



## Elaboración de una nueva versión digital del catálogo estelar "Córdoba Durchmusterung".



Daniel Esteban Severin<sup>1</sup>  
Diego Javier Sevilla<sup>2</sup>

Recibido: Diciembre 2008 – Aceptado: Abril 2009

<sup>1</sup>Jefe de Trabajos Prácticos (dedicación exclusiva) en Depto. de Matemática de la EFB | FCEIA | Universidad Nacional de Rosario. Investigador Asistente de CONICET. <http://www.fceia.unr.edu.ar/~daniel>,

<sup>2</sup>Jefe de Trabajos Prácticos (dedicación exclusiva) en Depto. de Física y Química de la EFB | FCEIA | Universidad Nacional de Rosario

E-mail: [daniel@fceia.unr.edu.ar](mailto:daniel@fceia.unr.edu.ar)

(Flamsteed), BD -16° 1591 (Bonner Durchmusterung), GC 8833 (Boss General Cat.), HD 48915 (Henry Drapper), HIP 32349 (Hipparcos), SAO 151881 (Smithsonian Astrophysical Obs.) y WDS J06451-1643 (Washington Double Star), entre otras. Actualmente los catálogos son generados automáticamente a partir de algoritmos que reducen las mediciones y publicados en versión digital a través de Internet. Por su parte los catálogos más antiguos, originalmente publicados en forma impresa, han sido transcritos/digitalizados manualmente. Uno de ellos, el que abordaremos en este trabajo, es el *Córdoba Durchmusterung* (CD), con más de 600000 estrellas entre la declinación -22° y el polo sur, realizado en el Observatorio Nacional Argentino a fines del siglo XIX y comienzos del XX <sup>[1]</sup>. A pesar de la antigüedad de este catálogo, el mismo sigue teniendo vigencia. En particular, existen muchas estrellas cuya designación principal en *SIMBAD* pertenece a este catálogo, por ejemplo: CD -23° 8634 o CD -23° 8645.

En la traducción de catálogos antiguos a formato digital resultan frecuentes los errores de *typo*. En el caso de CD, la transcripción fue llevada a cabo por un consorcio de varias instituciones lideradas por la *National Space Science Data Center* de la *NASA*, y posteriormente almacenada en *VizieR* como el catálogo I/114. Un ejemplo de error de *typo* en esta versión digital del catálogo es el que se manifiesta para la estrella CD -26° 16170, la cual figura con magnitud 6.5, pero en la versión impresa original la magnitud es de 9.5 <sup>[2]</sup>. Otra desventaja del catálogo I/114 es que no fueron transcritas anotaciones incluidas en la versión impresa, a saber: si la estrella a) se observa como doble o múltiple, o b) si muestra coloración. Además, el catálogo digital no brinda referencias cruzadas, ya que en el momento de su confección no existían otros catálogos lo suficientemente completos como para poder identificar cada estrella de CD. Algunas de las estrellas del CD han sido identificadas, como referencias cruzadas de otros catálogos como SAO o PPM.

Un ejemplo de una estrella no identificada anteriormente es CD -22° 2. Si se precesa y corrige por movimiento propio la posición de la estrella TYC 5844-275-1 a B1875.0 (la época

de CD), y se calcula la separación angular entre las coordenadas de ésta y la de CD -22° 2, esta distancia resulta inferior a 4 segundos de arco. Dado que este valor cae dentro de la precisión de CD, y que la siguiente estrella más cercana a CD -22° 2 es TYC 5844-246-1 con una separación angular de 593 segundos de arco, la identificación entre ambos registros resulta clara. En la actualidad sólo algunas de las estrellas del CD, aproximadamente la cuarta parte, han sido identificadas.

Contar con una identificación cruzada entre el catálogo CD y otros catálogos es importante por varios motivos. En principio, porque ayuda a detectar errores en la confección digital. Por ejemplo, el error antes mencionado de la estrella CD -26° 16170 sólo fue detectado al ser identificada con PPM 273886, ya que en el catálogo PPM dicha estrella figura con magnitud V de 10, con discrepancia notable respecto a la versión digital del CD, pero cercano al valor dado en la versión impresa. La identificación cruzada, además, posibilita comparar observaciones obtenidas en diferentes tiempos, permitiendo detectar cambios en parámetros como la luminosidad en estrellas. Por otro lado, objetos que no puedan ser identificados con contrapartes en catálogos más modernos también resultan de interés, ya que podrían tratarse de objetos como asteroides, novas, supernovas, etc. que resultaron visibles entonces en aquel sector de la bóveda celeste. Finalmente, el catálogo CD fue extensamente utilizado a lo largo del Siglo XX y existen publicaciones que identifican a las estrellas con su designación. Un ejemplo es el uso intensivo de CD y los otros *durchmusterung* en la creación de cartas estelares para poder realizar estimas visuales de estrellas variables <sup>[3]</sup>.

Existen antecedentes de digitalización de catálogos originalmente impresos e identificación cruzada con catálogos modernos <sup>[4]</sup>. En particular la *Uranometría Argentina*, también confeccionada por el Observatorio Nacional Argentino, fue transcrita de manera independiente por dos grupos diferentes. Por un lado, la versión de Paolantonio y Minniti realizada en el marco de un proyecto de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba <sup>[5]</sup>, y por otro, la versión de Frederick Pilcher, que se puede acceder como catálogo V/135A en *VizieR*. En esta última además se brindan referencias cruzadas a los catálogos HD y SAO.

Con la ayuda de un grupo de astrónomos aficionados, nos proponemos elaborar una nueva versión del catálogo Córdoba Durchmusterung que corrija errores de typo presentes en el catálogo digital I/114, e incorpore la información faltante que figura en el catálogo impreso (estrellas de color y dobles) así como una identificación cruzada con otros catálogos más modernos. Debido a la magnitud del trabajo presente, y con la finalidad de evaluar la factibilidad del mismo, en este artículo sólo abordamos una franja del cielo correspondiente a las estrellas de CD con declinación -23° (B1875.0). El trabajo que aquí presentamos fue realizado en 3 etapas. En la primera, se transcribió la lista de las estrellas de color y dobles correspondientes a la zona del cielo mencionada. En la segunda, se realizó la comparación entre los registros de CD y otros catálogos para los cuales existe una identificación cruzada entre CD y éstos. En la tercera, se abordó el problema de la identificación cruzada de los registros de CD con los de un catálogo moderno, mediante la implementación de un algoritmo que resuelve dicha identificación.

El presente artículo se divide de la siguiente manera. En la Sección 2, se detallan diferentes aspectos del catálogo CD así como la conversión entre su escala de magnitud y las utilizadas en la actualidad. En las Secciones 3, 4 y 5, se desarrollan las etapas enunciadas. En la Sección 6 se explica el formato del catálogo digital confeccionado. Finalmente, en la Sección 7, se dan conclusiones y trabajos futuros.

## El catálogo Córdoba Durchmusterung

Los llamados *durchmusterung* consisten en 3 catálogos estelares que barren completamente la bóveda celeste hasta la magnitud 10 aproximadamente <sup>[1]</sup>. El Bonner Durchmusterung (BD) abarca estrellas desde el polo norte hasta la declinación  $-1^\circ$ . El segundo, Südlicher Durchmusterung (SD), abarca las declinaciones  $-2^\circ$  a  $-23^\circ$ . Por último, el CD abarca desde la declinación  $-22^\circ$  hasta el polo sur y cuenta con 613778 estrellas y 175 objetos no estelares. La superposición de dos grados entre el SD y el CD tiene por objeto empalmar adecuadamente ambos catálogos.

Para identificar una estrella de CD, se utiliza la notación CD  $-xx^\circ$  *yyyyy* donde *xx* corresponde a los grados de declinación de la estrella e *yyyyy* es una enumeración en orden ascendente según la ascensión recta entre las estrellas con mismo grado de declinación. El catálogo da posiciones y brillos hasta la magnitud 10 (según la escala utilizada en aquella época). Las posiciones se reportan para la época B1875.0, siendo la ascensión recta dada en horas, minutos y segundos con una precisión de  $0^s.1$ , y la declinación en grados y minutos de arco, con precisión de  $0'.1$ . La precisión para la magnitud visual es 0.1. La versión impresa del catálogo cuenta además con información de las estrellas que presentan color (amarillo – rojo) y que se observan como dobles o múltiples, es decir, con una separación angular pequeña para el aumento utilizado (15x). Esta información es dada en notas a pies de páginas. La versión impresa además cuenta con referencias cruzadas a catálogos de esa época <sup>[2]</sup>.

Varias correcciones al catálogo impreso han sido publicadas. Nosotros utilizamos la *corrigenda* del mismo volumen <sup>[2]</sup>, las reportadas por Thome <sup>[6]</sup> y por la *NSSDC* <sup>[7]</sup>.

Para la franja del cielo correspondiente a la declinación  $-23^\circ$ , la cantidad de registros de CD es de 18133 estrellas y 5 objetos no estelares (los cuales se distinguen con la leyenda *neb.* en el campo donde debería reportarse la magnitud estelar):  $-23^\circ$  542,  $-23^\circ$  1186,  $-23^\circ$  13481,  $-23^\circ$  14459 y  $-23^\circ$  17047. Estos registros no deben ser considerados para la identificación cruzada con otros catálogos estelares.

Para estrellas entre  $-22^\circ$  a  $-32^\circ$ , el error *probable* es  $\pm 0^s.42$  en ascensión recta y  $\pm 0'.23$  en declinación <sup>[2]</sup>. Esto equivale a  $\sigma_{\alpha^*} = 9.3$  arcsec y  $\sigma_{\delta} = 20.5$  arcsec en términos de desviación estándar, siendo  $\alpha^* = \alpha \cdot \cos(\delta)$ . En observaciones posteriores, para otros sectores del cielo, el error pudo ser reducido.

Para estimar el error en el brillo, es necesario convertir entre la antigua escala visual empleada en CD y la escala de magnitudes Johnson V. En este trabajo ajustamos por mínimos cuadrados un polinomio cuadrático a puntos que representan las magnitudes de CD respecto a la que figura en otros catálogos, para casos en los que ya se tiene identificación cruzada. Para ello se utilizaron los catálogos *Hipparcos* y *PPM*, que son descritos con más detalle en la Sección 4. Para *Hipparcos*, se consideraron por separado tanto las estrellas no variables (VarFlag=1), como todos los registros, pero se descartó HIP 67620 dado que la diferencia de magnitud respecto a CD  $-23^\circ$  11328 es 3.5

2240

unidades, demasiado alta para tratarse de un error estadístico. Los números de estrellas utilizadas en cada ajuste fueron 763 (HIP sin variables), 809 (HIP) y 3927 (PPM). Por un lado, se observó que la diferencia entre los ajustes para HIP son despreciables, es decir, que el descartar estrellas variables tiene poco efecto en el resultado. Por otro lado, se observó que los ajustes con HIP no resultan adecuados para estrellas de alta magnitud, dado que dicho catálogo no cuenta con suficientes estrellas débiles. Concluimos que el ajuste más adecuado se logra usando PPM. Si  $f(m) = Am^2 + Bm + C$  es la función que, dada la magnitud  $m$  de PPM en Johnson V, devuelve en  $f(m)$  la magnitud en la escala de CD, los coeficientes obtenidos en la regresión son:  $A = -0.01335368$ ,  $B = 1.076636$ ,  $C = 0.2249828$ , siendo el error cuadrático medio  $\sigma_m = 0.2759$ .

### Incorporación de estrellas de color y dobles

En esta etapa, se leyeron los pies de las pág. 56 a 117 del catálogo impreso <sup>[2]</sup>, justamente las que corresponden a la declinación  $-23^\circ$ , y se anotaron las estrellas de color y dobles en archivos de texto. Posteriormente, con un algoritmo sencillo se incorporaron estos datos al catálogo digital.

La cantidad de estrellas de color asciende a 90, mientras que la de dobles asciende a 194. Además hallamos un error evidente en el catálogo impreso: en la pág. 65, CD  $-23^\circ$  3539 está marcada como de color. Sin embargo, esto es imposible ya que aquella estrella no se encuentra en dicha página. Por otro lado, la estrella CD  $-23^\circ$  2539 (que sí se encuentra en esa pág.) es brillante y de color naranja, características suficientes como para ser marcada como de color. Esta estrella se identifica en *SIMBAD* como HD 34087, con magnitud V de 7.4 y cuyo índice de color es  $B - V = 1.03$ .

Durante nuestra lectura también observamos que, en la pag. 105, la estrella CD  $-23^\circ$  14537 figura, al pie de página, como *cúmulo*. Efectivamente, su posición coincide con el cúmulo globular Messier 22 (NGC 6656) aunque desconocemos porqué fue reportada como estrella en vez de como un objeto no estelar.

Además de la tarea realizada, también consideramos la posibilidad de transcribir las referencias cruzadas presentes en el catálogo CD. Estas referencias tampoco existen en la versión digital I/114. En la versión impresa, aparecen como un campo adicional que puede ser una combinación de 2 letras que hacen referencia a 5 catálogos antiguos donde esta estrella tiene designación. Sin embargo, transcribir todas estas entradas insumía mucho tiempo y sólo pasamos 266 entradas correspondientes a 8 páginas del catálogo impreso.

## Comparación con otros catálogos y corrección de errores de typo

Existen catálogos cuyas estrellas tienen referencias cruzadas a CD. Entonces, es posible aprovechar esta información para comparar sus parámetros con los registrados en CD y así poder hallar errores en éste último. Nuestra hipótesis se basa en que, si asumimos que cada parámetro de CD se distribuye normalmente con una desviación estándar de  $\sigma$ , hay una gran probabilidad de que un error de typo en la versión digital se manifieste al compararse dicho parámetro con otro de referencia que se supone exacto (aquel que se obtiene de otro catálogo), y que su diferencia sea superior a los  $3\sigma$  ya que sólo un 0.3% de estas diferencias serían debido a la naturaleza del parámetro.

La modalidad de trabajo es la siguiente: se elige un catálogo, se barren todos los registros que estén identificados con una estrella CD  $-23^\circ$ , se precesa y corrige por movimiento propio la posición de la estrella a B1875.0 y se ajusta su magnitud a la escala de magnitudes de CD, luego un algoritmo enumera los casos en donde la diferencia de ascensión recta, declinación o magnitud con CD supera un umbral previamente determinado, y finalmente se comparan manualmente la versión digital de CD con la impresa sobre dichos registros. Denotamos con  $u_{\alpha^*}$ ,  $u_\delta$  y  $u_m$ , a los respectivos umbrales tomados para cada catálogo. Dichos umbrales fueron escogidos de manera de balancear la cantidad de casos en los distintos parámetros. A continuación describimos aquellos catálogos empleados en la comparación:

- *Hipparcos* (HIP). Este catálogo astrométrico, que se puede obtener en *VizieR* como I/239, es muy preciso y casi la totalidad de sus registros contienen una referencia cruzada a CD. La desventaja radica en que tiene pocos registros comparados con CD: 811 en total. Luego de descartar HIP 62292 (que se identifica con CD  $-23^\circ$  10709) debido a que faltan parámetros para calcular la posición de la estrella para 1875, y descartar aquellas que superan los umbrales  $u_{\alpha^*} = 23.4$  arcsec,  $u_\delta = 51.9$  arcsec,  $u_m = 0.84$ , nos quedan 50 registros. Los verificamos manualmente y hallamos un error en I/114 para CD  $-23^\circ$  9296.
- *Positions and Proper Motions* (PPM). Este catálogo, cuyas identificaciones en *VizieR* son I/193 (South) y I/208 (Supplement), es menos preciso que HIP pero es el que más cantidad de registros con referencias cruzadas a CD contiene. Si nos restringimos a aquellas indexadas con estrellas de CD  $-23^\circ$  obtenemos 4583 registros (que equivale al 25% de las estrellas de CD comprendidas en esa faja). De ellas, 656 no poseen magnitud visual por lo que sólo son comparadas por posición. Los umbrales escogidos son  $u_{\alpha^*} = 34$  arcsec,  $u_\delta = 62$  arcsec,  $u_m = 0.75$ . Nos encontramos con 84 registros a verificar, de los cuales se hallan errores en las estrellas  $-23^\circ$  8777,  $-23^\circ$  9296 y  $-23^\circ$  12414. También se halla diferencia en  $-23^\circ$  5830, pero esta estrella figura corregida en la corrigenda de la *NSSDC* <sup>[7]</sup>, por lo que el registro en el catálogo digital es el correcto.

Observamos también que los campos de  $-23^\circ$  12413 y  $-23^\circ$  12414 están invertidos en el catálogo digital. Quizás puede deberse a un intento pasado de corregir el orden en ascensión recta, pero este orden no debería corregirse ya que la identificación cruzada sería incorrecta.

- *Córdoba A* (AGK). Este catálogo fue realizado en el Observatorio Nacional Argentino a principios del siglo XX como parte de un proyecto internacional denominado *Astronomische Gesellschaft Katalog* y contiene 3389 registros con referencias cruzadas a estrellas CD  $-23^\circ$ . La ventaja de utilizar este catálogo para comparar es que las coordenadas fueron medidas en casi la misma época que CD por lo que no se

2242

propaga tanto error debido al movimiento propio de las estrellas, y además las magnitudes están en una escala similar a la de CD, por lo que no es necesario realizar un cambio de escala para compararlas. La versión digital de este catálogo se puede acceder ingresando a <http://dc.zah.uni-heidelberg.de/ariqfh/katkat/q/form> y colocando 986 en el campo *Teleki*.

Un inconveniente al utilizar este catálogo fue que, en cada registro, figura el número de identificación de CD sin la declinación, por lo que ésta última debe inferirse eligiendo aquella estrella de misma identificación de CD más cercana en posición a la de AGK. Esto se logra con un algoritmo sencillo. Por ejemplo, AGK 11856 se identifica con una estrella CD de numeración 13169. Como la distancia entre AGK 11856 y CD  $-23^{\circ}$  13169 (21 arcsec) es inferior a la existente entre AGK 11856 y CD  $-24^{\circ}$  13169 (461 arcsec), podemos garantizar que AGK 11856 se identifica efectivamente con CD  $-23^{\circ}$  13169. Elegimos los umbrales  $u_{\alpha^*} = 34$  arcsec,  $u_{\delta} = 62$  arcsec,  $u_m = 0.96$ , y nos encontramos con 107 registros a verificar, hallando errores en:  $-23^{\circ}$  8777,  $-23^{\circ}$  9296,  $-23^{\circ}$  9705 y  $-23^{\circ}$  12414.

- *Cape Photographic Durchmusterung* (CPD). Este catálogo, en *VizieR I/108*, es de la misma época que CD y también puede utilizarse para comprobar errores en sus posiciones sin necesidad de corregir movimientos propios. El catálogo en sí no contiene referencias a las estrellas de CD pero se conoce una identificación cruzada entre ambos catálogos realizada por Rappaport y Warren <sup>[8]</sup>, que se puede obtener de: <ftp://dbc.nao.ac.jp/DBC/NASAADC/catalogs/4/4019>

Tenemos 8091 registros identificados con CD  $-23^{\circ}$ . Debido a que la magnitud de CPD es fotográfica, no tenemos en cuenta este parámetro. Elegimos los umbrales  $u_{\alpha^*} = 64$  arcsec y  $u_{\delta} = 94$  arcsec, y nos encontramos con 392 registros a verificar, hallando errores en:  $-23^{\circ}$  5485,  $-23^{\circ}$  6942,  $-23^{\circ}$  7042,  $-23^{\circ}$  8777,  $-23^{\circ}$  11817 y  $-23^{\circ}$  12414.

## Identificación cruzada entre CD y otros catálogos

Si bien existen identificaciones cruzadas entre CD y otros catálogos como HIP, SAO o PPM, al tener CD más cantidad de estrellas que estos otros, inevitablemente existen estrellas de CD que no están actualmente identificadas. Como dijimos en la introducción, CD -22° 2 era un ejemplo. No obstante, en las últimas décadas se han confeccionado nuevos catálogos con una enorme cantidad de estrellas y otros objetos estelares (por ejemplo, USNO-B1.0 con 1042618261 registros), por lo que actualmente es posible realizar una identificación cruzada que le haga corresponder una referencia a cada estrella de CD.

El catálogo que utilizamos para realizar la identificación cruzada con CD es el *Positions and Proper Motions eXtended* (PPMX, en *VizieR I/312*), ya que tiene la suficiente densidad como para poder identificar cada estrella de CD, y a la vez tiene la suficiente precisión como para poder obtener posiciones confiables de sus estrellas para la época B1875.0. Cada registro de este catálogo contiene los siguientes parámetros: *posiciones* (RAJ2000, DEJ2000) y su *error* (e\_RAJ2000, e\_DEJ2000) para una *época* determinada (epRA, epDE), *movimiento propio* (pmRA, pmDE) y su *error* (e\_pmRA y e\_pmDE), *magnitud* en Johnson V (Vmag) y su *error* (e\_Vmag), entre otros <sup>[9]</sup>. Si nos limitamos a estrellas de magnitud a lo sumo 13.5 y con un rango de declinaciones que contenga todas las estrellas de CD -23° y 10 arcos de minutos más en ambos lados, obtenemos 119230 estrellas de PPMX.

La técnica habitual para realizar identificaciones cruzadas es la de considerar una métrica para comparar los parámetros de ambos catálogos, por ejemplo la separación angular entre las coordenadas de ambos registros (reducidas para una misma época). Posteriormente, se van asignando las estrellas que minimizan dicha separación angular con un algoritmo de *matching*. Este método es ampliamente utilizado hoy en día, en los llamados *Observatorios Virtuales* los cuales ofrecen, entre otras herramientas online, la de realizar identificaciones cruzadas (como, por ejemplo *X-Match* del Centro de Datos de Estrasburgo). Sin embargo, las métricas utilizadas en este tipo de herramientas sólo tienen en cuenta la posición, aunque en la literatura existan métricas más sofisticadas que contemplan la magnitud y otros parámetros físicos <sup>[10], [11]</sup>. Por otro lado, estas herramientas no devuelven una identificación uno-a-uno.

Para poder cumplir nuestro objetivo, fue necesario diseñar e implementar una herramienta informática que resuelva los inconvenientes mencionados, a saber:

- Que genere una identificación cruzada para la cual, por cada estrella de CD, se corresponda una única estrella de PPMX (o, en su defecto, ninguna) y, por cada estrella marcada como doble en CD, se correspondan dos estrellas PPMX (o, en su defecto, una o ninguna).
- Que, en el caso de ambigüedades, el algoritmo elija la asignación más "probable".
- Que, para el caso de estrellas no-dobles, la métrica se aproxime a la probabilidad de que a una estrella de CD le corresponda una de PPMX, teniendo en cuenta posición, brillo y desviaciones estándares de ambos.
- Que, para el caso de estrellas dobles, la métrica también tenga en cuenta la probabilidad de que un par de estrellas de PPMX sea considerado un candidato a ser correspondido con una estrella doble de CD. Es decir, la métrica debe tener en

2244



cuenta también la separación angular y la diferencia en magnitud de las componentes que se consideran en el par.

Para el último ítem, fue necesario determinar tanto la separación angular promedio como la diferencia en magnitud entre las componentes del par que se corresponde con una estrella doble en CD. Realizamos una inspección e identificación manual de las siguientes 7 estrellas dobles:  $-25^\circ 8168$ ,  $-30^\circ 8494$ ,  $-23^\circ 9730$ ,  $-25^\circ 10131$ ,  $-25^\circ 10846$ ,  $-25^\circ 11101$ ,  $-24^\circ 11610$ , y observamos que la separación angular entre las componentes suele variar entre los 15 arcsec y 1 arcmin, con un promedio de 34.9 arcsec y desviación estándar de 13.65 arcsec. Respecto a la diferencia entre las magnitudes, es claro que si ésta es pronunciada, la estrella más débil es difícil de observar visualmente. Si asumimos que esta diferencia se comporta según una distribución normal centrado en cero, estimamos su desviación estándar en 0.915.

Por otra parte, el catálogo CD no posee información del error cometido para cada estrella ni en posición ni en brillo. En este sentido, decidimos asignar un error constante para todas las estrellas de CD. En lo que respecta a magnitud, asumimos como error el brindado por el ajuste cuadrático realizado en la Sección 2, es decir  $\sigma_m = 0.2759$ . Respecto a la posición, podíamos optar por usar los valores sugeridos por Thome (los cuales fueron calculados en base a una comparación con el catálogo SD de Schönfeld que también posee poca precisión en las posiciones) o podíamos comparar las posiciones con un catálogo moderno. Optamos por esto último, utilizando el catálogo PPM. Descartamos 3 estrellas marcadas como "problemáticas astrométricamente" en PPM, más otras 253 estrellas para las cuales la diferencia en ascensión recta o declinación superan los 2 arcos de minutos respecto a CD. De los 4327 registros restantes, obtuvimos  $\sigma_{\alpha^*} = 10.15$  arcsec y  $\sigma_\delta = 22.74$  arcsec. Obsérvese que nuestros valores son similares a los obtenidos por Thome. El catálogo PPM fue elegido justamente por la cantidad de estrellas que tienen referencias cruzadas a CD, y también por su precisión. Por ejemplo, la estrella con mayor incertidumbre, es decir aquella que maximiza  $\sigma_{\alpha^*}^2 + \sigma_\delta^2$ , donde  $\sigma_{\alpha^*}$  y  $\sigma_\delta$  son las desviaciones estándares propagadas a la época B1875.0, es la PPM 266698 con  $\sigma_{\alpha^*} = 0.57$  arcsec y  $\sigma_\delta = 0.61$  arcsec (esto se debe, principalmente, a su alta incertidumbre en el movimiento propio, la que se propaga a razón de 6.3 y 6.2 mas/año, a lo largo de 98 años, pues  $E_p = 1973$ ). Nótese que estas cifras son relativamente pequeñas respecto al error que se manifiesta el catálogo CD.

En lo que sigue se describe brevemente el algoritmo utilizado, el cual fue implementado en C++. Los detalles técnicos serán objeto de una publicación futura. Cabe mencionar que el algoritmo no se limita a la identificación entre los catálogos CD y PPMX, si no que puede ser aplicado para generar referencias cruzadas entre dos catálogos cualesquiera de diferentes épocas que contengan información sobre posiciones y brillos, sus respectivos errores y, eventualmente, que en uno de ellos puedan existir estrellas marcadas como dobles.

En primer lugar, se precisan y corrigen por movimientos propios las posiciones de las estrellas de PPMX hasta la época besseliana de 1875 utilizando la rutina *wcscnp* de la biblioteca WCSTools<sup>[12]</sup>. También se propaga el error<sup>[13]</sup>, es decir se calculan las desviaciones estándar en ascensión recta y declinación a partir de los parámetros presentes en el catálogo PPMX.



Luego se construye lo que, en Matemática Discreta, se conoce como un *grafo bipartito*. Cada estrella de CD y PPMX se representa con un vértice del grafo, y por cada par de estrellas  $(x,y)$  donde  $x$  es una estrella de CD e  $y$  una estrella de PPMX, se considera una arista entre sus vértices sólo si la separación angular entre  $x$  e  $y$  es inferior a una tolerancia (en nuestro caso, 6 arcos de minuto). Para evitar errores de identificación, también se agregaron aquellas estrellas de CD cuya declinación es  $-22^\circ$  y  $-24^\circ$ . Con este criterio, el grafo contiene 171904 vértices y 320542 aristas.

Con un algoritmo similar al de *Kruskal*, es fácil descomponer este grafo en sus componentes conexas. En nuestro caso, la cantidad de componentes ascendió a 31092. De éstas, 22862 son sólo vértices aislados, 3960 consisten en un vértice de CD y uno o más de PPMX cuya identificación es trivial (a dicha estrella de CD se le asigna directamente la de mayor probabilidad de PPMX) y 4270 consisten en 2 o más vértices de CD y varios de PPMX. La componente conexa más grande se corresponde con una zona densamente poblada de estrellas de la vía láctea, con 10249 estrellas de CD (de las cuales 71 son dobles), 33071 estrellas de PPMX y 133162 aristas.

La identificación cruzada para una componente conexa dada se realiza de la siguiente manera. En caso de que existan estrellas dobles presentes en la componente conexa, se calculan los pares de estrellas de PPMX  $(y_1, y_2)$  candidatos a ser dobles, donde  $y_1$  es considerada la principal e  $y_2$  la secundaria. El criterio para elegirlos se basa en que la separación angular entre  $y_1$  e  $y_2$  sea inferior a 80 arcsec, y que el brillo de  $y_1$  no pueda ser más débil que el de  $y_2$  en 1 magnitud. Con este criterio, en la componente conexa más grande, se obtuvieron 1609 pares.

Luego, por cada arista  $(x,y)$  con  $x$  estrella no-doble se calcula la probabilidad de que  $x$  sea correspondida con  $y$ , y por cada arista  $(x,y_1)$  con  $x$  estrella doble, se calcula la probabilidad de que  $x$  sea correspondida con un par  $(y_1,y_2)$  para todo  $y_2$ . También se calculan las probabilidades de que una estrella  $x$  no sea correspondida con ninguna de PPMX, o que una estrella doble  $x$  sólo sea correspondida con una única estrella de PPMX. La búsqueda de la asignación que maximiza la probabilidad se efectúa mediante la optimización de un Modelo de Programación Lineal Entera que es resuelto por el solver GuRoBi 6.0.2 (<http://www.gurobi.com>). A fin de que la función objetivo del modelo sea lineal, se trabaja con los logaritmos de las probabilidades calculadas. El modelo correspondiente a la resolución de la componente conexa más grande contiene 144167 variables binarias y 43320 restricciones lineales, pero es resuelto en menos de 4 segundos de tiempo.

El tiempo transcurrido durante todo el proceso fue de 1114 segundos, sobre 1 hilo de ejecución, en una computadora equipada con un Intel i5 2.67Ghz.

## Presentación y formato del nuevo catálogo

La versión digital del nuevo catálogo fue generada en forma de dos archivos de texto. El primer archivo, *cd.txt*, contiene la nueva versión del catálogo CD, incluyendo todas las estrellas y objetos no estelares de la versión impresa para la declinación  $-23^\circ$  en un formato similar al utilizado en el catálogo I/114. A continuación presentamos el formato, en el estándar utilizado para submitir catálogos en *VizieR*:

Byte-by-byte Description of file: cd.txt

Bytes	Format	Units	Label	Explanations
1-	2 A2	---	---	[CD] The catalog prefix
3-	5 I3	deg	zone	[-23] The declination zone
6-	10 I5	---	num	The number of the star within the zone
	11 A1	---	suppl	[ ] Corrigenda flag: not needed here
12-	15 F4.1	mag	mag	*Estimated visual magnitude
16-	17 I2	h	RAh	Hours of right ascension, 1875
18-	19 I2	min	RAm	Minutes of right ascension, 1875
20-	23 F4.1	s	RA s	Seconds of right ascension, 1875
	24 A1	---	DE-	[-] Sign of declination
25-	26 I2	deg	DEd	Degree of declination, 1875
27-	30 F4.1	arcmin	DEm	Minutes of declination, 1875
	31 A1	---	dpl	*[D ] Double flag
	32 A1	---	color	*[C ] Color flag

**Note on mag:**

20.0 = neb; 30.0 = var

**Note on dpl:**

D = the star is double visually

**Note on color:**

C = the star is yellow or red

El segundo archivo, *cross.txt*, tiene almacenada las identificaciones cruzadas generadas entre CD y PPMX en el siguiente formato:

Byte-by-byte Description of file: cross.txt

Bytes	Format	Units	Label	Explanations
1-	2 A2	---	---	[CD] The catalog prefix
3-	5 I3	deg	zone	[-23] The declination zone
6-	10 I5	---	num	The number of the CD star within the zone
	11 A1	---	suppl	[ ] Corrigenda flag: not needed here
	12 A1	---	type	*[SUAB] Type of identification
13-	18 F6.2	arcsec	dist	Angular separation between position of CD star and position of PPMX star
19-	22 A4	---	---	[PPMX] The catalog prefix
23-	37 A15	---	PPMX	Name of PPMX star (HHMMSS.S+DDMMSS)
	38 A1	---	m_PPMX	[pf] if name of PPMX star is identical

**Note on type:**

S = CD star is single

U = CD star is double, but it was identified with a unique PPMX star

A = PPMX star corresponds to the component A of CD star

B = PPMX star corresponds to the component B of CD star

Además de los archivos mencionados, hemos colocado en el mismo sitio los siguientes archivos adicionales: *color.txt* y *dpl.txt*, que contienen la lista de estrellas de color y dobles realizadas en la primer etapa; *ref.txt*, que contiene una lista parcial de referencias realizadas en la primer etapa; *comp\_hip.txt*, *comp\_ppm.txt*, *comp\_agk.txt* y *comp\_cpd.txt*, que contienen la lista de estrellas de CD generadas automáticamente en la segunda etapa; *changes.txt*, que contiene la lista de cambios realizados al catálogo digital, con el formato utilizado en I/114/cdchg.

## Conclusiones

En este artículo describimos el inicio de un trabajo mayor, que consiste en la elaboración de una nueva versión del catálogo Córdoba Durchmusterung, que sea lo más fiel posible al catálogo impreso original (es decir, con la menor cantidad de errores de transcripción posibles y con toda la información que provea la versión impresa). Para tener una idea de la magnitud de este trabajo, abordamos sólo una franja del cielo, concluyendo lo siguiente:

Respecto a la primera etapa, fue simplemente una tarea de transcripción de los datos faltantes, que se realizó en 2-3 horas de trabajo, lo que hace factible abordar el resto del catálogo CD. Además, la incorporación del conjunto de estrellas dobles al catálogo hizo que fuera posible realizar la identificación cruzada de las estrellas dobles de manera adecuada. En lo que respecta a transcribir las referencias presentes en CD, observamos que insume mucho tiempo y la desestimamos.

En la segunda etapa, observamos que el uso de varios catálogos de comparación hizo que se encuentren una y otra vez errores en las mismas estrellas, contribuyendo cada catálogo muy pocos errores adicionales. Concluimos que, para abordar el resto del catálogo CD, solamente la comparación con el catálogo PPM sería suficiente.

La tercera etapa es, quizás, la más interesante ya que abre una nueva línea de investigación que es la de identificar estrellas dobles. En este sentido, quedan varios trabajos a futuro que se pueden realizar, a saber:

- Explorar distintas métricas de la literatura o proponer métricas propias, y probarlas sobre identificaciones ya realizadas para comprobar su porcentaje de éxito. También, ver la posibilidad de "aprovechar" el conjunto de estrellas de color para mejorar la calidad de la métrica.
- En el caso específico de CD, utilizar una muestra mayor de estrellas dobles con el fin de inferir los parámetros "separación angular" y "diferencia en magnitud". En un caso general, proponer nuevos parámetros para llevar a cabo identificaciones cruzadas de estrellas con 3 o más componentes.
- Adaptar la actual herramienta para hallar la identificación cruzada en catálogos arbitrarios.
- Es fácil ver que el problema de la identificación cruzada es *polinomial* cuando no hay que identificar estrellas dobles. En este caso, una instancia de la misma puede resolverse eficientemente mediante un algoritmo de matching como, por ejemplo, el Algoritmo Húngaro <sup>[14]</sup>. Sin embargo, en el caso de que uno de los catálogos posea estrellas dobles, queda pendiente estudiar la *complejidad computacional* de este problema y, según resulte ésta, abordarlo luego con las técnicas disponibles para aquella clase de complejidad.

## Agradecimientos

Agradecemos a la Srita. Andrea Jimenez, Sr. Catriel Caruso, Srita. Noelia Acosta Pedemonte, y a nuestros colegas Guillermo Ibañez y Silvia Morales por colaborar en este trabajo. También agradecemos especialmente al Prof. Santiago Paolantonio por contestar nuestras consultas respecto al tema.

## Referencias

- (1) MINNITI, E.; PAOLANTONIO, S. Córdoba Estelar. Desde los sueños a la Astrofísica. Historia del Observatorio Nacional Argentino. *Editorial Universidad Nacional de Córdoba*, 2013.
- (2) THOME J. M. Zonas de Exploración (Córdoba Durchmusterung): declinación  $-22^{\circ}$  a  $-32^{\circ}$ . *Resultados del Observatorio Nacional Argentino*, 1892, Vol. 16.
- (3) GLASBY, J. S. The Variable Star Observer's Handbook. 1st. Edition. *Sidgwick and Jackson*, 1971.
- (4) LEQUEUX J. From Flamsteed to Piazzini and Lalande: new standards in 18th century astrometry. *Astronomy & Astrophysics*, 2014, Vol. 567, A26 pp. 1--9.
- (5) MINNITI, E.; PAOLANTONIO, S. Uranometría Argentina 2001. Historia del Observatorio Nacional Argentino. Edición digital. <http://historiadelastronomia.wordpress.com/documentos/uranometria-argentina/uabicentenario>
- (6) THOME J.M. List of additional errors found in the Cordoba catalogues. *The Astronomical Journal*, 1896, Vol. 16, pp. 30.
- (7) NASA Reference Publication 1299. Cordoba Durchmusterung, volume 1-4. *Technical Report, NASA Goddard Space Flight Center*, 1993.
- (8) RAPPAPORT, B. N.; WARREN, W. H., Jr. Cross-identification of the Cordoba and Cape Photographic Durchmusterung. *Bulletin of the American Astronomical Society*, 1986, Vol. 18, pp. 897.
- (9) ROESER, S.; SCHILBACH, E.; SCHWAN, H.; KHARCHENKO, N. V.; PISKUNOV, A. E.; SCHOLZ R.-D. PPM-Extended (PPMX), a catalogue of positions and proper motions. *Astronomy & Astrophysics*, 2008, Vol. 488, pp. 401--408.
- (10) SUTHERLAND, W.; SAUNDERS, W. On the likelihood ratio for source identification. *Monthly Notices of the Royal Astronomical Society*, 1992, Vol. 259, pp. 413--420.
- (11) BUDAVARI T.; SZALAY A. S. Probabilistic Cross-Identification of Astronomical Sources. *The Astrophysical Journal*, 2008, Vol. 679, pp. 301--309.
- (12) *WCSTools*: Image World Coordinate System Utilities. Harvard-Smithsonian Center for Astrophysics. <http://tdc-www.harvard.edu/wcstools>
- (13) KOVALEVSKY J.; SEIDELMANN P. K. Fundamentals of Astrometry. *Cambridge University Press*, 2004.
- (14) COOK W. J.; CUNNINGHAM, W. H.; PULLEYBACK, W. R.; SCHRIJVER A. Combinatorial Optimization. *Wiley-Interscience*, 1998.