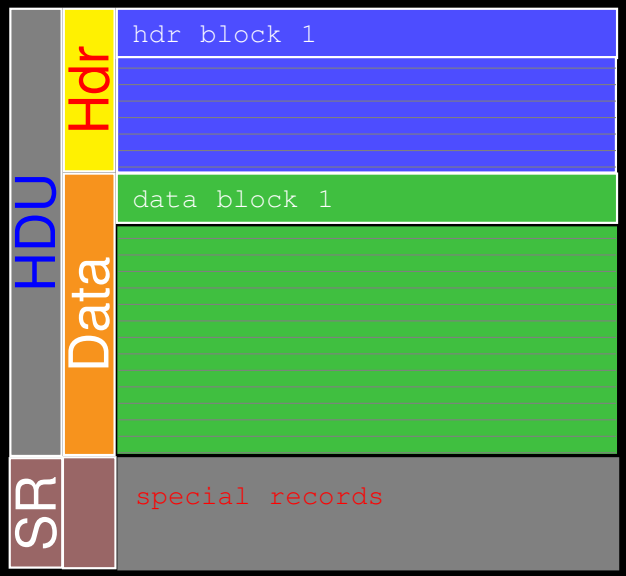


A FITS file is made of 2880-byte records called *FITS blocks* divided between a header and a data area.

why 2880 ?

SIMPLE□□=□T – 6.b

Basic FITS (1979-81)



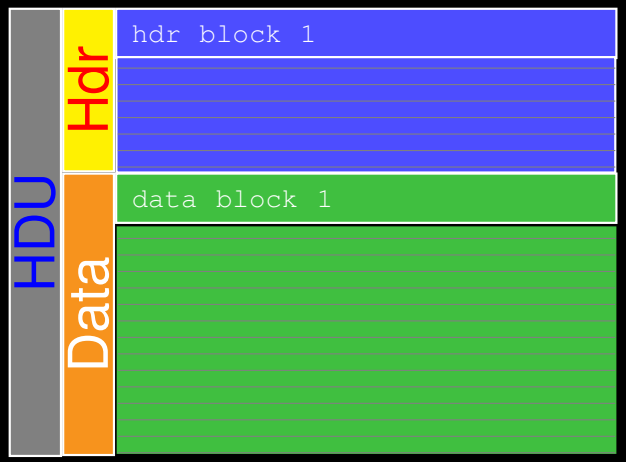
A FITS file is made of 2880-byte records called *FITS blocks* divided between a header and a data area.

why 2880 ?

All what follows the (last) HDU is intended for (testing of) future development, reserved to IAUFWG . . . forget about it !

SIMPLE□□=□T – 6.c

Basic FITS (1979-81)

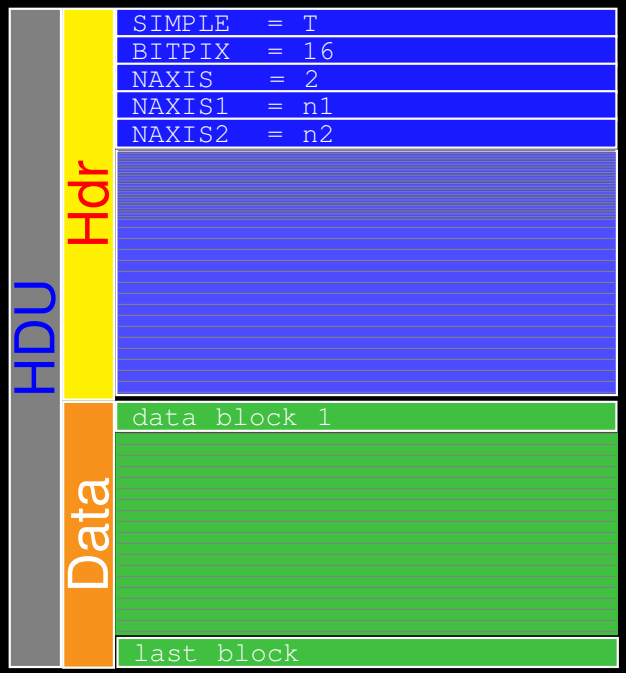


A FITS file is made of 2880-byte records called *FITS blocks* divided between a header and a data area.

why 2880 ?

SIMPLE□□=□T – 6.d

Basic FITS (1979-81)



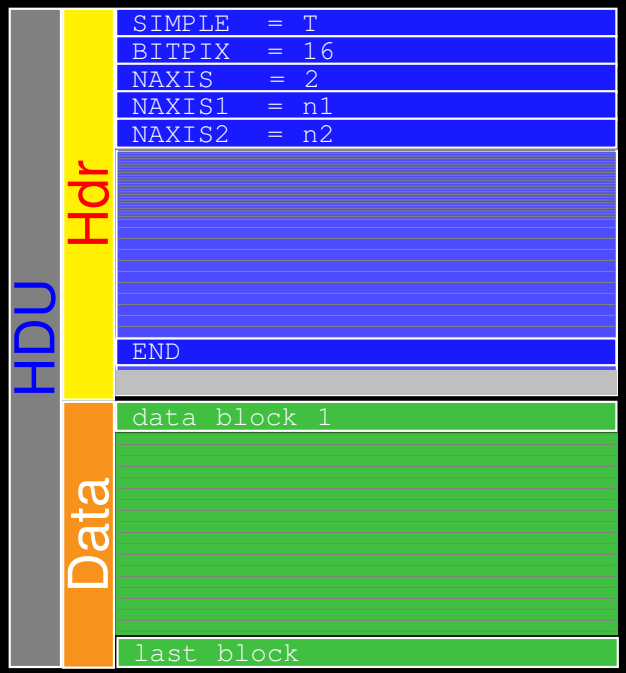
The header is made of keywords contained in a card image of 80 characters. One header block contains up to 36 kwds.
Kwds may contain numeric, boolean or string values, or be valueless (HISTORY, COMMENT, END), and use a restricted subset of ASCII.

why card images ?

why ASCII ?

SIMPLE□□=□T – 6.e

Basic FITS (1979-81)

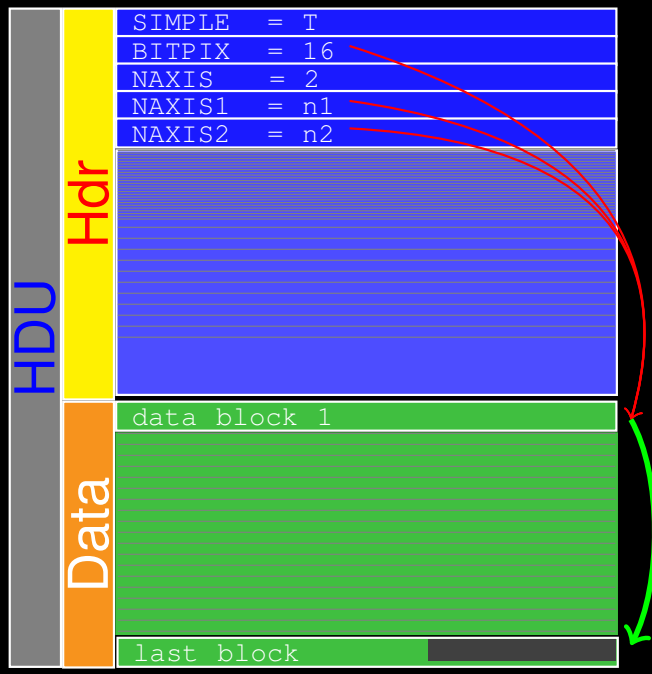


The header terminates with an END kwd.

The rest of the last block is filled with blanks.

SIMPLE□□=□T – 6.f

Basic FITS (1979-81)

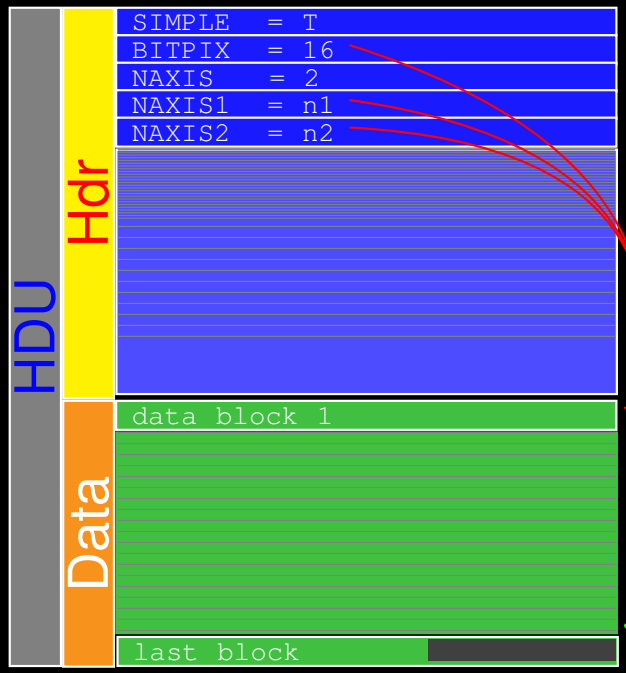


Data occupy (in byte)
|BITPIX|*(NAXIS1*...*NAXISm))/8
a space depending on the *dimensions* and the *type* of the image.

The rest of last block is filled with null (x'00')

SIMPLE□□=□T – 6.g

Basic FITS (1979-81)



The rest of last block is filled with null (x'00')

Data occupy (in byte)
|BITPIX|*(NAXIS1*...*NAXISm))/8
a space depending on the *dimensions* and the *type* of the image.

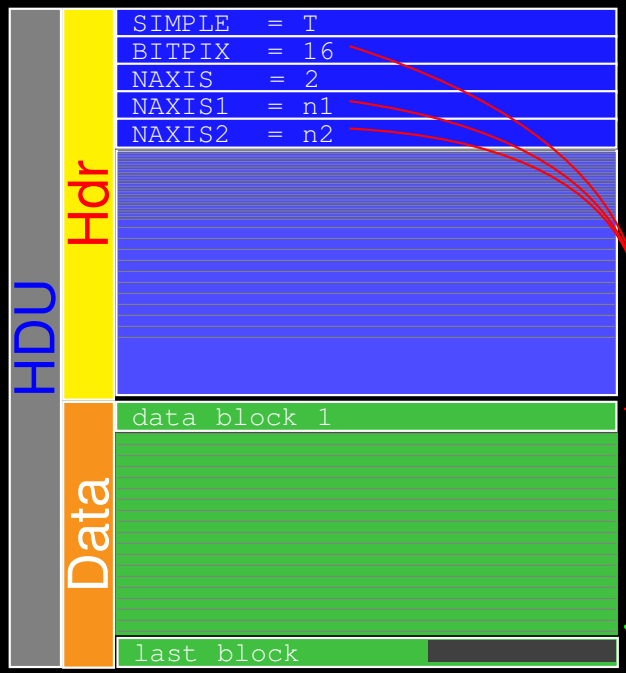
Images can have NAXIS=1...999
NAXIS=2 Usually 2-d
NAXIS=1 But also 1-d spectra

sometimes NAXIS=2 NAXIS1=nx NAXIS2=1

NAXIS=3 Data cubes (stack)
NAXIS=2 Also 2-d stacks of 1-d spectra

SIMPLE□□=□T – 6.h

Basic FITS (1979-81)



The rest of last block is filled with null (x'00')

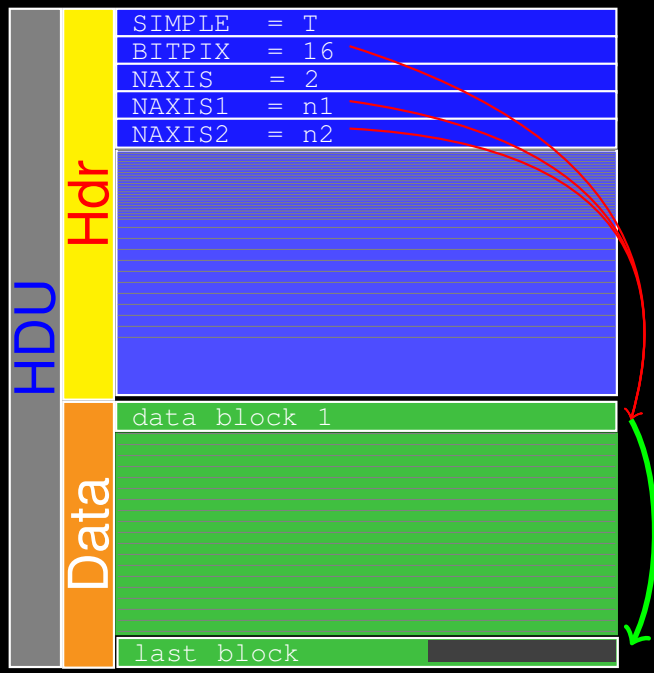
Images are allowed as

- BITPIX=8 unsigned byte
- BITPIX=16 16-bit integer (2-complement)
- BITPIX=32 32-bit integer
- BITPIX=64 64-bit integer
- BITPIX=-32 32-bit real (IEEE)
- BITPIX=-64 64-bit real (IEEE)

why 2-complement or IEEE ?

SIMPLE□□=□T – 6.i

Basic FITS (1979-81)



The rest of last block is filled with null (x'00')

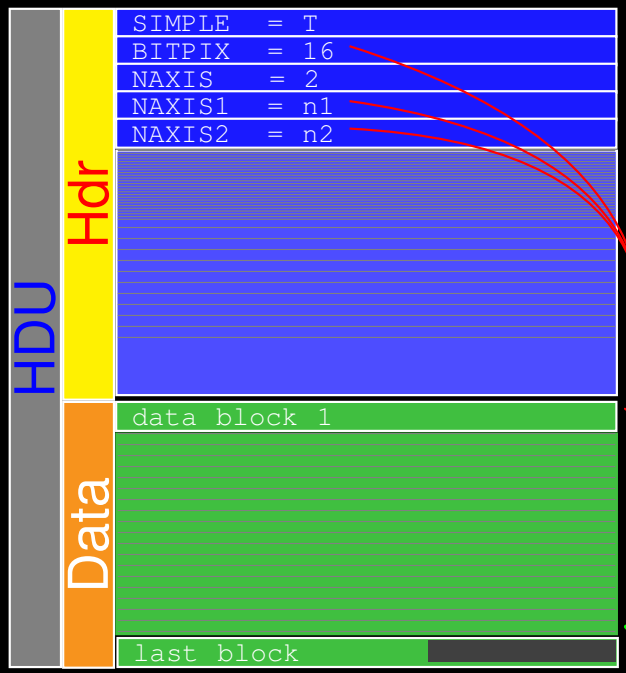
Images are allowed as

- BITPIX=8 unsigned byte
- BITPIX=16 16-bit integer (2-complement)
- BITPIX=32 32-bit integer
- BITPIX=64 64-bit integer
- BITPIX=-32 32-bit real (IEEE)
- BITPIX=-64 64-bit real (IEEE)

why 2-complement or IEEE ?

SIMPLE□□=□T – 6.j

Basic FITS (1979-81) + Floating Point (1990)



Images are allowed as

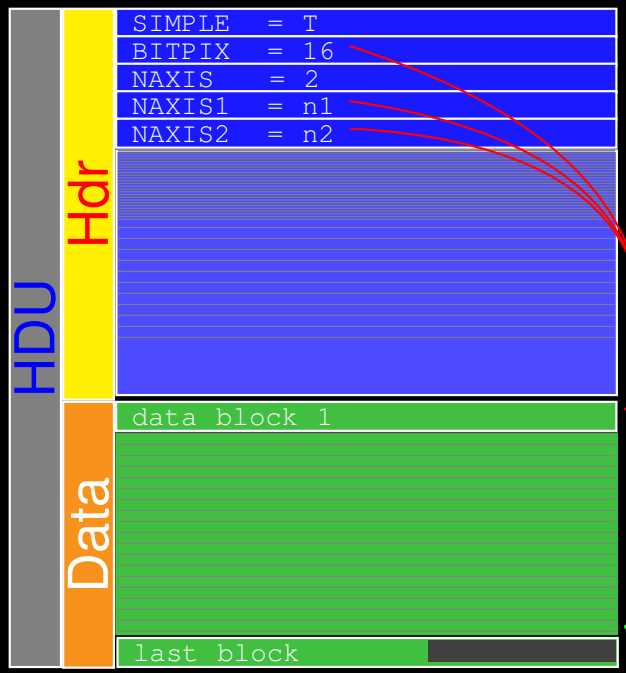
- BITPIX=8 unsigned byte
- BITPIX=16 16-bit integer (2-complement)
- BITPIX=32 32-bit integer
- BITPIX=64 64-bit integer
- BITPIX=-32 32-bit real (IEEE)
- BITPIX=-64 64-bit real (IEEE)

why 2-complement or IEEE ?

The rest of last block is filled with null (x'00')

SIMPLE□□=□T – 6.k

Basic FITS (1979-81) + Floating Point (1990) + 64-bit integer (2005)



Images are allowed as

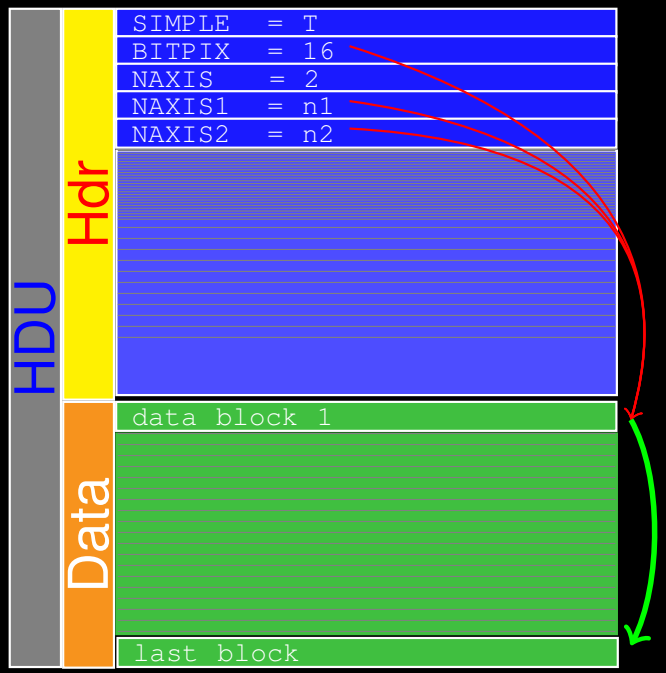
- BITPIX=8 unsigned byte
- BITPIX=16 16-bit integer (2-complement)
- BITPIX=32 32-bit integer
- BITPIX=64 64-bit integer
- BITPIX=-32 32-bit real (IEEE)
- BITPIX=-64 64-bit real (IEEE)

why 2-complement or IEEE ?

The rest of last block is filled with null (x'00')

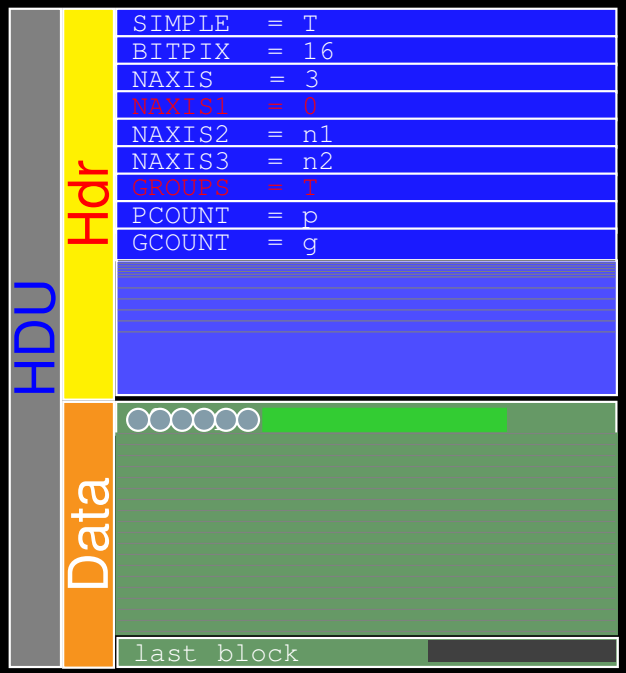
SIMPLE□□=□T – 6.I

Basic FITS (1979-81) + Floating Point (1990) + 64-bit integer (2005)



SIMPLE□□=□T – 6.m

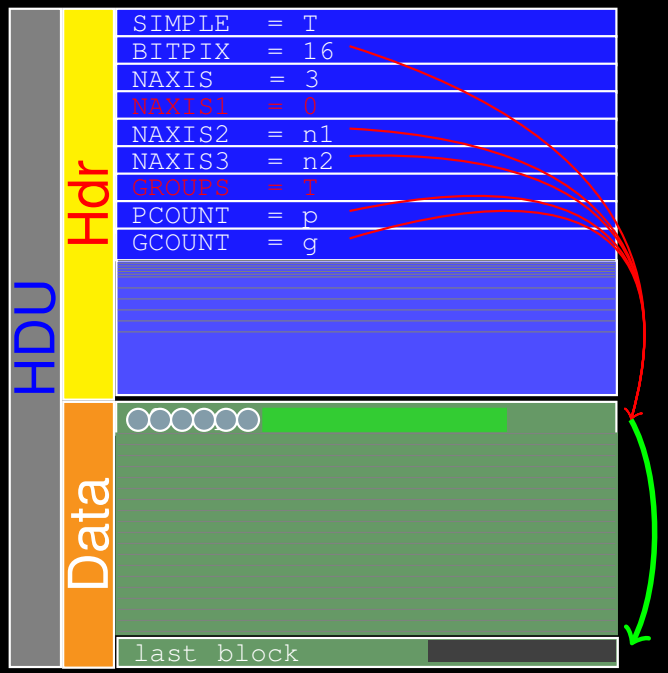
Random Groups (1981)



An obscure structure used by radioastronomers (now deprecated for other usages) . . .

SIMPLE□□=□T – 6.n

Random Groups (1981)



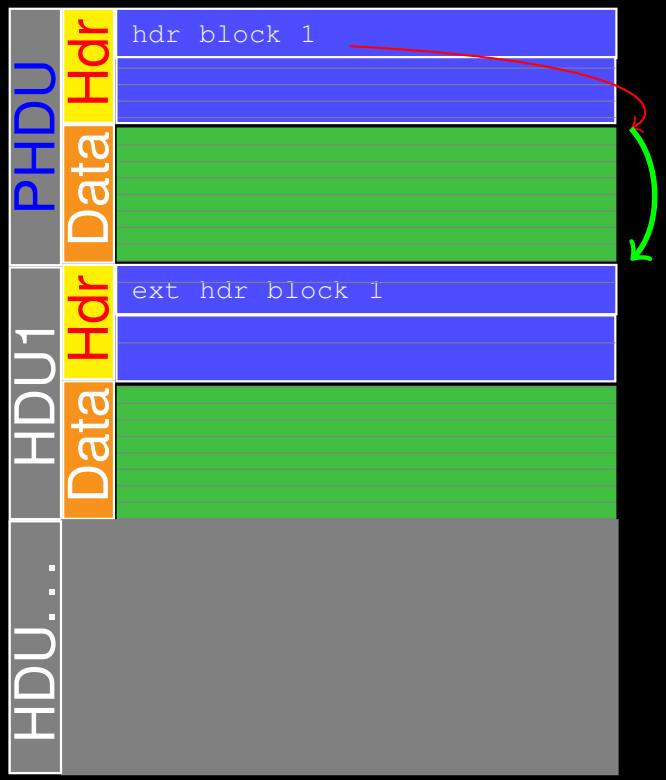
An obscure structure used by radioastronomers (now deprecated for other usages) . . . allowing to repeat g times sequences of p parameters and a mini-data-array with dimension NAXIS-1=m-1

Data occupy (in byte)

|BITPIX|*GCOUNT*
(PCOUNT+(NAXIS2*...*NAXISm))/8

SIMPLE□□=□T – 6.0

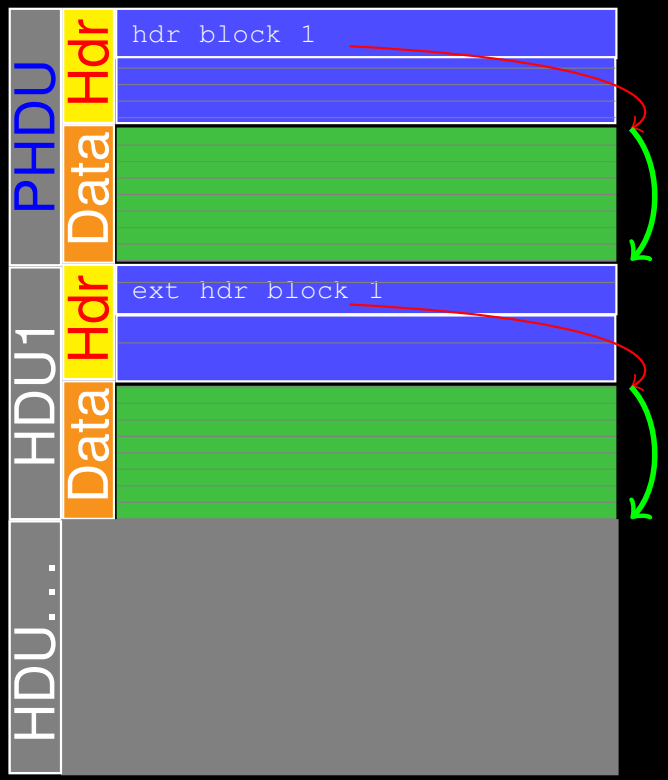
Generalized ("conforming") extensions (1988-)



After the first ("primary") HDU one can insert another

SIMPLE□□=□T – 6.p

Generalized ("conforming") extensions (1988-)

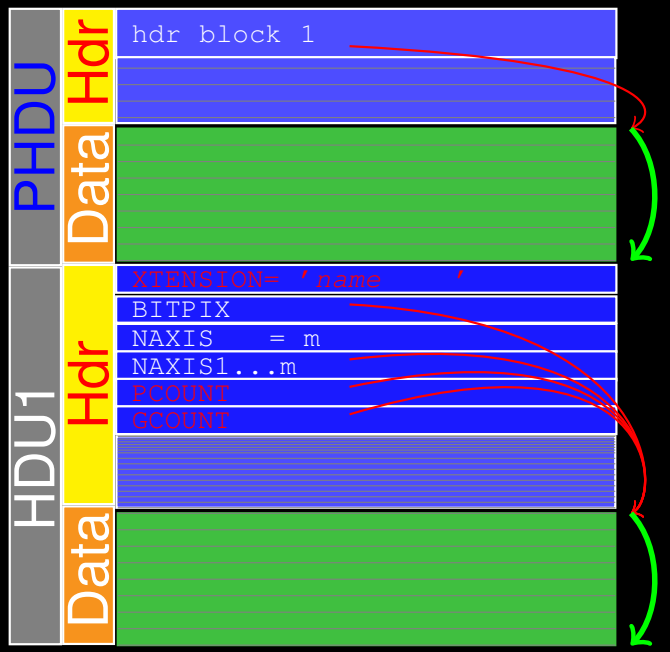


After the first ("primary") HDU one can insert another
and potentially more again . . .

Let's concentrate on the first HDU and forget the other (they are all equivalent).

SIMPLE□□=□T – 6.q

Generalized ("conforming") extensions (1988-)



In an extension data occupy (in byte)
|BITPIX|*GCOUNT*
(PCOUNT+(NAXIS1*...*NAXISm))/8

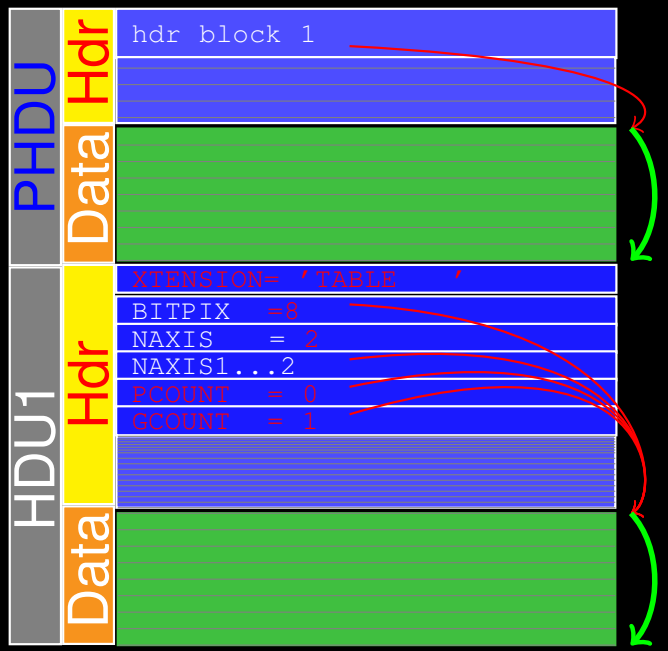
Not all extensions are *part* of the standard, but names shall be registered with IAUFWG.

There are 3 types of *standard extensions*

- TABLE extension
- IMAGE extension
- BINTABLE extension

SIMPLE□□=□T – 6.r

ASCII TABLE extensions (1988)



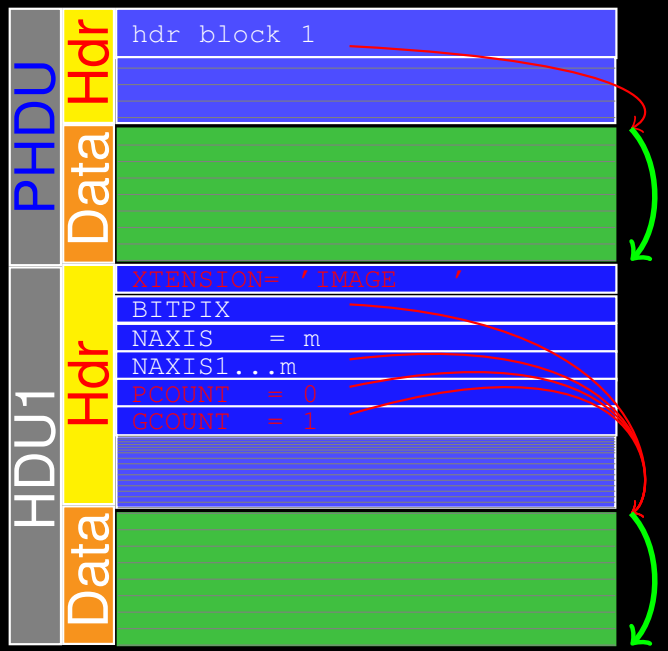
In an extension data occupy (in byte)
|BITPIX|*GCOUNT*
(PCOUNT+(NAXIS1*...*NAXISm))/8

Not all extensions are *part* of the standard, but names shall be registered with IAUFWG.

TABLE extension
contains an ASCII tables
not terribly efficient
(historically) used for catalogues

SIMPLE□□=□T – 6.s

IMAGE extensions (1994)



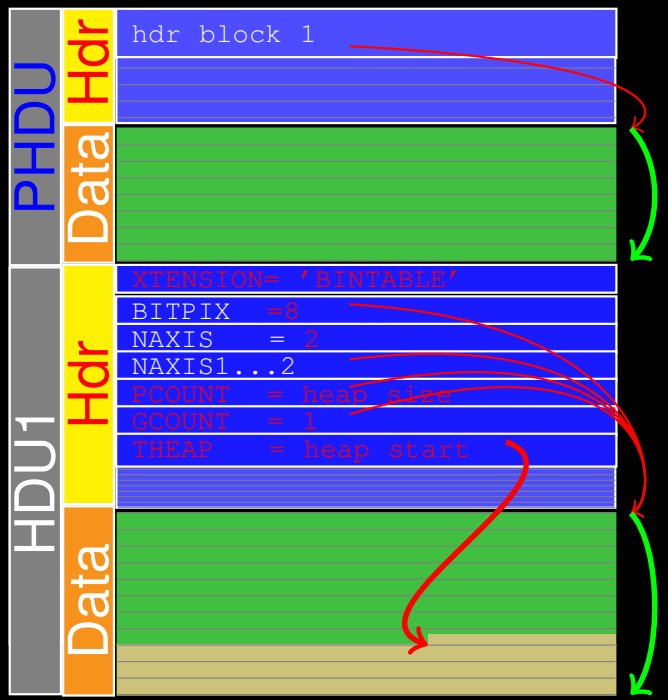
In an extension data occupy (in byte)
|BITPIX|*GCOUNT*
(PCOUNT+(NAXIS1*...*NAXISm))/8

Not all extensions are *part* of the standard, but names shall be registered with IAUFWG.

IMAGE extension
Each extension contains a normal m-dimensional image
developed initially for IUE

SIMPLE□□=□T – 6.t

Binary table BINTABLE extensions (1995)



In an extension data occupy (in byte)
|BITPIX|*GCOUNT*
(PCOUNT+(NAXIS1*...*NAXISm))/8

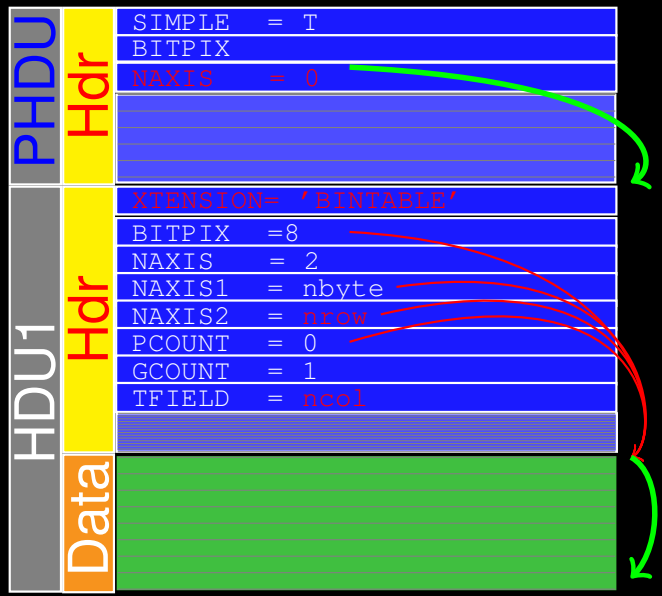
Not all extensions are *part* of the standard, but names shall be registered with IAUFWG.

BINTABLE extension
Each estension contains a binary table with NAXIS2 rows and TFIELD columns.
Each row of NAXIS1 bytes
columns may have a n-dimensional depth
fixed or variable

extremely flexible

not only images . . . – 7.a

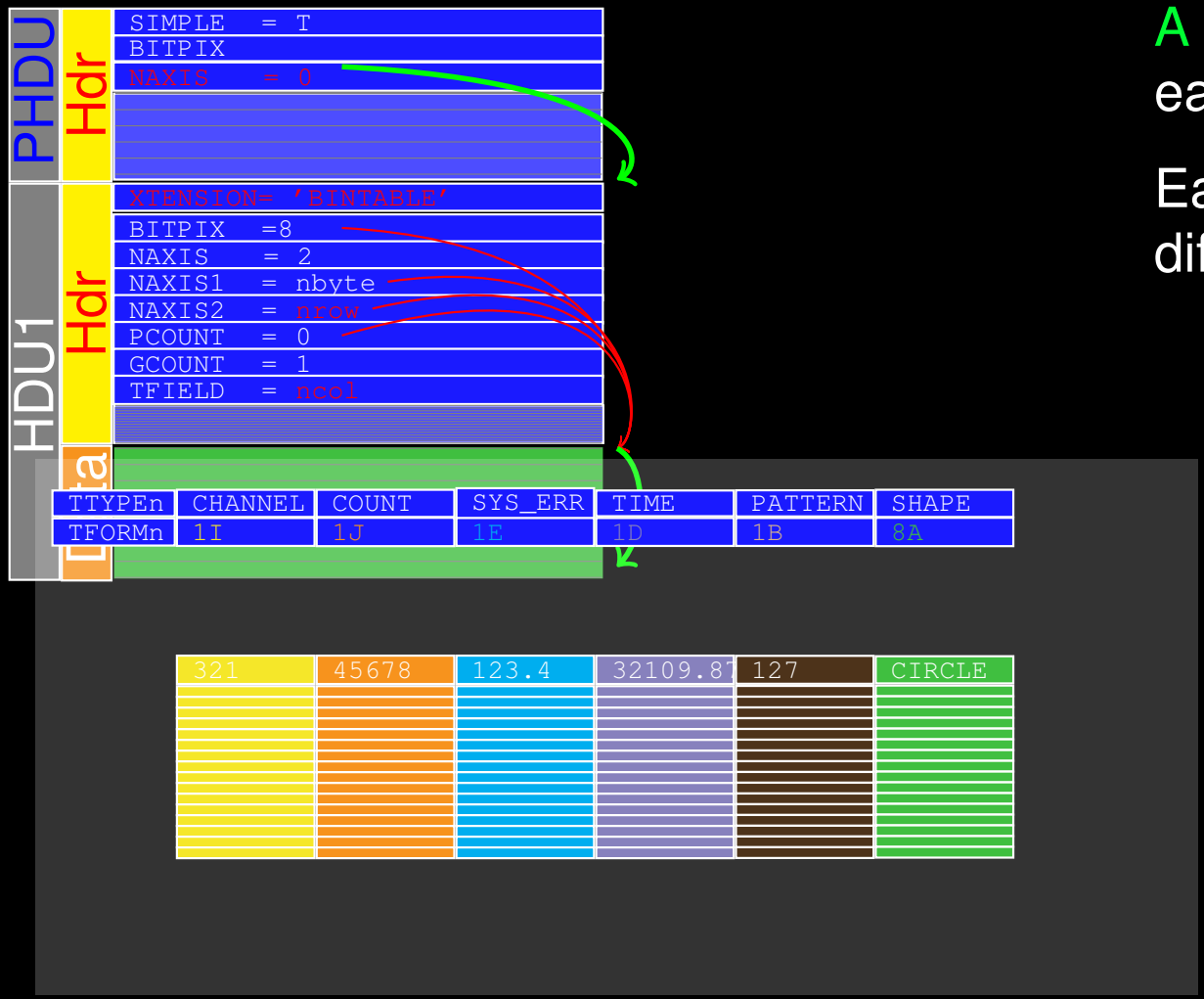
A FITS file may be composed only by extensions



In order to have only extensions the Primary HDU shall not contain a data array (but have a valid, more or less "rich", header)

not only images . . . – 7.b

A FITS file may be composed only by extensions



A simple table

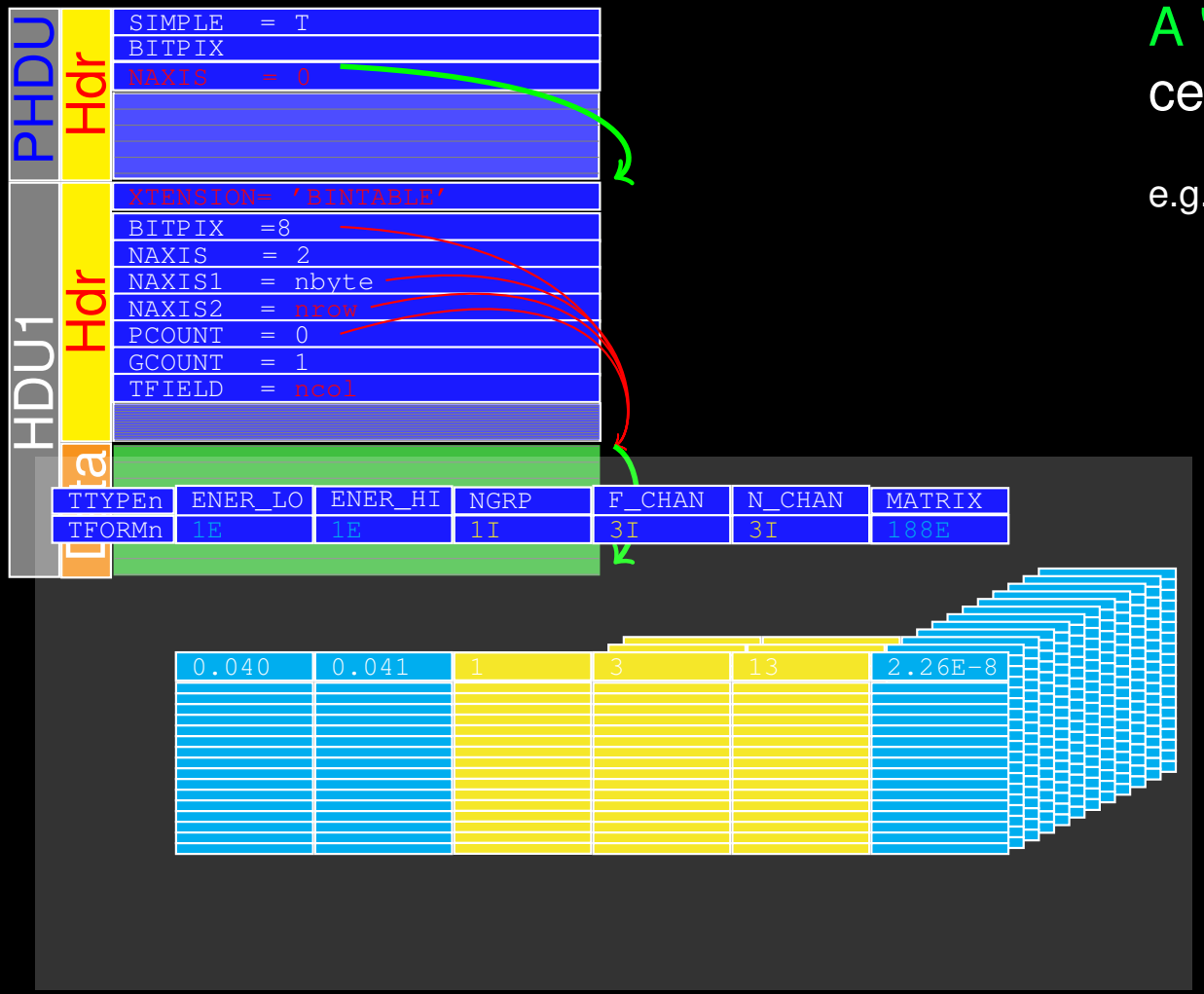
each cell is a scalar

Each column may be of a different type

- B unsigned byte
- I 16-bit integer
- J 32-bit integer
- K 64-bit integer
- E 32-bit real
- D 64-bit real
- A character string
- C,M 32- or-64-bit complex
- L,X logical or bit array
- P,Q 32- or 64-bit array descriptor

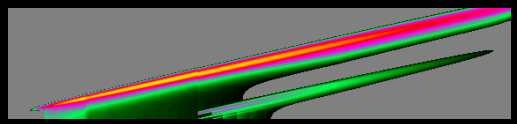
not only images . . . – 7.c

A FITS file may be composed only by extensions



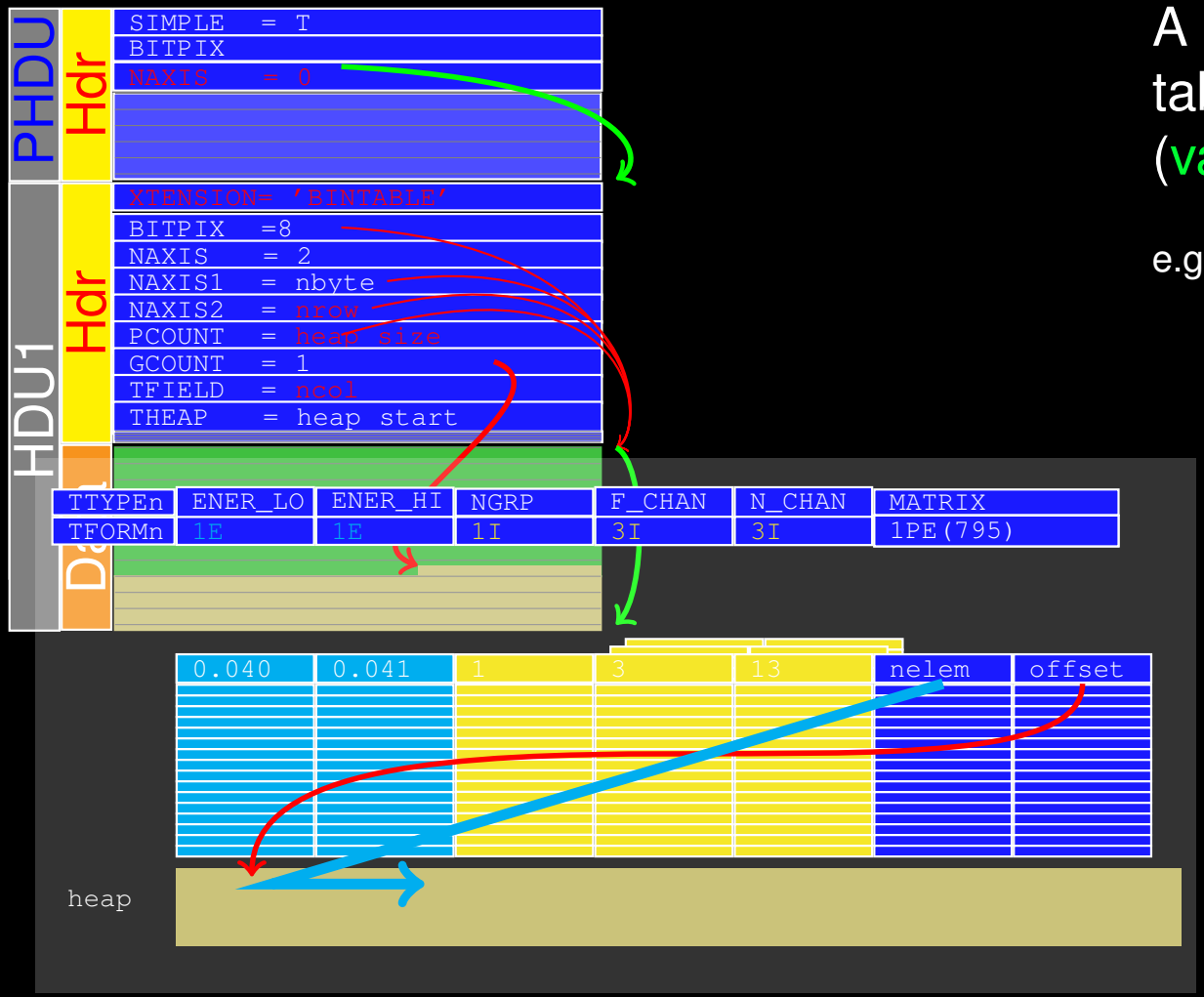
A "tridimensional" table
cells may have a fixed depth

e.g. spectral response matrices



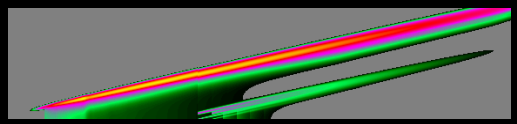
not only images . . . – 7.d

A FITS file may be composed only by extensions



A complex "tridimensional" table
(variable length arrays)

e.g. spectral response matrices



What are data for us: – 8.a

The FITS format copes with:

- images
- but not only images
- metadata

What are data for us: images – 8.b

A (FITS) 2-d image is for members of the astronomical community any data array z(x,y) where:

- z is any *physical quantity* . . .
- . . . usually linked to flux or brightness
  e.g. a flux in $erg/cm^2/s$, a magnitude, a number of photons
  but also ADU on a CCD, or density on a photographic plate
- x,y usually depend on mm positions on some detector focal plane . . .
- . . . and usually map to sky coordinates
- . . . but one coordinate can be a *dispersion axis* (λ)
  e.g. for multi-object spectrographs
- but z(x,y) can be anything more or less exotic
  (e.g. χ^2 vs spectral index and N_H, detector response matrices,
  Fourier transforms or periodograms vs energy . . .)

What are data for us: not only images – 8.c

- Images can have $n \neq 2$ dimensions
- a spectrum $f(\lambda)$ or any histogram can be saved as a 1-d image (but also in tabular form)
- e.g. associate a convenience pixel-by-pixel quality flag to an image as a second plane (NX*NY*2)
  (but also as a separate IMAGE extension)
- e.g. associate a time sequence of images (third axis in the *data cube* is time)
  but time profiles and time-tagged photon lists are binary tables
- associate different physical parameters on the same spatial frame (e.g. Stokes I Q U V)
- binary tables can be used for many purposes
  spectra, time profiles, catalogues ...
- alternate usage of binary tables vs images
  (e.g. response matrices, images as table cells,
  tables as support structures for image compression (e.g. tiled image convention

What are data for us: metadata – 8.d

For what concerns header keywords (*metadata*)

- standard shall be strictly followed for mandatory kwds
- freedom to develop (and document) your own convention
- WCS is important for us to map z,x,y to physical units
  not just and not only x,y to sky coordinates
  e.g. one can map pixels to linear coordinates (mm)

  An image can have up to 27 alternate WCSs

- separate keywords *required by software* from those intended as *human readable*
- usage of commentary (valueless) keywords and/or keyword comments
- inheritance of headers across extension HDUs
- headers are not easily extendable (without rewriting data which follows):

  store metadata in header of a final, dataless extension or use ad-hoc extensions

Extensions vs conventions – 9.a

- *a standard extension is a conforming extension whose organization and content are completely specified in [the] standard*

  only TABLE, IMAGE, BINTABLE

- registered extensions are not approved/endorsed by IAU FWG: in this respect they are just another *convention*
  FOREIGN is the only defined, implemented and separate extension

  IUEIMAGE and A3DTABLE are old versions of IMAGE and BINTABLE

- conventions may or may not be *registered* with IAU FWG (for documentary purpose only)
- conventions specify a set of FITS header keywords, and optionally, other data structures within a FITS file

Extensions vs conventions – 9.b

http://fits.gsfc.nasa.gov/fits_registry.html

The Registry of FITS Conventions

Maintained by the IAU FITS Working Group

(See also the local FITS conventions , the registered extension names, and keyword dictionaries).

The Registry of FITS Conventions provides a central and authoritative repository for documenting conventions that have been developed by the FITS user community for storing and transmitting various types of information in FITS format data files. A FITS convention is defined as a set of related FITS header keywords, and optionally, other data structures within FITS tables, FITS images, or other types of conforming FITS extensions that are to be used for a specific purpose. The IAU FITS Working Group is responsible for the this Registry and for the rules and procedures for entering new conventions into it.

These conventions are not necessarily recommended by the IAU FITS Working Group for reuse in new applications. The registration process is mainly designed to ensure that the documentation about the existing FITS convention meets a minimum level of completeness and clarity. A separate and more rigorous review process is required before a FITS convention is endorsed by the IAUFWG and is approved as part of the FITS Standard.

Conventions currently under review for inclusion in the Registry:

- TPV World Coordinate System is a non-standard convention that builds on the standard TAN projection by adding a general polynomial distortion correction.
- ZPX World Coordinate System is a non-standard convention that builds on the standard ZPN projection by adding a general polynomial distortion correction.
- Spatial Region File convention defines a spatial region of a 2-dimensional image. The region file is often used to define an area that is to be included or excluded from certain data processing operations on the image. The region is specified as the union or intersection of geometric shapes, such as 'circle' or 'rectangle'.
- SDFITS binary table convention for interchange of single dish data in radio astronomy.
- Green Bank Keyword convention for recording parameters related to images that are stored in a vector column of a FITS binary table.

Registered Conventions:

- FITS Interferometry Data Interchange (FITS-IDI) Convention for the interchange of data recorded by interferometric telescopes, particularly at radio frequencies and very long baselines.
- ESO HIERARCH Keyword convention uses a hierarchical structure to define the keyword name. This convention can be generalized to support keyword names longer than 8 characters or containing characters that would not be allowed in a standard FITS keyword name.
- Substring Array Convention for Binary Tables may be used to specify that a character array field (TFORMn = 'rA') consists of an array of either fixed-length or variable-length substrings within the field.
- Simple Imaging Polynomial convention provides a convenient means of representing non-linear geometric distortion of the coordinate system as polynomials in FITS header keywords.
- TNX World Coordinate System is a non-standard coordinate system for evaluating celestial coordinates from the image pixel coordinates. It follows the FITS conventions for undistorted tangent plane projections but adds a non-linear distortion term to the evaluation.
- Euro3D: An interchange data format for integral field spectroscopy in which 1-dimensional spectra are obtained at multiple positions over a 2-dimensional spatial field of view.
- CONTINUE Long String Keyword convention for writing string keyword values that are longer than the 68-character limit of a single FITS keyword.
- OIFITS: A Data Standard for Optical Interferometry.
- CHECKSUM keyword convention for verifying the integrity of FITS HDUs.
- Multi-Beam FITS (MBFITS) data format for single-dish mm/submm telescopes.
- Column Limits keywords (TLMINn/TLMAXn & TDMINn/TDMAXn).
- Tiled Image Compression divides image into a grid of tiles, and stores the compressed tiles in a variable length array column of a binary table.
- INHERIT keyword indicates that a HDU should inherit the primary header keywords.
- Hierarchical Grouping for defining hierarchical associations of HDUs.
- FOREIGN file encapsulation for wrapping other types of files in FITS.

What can FITS do for BAV ? – 10.a

. . . and what does BAV plan to do with FITS ?

- just store or distribute documents ?
- document data layout
- technicalities about colour
- data compression
- metadata

What can FITS do for BAV ? – 10.b

- Does BAV intend to *distribute* documents in FITS ?
  Or just use as a (long-term, stable, permanent) *deep store* ?
  and eventually convert to "temporary" (fashionable) distribution formats ?
  IIIF modular approach (http://lib.stanford.edu/iiif) ?
- Does BAV plan to keep one scanned document page per file or per `IMAGE` extension ?
  this might allow a single set of shared metadata
  but files may become large (⇒ compression ▸)
- a *data cube* minimizes the overhead of header keywords
  w.r.t. separate `IMAGE` extensions (which can be more elegant)
- What kind of (custom) viewer will the typical scholar use to view FITS ?
  astronomical ones may not be adequate
  e.g. visualization of data cubes usually occurs one 2-d layer at a time (but see DS9 7.0)

What can FITS do for BAV ? – 10.c

How will be document scans be arranged ?

- always colour (multispectral scan) or only for "colorful" documents ?
- store as 8-bit RGB data cube ?

  see previous considerations on keyword overhead
- *FITS, unlike "picture" formats like jpg and tiff, has no* native *convention for storing colour information. It would likely be necessary to define new FITS conventions (with our help, hopefully) to encode the color information.*
  - in no case z(x,y) or image dimensionality has for astronomical images a native interpretation in terms of "typographic" colours (RGB, CMYK)
  - ITTs and LUTs (colormaps) are just used for display purposes as a graphical artifact at discretion of the user
  - e.g. false colour images where each "RGB colour" is intensity in one energy band ? (see CHANDRA example)
- interested in linear WCS (pixel/mm) ?

What can FITS do for BAV ? – 10.d

BAV is projecting a huge data volume (even by astronomical standards) so data compression will be of great interest.

We recommend benchmark studies with some actual manuscript image scans to see what technique is most effective.

Our tiled-image convention (layered onto `BINTABLE`) might give the best compression, but it has the *drawback that the format is complex* and they would need to rely on *existing software libraries like CFITSIO* to read the images.

This could be an issue in long term preservation of the images.
(however it is documented in journals)

convention registry entry	Tiled Image Compression Convention
PASP, 121, 414 (2009)	Lossless compression algorithms (Pence, Seaman & White 2009)
PASP, 122, 1065 (2010)	Lossy compression algorithm (Pence, White & Seaman 2010)

What can FITS do for BAV ? – 10.e

FITS is a metadata-rich format and document curation and preservation requires metadata handling in addition to imaging.

- however FITS kwds are currently limited (by OFAF) to a subset of ASCII (and 68-char limit)
  less problems for dedicated extensions
- Libraries (including BAV) may need to use other standards like PREMIS or non-ASCII (UNICODE) stuff
- BAV may want to include other metadata (e.g. from TIFF) in header or extension

- the `FOREIGN` non-standard extension can encapsulate any format but . . .
  The essence of the idea to use FITS is as a very stable, well defined, easy to reverse engineer format for very deep storage over the decades. `FOREIGN`, on the other hand, is a way to insert a third-party standard into a FITS stream. If that third-party standard is not as stable as FITS, the permanence is sacrificed.
- tradeoff study on ease of use (and manpower needed) in defining own convention and writing eventual converters

What can FITS do for BAV ? – 10.f

IAU FWG is willing to help

We have established a small technical panel

- W.Pence (NASA GSFC)
- L.Chiappetti (INAF IASF-MI)
- A.Dobrzycki (ESO)
- R.Seaman (NOAO)
- ...

. . . and I could set up a *mailman* exploder to discuss via e-mail with BAV-designated people.

SIMPLE = T

```
SIMPLE  =                    T
BITPIX  =                    8
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
PCOUNT  =                    0
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
PCOUNT  =                    0
GCOUNT  =                    1
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
PCOUNT  =                    0
GCOUNT  =                    1
TFIELDS =                    2
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
PCOUNT  =                    0
GCOUNT  =                    1
TFIELDS =                    2
TFORM1  = 'J       '
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
PCOUNT  =                    0
GCOUNT  =                    1
TFIELDS =                    2
TFORM1  = 'J       '
TFORM2  = 'D       '
```

```
SIMPLE  =                    T
BITPIX  =                    8
NAXIS   =                    0
EXTEND  =                    T
END
XTENSION= 'BINTABLE'
BITPIX  =                    8
NAXIS   =                    2
NAXIS1  =                   12
NAXIS2  =                  528
PCOUNT  =                    0
GCOUNT  =                    1
TFIELDS =                    2
TFORM1  = 'J       '
TFORM2  = 'D       '
HISTORY Terminated at 2012-07-05T11:45:00
END
```

SIMPLE = T

HISTORY Terminated at 2012-07-05T11:45:00
END

What is BEAMER and what is TikZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!

TikZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TikZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

What is BEAMER and what is TIKZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!
TIKZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TIKZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

What is BEAMER and what is TIKZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!
TIKZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TIKZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

What is BEAMER and what is TIKZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!
TIKZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TIKZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

What is BEAMER and what is TIKZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!
TIKZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TIKZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

What is BEAMER and what is TIKZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!
TIKZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TIKZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

What is BEAMER and what is TIKZ ?

and how was this presentation prepared ?

BEAMER is a *LaTeX class* which allows to produce "interactive" presentations in PDF format ... which means they are intrinsically portable!
TIKZ is a *LaTeX package* which allows to design graphics in LaTeX (no WYSIWYG but high reproducibility)

- BEAMER can be downloaded from bitbucket
- TIKZ can be also downloaded from sourceforge
- there is a nice manual
- some examples are available online
- there are also online resources which discuss and compare the merits of other LaTeX and PDF presentation preparation tools, like here (Berlin) and here (Akron)

Happy birthday, FITS !

another form of *preserved document*

Newsgroups: sci.astro.fits
From: dwells@azalea.cx.nrao.edu (Don Wells)
Subject: Happy Birthday, FITS! [LONG]
Organization: National Radio Astronomy Observatory, Charlottesville, VA.
Date: Sun, 29 Mar 1992 05:48:47 GMT

Happy Birthday, FITS!

I consider March 28th to be the birthday for FITS, because it was on March 28, 1979, that Eric Greisen and I completed the Basic FITS Agreement. FITS is now a teenager - today is its 13th birthday.

-- Coincidences of Historic Events --

Afficionados of historical trivia will want to know that *three* famous events happened on March 28, 1979:

* Birthdate of the Basic FITS Agreement
* Conservatives win British election; Margaret *Thacher* new Prime Minister
* Nuclear power plant accident at Three Mile Island, Pennsylvania, releases radioactivity

full articles GSFC NRAO 18th back

Once upon a time – a

Some FITS features depend on the epoch when they were defined
(and *once FITS forever FITS*)

why 2880 byte ?

2880 is the least common multiple of 12, 16, 18, 24, 32, 36, 48, 60 (64) and in 1979 there were computers with memory words of such *bit* lengths.

2880 *byte* records could be reasonably easily unpacked on all machines where they could not be read natively.

A 1:1 straight blocking was initially enough efficient for the usage of magnetic *tapes*, although later a larger blocking was used e.g. 10:1 (28800 byte).
See (deprecated) blocking agreement (1994)

back

Once upon a time – b

Some FITS features depend on the epoch when they were defined
(and *once FITS forever FITS*)

why "card images" ?

In the last standard just called "*keyword records*".

Punched cards were 80-character long

Each header block is composed by 36 card images (2880=36*80)

Some mandatory keywords still have compulsory fixed format

back

Once upon a time – c

Some FITS features depend on the epoch when they were defined
(and *once FITS forever FITS*)

why ASCII, 2-complement, IEEE ?

Because FITS people were luckly . . . or far-sighted

Today all this is (still) obvious (but Unicode . . .), but before the '80s people used EBCDIC or other, 1-complement, and proprietary floating point format.

Initially FITS used only integers, scaled with `BSCALE` and `BZERO`. The floating point agreement dates to 1990 . . . since 1981 almost all *new* systems use IEEE.

back

Once upon a time – d

Some FITS features depend on the epoch when they were defined
(and *once FITS forever FITS*)

why ASCII, 2-complement, IEEE ?

Because FITS people were luckly . . . or far-sighted

Today all this is (still) obvious (but Unicode . . .), but before the '80s people used EBCDIC or other, 1-complement, and proprietary floating point format.

Initially FITS used only integers, scaled with `BSCALE` and `BZERO`. The floating point agreement dates to 1990 . . . since 1981 almost all *new* systems use IEEE.

and why *big endian* ?
because not all doughnuts come out with a hole !
Big endian order was prevailing on Unix (e.g. Sun) even if it requires byteswap on Intel machines (Linux, Windows . . .) etc. However it is the *network byte order*.

back

Once upon a time – e

Some FITS features depend on the epoch when they were defined (and *once FITS forever FITS*)

deprecated features

Remain valid but shall not be used in new structures

- Random Groups except for radio interferometry
- DATE format DD/MM/YY after 2000
- "implicit decimal point" format
- CROTAj keywords
- minor obscure variants of WCS keywords

back

A fantasy header

```
SIMPLE  =                    T
BITPIX  =                    8
...
DATE    = '2012-06-02T15:20:00'          / Date file was written
ORIGIN  = 'BAV'                          / Creating institution
INSTRUME= 'Metis DRS 5070'               / scanning instrument
CALPASS = T                              / calibration passed
...
TITLE   = 'Almagestum'                   / Common title
TITLE2  = 'Mathematike Syntaxis'         / Alternate title
AUTHOR  = 'Claudius Ptolomaeus'          / Author
ORIGLANG= 'Greek'                        / Original language
LIBCODE = 'T CONS 456'                   / Library code
LIBROOM = 'I Graecorum'                  / Library room
LIBSHELF= 'IV gradus'                    / Library shelf
ACQUIRED= '1298-06-29'                   / Date of acquisition
ACQUIRER= 'Paulus Ariminensis'           / Responsible of acquisition
...
COMMENT   manuscript is badly burnt
HISTORY   original scan on 2012-05-29 at 14:22
HISTORY   calibration on 2012-05-30 at 09:47
HISTORY   conversion to FITS on 2012-06-01 at 11:23
...
```

◂ back