

El gene, la forma, el virus y la idea: una perspectiva personal de la biología matemática

Jorge X. Velasco Hernández

Departamento de Matemáticas

Universidad Autónoma Metropolitana-Iztapalapa

Apdo. Postal 55-534

09340 México, D.F.

México

jxvh@xanum.uam.mx

1 Introducción.

Una tarde tranquila, no hace mucho tiempo, Juan José Rivaud me invitó a escribir un artículo sobre el desarrollo de la biología matemática en el siglo XX. Acepté. Todavía no se si hice bien, primeramente, porque hoy ignoro si el siglo XX ya terminó y estoy escribiendo para la *Miscelánea Matemática* del siglo XXI, o si todavía andamos en el viejo y conocido siglo XX que nos vió nacer y que, aún en su último año, puede darnos biomatemáticas sorpresas; en segundo lugar porque cuando me senté frente a la computadora, con mi taza de café sobre la mesa, y empecé a planear el artículo me dí cuenta de lo formidable de la tarea: debía dar un panorama histórico escueto pero informativo, una opinión sobre las áreas de la biomatemática más prometedoras para la investigación en los próximos años, además de conservar la filosofía editorial de la revista. La palabra estaba dada sin embargo y además el reto era interesante. He aquí el resultado. El lector juzgará la calidad y generalidad de la obra y a su juicio me atengo.

Antes, sin embargo, de entrar en materia, quiero dejar sentado mi punto de vista y concepción sobre la disciplina que reseño y en la que trabajo. Sin pretender definir lo que pudiera ser la “biología

matemática”, es un hecho que la aplicación de matemáticas en el estudio de problemas biológicos, tanto de carácter práctico como teórico, ha cobrado importancia en los años recientes desarrollándose fundamentalmente al amparo de colaboraciones multidisciplinarias entre científicos de diversas áreas biológicas y matemáticos interesados en aplicar sus métodos a problemas que pudiéramos llamar no clásicos (coloquialmente, problemas que no provienen de la física). En las páginas que siguen llamaré *biología matemática* a aquellas áreas de la biología experimental, de campo o teórica en donde se han aplicado de manera relevante métodos matemáticos. Dejo para otra ocasión la discusión de lo que pudiera ser la ciencia con el mismo nombre.

Hace 30 años, la biología matemática era apenas algo más que de interés tangencial para la mayoría de los biólogos, cualquiera que fuese su disciplina particular. Hoy, en muchas áreas, los desarrollos dictados por la biología matemática guían programas de investigación, experimentos y trabajos de campo y en muchas áreas la biología matemática está indisolublemente ligada a prácticas experimentales. En biología, desde mi perspectiva personal no hay leyes, al menos no las hay en el sentido entendido en la física y, quizás por ello, cada problema atacado desde una perspectiva matemática es un problema único de modelación, de construcción de modelos, de usos de marcos de referencia biológicos. En biología no existen equivalentes de la ecuación de Schroedinger ni de las ecuaciones de Newton de la mecánica celeste simple y llanamente porque en biología la historia del objeto de investigación importa mucho y no puede ser despreciada. Cada problema atacado es particular y casi único, y en consecuencia la existencia de leyes generales formalizadas en lenguaje matemático son prácticamente inexistentes. Siempre me ha sorprendido la forma en que las matemáticas y la biología se han integrado finalmente en este siglo pues sus formas de atacar problemas parecen, si no contradictorias, sí distintas: las matemáticas enfatizan y valoran el rigor y la precisión, sus resultados definitivos no cambian y se mantienen sólidos por miles de años: el ejemplo clásico son los *Elementos* de Euclides. Por el contrario, en biología nuestro conocimiento de los sistemas que estudiamos es muy reciente y las teorías biológicas evolucionan de manera rápida. Es quizás por esta diferencia que la aplicación de las matemáticas a la biología toma esencialmente dos caminos: por un lado la aplicación rutinaria de técnicas conocidas; por el otro, el desarrollo de nuevos métodos necesarios para el análisis de sistemas biológicos. En todo caso el problema central de la biología matemática es, en mi opinión, el siguiente: dado un problema

biológico ¿qué enfoque metodológico matemático o cuantitativo puede coadyuvar a su solución? Este planteamiento inmediatamente indica la necesidad del trabajo multi- e interdisciplinario, que integre especialistas de diferentes disciplinas biológicas y matemáticas que sean capaces de subordinar sus intereses disciplinarios naturales a la resolución de un problema común de naturaleza biológica.

En las páginas que siguen presentaré una imagen escueta y parcial de lo que ha sido la biomatemática en el siglo XX. La temática es amplísima y obviamente el espacio con el que cuento y el tiempo del que dispuse para escribir lo que el lector ahora tiene en sus manos son insuficientes para dar una visión panorámica general de esta nueva disciplina. Hay varias áreas importantes que no menciono, como las aplicaciones de teoría de nudos a la estructura tridimensional del ADN y de las proteínas, los avances impresionantes y maravillosos sobre máquinas moleculares, la F_0F_1 ATP sintasa o la V-ATPasa, por ejemplo; la computación con DNA, la fisiología del músculo cardiaco especialmente el estudio del origen de la fibrilación ventricular, lo que ha dado en llamarse complejidad, y otras muchas que menciono solo de pasada. En particular es innegable que la capacidad de autoorganización y la complejidad de los sistemas biológicos sean estos celulares, fisiológicos o ecológicos, constituyen características relevantes y definitorias de éstos y que estas mismas características hacen que la matemática de lo biológico sea la matemática de la dinámica no lineal. Afortunadamente hay en la literatura popular y especializada una abundancia de obras sobre estos temas que me permite en estas páginas concentrarme en otros aspectos menos populares pero de igual importancia que, por otro lado, ocurren en sistemas autoorganizados con un alto nivel de complejidad y, por supuesto, presentan dinámicas no lineales.

Tuve que tomar decisiones para poder condensar en un artículo de unas páginas lo que bien podría caber en un libro entero. Los temas elegidos reflejan mis gustos y sesgo. Invito al lector a explorar conmigo, esta perspectiva personal de la biología matemática.

2 La síntesis darwiniana.

Charles Darwin sintió siempre la necesidad de proporcionar una explicación causal para la evolución de las especies [24]. Darwin estudió en Cambridge en donde era casi obligado considerar que la grandeza

de Newton provenía de haber proporcionado un mecanismo causal al copernicanismo. Para Newton la clave fue la atracción gravitacional; para Darwin la selección natural sería el equivalente biológico de la gravedad, como una fuerza que dirigía y empujaba al mundo orgánico como la gravedad lo hacía con el mundo físico [24]. Darwin, aunque lejos de proporcionar un fundamento axiomático a su teoría, siempre buscó algo que confiriera a su teoría un carácter hipotético-deductivo, la marca de la buena ciencia en su época. El argumento de Darwin es como sigue: el crecimiento malthusiano de la población, es decir geométrico, entra en conflicto con el crecimiento aritmético de los recursos de que los organismos disponen. Esto conduce a la lucha por la existencia y a la selección natural como causa explicativa de la sobrevivencia diferencial de dichos organismos. Esta forma de pensar Darwiniana encontró su reivindicación muchos años después en las manos de los que podemos llamar titanes de la síntesis darwiniana. Uno de estos titanes fue Haldane. J.B.S. Haldane inicia así su artículo sobre la teoría matemática de la selección natural [6]:

“Una teoría satisfactoria de la selección natural debe ser cuantitativa. Para fincar la visión de que la selección natural es capaz de dar cuenta de los hechos conocidos de la evolución, debemos demostrar no únicamente que puede causar el cambio en las especies sino que puede causar ese cambio a una tasa que dé cuenta de las transmutaciones pasadas y presentes. En cada caso dado debemos especificar lo siguiente

1. El modo de heredabilidad del carácter considerado
2. El sistema de apareamiento en el grupo de organismos estudiado
3. La intensidad de la selección
4. Su incidencia (e.g. en ambos sexos o sólo en uno de ellos)
5. La tasa a la que la proporción de individuos que expresan el carácter se incrementa o disminuye.”

La conclusión de Haldane fue que debía entonces ser posible obtener una ecuación que conectara (3) y (5). Este artículo fue fundamental para el desarrollo de una teoría evolutiva cuantitativa que permitiera la comparación y verificación de evidencia factual, experimental o de campo, con las explicaciones y predicciones de la teoría.

La publicación de la obra de Charles Darwin *El Origen de las Especies* en 1859 [2] marcó un hito en el desarrollo de biología en general. No es este el lugar para hablar de la importancia de la teoría de la selección natural para explicar el proceso evolutivo. Lo que aquí nos interesa es señalar que por muchos años la teoría evolutiva fue cualitativa, fundamentalmente basada en argumentos verbales carentes de evidencia cuantitativa [22]. El descubrimiento de las llamadas leyes de Mendel permitió que una teoría evolutiva genética y cuantitativa se desarrollara, obteniendo su fundamento más temprano con los trabajos de R.A. Fisher quien, empezando 1918, propuso que las poblaciones biológicas en evolución podían ser siempre caracterizadas por la distribución estadística de los genes en la población y, en consecuencia, la historia evolutiva de una población sería el recuento de los cambios en las distribuciones de estos genes. De acuerdo a Provine [22] esta idea es quizás la más importante contribución de Fisher a la teoría evolutiva. La matemática que usó Fisher, quién tenía una formación sólida en matemáticas, física y astronomía, en estos trabajos, no fue del agrado de otro muy destacado evolucionista, Sewall Wright. Wright decidió construir su propia teoría evolutiva cuantitativa desarrollando su propio método, una forma de regresión lineal múltiple [22]. Su enfoque, publicado en 1931, fue sumamente importante para el estudio de la evolución en la naturaleza y para la teoría de cruzamiento en especies animales. El otro iniciador ya mencionado de la teoría evolutiva temprana fue J.B.S. Haldane que en una serie de artículos explícitamente intentó cuantificar las teorías de selección natural y artificial. Estos fueron los pioneros y gracias a ellos la genética de poblaciones comenzó a ser fuente de problemas interesantes para los matemáticos. No obstante esta riqueza, durante la primera mitad del siglo los matemáticos que se adentraron en su estudio no fueron muchos hasta que alrededor de 1950 la situación empezó a cambiar. Uno de los problemas más interesantes fue, y aún continua siendo, el de fundamentar qué tan apropiado es el uso de métodos de aproximación de difusiones en problemas de genética poblacional. William Feller fue uno de los fundadores en esta área. Feller estudió como el paso de una cadena de Markoff de los cambios en las frecuencias génicas, a un proceso de difusión podía hacerse de manera adecuada además de estudiar en forma matemática rigurosa las suposiciones subyacentes en modelos estocásticos importantes de genética de poblaciones. Finalmente Feller obtuvo los valores propios de una cadena de Markoff para el caso de una mutación reversible entre dos alelos [3].

La teoría moderna de genética poblacional se basa en el trabajo de los autores mencionados (y muchos más) que sentaron las bases metodológicas y teóricas del área. En la actualidad, los problemas de genética poblacional tienen que ver con aspectos de biología de la conservación y el uso de organismos transgénicos por ejemplo, que han resultado en la aparición de problemas matemáticos importantes que involucran loci múltiples y efectos del espacio, además del manejo de inmensas bases de datos genéticos.

En sistemas evolutivos la búsqueda de principios de optimización y su aplicación ha sido controversial debido a factores que tienen que ver con el papel de la frecuencia que diferentes genotipos tienen sobre la adecuación (fitness) relativa de los organismos. La alternativa exitosa a estos principios de optimización ha sido la de enfocar el problema de la evolución y adaptación de organismos como un problema en teoría de juegos [12]. Esta es un área de desarrollo sobre la que nos extenderemos a continuación. Maynard Smith desarrolló la idea de Estrategia Evolutiva Estable o ESS por sus siglas en inglés la cual se ha aplicado de manera muy importante al estudio del comportamiento altruista. La pregunta central aquí es cómo la cooperación entre individuos puede evolucionar cuando los procesos de selección natural actúan a nivel individual [16]. Considere el lector una población que se aparea aleatoriamente y que puede subdividirse en grupos de caracteres, cada grupo con n individuos. Suponga el lector que la selección natural ocurre mientras los organismos se encuentran en sus respectivos grupos de carácter. La adecuación de cada individuo depende de su propio genotipo y de los genotipos de los individuos pertenecientes a su mismo grupo. Una vez que la selección ha actuado, los individuos sobrevivientes se aparean al azar y generan sus gametos para la generación siguiente. Suponga el lector que existen solamente dos genotipos en la población, A y a . Los individuos A son altruistas en el sentido de que tienen algún comportamiento que disminuye su propia adecuación; es decir, ser altruista tiene un costo c para estos individuos. Por otro lado, el mismo comportamiento incrementa la adecuación de otros miembros de su grupo; es decir estos individuos obtienen un beneficio b debido al comportamiento altruista. Definamos W_0 como la adecuación en la ausencia de cualquier interacción. Definamos s como la adecuación adicional que un individuo cooperativo, es decir altruista, obtiene si el individuo a quien ayuda también coopera. La matriz de ganancia o beneficio es la siguiente:

		no cooperar	cooperar
no cooperar	W_0		$W_0 + b$
cooperar	$W_0 - c$		$W_0 + b - c + s$

La interpretación de esta matriz puede darse en dos formas: si los grupos que se encuentran interactuando consisten cada uno de un solo individuo, las entradas de la matriz representan las adecuaciones de un individuo adoptando alguna estrategia de la primera columna si su contrincante opta por alguna estrategia de la primera fila. Si los grupos de carácter están constituidos por más de un individuo, entonces las entradas de la matriz representan las adecuaciones de un individuo que adopta alguna estrategia de la primera columna *si todos los demás miembros de su grupo* adoptan estrategias en la primera fila

Consideremos ahora el siguiente ejemplo: $W_0 = 2$; $b = 2$; $c = 1$; $s = 0$. La matriz de ganancias es

		no cooperar	cooperar
no cooperar	2		4
cooperar	1		3

La matriz anterior representa el dilema del prisionero, pues independientemente de lo que el otro individuo haga, lo que conviene es no cooperar, por consiguiente, una decisión racional sería que *ambos* individuos eligieran no cooperar obteniendo un beneficio igual a 2. Sin embargo, si los individuos no tomaran decisiones racionales y eligieran cooperar su ganancia sería 3. He ahí el porqué de la paradoja. Lo de prisionero le viene porque se supone que ambos individuos son dos delincuentes en la cárcel: si ambos no confiesan salen libres, pues no hay pruebas en su contra; si ambos confiesan salen libres y además reciben un premio por ello. Si solo uno confiesa, ése sale libre, recibe un premio todavía mayor por su traición, y el otro se queda encerrado. Una forma de escapar de la paradoja es suponer que este juego es jugado repetidamente por ambos contendientes. Este nuevo juego se llama *del juego iterado del dilema del prisionero*. En este caso la estrategia de empezar cooperando y después contestar con la misma respuesta dada por el contrincante, es todavía mejor que la de no cooperar (en inglés esto se llama una estrategia tit-for-tat; en México diríamos ojo por ojo, diente por diente). La investigación en esta área incluye modelos con memoria es decir, se recuerdan comportamientos pasados y con efecto explícito del espacio [12].

3 Biología de poblaciones: Lotka y Volterra.

Empezaré esta sección citando a uno de los clásicos: Vito Volterra

“El Dr. Umberto D’Ancona me ha hablado muchas veces de su colección de datos sobre pesquerías durante la época de la guerra. Me ha preguntado si sería posible explicar matemáticamente sus resultados sobre las variaciones temporales en la composición de las especies. Su pregunta me llevó a plantear el problema en las páginas que siguen [...]. Ambos hemos trabajado independientemente y estamos complacidos con las coincidencias de nuestros resultados obtenidos por análisis y por observación, respectivamente. Al pescar, el hombre perturba el estado natural de la variación de la presa y del depredador, incrementando la cantidad de la especie depredada y decreciendo la del depredador...” (Vito Volterra, *Variazioni e fluttuazioni del numero d’individui in specie animali conviventi*, [29]).

En esta monografía Volterra incorpora y analiza problemas biológicos que más tarde se convertirían en fundamentales de la ecología teórica; por ejemplo, modelos con n especies en los que los coeficientes de las ecuaciones son dependientes de la densidad de población o funciones periódicas del tiempo, modelos depredador presa formulados en ecuaciones integrodiferenciales con efectos de retardo en las interacciones.

Las ecuaciones clásicas de depredador presa estudiadas por Volterra en la monografía citada son las que hoy aparecen en cualquier texto de ecología

$$\frac{dN_1}{dt} = N_1(\epsilon_1 - \gamma_1 N_2), \quad (1)$$

$$\frac{dN_2}{dt} = N_2(-\epsilon_2 + \gamma_2 N_1), \quad (2)$$

donde N_1 y N_2 son las densidades de ambas poblaciones y otros parámetros son constantes positivas. El razonamiento de Volterra es que si bien la ecuación (1) es correcta, la ecuación (2) es criticable pues la energía que usan los individuos depredadores para reproducirse y aumentar la densidad de la población no proviene instantáneamente de las presas capturadas sino de aquellas que fueron capturadas en ocasiones

anteriores. Las ecuaciones integrodiferenciales que Volterra obtiene al generalizar el modelo anterior son

$$\begin{aligned}\frac{dN_1}{dt} &= N_1(t) \left(\epsilon_1 - \gamma_1 N_2(t) - \int_{-\infty}^t F_1(t - \tau) N_2(\tau) d\tau \right), \\ \frac{dN_2}{dt} &= N_2(t) \left(-\epsilon_2 + \gamma_2 N_2(t) - \int_{-\infty}^t F_2(t - \tau) N_1(\tau) d\tau \right).\end{aligned}$$

Alfred James Lotka publicó su libro *Elements of Physical Biology* en 1925 [14] donde aplica principios de la física al estudio de sistemas biológicos y enfatiza el estudio de las relaciones entre los organismos desde un punto de vista energético. De acuerdo con Kingsland [11], el trabajo de Lotka se adelantó y anticipó las ideas que Ludwig von Bertalanfy desarrollaría 25 años más tarde sobre la teoría general de sistemas. Sin embargo, de acuerdo a la misma autora, Lotka se mostró decepcionado de que los biólogos, y no los físicos, fueran quienes consideraban las ideas de su libro interesantes. Azares del destino. El libro de Lotka estaba destinado a convertirse en uno de los pilares de una de las subdisciplinas biológicas más activas en investigación teórica usando modelos matemáticos: la ecología de poblaciones. Aunque Vito Volterra y Alfred Lotka coincidieron en muchos de sus resultados, publicados casi simultáneamente, existen sin embargo importantes diferencias en sus enfoques. El enfoque de Lotka siempre estuvo comprometido con los datos que podían validar sus hipótesis y aunque desarrolló y planteó modelos de gran generalidad siempre estuvo interesado en predicciones cuantitativas de sus resultados; por el contrario Volterra se centró en explicaciones generales de carácter cualitativo y generalizaciones matemáticas [25]. Sin embargo, y como ha ocurrido muchas veces en la historia de la ciencia, la obra escrita de Volterra refleja su formación profesional de matemático pero no su enfoque particular al estudio de los sistemas ecológicos. Vito Volterra estuvo muy asociado con entomólogos, ictiólogos y otros biocientíficos, y siempre se mantuvo al tanto de sus ideas y resultados experimentales, tanto así que intentó adecuar sus ecuaciones a diferentes procesos ecológicos prácticos. Fue su yerno, Umberto D'Ancona, el biólogo, quien prefería el valor teórico del trabajo de Volterra.

El enfoque de Lotka paradójicamente, precisamente por buscar generalidad teórica no necesariamente matemática en su ataque a problemas biológicos, sirvió de punto de partida para que A. Kolmogoroff

usara la forma general de los modelos de Lotka

$$\begin{aligned}\frac{dx}{dt} &= f(x, y)x, \\ \frac{dy}{dt} &= g(x, y)y,\end{aligned}$$

para, especificando condiciones cualitativas generales sobre las funciones f y g , demostrar los posibles comportamientos cualitativos de las soluciones del sistema de ecuaciones que incluyen nodos, focos y ciclos límite.

En esta área de biología de poblaciones, donde Lotka y Volterra hicieron sus contribuciones más importantes, el estudio de modelos sencillos constituye un claro ejemplo de la influencia positiva mutua que puede existir entre matemáticas y biología. La investigación de estos modelos simples mostró la existencia de sucesiones de bifurcaciones universales en las iteraciones de mapeos unidimensionales (e.g. bifurcaciones sucesivas con doblamiento de período).

En ecología el verdadero renacimiento de las matemáticas se llevó a cabo al final de la década de los 50s con Robert H. McArthur. McArthur tuvo una formación de matemático pero su tesis doctoral la hizo en el estudio de nicho ecológico en ciertas especies de aves muy cercanamente relacionadas. McArthur desarrolló posteriormente muchas áreas de trabajo: le dió nuevo impulso a la teoría de depredador-presa, de competencia por recursos, a la teoría del nicho y de la similitud limitante; la teoría de biogeografía de islas también tuvo un importante impulso bajo McArthur. No todos sus intentos fueron exitosos y algunas de sus ideas, como su teoría de similitud limitante y teoría de nicho, sufrieron serias críticas. Sin embargo el ímpetu general que este hombre tuvo sobre la teoría fue muy importante. A su muerte, muchos de sus colaboradores y exalumnos desarrollaron muy importantemente las investigaciones de McArthur. Entre ellos el más destacado es quizás Robert M. May quien en 1974, dos años después de la muerte de McArthur, escribe una obra fundamental para el desarrollo de la ecología teórica: *Stability and Complexity in Model Ecosystems* [15]. Esta obra es un compendio general de lo que en esos años se tenía en teoría ecológica fundamentalmente en referencia a sistemas competitivos de n especies, sistemas depredador presa y respuestas funcionales, efecto del ambiente variable y alguna discusión sobre modelos estocásticos dinámicos. El grueso del libro fundamenta sus resultados usando la herramienta de las ecuaciones diferenciales ordinarias. Poco tiempo después, en 1976

el mismo May estudiando la ecuación logística discreta

$$x_{t+1} = rx_t(1 - x_t/K),$$

donde r es la tasa intrínseca de crecimiento y K es la capacidad de carga del ambiente, encuentra que esta sencilla relación de recurrencia es capaz de una gama increíblemente compleja de comportamientos. El impacto que este descubrimiento tuvo en ecología es comparable con aquel que tuvo en la hoy llamada teoría del caos.

En años recientes, debido a la disponibilidad de computadoras de gran potencial numérico, estos grandes temas de la ecología, depredación, competencia, etc, están siendo revisados al introducir el componente espacial explícito dentro de los modelos. El espacio es un invariante general en las interacciones ecológicas y forma parte de un tema de estudio muy importante en la biología matemática moderna: el efecto de la heterogeneidad, ya sea espacial, fisiológica, ambiental, etc., en las interacciones y evolución de los sistemas biológicos. La inclusión del espacio en ecología ha promovido la aplicación de técnicas estocásticas, autómatas celulares, sistemas de partículas interactuantes, ecuaciones diferenciales parciales y estocásticas, procesos de nacimiento y muerte, a sistemas con una gran cantidad de componentes, impensables en la época de Lotka y Volterra.

4 Kermack y McKendrick: la epidemiología matemática.

La aplicación de las matemáticas a la epidemiología puede trazarse al menos hasta el año de 1760 cuando Daniel Bernoulli publicó un pequeño tratado sobre la epidemia de peste que en ese entonces se abatía sobre Europa. En este siglo que termina, el interés por la aplicación de métodos cuantitativos a la biología floreció como consecuencia del éxito de éstos en la física y en particular en biofísica y bioquímica. Una observación temprana que fue rápidamente puesta en términos cuantitativos fue la de que las enfermedades infecciosas se transmiten por contacto entre un individuo susceptible y uno enfermo infeccioso. Hamer en 1906 formuló la ley de acción de masas que establece que el número de contactos infecciosos, es decir que producen enfermedad, por unidad de tiempo es proporcional al número total de contactos entre individuos infecciosos y sanos. Después Ronald Ross en el apéndice a su libro *The Prevention of Malaria* en la edición de 1911, formuló

un modelo matemático sencillo como apoyo a su argumentación de que para erradicar el paludismo era suficiente con disminuir la población de mosquitos a un nivel bajo, sin necesariamente extinguirlos. Más tarde, en 1927 Kermack y McKendrick formularon un modelo matemático bastante general y elaborado para describir la epidemia de peste que sufriera la India en 1906.

Ronald Ross ganó en 1902 el premio Nobel por sus descubrimientos sobre la etiología del paludismo. Sus estudios le llevaron a plantear la necesidad de reducir la población de mosquitos *Anopheles*, transmisores del paludismo, para controlar la enfermedad. En un ejemplo más de la importancia de las creencias, de la visión del mundo que se tiene en el desarrollo de la ciencia esta idea de Ross, tan natural ahora, fue rechazada porque la visión dominante en aquellos tiempos era que la incidencia de la enfermedad no estaba correlacionada con la densidad de mosquitos en el área [11]. Para argumentar su punto Ross desarrolló lo que él llamó *método a priori* que consistía en, primero, postular hipótesis sobre las causas de un brote epidémico; luego construir un conjunto de ecuaciones diferenciales para describir la situación postulada; después deducir de los resultados del análisis de las ecuaciones diferenciales las implicaciones epidemiológicas y, finalmente, comparar estas predicciones con los datos y observaciones disponibles. Como puede verse este es el programa normal de construcción de un modelo matemático, novedoso en epidemiología en aquella época en la cual la estadística era reina, caracterizado por un método distinto donde, a partir de datos y observaciones, se ajustan curvas y distribuciones.

No es este el lugar para relatar la historia de la epidemiología matemática. Permítame el lector brincarme casi dos décadas para describir el trabajo, publicado en 1927, de W.O. Kermack y A.G. McKendrick, ambos del Laboratory of the Royal College of Physicians en Edinburgo, quienes desarrollaron las ideas de Ronald Ross. En particular estos dos investigadores estaban interesados en explicar la típica curva epidémica acampanada.

El modelo general [10] es construido dividiendo el tiempo en un conjunto de intervalos de tamaño igual y constante, donde los contagios ocurren únicamente en los instantes cuando se pasa de un intervalo a otro. Tomando el tamaño de estos intervalos como la unidad de tiempo, el número de individuos infectados al tiempo t que han estado infectados por θ intervalos de tiempo es $v_t(\theta)$. Note el lector que θ es lo que llamamos ahora edad de la infección. El número total de enfermos al tiempo t es $y_t = \sum_{\theta=0}^t v_t(\theta)$. Si definimos $v_t = v_t(0)$, como el número de individuos que se infectan por primera vez durante la transición del

intervalo $t - 1$ al intervalo t , imponiendo las condiciones iniciales

$$v_0(0) = v_0 + y_0,$$

y denotando por $\psi(\theta)$ la tasa de curación o muerte de los enfermos con edad de infección θ , entonces el número de individuos con edad de infección θ que dejan de ser infecciosos al final del intervalo t , es $\psi(\theta)v_t(\theta)$; por consiguiente

$$\begin{aligned} v_t(\theta) &= v_{t-1}(\theta - 1) [1 - \psi(\theta - 1)] \\ &= v_{t-2}(\theta - 2) [1 - \psi(\theta - 1)] [1 - \psi(\theta - 2)] \\ &= v_{t-\theta} B(\theta), \end{aligned}$$

donde

$$B(\theta) = [1 - \psi(\theta - 1)] [1 - \psi(\theta - 2)] \cdots [1 - \psi(0)].$$

Como el número de personas por unidad de área que se contagian en el intervalo t satisface $v_t = x_t \sum_{\theta=1}^t \phi(\theta)v_t(\theta)$, donde x_t es el número de individuos susceptibles y $\phi(\theta)$ es la tasa de infección cuando la edad de la infección es θ , entonces podemos escribir

$$x_t = N - \sum_{\theta=0}^t v_t(\theta),$$

con N el número total de individuos en la población. Kermack y McKendrick definen después z_t como el número de individuos que han sido removidos, por inmunidad o muerte, de la dinámica epidemiológica, para obtener que

$$N = x_t + y_t + z_t$$

para todo t . Definiendo $A(\theta) = \phi(\theta)B(\theta)$ y $C(\theta) = \psi(\theta)B(\theta)$, $B(\theta) = \exp\left(-\int_0^\theta \psi(s)ds\right)$, obtienen el sistema de ecuaciones integrodiferenciales

$$\begin{aligned} \frac{dx_t}{dt} &= -x_t \left(\int_0^t A(\theta)v_{t-\theta}d\theta + A(t)y_0 \right), \\ y_t &= \int_0^t B(\theta)v_{t-\theta}d\theta + B(t)y_0, \\ \frac{dz_t}{dt} &= \int_0^t C(\theta)v_{t-\theta}d\theta + C(t)y_0. \end{aligned}$$

La forma más conocida del sistema de ecuaciones anterior es el que se obtiene considerando $\phi \equiv k$, $\psi \equiv \gamma$ constantes,

$$\begin{aligned}\frac{dx_t}{dt} &= -kx_t y_t, \\ \frac{dy_t}{dt} &= kx_t y_t - \gamma y_t, \\ \frac{dz_t}{dt} &= \gamma y_t.\end{aligned}$$

La generalidad del modelo original de Kermack y McKendrick y el poder explicativo de su simplificación más conocida fueron durante todo el siglo XX pilares fundamentales para el desarrollo de la llamada epidemiología teórica y matemática. Más aún, el estudio matemático de la dinámica poblacional de enfermedades infecciosas con el advenimiento del SIDA, ha estado fundamentalmente basado en el modelo que Kermack y McKendrick publicaron en 1927.

Largo ha sido el camino recorrido desde los trabajos pioneros de Ross y Kermack y McKendrick. Un problema fundamental en la epidemiología actual es cómo la transmisibilidad de un patógeno depende del espacio social o geográfico y cómo incorpora y con cuánto detalle estos factores espaciales en las tasas de contacto de los modelos que representan estos procesos. La naturaleza compleja de estos sistemas hace necesaria la aplicación de modelos estocásticos a su estudio particularmente en aspectos prácticos y de gran importancia en epidemiología: diseño de estrategias de vacunación. La pregunta teórica de interés aquí es qué tanta agregación espacial debe incorporarse y que tan simple debe ser un modelo para que pueda proporcionar resultados de aplicabilidad práctica e inmediata. El problema no es trivial de ninguna manera.

Una de las políticas de salud pública más importantes es la del diseño de programas de vacunación. Las enfermedades infantiles son los ejemplos típicos en los que las vacunas son altamente eficientes para prevenir brotes epidémicos. Existen otras como la tuberculosis o el SIDA para las que existen o se están investigando vacunas. En todo caso, el objetivo de una vacuna es, además de proporcionar inmunidad a la persona que la recibe, la de proteger a la población ya sea previniendo brotes epidémicos o erradicando la enfermedad. En todo caso, el resultado inmediato de una campaña de vacunación efectiva es la reducción de la prevalencia o de la incidencia de la enfermedad. Desde el punto de vista de la epidemiología matemática lo que la vacunación de una

población susceptible logra es reducir el valor del número reproductivo básico R_0 , definido como el número de casos secundarios producidos por un enfermo primario típico en una población susceptible, por debajo de un umbral dado, y evitar que rebase este umbral. Toda campaña de vacunación tiene costos asociados que están, por ejemplo, en función de la edad del segmento de la población a la que la campaña está dirigida (llamada de pacientes, propaganda, equipo, personal, etc.) y del hecho de que la eficacia de la vacuna puede ser baja o puede aplicarse a personas que ya estén inmunizadas o están enfermas. El problema sobre el que nos centraremos ahora será el del diseño de programas de vacunación eficientes.

Para lograrlo tenemos que caracterizar y definir varias cosas. Primeramente debemos dar una definición adecuada de un número reproductivo en presencia de vacunación; seguidamente será necesario cuantificar el decremento en el número reproductivo logrado por la estrategia de vacunación y, finalmente, habrá que asociar la estrategia de vacunación a sus costos que, en el ejemplo que trataremos en esta sección, están asociados a la edad de los individuos.

Esta estructura no es compatible con enfermedades infecciosas venéreas de transmisión heterosexual, ni para la poliomielitis ni enfermedades macroparasíticas, sin embargo es todavía lo bastante general como para ilustrar el enfoque que nos interesa. Primeramente definiremos el número reproductivo de vacunación como $R(\varphi)$ donde $\varphi \geq 0$ es la tasa a la que individuos susceptibles son vacunados. Este nuevo parámetro satisface

$$R(0) = R_0 \text{ y } R(\varphi) \leq R_0 \text{ si } \varphi \geq 0$$

es decir, que en ausencia de vacunación este nuevo parámetro es exactamente el número reproductivo de la enfermedad y que en presencia de cualquier tipo de estrategia de vacunación, siempre se obtiene un parámetro umbral menor que el valor del número reproductivo básico. En otras palabras, la vacuna en el peor de los casos deja todo como si no se hubiera aplicado. Karl Hadeler de la Universidad de Tubingen encontró, en 1993, que el número reproductivo de vacunación puede descomponerse en dos sumandos, uno que es exactamente igual al número reproductivo básico, y otro que incorpora el efecto de la vacunación en la reducción de éste último

$$R(\varphi) = R_0 - \frac{F(\varphi)}{N},$$

donde N es la población total y F mide la reducción en el número reproductivo básico que puede alcanzarse usando la estrategia de vacunación

φ . Además este término es función de la diferencia en las tasas de infección de individuos vacunados y no vacunados. Aclaremos brevemente el comentario anterior. Una vacuna tiene un parámetro llamado eficacia que mide el porcentaje de individuos que aun siendo vacunados, todavía pueden contraer la enfermedad. Una buena vacuna tiene una alta eficacia, es decir, el porcentaje de gente desprotegida aunque vacunada es muy pequeño, y aún esa gente desprotegida es menos susceptible a enfermarse que una gente que no ha sido vacunada. El objetivo entonces de una estrategia de vacunación es la de incrementar F para disminuir $R(\varphi)$.

Para modelar esta situación deberemos, como hemos hecho en cada uno de los ejemplos discutidos, proponer varias simplificaciones. Las que haremos ahora son las siguientes. Supondremos que la población se encuentra en equilibrio demográfico con una estructura de edades estacionaria. Definiremos además un parámetro umbral de referencia que será denotado por R que es el número reproductivo de vacunación cuando toda la población, sin excepción alguna, ha sido vacunada. Supondremos además que los costos totales de la estrategia de vacunación dependen linealmente del número de vacunaciones efectuadas. Estos costos totales serán representados por $C(\varphi)$.

Hadeler diferenció dos casos. El primero ocurre cuando solamente se vacuna individuos susceptibles y el segundo cuando se vacuna cualquier individuo independientemente de su estado inmunológico. Hadeler definió además esencialmente dos posibles escenarios de optimización: i) lograr una meta con costo mínimo (por ejemplo reducir $R(\varphi)$ hasta un valor R^* minimizando $C(\varphi)$); ii) lograr un óptimo con gasto limitado (por ejemplo, minimizar $R(\varphi)$ sujeto a la condición $C(\varphi) \leq c$). En ambos casos, la diferencia estribando en la dificultad técnica del análisis, la estrategia óptima de vacunación resulta una de las dos siguientes:

- vacunar a toda la población exactamente a la edad A,
- vacunar parte de la población a la edad A y el resto a la edad B.

Hadeler proporciona los valores de las edades A y B en general. Es interesante señalar que el problema de vacunación en edades óptimas rinde el mismo tipo de resultados que el problema de determinar las edades óptimas de pesca en la explotación de recursos pesqueros. De hecho, la estructura formal de los problemas de explotación y manejo de recursos son muy similares en estructura matemática a los de vacunación epidemiológica y dinámica metapoblacional en ecología.

A partir de 1976 con el inicio de la pandemia del SIDA, la epidemiología matemática, que hasta esas fechas había jugado un papel marginal en el desarrollo de la biología, cobró un ímpetu sin precedentes [1]. Aparte de los trabajos de Ross, Kermak y McKendrick y Lotka sobre paludismo y peste, otros investigadores habían estudiado la dinámica poblacional de enfermedades infecciosas ya sea transmitidas por organismos vectores o directamente transmisibles. Es importante señalar que es posible clasificar a las enfermedades infecciosas, esencialmente desde el punto de vista de las escalas temporales a las que ocurre la transmisión, como macro y microparasíticas. Las enfermedades microparasíticas son las producidas por microorganismos como bacterias, virus y protozoarios: enfermedades macroparasíticas encuentran su ejemplo más importante en las infecciones por lombrices intestinales. Con el advenimiento del SIDA todos estos resultados teóricos comenzaron a desarrollarse de manera vertiginosa hasta llegar al nivel del día de hoy, cuando la epidemiología matemática es un área con gran productividad y que ha contribuido de manera muy importante al desarrollo de la ecología teórica. Los problemas fundamentales a resolver, además de los grandes problemas teóricos mencionados en un párrafo anterior, son en nuestros países el control de la transmisión de enfermedades infecciosas como el paludismo, el dengue, la enfermedad de Chagas, la oncocercosis, la leishmaniasis, etc, que no existen y por lo tanto no constituyen problemas relevantes en los países que son la meca de la investigación en el área. Este es el gran reto para los epidemiólogos matemáticos de nuestros países, desde mi punto de vista.

5 Morfogénesis: Turing.

En el área del estudio de la forma, en la cual el trabajo de Alan Turing, Rene Thom, James Murray y tantos otros puede inscribirse, resalta una figura pionera: D'Arcy W. Thompson. En su libro *Sobre el crecimiento y la forma* [27], este autor aplica transformaciones geométricas al estudio de la forma y estructura de tejidos, espículas, esqueletos, cuernos, conchas, hojas, huevos, etc. Escribe Thompson con un lenguaje casi poético en su libro citado [capítulo IX, segunda sección: Matemáticas y Forma] :

“Podría pensarse que las definiciones matemáticas son demasiado estrictas y rígidas para el uso corriente, pero su rigor está combinado con una libertad casi infinita. La definición exacta de una elipse nos introduce a todas las elipses del

mundo, la definición de una sección cónica amplía nuestro concepto y una curva de orden superior aumenta aún más nuestro campo de libertad... Por una razón o por otra, existen muchas formas orgánicas que no podemos describir, y menos aún definir en términos matemáticos... Ni siquiera nos molestamos en buscar una fórmula que defina este o aquel pez, o un determinado cráneo de un vertebrado. Pero ya podemos utilizar el lenguaje matemático para describir e incluso definir en términos generales, la concha de un caracol, el retorcimiento de un cuerno, el contorno de una hoja, la textura de un hueso, la estructura de un esqueleto, las líneas de corriente de un pez o un ave, el fino encaje de las alas de un insecto. Para hacerlo debemos aprender de los matemáticos a eliminar y descartar; mantener en la mente el tipo y aislar el caso particular con todos sus accidentes; y a encontrar en este sacrificio de lo superfluo y conservación de lo importante una de las excelencias del método matemático”.

D’Arcy Thompson reconoció en cada forma biológica específica la deformación o transformación de otra debida a cambios de sistemas coordenados. Supongamos que dibujamos el perfil de un organismo cualquiera en el plano cartesiano. Si este sistema de referencia lo deformamos alterando la dirección de los ejes, cambiamos ángulos rectos en agudos, cambiando la razón de la escala entre x y y , o realizamos transformaciones más complicadas, es posible obtener, por ejemplo, de manera sucesiva los anfípodos: *Harpinia plumosa*, *Stegocephalus inflatus* e *Hyperia galba*.

En 1936 Joseph Needham escribía [19]:

“Es asombroso que [D’Arcy Thompson] haya tenido tan pocos seguidores [...] nadie ha extendido su trabajo a la embriología. La embriología tiene una gran ventaja [...] que ofrece la posibilidad de una correlación cercana con estudios fisiológicos y bioquímicos sobre el desarrollo. Sin embargo, esta notable oportunidad para la comparación entre datos morfológicos y físico-químicos no haya sido aprovechada por nadie”.

Pasaron todavía 20 años para que alguien, un brillante matemático que trabajó en la guerra descifrando códigos secretos alemanes, tomara el

desafío. En 1952 Alan Turing publicó el artículo, ahora clásico, *The chemical basis of morphogenesis* [28] en el que propuso que especies químicas llamadas por él morfógenos pueden reaccionar y difundirse de tal forma que inducen la formación de patrones espaciales en la concentración de los morfógenos. De acuerdo a James Murray [18]

“...Lo que es impresionante acerca del artículo seminal de Turing es que, con muy pocas excepciones, le llevó al mundo matemático más de 20 años darse cuenta de la riqueza en problemas fascinantes presentes en su teoría. Lo que es aún más sorprendente es que pasaron casi 30 años para que un número significativo de biólogos experimentales tomaran en serio sus implicaciones y su potencial aplicativo en biología del desarrollo, ecología y epidemiología.”

Como puede verse en el desarrollo histórico del estudio matemático de la morfogénesis, ha sido considerado como sorprendente que tanto conocimiento experimental e intentos importantes de modelación de la forma, no hayan sido vistos por los matemáticos aplicados ni biólogos teóricos como una oportunidad sin igual para trabajo interesante.

El artículo de Turing es realmente bueno y es un placer leerlo. Turing inicia su modelo de morfogénesis con el caso de un anillo de N células idénticas con únicamente dos morfógenos denotados por X e Y respectivamente. La concentración del morfógeno X en la célula r es denotada por X_r ; análogamente para Y . En este anillo de células cada una de ellas intercambia materiales por difusión con sus dos células adyacentes a una tasa de difusión μ y ν respectivamente. Las tasas de reacción química de ambos morfógenos son $f(X, Y)$ para el morfógeno X , y $g(X, Y)$ para el morfógeno Y . El sistema de ecuaciones diferenciales (la discretización espacial de una ecuación de reacción difusión) que describe este proceso es

$$\begin{aligned}\frac{dX_r}{dt} &= f(X_r, Y_r) + \mu(X_{r+1} - 2X_r + X_{r-1}), \\ \frac{dY_r}{dt} &= g(X_r, Y_r) + \nu(Y_{r+1} - 2Y_r + Y_{r-1}).\end{aligned}$$

$$r = 1, \dots, N.$$

Turing supone que el sistema anterior tiene un estado de equilibrio homogéneo con $X = X^*$ y $Y = Y^*$ tal que $f(X^*, Y^*) = g(X^*, Y^*) = 0$. Suponiendo que el sistema no se encuentra demasiado lejos del punto

de equilibrio Turing obtiene las soluciones del sistema linealizado que le permiten deducir las condiciones que generan inestabilidad espacial.

En el caso en el que uno quiera considerar un anillo continuo de tejido, el sistema de ecuaciones que se obtiene es de la forma

$$\begin{aligned}\frac{\partial X}{\partial t} &= f(X, Y) + D_1 \frac{\partial^2 X}{\partial x^2}, \\ \frac{\partial Y}{\partial t} &= g(X, Y) + D_2 \frac{\partial^2 Y}{\partial x^2},\end{aligned}$$

donde D_1, D_2 con los coeficientes de difusión y x es el ángulo que describe la posición de un punto en el espacio. Este sistema puede verse como un caso límite del modelo discreto que Turing describe y estudia ampliamente en el artículo citado. Turing, en el análisis que desarrolla en su trabajo, observó que si en la ausencia de difusión ($D_1 = D_2 = 0$), la concentración de morfógenos tiende a un estado estacionario uniforme linealmente estable es posible que se generen patrones espacialmente heterogéneos debido a que la difusión desestabiliza el estado estacionario. Este resultado es muy importante puesto que generalmente, la difusión es un proceso estabilizador. En resumen diremos que un sistema de reacción-difusión exhibe inestabilidad dirigida por la difusión o inestabilidad de Turing si el estado estacionario homogéneo es estable ante perturbaciones pequeñas en la ausencia de difusión, pero inestable ante pequeñas perturbaciones espaciales cuando la difusión no es cero.

A continuación describiremos de manera muy escueta las condiciones generales para la existencia de inestabilidad de Turing. El lector interesado en los detalles deberá consultar, por ejemplo, Murray [17]. El modelo general en forma no dimensional de reacción difusión con condiciones a la frontera de no flujo, pues estamos interesados en la *autoorganización* de patrones sin efectos externos, es

$$\begin{aligned}u_t &= \gamma f(u, v) + \nabla^2 u, \\ v_t &= \gamma g(u, v) + \delta \nabla^2 v, \\ (n \cdot \nabla) \begin{pmatrix} u \\ v \end{pmatrix} &= 0, \quad r \text{ en } \partial B; \quad u(r, 0), v(r, 0) \text{ dados,}\end{aligned}$$

donde ∂B es la frontera cerrada del dominio B donde ocurre nuestro proceso de reacción difusión; n es un vector normal a ∂B ; u y v representan las concentraciones reescaladas de los morfógenos X y Y respectivamente. El estado estacionario homogéneo es la solución positiva del sistema

$$f(u^*, v^*) = g(u^*, v^*) = 0.$$

En ausencia de variación espacial se debe resolver un sistema de ecuaciones diferenciales

$$\begin{aligned} u_t &= \gamma f(u, v), \\ v_t &= \gamma g(u, v). \end{aligned}$$

de cuyo análisis, una vez linealizado, se pueden obtener condiciones para la estabilidad local del punto de equilibrio asociado con un espacio homogéneo. Denotemos por λ a los valores propios de la matriz jacobiana del sistema. Una vez realizado este análisis consideremos el sistema linealizado completo, es decir, cuando la difusión está presente que puede escribirse como

$$w_t = \gamma Aw + D\nabla^2 w, \quad (3)$$

donde A es una matriz constante y D es una matriz diagonal positiva.

Denotemos por $W(r)$ a la solución del problema de valores propios

$$\nabla^2 W + k^2 W = 0, \quad (n \cdot \nabla) W = 0 \text{ para } r \text{ en } \partial B.$$

$W(r)$ es la eigenfunción correspondiente al número de onda k . Las soluciones de (3) tienen la forma

$$w(r, t) = \sum_k s_k e^{\lambda t} W_k(r),$$

donde las constantes s_k se determinan por la expansión de Fourier de las condiciones iniciales en términos de $W_k(r)$. Requiriendo soluciones no triviales de $W_k(r)$ se obtiene que los valores propios λ son función de los números de onda k ; es decir $\lambda = \lambda(k)$. Para que el estado estacionario homogéneo sea inestable ante perturbaciones espaciales es necesario que $\text{Re}\lambda(k) > 0$ para alguna $k \neq 0$. La relación entre λ y k se conoce como relación de dispersión y en general depende del valor de $\delta = D_1/D_2$. En la gráfica (tomada de Murray, [17]) se representa típicas relaciones de dispersión para diversos valores de δ .

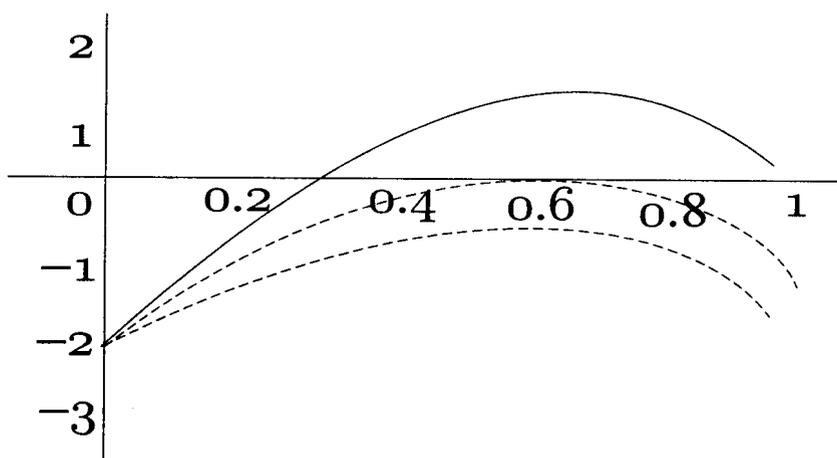


Gráfico k^2 versus $\text{Re } \lambda(k)$ para tres valores de δ .

La línea continua representa valores de k^2 que implican a la existencia de eigenfunciones inestables que evolucionan hacia estados estacionarios no homogéneos. Esta línea corresponde a un valor del coeficiente de difusión $\delta > \delta_c$ crítico (línea a trazos). Para $\delta < \delta_c$, $\text{Re } \lambda(k)$ es siempre negativa para cualquier valor de k^2 y, en consecuencia, no se producen patrones espaciales no homogéneos. La confirmación y aplicación de estos resultados teóricos a la embriología y al estudio de patrones de la piel en animales y plantas ha sido uno de los éxitos de la biología matemática contemporánea.

He reseñado con algún detalle el trabajo de Turing pues, en mi opinión, es un bello ejemplo de la manera de pensar con matemáticas sobre sistemas biológicos. Turing no solo reencontró el estudio matemático de la forma y su generación en biología, sino que el tiempo demostró también que abrió camino para la interacción real y continua entre los embriólogos y sus resultados experimentales y la matemática aplicada. Las ecuaciones de reacción difusión y otras estrechamente relacionadas con ellas, se han utilizado para modelar efectos quimiotácticos celulares, es decir, procesos de "atracción" en vez de difusión. La quimiotaxis es un fenómeno muy común en la naturaleza que se observa desde células hasta organismos multicelulares. Otras aplicaciones relevantes de este tipo de estructuras matemáticas ocurren en el estudio de tumores y cánceres y en el estudio de los patrones de manchas en la piel de diferentes organismos. En particular este último aspecto tiene muy importantes consecuencias para la teoría biológica. Maynard Smith, otra figura central en el desarrollo de la biología matemática y teórica de este siglo, ha dicho que el objetivo central de cualquier teoría de la evolución biológica es el de dar cuenta de las adaptaciones complejas de

los organismos a su ambiente. No voy a meterme aquí a discutir el serio y polémico problema de definir que es una adaptación ni a caracterizar la complejidad. Baste decir que los sistemas biológicos son sistemas lejos del equilibrio que tienen la capacidad de autoorganizarse. La formación de patrones en la piel de los animales, por ejemplo, muchas veces constituye adaptaciones complejas que afectan significativamente la adecuación del organismo.

No puedo concluir esta sección sin mencionar el trabajo de René Thom [26] que causó gran revuelo con su teoría de las catástrofes. Thom, basando sus estudios en las teorías de C. Waddington, desarrolló su teoría matemática de la morfogénesis y la generalizó bajo el nombre ya mencionado. Las aplicaciones de la teoría de las catástrofes a diversas áreas de la ciencia fueron poco aceptadas puesto que muchas veces forzaban la realidad a satisfacer los postulados "catastróficos" de la teoría de Thom. Sin embargo, desde el punto de vista matemático, la teoría de Thom fue y sigue siendo un área importante de investigación geométrica y topológica. Su impacto sobre la biología matemática fue sin embargo poco relevante.

6 Hodgkin y Huxley: ritmos biológicos.

El tiempo es esencial para los organismos. Podría decirse que la vida es rítmica y que sin ritmos la vida cesa. ¿Cómo se generan los ritmos vitales? Winfree [31] ha escrito una obra fundamental al respecto de la que pronto gozaremos de una nueva edición. El tema es amplísimo y fascinante. Algunos de los ritmos son generados por células aisladas o por células interconectadas que pueden generar algún tipo de actividad oscilatoria. Células con esta capacidad se denominan marcapasos. A continuación reseñaremos uno de los más importantes resultados sobre estos.

En el año de 1952 se publicaron una serie de artículos en el *Journal of Physiology* en los que se determinaba la dinámica de las conductancias iónicas que generan el potencial de acción. Los artículos estaban firmados por Alan L. Hodgkin y Andrew F. Huxley [8]. En 1963, gracias a este trabajo, se les otorgó el Premio Nobel en Fisiología y Medicina. Hodgkin y Huxley desarrollaron un modelo matemático explícito basado en sus descubrimientos e hipótesis; el modelo fue ajustado a los datos experimentales de conductancia generados por ellos mismos y así predijeron la evolución temporal del potencial de membrana durante un impulso nervioso [23]; [9]. Hoy sabemos que el valor del potencial

de membrana se encuentra entre -70 mV en estado no excitado, y $+40$ mV en el pico del potencial de acción. El éxito de Hodgkin y Huxley en la medición del potencial de membrana se debió al desarrollo del llamado *voltage clamp*. Esta técnica fija el potencial de membrana y luego mide la corriente que debe ser suministrada para mantener un voltaje constante. Sin entrar en detalles señalaremos que Hodgkin y Huxley supusieron que la corriente inicial de entrada es debida fundamentalmente a iones Na^+ , mientras que la corriente de salida es debida a iones K^+ . Con estas suposiciones Hodgkin y Huxley fueron capaces de separar la corriente iónica total en sus partes iónicas constitutivas.

Los detalles completos de la derivación de las ecuaciones diferenciales que pondremos a continuación los encontrará el lector en Keener y Sneyd [9]. Aquí nos limitaremos únicamente a listar el sistema y comentar algunos de sus desarrollos importantes. Las ecuaciones de Hodgkin y Huxley son .

$$\begin{aligned} C_m \frac{dv}{dt} &= -\bar{g}_K n^4 (v - v_K) - \bar{g}_{\text{Na}} m^3 h (v - v_{\text{Na}}) - \bar{g}_L (v - v_L) + I_{\text{app}}, \\ \frac{dm}{dt} &= \alpha_m (1 - m) - \beta_m m, \\ \frac{dn}{dt} &= \alpha_n (1 - n) - \beta_n n, \\ \frac{dh}{dt} &= \alpha_h (1 - h) - \beta_h h, \end{aligned}$$

donde $\alpha_m, \beta_m, \alpha_h, \beta_h, \alpha_n$ y β_n son funciones no lineales del voltaje. Además

$$\begin{aligned} \bar{g}_{\text{Na}} &= 120, \bar{g}_K = 36, \bar{g}_L = 0.3, \\ v_{\text{Na}} &= 115, v_K = -12, v_L = 10.6 \end{aligned}$$

El potencial v es la desviación del estado de reposo $V = V_r + v$ y está medido en mV, la densidad de corriente tiene unidades de $\mu\text{A}/\text{cm}^2$, conductancias tienen unidades de mS/cm^2 y las capacitancias unidades de $\mu\text{F}/\text{cm}^2$. Es muy importante señalar que el trabajo de estos dos autores es un ejemplo relevante de una investigación en donde métodos experimentales, métodos matemáticos, ajuste de curvas y teoría biológica se combinan para resolver un problema biológico de fundamental importancia. Las variables m, n y h son variables fenomenológicas, es decir, son variables que han sido incluídas para ajustar observaciones experimentales. La variable m está asociada con la apertura del canal para el sodio; la variable n , a su vez, esta asociada con la apertura del

canal de potasio y, finalmente, la variable h se asocia con la obstrucción del canal de sodio.

El estudio de las ecuaciones anteriores está lejos de estar terminado pero han sido muy importantes para el estudio de las propiedades de marcapasos fisiológicos y el enfoque tomado por los autores ha servido de modelo para otros casos. En particular, una pregunta básica es la posibilidad de que los marcapasos dejen de producir oscilaciones rítmicas para dar lugar a comportamientos irregulares. En el corazón el caso de la fibrilación ventricular es un ejemplo clásico aunque no por ello completamente estudiado. Winfree [32] ha señalado algunos fenómenos que requieren de atención y que están relacionadas con medios biológicos excitables, la mayoría de ellos relacionados con la fisiología del músculo cardiaco, particularmente con la aparición de fibrilación ventricular. Remito al lector al artículo citado para mayor información y referencias. Otro ejemplo, que ha atraído la atención de muchos investigadores, es el de la actividad rítmica de las células β pancreáticas que están asociadas con la secreción de insulina. Las investigaciones realizadas con modelos realistas pero sumamente complejos indican que los marcapasos son capaces de comportarse caóticamente y que la presencia de caos en sistemas fisiológicos que generan ritmos tipo marcapaso pudiera ser un fenómeno muy común en la naturaleza [5]. En resumen podemos concluir que las ecuaciones de Hodgkin y Huxley han sido fundamentales para el desarrollo del estudio de medios biológicos excitables y que poseen, al igual que modelos análogos para otros ritmos fisiológicos, una riqueza muy grande en comportamientos no lineales. El trabajo matemático sobre estos modelos de Hodgkin y Huxley y otros modelos relacionados, se ha centrado en la descripción en el plano fase de la excitabilidad, el análisis de oscilaciones no lineales, efectos de acoplamiento de osciladores, caracterización de soluciones tipo onda viajera y propagación en cables no homogéneos [23].

7 El ADN y la clave genética.

Para la discusión de este punto nos basaremos en un artículo de Michael S. Waterman [30] al que referimos al lector para más detalles. En 1953 James Watson y Francis Crick propusieron la estructura de doble hélice del ácido desoxiribonucleico o ADN además de proponer el mecanismo de replicación o copiado de esta molécula basado en los pares adenina-timina y guanina-citosina (A-T, G-C). Los últimos 45 años de la biología molecular se han basado sobre este descubrimiento. La propiedad

fundamental del ADN es que puede codificar información genética. El ADN es una sucesión de bases o nucleótidos. Si la cadena de nucleótidos consta de n elementos, entonces existen 4^n posibilidades de ordenarla: la complejidad de la molécula de ADN es inmensa. Por ejemplo, si la cadena tuviera 1000 nucleótidos, los posibles rearrreglos de la misma son del orden de 10^{600} , muchas más que el número de partículas elementales en el universo. Además, la manera en que la información se codifica hace uso únicamente de tres bases por aminoácido, donde los tripletes que codifican sucesivos aminoácidos en una proteína están ordenados secuencialmente en la molécula de ADN. Un problema básico de la genética molecular es la secuenciación de moléculas de ADN. El problema es monumental pero existen propiedades interesantes de la molécula que lo hacen tratable. Una de estas propiedades, en células eucariontes, es la existencia de sucesiones de bases en el ADN que no codifican nada; estas sucesiones se llaman intrones. Los segmentos que sí codifican se denominan exones. La explicación evolutiva de la existencia de intrones en el ADN eucarionte es un problema abierto todavía. Pero pasemos a un problema de más directo interés en matemáticas. Los genes codificados por el ADN no son fáciles de reconocer. Existen en la estructura del ADN tripletes específicos que indican la terminación de una cadena de aminoácidos, además del triplete ATG que indica el inicio de un gene. Los problemas matemáticos de interés en esta área son:

1. Mapeo del ADN
2. Comparación de cadenas
3. Determinación de genomas, es decir, todo el ADN de un organismo.

Una de las más interesantes áreas de desarrollo en el estudio del ADN que quiero comentar aquí se refiere a la computación con ADN. Para ello me basaré en los artículos de Ogihara y Ray [20] y Liu et al [13] publicados en *Nature* este año. Es sabido que la dificultad para encontrar soluciones a problemas matemáticos se clasifica por la velocidad a la cual los mejores algoritmos pueden obtener soluciones. Los problemas difíciles son los NP cuyos tiempos crecen exponencialmente con el número de variables involucradas. El problema de caminos o rutas de Hamilton consiste en encontrar la ruta de vuelo de un avión que conecte varias ciudades y en donde cada ciudad sea visitada solamente una vez, como dice la canción. Este es un problema NP de los más difíciles.

En 1994 Leonard Adleman propuso un método basado en el ADN que era polinomial. Un problema NP puede verse como la búsqueda de una solución que satisface simultáneamente cláusulas lógicas compuestas, cada una, de tres variables y que pueden ser verdaderas o falsas y que se encuentran conectadas por conjunciones “o”, por ejemplo el problema resuelto por Liu et al (2000) está definido por las cláusulas

$$(x_1 \text{OR } x_2 \text{OR } \bar{x}_3) \text{ AND } (\bar{x}_1 \text{OR } x_2 \text{OR } \bar{x}_3). \quad (4)$$

Este tipo de problema se denomina 3-SAT y, el resuelto por Liu et al es uno de los más sencillos de este tipo (que es NP). El método de Liu et al se inicia con una cadena binaria que representa las variables en una fórmula 3-SAT y que puede ser a su vez representada por una única secuencia de nucleótidos. Para ejemplificar la cadena TGCGG podría representar la cadena binaria 001. Para cada n variables existen 2^n respuestas, denominadas cadenas de Watson. Para cada cadena de Watson existe una cadena complementaria o cadena de Crick. El objetivo del método es identificar aquellas cadenas de la biblioteca de cadenas (en el caso de Liu con tres variables su biblioteca es de 8 cadenas de Watson) que satisfacen las cláusulas de su problema 3-SAT. Ilustremos el método de la siguiente manera. La biblioteca de 8 cadenas binarias para el problema 3-SAT citado es

x_1	x_2	x_3
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

El método se inicia asociando cadenas de ADN a cada posible respuesta en una superficie tratada especial:

Cadena de ADN	Cadena binaria	Superficie
<i>ATGCC</i>	0 0 0	1
<i>TGCGG</i>	0 0 1	2
<i>AAGCG</i>	0 1 0	3
<i>CCTAT</i>	0 1 1	4
<i>TAGAC</i>	1 0 0	5
<i>GGATT</i>	1 0 1	6
<i>CTTCG</i>	1 1 0	7
<i>GTTAT</i>	1 1 1	8

Luego, cadenas complementarias que satisfacen la primera cláusula en (4) se añaden a la solución adhiriéndose a las cadenas 1 y 3-8. La cadena sin complemento que queda, la 2, es desnaturalizada por enzimas. Después, las cadenas complementarias se eliminan y la superficie es lavada. Este ciclo es repetido para la segunda cláusula en (4) resultando en la destrucción de la cadena 6. Las cadenas restantes son leídas proporcionando la solución correcta del problema:

000, 010, 011, 100, 110, 111.

Este resultado es importante pero su ampliación para problemas 3-SAT más complejos es poco realista por ahora: existen problemas serios en la corrección de errores asociados con la química del ADN, se requieren una cantidad muy grande de cadenas de ADN para la técnica, aparte de la necesidad del diseño y construcción de cadenas de ADN únicas, y finalmente, el crecimiento exponencial de cadenas de ADN necesarias para resolver problemas 3-SAT.

8 Epílogo.

Hace no muchos años enfocar problemas biológicos desde una perspectiva matemática o computacional era poco importante para la biología y quienes se dedicaban a ello escribían esencialmente para sí mismos. Ahora, la perspectiva ha cambiado radicalmente. Las matemáticas y la computación están presentes, por ejemplo, en el análisis de las secuencias de ácidos nucleicos y otras macromoléculas y en la determinación de tasas de replicación de virus letales como el SIDA. Cito estas áreas porque en ellas es donde en años recientes la aplicación de técnicas matemáticas de modelación, analítica o computacional, han tenido resultados sorprendentes y muy importantes. La comunidad médica ha sido particularmente reticente al uso de matemáticas en la investigación clínica, sin embargo, como señalaré más adelante, el uso de modelos ha resuelto recientemente uno de los problemas más importantes relativos a la naturaleza del VIH, y éstos métodos son ahora incorporados en los protocolos de trabajo. Otras áreas, como la genética de poblaciones y la ecología, tienen una tradición de aplicación de herramientas matemáticas de casi un siglo. En años recientes, la disponibilidad de supercomputadoras ha hecho posible la elaboración y uso de modelos muy detallados en ecología de comunidades y epidemiología, por ejemplo. El reto actual en esta área de interfase entre biología y matemáticas, sean estas aplicadas al ADN, al sistema inmune, al ciclo celular, poblaciones

o bosques, es el de desarrollar modelos mecánicos que se basen en lo que es conocido sobre el sistema tanto a nivel experimental o de campo, como teórico y usar el poder predictivo y explicativo del modelo construido para analizar y explicar el comportamiento emergente en términos de las interacciones individuales de sus componentes [12].

En inmunología los avances en la aplicación de matemáticas a problemas en enfermedades infecciosas han sido notables. Alan Perelson y David Ho han determinado la tasa de replicación del virus del SIDA en 1996 [21]. El trabajo publicado sobre el tema fue el más citado en *Science* en ese año además de otorgarle el título del hombre del año a David Ho. En la determinación de la tasa de replicación del virus del SIDA los modelos matemáticos jugaron un papel relevante. A un grupo de pacientes se les suministró una droga, ritonavir, que inhibe la replicación del virus en el cuerpo. El análisis, en donde se hizo uso de modelos en ecuaciones diferenciales no lineales y estimación de parámetros estadística, determinó que la tasa de producción de virus es mayor a 10 mil millones de partículas virales por día, que las partículas libres de virus eran eliminadas del cuerpo en aproximadamente 6 horas y que la vida media de las células T infectadas productoras de viriones es de 1.5 días: el SIDA en consecuencia no es una enfermedad lenta. Las consecuencias de este resultado, en el cual los modelos matemáticos jugaron un papel fundamental, indica que la evolución del virus del SIDA es extremadamente rápida, y que las posibilidades de virus mutados resistentes a los fármacos disponibles ahora es una realidad con consecuencias graves para la salud pública.

Este resultado no solo es importante en cuanto al conocimiento del virus, sino también en la opinión y consideración en que la comunidad médica tiene ahora el uso de modelos matemáticos. En mayo de 1999 visité la Universidad de California en San Francisco y tuve la oportunidad de escuchar una conferencia dada por David Ho. En el argumento de su charla algo me llamó la atención. Para justificar el protocolo de ciertos experimentos, Ho argumentó que sus suposiciones eran razonables pues coincidían con los resultados del modelo matemático de su artículo publicado en *Science*. En la sesión de preguntas y respuestas, donde yo esperaba críticas a la aplicabilidad del modelo matemático, no escuche ninguna: el uso de modelos matemáticos en ecuaciones diferenciales se tomó con la misma actitud que la descripción de un resultado experimental usando PCR. De hecho, Alan Perelson [*comunicación personal*] afirma que cada uno de los grupos importantes en investigación de SIDA en Estados Unidos tiene ya o está buscando un “modelador

Debo subrayar que el potencial científico de las interacciones entre matemáticas y biología puede ser desarrollado únicamente si se le fomenta y promueve entre nuestros estudiantes tanto de matemáticas como de biología. Las interacciones entre individuos que han dado lugar a los desarrollos del siglo XX se deben a que problemas interesantes han existido y que mucho esfuerzo e interés se ha colocado en el establecimiento de interacciones inter y multidisciplinarias. Es necesario fomentar esta perspectiva en las escuelas e institutos de estudios superiores. La Sociedad Matemática Mexicana en ese aspecto ha contribuido de manera importante en nuestro país para lograr este objetivo: desde hace varios años, existe una sesión dedicada al área en su Congreso Nacional, que es además muy popular entre los estudiantes y que ha permitido la divulgación del trabajo nacional entre el público universitario del país.

Finalmente, no quiero concluir sin advertir que en México la biología matemática se ha desarrollado de manera desbalanceada. Un número sustancial de los investigadores que realizan este tipo de investigación son matemáticos de formación. Varios de ellos tienen amplia experiencia y conocimiento de problemáticas biológicas. Stephen Smale en algún libro dijo que la biología teórica debía desarrollarse en los departamentos de matemáticas. Yo disiento totalmente de tal afirmación. En principio biología matemática no es biología teórica, como claramente la obra fundamental de Darwin demuestra; en segundo lugar la biología teórica debe desarrollarse desde la biología a través de la constitución de grupos multidisciplinarios que incorporen entre sus miembros además de biólogos a matemáticos. La biología matemática es una forma de atacar problemas biológicos [7]. El desarrollo de la biología matemática desde la perspectiva de los biólogos seguramente enriquecerá y promoverá aún más los avances que en nuestro país se han logrado: la presencia de la biomatemática mexicana a nivel internacional empieza a ser significativa. Esperemos que este siglo que empieza (no se si este año o el siguiente) vea consolidarse la forma de ver el mundo y hacer investigación matemática desde una perspectiva biológica.

Agradecimientos.

Este artículo fue financiado por CONACYT. No puedo concluirlo sin agradecer a dos personas su apoyo y camaradería: a Juan José Rivaud por la invitación a escribir esta reseña. Realmente ha sido un reto.

Independientemente de su éxito o el fracaso, a Juanjo le debo haberme planteado el problema cuyo intento de resolución disfruté plenamente (además de agradecerle su paciencia, que tiende a infinito, y que puse a prueba con mis continuos retrasos voluntarios e involuntarios); a Gerardo Hernández, biólogo y matemático, compañero de tantos años y de tantas batallas, agradezco también sus críticas y siempre sesudos comentarios. Los errores y faltas contenidos aquí son, por supuesto, absolutamente de mi responsabilidad. En el año 2000, *saevis tranquillus in undis*.

Referencias

- [1] R. Anderson and R. May. *Infectious diseases of humans*. Oxford University Press, Oxford, 1991.
- [2] C. Darwin. *The origin of species by means of natural selection or the preservation of favored races in the struggle for life*. Facsimil, 1859. Random House, New York, 1991.
- [3] W. Feller. Diffusion processes in genetics. *Proceedings of the Second Berkeley Symp. Math. Stat. Prob.*, 1:227–246, 1951.
- [4] E. Garber. *The language of physics: the calculus and the development of theoretical physics in Europe, 1750-1914*. Birkhauser, Boston, 1999.
- [5] L. Glass and M. Mackey. *From clocks to chaos: the rythms of life*. Princeton University Press, Princeton, 1988.
- [6] J. Haldane. A mathematical theory of natural and artificial selection I. *Transactions of the Cambridge Philosophical Society*, 23:19–41, 1924.
- [7] G. Hernandez and J. Velasco-Hernandez. *El manantial escondido: un acercamiento a la biología teórica y matemática*. Fondo de Cultura Económica, México, 1999.
- [8] A. Hodgkin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.
- [9] J. Keener and J. Sneyd. *Mathematical Physiology*. Springer Verlag, New York, 1998.

- [10] W. Kermack and A. McKendrick. Contributions to the mathematical theory of epidemics I. *Proceedings of the Royal Society A*, 115:700–721, 1927.
- [11] S. Kingsland. *Modeling Nature: episodes in the history of population ecology*. The University of Chicago Press, New York, 1985.
- [12] S. Levin, B. Grenfell, A. Hastings, and A. Perelson. Mathematical and computational challenges in population biology and ecosystem science. *Science*, 275:334–343, 1997.
- [13] Q. Liu, L. Wang, A. Frutos, A. Condon, R. Corn, and L. Smith. Dna computing on surfaces. *Nature*, 403:175–179, 2000.
- [14] A. Lotka. *Elements of Mathematical Biology*. Dover, New York, 1956.
- [15] R. May. *Stability and Complexity in model ecosystems*. Princeton University Press, Princeton, 1974.
- [16] J. Maynard-Smith. *Evolution and the theory of games*. Cambridge University Press, Cambridge, 1982.
- [17] J. Murray. *Mathematical Biology*. Springer Verlag, New York, 1989.
- [18] J. Murray. Turing’s theory of morphogenesis-its influence on modelling biological pattern and form. *Bulletin of Mathematical Biology*, 52:119–152, 1990.
- [19] J. Needham. *Order and life*. The MIT Press, Cambridge, 1936.
- [20] M. Ogiwara and A. Ray. DNA computing on a chip. *Nature*, 403:143–144, 2000.
- [21] A. Perelson. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271:1582, 1996.
- [22] W. Provine. Population genetics. *Bulletin of Mathematical Biology*, 52:201–207, 1990.
- [23] J. Rinzel. Electrical excitability of cells, theory and experiment: review of the Hodgkin Huxley foundation and an update. *Bulletin of Mathematical Biology*, 52:5–23, 1990.

- [24] M. Ruse. Darwin and the philosophers: epistemological factors in the development and reception of the theory of the Origin of Species. *Creath, R and Maienschein, J: Biology and Epistemology. Cambridge University Press*, pages 28–52, 2000.
- [25] F. Scudo and J. Ziegler. *The Golden Age of Theoretical Ecology*. Springer Verlag, New York, 1978.
- [26] R. Thom. *Structural stability and morphogenesis*. Benjamin Cummings, London, 1975.
- [27] D. Thompson. *On growth and form*. Cambridge University Press, Cambridge, 1942.
- [28] A. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society (part B)*, 237:37–72, 1953.
- [29] V. Volterra. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. Premiate Officine Grafiche C Ferrari, Venezia, 1927.
- [30] M. Waterman. Genomes, maps, and sequences. *Levin, SA: Frontiers in Mathematical Biology. Biomathematics. Springer Verlag*, 100:28–52, 1995.
- [31] A. Winfree. *The geometry of biological time*. Springer Verlag, New York, 1990.
- [32] A. Winfree. Puzzles about excitable media and sudden death. *Levin, SA: Frontiers in Mathematical Biology. Biomathematics. Springer Verlag*, 100:139–158, 1995.