English translation V2.1

GRAPPA3E

GAIA EDR3 Restricted for Astrometry and Photometry Provided by/for "Amateurs"

Edition "Perseverance"[1]

The GRAPPA3E database comes from the EDR3 delivery of ESA's GAIA mission. The content of this delivery is described here.

Acknowledgments

Many thanks to the GAIA teams who made these informations available. A big thank you also to Dave Herald for giving access to OCCULT4's data and Paolo Tanga for his advices in using GAIA data.

If you intend to use this database, please, read first : Acknowledgement and credit_and_citation.

Table of contents

1 Many unprecedented computer problems had to be fixed to create this version of GRAPPA. Finally, the last action taken to compile the database for the last time was done on February 21, 2021 around 9:40 pm, when the "Perseverance" rover successfully landed on Mars!

Introduction

The main goal of GRAPPA3E is to make the essential data of the GAIA catalog available offline. Thus downloaded to a workstation, no internet link will be necessary to access to the high accuracy of the astrometric and photometric data of the full Early 3 version catalog. Access to the data will be accelerated and will allow it to be used more intensively.

Contents

In order to build the GRAPPA3E database, the GAIA EDR3 Archive database source was used and partially translated by condensing in binary and structured form the contents of some 3386 compressed csv source files. In order to accelerate the search by position, the corresponding records were no longer redistributed in the ascending source_id order but by zone in the sky of 1° x 1° in right ascension and declination. All of the 1,811,709,771 GAIA EDR3 sources are recorded in the database. Some other additional files were used in order to establish links to the Hipparcos2 and Tycho2DSC catalogs. The GrappaVar.dat file has been built from Gaia DR3 data and especially from files in the *vari_classifier_definition*, v*ari_classifier_class_definition*, *vari_classifier_result* et *vari_summary*.

Sources organisation

The GRAPPA3E files are divided into 1 ° declination bands, except for the 2 polar zones above 85 ° absolute declination where all the stars are grouped into two unique files, one for the north polar cap and the another for the southern one.

The database is structured in 170 directories. The directories are named 5 through 174. The directory name is the declination band plus 90.

In each directory, the stars are recorded in 360 files, each file therefore has a width of 1 ° in Right Ascension (R.A.). Files are named according to the area they represent.

Thus, the **22**-**25** file stored in the "**25**" directory is the area where the stars have a Right Ascension in the interval [**22** °; 23° [and a Declination in the interval [-65 °; -64 °[(because **25**° - 90° = -65 °). See the Data access paragraph.

The polar zones are named N85-90 and S85-90 according to the range of declination they represent.

Therefore, there is a total of 170x360 + 2 = 61202 data files for the GAIA sources.

Star encoding...

In the 170x360+2=61202 data files, the stars are all sorted sequentially according to their Right Ascension (R.A.). Each record is 52 bytes in length. The structure of the record is the same for all data files and can be written in Pascal language like this:

```
TGrappa3ESource = bitpacked record
                          //         Size
                          // Unit   (byte) Definition
     Source_id : QWord;   //           8    Unique within a particular Data Release. See here
     RA        : int32;   // mas       4    R.A. truncated to 1mas precision truncation: +10.123 = 10
     RA2       : byte;    // 4µas      1    R.A. component <1mas. Units of 4 µas. To be added to RA. 30 = 0.120mas
     e_RA      : uint16;  // 10µas     2    Uncertainty in RA units of 10 µas
     DEC       : int32;   // mas       4    Declination+90° truncated 1 mas precision truncation: +10.123 = 10
     DEC2      : byte;    // 4µas      1    DEC component <1mas. Units of 4 µas. To be added to DEC. 30 = 0.120mas
     e_DEC     : uint16;  // 10µas     2    Uncertainty in Dec, units of 10 µas
     parallax  : uint16;  // 12.5µas   2    This unit permits to code Proxima Centauri's parallax
     e_parx    : uint16;  // 10µas     2    Uncertainty in Parallax units of 10 µas
     PM_RA     : int32;   // µas/yr    4    Annual proper motion in RA. Value has been multiplied by Cos(declination)
     e_PM_RA   : uint16;  // µas/yr    2    Uncertainty in Proper motion in RA
     PM_DEC    : int32;   // µas/yr    4    Annual proper motion in Dec
     e_PM_DEC  : uint16;  // µas/yr    2    Uncertainty in Proper motion in Dec
     RUWE      : byte;    //   0.1     1    Value for ruwe indicator See here
                          //                RUWE is set to 0 for sources with only two parameters (only RA & DEC)
     has_RV,              //          1/8   Set if there is a radial velocity for the source
     has_PM_parallax_RUWE,//          1/8   Set if there are PM, parallax and RUWE values for the star
     has_Hipparcos2_id,   //          1/8   Set if an Hipparcos2 crossmatch has been found
     has_TYCHO2DSC_id,    //          1/8   Set if an Tycho2DSC crossmatch has been found
     has_G,has_BP,has_RP, //        3x1/8   Set if magG, magBP, magRP are given respectively
     duplicated_source : boolean;//
     G,BP,RP   : uint32;  // 0.1mmag 3x4   Values for magnitude in the G, BP and RP bands respectively
                          //                including e_G, e_BP, e_RP. Coded as described below.
   end;
```

bitpacked record are record "compacted" on bit scale, that is to say, ordinal types are align on bit boundaries, not on byte, or 32 or 64bits words. The 8 flags are defined only on one byte, record are 52 bytes long or 13 four bytes words in total.

Used data types :
The data types given above, using the Pascal language, are :

QWord : 64bits unsigned integer
int32 : 32bits signed integer (only positive values have been recorded so uint32 is compatible)
uint32 : 32bits unsigned integer
uint16: 16bits unsigned integer
byte : 8bits positive integer

And decoding...

All values refer to Epoch J2016.0 TCB.

Source_ID

The *source_id* is a unique identifier of the source

Right Ascension

Right Ascension and error are decoded as follow, using the Pascal language, for instance R.A. in degree, uncertainty in *mas*:

value:=(RA+RA2/250)/3600000;
uncertainty:=e_RA/100;

Declination

Declination and error are decoded as follow, using the Pascal language, for instance declination in degree, uncertainty in *mas*:

value:=(DEC+DEC2/250)/3600000-90;
uncertainty:=e_DEC/100;

Proper Motions, Parallax and RUWE

The proper motion in R.A. and Declination, the parallax of a source and the RUWE (Renormalised unit weight error) value are given when the *has__PM_parallax_RUWE* flag is ON (TRUE=1). Note that the given **proper motion in R.A. is in fact μa* = μa . cos (declination)**.

For proper motion and error in R.A, in mas:

value:=PM_RA/1000;
uncertainty:=e_PM_RA/1000;

For proper motion and error in Declination, in mas:

value:=PM_DEC/1000;
uncertainty:=e_PM_DEC/1000;

For Parallax and error, in mas:

*value:=parallax*0.0125;*
uncertainty:=e_parallax/100;

For RUWE:

value:=RUWE/10;

Magnitudes

Magnitudes and respective uncertainties are coded in the same int32 (non signed integer on 32bits) field when the relevant flag is ON:

The decoding of magnitudes can be easily done. For instance for G magnitude, using the pascal language:

```
if has_G then begin
   Gmagnitude:=(G AND $3FFFF)/10000;
   Guncertainty:=((G AND $FFFC0000) SHR 18)/10000;
end;
```

Where SHR is the Shift Right operator (or a binary division by 2^18 in this case).

The magnitude is provided by the least significant 18bits of the 32-bit word, so only one logical operation is needed to get its value with a unit of 0.1mmag. The uncertainty is given by the 14 most significant bits of the word.

The same algorithm can be used for BP and RB magnitudes and associated uncertainties.

NOTE: the uncertainties on the magnitudes are not provided in the GAIA EDR3 catalog but have been calculated from the fluxes and errors on the fluxes considering that the reference magnitude is zero (Zero Point in the Vega scale as indicated here for magnitudes). The maximum uncertainty given by the minimum flux has been taken into account.

Radial velocity

The radial velocity of a source is given in a separated file when the *has_RV* flag is ON. Details will be given below.

Hipparcos2 crossmatch

A matched star in the Hipparcos2 catalogue exist for the source when the *has_hipparcos2_id* flag is ON. Details will be given bellow. See hipparcos2_best_neighbour for more information.

TYCHO2DSC crossmatch

A matched star in the Tycho2TDSCmerge catalogue exist for the source when the *has_tycho2dsc_id* flag is ON. Details will be given bellow. See tycho2tdsc_merge_best_neighbour for more information.

Associated files

There are three associated files to the GRAPPA3E database. This files are:

Radial Velocity file

As "only" 7,209,831 sources in the GAIA EDR3 catalog have a defined radial velocity (in fact the same value as in DR2 release), this value has been stored outside the main database. This saves disk space. The **VR.dat** file gives this information, coded as following:

```
TVR = packed record          //    Unity Size(byte) Definition
      source_id: QWord;      //      -     8
      VR : int16;            //    km/s    2       Radial velocity (int16 = -32768..32767)
      e_VR: byte;            //    km/s    1       Uncertainty in Radial Velocity
   end;
```

Where:
Radial velocity is given by the *VR* field in km/s with a signed integer value. Its uncertainty *e_VR* is given in the positive one byte value.
Packing record on byte boundaries means that each new element of a structured type starts on a byte boundary, not depending on processor register size.

The **VR.dat** file is sorted on the GAIA EDR3 *source_id* field in the ascending order.

Hipparcos2 crossmatch file

The **Hipparcos2BestNeighbour.dat** file contains all crossmatches found between GAIA EDR3 sources and Hipparcos2 objets. The search is not symmetric and several GAIA sources can be associated to an Hipparcos2 object. A good crossmatch respects the postion and error of GAIA and the external catalogue. The file has the following structure :

```
THipparcos_index = packed record
      source_id : QWord;
      original_ext_source_id : longint;
      angular_distance : single;
      number_of_neighbours,xm_flag : byte;
   end;
```

Where :
source_id is the GAIA EDR3 source identifier.
original_ext_source_id is an unique identifier in the Hipparcos2 catalogue.
angular_distance is the resulting separation of the best Hipparcos neighbour and the GAIA EDR3 source. The unit is arcsec. The *single* data type is the 4 bytes IEEE floating point.
The *number_of_neighbours* gives the number of neighbour in the GAIA catalogue.
The field *xm_flag* gives details on the crossmatch algorithm. See here for more informations.

Packing record on byte boundaries means that each new element of a structured type starts on a byte boundary, not depending on processor register size.

The **Hipparcos2BestNeighbour.dat** file is sorted on the GAIA EDR3 *source_id* field in the ascending order.

Tycho2DSCmerge crossmatch file

The **Tycho2tdscMergeBestNeighbour.dat** file contains all crossmatches found between GAIA EDR3 sources and the Tycho2DSCmerge objets. The search is not symmetric and several GAIA sources can be associated to an Tycho2DSCmerge object. A good crossmatch respects the position and error of GAIA and the external catalogue.

The file has the following structure:

```
TTycho_index = packed record
      source_id : QWord;
      original_ext_source_id : packed array[1..11] of char;
      angular_distance : single;
      number_of_neighbours,xm_flag : byte;
      tycho2tdsc_merge_oid : longint;
   end;
```

Where
source_id is the GAIA EDR3 source identifier.
original_ext_source_id is a unique alphanumeric identifier in the Tycho2DSCmerge catalog.
angular_distance is the resulting separation of the best Tycho2DSCmerge neighbour and the GAIA EDR3 source. The unit is arcsec. The *single* data type is the 4 bytes IEEE floating point value.
The *number_of_neighbours* gives the number of neighbour in the GAIA catalog.
The field *xm_flag* gives details on the crossmatch algorithm. See here for more informations.
tycho2tdsc_merge_oid field gives a complementary numerical source identifier in the external catalog.

The **Tycho2tdscMergeBestNeighbour.dat** file is sorted on the *source_id* field in the ascending order.

UCAC4 crossmatch file

Thanks to Dave Herald that gave the possibility to accesss to the Occult GAIA EDR3 data, the crossmatches with UCAC4 stars brighter than mag G=16 is available through the **UCAC4-16_crossmatch.dat** file. This file has the following structure :

```
TGrappaUCACIndex = packed record //  Size
      Source_ID : QWord;   //          8    Unique within a particular Data Release.
      CatNum    : uint32;  //          4    Sequential number in catalogue zone (or number in
catalogue)
      match     : byte;    //          1    +1 if UCAC4 identifier based on a poor match
                           //               +2 if Gaia does not contain a proper motion for the star,
                           //               with the UCAC proper motion used to made crossmatching
end;
```

`CatNum : zzznnnnnnn` where `zzz` gives the zone and `nnnnnnn` a sequential number in the zone.

The file is sorted on the ascending order of the Source_ID field.

ATTENTION : due to the lack of place in the GRAPPA3E data structure, there is no way to directly know in GRAPPA if the GAIA source is crossmatched with an UCAC4 star. At present, a search in the **UCAC4-16_crossmatch.dat** file is the only way.

Variability

The GRAPPAVar.dat file contains 9543807 records of 50 bytes under a compact form of the following informations which permit to know if a soucre is seen as variable after Gaia analysis.

The file structure is given in the Pascal laguage as :

```
TGRAPPAVar = bitpacked record
                          //  Unit   Size  Definition
                          //        (byte)
      Source_ID : QWord;  //          8    See https://dms.cosmos.esa.int/COSMOS/doc_fetch.php?
id=2779219
      RA        : int32;  //   mas    4    R.A.truncated to 1mas precision truncation: +10.123 = 10
      DEC       : int32;  //   mas    4    Declination+90° with 1 mas precision truncation: +10.123 =
10
      in_vari_classifier_result,   // 1/8   Set if present in vari_clssifier_result table
      in_vari_rrlyrae,             // 1/8   Set if present in vari_rrlyrae table
      in_vari_cepheid,             // 1/8   Set if present in vari_cepheid table
      in_vari_planetary_transit,   // 1/8   Set if present in vari_planetary_transit table
      in_vari_short_time_scale,    // 1/8   Set if present in vari_short_time_scale table
      in_vari_long_period_variable,// 1/8   Set if present in vari_long_period_variable table
      in_vari_eclipsing_binary,    // 1/8   Set if present in vari_eclipsing_binary table
      in_vari_rotation_modulation, // 1/8   Set if present in vari_rotation_modulation table
      in_vari_ms_ocillator,        // 1/8   Set if present in vari_ms_ocillator table
      in_vari_agn,                 // 1/8   Set if present in vari_agn table
      in_vari_microlens,           // 1/8   Set if present in vari_microlens table
      in_vari_compact_companion,   // 1/8   Set if present in vari_compact_companion table
      dummy1,dummy2,dummy3,dummy4 : boolean;

      num_selected_g_fov,                                              // Nb de mesures Gaia G
      min_mag_g_fov,max_mag_g_fov,mean_mag_g_fov,median_mag_g_fov,     // magnitude*1000
      num_selected_bp,                                                 // Nb de mesures Gaia Gbp
      min_mag_bp,max_mag_bp,mean_mag_bp,median_mag_bp,                 // magnitude*1000
      num_selected_rp,                                                 // Nb de mesures Gaia Grp
      min_mag_rp,max_mag_rp,mean_mag_rp,median_mag_rp  : smallint;     // magnitude*1000

      best_class_score : byte;     //  1
      best_class_name  : shortInt; //  1     Type vari_class

   end;
```

the coding of *best_class_name* is defined by:

```
max_classe = 25;
vari_class : Array[1..max_classe] of string = ('"ACV|CP|MCP|ROAM|ROAP|
SXARI"','"ACYG"','"AGN"','"BCEP"',
                              '"BE|GCAS|SDOR|WR"','"CEP"','"CV"','"DSCT|GDOR|SXPHE"',
                              '"ECL"','"ELL"','"EP"','"LPV"','"MICROLENSING"','"RCB"','"RR"',
                              '"RS"','"S"','"SDB"','"SN"','"SOLAR_LIKE"','"SPB"','"SYST"','"WD"',
                              '"YSO"','"GALAXY"');
```

The field *best_class_score* is a value comprised between the values 0 and 255. It gives the probability of a source to be in the given *vari_class* with the name given by *best_class_name*.

The *source_id*, *RA* and *DEC* data are include in this file because no information of variability are given in the main database Grappa3. They permits to retrieve easily if a source is variable directly from its *source_id* or its coordinates.

Data Download

These data are available on the IMCCE FTP site (https://ftp.imcce.fr/pub/catalogs/GRAPPA3E/) at the Paris Observatory, France, or on my FTP site at ftp://uaib24cesson.ddns.net/. Open this link with a FTP clien like Filezilla for security reasons (TLS 1.3 needed...) and use login **grappa** and password **grappagrappa**. Eleven 7zip compressed files make up the database:

- 10 source data files that can be downloaded and installed according to needs and useful areas. The numbers in the names of the files correspond to the directories of the database and are therefore associated with the declination of the sources:

Filenames	Areas from [δ_{min}	up to δ_{max} [
GRAPPA3E_005-025+Zone Sud.7z	-90	-64
GRAPPA3E_026-033.7z	-64	-56
GRAPPA3E_034-044.7z	-56	-45
GRAPPA3E_045-054.7z	-45	-35
GRAPPA3E_055-062.7z	-35	-27
GRAPPA3E_063-071.7z	-27	-18
GRAPPA3E_072-085.7z	-18	-4
GRAPPA3E_086-105.7z	-4	+16
GRAPPA3E_106-127.7z	+16	+38
GRAPPA3E_128-174+Zone Nord.7z	+38	+90

- A complementary file named *GRAPPA3E_Complements.7Z* which must be downloaded. It contains the radial velocity and the crossmatch data with the ***Hipparcos2, Tycho2DSCmerge*** and ***UCAC4*** catalogs. The file GRAPPAvar.dat is given in a compressed .7Z form.

All the data must be unzipped in the same directory before use in order to restore the structure of the database. Depending the software using the data, it is possible not to download or delete the directories concerning declinations not needed for instance. The full database need around 88Go on disk.

An example of a program using the database is provided with its sources.

Access to database

In addition to certain software that already includes direct access to the GRAPPA E3 database (e.g. PRISM V11, Tycho-Tracker), the *ExtracteurGRAPPAE3.exe* utility provided on the ftp site allows you to search for sources in the GRAPPA catalog and generate lists of objects with more or less informations (source_ID, astrometric, photometric, proper motions, parallaxes…) and according to some criteria (area on the sky, source_id, list of source_id, limiting magnitude). The output is a plain text that can be used thereafter.

After launching *ExtracteurGRAPPAE3.exe* utility without parameter or with -h or –help, we obtain some help on the syntaxe :

```
ExtrateurGRAPPAE3 V1.0 par Marc SERRAU 2022.

Usage: ExtracteurGRAPPAE3.exe -h|[-b "Ramin,RAmax,DECmin,DECmax"]|[-s source_id]|[-l fichier][-m
Glimite] [-a][-p][-v]|[-f][-g chemin]

Usage: ExtracteurGRAPPAE3.exe --help|[--box="Ramin,RAmax,DECmin,DECmax"]|[--source_ID=identificateur]|
[--liste=fichier][--magGlimite=Glimite][--astro][--photo][--mvt]|[--full][--grappa="chemin"]

b ou box        : extrait toutes les source_id dans la boite de coordonnees
                  indiquee en degres decimaux. 0 < RA < 360 et -90 < DEC < +90
s ou source     : extrait les donnees pour un source_id particulier.
l ou liste      : extrait les donnees pour la liste de source_id donnee dans le fichier indique.
m ou magGlimite : filtre la réponse avec une magnitude G limite
a ou astro      : donne les donnees astrometriques connues.
p ou photo      : donne les donnees photometriques connues.
v ou mvt        : donne les mvt propres, parallaxe et RUWE connus.
f ou full       : donne toutes les donnees connues.
g ou grappa     : indique l'adresse de la base GRAPPAE3 si different du repertoire courant.
d ou debog      : mode debogueur.

Exemple : ExtracteurGRAPPAE3.exe -b "12.5,12.6,-0.1,0.1" -f --magGlimite=18
Exemple : ExtracteurGRAPPAE3.exe --box="12.5,12.6,-0.1,0.1" --full
Exemple : ExtracteurGRAPPAE3.exe -s "12.5,12.6,-0.1,0.1" -f
Exemple : ExtracteurGRAPPAE3.exe -a -p  --source_ID=4508084161667875840

Licence : GNU General Public License.
          This license is available on the World Wide Web at http://www.gnu.org/copyleft/gpl.html
```

Exemple of output with the following command :

```
>ExtracteurGRAPPAE3.exe --box="0,360,89.99,90" --full
Source_ID,RA,eRA,DEC,eDEC,G,eG,G_BP,eG_BP,G_RP,eG_RP,Parallax,eParallax,RUWE,PM_RA,ePM_RA,PM_DEC,ePM_DEC
1729382222550672128,241.926341524444470,0.160,89.990051966666670,0.170,19.1207,0.0022,20.1831,0.0910,18.1041,0.0172,0.
7375,0.1900,0.90,0.6690,0.2070,-10.2050,0.2020
```

This exemple give the most northern star in the sky known by GAIA. The header gives a short description of the contents :

Source_ID : source indentifier
RA : Right Ascension in decimal degree, eRA : incertitude on RA in mas
DEC : Declination in decimal degree , eDEC : incertitude on DEC in mas
G, G_BP et G_RP : magnitudes in the GAIA system and eG, eG_BP and eG_RP incertitude on magnitudes (-1 if undefined values)
Parallax and eParallax in mas (null if unknown)
RUWE : astrometric quality indicator
PM_RA and PM_DEC : proper motions on RA and DEC in mas/yr (ePM_RA and ePM_DEC : incertitudes in mas/yr). (Null if unknown)

This tool was compiled with the Lazarus/FPC and sources are provided under GPL license.

Conclusion

Although great attention was given to the realization of this work and that many checks have been carried out, the author declines any responsibility for the possible consequences of the use of these data.

These data and information are free of use. Any suggestion is welcome and can be directed to the author.

Thank you for using GRAPPA3E, to consume without moderation!

Done at Cesson (France, 77) on February 19, 2021.

Marc SERRAU – *marc.serrau2@free.fr*