Revista de Educación Matemática
Volumen 37, N° 2 (2022), páginas 3 – 26
Unión Matemática Argentina - Famaf (UNC)

Artículo de Matemática

CRIPTOGRAFÍA, LITERATURA Y EDUCACIÓN INTEGRAL: "VIAJE AL CENTRO DE LA TIERRA" DE JULIO VERNE

Vicente Jara-Vera y Carmen Sánchez-Ávila

Resumen. Ofrecemos en este artículo un ejemplo de las posibilidades de la Criptografía, como rama de la Matemática Aplicada, para abarcar de manera grata y fructífera numerosas cuestiones imprescindibles de la Educación. Y cuando decimos Educación queremos resaltar su semántica más amplia y extensa, tan necesaria y tan añorada muchas veces hoy en día, sin perder la necesaria precisión de la especialización, agrupando en este estudio cuestiones relativas a áreas como la Matemática, la Seguridad de las Comunicaciones, la Literatura, la Lingüística, o las Lenguas Clásicas. Nos hemos apoyado para tal efecto en una famosa obra literaria del genial escritor francés Julio G. Verne, "Viaje al centro de la Tierra". A partir únicamente del enigmático mensaje que aparece al inicio de la novela ofrecemos detalladamente y por medios criptoanalíticos la solución, mostrando paso a paso los diversos acercamientos para resolver de manera general un texto desconocido. Finalmente, este ejemplo nos sirve para detallar algunos elementos fundamentales del Álgebra y de la Teoría de la Comunicación que son propios de la Criptología.

Palabras clave: Criptoanálisis, Criptografía, Educación, Literatura.

Abstract. We offer in this paper an example of the possibilities of Cryptography, as a branch of Applied Mathematics, to cover pleasantly and fruitfully many essential questions of Education. And when we say Education we want to highlight its broader and more extensive semantics, so necessary and so longed for many times today, without losing the necessary precision of specialization, grouping in this study issues related to areas such as Mathematics, Communications Security, Literature, Linguistics, or Classical Languages. For this purpose, we have relied on a famous literary work by the great French writer Jules G. Verne, "Journey to the Center of the Earth". Based solely on the enigmatic message that appears at the beginning of the novel we offer the solution in detail and by cryptanalytic means, showing step by step the various approaches to generally resolve an unknown text. Finally, this example helps us to detail some fundamental elements of Algebra and Communication Theory that are typical of Cryptology.

Keywords: Cryptanalysis, Cryptography, Education, Literature

§1. Introducción

Toda novela es un testimonio cifrado (*Vargas-Llosa*, *1971*).

Sin negar la afirmación mencionada por el premiado con el Nobel de Literatura, vamos a circunscribirnos mucho más, en concreto, a algunas novelas muy significativas donde el cifrado es mucho más patente. Nuestra intención es llamar la atención sobre esas obras, las cuales suelen quedar ubicadas en aquellos momentos de la vida donde los niños y jóvenes atisban la adultez, con la pretensión de rescatarlas también, sin olvidar aquellos, para una edad más madura y reflexiva. La justificación está en que los tonos y coloridos que proyectan ofrecen a este lector más reflexivo aspectos inéditos que con casi total seguridad pasó por alto en aquella época. Además, en las tareas educativas y docentes pueden utilizarse muy acertadamente, sirviendo de textos motivadores y ofreciendo una gran transversalidad, abarcando desde ellas multitud de disciplinas, no solo la Literatura o la Lingüística, o las Lenguas Clásicas, sino también la Matemática, la Criptología y la Teoría de la Seguridad y las Comunicaciones.

Han sido muchos los literatos que se sintieron atraídos por la Criptografía, o al menos, la usaron como elemento de provecho, por aquello del misterio que lleva aparejado, para algunas de sus novelas(Kahn, 1996). Sin ser exhaustivos, podemos indicar algunos de ellos.

Así, el británico William Makepeace Thackeray (1811-1863), quien en su novela "The History of Henry Esmond" (1852) menciona un cifrado de transposición con rejilla de Cardano: una permutación de las letras del texto, desordenándolas según una serie de agujeros dispuestos de manera confusa en una plantilla.

Otro es el irlandés Bram Stoker (1847-1912), quien en "The Mystery of the Sea" (1902) muestra un complicado sistema de cifrado y esteganografía con el sistema de Francis Bacon, sustitución múltiple biliteral, obligándose a ofrecer al lector varios apéndices explicativos.

Mencionemos al inglés Henry Rider Haggard (1856-1925), el cual en "Colonel Quaritch, V.C.: A Tale of Country Life" (1888) ofrecerá un cifrado con letras nulas, escondiendo así el texto de interés.

El escritor estadounidense Robert William Chambers (1865-1933) escribirá "The Tracer of Lost Persons" (1906), donde aparecerá un cifrado de sustitución sencillo de signos a números, que serán luego los ordinales de las letras alfabéticas del mensaje oculto.

No podía faltar en este breve listado la inglesa Agatha Christie (1890-1976), quien en "The Four Suspects" (1932) diferentes nombres de dalias generan un mensaje secreto.

Sin embargo, de entre todos los autores, los que más resaltan son sin duda alguna Poe, Verne y Doyle.

De la pluma del estadounidense Edgar Allan Poe (1809-1849) salió el famoso relato "The Gold-Bug" (1843), donde encontramos un sencillo cifrado de sustitución monográmica monoalfabeto: es decir, donde cada letra del texto en claro es sustituida por otro signo y siempre por el mismo otro signo (Poe, 1843).

De entre los grandes novelistas en lengua francesa hemos de citar a Jules Gabriel Verne (1828-1905), conocido en los países iberoamericanos como Julio Verne, quien se servirá en cuatro de sus novelas de la Criptografía. En la famosa "Voyage au centre de la Terre" (Verne, 1864) encontramos un mensaje cifrado que esconde el lugar por el cual se puede acceder al centro del planeta Tierra. En la segunda novela, "Les enfants du Capitaine Grant" (Verne, 1868), los hijos de un capitán de navío lanzan una petición de socorro al mar en una botella que será posteriormente tragada por un tiburón martillo. La verdad es que aquí no estamos ante un mensaje cifrado, sino simplemente muy deteriorado, con un mismo texto de auxilio en inglés, alemán y francés, que a trozos, de uno y otro idioma, permiten recomponerlo en gran parte. Unos años más tarde escribirá Verne "La Jangada" (Verne, 1881), situada en el Amazonas, donde un texto cifrado de sustitución polialfabética (ahora no siempre la misma letra es sustituida por el mismo símbolo) es pieza fundamental para prevenir un matrimonio no deseado y conseguir una exoneración de delitos acaecidos en el pasado. La cuarta novela es "Mathias Sandorf" (Verne, 1885), donde un mensaje cifrado interceptado lleva a impedir una revuelta revolucionaria.

No podía faltar Sherlock Holmes en este tipo de enigmas, y así, el escocés Arthur Conan Doyle (1859-1930) lo situó enfrentándose a varios de estos textos cifrados. En su novela "The 'Gloria Scott' " (Doyle, 1893) encontramos un mensaje cifrado (y esteganográfico) escondido en un texto mayor con inserción de palabras nulas, pues para leer el correcto hay que tomar las terceras palabras del mensaje. El segundo caso es el más famoso, "The Dancing Men" (Doyle, 1905), donde aparece un cifrado de sustitución por figuritas humanas en diversas posiciones de brazos y manos y en ocasiones con o sin banderines. Finalmente, en el relato "The Valley of Fear" (Doyle, 1915) el famoso detective consigue acertar en la suposición del libro de código usado por su némesis Moriarty, solucionando un cifrado extremadamente complejo.

En este artículo nos centraremos en la primera novela mencionada de Julio Verne, "Viaje al centro de la Tierra", por recoger los cifrados de sustitución (al menos monoalfabético) y de transposición, que son componentes fundamentales de los métodos criptológicos (Jara-Vera y Sánchez-Ávila, 2021). Esperemos sirva para apetecer zambullirse en el resto de relatos aquí mencionados.

§2. El cifrado de "Viaje al centro de la Tierra"

En la muy conocida novela que nos ocupa, casi al inicio de la misma, durante la conversación del profesor Otto Lidenbrock, especialista en mineralogía, con su

sobrino acerca de un preciado libro del escritor islandés del siglo XII-XIII Snorri Sturluson, recién adquirido en una vieja librería, se desprende de su interior un pergamino con un extraño mensaje (Fig. 1). Este, sin ninguna pista más, será el objeto de nuestra investigación. Vamos a intentar resolverlo de la manera más difícil, es decir, sin tener en cuenta ninguno de los aspectos relativos al libro en el cual estaba escondido o su autor, es decir, intentaremos el descifrado basándonos en el solo texto.

2.1. Criptoanálisis.

FIGURA 1. Mensaje original de la novela (Verne, 1864).

Este extraño mensaje parece consistir en un conjunto amplio de signos fácilmente separables. Analizando la lista de elementos y entendiendo que es un texto manuscrito donde los elementos no tienen que ser exactamente iguales, podemos contar un total de 136 signos que podemos agrupar por su aparente similitud en diversos conjuntos (Fig. 2).

FIGURA 2. Agrupación inicial de signos por su similitud.

Una mirada más precisa sobre la lista de elementos nos lleva a fijarnos en el parecido entre algunos de ellos. En concreto, los de la (Fig. 3).

FIGURA 3. Algunos aparentes parecidos entre signos.

Los pares son similares a los impares respectivamente, pero con un cierto marcado en las partes finales de los trazos tal y como se haría con las letras mayúsculas en las lenguas provenientes del latín.

Y es también remarcable que haya algunos signos, los menos comunes, que son como tildes, comas o puntos, como si fueran signos de puntuación, como apreciamos en la parte final del listado de la (Fig. 2). Hemos de ver qué significado pueden tener. Porque aún no podemos afirmar que sea un texto cifrado, por ahora no sabemos nada del mismo, o casi nada. Decir que curiosamente, dentro de la historia de la Criptografía, al menos en la Criptografía clásica, han sido habitualmente signos que no han solido cifrarse, los signos de puntuación, dándose por supuesto en la lectura del texto descifrado, que podía irlos colocando por el contexto.

Al repasar la lista de frecuencias (Fig. 2) apreciamos que la cantidad de signos diferentes es pequeña. Incluso si consideramos signos diferentes los indicados previamente (Fig. 3) y no pensamos en signos de puntuación insertos, tendríamos un total de 25, que coincide muy bien con una estructura lingüística vocálico-consonántica, y no con una estructura silábica, que precisaría de varias veces más de signos, ni estructuras intermedias entre ambas. Este es el primero de los análisis que tenemos que realizar, ver la estructura de la lengua subyacente.

En el caso más restrictivo de no contar como distintos los cuatro signos de la (Fig. 3) ni tampoco los dos finales del listado de la (Fig. 2) tendríamos un total de 19 elementos, lo que es algo menos de lo esperado en una lengua vocálico-consonántica, si bien el texto bajo estudio (Fig. 1) no es muy largo y podemos esperar que no todos los signos, por ejemplo, de un alfabeto amplio, como {a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z}, se hayan utilizado.

Otro aspecto que salta a la vista es la curiosa exposición de elementos, agrupados en lo que parecen ser tres columnas distintas y separables. Atendiendo a la cantidad de unidades o elementos que aparece en cada línea de cada columna tenemos un recuento altamente regular (Fig. 4).

7	7	7
7	7	7
7	7	7
7	6	6
6	6	6
6	6	6
6	6	6

Figura 4. Disposición estructural de los signos del mensaje.

A partir de este momento vamos a intentar suponer la lengua en la cual está escrito el texto. Un primer grupo de lenguas en las que podemos pensar es el gran Filo Indoeuropeo, o en los Filos Urálico, Caucásico y Altaico. Con esto completaríamos todo el arco de Europa y el Asia Occidental.

Existe un invariante muy usado en Criptografía, el denominado Índice de Coincidencia, el cual nos permite acotar la posible lengua base del mensaje, ya que nos

da la frecuencia relativa de encontrar en dos textos el mismo carácter en la misma posición (Friedman, 1987; Bauer, 2007), un dato que se mantiene con gran regularidad en los lenguajes a pesar del paso del tiempo (Jara-Vera y Sánchez-Ávila, 2019), estando definido como

$$I_c = \frac{\sum_{i=1}^{m} q_i(q_i - 1)}{Q(Q-1)},$$

donde Q es el número de elementos de la secuencia literal, q_i las respectivas apariciones de cada grafema en la secuencia literal y m la cantidad de elementos alfabéticos. La expresión evalúa la probabilidad de elegir dos caracteres iguales de forma aleatoria, de ahí el producto $(\frac{q_i}{Q})(\frac{q_i-1}{Q-1})$, que proviene del cálculo de las combinaciones $C_{q_i,2} = \frac{1}{2}q_i(q_i - 1)$ dividido entre $\frac{1}{2}Q(Q-1)$, agrupadas en un sumatorio sobre cada uno de los caracteres o elementos alfabéticos.

El índice I_c mide la variación de la frecuencia de aparición de los caracteres en un texto, y como hemos dicho, suele estar muy acotado para cada lenguaje. Además, y es lo que a nosotros nos interesará, este valor se mantiene aun cuando el texto haya sido cifrado por transposición o sustitución monoalfabética, e incluso permite comprobar si el mensaje ha sido cifrado usando un modo polialfabético, pues en este último caso ofrece un valor cercano a cero.

Para nuestro texto cifrado el valor del I.C. sin considerar los valores mayúscula es de $0{,}071203777$, mientras que en el caso de tenerlos en cuenta es de $0{,}072287763$. Estos valores nos llevan a pensar en que no se han aplicado cifrados polialfabéticos, y en todo caso solo permutaciones o sustituciones monoalfabéticas. Sin embargo, los valores obtenidos son muy comunes en multitud de lenguas, incluso de diversos Filos lingüísticos diferentes (Jara-Vera, 2016), por lo que prescindiremos de este invariante y buscaremos otros modos de atacar nuestro texto.

A pesar de lo poco que sabemos hasta ahora de nuestro mensaje secreto, es claro no obstante, que los signos gráficos del mensaje nos obligan a considerar las runas como un lenguaje que quizás resuelva el significado de lo que tenemos delante, o cuanto menos, ayuden en algo. La grafía es de alfabeto rúnico, si bien esta denominación, más que única, corresponde a un conjunto de alfabetos diversos más o menos similares que sirvieron de base para escribir lenguas germánicas durante la Edad Antigua y Media, siendo desplazadas poco a poco por las formas latinas a partir del siglo VIII. Las inscripciones más antiguas se remontan a la mitad del siglo II, habiendo persistido su uso hasta el siglo XX en algunas zonas rurales de Suecia. Tratando la diversidad de subfamilias, hemos de citar el futhark antiguo, que se desarrolló desde el siglo II hasta el IX. Entre tanto, en el siglo V surgió la variante anglosajona o futhorc, presente hasta el siglo XII. También, entre el siglo IX y el XII se desarrolló el futhark escandinavo, del que se derivarían las formas danesas, las sueco-noruegas, las islandesas, las de Hälsingland, las marcomanas, las medievales y las dalecarlianas (Arntz, 2007).

Sin entrar en los detalles de las variantes y considerando las más cercanas en su trazado a las de nuestro texto, hemos de quedarnos con seis posibilidades (Fig. 5). Recogemos cada uno de los signos del mensaje que estamos analizando y damos la transliteración entre corchete y el sonido fonético IFA correspondiente en los casos en que existe igualdad. En los que no existe igualdad damos la runa más cercana con sus opciones transliterada y fonética.

	futhark antiguo	futhorc anglosajón	futhark escandinavo	runas marcomanas	runas medievales	runas dalecarlianas
ᚨ	ᚷ: [g], /g/ ᚾ: [n], /n/	ᚷ: [g], /g/ ᚾ: [n], /n/	[a], /a(:)/	[n], /n/	[æ], /ɛ/,/æ/	[a], /a(:)/
ᚩ		[k], /k/	[ʀ], /ɹ/, /r/		[y], /y/	[o], /o/
ᛏ					[t], /t/, /d/	[t], /t/, [e], /e/, [n], /n/
ᛏ	[t], /t/	[t], /t/	[t], [d] /t/, /d/	[t], /t/	[t], /t/, /d/	[y], /y/
ᛋ		[s], /s/	[s], /s/	[s], /s/	[s], /s/	[s], /s/, [c]
ᚴ		[k], /k/			[n], /n/	[n], /n/
ᛁ	[i], /i(:)/	[i], /i(:)/	[i], /i(:)/	[i], /i(:)/	[i], /i/, /j/, /e/	[s], /s/, [i], /i/
ᚴ			[k], /k/, [g], /g/	[ch]	[k], /k/, [g], /g/	[k], /k/, [y], /y/
ᚢ	[u], /u(:)/	[u], /u(:)/	[u], [o], /u(:)/, /y(:)/, /ɔ(:)/, /w/	[r], /r/, [u], /u/	[u], /u/, /w/, /v/	[u], /u/
ᛚ	[l], /l/	[l], /l/	[l], /l/	[l], /l/	[l], /l/	[l], /l/
ᛒ						[b], /b/
ᛉ	[z], /z/	[z], /z/	[m], /m/	[y], /y/	[m], /m/	[m], /m/
ᚠ	ᚠ: [f], /f/	ᚠ: [f], /f/	ᚠ: [f], /f/, [v], /v/	ᚠ: [f], /f/	ᚠ: [f], /f/, [v], /v/	[f], /f/
ᚫ	[a], /a(:)/	[æ], /æ(:)/	[ą], /ą̃/, [o], /o(:)/	[a], /a(:)/		[f], /f/
ᛒ	[b], /b/	[b], /b/	[b], /b/, [p], /p/	[b], /b/	[b], /b/	[b], /b/
ᚱ			ᚴ: [k], /k/, [g], /g/		ᚴ: [k], /k/, [g], /g/	[k], /k/, [y], /y/
ᚾ						
ᚼ		/ǂ/: [io], /jo/	/ǂ/: [h], /h/	[k], /k/	/ǂ/: [h], /h/	[m], /m/ /ǂ/: [h], /h/, [å]
ᛁ						
ᛋ				[s], /s/		
ᛏ						
ᛁ				[i], /i(:)/		
ᛉ						

FIGURA 5. Tabla de correspondencias de los elementos del mensaje con distintas subfamilias de runas.

A la vista de la (Fig. 5) no podemos decantarnos por una variante lingüística con claridad. Incluso la variante dalecarliana, que es la que más correspondencias tiene

mantiene varios signos sin clara identificación. Por otro lado, como ya indicamos (Fig. 3), parece que los cuatro signos últimos se corresponden a letras ya presentes.

Sin embargo, al hacer una transliteración el texto no tiene sentido alguno, ya se lea columna tras columna, de izquierda a derecha o viceversa; o bien se lea fila a fila sin considerar la división en columnas, tanto a derecha como a izquierda. Incluso parece que algunos de los signos han sido inventados, como ocurre con la 7ª y 5ª fila de la (Fig. 5) contando desde abajo. Todo lo cual nos hace dudar de que sea un texto escrito en un verdadero lenguaje rúnico y quizás sea más bien una utilización de estos signos junto con otros inventados similares escondiendo un texto en otra lengua. Estaríamos entonces claramente no ante un texto escrito en alguna lengua originada en el antiguo norte de Europa, sino ante un texto que podemos calificar como cifrado, en toda la extensión del término, un texto que esconde su significado por algún método o sistema. ¿Pero cuál es el sistema utilizado? Es más, ¿cuál sería el lenguaje del texto origen?

Volviendo a la estructura tan regular que hemos recogido en la (Fig. 4), hemos de decir que no surge esta disposición haciendo una escritura desde arriba hacia abajo, sino más bien al modo de ir rellenando desde arriba hacia abajo, línea a línea, y de izquierda a derecha (o quizás también en bustrofedon desde la esquina superior derecha) elemento a elemento, signo a signo, quedándonos sin más signos al llegar a la cuarta línea de la primera columna. De entre estas dos opciones consideraremos la más cercana a nuestros lenguajes (Fig. 6), volviendo si no obtuviéramos nada a la opción de bustrofedon.

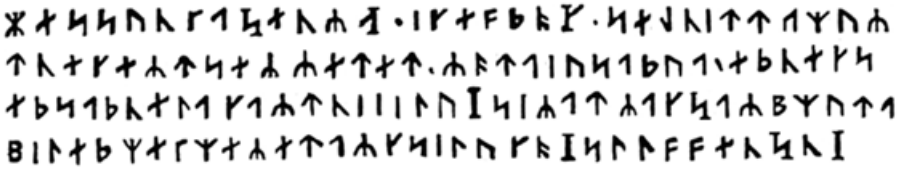

Figura 6. Posible recolocación del texto.

Intentando justificar si esta es la verdadera ordenación original del mensaje podemos fijarnos en los dos signos que aparentan ser como puntos. Notamos que delante de ellos lo que tenemos son dos signos que hemos considerado como posibles formas de letras mayúsculas. Pero si fueran puntos deberían estar detrás de ellos las mayúsculas y no delante. Por otro lado, el último signo es de nuevo una posible forma mayúscula. Estos aspectos tendrían sentido si el texto fuera no tanto el de la (Fig. 6) sino su inverso, el escrito de atrás hacia adelante (Fig. 7).

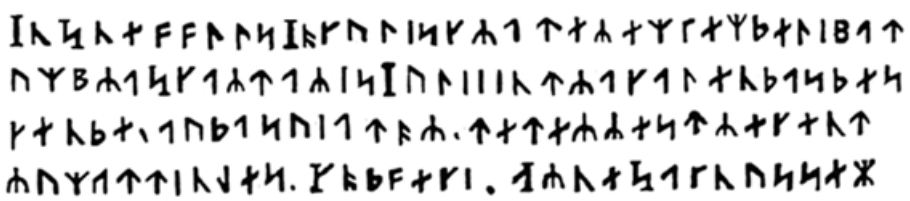

Figura 7. Más que posible recolocación del texto.

Volviendo a aplicar la transliteración con las diferentes posibilidades (Fig. 5) con esta nueva estructura no parece obtenerse ningún sentido dentro de las lenguas que les corresponderían, en especial las lenguas de la Familia Germánica, Subfamilia Septentrional, del Filo Indoeuropeo, como el islandés, danés, noruego, sueco o elfdaliano (o dalecarliano).

Con estos resultados podemos decir que nuestro texto ha sido sometido tanto a un cifrado de sustitución utilizando signos de aspecto rúnico, y además a un cifrado de transposición, con la colocación elemento a elemento en los distintos bloques y columnas junto con la lectura inversa. Queda por ver ahora la correspondencia de cada signo, de cada runa o pseudo-runa, con alguna determinada letra del lenguaje original del texto, que aún desconocemos.

De ahí que un siguiente cometido sea el análisis de frecuencias de los signos presentes, el cual suele ofrecer numerosas indicaciones de la lengua originaria, pues es un dato suficientemente ligado a cada lengua y mantenido diacrónicamente(Jara-Vera y Sánchez-Ávila, 2019). Calcularemos las dos opciones siguientes (Fig. 8), ambas sin tener en cuenta los signos que parecen ser puntos y comas, que tampoco son demasiados para alterar mucho los porcentajes: a) sin considerar los cuatro signos que parecen ser versiones mayúsculas; b) considerándolos y asimilándolos a sus posibles versiones minúsculas.

Los datos obtenidos son muy corrientes en multitud de lenguas, pudiendo inicialmente ceñirnos al ámbito Indoeuropeo o tal vez del Urálico, como Filos más probables, y más extraordinariamente, podríamos pensar en lenguas de los Filos Altaico o Caucásico. Por lo tanto, podemos suponer un cifrado de sustitución donde cada signo de las runas y pseudo-runas han sido tomados como reemplazos de un signo de un alfabeto, posiblemente un alfabeto latino.

Considerando los primeros 7 signos más habituales del texto, que suponen alrededor de algo más de la mitad del total del mismo, un 59,56 %, vamos a hacer las sustituciones con un conjunto amplio de lenguas, como son las siguientes: español, portugués, italiano, francés, inglés, alemán, holandés, danés, sueco, noruego, islandés, finés, húngaro, polaco, griego, latín y esperanto(Jara-Vera, 2016; ETA, 2021).

Al considerar los signos más habituales en las diversas lenguas que estamos considerando como hipótesis, la más repetida es una vocal, ya sea una "e" (ej.

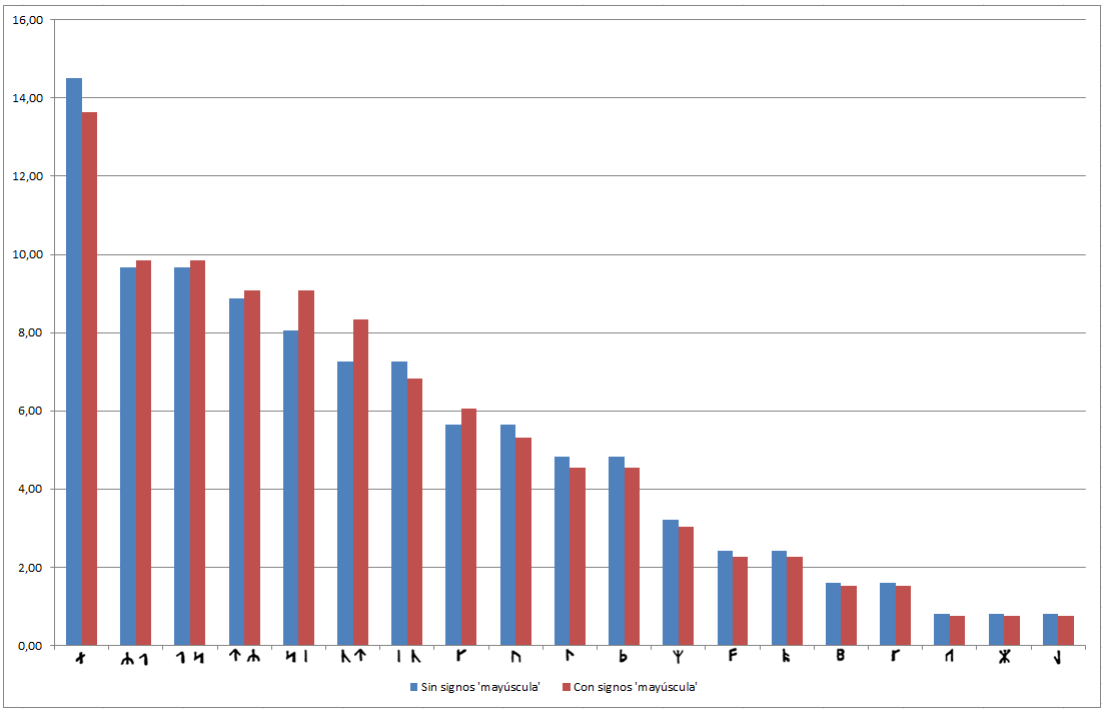

FIGURA 8. Frecuencias de los signos.

español), una "a" (ej. griego) o una "i" (ej. polaco). Viendo la (Fig. 5), este dato nos hace pensar como una posibilidad quizás elegida por el criptógrafo el darle al signo más repetido algunas de las opciones del futhark escandinavo, medieval o dalecarliano, lo que vendría a ser una letra como la "a" o la "e", si bien podría haber tomado otra elección. También es habitual en las lenguas consideradas que el segundo signo más repetido o bien sea una vocal (ej. español) o bien "t" (ej. inglés), "n" (ej. danés), "s" (ej. francés), "r" (ej. noruego), lo que nos hace pensar, si es que hubiera tenido en cuenta valores transliterados, en el futhark escandinavo, el medieval o el dalecarliano. De igual forma, para la tercera letra más repetida podemos también quedarnos con letras que tenemos en la (Fig. 5), como son "e", "t", "n", que encontramos en esas posiciones en varias lenguas, como el esperanto, húngaro o sueco, respectivamente.

Con este procedimiento podemos pensar en la hipótesis de la transliteración de los diversos signos, y con ello, considerando los siete más habituales, probaremos las siguientes opciones:

$$(a/e) - (r/y/o) - (e/t/n) - (t/d/y) - (s) - (k/n) - (i/j/s)$$

Por otro lado, debido a la presencia de la letra "a" en las lenguas que estamos considerando en los primeros lugares por su elevada frecuencia, también consideraremos en todos los lugares su posible existencia como opción posible.

Al realizar las diversas sustituciones hay en concreto una que parece ofrecer una cierta presencia de significado: e-r-a-t-s-n-i. Y que en gran parte parece acercarse a la (Fig. 5). La recogemos en la (Fig. 9).

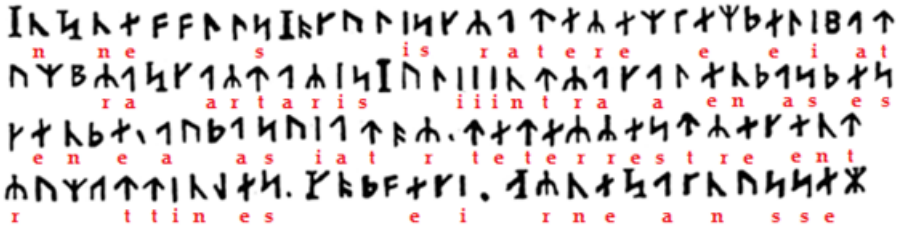

FIGURA 9. Primera aproximación a la solución, con 7 letras (minúsculas).

En la 3ª fila podemos suponer la palabra "terrestre", que tenemos en el español. Si añadimos tres de los signos que suponemos ser formas mayúsculas, y les damos los posibles valores de sus formas minúsculas consideradas, "a", "s", "i", tendríamos:

InSne - - - - sI - - - - is - ratere - - e - - e - i - at
- - - raS - artarisI - - iiintra - a - en - as - es
- en - e - a - - as - iat - r - teterrestre - ent
r - - - ttin - es - - - - - e - i - ArneSa - n - sse -

No obstante, la posibilidad de la escritura en español se desvanece por la presencia de otras formas ajenas que podemos ver. Pero parece ser acertada la suposición de considerar la transliteración de los signos. Incluso podemos añadir alguno más del conjunto de las runas de la (Fig. 5) que sea el mismo en todas las variantes y que no sea una opción posible para varios signos, como es el caso de la "l":

InSne - - llsI - - - lis - ratere - - e - - eli - at
- - - raS - artarisI - liiintra - alen - as - es
- en - e - a - - as - iat - r - teterrestre - ent
r - - - ttin - es - - - - - e - i - ArneSa - n - sse -

El texto parece ir cobrando sentido, si bien no es sencillo, apareciendo también secuencias extrañas como "iii". Aún está pendiente conocer la lengua origen del texto, mientras que la posibilidad de que fuera español ha desaparecido. Por otro lado, la última línea, tras lo que parece realmente un punto, parece mencionar, al suponer una mayúscula inicial, *Arne Sa...*, lo cual nos lleva a pensar en una firma final, un nombre y un apellido, o un título, siendo *Arne* un nombre masculino habitual escandinavo, coherente además con el uso de las runas.

Todo lo anterior parece confirmar los signos en mayúscula y el uso de puntos y comas. Lo cual, por otro lado, no es una buena práctica porque facilita la labor al criptoanalista que desee conocer el texto de manera no autorizada. Con ello, si vamos a la primera línea en el inicio, tendríamos el comienzo *In Sne...* lo

que nos acerca a lenguas como el inglés, pero que descartamos por la aparición de la expresión *terrestre*, llevándonos a pensar en el latín, que sí tiene la forma "terrestris/terrestre", y además la preposición "in".

Veamos a continuación las frecuencias de la lengua latina. Para ello tomaremos en consideración el análisis de diversas obras, en concreto diez, a saber, una sátira de Juvenal, fragmentos de la "Vulgata" del evangelio de san Lucas, también de "Confesiones" de san Agustín, poemas de Venancio Fortunato, de "Etimologías" de san Isidoro, del "Carmina Burana", de himnos de Pedro Abelardo, de un sermón de san Antonio de Padua, así como algunos textos de la "Summa de Teología" de santo Tomás de Aquino, junto a algunos versos de Petrarca (Jara-Vera, 2016), lo cual nos da un abanico de obras en prosa y en verso desde el siglo I hasta el XIV (Fig. 10).

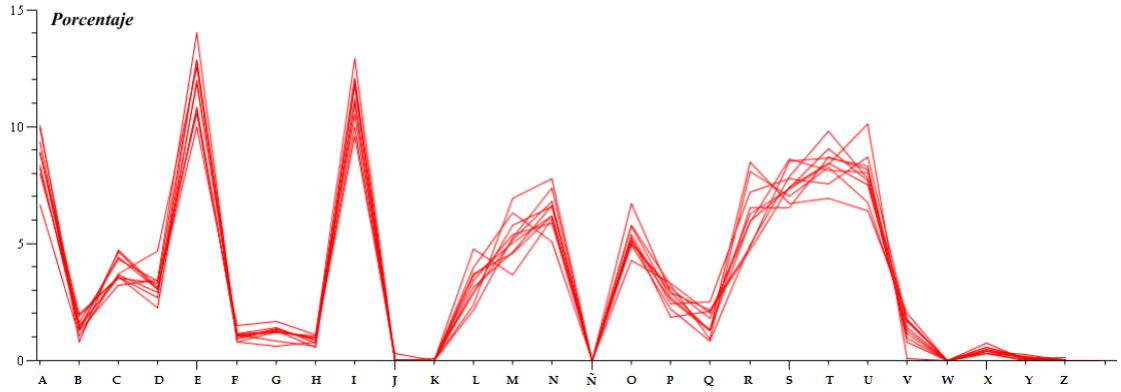

FIGURA 10. Frecuencias de las letras en la lengua latina.

La solución última dada para el mensaje cifrado, a pesar de que el texto no sea muy extenso, arroja una cierta coherencia con las frecuencias latinas. Así, hemos colocado numerosas letras de entre las más frecuentes: *e-a-s-r-i-t-n-*-*-l-**, con una presencia de *18-13-13-12-12-11-9-8-7-6-6*.

Echamos en falta, viendo las frecuencias latinas, la necesidad de localizar los valores "m", "u", "o" en nuestro texto, siendo valores con esperada presencia. Podemos suponer que se encuentran en los lugares que hemos marcado con asteriscos en la secuencia previa, y que corresponden a los signos rúnicos pendientes (Fig. 11).

FIGURA 11. Tres signos a los que buscar su significado.

Considerando la (Fig. 11), la suposición de que el signo situado a la izquierda sea una "m" no parece tener sentido al sustituirse, generándose expresiones absurdas,

como "lismratere". Lo mismo si suponemos los valores "u" y "o". Tampoco parece que el signo de la derecha encaje con ninguna de las opciones "m", "u", "o", mirando por ejemplo, la secuencia "alen*as*es". Para el signo central encontramos que el valor "u" ofrece altas posibilidades, quedándonos el siguiente texto, (Fig. 12), que mostramos con los signos originales.

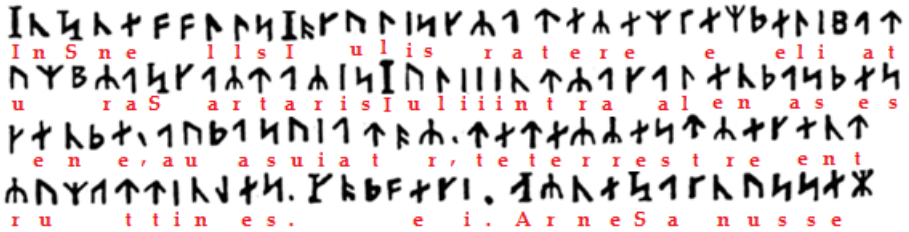

FIGURA 12. Solución parcial con 9 letras.

Centrémonos ahora en el signo situado a la izquierda de la (Fig. 11), con 7+1 apariciones (minúsculas+mayúscula) en el texto. Mirando las frecuencias de la lengua latina, y eliminando las opciones colocadas y las que hemos descartado para él, podemos pensar en "p" y en "c". Es claro al hacer algunas sustituciones que el valor que ofrece varios sentidos en el texto es "c". Además, esto es coherente con la (Fig. 5), donde una de las opciones era el valor "k". El texto queda de esta forma:

InSne - - llsI - culiscratere - - e - - eli - at
u - - raScartarisIuliiintracalen - as - es
cen - e, au - asuiat - r, teterrestrecent
ru - - ttin - es. C - - - eci. ArneSa - nusse -

Para el signo situado en la derecha (Fig. 11), la letra más inmediata viendo la (Fig. 5) de runas y pseudo-runas es la "b", pero no obtenemos un sentido claro con esta asignación. También parece que no corresponde a un valor vocálico, sino consonántico. Viendo las letras que nos quedan pendientes y su no escasa frecuencia parece encajar con "d", llegando al mensaje mostrado en (Fig. 13).

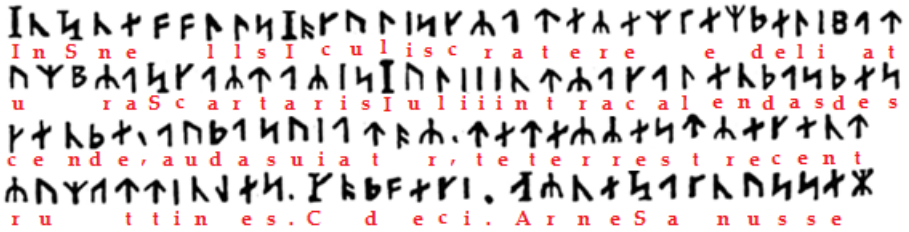

FIGURA 13. Solución parcial con 11 letras.

Ya es fácil darse cuenta por el sentido que tenemos las palabras "descende, audas viator", lo que nos indica que en ocasiones el valor "u" hace las veces de "u" o de

"v", lo cual es propio del latín, además de resolver el valor para el tercero de los signos de la (Fig. 14) al que le asignamos la "o", por otro lado coherente con la variedad futhark escandinava.

Figura 14. Cinco signos a los que buscar su significado.

Además, tenemos en el lenguaje latino algunas letras que deberían de aparecer (Fig. 10): en especial "m" y "p", con porcentajes no del todo escasos, en torno al 5 %-3 %; con en torno al 1 % podemos esperar los valores "b", "f", "g", "h", "q". Viendo además la (Fig. 5) podemos hacer las suposiciones siguientes para la (Fig. 14): pensar en "m" para el primero o el quinto de los signos; "f" para el segundo; "p" para el cuarto. Al ir sustituyendo podemos comprobar que son adecuadas las suposiciones excepto para el caso de "p", que no parece ser correcto. Por otro lado, para la "m" es llamativo que corresponda bien para ambos signos, el primero y el quinto. Otra opción, y por el doblado de signo, arriba y abajo, frente al primero, es suponer que el quinto corresponda al valor "mm". El texto quedaría así:

InSnefflsIoculiscraterem - emdeli - at
um - raScartarisIuliiintracalendasdes
cende, audasuiator, teterrestrecent
rum - ttin - es. Codfeci. ArneSa - nussemm

En este momento podemos intentar para el cuarto de los signos de la (Fig. 14) el que la (Fig. 5) ofrecía también, el valor "b", lo cual parecer ser correcto.

Figura 15. Dos signos parecidos.

Hay dos signos parecidos, uno de los cuales ya resuelto (Fig. 15), el de la derecha, con valor "c". Para el otro, el de la izquierda, intentando las opciones "k", "g", "y", presentes en la lista de la (Fig. 5), vemos que parece ser el valor "k", cercano al valor "c", dado al otro signo tan parecido a él, por lo que fonéticamente tienen una gran semejanza. No obstante, mirando el mensaje, todo es algo más confuso y no tan sencillo, pues el signo de la derecha en ocasiones obliga a "c", como en "craterem", pero en otras a "qu", como parece sugerir su forma mayúscula en "Cod feci", debiendo ser en latín "Quod feci". Esto hace que el signo a la izquierda de la (Fig. 15), que tenemos en dos ocasiones en el texto, en una sea "k", para "Saknussemm", más propio de las lenguas escandinavas, y en otra sea "qu", para "quem".

Así, llegamos al texto siguiente, que permite entender el llamativo "iii", y donde podemos ver que hay un error con la preposición "te" en "...te terrestre", debiendo decir "et terrestre":

In Sneffls Ioculis craterem quem delibat
umbra Scartaris Iulii intra calendas des
cende, audas viator, te [=et] terrestre cent
rum - ttin - es. Quod feci. Arne Saknussemm

¿A qué corresponden los dos extraños signos de la (Fig. 16), pseudo-runas pendientes, que en la (Fig. 5) no concuerdan con runas reales existentes? ¿Acaso con alguno de los valores no utilizados como "g", "h", "p", "x", "y", "z"? No parece ser el caso para el primero de ellos, que parece pedir una vocal, si bien para el segundo podemos pensar en la forma latina "-ttinges", y por lo tanto en una "g", forma verbal que tenemos en "attingo, attigi, attactum", que significa "alcanzar, tocar ligeramente", y nos da la vocal "a" para el primero.

Figura 16. Últimos signos pseudo-runas.

Indicar además que la palabra latina "audas" corresponde más concretamente a "audax", habiéndose así sustituido tanto la "s" como la "x" por el mismo signo, lo cual es un caso extraño en Criptografía, por llevar a dificultades de recuperar el texto original, si bien solo se ha usado en esta ocasión y con letras cercanas fonéticamente. De esta forma, el texto queda tal y como recoge la (Fig. 17).

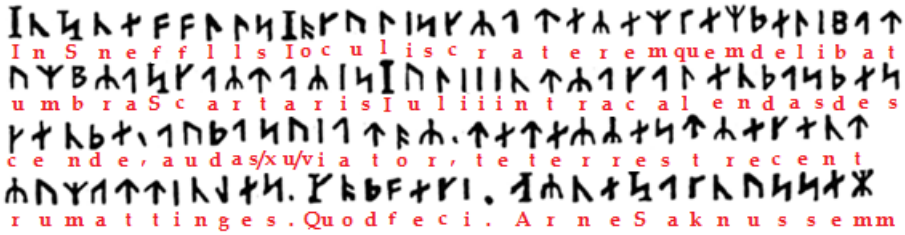

Figura 17. Solución del mensaje.

Sin entrar en detalles de traducción, podemos reconocer la forma vocativa de "audax viator" con la coma precedente, el verbo al final, "descende", en presente de imperativo, necesitando de un acusativo, que encontramos en "craterem", el cual lleva una oración subordinada relativa con el mismo caso en "quem", la cual tiene un verbo en presente de indicativo ligado al complemento temporal "Iuliis intra calendas", donde "intra" adquiere un significado de "antes de", en lugar del habitual "dentro de", por la razón de referirse a una medida de tiempo, las calendas

o primer día de mes; además, tenemos el complemento de lugar con la preposición "in" para "Sneffls Iouculis"; seguida de una coma tenemos una oración unida por la conjunción copulativa "et" y un verbo en futuro imperfecto de indicativo, "attinges". Finalizando el texto tenemos en "Quod feci. Arne Saknussemm" un falso relativo, debido a la presencia del punto, que separa ambas sentencias y rompe la subordinación; así "quod" adquiere el sentido de pronombre anafórico "is/ea/id", y se refiere a la frase precedente, y el verbo es un pretérito perfecto de indicativo en 1ª persona del singular. De esta forma el mensaje,

In Sneffls Ioculis craterem quem delibat umbra Scartaris Iulii
intra calendas descende, audas[=x] viator, te [=et] terrestre centrum attinges.
Quod feci. Arne Saknussemm

tendría como traducción ajustada desde el latín al español:

Audaz viajero, desciende el cráter que la sombra del Scartaris roza
antes del 1 de julio en el Sneffls Ioculis, y alcalzarás el centro de la tierra.
Lo cual he hecho. Arne Saknussemm

En la novela las cosas son más sencillas: primeramente se toma algunas licencias suponiendo que es rúnico islandés y con ello el sabio profesor Lidenbrock copia de manera directa las letras latinas. Por otro lado, revisando el antiguo libro donde estaba el pergamino encuentra en un margen escrito el nombre "Arne Saknussemm" con los mismos símbolos rúnicos. Junto a ello, de manera inmediata supone que el texto está escrito en latín, por ser la lengua culta por antonomasia. A partir de aquí, de forma sencilla acabarán, tanto él como su sobrino, ordenándolo y leyéndolo de manera inversa (Fig. 18).

In Sneffels Yoculis craterem kem delibat
umbra Scartaris Julii intra calendas descende,
audas viator, et terrestre centrum attinges.
Kod feci. Arne Saknussemm.

Descends dans le cratère du Yocul de
Sneffels que l'ombre du Scartaris vient
caresser avant les calendes de Juillet,
voyageur audacieux, et tu parviendras
au centre de la Terre. Ce que j'ai fait.
Arne Saknussemm.

FIGURA 18. Solución al texto en latín y francés (publicación original de la novela (Verne, 1864)).

Como podemos ver al comparar la solución obtenida con la dada por Verne (Fig. 18) hay algunos errores y diferencias, pues en lugar de "Sneffls" aparece

"Sneffles", cambiando una "l" por una "e". También el cambio de "I" por "Y" desde "Ioculis" a "Yoculis", si bien comprensible por la similitud fonética. También el error ya mencionado de "te" por "et". Junto a ello, mencionar la inexistencia de la palabra "Ioculis", que proviene de una voz islandesa, "Jökull", con significado de "glaciar", y que suele formar parte sufija de muchos picos de la región, y que Verne traduce como "ventisquero" (Lexilogos, 2021). Así, el nombre en islandés es "Snæfellsjökull", que es verdaderamente un volcán islandés. Indicar que "Scartaris" no es un pico o monte existente, sino que parece una invención de Verne. No obstante, Arne Saknussemm está basado en el del erudito islandés Árni Magnússon (1663-1730) (Margot, 1980).

2.2. Algunas formalizaciones. Demos algunas definiciones algo más formales de los tipos de cifrados que acabamos de criptoanalizar.

Sea $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$ un criptosistema, formado por el conjunto de textos en claro $\mathcal{P}$, textos cifrados $\mathcal{C}$, espacio de claves $\mathcal{K}$, así como las operaciones de cifrado $\mathcal{E}$ y descifrado $\mathcal{D}$.

En general $\mathcal{P} = \mathcal{C}$, habiendo previamente un alfabeto, más o menos amplio,

$$L = \{A, B, C, ..., Z, a, b, ..., z, 0, ..., 9, ...\},$$

incluso con algunos signos de puntuación o no, con cardinal $\#L = m$, capaz de constituir los diferentes elementos de $\mathcal{P}$ y $\mathcal{C}$. En nuestro caso

$$L_1 = \{A, B, ..., G, I, K, ..., N, O, ..., U, V, X, Z, a, b, ..., g, i, k, ..., n, o, ..., u, v, x, z, ., ,\},$$

con $\#L_1 = 46$, suponiendo un alfabeto latino de 22 letras, además del punto y la coma.

Para el caso de las runas hemos encontrado una asignación para las letras como la que mostramos en la (Fig. 19). No sabemos si hay más casos como el de la forma doble "mm". Suponiendo que no, y echando en falta las asignaciones "p" y "z", que suponemos que existan, tendríamos 21 signos para las minúsculas, con otros tantos para las mayúsculas, y al menos dos valores más para la coma y el punto, resultando así un cardinal, si no hay más elementos, de $\#L_2 = 44$.

En los cifrados simétricos tanto la operación de cifrado e como de descifrado d hacen uso de una misma clave k. En los cifrados asimétricos no son iguales dichas claves si bien están relacionadas matemáticamente entre sí. Con ello, de manera general tendríamos:

$$e_k : \mathcal{P} \to \mathcal{C},$$

$$d_k' : \mathcal{C} \to \mathcal{P},$$

a		g		n		t	
b		i		o		u	
c		k		p	¿?	v	
d		l		q[u]	(mayúsc.)	x	
e		m		r		z	¿?
f		mm		s			

FIGURA 19. Asignación de las runas y pseudo-runas (solo minúsculas, con una excepción).

de tal forma que

$$d_k{}'(e_k(x)) = x,\ \forall x \in \mathcal{P},\ \ k, k' \in \mathcal{K},\ \ e \in \mathcal{E},\ \ d \in \mathcal{D}.$$

2.2.1. Cifrados de sustitución y de permutación. Vayamos, dicho lo anterior, a formalizar los cifrados que hemos encontrado en nuestro texto. Hemos tenido tanto un cifrado de sustitución, donde las letras latinas eran sustituidas por runas o pseudo-runas, y dos cifrados de permutación o transposición, donde las letras eran cambiadas de lugar o de posición, por un lado, escribiendo el texto original al revés, de atrás hacia adelante, y por otro lado, colocando en la estructura tabulada de tres columnas y siete filas (Fig. 4) una letra tras otra en cada posición. Podemos intercambiarlos entre sí, procediendo primero con la sustitución y luego con las permutaciones o al revés. Quizás lo más cómodo en el procedimiento sea empezar con la transposición que construye el texto en sentido inverso, seguir con la que lo coloca en la estructura de columnas y filas, y finalmente hacer la sustitución de las letras por signos rúnicos.

De manera general la transposición no es sino una permutación algebraica, que podemos expresar del modo conocido:

$$e_\pi(x_1, x_2, ..., x_n) = (x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(n)}), \tag{2.1}$$

para el caso de un mensaje de longitud n, siendo la función una biyección, y que en el caso de aplicarse monográmicamente sobre un texto de longitud n la podemos expresar como

$$\mathcal{P}_n^{(1)} \leftrightarrow \mathcal{C}_n^{(1)},$$

donde para nuestro ejemplo $\mathcal{P} = \mathcal{C}$ y el alfabeto es L_1.

En el caso de la permutación consistente en escribir el texto en sentido inverso, π_1, tendríamos que (2.1) debe de cumplir que $\forall\, i \in \{1, 2, ..., n\}$:

$$\pi_1(i) :\ n + 1 - i,$$

es decir,

$$e_{\pi_1}(x_1, x_2, ..., x_n) = (x_n, x_{n-1}, ..., x_1),$$

donde $n = 136$, $\mathcal{P} = \mathcal{C}$ y el alfabeto es L_1.

Para el segundo cifrado de transposición, que coloca el actual texto en la estructura de columnas y filas (Fig. 4), transposición más compleja de formalizar, podemos expresarla, haciendo una lectura de arriba hacia abajo, y de izquierda a derecha, donde la ecuación (2.1) se expresa para todo $i \in \{1, 2, ..., n\}$, con $n = 136$, como

$$\pi_2(i) : \left(\left\lfloor \tfrac{i+71}{71} \right\rfloor \pmod 2\right) \left\lfloor \tfrac{i+6}{7} \right\rfloor + \left\lfloor \tfrac{i}{71} \right\rfloor \left\lfloor \tfrac{i-5}{6} \right\rfloor + \\ + 21\Big(i - \Big(\left(\left\lfloor \tfrac{i+71}{71} \right\rfloor \pmod 2\right)(7\left\lfloor \tfrac{i-1}{7} \right\rfloor + 1) + \left\lfloor \tfrac{i}{71} \right\rfloor (6\left\lfloor \tfrac{i-71}{6} \right\rfloor + 71)\Big)\Big),$$

quedando expresado en algunos pocos términos de la siguiente forma:

$$e_{\pi_2}(x_1, x_2, x_3, ..., x_8, x_9, ..., x_{135}, x_{136}) = (x_1, x_{22}, x_{43}, ..., x_2, x_{23}, ..., x_{105}, x_{126}),$$

así como gráficamente en todos sus valores en la (Fig. 20).

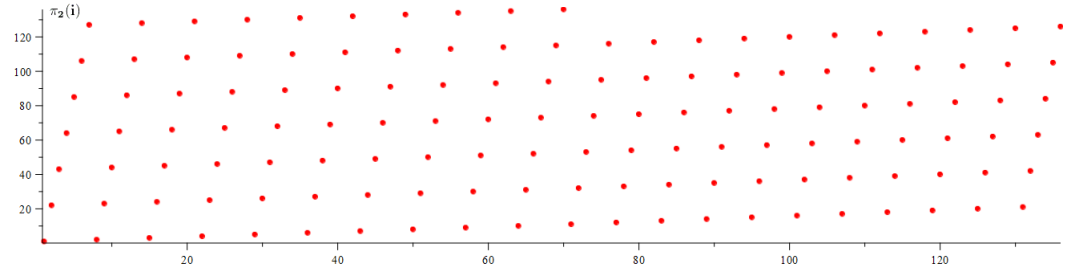

FIGURA 20. Función $\pi_2(i)$.

Sus funciones inversas de descifrado son respectivamente $d_{\pi_1^{-1}}$ y $d_{\pi_2^{-1}}$, donde $\pi_1^{-1}(i) = \pi_1(i)$ y

$$\pi_2^{-1}(i) : \left\lfloor \tfrac{i+20}{21} \right\rfloor + 6\Big((i-1) \pmod{21} - 10\Big) \left\lfloor \frac{(i-1) \pmod{21}+1}{12} \right\rfloor + \\ +7\Big(11 \left\lfloor \frac{(i-1) \pmod{21}+1}{12} \right\rfloor + \Big((i-1) \pmod{21}+1\Big)\Big(\left\lfloor \frac{((i-1) \pmod{21})+13}{12} \right\rfloor \pmod 2 - 1\Big)\Big),$$

con

$$d_{\pi_1^{-1}}(x_1, x_2, ..., x_n) = (x_n, x_{n-1}, ..., x_1)$$

$$d_{\pi_2^{-1}}(x_1, x_2, x_3, ..., x_8, x_9, ..., x_{135}, x_{136}) = (x_1, x_8, x_{15}, ..., x_{50}, x_{57}, ..., x_{63}, x_{70}),$$

siendo la representación gráfica de valores para esta última transposición la mostrada en la (Fig. 21).

En cuanto a la sustitución tendríamos

$$\sigma : \mathcal{P}_n^{(1,2)} \to \mathcal{C}_n^{(1)},$$

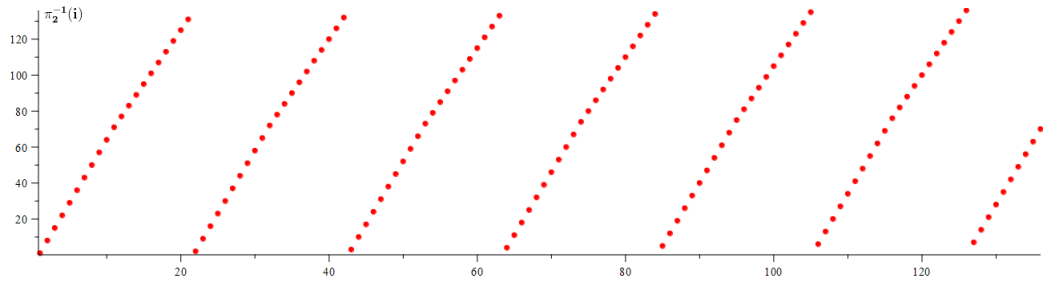

FIGURA 21. Función $\pi_2^{-1}(i)$.

donde $\mathcal{P} \neq \mathcal{C}$, al ser el conjunto de sustitución $\mathcal{C}$ el de runas y pseudo-runas, con su alfabeto L_2, diferente al inicial de letras de grafía latina, L_1.

Tal y como vimos en nuestro criptoanálisis nuestro texto no ha sido una biyección ni tampoco ha sido incluso una función (Fig. 19). En la mayoría de las ocasiones, es verdad que un elemento del conjunto origen ha sido sustituido siempre por un elemento, el mismo, del conjunto destino; pero en otras, ha habido más de una opción, como en el caso de "a"; o bien en otros casos, dos elementos diferentes como "s" y "x" han tenido la misma imagen; e igualmente a "k" (monograma) y a "qu" (bigrama) se les ha asignado la misma imagen. Y como hemos acabado de mencionar, no solo con una asignación monográmica, sino en ocasiones poligrámica: tomándose del conjunto origen uno, o más raramente dos elementos, pero solo uno siempre del conjunto destino, como ocurre con el caso "mm" o el antes mencionado "qu". Todo lo dicho lleva a que nuestra relación no sea una función, permitiendo la Criptografía relaciones y asignaciones altamente elaboradas, en muchos casos muy complejas para dificultar el ataque criptoanalítico, dejando ciertas ambigüedades que pudieran darse en las asignaciones a la posterior resolución con apoyo de las redundancias del lenguaje y el conjunto finito de las palabras posibles de la lengua.

Con todas las funciones de cifrado y descifrado definidas, tendríamos las composiciones

$$P_1 \xrightarrow{e_{\pi_1}} P_1 \xrightarrow{e_{\pi_2}} P_1 \xrightarrow{e_\sigma} C_2$$

$$C_2 \xrightarrow{d_{\sigma^{-1}}} P_1 \xrightarrow{d_{\pi_2^{-1}}} P_1 \xrightarrow{d_{\pi_1^{-1}}} P_1$$

donde los subíndices de los conjuntos indican los alfabetos a los que pertenecen, L_1 o L_2.

Otro aspecto que queremos resaltar es la escasa longitud del texto, incluso con presencia de palabras ajenas a la lengua en la que está escrito, el latín, como "Sneffls", "Ioculis", "Scartaris" o "Arne Saknussemm", que llevan a variar las frecuencias esperadas de las letras del idioma, dificultando un ataque de frecuencias.

Hay una pregunta que antes de finalizar deberíamos responder. Se trata de saber si nuestra solución es la correcta, si es única o hay alguna otra más que pueda obtenerse del texto cifrado original. Para atender a este asunto se define la Distancia de Unicidad,

$$D_U = \frac{H(\mathcal{K})}{D},$$

cociente entre la Entropía del espacio de claves y la Redundancia del lenguaje. Esta última, $D = R - r$, es la diferencia entre el Ratio Absoluta del lenguaje y la Ratio Real (Guerrero, 2009; Shannon, 1950, 1951).

Considerando que en ambas transposiciones operamos sobre L_1 y que en la sustitución es también éste nuestro lenguaje del conjunto origen, tendríamos un total de 46 valores, si consideramos distintos los signos mayúscula y minúscula junto a los signos de puntuación. Con ello y a partir de la definición de $R = \log_2(\#L)$, obtendríamos $R = 5{,}523$. En lenguajes como las lenguas romances tenemos una Ratio Real en el rango $1{,}2 \leq r \leq 1{,}5$. Así los valores obtenidos para la Redundancia estarían en el rango $4{,}323 \leq D \leq 4{,}023$, que podemos ponderar como $D = 4{,}173$.

Calculemos ahora la entropía del espacio de claves. Tengamos en cuenta que la composición de los dos cifrados de transposición es otro cifrado de transposición. El cálculo consistirá en toda las permutaciones posibles, es decir,

$$\log_2\left(\frac{136!}{18!12!12!11!10!9!9!7!7!6!6!4!3!3!2!2!1!1!1!3!1!3!1!2!2!}\right),$$

y la Distancia de Unicidad final es de $123{,}626$ caracteres.

Para el caso de un cifrado de sustitución monográmica monoalfabeto $H(\mathcal{K})$ viene dado como $\log_2(\#L_1!)$, resultando un valor de $45{,}964$. No obstante, nuestro cifrado de sustitución tenía en un par de ocasiones bigramas con "mm" o "qu", o una doble asignación para la "a", o la misma para la "s" y la "x", o la "k", la "c" y la "qu". No parece que haya más, pues son las que podríamos esperar por los parecidos fonéticos, por lo que podríamos aumentar el anterior valor ligeramente, sin entrar en más afinados cálculos, pongamos por caso, $\log_2((\#L_1 + 5)!) = 52{,}691$. Todo ello sin necesidad de suponer más parejas ni otras posibilidades de relaciones entre conjunto origen y destino.

Considerando finalmente la suma de ambas distancias de unicidad, las de transposición y de sustitución, tendríamos un valor total de $176{,}317$ caracteres. Como nuestro texto tiene 136 signos, un valor menor que el obtenido de unos 176 o 177 caracteres, no tenemos unicidad. Podría por lo tanto haber otra transposición distinta a las obtenidas (conjuntamente) y otra sustitución que desde el texto cifrado diera un mensaje con sentido. No obstante, la coherencia de la sustitución encontrada con las runas diversas, así como las dos transposiciones, una inversa y otra según la colocación ordenada y secuencial en la tabla por el patrón percibido, parece

indicar claramente que la solución encontrada sí es la correcta y la que generó el mensaje.

Para finalizar, y más en un artículo orientado al ámbito docente y educativo, mencionar la necesidad de un cierto conocimiento, a veces profundo y no solo somero, de la lengua origen del texto mensaje, en este caso el latín, lengua clásica por antonomasia, para poder ir adivinando las palabras que puedan ir apareciendo en sílabas y pequeños fragmentos. Cuánto más en casos como este donde hay varias opciones para un mismo signo, e incluso errores (tal vez intencionados para complicar el descifrado) en la generación del texto.

§3. Comentarios finales

No hemos querido decir apenas nada más del argumento de esta novela, ni del resto que hemos mencionado al inicio, todas ellas muy recomendables y entretenidas, al tiempo que permiten apreciar el interés por los mensajes secretos en multitud de autores de aventuras y misterio, llevando al lector a participar de la intriga intentando descifrarlos por su cuenta, lo que aquí hemos expuesto de forma detallada para la novela "Viaje al centro de la Tierra".

Leamos, o incluso releamos, algunas de estas obras tras haber dejado la adolescencia ya hace tiempo pero esta vez con ojos más agudos e interés científico-humanístico, para indagar con perspicacia, como ocurre con Julio Verne, la amplia variedad de conocimientos en diversas disciplinas que salpican sus obras, desde la matemática, pasando por las diversas ingenierías, a la geografía, la botánica, o la lingüística.

En lo que a nuestra especialidad se refiere, la Matemática Aplicada en Criptografía, creemos que puede ser muy sugerente iniciar a los más jóvenes en esta disciplina con este tipo de ejemplos, no tan ingenuos ni sencillos, sino incluso arduos, y potencialmente muy motivadores y efectivos para introducir diversidad de conceptos y métodos de esta área en los procesos educativos, al tiempo que se les ofrece un marco intelectual integral amplio en muchas otras materias del conocimiento.

Contribuciones

Conceptualización: VJV. Metodología: VJV. Investigación: VJV. Supervisión: VJV, CSA. Validación: VJV, CSA. Escribió el artículo: VJV.

Bibliografía

Arntz, H. (2007). *Handbuch der Runenkunde*. Leipzig, Deutschland: Lempertz.

Bauer, F. L. (2007). *Decrypted Secrets: Methods and Maxims of Cryptology*. Berlin, Deutschland: Springer.

Doyle, A. C. (1893). *The Memoirs of Sherlock Holmes*. London, England: George Newnes.

Doyle, A. C. (1905). *The Dancing Men*. London, England: George Newnes.

Doyle, A. C. (1915). *The Valley of Fear*. New York, United States of America: George H. Doran Company.

ETA. (2021). *LetterFrequency.org*. Descargado de http://letterfrequency.org/letter-frequency-by-language/

Friedman, W. F. (1987). *The Index of Coincidence and its Applications in Cryptanalysis*. Laguna Hills (CA), United States of America: Aegean Park Press.

Guerrero, F. G. (2009). A New Look at the Classical Entropy of Written English. *Transactions on Information Theory*, 1–15.

Jara-Vera, V. (2016). *Contexto, Criptoanálisis y propuesta de solución de la inscripción de la talla (original) de la Virgen de Candelaria de Tenerife (Canarias, España)* (Tesis Doctoral no publicada). ETSIT, Universidad Politécnica de Madrid, Madrid, España.

Jara-Vera, V., y Sánchez-Ávila, C. (2019). Graphemic-phonetic Diachronic Linguistic Invariance of the Frequency and of the Index of Coincidence as Cryptanalytic Tools. *Plos One*, *14*(3), 1–31.

Jara-Vera, V., y Sánchez-Ávila, C. (2021). Some Notes on a Formal Algebraic Structure of Cryptology. *Mathematics*, *9*(2183), 1–28. Descargado de https://doi.org/10.3390/math9182183

Kahn, D. (1996). *The Code-breakers: The Comprehensive History of Secret Communication from Ancient Times to the Internet*. New York, United States of America: Scribner.

Lexilogos. (2021). *Icelandic Dictionay.* Descargado de https://www.lexilogos.com/english/icelandic_dictionary.htm

Margot, J. M. (1980). Comment se Documentait Jules Verne. *Bulletin de l'Association Bibliothécaires Genevoise*, *4*, 5–12.

Poe, E. A. (1843). The Gold-Bug. *Dollar Newspaper*, *28 de junio*, 1–4.

Shannon, C. E. (1950). The Redundancy of English. *Cybernetics*, 248–272.

Shannon, C. E. (1951). Prediction and Entropy of Printed English. *Bell System Technical Journal*, *30*(50), 47–51.

Vargas-Llosa, M. (1971). *García Márquez: historia de un deicidio*. Barcelona, España: Seix Barral.

Verne, J. (1864). *Voyage au Centre de la Terre*. Paris: France: Pierre-Jules Hetzel.

Verne, J. (1868). *Les Enfants du Capitaine Grant*. Paris: France: Pierre-Jules Hetzel.

Verne, J. (1881). *La Jangada*. Paris: France: Pierre-Jules Hetzel.

Verne, J. (1885). *Mathias Sandorf*. Paris: France: Pierre-Jules Hetzel.

Vicente Jara-Vera
Dpto. de Matemática Aplicada a las Tecnologías de la Información y las Comunicaciones.
ETS de Ingenieros de Telecomunicación.
Universidad Politécnica de Madrid (*España*)
(✉) *vicente.jara@upm.es*

Carmen Sánchez-Ávila
Dpto. de Matemática Aplicada a las Tecnologías de la Información y las Comunicaciones.
ETS de Ingenieros de Telecomunicación.
Universidad Politécnica de Madrid (*España*)
(✉) *carmen.sanchez.avila@upm.es*

Recibido: *7 de agosto de 2021*.
Aceptado: *17 de julio de 2022*.
Publicado en línea: *6 de setiembre de 2022*.