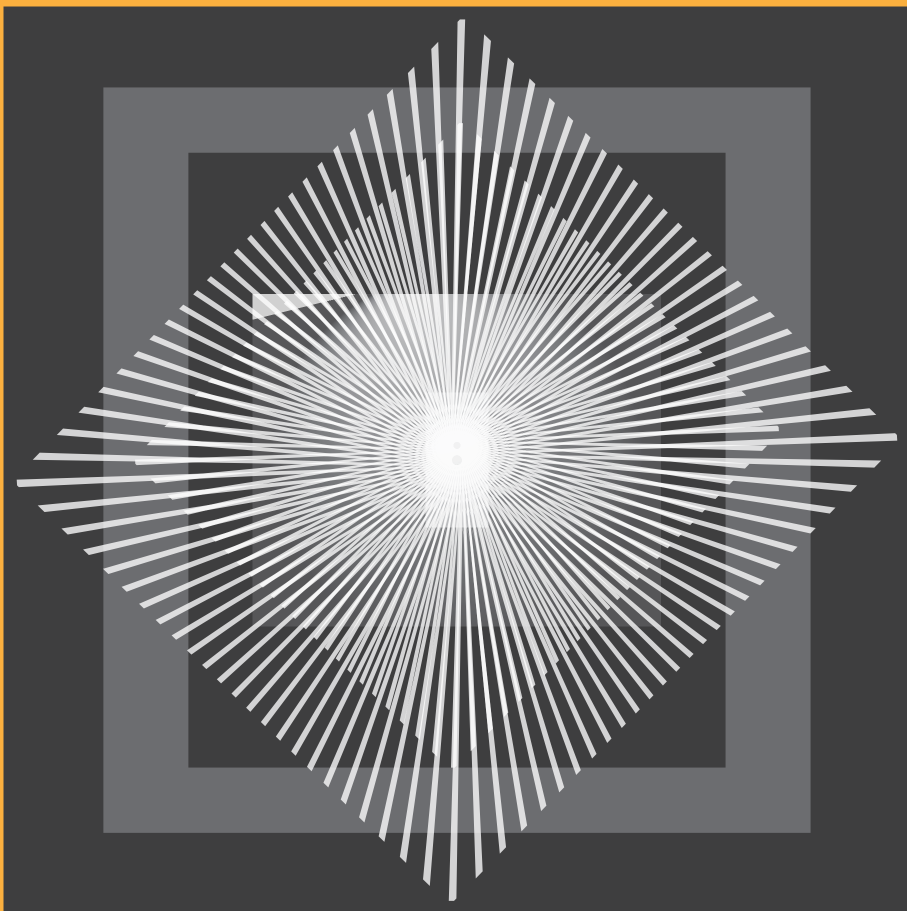
20. COLOQUIO DEL DEPARTAMENTO DE MATEMÁTICAS

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL I.P.N.

CONFERENCIAS GENERALES

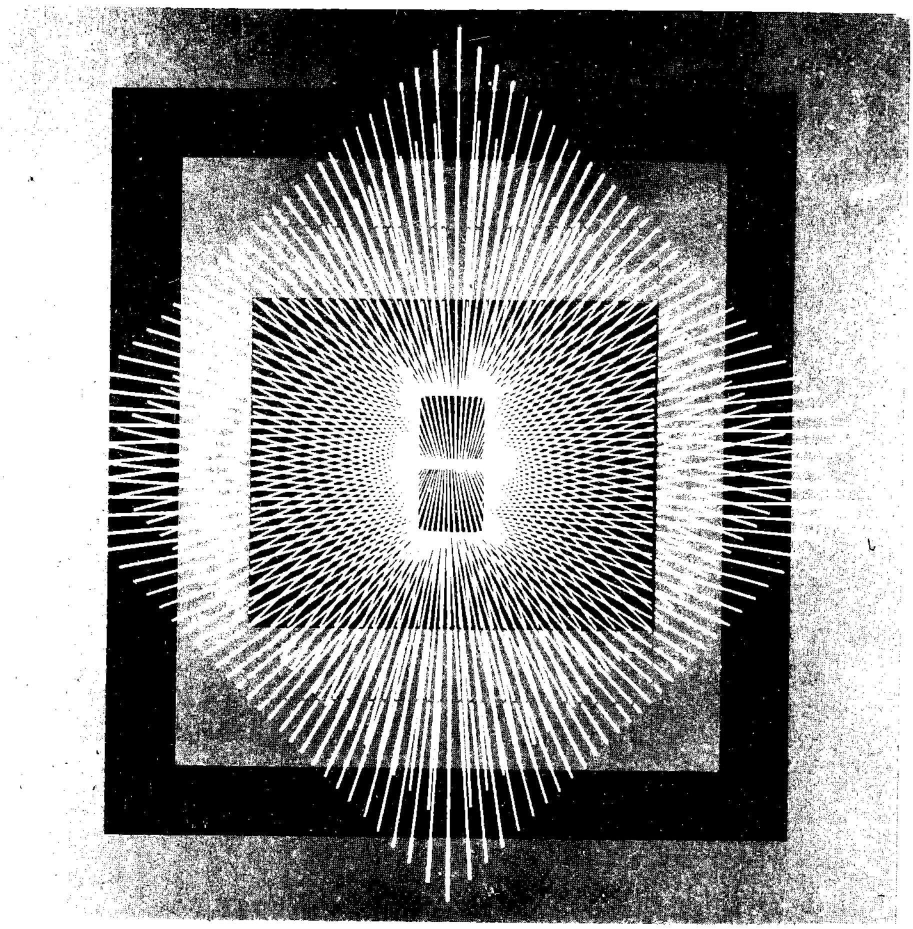
COMPILADOR: JUAN JOSE RIVAAUD



OAXTEPEC, MORELOS AGOSTO DE 1981

CONFERENCIAS GENERALES

COMPILADOR: JUAN JOSE RIVAUD



OAXTEPEC, MORELOS   AGOSTO DE 1981

CONFERENCIAS GENERALES

COMPILADOR: JUAN JOSE RIVAUD

2º COLOQUIO DEL DEPARTAMENTO DE MATEMATICAS
CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL IPN
Oaxtepec, Morelos
10 al 28 de Agosto de 1981

Patrocinadores:
Secretaría de Educación Pública
Consejo Nacional de Ciencia y Tecnología
Instituto Politécnico Nacional
Centro de Investigación y de Estudios Avanzados del IPN

EL PRESENTE VOLUMEN CONTIENE LAS CONFERENCIAS IMPARTIDAS DURANTE EL 2o. COLOQUIO DEL DEPARTAMENTO DE MATEMATICAS DEL CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITECNICO NACIONAL. SU EDICION CONSTA DE 500 EJEMPLARES.

CONTINUAMOS DE ESTA MANERA LA LABOR INICIADA EN EL PRIMER COLOQUIO, ESPERANDO QUE LA PRESENTE OBRA AYUDE A LA DIFUSION DE LA MATEMATICA Y SUS APLICACIONES.

ENRIQUE RAMIREZ DE ARELLANO
Coordinador

CONTENIDO

Este volumen consta de las Conferencias Generales dictadas durante el 2º Coloquio del Departamento de Matemáticas del Centro de Investigación y de Estudios Avanzados del I.P.N., celebrado en el Centro Vacacional del I.M.S.S. en Oaxtepec, Morelos, del 10 al 28 de Agosto de 1981, en el orden en que fueron presentadas.

El propósito de dichas conferencias fué el propiciar la discusión entre los participantes del Coloquio, de temas y tópicos científicos cercanos a las matemáticas, o de algunos aspectos de éstas que rara vez se tiene la oportunidad de ver en un curso general. Con objeto de no cansar a los participantes se procuró que hubiese diversidad en los temas tratados; por esta razón hay poca relación entre ellos.

Damos las gracias a todos los conferencistas por su valiosa colaboración.

Juan José Rivaud
Compilador

APLICACIONES DE LA INVESTIGACION DE OPERACIONES

Juan Prawda*

SUMARIO

Este trabajo pretende ilustrar sucintamente, para el lector neófito, algunas de las aplicaciones más populares de la Investigación de Operaciones y describir la naturaleza de las limitaciones más serias que existen para su uso.

La Investigación de Operaciones es una herramienta de la planeación o planificación, consistente en la utilización del método científico por grupos interdisciplinarios en las organizaciones productoras de bienes y servicios, para que éstas puedan ser más eficientes y efectivas**.

Una limitante práctica para utilizar esta herramienta en la realidad está asociada con los diversos enfoques de planificación que practiquen los decisores.

La planificación es un proceso anticipatorio de asignación de recursos para el logro de ciertos fines.

Un proceso da a entender un conjunto de actividades que se ejecutan en forma dinámica, no estática; no se dan una sola vez, se desarrollan en forma permanente y continua. Se puede planificar a corto plazo (programar recursos de inmediato), a mediano y largo plazo (planes sexenales, decenales o quinquenales).

El concepto **anticipatorio** significa que la planificación debe realizarse antes de que algo suceda para:

– Aminorar los efectos negativos que pudieran ocurrir en el futuro (por ejemplo, los resultantes del congestionamiento humano del Valle Metropolitano

* Director de Estudios y Proyectos, Dirección General de Planeación, S.E.P.; Profesor de Graduados, SEGICYT-UPIICSA-IPN.
** En este trabajo se asocia la eficiencia al consumo de recursos y la eficacia al alcance de fines.

de la ciudad de México, o los inducidos por producir más maestros de primaria de los que requiere el país).

- Aprovechar coyunturas favorables que pueda proporcionar el futuro (por ejemplo la bonanza petrolera, el uso de alguna tecnología avanzada).

Por recurso se entiende cualesquiera de los siguientes: humanos, materiales, financieros, tecnológicos o de conocimiento y el tiempo. Como estos son siempre limitados se justifica desarrollar metodologías que permitan hacer una buena asignación de ellos.

Por fines se entiende los ideales, objetivos y metas que persigue en el tiempo y en el espacio cualquier organización productora de bienes y servicios.

Independientemente de los diferentes enfoques de planeación que limitan o apoyan el uso de la Investigación de Operaciones, la **metodología general** de la planificación está constituida por la:

- Elaboración de un **diagnóstico** del presente, para identificar los aciertos y problemas asociados al funcionamiento de una organización cualquiera.
- Elaboración de una **proyección de referencia**, para pronosticar el funcionamiento de la organización a corto y mediano plazo, suponiendo que no va a haber cambios significativos en las tendencias estructurales de la misma. La proyección de referencia permite identificar los futuros problemas de la organización, bajo el supuesto que el crecimiento futuro se asemejará estructuralmente al del pasado.
- Definición de fines (ideales, objetivos, metas) para estructurar las condiciones futuras deseables de la organización. Los fines pueden clasificarse en ideales, objetivos y metas. Los **ideales** son estructuras tecnológicamente factibles de ser alcanzadas y con probabilidad de sobrevivir una vez que se consigan, pero que dificilmente pueden lograrse; son faros de luz que guían el crecimiento y desarrollo de una organización a muy largo plazo. Los **objetivos** son ideales más pragmáticos, alcanzables y medibles. Las metas son objetivos cuantificados en el tiempo y en el espacio.
- Definición de medios (políticas, estratégias, programas, tácticas, acciones), que supuestamente conducirán al sistema de su estado presente al ideal

propuesto. **Las políticas** son reglas a respetarse durante el proceso de cambio. **Las estratégias** indican el modo de empleo de los medios dada una serie de políticas y fines a alcanzar. **Los programas** son un conjunto de acciones que puestos en práctica supuestamente conducen al sistema de un estado presente a un futuro idealizado o concebido. **Las tácticas** indican el modo de ejecutar los programas.

Las acciones que configuran los programas consumen recursos, los cuales tienen que presupuestarse y ponerse en marcha a través de una calendarización (esto es programar).

— Elaboración de mecanismos de **evaluación y control**, para medir en forma permanente los logros alcanzados y compararlos con los deseados. Cuando se encuentran diferencias no tolerables entre lo deseado y lo logrado se identifican las causas para corregirlas (esto constituye un control cibernético, no policiaco).

Existen muchos enfoques de planeación pero independientemente del enfoque, los pasos metodoloógicos de todos ellos son los que se acaban de mencionar.

Es conveniente mencionar brevemente las características de algunos enfoques, especialmente el de aquel que utiliza a la Investigación de Operaciones como una herramienta de decisión.

Los principales enfoques de planeación son: **racionalistas** (optimizantes, comprensivos, satisfacientes), **incrementalistas, exploración mixta, adaptativos, innovativos, por ideales y participativos** (radicales, por apoderado e inductivos). Sin detallar cada uno de estos enfoques*, se resalta a continuación las características principales del asociado al tema de este artículo.

La planificación racionalista enfatiza la selección de medios para el alcance de ciertos fines de acuerdo con un proceso racional, entendido como aquél que cumple con ciertas reglas lógicas de transitividad y comparabilidad. Decidir racionalmente requiere enlazar el estado de ambigüedad (lo que plantea la necesidad de decidir) expresado por un conjunto de alternativas, por el acto de decidir, mediante un conjunto de operaciones deductivas.

* El lector interesado puede consultar el trabajo de J. Elizondo "Algunos enfoques de planeación", Proyecto 7188, Instituto de Ingeniería, UNAM, Nov. 1978, para mayor información.

Karl Manheim[1], sociologo alemán y pionero de los trabajos de investigación de la racionalidad en el contexto de la planeación, distinguió en sus trabajos dos tipos de racionalidad: la funcional, cuyo objetivo es emplear eficientemente los medios para alcanzar ciertos fines y la sustancial, definida como la capacidad para comprender situaciones complejas y decidir sobre los mismos fines.

Existen 3 enfoques de planeación racionalista: el optimizante, el comprensivo y el satisfaciente.

La planificación optimizante, planteada entre otros por R. Ackoff[2], se apoya en el desarrollo de modelos matemáticos, de simulación deductivos y en la disponibilidad de computadoras. Los modelos representan los estratos de ambigüedad y conocimientos mediante variables cuantitativas, cuya manipulación lógica, a través de una serie de reglas conocidas como algoritmos, consigue obtener el valor óptimo de una función, llamada objetivo, que a su vez modela la medida en que se logran los fines del plan. La Investigación de Operaciones, los modelos econométricos, la probabilidad, la estadística, los pronósticos, la ingeniería industrial, etc., son herramientas que utiliza este tipo de enfoque. En los dos volúmenes de Prawda[3], pueden entre otras referencias, consultarse una gran variedad de modelos (determinísticos, estocásticos y bajo total incertidumbre) que utiliza este enfoque de planeación.

El éxito de este tipo de planeación depende de la factibilidad de introducir en el modelo todos los aspectos que sean relevantes, lo cual no es posible en la mayor parte de las situaciones problemáticas. Por esa razón, el planificador tiende a tomar en cuenta sólo variables que pueden manejar matemáticamente. Otro sezgo consiste en plantear el problema según modelos conocidos y resueltos.

Dado el nivel de las técnicas actuales no es posible optimizar situaciones complejas.

Este enfoque de planeación ha sido severamente criticado por los siguientes motivos:

- La teoría establece que todas las alternativas de acciones están dadas, pero su número puede ser tan grande que resulta prácticamente imposible tomarlas todas en cuenta.
- No es posible conocer todas las consecuencias de cada alternativa, pués se ramifican at infinitum.

— No todos los fines perseguidos con los cuales se comparan las consecuencias de cada alternativa son cuantificables ni comparables entre sí.

Los defensores de este enfoque de planificación señalan dos grandes ventajas del mismo:

— Existen ciertos problemas específicos que pueden ser resueltos eficientemente con estas herramientas (por ejemplo, los de distribución de productos básicos en Conasupo; determinación de los niveles de inventario de productos variados en una empresa comercial; el número de bombas que requiere un gasoducto u oleoducto para mantener un mínimo de presión, etc.).
— El intento de utilizar estas herramientas de planificación, aún en experiencias fallidas, siempre deja como subproducto, mayor entendimiento del sistema por planear.

El resto de los enfoques de planeación que aquí no se discuten, pero que se orienta al lector a las referencias correspondientes[4], han sido respuestas a las limitantes de aplicar en forma universal la planeación racional optimizante. De ahí se puede deducir que existen una serie de campos o sistemas, extremadamente complejos, como los sociales y políticos, donde la Investigación de Operaciones aún no se puede utilizar.

La segunda limitante para aplicar la Investigación de Operaciones está asociada con los elementos de su definición, es decir, el uso o no de la metodología científica para decidir, la participación o no de grupos interdisciplinarios en la solución de los problemas de la organización, así como la misma estructura orgánica del sistema bajo consideración.

A continuación se menciona brevemene porque estos elementos pueden ser un obstáculo para la aplicación de la Investigación de Operaciones.

La metodología científica es la instrumentación secuencial de la:

— Observación de un sistema.
— Formulación de un problema.
— Formulación de hipótesis (construcción de un modelo).

— Aceptación o refutación de las hipótesis planteadas (solución del modelo).
— Verificación o generalización de los resultados.

Si en una organización productora de bienes y servicios se improvisan decisiones descartando la metodología científica, es altamente probable que no se utilizará la Investigación de Operaciones u otra herramienta equivalente para decidir*.

Si en una organización productora de bienes y servicios no se conjugan grupos interdisciplinarios para la resolución de problemas, seguramente tampoco se utilizará una herramienta como la Investigación de Operaciones. Por lo general el ataque de un problema los realizan grupos unidisciplinarios o multidisciplinarios.

La estructura orgánica más frecuente que presentan los sistemas productivos de bienes y servicios adopta una jerarquía piramidal, donde se confunden fines con medios y donde estos últimos empiezan a ser más importantes que los primeros. Existen muchas organizaciones con estructura piramidal que utilizan en forma limitada la Investigación de Operaciones. Otras estructuras, como la organización matricial (NASA: National Aereonautics and Space Administration de los Estados Unidos) o circular[5], hacen uso más extensivo de esta herramienta.

Hasta aquí se ha presentado un bosquejo muy sintético de las condiciones que limitan definitivamente la aplicación de la Investigación de Operaciones. Se puede resumir, que en aquellas organizaciones productoras de bienes y servicios donde no se utilicen enfoques de planeación racional optimizantes, donde no impere la metodología científica para la toma de decisiones, donde no existan grupos interdisciplinarios y donde las estructuras orgánicas bloqueen el proceso de toma de decisiones, la probabilidad de usar estos instrumentos será muy baja o nula. Lo anterior no indica que la Investigación de Operaciones no se utilice y precisamente a continuación se ilustrará muy brevemente algunas de sus aplicaciones.

En los países desarrollados la Investigación de Operaciones se ha utilizado extensamente en los servicios terciarios (grandes consorcios comerciales, transportes aéreos, terrestres, ferroviarios y marítimos, etc.), en el sector público (defensa, inteligencia, administración espacial, educación, etc.), en la industria petrolera, petroquímica, de la

* Un gran número de organismos del sector público y privado de México, ni conocen, ni se interesan por conocer u aplicar herramientas de esta naturaleza. De ahí uno de los orígenes de problemas ancestrales aún sin resolver.

construcción, de comunicaciones y telefonía, electrónica y de computación, siderúrgica, textil, automotríz, de bienes de capital y sobre todo en el sector académico y de investigación básica y aplicada.

En México existen ciertas empresas del sector público y privado que utilizan esta herramienta. Sobresalen el Instituto Mexicano del Petróleo, Teléfonos de México, Comisión Federal de Electricidad, Petróleos Mexicanos, Instituto Mexicano del Seguro Social, dentro del sector público central o paraestatal y otras, como Aurrerá, o los grupos Alfa y Visa, del sector privado.

Empecemos analizando la aplicación de los **modelos lineales determinísticos** que son los más populares, sin ser necesariamente los más efectivos. Estos se definen como:

$$
\begin{aligned}
\text{Opt } Z &= cX \\
\text{sujeto a } \; AX &\leqslant b \\
X &\geqslant 0,
\end{aligned}
$$

$X \in E^n$, $A = a_{ij}$, $i = 1,\ldots,m$; $j = 1,\ldots,n$. El detalle de cada componente de la definición anterior se puede consultar, entre otros, en el Prawda[6].

Este modelo se ha utilizado básicamente para resolver problemas de asignación, distribución y transporte.

Adolecen estos modelos de dos grandes defectos: consideran que las relaciones de los medios para alcanzar los fines y la explicitación de los mismos son funciones lineales y que todos los elementos se conocen con certeza. En la realidad y por lo general, ninguno de los dos supuestos son ciertos.

Los modelos lineales, no sólo tienen limitaciones, presentan ciertas virtudes: se pueden resolver fácilmente y los paquetes de computadora que existen hoy en día pueden llegar a resolver problemas lineales con varios millones de variables y varias miles de restricciones.

Existen ciertas situaciones donde las relaciones funcionales son de naturaleza lineal y se conoce toda la información con exactitud y certeza. Por ejemplo, el autor del presente artículo diseñó un modelo de distribución de trigo para CONASUPO que se operó con bastante éxito durante el período 1971-1978.

Otro enfoque más práctico para resolver problemas lineales determinísticos es el de redes.

La formulación general de un problema de redes, cuyo detalle se puede consultar entre otros, en Prawda[7], es:

$$\text{Mín } Z = \sum_{A_{ij}} c_{ij} X_{ij}$$

sujeto a

$$\sum_{N_i} X_{ij} - \sum_{N_k} X_{jk} = \begin{cases} -v, & \text{si } N_j = N_s \\ 0 & \text{si } N_j \neq N_s, N_t \\ v, & \text{si } N_j = N_t \end{cases}$$

$$0 \leqslant \ell_{ij} \leqslant X_{ij} \leqslant u_{ij}, \text{ para toda } A_{ij}$$

$$v \geqslant 0.$$

Los problemas tradicionales que se pueden resolver con el enfoque de redes son: de flujo máximo, costo mínimo, flujo máximo a costo mínimo, expansión, árboles de comunicación, cadenas múltiples, multiflujos, redes de actividad (ruta crítica y PERT), síntesis de redes.

Todo problema lineal puede resolverse indistintamente por programación lineal o por redes, pero su formulación es más sencilla con este último enfoque.

La programación dinámica, introducida por Richard Bellman[8], es un enfoque versátil que permite resolver problemas lineales y no lineales, discretos y continuos, determinísticos y estocásticos, univariables y multivariables. Desgraciadamente adolece de la llamada maldición de la dimensionalidad, correspondiente a un requerimiento exponencial de memoria en una computadora en función del número de variables del problema. Si bien es cierto que con las tecnologías actuales de las computadoras se ha logrado aumentar considerablemente la dimensión de los problemas a resolver mediante este enfoque, también lo es que se sigue teniendo limitantes serias.

Una función recursiva típica del enfoque de programación dinámica, cuyo detalle se puede consultar, entre otros en el Prawda[9], es el siguiente:

$$g_n(X_n) = \text{Opt} \quad r_n(X_n, D_n) + g_{n-1}(X_{n-1}) \quad , n = n, n-1, \ldots$$

$$X_{n-1} = f_n(X_n, D_n), \ \forall n,$$

$$g_0(X_0) \equiv 0.$$

Otro modelo lo constituye la programación no lineal o programación matemática[10] que se define como

$$\text{Opt} \ f(X)$$

$$\text{sujeto a} \quad g_i(X) \leqslant 0, \ i = 1, \ldots, m$$

$$X \in E^n,$$

donde $f(x)$ y las m funciones $g_i(x)$ son estructuras no lineales de un vector x en el espacio n-dimensional.

La programación no lineal ha servido para resolver problemas determinísticos donde las relaciones funcionales de los fines que persigue la organización, así como la estructuración de los medios para alcanzar esos fines (restricciones), son no lineales. Estos modelos se han utilizado, entre otros, para calcular tasas internas de retorno en problemas financieros[11], en diseño ingenieril[12], en la industria petrolera y petroquímica y en la elaboración de modelos econométricos[13].

El autor de este trabajo y un alumno doctoral suyo, utilizaron un modelo de programación no lineal para asignar una variedad de recursos humanos a futuro en un hospital de Nueva Orleans, U.S.A.[14].

Los problemas enteros o combinatorios[15] se definen como

$$\text{Opt } Z = cX + dY$$

$$\text{sujeto a } AX + BY - b$$

$$X \geqslant 0, \text{ entero}$$

$$Y \geqslant 0,$$

$$X \in E^n, Y \in E^m.$$

Los modelos enteros han servido para resolver problemas de secuenciación[16], (especialmente en grandes fábricas que utilizan máquinas herramientas), de localización[17], (estaciones de bomberos, servicios médicos de emergencia, recolección de basura, tratamiento de aguas negras), ruteo de transportes[18], asignación de recursos financieros en diferentes inversiones y balance de líneas de producción[19].

Otras de las aplicaciones populares de la Investigación de Operaciones la constituyen los modelos de inventarios. Estos determinan la frecuencia y el volumen razonable de adquisición de un producto en función del balance de dos costos, uno de almacenamiento que aumenta con la cantidad de producto adquirido y otro, penal o de oportunidad, que disminuye inversamente al nivel del inventario[20]. Estos modelos se han desarrollado aún para el caso donde la demanda de un producto es estocástico.

La Investigación de Operaciones ha servido para describir las líneas de espera o colas[21] en función de dos costos: uno social (asociado a la espera) y otro de servicio (asociado al gasto corriente y de inversión del mismo).

Estos modelos se han utilizado para determinar un número razonable de salas de parto en un hospital, o el de camas en un pabellón, o de unidades de emergencia, de controladores aereos en un aeropuerto, de gruas mecánicas en un puerto, de cajeras en un banco o en una tienda de autoservicio en función de la hora del día, del día de la semana y del mes del año que se trate, de operadoras en una central telefónica, de garitas en una autopista de peaje, etc.

La Investigación de Operaciones ha generado modelos que determinan la frecuencia de reemplazo de ciertos equipos (alumbrado público, relays en elevadores, etc.), de mantenimiento preventivo y/o correctivo y ha permitido estructurar

diseños de confiabilidad de sistemas eléctricos, electrónicos y mecánicos[22].

Se han desarrollado muchos modelos de Investigación de Operaciones en coyunturas determinísticas y estocásticas, es decir, cuando la información que requiere el modelo se conoce respectivamente con certeza o se le puede asociar una distribución de probabilidad conocida. Existen sin embargo algunos modelos que se han desarrollado en coyunturas de total incertidumbre, donde no se conoce la información, ni se le puede asociar una distribución de probabilidad. El **análisis de decisiones** ha sido el modelo clásico bajo estas circunstancias.

Una situación real es la que confrontan frecuentemente los funcionarios superiores de Pemex, quienes deben decidir a priori cursos de acción en función a supuestos que una capa geológica contenga suficiente petróleo, poco petróleo, solamente gas o se encuentre totalmente seca.

Otra gran limitante para aplicar modelos de Investigación de Operaciones es la habilidad de formular la realidad que uno observa sin deformarla. De ahí que casi el 85 de las aplicaciones de Investigación de Operaciones se hacen a través de procesos de **simulación**[23], donde más que construir un modelo matemático de la realidad, se aparenta en una computadora lo que se observa[24].

Las aplicaciones de la simulación han sido muy variadas y muy extensas[25] y sólo se anota en este documento las hechas por el autor y un alumno doctoral suyo, para formular políticas de vacunación de perros callejeros en la Ciudad de Cali, Colombia[26].

Los modelos de **planeación corporativa** de los grandes emporios farmacéuticos, industriales, financieros y periodísticos (Ceiba-Geigy, Anheusser-Bush, Citibank, New York Times), han sido híbridos que conjugan la simulación con modelos matemáticos inductivos o deductivos, normativos y descriptivos[27].

Todo lo que se inicia debe concluir y por razones de espacio y de respeto al lector, se finiquita aquí este ensayo.

Existe mucha bibliografía escrita al respecto. Sólo el autor de este trabajo ha publicado más de 2 mil páginas sobre este campo. Sólo me queda invitar al lector interesado y motivado, a hojear más profundamente algo del acervo científico de la Investigación de Operaciones.

BIBLIOGRAFIA

1. *Manheim, K.,* Man and society in an age of reconstruction. Harcourt Brace and Company, 1951.
2. *Ackoff, R.,* A concept of corporate planning. J. Wiley, 1970.
3. *Prawda, J.,* Métodos y modelos de Investigación de Operaciones. Tomo I (Modelos determinísticos), Tomo II (Modelos estocásticos), publicados respectivamente por Editorial Limusa en 1976 y 1980.
4. Planeación racional comprensiva:

   Banfield, E., Politics planning and the public interest. The Free Press, 1955.

   Bolan, R., Emerging views of planning. American Institute of Planners Journal, Vol. 33, 1967.

   Bolan, R., Community decision behavior: The culture of planning. American Institute of Planners Journal, Vol. 35, 1969.

   Planeación racional satisfaciente:

   March, J.Q. y Simon, H., Organizations. J. Wiley, 1959.

   Incrementalismo disjunto:

   Braybrooke, D. y Lindblom, A strategy of decisions. Free Press, 1963.

   Exploración mixta:

   Etzioni, A., The active society. A theory of societal and politcal processes. Free Press, 1968.

   Adaptativa:

   Ackoff, R., A concept of corporate planning. J. Wiley, 1970.

   Innovadoras:

   Friedman, J., Retracking America. A theory of transactive planning. Doubleday Publishing Co., 1973.

   Ideales:

   Ozbelckan, H., The emerging methodology of planning. Fields Within Fields, Vol. 10, pp 63-80, 1973.

   Participativo:

   Ackoff, R., Redesigning the future. J. Wiley, 1974.

Ackoff, R., National development planning revisited. Operations Research, Vol. 25, 1977.

Participativo radical:

Grabow, S. y Heskin, A., Foundations for radical concept of planning. Journal of the American Institute of Planners, Vol. 39, 1973.

Participativo por apoderado:

Davidoff, P., Advocacy and pluralism in planning. Journal of the American Institute of Planners, Vol. 28, 1965.

5. *Davis, S.M. y Lawrence, P.R.,* Matrix. Addisson-Wesley, 1977.
   Ackoff, R., The art of problem solving. John Wiley, 1978.
6. *Prawda, J.,* Ibid, pp 67, Vol. I.
7. *Prawda, J.,* Ibid, pp 329, Vol. I.
8. *Bellman, R.,* Dynamic Programming. Princeton University Press, 1957.
9. *Prawda, J.,* Ibid, pp 437, Vol. 1.
10. *Prawda, J.,* Ibid, pp 615, Vol. 1.
11. *Taylor, G.A.,* Managerial and Engineering Economy. D. van Nostrand, 1964.
12. *Beightler, C. y Phillips, D.T.,* Applied Geometric Programming. J. Wiley, 1976.
13. *Lasdon, L. y Warren, A.,* Survey on nonlinear programming applications. Operations Research, Vol. 28-5, pp 1029-1073, Sept.-Oct., 1980.
14. *Warner, D. y Prawda, J.,* A mathematical model for scheduling nursing personnel in a hospital. Management Sciences, Vol. 19-4, pp 411-422, 1972.
15. *Prawda, J.,* Ibid, pp 503, Vol. 1.
16. *Prawda, J.,* Ibid, pp 631, Vol. II.
17. *Prawda, J.,* Ibid, pp 579, Vol. II.
18. *Prawda, J.,* Ibid, pp 584, Vol. I.
19. *Prawda, J.,* Ibid, pp 677, Vol. II.
20. *Prawda, J.,* Ibid, pp 93, Vol. II.
21. *Prawda, J.,* Ibid, pp 243, Vol. II.
22. *Prawda, J.,* Ibid, pp 695, Vol. II.
23. *Prawda, J.,* Ibid, pp 315, Vol. II.
24. *Naylor, T.H.,* Some statistics of simulation models. Simulations, pp 171-176, Junio 1975.

25. *Naylor, T.H.,* Simulation and gaming bibliography. Computing Reviews, Vol. 10-1, pp 61-62, 1969.

26. *Frerichs, R. y Prawda, J.,* A computer simulation model for the control of rabies in an urban area of Colombia. Management Sciences, Vol. 22-4, pp 411-421, 1975.

27. *Naylor, T.H.,* Corporate Planning Models. Addisson-Wesley, 1979.

Naylor, T.H., (editor) Simulation Models in Corporate Planning. Praeger Publishing Co., 1979.

EL MOMENTO ANGULAR EN EL UNIVERSO Y LA MORFOLOGIA GALACTICA

L. Carrasco*

Recientemente, una nueva regularidad en el Universo ha sido investigada. Esta ley establece la existencia de una correlación entre la masa (M) de los objetos astronómicos que conforman el zoológico universal y su cantidad de rotación, esta última medida por el momento angular (J) de los objetos

$$J = I\omega = \int_V r^2\omega(r)\rho(r)dV \tag{1}$$

donde ω representa una velocidad angular y I el momento de inercia. Esta relación fue señalada por primera ocasión por P. Brosche en 1963, quien estableció una ley del tipo:

$$J \alpha M^2 \tag{2}$$

Sin embargo, una revisión de los datos más confiables de que disponemos en la actualidad, señala que la ley más probable es realmente

$$j \alpha M^{2/3} \tag{3}$$

* Instituto de Astronomía, Universidad Nacional Autónoma de México

donde $j \equiv J/M$ (Carrasco, Roth y Serrano, 1981). Esta es una relación que se espera como consecuencia de la existencia de la 3a. ley de Kepler la cual relaciona los períodos con los tamaños de las órbitas de los planetas alrededor del Sol. Los datos actualizados que se utilizaron para sacar esta conclusión se resumen en la Figura No. 1, en donde se ha graficado el logaritmo del momento angular contra el logaritmo de la masa para una gran variedad de objetos astronómicos, como podemos juzgar de esta figura, la ley de la que hablamos es válida tanto para objetos pequeños como son los asteroides hasta cúmulos de estrellas y galaxias, siendo estas últimas conglomerados de 10^{11-12} estrellas !!! A partir de un análisis serio del problema, se puede interpretar este resultado como consecuencia directa de las condiciones físicas prevalecientes durante la formación de cualquier objeto astronómico; si partimos de la hipótesis de que el Universo como un todo está en rotación, entonces la formación de cualquier sistema, por medio de contracción gravitacional (contracción bajo su propio peso) tendría un exceso de momento angular y por otro lado podemos demostrar que la cantidad de rotación máxima que un esferoide de masa M y densidad media $\bar{\rho}$ puede tener está dado por:

$$j = \sqrt{G K_r K_g}\ \bar{\rho}^{-1/6} M^{2/3} \qquad (4)$$

donde G es la constante universal de la gravitación, K_r y K_g son factores que dependen de la geometría específica de los objetos.

Esta cantidad de rotación se opone a la creación de sistemas estables (desde el punto de vista mecánico) debido a la existencia de la fuerza centrífuga, la cual opera fundamentalmente en el plano ecuatorial (plano perpendicular al eje de rotación) de un sistema, oponiéndose a su contracción. Cuando se tiene un objeto en equilibrio roto-gravitacional, para poder contraerlo y fraccionarlo, cuando vamos por ejemplo de galaxias a estrellas individuales, tenemos que deshacernos de momento angular de manera tal que el nuevo equilibrio se alcanza, cuando la gravitación y la rotación cumplen con la relación (4) para el momento angular y la masa.

La relación (4) predice una ley muy específica para el momento angular y la masa de los objetos astronómicos. En particular predice que los sistemas aplanados debido a rotación j, deben seguir la relación:

$$j = \sqrt{G K_r K_g} \; (\pi h)^{-1/4} \, \bar{\rho}^{-1/4} \, M^{3/4} \tag{5}$$

Las relaciones 4 y 5 señalan claramente que la densidad media $\bar{\rho}$ de un sistema interviene en la condición de equilibrio mecánico, y puesto que los objetos graficados en la Figura 1 cubren un intervalo muy grande de densidades, como podemos juzgar de la Tabla 1 (por ejemplo $\bar{\rho}$ = 1 gramo/cc. para los planetas y $\bar{\rho} = 10^{-23}$ para las galaxias). El efecto de la densidad se manifiesta en la Figura 1, como un aumento de la dispersión de los puntos y como un aumento ficticio de la pendiente de la relación observacional. Debido a estas últimas razones, resulta conveniente definir una nueva cantidad $Q \equiv j\bar{\rho}^{-1/6}$, la cual deberá depender exclusivamente de la masa de un objeto y su geometría pero no de su densidad media. Una vez corregida la cantidad de rotación de un cuerpo por densidad, podemos ver de la Figura 2, que cuando graficamos $\log Z$ vs. $\log M$, la dispersión es menor y la pendiente también lo es. En la Tabla 2, resumimos los valores para las pendientes (α y β) que relacionan las cantidades j y Q con la masa de los objetos ($j \alpha M^{\alpha}$ $Q \alpha M^{\beta}$). En esta Tabla podemos ver que la pendiente debe ser menor a 1, y que cuando se compensa por densidad, la pendiente para todo tipo de objeto astronómico, queda comprendida en el intervalo [2/3,3/4], en excelente concordancia con la predicción dada por nuestra hipótesis de trabajo: El Universo como un Todo ROTA. Por tanto sus componentes parciales "nacen" con una cantidad máxima de rotación correspondiente al equilibrio.

De la Figura 2 hay que señalar que tanto las estrellas individuales, representadas por (X), como las galaxias elípticas (representadas por Δ) caen aproximadamente en un orden de magnitud por debajo de la relación general, sin embargo siguen la misma ley de pendiente 2/3-3/4. Esta deficiencia indica que estos objetos han perdido una fracción aproximadamente constante de su momento angular original en equilibrio. Esta pérdida de cantidad de rotación es probablemetne debida a condiciones ambientales. Esta deficiencia de momento angular en las galaxias elípticas es muy probablemente la causa fundamental de su diferencia morfológica con las galaxias espirales.

MORFOLOGIA DE LAS GALAXIAS

La caracterización morfológica de las galaxias, es en sí un problema muy complejo, puesto que no existen dos galaxias idénticas en el Universo. Sin embargo, podemos decir que existen dos grandes grupos dentro de este tipo de objetos astronómicos: Las primeras quedan tipificadas por las siguientes características generales:

i) Ausencia de gas interestelar,

ii) compuestas por estrellas muy viejas,

iii) son elipsoides, cuyos ejes tienen cocientes del orden de la unidad.

En contraste, las galaxias espirales las podemos tipificar por medio de las siguientes características:

i) Presencia de gas interestelar,

ii) compuestas por una mezcla de estrellas viejas y jóvenes,

iii) son sistemas aplanados (discos), cuyos ejes principales muestran cocientes grandes (10) para el radio del sistema dividido por la altura media (grosor) del disco (h).

Los investigadores dedicados al problema consideran que estas diferencias se deben básicamente a la eficiencia con que se forman las estrellas en los diferentes sistemas, siendo las galaxias elípticas mucho más eficientes formando estrellas que las galaxias espirales.

Las diferencias morfológicas se pueden explicar si suponemos que la facilidad o dificultad para formar estrellas en una galaxia determinada depende por ejemplo de su cantidad de rotación. Esta tendría que ser una relación inversa dado que las galaxias elípticas han sido muy eficientes formando estrellas, razón por la cual hoy día no tienen gas disponible para formar nuevas estrellas, ya que todo el gas fué procesado en estrellas desde tiempos comparables a la edad misma del Universo. Para investigar la relación entre rotación y morfología, Carrasco, Serrano y Roth (1981) han estudiado varias galaxias espirales, en particular nuestra Galaxia (la Vía Láctea) y han podido correlacionar la eficiencia de formación estelar con el inverso de la cantidad de rotación en diversas regiones de nuestra galaxia para las cuales tenemos información disponible acerca de la cinemática del gas. Esta relación se nota claramente en la Figura 3, en donde hemos graficado de manera esquemática, por un lado la tasa de formación de estrellas jóvenes (N*) en función de la distancia galactocéntrica (R) y por otro lado las condiciones de rotación del gas interestelar. Vemos pues que en las zonas en las cuales el gas interestelar rota lentamente, tenemos concidentemente la presencia de estrellas jóvenes, recientemente formadas. Quedan aún muchos problemas y detalles por resolver en el contexto rotación-morfología galáctica, pero podemos decir hoy día que hemos comenzado a dar pasos firmes en el entendimiento de la morfología de las galaxias.

DENSIDAD TIPICA DE OBJETOS EN EL ZOOLOGICO ASTRONOMICO

Tipo de Objeto	Densidad Media [gm/cm³]
Asteroides, Satélites y Planetas	1
Estrellas Binarias en Contacto	10^{-2}
Estrellas Binarias Separadas	10^{-6}
Cúmulos y Galaxias Elípticas	10^{-22}
Galaxias Espirales	10^{-24}
Universo	10^{-29}

TABLA 1

Objetos	$J \alpha M^{\alpha}$		$Q \alpha M^{\beta}$	
	α	r	β	r
Asteroides, Satélites, Planetas	0.66±0.06	0.955	0.66±0.04	0.976
Binarias en Contacto	0.71±0.05	0.947	0.76±0.02	0.992
Binarias Visuales	0.67±0.07	0.711	0.68±0.03	0.910
Binarias "TODAS"	0.096±0.096	0.09	0.69±0.02	0.94
Cúmulos, Bulbos y Elípticas	0.66±0.07	0.998	0.62±0.01	0.997
Espirales + Supercúmulo	0.84±0.05	0.93	0.74±0.03	0.98
TODOS	0.94±0.09	0.987	0.71±0.01	0.994

TABLA 2

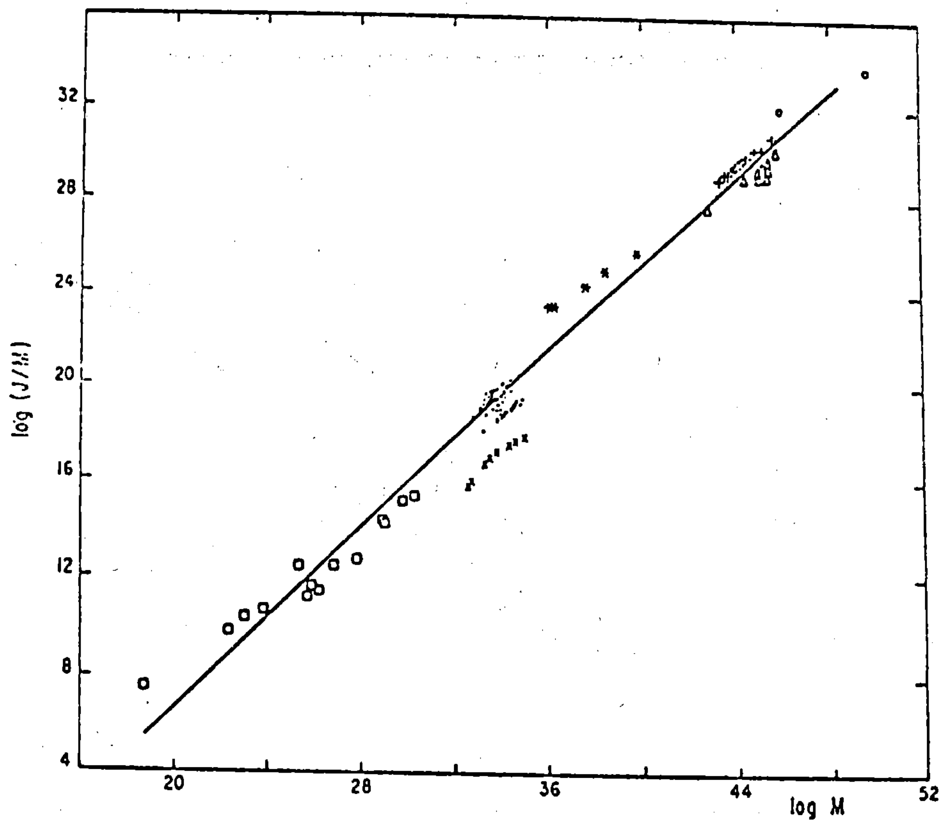


Fig. 1. Se muestra de manera gráfica la correlación que existe entre el momento angular específico (J/M) como función de la masa (en escala logarítmica) para una gran variedad de objetos astronómicos: (o) cumulos de galaxias, (+) galaxias espirales, (Δ) galaxias elípticas, (*) cumulos galácticos, (·) estrellas binarias (x) estrellas individuales, ( ) objetos de sistema planetario.

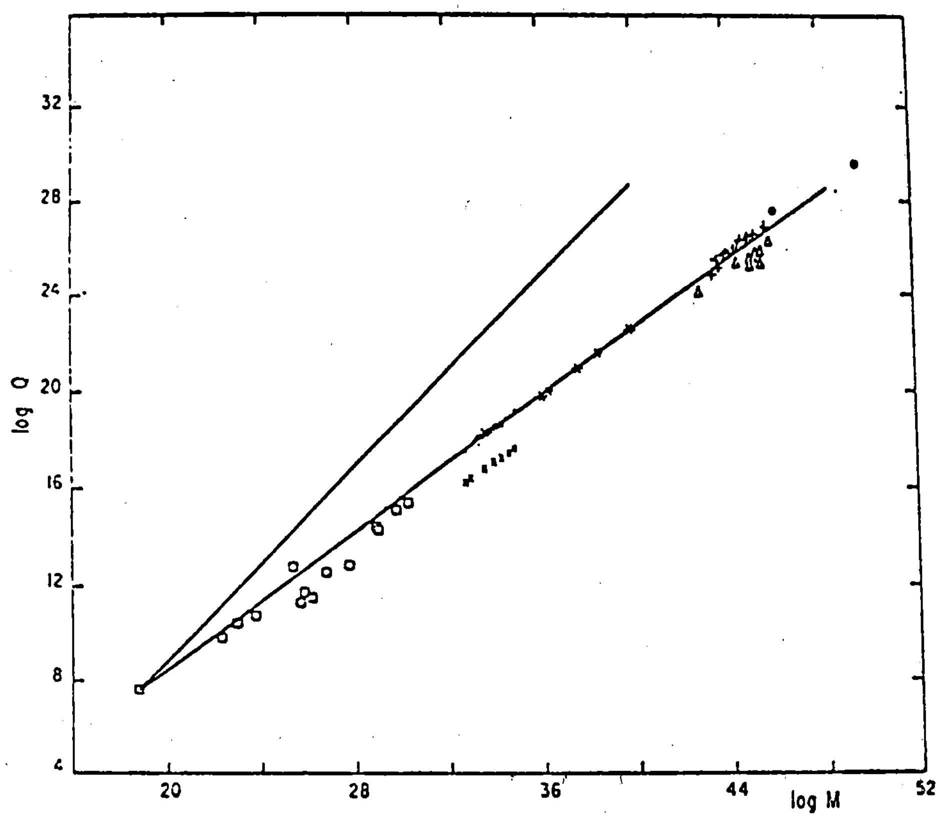


Fig. 2.
Es similar en código simbólico a la figura 1, sólo que se han graficado la cantidad $Q \equiv j\bar{p}^{\,1/6}$ en función de la masa de los mismos objetos.

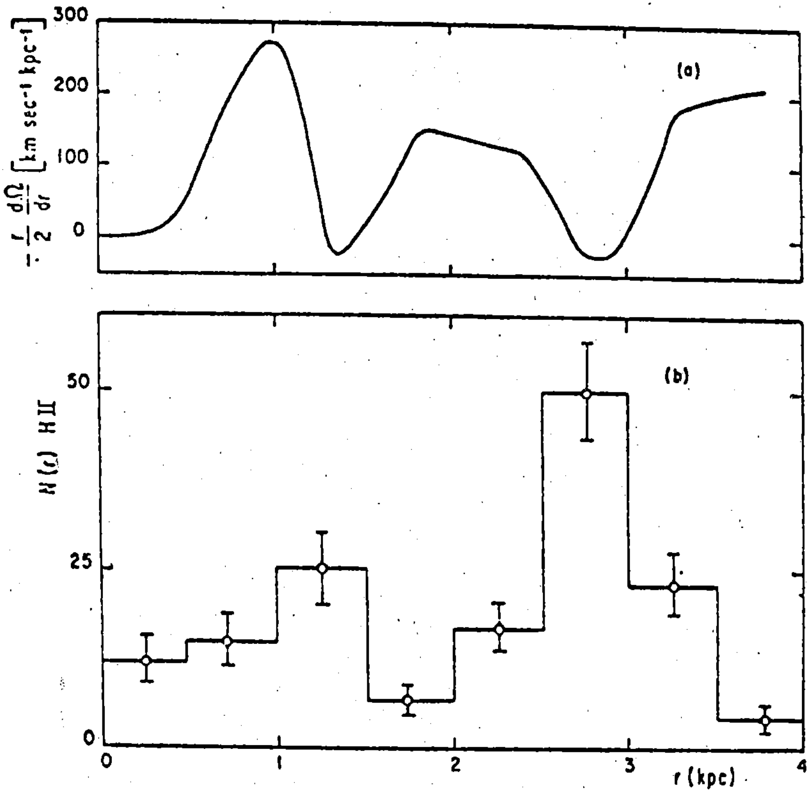


Fig. 3. a) Se muestra la distribución de momento angular específico como función del radio galactocéntrico, para la galaxia M51.

b) Se muestra la función de eficiencia de formación de estrellas en función del radio para la misma galaxia. Hay que hacer notar el gran grado de anticorrelación que existe entre estas dos cantidades.

ENFOQUES TERMODINAMICOS DE LA BIOLOGIA

Marcelino Cereijido*

Ya por el siglo XVII, poco después que Galileo ideara su termoscopio, y ni bien Jean Ray y el duque Fernando II de Toscana crearan los primeros termómetros con los que se iniciaban los rudimentos de la futura Termodinámica, se aplicaron dichos instrumentos al estudio de fenómenos netamente biológicos como la fiebre y las incubaciones. Sin embargo cuesta creer que una ciencia como la Termodinámica, desarrollada originalmente para entender gases, calderas y máquinas muy alejadas de los seres vivos, le diera luego tantos dolores de cabeza a la Biología (y viceversa). Es recién en nuestros días que esa historia de discrepancias y frustraciones está teniendo un final feliz. Pero esa interacción, desesperante por momentos, llevó a físicos, matemáticos y biólogos a revisar sus conceptos, perfeccionar sus herramientas, profundizar sus campos, y cambiar no sólo a dichas ciencias, sino a nuestra visión del mundo, creencias y filosofías. En esta breve exposición me propongo hacerles un bosquejo de dicha interacción y del estado actual de los problemas.

En los primeros tiempos, mientras pensadores como Francis Bacon y químicos como Joseph Black empezaban a distinguir entre calor y temperatura; o físicos como Benjamín Thompson atacaban la teoría del calórico; o ingenieros militares como Sadi Carnot aprendían a distinguir entre las interacciones de los sistemas y sus cambios internos de estado; o mientras James Prescott Joule, Julius Robert von Mayer y William Thompson formalizaban las ideas sobre trabajo y cambios de estado, la Termodinámica y la Biología no tuvieron mucho que ver, o por lo menos no discreparon. Los problemas aparecieron recién a mediados del siglo pasado cuando Rudolf Julius Clausius postuló claramente sus dos leyes, y sobre todo después de su enunciado "Entropie strebt eine Maxium zu". Desde entonces podemos dividir las relaciones entre Termodinámica y Biología en dos grandes ramas. La primera es la que

* Centro de Investigación y de Estudios Avanzados del IPN, México, D.F., México.

pasa por J.C. Maxwell, J. Williard Gibbs, Max Planck, E.A. Guggenheim y Walther Nernst y que dió a los biólogos los fundamentos necesarios para estudiar las reacciones químicas del metabolismo, los cambios de potenciales eléctricos en las fibras nerviosas y el pH sanguíneo. Ahí no hubo problemas. Pero la otra rama, la que pasa por Henri Poincaré, Constantin Carathéodory y llega a Lars Onsager, Erwin Schroedinger, Ilia Prigogine y Harold Morowitz trajo problemas a la Biología y a la Termodinámica. Veamos porque con algunos ejemplos.

Ejemplo I. La Evolución Termodinámica y la Evolución Biológica.

La segunda mitad del siglo pasado fué testigo de un enorme adelanto biológico que terminó con el despegue total de la Biología de posiciones animistas y creacionistas y que la fué encaminando a transformarse en la ciencia quasi-exacta que hoy conocemos. Claro que la Química Biológica era considerada todavía y por mucho tiempo como una aplicación un tanto sucia de la Química más científica y formal del mundo no-biológico. Pero así y todo, fué en la segunda mitad del Siglo XIX cuando la enorme diversidad (pasada y presente) del mundo biológico pasó a ser comprendida como la evolución progresiva de seres unicelulares, luego de bichitos simples como las esponjas y los gusanos, peces y después —mucho después— nosotros. Esa evolución presentaba flagrantes contradicciones con la evolución del mundo físico que comenzaban a formalizar los termodinamistas, y en base a la cual los sistemas tienden a estados cada vez más desorganizados y caóticos. La evolución que presentaban los termodinamistas parecían contradecir la complejización filogenética de (ameba a Homo sapiens) o y ontogenética (de huevo fecundado al estado adulto). Los pobres biólogos que estudiaban la evolución estaban apretados en una tremenda prensa: por un lado los religiosos que se resistían a tirar el Génesis por la ventana, y por otro, los termodinamistas que, como hizo Lord Kelvin, limitaban sus postulados a "entidades inanimadas" dejando todo lo biológico afuera. Baste decir que en pleno 1935, M. Jacobs en un trabajo utilísimo sobre la difusión que escribió para aprovechar el tiempo de postración (se había roto una pierna y tenía que hacer reposo) incluyó la siguiente frase:

> According to the second law of thermodynamics the distribution of matter and energy in the universe tends to become less and less orderly and more and more of the sort that would result from the operation of the law of chance. From this point of view, even the simplest organism is an almost incredibly improbable accumulation of matter.

Ejemplo II. El transporte activo.

Hace más de un siglo, el fisiólogo Dubois Reymond descubrió que si montamos una piel de rana entre dos cámaras con solución salina *idéntica* entre ambos lados (Fig. 1a), se desarrolla una diferencia de potencial eléctrico entre ambas cámaras. Tratando de explicar dicho fenómeno, Galeotti postuló en 1904 que se debía a que la permeabilidad al sodio era mayor de afuera para adentro que de adentro para afuera. Esta explicación fué poco menos que ridiculizada por los físicos de entonces pues vieron en ella una violación de los sacrosantos principios de la Termodinámica. En efecto, la mayor permeabilidad del sodio hacia la derecha traerá como consecuencia un aumento de su concentración de ese lado y, si construimos una cámara circular (Fig. 1b) se originará un proceso difusivo neto en el sentido contrario a las agujas del reloj.... y se tendrá un perpetum móbile! Luego los biólogos observaron que el móbil no era tan perpetuo que digamos, porque a las pocas horas la piel se moría y todo cesaba. Razonaron que dicha preparación consumía energía del metabilismo para hacer el trabajo de transportar al sodio. Pero sus explicaciones volvieron a ser vanas. Ahora están violando el Principio de Curie, respondieron los termodinamistas: magnitudes de orden tensorial distinto no son acoplables. Dicho de otro modo: si las reacciones químicas son fenómenos escalares y los flujos son fenómenos vectoriales, uno no puede esperar que las reacciones químicas del metabolismo celular originen un flujo de izquierda a derecha. Pero cuarenta años después llegaron los isótopos radiactivos, se pudieron medir los flujos unidireccionales de sodio y se vió que, tal como esperara Galeotti, el flujo de sodio hacia la derecha es mayor que el flujo hacia la izquierda. Era cierto entonces que existía un flujo neto, y hoy somos muchos los biólogos que nos ganamos la vida estudiándolo. ¿Se debía aceptar entonces que los fenómenos vitales contradijeran las leyes de la Física?

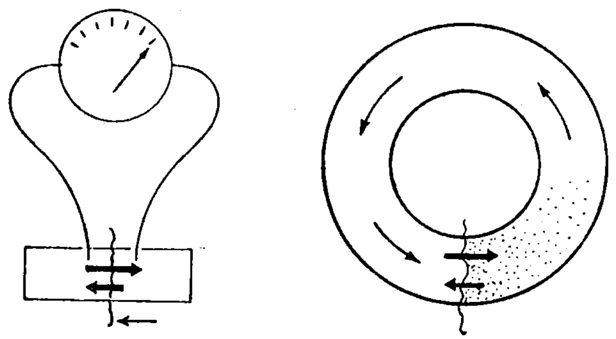


Fig. 1a. Piel de rana montada entre dos cámaras con solución salina idéntica en ambos lados. La interpretación de Galeotti era que la permeabilidad del Na de afuera (izquierda) hacia adentro era mayor que a la inversa y así se explicaba la diferencia de potencial entre ambos lados.

Fig. 1b. Razonamiento de los críticos de Galeotti: de ser cierta la interpretación de dicho investigador, el mayor flujo de sodio hacia la derecha aumentaría la concentración de ese ión y esto provocaría una difusión en el sentido contrario a las agujas del reloj.

Ejemplo III. El origen de la Vida. -

Antes de que apareciera la vida en la Tierra el mar era una especie de sopa que contenía gran cantidad de moléculas orgánicas "prebiológicas" de las que hoy constituyen los seres vivos. De manera que fué natural postular que, en tantos millones de años en el enorme volumen del mar prebiológico, una fluctuación fortuita hubiera dado origen a un ser muy primitivo, y que así comenzó la vida en el planeta. Toda una rama de la Biología se dedicó a estudiar las condiciones que dieron origen a los primeros aminoácidos, polímeros, etc. y de ahí al primer organismo creado al azar.

Pero Harold Morowitz hizo el siguiente cálculo. Tomó como base la bacteria *Escherichia coli*, uno de los seres más simples, y calculó la energía contenida en las uniones de sus moléculas. El cálculo no era complicado: sabiendo cuantas proteínas, cuantos azúcares, cuantos lípidos, cuantos fosfatos, ácidos desoxirribonucleico, etc. contiene una bacteria, y los tipos de enlaces que tienen los átomos que constituyen las moléculas, pudo hacer una estimación de la energía mínima que debe contener dicha bacteria para ser catalogada como algo vivo. Luego supuso que todos los átomos que componen una *Escherichia coli* están sueltos en el volumen que ocupa dicha bacteria, y se preguntó cuál es la probabilidad de que una fluctuación del estado de equilibrio diera, fortuitamente, una *E. coli*. La respuesta fué:

$$p = 10^{-10^{-10}}$$

A continuación calculó la probabilidad de que se haya dado una fluctuación así desde que existe el Universo (unos 10^{18} segundos). Supuso que en el Universo hubiera unos 10^{100} átomos, que todos estuvieran ensayando continuamente para hacer una bacteria y que la duración mínima de una unión molecular debería ser de 10^{-16} segundos para ser considerada "estable" y formar una estructura biológica. Dicho de otro modo, que los 10^{100} átomos del Universo podrían haber ensayado una vez cada 10^{-16} segundos durante 10^{18} segundos. El número de ensayos fué entonces: $10^{100} \times 10^{18} \times 10^{16}$, o sea unas 10^{134} veces. Es decir, que hubo unas 10^{134} oportunidades de obtener algo cuya probabilidad es $10^{-10^{-10}}$. Eso quiere decir que la probabilidad de que se haya dado una simplísima *Escherichia coli* desde que el mundo es mundo es algo así como $10^{-99\ 999\ 999\ 866}$ o sea: imposible. Y esa era, por aquel entonces, la mejor teoría disponible para explicar científicamente el origen de la vida en la Tierra. Pensar que estamos hablando de hace apenas veinte años.

En estos tres ejemplos se muestran entonces algunas de las dificultades que se encontraron al tratar de entender la vida como proceso material regido por la Termodinámica. Como esta ciencia estudia las condiciones que pueden asumir los sistemas materiales y los cambios que les puedan ocurrir ya sea espontaneamente o como resultado de interacciones, parecería que los organismos fueran entes aparte y en su estudio se debieran dejar de lado las explicaciones científicas que rigen en el mundo material físico. Desde el punto de vista ideológico esto era terrible y conste que no estamos entrando siquiera a los desarrollos de Boltzmann, Szilard, Brillouin y otros grandes de la Termodinámica Estadística y de la Teoría de Información para tratar de entender la tremenda acumulación informativa en el código genético: toda la explicitación para hacer un elefante contenida en un hilito molecular de ADN que es visible sólo con poderosísimos microscopios. Sin embargo algo debía andar mal, pues a nadie se le ocurría que los procesos fundamentales de la vida no fueran clara y completamente explicados por la Fisicoquímica. La solución tomó muchos años y esfuerzos. A continuación veremos entonces algunas de las etapas que llevaron a los conceptos actuales.

El abandono de los modelos de equilibrio.

Los primeros pasos de la Termodinámica fueron dados con los modelos de equilibrio. Los de la Fisiología también: no sólo comunmente (y a veces incorrectamente) aplicamos ecuaciones válidas sólo para estados de equilibrio, sino que el concepto de homeostasis es central al enfoque biológico. Típicamente, los fisiólogos nos ocupamos de como hace el organismo para mantener constante su glucemia, su pH, su contenido de sodio, su volumen sanguíneo. Pero es claro que los modelos de equilibrio no llevan más que a imposibilidades como la que ilustré más arriba con el cálculo de Morowitz. Con esos enfoques no podríamos explicar el fenómeno más evidente de la vida en la Tierra: su evolución a organismos cada vez más complejos. Peor aún: equilibrio significa detención de todo trabajo. Si nosotros mantuvieramos realmente el equilibrio estaríamos muertos: caería nuestra temperatura hasta la del medio ambiente, cesarían de contraerse nuestros músculos, de circular nuestra sangre, esas "Incredibly improbable accumulation of matter" de que hablaba M. Jacobs se disiparían y nuestro equilibrio consistiría lisa y llanamente en la muerte, es decir, la no-vida. Por eso se pasó a contemplar los modelos de Estados Estacionarios.

Los estados estacionarios: La Termodinámica de los Procesos Irreversibles.

Este enfoque termodinámico de los sistemas biológicos tuvo su mayor desarrollo en el estudio de las membranas biológicas (ver Cereijido y Rotunno, 1970). Se basó en dos consideraciones: 1) *la suposición de que los sistemas biológicos no están en equilibrio sino en estado estacionario.* Una neurona, por ejemplo, no mantiene constante su nivel de potasio porque su membrana sea impermeable a este ión, sino porque hay mecanismos complejísimos que mantienen su flujo de entrada igual al flujo de salida. Este proceso ocasiona gastos de energía metabólica que, también, se mantienen estables gracias a una constelación de enzimas que regulan una maraña de reacciones químicas en las que el influjo de reactivos está celosamente equilibrado con el eflujo de productos y catabólitos. La constancia no se debe entonces a un equilibrio sino a un estado estacionario. En esta situación se disipa energía y se produce entropía (S).

El primer paso fué entonces tomar el cambio total de entropía en dichos procesos (dS) y discriminar entre el cambio debido a la producción interna (d_iS) y el debido a la interacción con el medio (d_eS) (Guggenheim, 1950).

$$dS = d_iS + d_eS$$

utilizando la energía libre de Gibbs (G) se llegó a la ecuación:

$$d_iS = -\frac{dG}{T}$$

Puesto que se estaba estudiando procesos, se pasó a estudiar el *cambio* de energía libre de Gibbs, es decir, se introdujo el *tiempo* en las ecuaciones:

$$\frac{d_iS}{dt} = -\frac{1}{T}\frac{dG}{T}$$

como los organismos biológicos son fundamentalmente máquinas químicas, un punto importante fué expresar los procesos químicos como *flujos.* (Ver de Donder and van Rysselberghe, 1936 y Lewis and Randall, 1961). De modo que, en el caso de las reacciones se tomó:

$$T\frac{d_iS}{dt} = A\frac{d\xi}{dt}$$

donde $d\xi/dt$ es la velocidad de la reacción (reaction rate) y A es una "fuerza termodinámica" derivada de la afinidad. De manera que tanto los flujos de sustancias, como la marcha de las reacciones químicas del metabilismo pudieron ponerse ahora en la forma de productos de flujos (J_i) y fuerzas (X_i). La suma del producto de todos los flujos y fuerzas presentes en un sistema, son iguales a lo que Lord Rayleigh (1973) llamó "función disipación" (σ)

$$\sigma = \Sigma JX$$

o también

$$\sigma = T\frac{d_iS}{dt} = J_iX_i$$

veremos en un momento qué ventajas trajo este tratamiento de flujos.

La segunda consideración en el enfoque biológico de los sistemas biológicos fué considerar que *cada flujo presente en un sistema depende de todas las fuerzas presentes.* Clasicamente se tomaba cada flujo como dependiente de su *fuerza conjugada*: el flujo hidráulico depende del gradiente de presión (ΔP), el de electricidad de la diferencia de potencial ($\Delta\Psi$), el de calor del gradiente térmico (ΔT), etc. Onsager (1931) en cambio, desarrolló un tratamiento en el que un flujo determinado puede ser producido por su fuerza clásica y además por otras que son, en realidad, las fuerzas conjugadas a otros flujos. Los ejemplos más familiares de estos fenómenos son por supuesto los de termoelectricidad (ej: el efecto Peltier), donde un gradiente térmico da origen a una corriente, y una diferencia de potencial trae aparejado un flujo de calor. Se tiene entonces:

$$
\begin{aligned}
J_1 &= L_{11}X_1 + L_{12}X_2 + \ldots + L_{1n}X_n \\
J_2 &= L_{21}X_1 + L_{22}X_2 + \ldots + L_{2n}X_n \\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
J_n &= L_{n1}X_1 + L_{n2}X_2 + \ldots + L_{nn}X_n
\end{aligned}
$$

o, en forma más condensada:

$$J_i = \sum_{j=i}^{n} L_{ij}X_j$$

donde los J_i son flujos, los L_{ii} son los cieficientes que asocian cada flujo a su fuerza conjugada X_i, y los L_{ij} son los *coeficientes de acople* a una fuerza distinta (X_j).

Estos coeficientes de acople son los que explican, por ejemplo, el acople del flujo de sodio a través de una membrana de piel de rana a las reacciones químicas del metabilismo celular, y que terminaron con la aparente paradoja del Ejemplo II que les diera al comienzo de esta presentación, cuando no se entendía que las reacciones químicas del metabolismo originaran un flujo neto de sodio(*).

(*) Dicho sea de paso: otro de los puntos que se aclaró es que la aplicación del Principio de Curie suponía incorrectamente que la piel de rana era un medio isotrópico y que las reacciones químicas eran escalares (cosa cierta a escala macroscópica pero no cuando las moléculas están todas y cada una orientadas y las adenosintrifosfatasas tienen sus sitios de Na y ATP mirando para el mismo lado).

Este enfoque cobró tanto o más interés entre los biólogos cuando Ilia Prigogine demostró que si un sistema está cerca de su equilibrio, los flujos y fuerzas se combinan de tal modo que tienden a minimizar la prducción de entropía (σ). Esto quería decir que, ante una perturbación, aparecerían en el sistema gradientes orientados de tal forma, que los flujos originados tenderían a disipar dicho gradiente y restablecer el equilibrio, y que esta tendencia sería tanto mayor cuando mayor fuera el grado de alejamiento (δ). (Fig. 2).

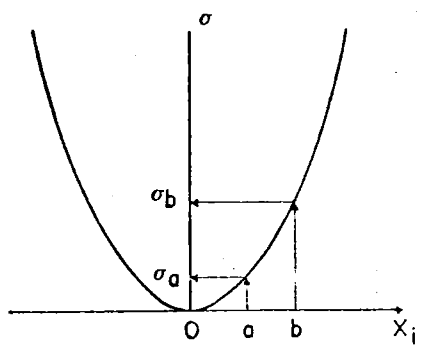


Fig. 2. Relación entre la producción de entropía por unidad de tiempo (σ) y las fuerzas (X_i) provocadas por el alejamiento del equilibrio.

El parecido de estas situaciones con la regulación biológica (homoestasis) fué realmente entusiasmante.

Pero el entusiasmo no duró mucho. Las nuevas ecuaciones ayudaban a estudiar mejor los fenómenos de estado estacionario cercanos al equilibrio pero, tal como lo ilustramos en el caso de la *Escherichia coli*, los sistemas biológicos están alejadísimos del equilibrio, tan alejados, que no se podría llegar a ellos a partir de una fluctuación de estado de equilibrio. Pero la situación era peor todavía: aun en reposo los sistemas biológicos no son estables: el corazón late, las neuronas disparan potenciales de acción regularmente, las células entran en crisis que las lleva a dividirse con regularidad, los organismos no son eternos: se mueren después de una duración bastante característica. Además, no seamos tan simplistas, si todo se mantuviera en un estado estacionario no evolucionaríamos hacia formas más complejas, la fecundación no sería seguida del embrión, el feto, el niño y el adulto; la vida no hubiera dado origen a monos y hombres.

Crisis y desequilibrios.

Como vimos en el ejemplo de la *E. coli*, la vida no está cerca del equilibrio, sino muy lejos. Justamente eso da vueltas nuestros conceptos: la vida se desarrolla en los desequilibrios al borde de las crisis. Una pequeñísima irritación de nuestra nariz provoca un estornudo con un trabajo desproporcionado a la energía del estímulo; una ínfima cantidad de noradrenalina trae cambios circulatorios exagerados en comparación con la energía de la interacción de dicha sustancia con los receptores. Esa es una característica fundamental del estado biológico: nuestros sistemas parecen estar siempre al borde de un cambio drástico, siempre al borde del colapso, siempre esperando la gota de agua que derramará el vaso. Es que la vida parece alejarse siempre del estado de equilibrio y llevar a sus sistemas y subsistemas al borde de una crisis, donde un pequeñísimo cambio o grado de alejamiento, ya no tenderá a restablecer el equilibrio, sino a provocar un cambio a otro estado. Prigogine hizo notar que estas situaciones tienen particularidades apasionantes. Tomemos el ejemplo que él mismo dió: el fenómeno de Bénard.

Si ponemos una caja de Petri con agua sobre una fuente de calor homogéneo, llegará un momento que el agua se ordenará en celdas hexagonales como la esquematizada en la Figura 3. Este ordenamiento es impensable cerca del equilibrio, donde el movimiento de las moléculas es caótico. Pero al intercalar al sistema en un flujo de calor desde la fuente al agua y de ahí al medio, se obtiene esa estructura con toda regularidad. Lo que era improbable-imposible cerca del equilibrio, es una ley causal en una crisis. Prigogine hizo notar que dicho fenómeno se debe a la no-linealidad. En efecto, cuando estamos cerca del equilibrio la relación entre flujos y fuerzas es más o menos lineal pero, al alejarnos, la linealidad se pierde y se llega a puntos críticos. Hay fenómenos físicos que mantienen una cierta linealidad aun en situaciones muy alejadas del equilibrio (frente a grandes fuerzas) y otros que la pierden a poco de alejarse de este punto. Los procesos químicos son de este último tipo: llegan enseguida a puntos críticos e inestabilidades. Como los sistemas biológicos son máquinas químicas y están muy alejadas del equilibrio, están propensas a la crisis y los desequilibrios.

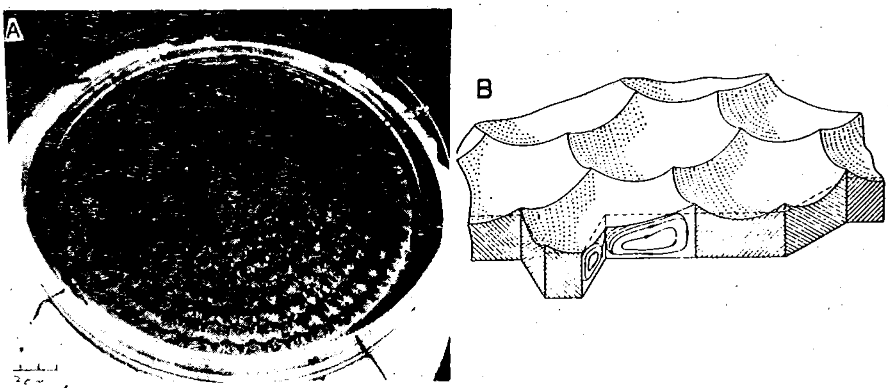


Fig. 3. A) Fenómeno de Bénard en un líquido uniformemente calentado desde abajo. B) Diagrama mostrando el movimiento convectivo en cada celda.

Esa situación tiene un aspecto interesante que, dicho sea de paso, también está ilustrado por el fenómeno de Benard. Una crisis significa que el sistema ya no tratará de restablecer el equilibrio primitivo. Los flujos y fuerzas que mencionábamos en la sección anterior ya no estarán orientadas a disipar los gradientes. La estructura misma del sistema cambiará y se tenderá a alcanzar otro equilibrio (Fig. 4). Gracias al cambio de estructura, el sistema tiende a un nuevo equilibrio donde disipará menos entropía. Prigogine hace notar que toda estructura es hija de un desequilibrio (y todo cambio de estructura también).

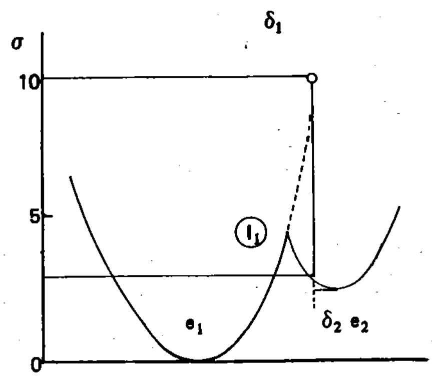


Fig. 4. Transición de un equilibrio (e_i) a otro (e_2). Cuando el sistema está en el estado indicado por la línea vertical, está alejado (δ_2) del equilibrio e_2, y por lo tanto su disipación de entropía (σ) es 3 unidades arbitrarias. Si no hubiera hecho la transición en la inestabilidad I_i, su producción de entropía hubiera seguido aumentando (línea de puntos), estaría muy alejado (δ_1) de su equilibrio e_1, y produciría 10 unidades arbitrarias.

Como esto implica un cambio conceptual muy importante, sobre todo para la Biología (ya veremos porque), se hicieron esfuerzos por desarrollar las bases teóricas de dichos fenómenos, y se trataron de obtener fenómenos de transición no ya en sistemas hidrodinámicos como en el caso de Bénard, sino en sistemas químicos. Así hoy son famosos los trabajos del grupo de Prigogine y los experimentos de Zhabotinski y otros (Fig. 5). Se obtuvieron procesos químicos que se mantienen gracias a un *flujo* contínuo de reactivos y productos y que, cuando se los perturba y aleja del equilibrio adoptan estructuras (formas espaciales) características y que tienen una relación estricta con el tipo de perturbación sufrida (tienen historia: memoria de la perturbación).

Fig. 5. **Estructuras disipativas en una caja de Petri. El líquido, de 0.1 cm de profundidad, era homogéneo al principio pero, al iniciar la reacción de Belusov-Zhabotinski, se ordenó espontaneamente en espirales que indican una segregación de los componentes moleculares participantes.**

Aquí deberíamos mencionar a Morowitz, quién formuló un teorema para un sistema intercalado en el flujo de energía entre una fuente a un sumidero. Todo flujo de energía —dijo— origina en ese sistema intermedio por lo menos un ciclo material. Podemos ilustrar esto fácilmente con el ciclo del agua en la Biósfera: el paso de energía solar (Sol → Biósfera → espacio extraterrestre) hace que el agua se evapore, se formen nubes, llueva y nieve, se formen ríos y el agua vuelva al mar. Notemos que no es el *suministro* de energía lo que crea este ciclo del agua y su organización (nubes, hielo, agua), sino el *flujo*. Tan importante es la provisión de calor como su disipación. De lo contrario el agua se evaporaría totalmente (o se congelaría) y el ciclo cesaría.

¿Y qué tiene que ver todo esto con la vida? Mucho. Ya hemos dicho que los organismos son sistemas químicos, que los sistemas químicos llegan fácilmente a crisis y que las crisis originan estructuras en las que el sistema pasa a tener ordenamientos y propiedades distintas e imposibles a nivel del equilibrio.

Cerca del equilibrio la tendencia es a la desorganización y al caos. Lejos del equilibrio los sistemas se organizan. Esta organización no es hija de la "negaentropía" ni alguna de las otras rarezas semánticas que se han usado: simplemente se debe al flujo de energía y materia a su través. Ya Schroedinger había mostrado que la tremenda organización biológica se hacía a expensas de un crecimiento de entropía mucho mayor en el medio. Ahora sabemos que ese gasto energético que complejiza más y más al mundo biológico, es producto del *flujo* de energía solar y no del mero suministro de energía. Si a esto agregamos que un sistema no tiene solamente *una* crisis, con *una* transición a *una* estructura nueva, sino que puede seguir siendo perturbado y cambiar a otras y otras, tenemos entonces un marco conceptual que nos permite comprender que el flujo de materia y energía pueda resultar en una serie de estructuras de complejidad creciente, Por otra parte esos cambios cíclicos tienen dimensiones, formas y duraciones determinadas, y estas son también características de los sistemas biológicos.

Pero a estas alturas estamos hablando de avances recientes de la Termodinámica y de la Biología. Los desarrollos teóricos y la información experimental están aún en pañales y va a requerirse un enorme esfuerzo de elaboración. Pero por lo menos ya no hay discrepancias. Por otro lado, sería adecuado resaltar que en el interín la Biología Molecular y la Genética han hecho avances enormes en descifrar la formación de las primeras moléculas-códigos (ADN, ARN) las primeras enzimas, y pronto comprenderemos no ya como se formaron sistemas químicos estructurados complejos y cíclicos (estructuras disipativas), sino y muy en particular aquellos sistemas que pueden ser considerados precursores de los primeros organismos en la Tierra. (Ver Manfred Eigen, 1971)

Agradezco a la Srita. Maricarmen De Lorenz la eficiente ayuda secretarial. Apoyado por donativos del PNCB del CONACyT.

REFERENCIAS

Caratheodory, C. Untersuchungen über die Grundlagen der Thermodynamik. Math. Annln., 1909, 67: 355.

Cereijido, M. Orden, equilibrio y desequilibrio: Una introducción a la Biología. Nueva Imagen, 1978.

Cereijido, M. and Rotunno, C.A. Introduction to the study of biological membranes. Gordon & Breach (London) 1970.

de Donder, T. and van Rysselberghe, P. Thermodynamic theory of afinity. Stanford University Press, Standford 1936.

Eigen, M. Molecular self-organization and the early stages of evolution. Q. Rev. Biophys., 1971, 4: 149.

Galeotti, G. Ricerche di elettrofisiologia secondo i criteri dell' elettrochimica. Z. Physiol., 1907, 6: 99.

Glandsdorff, P. and Prigogine, I. Thermodynamic theory of structure, stability and fluctuations. Wiley, Nueva York, 1971.

Guggenheim, E.A. Thermodynamics. Amsterdam, North Holland, 1950.

Jacobs, M.H. Diffusion Processes. Ergeb. Biol. 1935, 12:1.

Katchalsky, A. and Curran, P.F. Nonequilibrium thermodynamics in biophysics. Cambridge, Mass. Harvard University Press.

Lewis, G.N. and Randall, M. Thermodynamics. McGraw-Hill, New York, 1961.

Morowitz, H.J. Energy flow in biology: biological organization as a problem in thermal physics. Academic Press, Nueva York, 1968.

Onsager, L. Reciprocal relations in irreversible processes. Phys. Rev. 1931, 37: 405.

Onsager, L. Reciprocal relations in irreversible processes II. Phys. Rev. 1931, 38: 2265.

Prigogine, I. Thermodynamics of irreversible processes. Ch. Thomas Press, Springfield, 1955.

Prigogine, I. Temps, structure et entropie. Bull. Acad. R. Belg. (Cl. Sci.) 1967, 53: 273.

Prigogine, I. and Nicolis, G. Biological order, structure and instabilities. Quat. Rev. Biophysics, 1971, 4: 107.

Rayleigh, Lord. On convective currents in a horizontal layer of fluid, when the higher temperature is on the under side. Philos. Mag., 1916, **32**: 529.

Schrödinger, E. What is life? Cambridge University Press, 1944.

Turing, A.M. The chemical basis of morphogenesis. Phil. Trans. Roy. Soc. B., 1952, **237**: 37.

Zhabotinski, A.M. Periodic course of oxidation of malonic acid in solution. Biophysics, 1964, **9**: 306.

QUE SON LAS MATEMATICAS APLICADAS

Simón Mochón

Introducción

Matemáticas aplicadas es una disciplina que trata de entender y explicar un fenómeno ya sea natural, social o de otro tipo por medio de métodos matemáticos. Aunque en los últimos años ha habido un resurgimiento de ésta rama de las matemáticas y se ha empezado a aplicar a áreas como biología y sociología, este enfoque de las matemáticas no es de ninguna manera nuevo y grandes matemáticos como Euler, Lagrange, Gauss, Riemann, Fourier y posiblemente Einstein entre otros, pueden ser catalogados como matemáticos aplicados.

El matemático aplicado al igual que el matemático puro está interesado en desarrollar teorías matemáticas, pero la actitud del matemático aplicado es diferente: El matemático aplicado está fundamentalmente interesado en construir teorías que describan cierto fenómeno y extenderlas en direcciones que tengan conexión con el problema que se está investigando. Además, el objetivo del matemático aplicado es la solución del problema. Para esto, tiene que desarrollar métodos matemáticos que por lo general es muy difícil de probar su validez. Por esta razón, el énfasis que da a estas teorías no es riguroso sino que es intuitivo. Esta diferencia en actitud puede ilustrarse mejor con un ejemplo: mientras que el matemático puro está interesado en probar si una serie converge o diverge y asigna una etiqueta a las series divergentes que dice "sin valor alguno", el matemático aplicado usa frecuentemente series divergentes para llegar al resultado correcto de un problema y en cambio las "series convergentes" a veces le resultan practicamente inservibles debido a que la convergencia es muy lenta (Trata de calcular por ejemplo e^{-10} con su serie de Taylor).

Otra diferencia fundamental entre un matemático puro y uno aplicado es que el matemático aplicado se ve forzado generalmente a hacer aproximaciones y simplificaciones ya que los problemas que resultan en su investigación son no lineales y por lo cual muy difíciles de resolver exactamente. El siguiente ejemplo muestra la manera de atacar un problema sencillo desde el punto de vista de un matemático aplicado. Supongamos que queremos encontrar las raíces de la ecuación $X^3 - X + \frac{1}{10} = 0$. Aunque hay fórmulas explícitas para las raíces de polinomios de tercer y cuarto grados, estas son muy complicadas (para polinomios de grado mayor o igual a cinco sabemos que no existe ninguna fórmula). El matemático aplicado en estas situaciones está interesado en derivar métodos lo suficientemente generales que puedan aplicarse a una gama muy grande de ecuaciones de cualquier grado. Desde luego hay métodos numéricos que se pueden usar en estos casos, pero mi interés aquí es el de presentar el tipo de análisis que desarrolla un matemático aplicado. Dos caminos son mostrados a continuación.

Método #1.- Lo más obvio que se puede hacer con la ecuación dada es despejar equis y escribir: $X = X^3 + \frac{1}{10}$. Una raíz de esta ecuación es un número que al elevarlo a la tercera potencia y sumarle un décimo resulta otra vez el mismo número. Lo que podemos tratar entonces es substituir un número X_0 en el miembro derecho de la ecuación y obtener otro número X_1 como resultado. Substituir ahora X_1 para obtener X_2, etc. En general $X_n = X_{n-1}^3 + \frac{1}{10}$. Si esta sucesión converge, será a una de las raíces. A continuación se enlistan cuatro sucesiones obtenidas de esta manera empezando con los números -1, 0, 0.9 y 1.

-1	0.	0.9	1.
-0.9	0.1	0.829	1.1
-0.629	0.1010	0.6697	1.431
-0.1489	0.1010	0.4004	3.030
0.0967		0.1642	27.928
0.1009		0.1044	↓
0.1010		0.1011	diverge
0.1010		0.1010	

Obviamente 0.1010 es una raíz. La última sucesión muestra que no todo número inicial produce una sucesión convergente. ¿Cómo podemos encontrar las otras dos raíces?. No ayuda tomar más números iniciales, ya que las sucesiones generadas o divergen o convergen a 0.1010. Lo que podemos hacer ahora es despejar la "otra" equis y escribir la ecuación como: $X = \sqrt[3]{X - \frac{1}{10}}$. Usando el mismo método anterior $X_n = \sqrt[3]{X_{n-1} - \frac{1}{10}}$ encontramos las sucesiones:

-1.	0.	1.
-1.0323	-0.4642	0.9655
-1.0423	-0.8263	0.9530
-1.0453	-0.9748	0.9484
-1.0463	-1.0243	0.9467
-1.0466	-1.0398	0.9461
-1.0466	-1.0446	0.9458
		0.9457

Por lo tanto las otras dos raíces son: -1.0466 y 0.9457. En este caso cualquier número inicial produce una serie que converge a una de estas dos raíces. (Esto se puede comprobar gráficamente).

Método #2.- Otra manera de proceder es suponer que el término $\frac{1}{10}$ en la ecuación $X^3 - X + \frac{1}{10} = 0$ es "pequeño" y por lo cual lo podemos eliminar sin alterar "mucho" el balance de los términos. Así, una aproximación de la ecuación original es la ecuación $X^3 - X = 0$ que tiene como raíces: -1, 0, 1. Esto desde luego no es suficiente y tenemos que encontrar un método que mejore esta primera aproximación. Estas dos ecuaciones se pueden escribir en general como $X^3 - X + \varepsilon = 0$, en donde la ecuación original tiene $\varepsilon = \frac{1}{10}$ y la ecuación aproximada tiene $\varepsilon = 0$. Las raíces de esta ecuación más general van a depender del parámetro ε y por lo tanto podemos expresarlas en series de potencias en ε de la siguiente manera:

$$r = a_0 + a_1\varepsilon + a_2\varepsilon^2 + a_3\varepsilon^3 + \ldots$$

sustituyendo esta forma en la ecuación general, obtenemos ecuaciones para los coeficientes a_0, a_1, a_2,... etc.:

$$a_0^3 - a_0 = 0 \qquad \text{define } a_0$$

$$3a_0^2 a_1 - a_1 + 1 = 0 \qquad \text{define } a_1$$

$$3a_0^2 a_2 - a_2 + 3a_0 a_1^2 = 0 \qquad \text{define } a_2$$

$$\vdots \qquad\qquad \vdots$$

La primera ecuación $a_0^3 - a_0 = 0$ tiene como soluciones $a_0 = -1,0,1$, que corresponden a la primera aproximación de las tres raíces. Es importante hacer notar que en este método es necesario que la ecuación que resulte al poner $\varepsilon = 0$ sea más fácil de resolver que la ecuación original. Tomando $a_0 = -1$ podemos calcular $a_1, a_2, \ldots$, etc. y encontrar la serie correspondiente a una de las raíces:

$$r = -1 - \frac{1}{2}\varepsilon + \frac{3}{8}\varepsilon^2 - \frac{1}{2}\varepsilon^3 + \ldots$$

si ahora tomamos el valor de ε como $\frac{1}{10}$ la primera aproximación de la raíz resulta -1 (primer término de la serie); la segunda aproximación será -1.05 (sumando los dos primeros términos) y mejores aproximaciones serán -1.0463 (tres términos) y -1.0467 (cuatro términos). Si tomamos ahora $a_0 = 0$ y 1 y seguimos el mismo método obtendremos las otras dos raíces. Este método que es llamado "teoría de perturbaciones" es muy importante y tiene aplicaciones en problemas mucho más complejos en ecuaciones diferenciales ordinarias y parciales.

Actividades de un matemático aplicado.

En la investigación de un problema hay tres facetas en las que un matemático aplicado puede estar interesado. La primera es la construcción o extensión de un modelo matemático que describa cierto fenómeno. La segunda es la aplicación y solución de este modelo a un problema en particular. Aquí puede utilizar métodos de solución ya existentes o desarrollar nuevos métodos apropiados al problema investigado. La tercera es la de justificar y generalizar estos métodos. Esta última faceta es de interés tanto a matemáticos aplicados como a matemáticos puros. En realidad muchas de las ramas en matemáticas puras han surgido de esta manera. Por ejemplo, la teoría de las distribuciones surgió de la necesidad de justificar el uso de la delta de Dirac que es tan importante en la solución de ecuaciones diferenciales.

El camino que un matemático aplicado sigue en la construcción y solución de un modelo matemático puede darse en los siguientes pasos:

1. Identificar un problema de interés y establecer la finalidad de su estudio. Esta puede ser entre otras: comprobar o desmentir una hipótesis que haya sido propuesta basada en observaciones experimentales; entender mejor un proceso; enmarcar en una manera ordenada y precisa una serie de datos acumulados de un cierto proceso; o simplemente extender los resultados obtenidos por otros investigadores.

2. Estudiar todo lo referente al problema. Datos experimentales; trabajos previos, etc.

3. Aislar las características importantes del proceso estudiado y así simplificar el problema. Esto desde luego sacrifica algo del realismo y exactitud del modelo pero en cambio se gana en claridad y facilidad de solución.

4. Decidir que área de las matemáticas es adecuada para describir el proceso (probabilidad, ecuaciones diferenciales, etc.) y formular el problema en términos matemáticos. En algunos casos esta formulación ya ha sido previamente establecida; así por ejemplo, en la predicción del tiempo, las ecuaciones de la mecánica de fluidos son usadas. En otros casos, analogía con un proceso similar puede ayudar para derivar las ecuaciones pertinentes. En el resto de los casos, las ecuaciones son derivadas de principios básicos. Es aquí donde la experiencia en la construcción de modelos es muy importante.

5. Encontrar la "solución" del modelo matemático construído. En general, no es posible encontrar la solución exacta del problema. En estos casos hay que recurrir a simplificaciones y aproximaciones para hacer el modelo manejable.

6. Comparar los resultados del modelo con datos experimentales y analizar las predicciones del modelo.

7. Indicar claramente las limitaciones del modelo y revaluar el problema.

8. Si es necesario y siempre lo es, mejorar o corregir el modelo ya sea incluyendo más características que se piensen importantes o tratando de suprimir simplificaciones que fueron hechas previamente.

Esto generalmente se hace ya sea con métodos analíticos que pueden mejorar una solución aproximada o con métodos numéricos (usando una computadora). Es importante hacer notar qué métodos numéricos, en contra de la creencia general, pueden dar una solución incorrecta del problema. Esto es particularmente cierto en la solución de ecuaciones diferenciales parciales. Por lo tanto es aconsejable comprobar la solución numérica con un método analítico cuando esto sea posible. Este punto puede ilustrarse con un ejemplo muy sencillo: la ecuación $\varepsilon y' + y = 0$ con condición inicial $y(0) = 0$ tiene como solución $y = 0$. Para obtener la solución numérica podemos usar el método de Euler y aproximar la derivada con: $y_n' \doteq (y_{n+1} - y_n)/h$. La ecuación entonces se transforma en $y_{n+1} = (1 - \frac{h}{\varepsilon})\, y_n$. Si ε es pequeño comparado con h, un error de truncación de la máquina puede amplificarse enormemente y dar una solución incorrecta.

Estos 8 puntos son ilustrados a continuación con un ejemplo. Es conveniente recalcar primero que esta lista es sólo un bosquejo del camino a seguir. Diferentes investigadores se concentran en diferentes aspectos del modelo matemático. Gran parte usan modelos bien establecidos y que son suficientemente generales para aplicarlos a varios problemas de interés (las ecuaciones de la mecánica de fluidos son por razones obvias, muy usadas). Otros se concentran en los aspectos teoréticos que han sido dejados sin resolver. Así, la gama de aplicación de un matemático aplicado puede ser desde un 0.1% hasta un 100%.

Ejemplo.

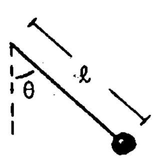
1. Problema: Predecir el movimiento de un péndulo y su período de oscilación.

(Este es un problema cuya solución es bien conocida y por lo cual puede parecer poco atractivo de analizar. Por otro lado, esta experiencia del problema es una ventaja, ya que hará la exposición más entendible).

2. Estudiar su comportamiento: El período es "independiente" de la amplitud en oscilaciones pequeñas e incrementa en oscilaciones grandes. Otro tipo de oscilaciones existen en donde el péndulo pasa por arriba.

3. Decidir las características importantes: La masa de la esfera es la que determina el movimiento. La masa del hilo puede despreciarse. Fricción no es importante. Una mosca volando cerca, no disturva el movimiento apreciablemente.

4. En este caso el modelo matemático es bien conocido: Ley de Newton, F = ma (Ecuacion Diferencial).



$$\ddot{\theta} + \frac{g}{\ell}\,\mathrm{sen}\,\theta = 0$$

5. No es posible resolver esta ecuación exactamente. Simplificación: Interesado en oscilaciones pequeñas. En términos matemáticos esto significa aproximar sen $\theta \doteq \theta$ (este proceso se llama linearización y es muy usado en matemáticas aplicadas). Con esto, el modelo simplificado es:

$$\ddot{\theta} + \frac{g}{\ell}\theta = 0$$

que tiene como solución una combinación lineal de senos y cosenos con período $T = 2\pi\sqrt{\frac{\ell}{g}}$.

6. Predicciones: Período independiente de la amplitud y proporcional a la raíz cuadrada de la longitud del péndulo. El péndulo oscilará para siempre con amplitud constante.

7. Limitaciones: Resultados válidos sólo para amplitudes pequeñas en las que no haya fricción. No se puede aplicar por ejemplo al movimiento de un columpio en donde fricción es importante.

8. Mejorar los resultados:

a) Estudiar la ecuación original $\ddot{\theta} + \frac{g}{\ell}\,\mathrm{sen}\,\theta = 0$ con métodos aproximados. En este caso hay varios métodos que se pueden emplear:

i) Plano de fase.- En este método las soluciones de la ecuación diferencial son representadas gráficamente por trayectorias en un plano que tiene como coordinadas la variable θ y su derivada $\dot{\theta}$. Para esto se define otra variable $V = \dot{\theta}$ y así se transforma nuestra ecuación diferencial de segundo orden en un sistema de dos ecuaciones de primer orden. El análisis para obtener el plano de fase para este ejemplo está fuera del objetivo de esta plática. Este método es muy útil para analizar ecuaciones diferenciales no lineales de segundo y tercer orden.

ii) Aproximaciones sucesivas.- Como ya obtuvimos la solución de la ecuación $\ddot{\theta}_1 + \frac{g}{\ell}\theta_1 = 0$ donde θ_1 denota la solución aproximada válida para amplitudes pequeñas podemos reescribir nuestra ecuación original de la siguiente manera

$$\ddot{\theta} + \frac{g}{\ell}\theta = \frac{g}{\ell}(\theta - \operatorname{sen}\theta)$$

en esta forma hemos pasado el término no lineal ($\operatorname{sen}\theta$) a la derecha que es la razón de la dificultad y hemos puesto en la izquierda la forma de la ecuación que podemos resolver. Podemos ahora proponer como una "mejor" aproximación de nuestro problema la función θ_2 que satisface la ecuación

$$\ddot{\theta}_2 + \frac{g}{\ell}\theta_2 = \frac{g}{\ell}(\theta_1 - \operatorname{sen}\theta_1)$$

y así continuar nuestro procedimiento. En general la enésima aproximación estará dada por:

$$\ddot{\theta}_n + \frac{g}{\ell}\theta_n = \frac{g}{\ell}(\theta_{n-1} - \operatorname{sen}\theta_{n-1})$$

En la primera etapa, la de encontrar la solución del problema, el matemático aplicado no está interesado en <u>probar</u> que ésta sucesión de funciones $\theta_1, \theta_2, \ldots$ converge. Su interés es comprobar que este método efectivamente le está proporcionando mejores aproximaciones en cada iteración. Esto desde luego no es tan fácil, pero intuición, corroboración con el problema u otro método pueden ser de gran ayuda. Además puede tener el respaldo de saber que el método propuesto tiene validez o fué útil en otros problemas.

iii) Métodos numéricos.- Este método, aunque muy valioso es a veces poco atractivo ya que los resultados están dados en listas y listas de números que es difícil de analizar y obtener una idea completa de la solución del problema.

b) Si ahora estamos interesados en investigar un problema más real como por ejemplo el movimiento de un columpio, sabemos que el columpio se detendrá después de cierto tiempo. En este caso fricción es importante y tenemos que modificar nuestro modelo anterior para que incluya esta propiedad. Esto se puede hacer añadiendo un término que represente la fuerza de fricción. Un nuevo modelo sería:

$$\ddot{\theta} + f\dot{\theta} + \frac{g}{\ell}\,\text{sen}\,\theta = 0$$

A este nuevo problema se le aplicarán nuevamente todos los procedimientos señalados anteriormente.

Este problema del péndulo es interesante porque muchos conceptos y métodos usados en matemáticas aplicadas pueden ser ilustrados. Por ejemplo, la pregunta: ¿por qué nunca vemos un péndulo "colgado" hacia arriba y siempre está hacia abajo? introduce el concepto de estabilidad que es fundamental en matemáticas aplicadas.

Extensión de las matemáticas aplicadas.

El trabajo de un matemático aplicado es el de construir puentes que conecten las matemáticas con otras ciencias. Una pregunta importante en este respecto es: ¿de dónde a dónde se pueden construir estos puentes? o más explícitamente, ¿qué área de las matemáticas son aplicables? y ¿qué problemas científicos pueden formularse matemáticamente?

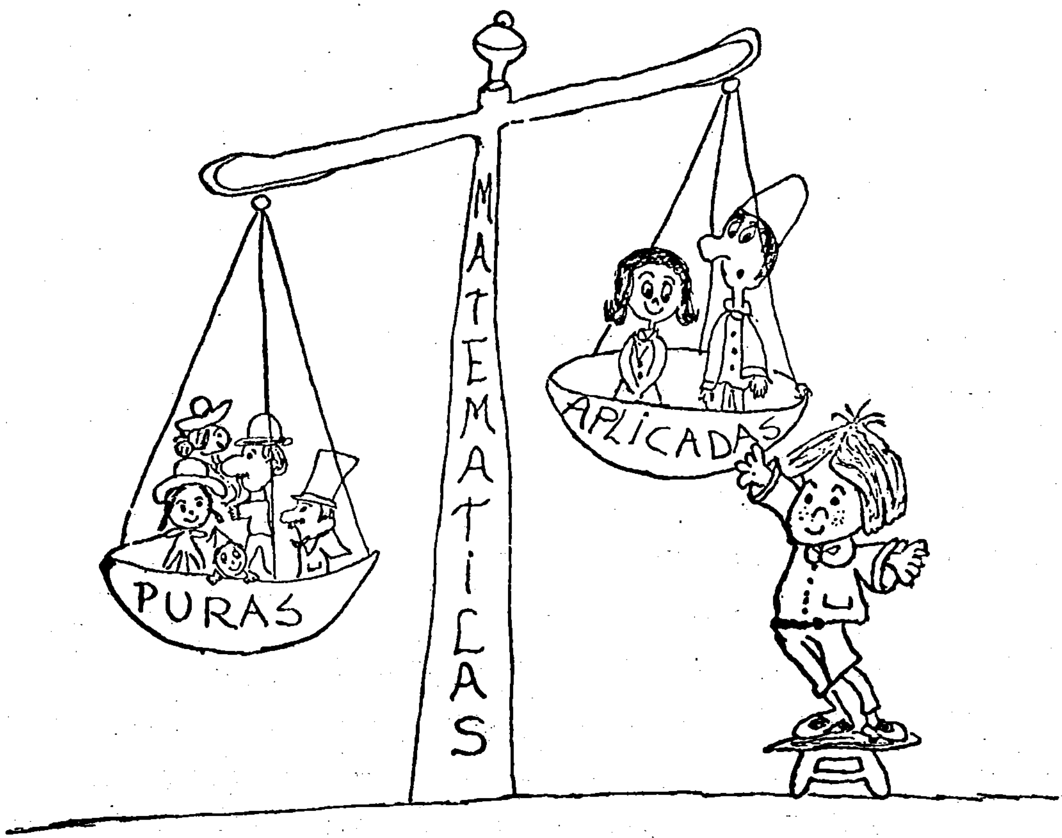
Tradicionalmente se piensa a probabilidad y estadística como las ramas aplicadas de las matemáticas. Otras áreas de las matemáticas han sido aplicadas pero en un grado mucho menor. Ecuaciones diferenciales es la excepción. Esta área ha sido usada en una infinidad de problemas científicos ya que se puede decir que ecuaciones diferenciales es la forma natural de expresar estos problemas. Es bien conocido, por ejemplo, que muchos fenómenos naturales en la física y química han sido descritos por medio de ecuaciones diferenciales. En general, modelación matemática está centrada en ecuaciones diferenciales. Es por esto que un matemático aplicado debe tener un conocimiento profundo de esta rama de las matemáticas.

La segunda pregunta es muy extensa y por lo cual muy difícil de contestar completamente. En ella están contenidas preguntas como: ¿puede ponerse cualquier problema en un contexto matemático? o ¿es posible que la influencia de las matemáticas llegue a áreas como Historia o Arqueología?. Desde luego, muchos problemas requieren de experimentación para encontrar su solución. Pero casi todos los problemas (excepto por un conjunto de medida cero) pueden ser ayudados de una manera u otra por las matemáticas. La cantidad e importancia de esta ayuda dependerá en gran parte de la experiencia, interés y agudez del investigador. En este momento lo único que puedo hacer es dar una lista pequeña para mostrar la variedad de problemas que se han formulado matemáticamente. Modelos de la estructura y forma de galaxias, modelos del tráfico de automóviles, modelos de guerras, modelos económicos, modelos del crecimiento de árboles, modelos del músculo, modelos del crecimiento de tumores, modelos de caminar, etc.

Se puede pensar de estos ejemplos que matemáticas aplicadas incluye a todas las ciencias. En realidad matemáticas aplicadas son una parte de las matemáticas con un enfoque en otras ciencias.

Conclusiones.

En esta plática se dió un bosquejo de las actividades de un matemático aplicado. La intención es dar a conocer este "nuevo" campo de acción de un matemático. Aunque hay investigadores en matemáticas interesados en aplicaciones, la balanza entre matemáticas puras y aplicadas está muy desequilibrada. Esto se refleja obviamente en la enseñanza de las matemáticas. Lo que se necesita es formar matemáticos interesados en la investigación y enseñanza de las diferentes áreas de matemáticas aplicadas. Entre otras, se deben cultivar las áreas: Análisis numérico y computación, ecuaciones diferenciales ordinarias y parciales, teoría de perturbaciones, probabilidad y esta-


MATEMATICAS
PURAS
APLICADAS

dística, mecánica de sólidos y fluidos (mecánica de lo continuo), aplicaciones a la biología y medicina, teoría de gráficas, investigación de operaciones, etc. Las áreas a escoger dependerán del departamento interesado pero es importante recordar que el énfasis que se debe dar es dirigido a aplicaciones.

Se deben ofrecer cursos a nivel de carrera, maestría y doctorado en estas áreas para dar al estudiante de matemáticas una idea general de las ramas de las matemáticas que existen y así darle la oportunidad de escoger entre ellas. La educación de matemáticas aplicadas no debe dejarse hasta el nivel graduado porque se perderá el interés y el conocimiento básico del estudiante en esta rama. No solamente se deben empezar con nuevos cursos en matemáticas aplicadas sino que también es conveniente modificar cursos básicos como cálculo, ecuaciones diferenciales, variable compleja, haciéndolos más aplicados y seguramente más interesantes.

Una pregunta que seguramente estará en la mente de los estudiantes de matemáticas es, ¿qué es más difícil, matemáticas puras o aplicadas?. La complejidad en cada una de ellas es la misma ya que ésta está dictada por grandes investigadores en el área. Sin embargo, hay algunas diferencias. En matemáticas aplicadas, muchas veces el método de solución de un problema es sugerido por el problema real que se está investigando. Esta visualización dada por el problema real es siempre de gran ayuda al matemático aplicado. Por otro lado un matemático aplicado necesita tener un conocimiento muy amplio en varias áreas tanto teoréticas como aplicadas de las matemáticas y debe estar interesado en estudiar otras ramas científicas.


MATEMATICAS
PURAS
APLICADAS

Una diferencia importante es que un matemático aplicado tendrá más oportunidades de realizar su trabajo en el futuro ya sea, en la educación o fuera de ella. Matemáticas aplicadas proporciona al individuo métodos matemáticos, ideas y experiencia en desarrollar teorías en otras ramas científicas que le serán muy útiles tanto en empleos académicos como no académicos.

Quisiera terminar clarificando que matemáticas puras sirven también un propósito muy importante en el desarrollo científico y que estudiantes en matemáticas aplicadas deben aprender el rigor y formalismo de matemáticas puras al igual que el énfasis dado en matemáticas aplicadas. El mensaje que trato de comunicar es que estos dos tipos de razonamiento deben estar representados más uniformemente.

LA EXPONENCIAL DE UNA MATRIZ

Mauricio Gutiérrez

Consideraremos matrices cuadradas $(n \times n)$ de dos tipos

A = matriz con coeficientes numéricos.

$A = xB$, B matriz numérica.

Queremos definir en esta charla la exponencial e^A ó $\exp(A)$ de A y discutir algunos de los problemas que se presentan al tratar de calcularla.

La exponencial aparece tradicionalmente en cálculo como la inversa del logaritmo. Tal vez la primera definición intrínsica es la serie

$$e^a = 1 + a + \frac{1}{2} a^2 + \frac{1}{3!} a^3 \ldots$$

Si queremos una definición similar para e^A tenemos que tener una noción de límite entre matrices y ésta puede ser obtenida definiendo la norma de una matriz A como

$$\|A\| = \sup_{|v|=1} |Av|$$

donde v es un vector en R^n y $|\cdot|$ es la distancia euclídea. Para visualizar $\|A\|$, tenemos que $\{v \in R^n : |v| = 1\}$ es una esfera. Entonces los $\{Av : |v| = 1\}$ son un elipsoide con semieje mayor $\|A\|$.

Con esta noción la serie

$$e^A = I + A + \frac{1}{2}A^2 + \frac{1}{3!}A^3 + \dots \tag{1}$$

converge pues $\|I + A + \dots + \frac{1}{k!} A^k\| \leq \|I\| + \|A\| + \dots + \frac{1}{k!} \|A\|^k < \exp(\|A\|)$ y esta convergencia es absoluta.

CONSECUENCIAS ELEMENTALES.

1. Como la convergencia de (1) es absoluta, si $A = xB$ podemos diferenciar la serie término por término para obtener

$$\frac{d}{dx} (\exp(xB)) = B \exp(xB)$$

que caracteriza $\exp(xB)$.

2. Si $A_1A_2 = A_2A_1$ entonces

$$\exp(A_1 + A_2) = \exp(A_1)\exp(A_2)$$

pero esto es falso si A_1 y A_2 no conmutan.

En particular

$$\exp((x+y)B) = \exp(xB)\exp(yB).$$

3. Finalmente, $\exp(P^{-1}AP) = P^{-1}(\exp(A))P$ y esta propiedad nos da la primera indicación de cómo calcular $\exp(xB)$.

Si B es una matriz diagonal

$$B = \begin{bmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_n \end{bmatrix} = \mathrm{diag}[b_1,\dots,b_n]$$

entonces $\frac{1}{k!}x^kB^k = \mathrm{diag}\left[\frac{x^k}{k!}b_1^k,\dots,\frac{x^k}{k!}b_n^k\right]$ y por lo tanto

$$\exp(xB) = \mathrm{diag}[\exp(xb_1),\dots,\exp(xb_n)].$$

Entonces si $P^{-1}AP = B$,

$$\exp(xA) = P\exp(xB)P^{-1}.$$

EJEMPLO 1.

$$A = \begin{pmatrix} -49 & 24 \\ -64 & 31 \end{pmatrix}$$

Como

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & -17 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}^{-1}$$

(2) $$\exp(A) = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} e^{-1} & 0 \\ 0 & e^{-17} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}^{-1} =$$

$$\begin{pmatrix} -0.73 & 0.55 \\ -1.47 & 1.10 \end{pmatrix} \text{ (aprox.)}$$

¿Qué cálculos se necesitan para computar $\exp(A)$?

Primero, para diagonalizar A es necesario hallar

$$\det(yI - A) = (y+1)(y+17).$$

Por lo tanto, $A+I$ y $A+17I$ tienen rango 1 y se pueden encontrar soluciones no triviales a las ecuaciones

(a) $$(A+I)\vec{v} = 0 \quad y$$

(b) $$(A+17I)\vec{v} = 0$$

y estas son, para (a) $\vec{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$; y para (b) $\vec{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Esto puede hacerse por el método de Gauss o la regla de Cramer.

Lo que nos lleva a la expresión (2) que aun exige que invertamos una matriz.

Si tuvieramos un computador, ¿no sería mejor calcular un número inmenso de términos de (1) para encontrar $\exp(A)$? ¡No! Los errores de redondeo son gravísimos. Con exactitud de 10^{-7} una computadora trató el Ejemplo 1 por ese método. Después de 60 sumandos, los cinco primeros decimales cesaron de cambiar o sea, la serie "convergió" a menos de 10^{-6} en cada coeficiente. La respuesta fue

$$\begin{pmatrix} -22.25 & -1.43 \\ -61.49 & -3.47 \end{pmatrix} \quad !$$

Es importante, pués, aprender algunos métodos algebraicos para crear buenos algoritmos que calculen (1).

EJEMPLO 2. Sea

$$N = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Calcular $\exp(xN)$.

Como

$$N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

y $N^3 = 0$, la serie (1) es una suma finita y

$$\exp(xN) = I + xN + \frac{x^2}{2} N^2 = \begin{pmatrix} 1 & x & \frac{x^2}{2} \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}$$

El ejemplo canterior es más general de lo que parece; en efecto, dada una matriz A (3×3), es siempre posible encontrar un cambio de base P tal que $P^{-1}AP = A_k$, $k = 1,2$ ó 3 y donde

$$A_1 = \mathrm{diag}(a,b,c), \quad A_2 = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 1 \\ 0 & 0 & b \end{pmatrix}, \quad A_3 = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 1 \\ 0 & 0 & a \end{pmatrix}$$

donde a,b,c no son necesariamente distintos.

Si $k = 1$, la propiedad 3 nos dice que

$$\exp(xA) = P\,\mathrm{diag}(e^{ax},e^{bx},e^{cx})P^{-1}.$$

Veamos qué pasa si $k = 3$ (el caso $k = 2$ es parecido pero más fácil). Tenemos

$$A_3 = aI + N \quad y \quad (aI)N = N(aI).$$

Por la propiedad 2,

$$\exp(xA_3) = e^{ax}\exp(xN) = \begin{pmatrix} e^{ax} & xe^{ax} & \frac{x^2}{2}e^{ax} \\ 0 & e^{ax} & xe^{ax} \\ 0 & 0 & e^{ax} \end{pmatrix}$$

de donde se calcula $\exp(xA)$ por la propiedad 3.

Desafortunadamente, el cálculo de P y de A_k es largo y muy inestable. El menor error numérico puede alterar P y A_k radicalmente (por ejemplo, puede cambiar una matriz del tipo A_1 en una del tipo A_2).

Por ello es conveniente tratar de reducir la serie (1) a una suma finita. Lo que puede obtenerse gracias al Teorema de Cayley y Hamilton:

Si $\det(yI - A) = y^n + a_{n-1}y^{n-1} + \ldots + a_0$, entonces

$$A^n + a_{n-1}A^{n-1} + \ldots + a_0 I = 0.$$

En el Ejemplo 1, nos da:

$$A^2 + 18A + 17I = 0.$$

Por lo tanto, la n-ésima potencia de A es

$$A^n = -a_0 I - a_1 A - \ldots - a_{n-1}A^{n-1},$$

esto es, una combinación lineal de $I, A, \ldots, A^{n-1}$. Con ello podemos calcular A^{n+k}, $k \geq 0$ como una combinación de $I,\ldots,A^{n-1}$:

$$\begin{aligned} A^{n+1} &= -a_0A - a_1A^2 - \ldots - a_{n-1}A^n = \\ &= -a_0A - a_1A^2 - \ldots - a_{n-1}(-a_0I - a_1A - \ldots - a_{n-1}A^{n-1}) \\ &= a_0a_{n-1}I + (a_1a_{n-1} - a_0)A + \ldots + a_{n-1}^2A^{n-1} \end{aligned}$$

y en forma análoga A^{n+2}, A^{n+3}, etc.

Por lo tanto

$$\text{(3)} \qquad \exp A = \sum_{j=0}^{n-1} b_jA^j \qquad (A^o = I)$$

donde las b_j son series en los coeficientes $a_0,a_1,\ldots,a_{n-1}$ que a su vez dependen de las entradas de la matriz A.

El problema consiste en como calcular dichos coeficientes b_j. Para ello procederemos heurísticamente:

Primero generalicemos la expresión (3) considerando xA en lugar de A. En este caso en forma completamente análoga, obtenemos:

$$\text{(4)} \qquad \exp(xA) = \sum_{j=0}^{n-1} b_j(x)A^j$$

donde ahora las $b_j(x)$ son funciones de x.

Derivando ambos lados de la expresión (4) y sustituyendo (4), tenemos:

$$\sum_{j=0}^{n-1} b_j'(x)A^j = A\exp(xA) = A\sum_{j=0}^{n-1} b_j(x)A^j =$$

$$= \left(\sum_{j=0}^{n-2} b_j(x)A^{j+1}\right) + b_{n-1}(x)A^n =$$

$$= \left(\sum_{j=0}^{n-2} b_j(x)A^{j+1}\right) - b_{n-1}(x)\sum_{j=0}^{n-1} a_jA^j =$$

$$= -a_0b_{n-1}(x)I + \sum_{j=0}^{n-2} (b_j(x)-a_{j+1}b_{j-1}(x))A^{j+1}.$$

Igualando coeficientes, obtenemos el siguiente sistema lineal de ecuaciones diferenciales ordinarias:

$$b_0'(x) = -a_0b_{n-1}(x)$$

$$b_j'(x) = b_{j-1}(x) - a_jb_{n-1}(x) \qquad (j = 1,\dots,n-1)$$

que escrito en forma matricial es:

$$\begin{pmatrix} b_{n-1}(x) \\ b_{n-2}(x) \\ \cdot \\ \cdot \\ \cdot \\ b_0(x) \end{pmatrix}^1 = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 & -a_{n-1} \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \\ 0 & 1 & & & & \cdot \\ 1 & 0 & & & & \cdot \\ 0 & 0 & \dots & 0 & 0 & -a_0 \end{pmatrix} \begin{pmatrix} b_{n-1}(x) \\ \cdot \\ \cdot \\ \cdot \\ b_0(x) \end{pmatrix}$$

Puesto que $I = e^{oA} = \sum_{j=0}^{n-1} b_j(0)A^j$, además sabemos que $b_0(0) = 1$ y $b_1(0) = \ldots = b_{n-1}(0) = 0$.

Fácilmente se ve que si

$$C = \begin{pmatrix} a_1 & a_2 & a_{n-1} & 1 \\ a_2 & a_3 & 1 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ a_{n-1} & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

e $Y(x) = (y(x), y'(x), \ldots, y^{(n-1)}(x))^t$.

Entonces

$$CY = (b_0(x), \ldots, b_{n-1}(x))^t$$

donde $y(x) = b_{n-1}(x)$ satisface la ecuación

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1y' + c_0y = 0$$

con valores iniciales

$$y(0) = \ldots = y^{(n-2)}(0) = 0 , \quad y^{(n-1)}(0) = 1.$$

Reduciendo así nuestro problema al resolver una ecuación diferencial lineal con coeficientes constantes de orden n.

EJEMPLO 3. Si

$$\det(yI - A) = (y-a)^3$$

tenemos

$$a_1 = 3a^2, \quad a_2 = -3a \quad y \quad y(x) = (C_1 + C_2 + C_3x^2)e^{ax}.$$

Resolviendo las ecuaciones lineales obvias para los C, encontramos $C_1 = C_2 = 0$, $C_3 = \frac{1}{2}$ y $y = \frac{1}{2}x^2e^{ax}$

$$Y = \frac{1}{2}e^{ax}(x^2, ax^2 + 2x, a^2x^2 + 4ax + 2)^t$$

$$C = \begin{pmatrix} 3a^2 & -3a & 1 \\ -3a & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

y

$$\exp(xA) = \frac{1}{2}e^{ax}\{(a^2x^2 - 2ax + 2)I + (-2ax^2 + 2x)A + x^2A^2\}.$$

Un cálculo parecido, y algo más sencillo, es:

Sean $a_1, \ldots, a_n$ los valores propios de la matriz A y sean

$$P_0 = I, \qquad P_k = P_{k-1}(A - a_k I)$$

y $r_1,\ldots,r_n$ funciones que satisfacen

$$r_1' = a_1 r_1 \qquad\qquad r_1(0) = 1$$

$$r_{k+1}' = a_{k+1} + r_k \qquad r_{k+1}(0) = 0.$$

Entonces

$$\exp(xA) = \sum_{k=0}^{n-1} r_{k+1} P_k.$$

En el ejemplo anterior

$$r_1' = ar_1$$

$$r_2' = r_1 + ar_2$$

$$r_3' = r_2 + ar_3$$

que, con los valores iniciales dados, resulta

$$r_1 = e^{ax}, \quad r_2 = xe^{ax}, \quad r_3 = \frac{x^2}{2} e^{ax}$$

y

$$\exp(xA) = \frac{1}{2} e^{ax}\{2I + 2x(A - aI) + x^2(A - aI)^2\}.$$

Este último resultado es general: Si A es una matriz cuadrada $n\times n$ cuyos valores propios son todos a entonces

$$\exp(xA) = e^{ax} \sum_{k=1}^{n-1} \frac{x^k}{k!} (A - aI)^k.$$

En efecto axI y $x(A - aI)$ conmutan, luego, por la propiedad 2,

$$e^{xA} = e^{axI} e^{x(A-aI)} = e^{ax} \sum_{k=0}^{\infty} \frac{x^k}{k!} (A - aI)^k$$

pero $(A-aI)^{n+k} = 0$ para $k \geq 0$.

Estos resulrados se pueden encontrar en

1. T. Apostol, Calculus, Blaisdell, 1969.
2. E. Putzer, Avoiding the Jordan Canonical Form, American Math. Monthly, 73 (1966), pp. 2-7.

CURVATURA Y GRAVITACION

Gerardo F. Torres del Castillo

Dentro de la Mecánica Vectorial y en muchas otras áreas de la física y de la geometría, se usa extensamente el hecho de que en el plano y en el espacio euclideano se pueden trasladar libremente los vectores de un punto a otro. Es decir, que dado un vector en algún punto del plano o del espacio, éste siempre se puede trasladar a cualquier otro punto manteniendo su dirección y magnitud.

El concepto de trasladar paralelamente o rígidamente un vector, es en realidad muy importante ya que, por ejemplo, en la definición de la aceleración de una partícula se involucra la necesidad de comparar los vectores de velocidad de la partícula en diferentes puntos del espacio. Esto es, que es necesario trasladar los vectores de velocidad de la partícula en dos instantes diferentes a un orígen común para evaluar su diferencia. Como es bien sabido, la aceleración de una partícula es o no cero dependiendo de que la partícula se mueva libremente o que sufra la acción de algún agente externo.

Por otra parte, el espacio en que vivimos es tridimensional en el sentido de que cada punto del espacio puede ser etiquetado o caracterizado por tres números reales. Es por ello que, usualmente, nuestro espacio físico se identifica con $\mathbb{R}^3$ (el conjunto de triadas de números reales) y la estructura de éste se utiliza para modelar muchas teorías físicas. Así por ejemplo, en la mecánica clásica, en electromagnetismo clásico y las demás teorías basadas en éstas, se emplea comunmente el concepto de vector de posición y se efectúan operaciones entre estos vectores basándose en las operaciones correspondientes definidas en $\mathbb{R}^3$. Sin embargo, debe tenerse en mente que la correspondencia establecida entre los puntos sean equivalentes en todos los sentidos.

Tomemos por ejemplo la superficie de una esfera. Los puntos de esta superficie pueden etiquetarse por una pareja de números reales (v.gr. longitud y latitud). Es claro que, a pesar de esto, la superficie de una esfera es distinta de $\mathbb{R}^2$, aún cuando una región suficientemente pequeña de la superficie de una esfera luzca como una parte del plano euclideano.

De hecho si uno considera un vector tangente a la esfera, digamos en un punto A del ecuador, apuntando hacia el norte y este vector se traslada a lo largo del ecuador, conservándolo tangente a la esfera, a otro punto B sobre el ecuador, uno espera que el vector después de trasladado siga apuntando hacia el norte. Si en cambio, este vector primero es llevado al polo norte a lo largo de un meridiano y después se le traslada al punto B a lo largo del meridiano que pasa por B, es fácil imaginar que el vector resultante no apuntará hacia el norte. Así pués, mientras que en $\mathbb{R}^2$ el resultado de trasladar paralelamente un vector de un punto a otro no depende de la trayectoria que se siga; en el caso de la superficie de una esfera no ocurre esto. Existen además otras diferencias entre la superficie de una esfera y el plano euclideano, como que la suma de los ángulos interiores de un triángulo sobre la superficie de una esfera es mayor que 180°.

Cuando el resultado de trasladar paralelamente un vector de un punto a otro depende de la trayectoria que se sigue, se dice que existe curvatura. Esta definición de curvatura de un espacio es "intrínseca" en el sentido de que no se hace referencia a cómo se ve este espacio desde "afuera".

Veamos ahora cual es la consecuencia de suponer que nuestro espacio físico sea curvo. El asumir que existe curvatura significa que, al modelar el espacio físico por $\mathbb{R}^3$, vamos a reemplazar la noción de transporte paralelo existente en $\mathbb{R}^3$ (la cual no posee curvatura) por alguna otra. Una partícula que, de acuerdo a la noción original de trasporte paralelo, se moviera libremente aparecerá, en general, acelerada si se usa otra noción de transporte paralelo. Recíprocamente, una partícula que parece moverse libremente de acuerdo a la nueva noción de transporte paralelo que introduzcamos; cuando se le considera usando el transporte de vectores usual de $\mathbb{R}^3$ se comportará como si sobre ella actuase una fuerza.

Así pués, el cambiar una definición de transporte paralelo por otra, equivale a introducir una fuerza. Ahora bien, ¿cuál es la naturaleza de esta fuerza?. Debido a que esta fuerza tiene por orígen la forma en que uno compara los vectores de velocidad de una partícula en diferentes instantes, es claro que la aceleración producida por esta fuerza no depende en absoluto de las características de la partícula en cuestión (masa, tamaño, carga eléctrica, etc.). ¿Existe en la naturaleza una fuerza con tal propiedad?. Sí, tal como fue observado por Galileo, la fuerza gravitacional imprime a todos los objetos la misma aceleración. Por lo tanto se puede adoptar una definición de transporte paralelo de vectores con la cual una partícula tiene aceleración cero si y sólo si cae libremente.

Con este punto de vista la gravitación adquiere un carácter geométrico. En la teoría general de la relatividad se interpreta de esta manera la gravitación, no tratando al espacio como un ente absoluto sino reconociendo que el espacio y el tiempo forman una sola entidad.

En una forma semejante, dentro de la Mecánica Cuántica, se puede interpretar la fuerza electromagnética. Acualtmente se trata de interpretar mediante teorías similares, llamadas genéricamente teorías de norma, las otras dos fuerzas fundamentales de la naturaleza, las cuales sólo actúan a nivel subatómico, llamadas fuerzas débiles y fuertes.

BIOLOGIA, FISICA Y LA TEORIA CUALITATIVA DE LAS ECUACIONES DIFERENCIALES

Rafael Pérez Pascual*

Si revisamos la historia de las matemáticas, sobre todo del renacimiento a nuestros días, encontramos la existencia de un fuerte vínculo entre el desarrollo de ideas y teorías matemáticas y el de ideas y teorías físicas. Así el cálculo diferencial nace estrechamente unido al nacimiento de la mecánica clásica, el estudio de las ecuaciones diferenciales parciales con la mecánica de los medios continuos y la teoría de distribuciones con la formalización de la mecánica cuántica.

En ocasiones las teorías matemáticas se desarrollan sin nexo aparente con las aplicaciones físicas, para más tarde resultar indispensables en el desarrollo de teorías físicas, como son los casos de las geometrías no euclideanas o la teoría de grupos. En otras ocasiones encontramos lo inverso, ideas que surgen del contexto de la física y que al ser tomadas por las matemáticas forman la base de nuevos desarrollos matemáticos, como es el caso de la teoría ergódica.

Ahora bien, si hacemos la misma abreviación para el caso de otras ciencias como la biología o la sociología notamos de inmediato que el caso es distinto.

La biología, por ejemplo, se ha desarrollado por caminos alejados de las matemáticas; son tan pocos los ejemplos que podríamos citar sobre vínculos de la biología y las matemáticas, como son numerosos en el caso de la física. La discusión en torno a las causas de esta situación es ya una discusión clásica y mucho se ha dicho, desde los argumentos que señalan a la biología como una ciencia poco desarrollada, hasta aseveraciones que intentan fundamentar posiciones vitalistas argumentando la imposibilidad de matematizar la biología.

En efecto, la biología es un ejemplo típico de un campo científico que se ha mostrado reacio a la matematización. Creemos que esto se debe, básicamente, a dos

* Instituto de Física, Universidad Nacional Autónoma de México.

cuestiones inherentes a los fenómenos biológicos.

En primer lugar los fenómenos biológicos no admiten una cuantificación tan precisa como los fenómenos físicos: la velocidad de caída de una piedra o la masa del electrón se pueden medir con gran precisión y consistencia, pero las variables biológicas, como la concentración de alguna substancia en la célula o el peso de los individuos de una especie, presentan una gran dispersión haciendo muy difícil el fijar características cuantitativas.

En segundo lugar los fenómenos biológicos son más complejos. Pero no podemos limitarnos a señalar que lo son, debemos también decir qué entendemos por "más complejo". En nuestra opinión, lo que distingue la "complejidad" del fenómeno biológico de la "simplicidad" del fenómeno físico es la posibilidad de construir teorías sobre dichos fenómenos basadas en ecuaciones lineales o la imposibilidad de hacer esto. En el caso de la física las teorías lineales son sumamente útiles y muchas veces fundamentales, tal es el caso de las ecuaciones de Maxwell o de la mecánica cuántica. En el caso de la biología siempre es necesario recurrir a teorías o descripciones no lineales.

La diferencia parece pequeña pero no lo es, un ejemplo lo muestra claramente: La ecuación de Schöedinger para el átomo de hidrógeno es

$$ih\frac{\partial\psi}{\partial t} = -\frac{h^2}{2m}\ \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{c}{v}\psi$$

Esta ecuación es lineal, esto es si ψ_1 y ψ_2 son soluciones también lo serán $\psi = a\psi_1 + b\psi_2$; esto conduce al hecho de que, a pesar del aspecto complicado de la ecuación y de que sus soluciones nos dan con grán detalle y aproximación el comportameinto del átomo de hidrógeno, sus soluciones son relativamente fáciles de encontrar y se expresan en términos de funciones clásicas ampliamente estudiadas desde el siglo pasado. Por otra parte, la ecuación de Van der Pol

$$\frac{d^2\theta}{dt} - \epsilon(1-\theta^2)\frac{d\theta}{dt} + \theta = 0$$

que aparece en el análisis de circuitos eléctricos con tríodos, es aparentemente más sencilla y es ordinaria sin embargo es mucho más "compleja", al grado de que a pesar de ser una ecuación sumamente estudiada y de que conocemos mucho sobre sus soluciones, no somos capaces de encontrarlas. La gran diferencia consiste en la no linealidad de la ecuación de Van der Pol.

Este ejemplo es típico, en el sentido de que las ecuaciones lineales pueden, en general, ser resueltas o en el peor de los casos se cuenta con poderosos métodos aproximados, en el caso no lineal prácticamente nunca se pueden encontrar soluciones y los métodos aproximados dejan mucho que desear, pudiendo incluso conducir a conclusiones erróneas.

En esta perspectiva parece ser una desgracia encontrarse con un problema en física, biología, ingeniería u otra ciencia o tecnología donde tengamos que planteamos formalismos matemáticos no lineales. En realidad lo es en cierto sentido pero no en otro. Es una desgracia en el sentido de que al toparnos con una ecuación no lineal, lo más probable es que no encontremos soluciones y que al recurrir a métodos aproximados o numéricos quede en duda su aplicabilidad al caso, además de que con estos métodos generalmente se pierden soluciones importantes y gran parte de la comprensión del fenómeno en cuestión.

No es una desgracia en el sentido de que la fenomenología que presentan las cuestiones no lineales es mucho más rica que en el caso lineal, abriendo, entonces, la puerta al estudio de situaciones complejas y a múltiples y variadas aplicaciones y controles de un mismo sistema.

Ahora bien, la aparición y el estudio de fenómenos no lineales no es nuevo, en realidad aparecen simultáneamente con los lineales, por ejemplo toda la mecánica de medios continuos es no lineal y podemos ver que presenta fenómenos complejos como la turbulencia. Otro ejemplo clásico tampoco resuelto plenamente, a pesar de que en los últimos dos siglos todo físico o matemático ha tratado de resolverlo, es el problema de tres cuerpos en la mecánica clásica.

La inmensa mayoría de las formulaciones matemáticas en las ciencias y la técnica tienen en el fondo una ecuación diferencial, por otra parte, la teoría de ecuaciones diferenciales es tan rica, que se puede decir que casi toda teoría matemática tiene relación, utiliza o puede ser aplicada a las ecuaciones diferenciales. Veamos pues algunas cuestiones sobre fenómenos no lineales, comenzando con un pequeño resumen de la teoría de ecuaciones diferenciales ordinarias.

Una ecuación diferencial ordinaria es una ecuación en la que aparecen derivadas, esto es, una ecuación del tipo

$$F(x, \frac{dx}{dt}, \frac{d^2x}{dt^2}, \ldots, \frac{d^nx}{dt^n}, t) = 0. \tag{1}$$

Una solución de esta ecuación es una función del tiempo x(t) tal que al ser substituída en (1) obtengamos una identidad.

Ejemplo: Ecuación diferencial

$$\frac{d^2x}{dt^2} + x = 0$$

Soluciones $x = 0,\ x = \cos t,\ x = \operatorname{sen} t,\ x = 5\ell^{it}.$

Como vemos, una ecuación diferencial puede tener muchas soluciones, pero si especificamos los valores iniciales, esto es el valor de x y sus derivadas hasta orden n−1 en t = 0 u otro valor y la función F es continua con derivada continua, entonces sólo hay una solución, o mejor dicho, al escoger una de entre las infinitas condiciones iniciales escogemos una y sólo una de entre las infinitas soluciones.

Ejemplo: Si en el ejemplo anterior especificamos

$$x(0) = 1 \quad y \quad \left.\frac{dx}{dt}\right|_{t=0} = 2$$

la única solución es

$$x = 2 \operatorname{sen} t + \cos t.$$

En términos generales dada una ecuación diferencial tal que la función que la define sea continua y tenga propiedades de diferenciación, dadas unas condiciones iniciales, existirá un rango de tiempos en donde hay una y sólo una solución. A este resultado se le conoce con el nombre de teorema de existencia y unicidad y es de gran importancia en la teoría de ecuaciones diferenciales.

La notación que usamos en la ecuación (1) aunque es utilizada con gran frecuencia no siempre es la más conveniente. Si definimos nuevas variables

$$x_1 = x,\ x_2 = \frac{dx}{dt}, \ldots, x_n = \frac{d^{n-1}x}{dt^{n-1}}$$

es fácil ver que la ecuación (1) se puede expresar como un sistema de n ecuaciones de primer orden acopladas, esto es

$$\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, x_2, \ldots, x_n, t) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2, \ldots, x_n, t) \\ &\ \ \vdots \\ \frac{dx_n}{dt} &= f_n(x_1, x_2, \ldots, x_n, t) \end{aligned} \tag{2}$$

Como decíamos esta forma es más general y en muchos sentidos más conveniente. Una de las ventajas consiste en que nos permite definir un espacio, llamado espacio fase, como el espacio de n dimensiones generado por las eneadas $(x_1, \ldots, x_n)$.

Cuando las funciones $f_1, f_2, \ldots, f_n$ no dependen explícitamente de t se simplifica la teoría, por el momento nos restringiremos a este caso, que se conoce con el nombre de autónomo.

En esta notación, cuando la ecuación es lineal y autónoma queda expresada en general por

$$\begin{aligned} \frac{dx_1}{dt} &= a_{11}x_1 + a_{21}x_2 + \ldots + a_{n1}x_n \\ \frac{dx_2}{dt} &= a_{12}x_1 + a_{22}x_2 + \ldots + a_{n2}x_n \\ &\ \ \vdots \\ \frac{dx_n}{dt} &= a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{nn}x_n \end{aligned} \tag{3}$$

donde las a_{ij} son constantes. Nótese que si expresamos la eneada $(x_1, x_2, \ldots, x_n)$ como un vector X en el espacio fase la ecuación (3) se puede escribir como

$$\frac{dX}{dt} = Ax \tag{4}$$

donde A es la matriz n x n de coeficientes a_{ij}.

Ahora si hacemos una transformación lineal de coordenadas en el espacio fase por medio de una matriz B

$$X' = BX \qquad X = B^{-1}X'$$

en las nuevas coordenadas la ecuación (4) quedará como

$$\frac{dX'}{dt} = BAB^{-1}X'$$

Un conocido teorema del álgebra lineal nos dice que si la matriz A tiene valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ distintos y reales, siempre existe una matriz real no singular B tal que BAB^{-1} es diagonal. Esto implica que, en ese caso, existen coordenadas tales que la ecuación (4) se reduce a

$$\frac{dx'_1}{dt} = \lambda_1 x'_1$$

$$\frac{dx'_2}{dt} = \lambda_2 x'_2$$

$$\frac{dx'_n}{dt} = \lambda_n x'_n$$

que es un sistema desacoplado cuya solución general es simplemente

$$x'_1 = c_1 \ell^{\lambda_1 t},\ x'_2 = c_2 \ell^{\lambda_2 t}, \ldots, x'_n = c_n \ell^{\lambda_n t}$$

en el caso en que se tengan valores propios complejos se puede separar el sistema (4) en un sistema de ecuaciones desacopladas por parejas y la solución general se obtiene por el mismo procedimiento.

Vemos así que el problema de las ecuaciones autónomas lineales se puede resolver en forma general en unas cuantas líneas haciendo uso de resultados muy conocidos del álgebra lineal. Este caso también nos ilustra la claridad que la notación (2) nos da sobre las ecuaciones diferenciales y sus soluciones.

En el caso no lineal la cuestión es totalmente distinta, pues la introducción de un término no lineal inmediatamente transforma un problema en el que encontramos la solución fácilmente en un problema para el cual es prácticamente imposible encontrar solución.

Esta situación se ha atacado por varios caminos y aquí deseamos hacer énfasis en uno de ellos, la teoría cualitativa de las ecuaciones diferenciales, fundada por Liapunov y Poincare a principios de este siglo y finales del pasado.

La idea básica consiste en olvidar la búsqueda de soluciones cuantitativas y procuramos métodos que nos den resultados de tipo cualitativo. Pensemos, por ejemplo, en un ingeniero que desea estudiar la estabilidad de un puente, puede en principio analizar con gran precisión el comportamiento estático y dinámico de cada parte del puente y obtener una gran cantidad de información cuantitativa sobre el puente para después sintatizar a partir de ella una respuesta a una pregunta cualitativa, ¿el puente es estable o no? Este es el camino generalmente usado, esto es, se hacen estudios cuantitativos exhaustivos en muchos casos para buscar respuestas cualitativas; pero podríamos imaginar un camino distinto, un método tal que, sin perder el rigor matemático, nos de la solución a nuestras preguntas cualitativas sin necesidad de pasar por la solución cuantitativa.

Veamos un ejemplo:

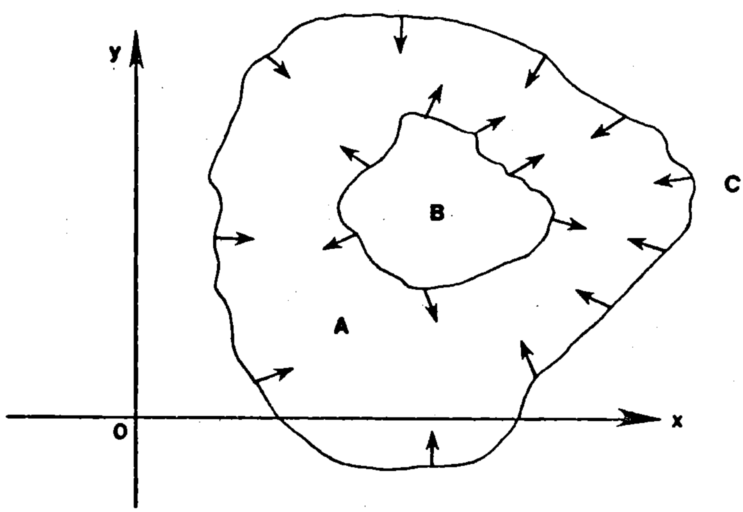
Tomemos una ecuación diferencial autónoma de orden dos o de dos dimensiones

$$\begin{aligned} \frac{dx}{dt} &= f(x,y) \\ \frac{dy}{dt} &= g(x,y) \end{aligned} \tag{5}$$

una solución son dos funciones $x(t)$ y $y(t)$ tales que satisfacen la ecuación, pero estas funciones se pueden interpretar como la representación paramétrica de una curva en el espacio fase (x,y), por tanto podemos ver las soluciones de (5) como una familia de curvas en el espacio (x,y) que no se intersectan y que llenan todo el espacio, como es fácil demostrar a partir del teorema de existencia y unicidad.

Llamando punto crítico o de equilibrio a aquel en que tanto f(x) como g(x) se anulan (nótese que un punto crítico es solución de la ecuación). Se ha demostrado que en el caso de dos dimensiones las soluciones tienen necesariamente que ser asintóticas a un punto crítico, a una curva cerrada o a infinito. El caso de las curvas cerradas es interesante pues corresponde a soluciones periódicas y en el caso de ser aisladas se conocen como ciclos límite.

Veamos ahora con este ejemplo el tipo de cuestiones que se plantea en la teoría cualitativa. Supongamos que hemos encontrado en el espacio fase una región del tipo A que se ilustra en la siguiente figura.



Esto es, una región que no contiene ningún punto crítico y tal que todas las curvas solución en su frontera apuntan hacia adentro, entonces es fácil demostrar que existe al menos una curva solución cerrada dentro de la región y por tanto una solución periódica, esto implica también que en la región B hay al menos un punto crítico. Nótese que este resultado no sería válido en tres o más dimensiones.

Este es el tipo de cuestiones que la teoría cualitativa ataca. En general son problemas sobre la existencia de soluciones periódicas, de superficies o variedades tales que las soluciones tiendan a ellas, del tipo de puntos críticos que tenemos y de la forma en que se conectan entre sí, de las fronteras entre las regiones de atracción de estos elementos críticos, de la estabilidad de estas propiedades ante el cambio de parámetros y en general de delucidar las llamadas propiedades topológicas de las curvas solución en el espacio fase.

Todas estas cuestiones están relacionadas con el desarrollo de ramas matemáticas como la topología, el análisis en variedades diferenciales y otras que en conjunto constituyen una de las ramas más dinámicas de la investigación matemática moderna.

Ahora bien, deciamos al principio que en nuestra opinión, disciplinas científicas como la biología no se han matematizado debido fundamentalmente a dos problemas, la dificultad o imposibilidad de establecer resultados experimentales cuantitativos y la fenomenología no lineal. Por otro lado, matemáticamente ha resultado sumamente poderosa la conjunción de las ecuaciones diferenciales no lineales con teorías de caracter cualitativo. De esta forma podemos ver que hay una semejanza o complementaridad entre las necesidades y determinantes biológicas con estas ideas y teorías matemáticas.

El resultado ha sido el surgimiento en los últimos años de la biología matemática. Permítasenos, aunque sea en forma somera, mencionar algunas de las ramas más dinámicas de esta nueva disciplina.

ECOLOGIA MATEMATICA

Es éste uno de los campos más conocidos desde que en los años veintes se inició su estudio con las famosas ecuaciones de Volterra.

Un sistema ecológico se representa por una serie de funciones del tiempo que denotan la densidad de individuos de cada especie de las que componen el sistema. El problema consiste en encontrar dichas funciones a partir del conocimiento de las relaciones entre las distintas especies. En realidad estas relaciones conducen al establecimiento de ecuaciones diferenciales, en general no lineales, cuyas soluciones son esas funciones del tiempo.

En este caso es claro que lo que nos interesa es el comportamiento general del sistema más que el número exacto de individuos de cada especie, interesan cuestiones como la existencia de oscilaciones, la estabilidad del sistema ante perturbaciones, si hay especies que se extingen al introducir otra especie, etc., etc. Este es precisamente el tipo de cuestiones que ataca la teoría cualitativa de ecuaciones diferenciales.

En el caso de la ecología matemática se ha podido demostrar, por ejemplo, la existencia de sistemas ecológicos muy generales en los que cambios pequeños en el número de individuos de una especie puede producir cambios cualitativos, como puede ser la extinción de una o varias especies.

BIOQUIMICA

Desde hace mucho tiempo se sabe que el comportamiento temporal de las concentraciones de las diversas substancias que intervienen en una reacción química está gobernado por una ecuación diferencial en general no lineal.

Por otro lado ha sido costumbre de los bioquímicos olvidar o ignorar este hecho. En los últimos años se ha dado un gran impulso al estudio matemático de las ecuaciones diferenciales que describen la cinética de procesos bioquímicos; entre los resultados más importantes citaremos dos:

En base a estudios matemáticos se predijo por Lotka hace ya casi seis décadas, la existencia de reacciones químicas oscilatorias. En los últimos años, no sólo se ha confirmado experimentalmente la existencia de dichas reacciones, sino que se ha encontrado que un gran número de importantes reacciones bioquímicas presentan este fenómeno; en la actualidad el estudio de las reacciones oscilatorias y sus interacciones es un campo de gran importancia en el que el análisis matemático es indispensable.

Se ha encontrado que muchas reacciones bioquímicas pueden presentar comportamientos radicalmente distintos dependiendo de la homogeneidad o heterogeneidad del sistema e incluso, en el caso heterogeneo, de su topología. Este hecho es de suma importancia para el futuro de la bioquímica ya que, en contradicción con la costumbre de la bioquímica de laboratorio, las reacciones bioquímicas en las células reales ocurren en sistemas heterogeneos y con topologías definidas, por lo que su riqueza es muy superior al de las mismas reacciones cuando se estudian en sistemas homogeneos. Por otro lado estos estudios requieren de avanzadas técnicas y teorías matemáticas y es uno de los casos en que el estudio de fenómenos biológicos ha inspirado resultados de caracter matemático.

FENOMENOS BIOLOGICOS PERIODICOS

Una de las características de los seres vivos es el gran número de periodicidades que presentan y su acoplamiento. Desde los cambios periódicos de 24 horas hasta el latir del corazón o las concentraciones hormonales, los seres vivos presentan un complejísimo sistema de periodicidades. De entre toda la variedad de estudios matemáticos para abordar esta problemática destaca el estudio de la estimulación periódica de un sistema de por si periódico. Esto nos conduce a lo que matemáticamente conocemos como un oscilador no lineal forzado que, a diferencia del caso lineal, presenta un comportamiento sumamente complejo, así como dificultades matemáticas muy serias.

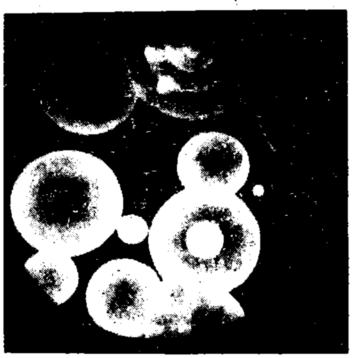
Uno de los fenómenos que se producen y que se ha encontrado en la última década, es la existencia de los llamados atractores extraños que son soluciones de ecuaciones diferenciales tan complejas que aparentemente reproducen el resultado de un fenómeno totalmente azaroso e irreproductible, siendo sin embargo, el resultado de un proceso totalmente determinista. En los últimos años este tipo de estudios parecen incluso apuntar hacia la solución de viejos problemas de la física como es el caso de la turbulencia.

MORFOGENESIS

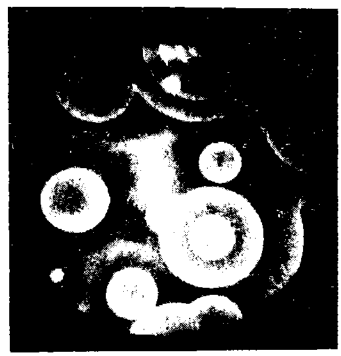
Finalmente mencionamos el caso de morfogénesis que es el estudio de los procesos naturales mediante los cuales a partir de un medio en el que no existia se genera una forma específica. Desde luego esto abarca una multitud de fenómenos pero sin duda dentro de esta multitud el caso de la generación de formas biológicas, esto es, del desarrollo embrionario que es de los más sorprendentes. ¿Cómo, a partir de una célula, se genera un organismo con formas y topologías tan complejas como lo es el animal adulto? ¿Hay leyes generales de los procesos morfogenéticos?. Estas y muchas otras preguntas son típicas en teoría de estos procesos.

Desde luego prácticamente no conocemos nada sobre este apasionante tema. Simplemente los intentos matemáticos más sencillos conducen de inmediato a la formulación de ecuaciones diferenciales parciales no lineales, y si en el caso de las ordinarias autónomas se han sentado bases para desarrollar una teoría, si en el caso de las ordinarias no autónomas en el inicio de su estudio nos encontramos con soluciones tan complejas como los atractores extraños, en el caso de las parciales no lineales, podemos decir que no existe una idea clara que permite iniciar un estudio sistemático.

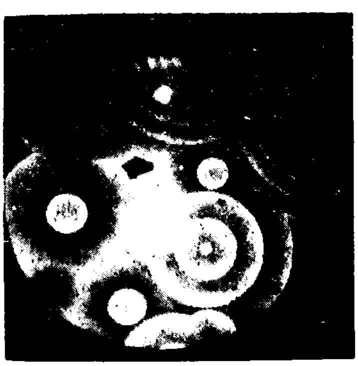
Así pues vemos como las ciencias naturales y las matemáticas, en esta ocasión con la participación de la biología, han recuperado ese estrecho nexo que durante varias décadas habían parecido perder. No cabe duda que este hecho es de gran importancia y repercutirá extensamente en el futuro de las ciencias, las matemáticas y la tecnología.



Time t_0



Time t_0 + 1 minute



Time t_0 + 2 minutes

Time t_0 + 3 minutes

BIBLIOGRAFIA

Hirsch, M.; Smale, S.: Differential equations, dynamical systems and linear algebra. Academic Press, N.Y. 1974.

May, R.M.: Model ecosystems, stability and complexity. Princeton U.P., 1974.

Maynard Smith, J.: Mathematical ideas in Biology. Cambridge U.P., 1974.

R. Rosen ED.: Fundations of mathematical Biology. Cuatro volúmenes, Academic Press.

Rosen, R.: Dynamical systems theory in Biology. Wiley 1970.

Spivak, M.: Calculus on manifolds. Benjamin 1965.

J. Moser ED.: Dynamical "Systems: theory and applications". Springer-Verlag 1975.

Y. Choquet-Brunat et all: Analysis, manifolds and physics. North Holland 1977.

SOBRE LOS PROBLEMAS DE LA CONFRONTACION TEORIA-EXPERIMENTO

Rubén G. Barrera*
Aurora Gallardo**

RESUMEN

En el presente trabajo analizamos, cómo los problemas estrictamente metodológicos del método experimental están contenidos en la problemática de la metodología de las ciencias sociales. Esto se realiza mediante un examen cuidadoso del método científico que muestra la necesidad de introducir el estudio de la historia social de la ciencia como herramienta de análisis. Los argumentos se ilustran con ejemplos tomados de la historia de la física.

*Instituto de Física, Universidad Nacional Autónoma de México.

**Sección Matemática Educativa, Centro de Investigación y de Estudios Avanzados del IPN.

1. Introducción

El poder de predicción que actualmente han logrado las ciencias experimentales, y en especial la física, las ha dotado de un prestigio metodológico tal, que han llegado a ser consideradas uno de los prototipos de cientificidad de nuestra época.

No es raro descubrir que un estudiante de la carrera de física, se encuentre firmemente convencido que los conocimientos que sus maestros le están transmitiendo, representan la forma de conocimiento más objetiva y racional lograda hasta ahora por la humanidad. Que a diferencia de las ciencias sociales y de las ciencias del hombre, la física posee un método seguro y firme para poder distinguir, por medio del experimento, entre las teorías falsas y las teorías verdaderas. Que el criterio de verdad se encuentra sustentado en la factibilidad de realizar y reproducir experimentos bajo condiciones controladas y en la posibilidad de exigir una coherencia entre las teorías propuestas y los hechos experimentales. Que es consciente de que los creadores de las teorías físicas, pueden estar fuertemente influenciados por diversas corrientes filosóficas o ideologías políticas, pero que sin embargo sus teorías van a confrontarse, a fin de cuentas, con el hecho experimental el cual si carece de ideología y de posición filosófica. Y que es precisamente dicha confrontación la que ha hecho posible el firme establecimiento de las llamadas "leyes físicas", las cuáles, especificada su área de aplicabilidad, representan un conocimiento de validez universal. Que bajo las mismas condiciones las leyes de Newton, por ejemplo, serán siempre válidas "aquí y en China".

Muchos están convencidos también, de que el conjunto de conocimientos en física, a diferencia del arte, por ejemplo, posee un carácter acumulativo, ya que las nuevas teorías deben incluir necesariamente, como caso particular, a las teorías anteriores de validez más restringida, siendo ésto último lo que define su mayor generalidad y un cierto sentido de progreso.

Es esta concepción tan deformada y tan extendida que se tiene de la física y de su metodología lo que nos ha llevado a escribir el presente artículo.

En él, nos proponemos realizar una revisión crítica de los problemas metodológicos a los que se enfrentan las ciencias experimentales, y en particular la física, con el fin de mostrar las graves dificultades que se presentan cuando se trata de concebir su proceso de adquisición de conocimientos, como un proceso objetivo con una dinámica condicionada esencialmente por contradicciones internas. A nuestra manera de ver, el problema reside en encontrar una explicación causal de la génesis y evolución de las teorías físicas; es decir, en llegar a comprender por que ciertas teorías subsisten a sus competidoras mientras otras son abandonadas.

Con el propósito de plantear el problema en toda su complejidad hemos estructurado el artículo en la siguiente forma:

En la segunda sección, hacemos una presentación esquemática de la manera "tradicional" de formular el método experimental. Esta consiste en la aplicación sistemática de cuatro pasos metodológicos: la observación o selección del problema físico, la postulación de una teoría o modelo, la experimentación y la verificación. El criterio de verificación reside en la concordancia entre teoría y experimento. En las siguientes dos secciones analizamos los problemas a que nos conduce la aplicación del método así formulado ilustrando, con diversos ejemplos, sus contradicciones internas y su incapacidad para explicar la evolución y el éxito de diversas teorías físicas. Exploramos también las distintas alternativas que han surgido para tratar de superar dichas dificultades pero que sin embargo respetan la rígida estructura del método en cuanto a su mecanismo metodológico interno. En la tercera sección planteamos las dificultades del criterio de verificación y analizamos la alternativa de sustituirlo por el criterio contrario: la refutación. La cuarta sección está dedicada a reconsiderar el criterio de refutación, pero basados ahora en un análisis cuidadoso del llamado "hecho experimental". Concluimos esta sección con la idea de que un método que consista en la aplicación estricta de un conjunto definido de reglas y de rígidos criterios de verdad para explicar el desarrollo de la física, se verá en serias dificultades al ser confrontado con el análisis histórico.

En la quinta sección nos servimos del concepto de actividad científica, como hilo conductor para conectar el marco conceptual de una teoría física con el entorno social donde ésta germina. La idea es mostrar que para poder conformar una explicación racional de la evolución conceptual de la física es necesario con siderarla no como una actividad intelectual aislada sino como parte de un proceso social. Es en este sentido que creemos necesario incorporar el análisis socio-histórico al marco metodológico de las ciencias experimentales. En la sexta y última sección presentamos nuestras conclusiones finales.

2. El Método Experimental

La forma "tradicional" de presentar el método experimental como un conjunto definido de reglas, lo hace aparecer como un método capaz de poder distinguir entre las teorías falsas y las teorías verdaderas.

Dichas reglas las podemos resumir en cuatro, a saber:

i) Observación o selección del problema

ii) Elaboración de una teoría o modelo

iii) Experimentación

iv) Verificación

La aplicación del método suele presentarse a lo largo de la siguiente línea de argumentación:

Es la curiosidad innata del hombre, la que lo conduce a observar el mundo físico que lo rodea y es su capacidad intelectual, la que lo lleva a buscar una explicación causal de dicho mundo.

Los diversos tipos de explicaciones que el hombre ha encontrado a lo largo de su historia, representan la aplicación de distintos métodos para la obtención del conocimiento. Dichos métodos definen distintas etapas en la evolución del pensamiento humano y así tenemos el pensamiento mágico, el religioso y el científico. Es éste último, el único que realmente puede ser considerado objetivo, sobre todo en el área de las ciencias experimentales, ya que en este caso el método está diseñado de tal forma que se deja a la misma naturaleza (a través del experimento) la toma de decisión sobre la validez o falsedad del conocimiento.

La aplicación de dicho método parte de <u>la observación</u> de la naturaleza. Es precisamente el hecho de observar, aunado a la curiosidad del hombre, lo que transporta a su conciencia la existencia de los fenómenos naturales que requieren de una explicación. Dicha explicación se elabora utilizando un conjunto de conceptos (marco teórico) los cuáles se articulan en base estricta a las reglas aceptadas de la lógica.

En el caso de un problema físico, la formulación de un modelo o teoría se basa primeramente, en la habilidad de captar los aspectos más significativos del fenómeno, para poder luego expresarlos en términos de conceptos físicos y finalmente articularlos de tal manera que sea posible lograr una explicación lógica y causal del fenómeno en consideración. Se busca, en general, que las propiedades de los conceptos físicos queden enmarcadas dentro de algún tipo de estructura matemática, con el fin de asegurar así que la coherencia lógica en la manipulación de los conceptos quede respaldada por la validez de las teorías matemáticas utilizadas.

A esta concatenación matemática de conceptos físicos, construída con la finalidad de explicar un cierto tipo de fenómenos es a lo que denominamos modelo matemático.

Se les llama modelos debido a que dichas construcciones nunca pueden ser exhaustivas, ya que de toda la riqueza y complejidad del fenómeno natural, se están seleccionando sólo un número reducido de conceptos mediante los cuáles se quiere entender sólo lo <u>esencial</u>, en el comportamiento de la naturaleza bajo circunstancias bien determinadas.

Y es precisamente la habilidad para extraer lo que de esencial tiene el fenómeno, así como la selección y estructuración más adecuada de conceptos, lo que le da a la física parte de su carácter creativo, no exento, por tanto, de un cierto placer estético.

Sin embargo, el uso de un formalismo matemático junto con una selección apropiada de conceptos asegura solamente la coherencia interna del modelo, mas no su validez como una interpretación correcta de la realidad. Para determinar esto último es necesario acudir al experimento.

Pero la posibilidad de corroborar la validez de un modelo sólo es posible si el modelo posee un poder predictivo. Es decir, si es capaz de predecir el comportamiento del sistema físico bajo circunstancias distintas a las del fenómeno que motivó su formulación, con la condición adicional de que sea posible, al menos en principio, reproducir dichas circunstancias en el laboratorio. Sólo de esta manera será posible poner el modelo a prueba, debido a que el método experimental basa su criterio de validez en la concordancia existente entre las predicciones del modelo y el experimento. Un modelo sin poder predictivo carece, por lo tanto, de valor científico. Obviamente las predicciones del modelo tampoco deben entrar en contradicción con hechos experimentales conocidos. Además el diseño de un experimento que ponga a prueba las partes más sensibles de uno o varios modelos, requiere también del ingenio y la capacidad creadora del experimentador.

Ahora bien, en el caso de que las predicciones del modelo y el experimento concuerden, se dice entonces que el modelo ha sido verificado por el experimento. En caso contrario será necesario entonces revisar el modelo, complicarlo, enriquecerlo o, de plano, sustituirlo por otro hasta lograr que sus predicciones estén en concordancia con los hechos experimentales. De esta manera se deja al experimento, que no es otra cosa que el comportamiento mismo de la naturaleza bajo condiciones controladas, el papel de árbitro en cuanto a la validez de un cierto modelo. Y es precisamente el hecho de que sea la naturaleza misma la que decida sobre la interpretación que el hombre hace de ella, lo que diferencía a las ciencias experimentales de otras ramas del conocimiento científico. En el caso de las ciencias experimentales la naturaleza juega un doble papel, como objeto de estudio y como árbitro en el criterio de verificación. Esto es lo que garantiza la objetividad del conocimiento adquirido ya que el comportamiento de la naturaleza es ajeno a las diversas motivaciones, prejuicios o emociones del investigador. Por otra parte, el hecho de que los nuevos modelos que se construyan no deban entrar en contradicción con los resultados experimentales conocidos garantiza, también, un claro sentido de progreso.

Dicho progreso está basado, por tanto, en la acumulación contínua de conocimientos, producto de una estrecha vinculación entre teoría y experimento. Así, el diseño de nuevos experimentos para probar uno o varios modelos y la construcción de nuevos modelos para interpretar resultados experimentales no explicados hasta entonces, constituyen los factores dinámicos primordiales en el avance de las ciencias experimentales.

Esta es, a grandes rasgos, la presentación que suele hacerse del método experimental y a la que hemos denominado "tradicional". Esta será, también, la presentación que tomaremos como base para realizar, en las secciones subsecuentes, un análisis crítico, tanto de las distintas fases del método así presentado, como de su dinámica real dentro de un contexto socio-histórico.

3. Verificación versus Refutación

En esta sección analizamos el criterio de validez del método experimental, basado en la verificación de los modelos teóricos por medio del experimento y expuesto en la sección anterior bajo lo que hemos denominado la presentación "tradicional".

Iniciamos la presente discusión, preguntándonos si en el caso de que exista concordancia entre teoría y experimento realmente es posible decir que el modelo ha sido verificado.

¿No es acaso posible que en otro arreglo experimental las predicciones del modelo pudieran entrar en conflicto con los resultados experimentales así obtenidos?. Y aunque hubiere un gran número de predicciones ya verificadas por el experimento, ¿sería acaso posible asegurar que alguna de las predicciones aún no probadas jamás entraría en desacuerdo con el experimento?.

Creemos que es fácil convencerse, que para poder verificar experimentalmente todas las predicciones posibles de modelos suficientemente ambiciosos, sería necesario realizar un número tan grande de experimentos que llegaríamos forzosamente a concluir que es materialmente imposible verificar dichos modelos, en el sentido estricto de la palabra. La única afirmación concebible sería el poder asegurar que un cierto modelo no ha sido refutado.

Por lo tanto, si en vez de tratar de verificar un modelo lo que tratamos es de refutarlo, vemos que entonces las dificultades mencionadas anteriormente desaparecen, ya que para refutar un modelo es necesario un sólo experimento. Es decir, que la no concordancia entre las predicciones del modelo y un solo hecho experimental es suficiente para refutar el modelo teórico y obligarnos a revisarlo críticamente, ya sea para modificarlo o para sustituirlo por otro modelo distinto.

En este sentido, el verdadero propósito de los experimentos no debe ser el de verificar los modelos sino muy por el contrario el de refutarlos[1]; y dado que la refutación es la única que nos obliga a revisarlos y mejorarlos, es la refutación el factor dinámico más importante para el avance de las ciencias experimentales.

Uno de los ejemplos más ilustrativos que nos muestra claramente la esencia del criterio de refutación, nos lo ofrece una de las leyes de la teoría electromagnética de Maxwell[2] sobre el comportamiento del campo magnético $\vec{B}$, y que expresada en lenguaje matemático toma la siguiente forma:

$$\nabla \cdot \vec{B} = 0 \qquad (1)$$

Esta ley es interpretada, a nivel macroscópico, diciendo que las líneas del campo magnético son siempre líneas cerradas. Aceptando el modelo actual sobre la estructura de la materia, esta misma ley se puede interpretar, a nivel microscópico, diciendo que no existen monopolos magnéticos, o sea, partículas con una sola carga magnética definida.

En este caso bastaría con diseñar un sólo experimento, que mostrara la existencia de monopolos magnéticos para que la ley fuera refutada. Esto dió origen a lo que optó por llamarse "la caza del monopolo", término bajo el cual se incluyen todos los esfuerzos encaminados a evidenciar la existencia de tales partículas[3]. Hasta ahora todos estos esfuerzos han sido vanos, lo que nos lleva simplemente a afirmar que la ley no ha sido refutada. Sin embargo, la actitud correcta del investigador en este campo seguirá siendo la de continuar tratando de refutar dicha ley, ya que sólo a partir de la refutación se podrán establecer nuevas leyes con el objetivo de que algún día lleguen, a su vez, a ser refutadas, generando así una dinámica en el avance del conocimiento.

Por otra parte, la refutación nos lleva también a plantear criterios de cientificidad, con respecto a la estructura de los modelos teóricos. Por ejemplo, es obvio que bajo este criterio sólo aquellos modelos que sean refutables poseen un verdadero carácter científico[1]; y este carácter es más firme mientras más refutable sea el modelo.

Es precisamente la resistencia de un modelo a una muy ámplia variedad de circunstancias de poder ser refutado lo que le otorga su fuerza. A este respecto, un modelo se considera débil, científicamente hablando, cuando es refutable sólo en condiciones muy especiales. Por ejemplo, existen ya diversas teorías[4] sobre el comportamiento electrodinámico de los supuestos monopolos magnéticos. Para que dichas teorías pudieran llegar a ser refutadas es necesario que, primero, se encuentren los monopolos y, segundo, que éstos puedan ser examinados bajo condiciones controladas. Dado que hasta ahora, no ha sido siquiera posible detectar monopolos, estas teorías se consideran científicamente más débiles que cualquier otra que pueda ser refutable hoy en día.

En conclusión, podemos decir que desde el punto de vista del criterio de refutación, el verdadero objetivo de la física no es el de llegar a construir el modelo correcto de la naturaleza, sino que muy por el contrario su objetivo será siempre el de destruir, el de refutar cualquier edificio teórico que pretenda haber comprendido el comportamiento de la naturaleza.

Por ejemplo, la teoría de los vórtices de Descartes[5] fue refutada -y eliminada- por el hecho de que los planetas se mueven en elipses y no en los círculos previstos por Descartes. La teoría de la gravitación de Newton[6] explicó con éxito los hechos disponibles en aquel entonces, tanto los que habían sido explicados por la teoría de Descartes como aquellos que refutaban dicha teoría. Por lo tanto la teoría de Newton sustituyó a la de Descartes. De manera análoga la teoría de Newton fué a su vez refutada por el movimiento anómalo del perihelio de Mercurio, hecho que fue explicado por la teoría general de relatividad de Einstein[7].

En este contexto, la honestidad científica consiste pues en especificar, de antemano, un experimento tal que si el resultado contradice la teoría, la teoría debe ser abandonada[1]. Es importante señalar que para el criterio de refutación la existencia de experimentos cruciales es fundamental, ya que una teoría puede ser refutada por un solo experimento que sirve como base lógica para una nueva teoría.

Tal es el caso del experimento de Michelson-Morley[8], del cual se dice[9] que refutó las teorías del éter y condujo a la teoría especial de la relatividad[10]. Muchos físicos y filósofos le dieron al experimento de Michelson-Morley un carácter crucial, significando el rechazo de la estructura newtoniana del espacio y del tiempo.

4. Experimento versus Teoría

En el desarrollo de la sección anterior sobre el criterio de refutación, hemos supuesto implícitamente que es posible comparar la teoría con el experimento. Esta suposición nos llevó a concluir que la discordancia entre las predicciones de la teoría y un hecho experimental implicaba necesariamente la refutación de la teoría. En esta sección analizamos en que sentido dicha suposición es realmente adecuada.

Nuestra duda aparece cuando nos preguntamos, si existe realmente un criterio metodológico que nos indique si los resultados experimentales han sido interpretados correctamente.

Todo experimento, por sencillo que parezca, requiere de un proceso de interpretación. Dicho proceso se encuentra sustentado, primero, sobre un conjunto de teorías que respaldan, por ejemplo, la validez de las lecturas de los aparatos de medición y, segundo, sobre la suposición de haber podido controlar <u>todos</u> los factores significativos del fenómeno en cuestión. Por ejemplo, el simple hecho de asomarse por un telescopio para medir la posición de una estrella en el firmamento, requiere de haber aceptado ya un modelo teórico sobre la emisión luminosa, otro sobre el comportamiento de la luz en el espacio interestelar y en su viaje a través de la atmósfera terrestre y otro más sobre las leyes de refracción y reflexión de la luz en su interacción con las lentes y espejos que componen el telescopio. No se diga ahora de todo el bagaje teórico que debemos aceptar como no cuestionable cuando se quiere interpretar lo que aparece en una fotografía tomada por un moderno microscopio electrónico.

Esto nos indica simplemente, que la observación directa y pura de la naturaleza es algo que realmente no existe y que toda observación por más directa que pueda parecernos, lleva ya en sí una fuerte carga interpretativa basada en modelos teóricos preconcebidos.

Por otra parte, el método experimental, tal y como fue presentado en la segunda sección, no nos ofrece ningún criterio para decidir si en un experimento se han tomado en consideración todos los factores relevantes en el comportamiento de un sistema físico dado o como llegar a determinar la existencia de posibles elementos perturbadores ocultos. Ahora bien, dado que es posible convencerse de que el hecho experimental puro no existe, esto nos hace dudar consecuentemente de la validez misma del criterio de refutación. Esto se debe a que las predicciones de un modelo no se están comparando con el comportamiento de la naturaleza como tal, sino más bien con una interpretación que se tiene de dicho comportamiento. Por consiguiente, las predicciones de un modelo no sólo se están confrontando con hechos generados por la realidad sino también con toda una serie de predicciones de otros modelos que sustentan la interpretación de dichos hechos y a la que llamaremos la teoría del experimento.

Esto nos lleva a pensar que cuando los resultados experimentales no concuerden con las predicciones de un modelo dado, el método experimental no nos provee de un criterio para concluir que el modelo ha sido refutado, ya que cabe la posibilidad de que lo equivocado no sea el modelo sino la teoría del experimento o partes de ella.

Esta falta de criterio para poder decidir si realmente un modelo ha sido refutado, nos explica la actitud de muchos físicos ante la falta de concordancia entre teoría y experimento. Recordemos simplemente lo que sucedió hace algunos años cuando fué reportado en la literatura científica[11)] un experimento, llevado a cabo en un globo en las regiones altas de la atmósfera, donde se daba evidencia sobre la existencia de monopolos magnéticos. Ante este hecho, la actitud de muchos físicos fue la de no considerar refutada una de las leyes de la electrodinámica de Maxwell[2)] ($\nabla . \vec{\beta} = 0$) sino muy por el contrario la de suponer que por alguna u otra razón el experimento había estado mal interpretado. Pero esta suposición, no sólo genera una actitud, sino también toda una línea de investigación y de trabajo encaminada al análisis exhaustivo y crítico de la teoría del experimento.

Otro ejemplo ilustrativo fue también la actitud de los científicos cuando se reportó que los movimientos planetarios de Urano[12)] y de Mercurio[13)] no concordaban con las predicciones de la Teoría de la Gravitación de Newton[6)]. Ante la evidencia de estos resultados experimentales, la mayoría de los físicos no consideraron refutada la teoría newtoniana, sino que más bien dudaron de la correcta interpretación de las observaciones.

Estas dudas generaron una nueva dirección en las investigaciones y se llegó al extremo de suponer[14,15)] la existencia de dos planetas ocultos, Neptuno y Vulcano, que por alguna razón no habían sido observados y que perturbaban los movimientos de Urano y Mercurio en forma tal que las observaciones pudieran entrar en concordancia con la teoría newtoniana. En cierta manera se estaba "forzando" a los hechos experimentales a concordar con la teoría, llegando a calcularse las órbitas que deberían tener los planetas ocultos para poder explicar las observaciones en base a la teoría de Newton.

Lo que estos ejemplos nos muestran es que aunque el método experimental, no ofrece un criterio para decidir si la discordancia entre modelo y experiencia se debe a fallas del modelo o a errores en la teoría del experimento, en la práctica parece que si existe forma de decir en que dirección proseguir el trabajo de investigación; ya sea considerando al modelo refutado y haciendo esfuerzos para mejorarlo o considerando al modelo como correcto y haciendo esfuerzos para encontrar fallas en la interpretación de la experiencia. Llegados a este punto nos preguntamos ¿qué criterio sigue el investigador para hacer esta decisión?.

Sin embargo, podríamos pensar que la pregunta no tiene mucho sentido ya que independientemente de la decisión que tome el investigador, ésta puede considerarse simplemente como la selección indispensable de una hipótesis de trabajo, que hace posible la continuación del trabajo de investigación en una u otra dirección. Y que realmente no importa cual sea esta dirección, ya que es posible que ambas direcciones deban explorarse y serán los resultados mismos de las investigaciones los que prueben si la hipótesis seleccionada fue efectivamente la correcta, ya que la forma como se desarrolla la actividad científica no es fundamental, lo importante son los resultados.

Los éxitos a los que ha llevado esta forma de trabajo pueden ser ilustrados por la interpretación fallida sobre la existencia de monopolos magnéticos. En su interpretación original los experimentadores reportaron[11] el descubrimiento de un monopolo magnético, sin embargo en las investigaciones posteriores no se dudó sobre la validez de la electrodinámica clásica (la cual prohibe la existencia de monopolos) sino más bien sobre la correcta interpretación del experimento. Por lo tanto, tomando como hipótesis de trabajo la validez de la electrodinámica clásica, las investigaciones señalaron inconsistencias[16] en la interpretación del experimento, mostrando que la partícula detectada no correspondía a un monopolo, corroborando así la hipótesis seleccionada.

Otro ejemplo, tal vez más espectacular, nos lo ofrece la suposición de que la disconcordancia entre las predicciones de la teoría de gravitación de Newton y el movimiento planetario de Urano se debía a una interpretación errónea de las observaciones. Como hemos mencionado anteriormente, esto llevó a suponer la existencia de un planeta oculto, Neptuno, que perturbaba la órbita de Urano. Ahora bien, cuando los astrónomos apuntaron sus telescópios hacia el lugar del firmamento donde el supuesto planeta debería estar, de acuerdo a los cálculos realizados, encontraron efectivamente un nuevo planeta[17]; corroborando así la suposición original y dando lugar a uno de los descubrimientos astronómicos más sorprendentes de la época.

Por otra parte cuando se apuntaron los telescópios hacia las cercanías del sol en busca de Vulcano, éste nunca apareció[18]. Sin embargo hay que tomar en consideración que debido a la brillantez del Sol, las observaciones en su cercanía son de muy difícil realización y siempre sería posible pensar que el supuesto planeta ha permanecido oculto a todos los intentos de observación realizados hasta la fecha. Por tanto la discordancia entre el movimiento planetario de Mercurio y las predicciones de la teoría newtoniana de gravitación sigue existiendo[6]. Esto no hizo, sin embargo, que la teoría newtoniana fuera considerada refutada y por más de un siglo, en la mente de muchos astrónomos, el movimiento planetario de Mercurio fue considerado simplemente como una anomalía que surgía de una interpretación incorrecta de las observaciones y que sino era Vulcano, podrían existir aún otros elementos perturbadores que no habían sido tomados en cuenta en el proceso de interpretación.

Por ejemplo, el no haber tomado en cuenta la no-esfericidad del Sol en los cálculos de la órbita de Mercurio podría también haber sido la causa de la no-concordancia entre teoría y experimento. Sin embargo al realizarse los cálculos de la influencia del momento cuadrupolar del tensor de inercia del Sol en la órbita de Mercurio, se concluyó[19] que esta influencia no era suficiente para subsanar las diferencias entre teoría y experimento. Es más, al parecer no se han encontrado, hasta la fecha, fallas en la interpretación de las observaciones del movimiento planetario de Mercurio que pudieran explicarlo en base a la teoría newtoniana de gravitación. Sin embargo, ¿es ésto suficiente para considerar refutada la teoría newtoniana?, ¿no es acaso posible pensar que sencillamente no se ha dedicado el esfuerzo suficiente o que se ha carecido del ingenio necesario para imaginar un mecanismo que haga posible interpretar las observaciones de acuerdo a la teoría de Newton? o ¿bajo que bases podría asegurarse que se han explorado todas las posibilidades de error en la interpretación de la experiencia?. Y aún en el caso de que se lograra construir otra teoría, cuyas predicciones hicieran desaparecer las discrepancias con la interpretación que se tiene de la experiencia, ¿podría asegurarse entonces que la teoría newtoniana ha sido refutada?.

Obviamente, la única forma de poder refutar una teoría es suponiendo como correcta la interpretación de la experiencia con la cual entra en desacuerdo. Y el problema que entonces surge es que el método experimental no nos provee de ningún criterio que nos asegure la validez de dicha suposición.

Esto nos lleva, desde un punto de vista estrictamente metodológico, a un callejón sin salida, ya que hemos podido llegar a concluir que las teorías no pueden llegar a ser ni verificadas, ni refutadas; conclusión que nos lleva a cuestionar la esencia misma del método en que están basadas las ciencias experimentales.

Ahora bien, dado que para tener la posibilidad de refutar una teoría o modelo es indispensable suponer, sin cuestionamiento, la validez de la teoría del experimento, el problema radi ca entonces en la dificultad que presenta la justificación de dicha suposición; y el problema se vuelve obviamente muy complejo ya que significa el replanteamiento del criterio mismo de verificación.

Llegados a este punto vemos claramente que vá a ser imposible justificar la validez de la teoría del experimento ya que caeríamos entonces en un círculo vicioso; ya que para justificarla llegaríamos a exactamente los mismos problemas sobre el criterio de verificación que los que han dado origen a la presente discusión.

Vemos entonces que es absolutamente necesario partir de algo que tendríamos que aceptar como no cuestionable; ya que de otra manera, si queremos partir de algo que podamos considerar como objetivamente correcto, caeríamos entonces en la discusión del viejo problema filosófico sobre la existencia de la verdad absoluta[20)].

Por otra parte hemos visto, que en la práctica es posible determinar una dirección al trabajo de investigación sólo cuando se acepta un conjunto de hipótesis no justificadas.

Por tanto, si queremos evitar abordar el problema de la verdad absoluta y tomamos en consideración la práctica del trabajo científico, una de las posibles soluciones al problema de verificación consistiría en considerar al conjunto de hipótesis no justificadas como parte de una decisión metodológica. Es decir, en incorporar al método experimental la decisión sobre que teoría y reglas de interpretación de la experiencia serían consideradas como conocimientos no cuestionable, como conocimiento no problemático y al que llamaremos el núcleo irrefutable[21].

El núcleo irrefutable sirve entonces como una especie de visión aceptada del mundo, que contiene todas las teorías o modelos fundamentales de la naturaleza y las reglas para la interpretación de la experiencia. De esta manera si es ahora posible realizar una comparación entre las predicciones de la teoría e interpretaciones aceptadas de la experiencia. El problema se reduce entonces a decidir sólo sobre la validez del conocimiento construido sobre dicho núcleo.

Por otra parte es necesario reconocer también que los modelos o teorías no se encuentran aislados unos de otros, sino que debido a sus múltiples interconexiones éstos forman una estructura conceptual más ámplia, capaz de poder extenderse a otras áreas de la física, con el fín de tratar de comprender la naturaleza de una manera coherente. Por lo tanto, cuando no exista concordancia entre las predicciones de un sólo modelo con la interpretación aceptada de la experiencia, esto no se toma como una refutación de dicho modelo, sino simplemente como una dificultad que hay que superar para preservar la consistencia interna de toda la estructura.

Desde este punto de vista, el trabajo del físico no consistirá entonces en la construcción de modelos aislados de la naturaleza con la pretensión de que, algún día, pudieran llegar a adquirir validez universal; sino más bien en el diseño de proyectos o programas de investigación[21]. Dichos programas de investigación consistirán entonces en todo un conjunto de modelos estructurados en torno a un principio de búsqueda o heurística, construido sobre la base de otra estructura de conocimientos no problemáticos, y a la que habíamos denominado núcleo irrefutable.

En el presente trabajo, llamaremos heurística[21)] al conjunto de hipótesis y modelos que definen tanto el programa de investigación como sus limitaciones propias, así como también, la definición de los problemas que serán considerados importantes dentro del programa.

Ahora bien, cuando existan programas de investigación diferentes diseñados con el propósito de resolver un mismo tipo de problemas, diremos entonces que se trata de programas rivales o competitivos.

Como ejemplo de programa de investigación, podemos referirnos a la Teoría de Electrones de Lorentz[22)]. Con dicho programa, Lorentz trataba de explicar todas las propiedades electromagnéticas de los cuerpos materiales en base a la teoría del éter electromagnético de Maxwell, suponiendo que la materia estaba compuesta por una multitud de pequeñas partículas cargadas (electrones) que obedecían las leyes dinámicas de la mecánica de Newton. Aceptando además, los conceptos clásicos sobre la estructura del espacio y del tiempo, Lorentz extiende su programa para tratar de concebir a la gravitación como un fenómeno transmitido a través del éter[23)].

En este caso el núcleo irrefutable de la Teoría de Electrones podría estructurarse de la siguiente forma:[24)]

(i) Las ecuaciones de Maxwell

(ii) La mecánica de Newton

(iii) La fuerza de Lorentz

(iv) Los conceptos clásicos de espacio y tiempo

(v) La estructura atómica de la materia[25]

Por otra parte, la heurística estaría definida por lo siguiente[24]:

Todos los fenómenos físicos están determinados por acciones que se transmiten a través del éter.

Ahora bien, dado que el programa de investigación es un ente dinámico, los criterios de verificación no pueden ahora formularse en términos de la veracidad o falsedad del programa, sino más bien sobre su estado de progreso; es decir, deben caracterizar tanto el estado presente del programa como sus prospecciones futuras.

En este sentido, cuando alguna de las predicciones del programa no concuerda con el experimento, este hecho no se toma, de ninguna manera, como una refutación sino simplemente como una dificultad que el programa tiene que superar. Respetando siempre el núcleo irrefutable, se tiene que buscar, dentro de la heurística del programa, nuevas posibilidades de explicación a la discordancia entre teoría y experimento. Esto se hace introduciendo nuevas hipótesis auxiliares o nuevos conceptos físicos, reformulando la estructura de algunos de los modelos del programa o sustituyéndolo por otros diferentes, o tomando en cuenta variables que habían sido consideradas no significativas.

Por tanto, tomando en consideración la estructura dinámica de un programa de investigación, se dice[21] que éste permanece en un estado progresivo cuando la mayoría de sus predicciones se ven confirmadas por el experimento, cuando ha demostrado poseer la suficiente flexibilidad para superar las distintas dificultades que le ha presentado la discordancia entre teoría y experimento, cuando las nuevas predicciones que han surgido de la superación de dichas dificultades han sido, a su vez, confirmadas por el experimento y cuando es capaz de extender contínuamente su campo de acción a otras áreas de la física. En este caso el investigador tiene una justificación racional de seguir comprometido con dicho programa.

Por otra parte, se dice[21] que un programa de investigación se encuentra estancado o en estado regresivo cuando la mayoría de sus nuevas predicciones han entrado en conflicto con el experimento, cuando se encuentra continuamente a la defensiva ideando nuevas hipótesis o modificando modelos cuyas predicciones vuelven a presentar dificultades al compararse con la experiencia, cuando se ve en la necesidad de echar mano de los descubrimientos de programas rivales y cuando su área de aplicación se va reduciendo día con día. En este caso el investigador debe abandonar su compromiso con dicho programa.

Por ejemplo, de acuerdo a Zahar[24], el programa de Lorentz se mantuvo en estado progresivo hasta 1905, y hasta entonces no hubo razón alguna para abandonar dicho programa en favor del programa rival de Einstein sobre la nueva teoría de la relatividad. Sin embargo, a partir de 1905, el programa de Lorentz entra en un estado regresivo; mientras que, por otra parte, en la década siguiente, el programa de Einstein progresa de manera contínua hasta culminar en la Teoría General de Relatividad[7] y su consiguiente verificación experimental[26].

Por tanto, durante el período comprendido entre 1905 y 1915, los investigadores tuvieron razones suficientes para abandonar la Teoría de Electrones de Lorentz e incorporarse al programa einsteniano. La decisión final, a favor del programa de Einstein, ocurre en 1915 con su explicación de las observaciones sobre la deflexión de la luz[27] y el corrimiento anómalo del perihelio de Mercurio[13].

Como vemos, esta nueva versión dinámica del criterio de verificación preserva la esencia de la contenida en la presentación tradicional; en el sentido, de que es el experimento el que, a fin de cuentas, decide sobre el estado de progreso de un cierto programa de investigación. Es, otra vez, la misma naturaleza, o más precisamente la interpretación aceptada que se tiene de hechos generados por ella, la encargada de hacer la decisión sobre la acción de abandonar un programa o de continuar comprometido con él.

Sin embargo, podríamos preguntarnos ¿si es acaso posible eliminar objetivamente un programa que ha sufrido fuertes reveses iniciales, por el hecho de que la mayoría de los investigadores lo abandonan para incorporarse al programa rival donde los éxitos son más fáciles?, ¿no es acaso posible que un programa estancado y aparentemente derrotado pueda recuperarse y entrar nuevamente, después de algún tiempo y con una heurística mejor articulada, en un estado progresivo?; y para que dicho programa vuelva a entrar a un estado progresivo ¿no es absolutamente necesario que existan algunos investigadores dedicados a desarrollar su heurística?, entonces ¿hasta que punto resulta irracional para un investigador continuar esforzándose en desarrollar la heurística de un programa estancado?.

Por otra parte, existen también problemas relacionados con la selección del núcleo irrefutable. Por ejemplo, podríamos preguntarnos ¿cómo es posible saber si el estado regresivo de un programa, no se debe a su heurística sino más bien a fallas en los modelos considerados dentro del núcleo irrefutable?, ¿bajo que condiciones le es permitido al investigador reconsiderar la selección del núcleo irrefutable de un programa estancado sin abandonar los esfuerzos por seguir articulando su heurística? o ¿hasta que punto es realmente posible comparar dos programas que tengan en común un cierto tipo de problemas, pero que se encuentren construidos sobre núcleos irrefutables diferentes?.

Para contestar estas preguntas vemos que sería indispensable precisar, con mayor nitidez y sobre todo con objetividad, las condiciones bajo las cuales se va a considerar que un programa ha sido derrotado en forma definitiva. Obviamente, el abandono por parte de muchos de sus seguidores o la aparición de dificultades aparentemente insuperables, no son razones suficientes para inferir que el programa jamás podrá recuperarse; podría ser que se encontrara simplemente en un período de hibernación o que requiriera de modificaciones en la estructura o constitución de su núcleo irrefutable.

Con el fin de ilustrar la complejidad de esta problemática, tomemos como ejemplo el caso del programa de Lorentz sobre la Teoría de Electrones, mencionado anteriormente y que de acuerdo a Zahar[23)] entra en estado regresivo a partir de 1905.

La heurística de dicho programa supone la existencia de un medio transmisor (denominado éter) de las acciones de todos los sistemas físicos. Supuestamente, dicho programa fue derrotado por el programa rival de la nueva Teoría de Relatividad de Einstein. El programa einsteniano negaba por completo la existencia del éter e introducía una estructura nueva al espacio y al tiempo que entraba en franca contradicción con los conceptos clásicos, tal y como eran considerados por Lorentz dentro del núcleo irrefutable de su programa.

Ambos programas tenían en común problemas relacionados con la electrodinámica de cuerpos en movimiento, pero partían de núcleos irrefutables distintos. Mientras que Einstein consideraba dentro del núcleo irrefutable de su programa:

(i) las ecuaciones de Maxwell,

(ii) la propagación de la luz en el vacío sin necesidad de un medio transmisor y

(iii) el hecho de que la velocidad de la luz en el vacío es la misma en <u>todos</u> los sistemas inerciales,

el programa de Lorentz incluía, como ya habíamos visto, además de las ecuaciones de Maxwell, la fuerza de Lorentz, los conceptos clásicos de espacio y tiempo, el atomismo y la mecánica de

Newton. Por su parte, la heurística del programa einsteniano era que:

> la dinámica de la materia se encuentra gobernada por la estructura geométrica del espacio-tiempo.

La extensión del programa einsteniano al campo de la gravitación culmina con la llamada Teoría General de la Relatividad, donde la acción de la fuerza de gravedad se interpreta como cambios en la estructura geométrica del espacio-tiempo causados por la presencia de masa. Como ya habíamos mencionado, sus predicciones sobre el corrimiento del perihelio de Mercurio y la deflexión gravitatoria de la luz estuvieron en concordancia con las observaciones, hechos que, según Sahar[24)], significaron la derrota definitiva del programa de Lorentz.

Esto indica que implícitamente se está suponiendo que resultaba imposible para el programa Lorentziano, llegar a explicar estos mismos hechos experimentales, por lo que lo razonable era aceptar la derrota, abandonar el programa e incorporarse al programa triunfador.

Sin embargo Lorentz pensaba precisamente lo contrario[28)], es decir, que las predicciones de la Teoría General de Relatividad podrían también obtenerse como una extensión apropiada de la Teoría de Electrones. Así, ya en 1900, en sus Considerations sur la Pesanteur[23)], Lorentz discute el papel del éter como el medio portador de las ondas gravitacionales, que fueron otra de las predicciones importantes de la Teoría General, aún no corroboradas por el experimento. Discute también[28)] esfuerzos infructuosos sobre la construcción de un modelo de atracción gravitatorio, concedido como una especie de interacción dipolar de origen electromagnético transmitida a través del éter, con la idea de aplicarlo, precisamente, al movimiento anómalo del perihelio de Mercurio.

Esto nos indica claramente que para un programa tan ambicioso como la Teoría de Electrones de Lorentz, las predicciones de la Teoría General de Relatividad no iban a ser consideradas como conflictivas, sino que representaban simplemente las dificultades propias que se le presentan a cualquier programa cuando éste trata de extender su campo de acción a otras ramas de la física.

Por otra parte, el incorporarse al programa rival no indica, necesariamente, que se ha aceptado la derrota final. El mismo Lorentz realiza contribuciones importantes[29)] a la Teoría General de la Relatividad sin que ello haya significado el abandono de sus convicciones sobre la existencia del éter y los conceptos clásicos del espacio y el tiempo. Muy por el contrario, Lorentz llega a sugerir[30)], que tal vez resultó más afortunado para el desarrollo de la física, que Einstein haya abandonado, temporalmente, el concepto de éter, para poder así desarrollar con mucho mayor libertad el complejo formalismo de la Teoría General sin tener que arrastrar el lastre metafísico de los conceptos clásicos.

Resulta interesante también, hacer notar, que las embestidas más fuertes al programa no vinieron, tal vez, del programa einsteniano; sino más bien de la multitud de inconsistencias que surgieron al tratar de construir una teoría atómica de la materia. Hay que recordar también, que dichas inconsistencias se comenzaron a resolver cuando se admitió que la mecánica newtoniana era incapaz de describir la dinámica de las partículas atómicas y se inició la construcción de una nueva mecánica: la mecánica cuántica.

Sin embargo, como el atomismo y la mecánica newtoniana se encontraban dentro del núcleo irrefutable del programa lorentziano, éste sufrió también los embates del programa cuántico, que no atacaban directamente su heurística, sino más bien partes muy sensibles de su núcleo irrefutable. Y no hay que olvidar, que cuando el programa einsteniano trató de extender su campo de acción a la dinámica de las partículas atómicas, tropezó también con dificultades tan graves, que hasta la fecha no ha sido posible resolverlas satisfactoriamente.

Ante este panorama, ¿podemos decir, con absoluta certeza, que la idea de un éter ha sido derrotada en forma definitiva como resultado de los hechos experimentales?. A este respecto resulta interesante citar las palabras de Sir Edmund Whittaker que aparecen en el prefacio a su monumental obra sobre la historia de las teorías del éter y la electricidad[31)]:

> "Se podrían decir algunas palabras sobre el Título Eter y Electriciad. Como todos sabemos el éter jugó un papel muy importante en la física del siglo diecinueve; pero en la primera década del siglo veinte, debido fundamentalmente a los intentos fallidos de medir el movimiento relativo de la tierra con respecto al éter y la aceptación del principio de que todos esos intentos siempre fallarían, la palabra 'éter' cayó de gracia y se hizo costumbre referirse a los espacios interplanetarios como 'vacíos'; concibiendo al

vacío como mera vacuidad, sin ninguna otra propiedad que la de propagar ondas electromagnéticas. Sin embargo, con el desarrollo de la electrodinámica cuántica, el vacío ha llegado a ser considerado el asiento de las oscilaciones del 'punto cero' del campo electromagnético, de las fluctuaciones del 'punto cero' de la carga y corriente eléctricas y de una 'polarización' correspondiente a una constante dieléctrica distinta de la unidad. Parece absurdo seguir llamando 'vacío' a una entidad tan rica en propiedades físicas y la histórica palabra "éter" podría, con toda propiedad, ser retenida."

Además, resulta también pertinente añadir que con esta nueva concepción del éter, se realizan hoy en día intentos[32)] para fundamentar los principios mismos de la mecánica cuántica, en base al acoplamiento electromagnético entre las partículas cargadas y el éter.

En conclusión, creemos haber ilustrado con estos ejemplos, primero, los graves problemas que se desprenden de la ausencia total de una metodología para la selección y estructuración de un núcleo irrefutable y, segundo, el tipo de dificultades metodológicas que se presentan cuando se quiere decidir sobre la derrota final de un cierto programa de investigación, y en especial la existencia de los llamados experimentos cruciales.

Esto nos deja la impresión que va a resultar extremadamente difícil, sino imposible, el diseñar una metodología capaz de comparar teoría y experimento sin caer, de una u otra manera, en decisiones metodológicas de carácter arbitrario.

Sin embargo, somos también conscientes de que las ciencias experimentales son una actividad que se ha venido desarrollando en forma sistemática desde hace varios siglos y que ha tenido muy fuerte impacto sobre la tecnología y, en general, sobre la cultura del mundo occidental, imprimiéndole, además, un cierto sentido de progreso. Resulta difícil, por tanto, pensar que dicho impacto haya sido el resultado de una actividad caótica carente por completo de bases metodológicas firmes.

Ante esta disyuntiva, mostramos en la siguiente sección que es indispensable tener que ampliar nuestro concepto de ciencia y considerarla ya no como una actividad intelectual encaminada a la búsqueda de un conocimiento objetivo del mundo, sino como una actividad social, practicada por un determinado sector de la sociedad comprometido a resolver un cierto tipo de problemas. De esta manera, podríamos basarnos en un análisis de la práxis de la actividad científica a través de la historia, para poder llegar a descubrir cuáles han sido <u>realmente</u> los fundamentos metodológicos que han guiado a los físicos a establecer y a aceptar los diversos modelos que han existido sobre la estructura y comportamiento de la naturaleza.

5. La Actividad Científica

En esta sección exploraremos la posibilidad de encontrar los criterios metodológicos que se utilizan para aceptar, rechazar o abandonar un cierto programa de investigación en base a un análisis de la actividad científica.

En primer término, es necesario reconocer que la actividad científica no es una actividad aislada sino una actividad social realizada por un determinado sector, al que llamaremos: la comunidad científica. Es un hecho, además, que ha sido la propia comunidad científica la que ha determinado, en la práctica, la aceptación o el rechazo de una teoría o de un programa de investigación.

Con esto queremos decir, que es precisamente el grado de intensidad de la actividad científica en el desarrollo y la aplicación de una determinada teoría o en la extensión de un cierto programa de investigación, lo que define el grado de aceptación o de rechazo que la comunidad científica les otorga.

No podemos negar, por ejemplo, que en este sentido, el programa de investigación de Einstein sobre la estructura del espacio-tiempo, fue ampliamente aceptado, mientras que el programa de Lorentz sobre la existencia de un éter, fue prácticamente abandonado. Sin embargo, como vimos en la sección anterior, no podemos decir que esta decisión de la comunidad científica fue una decisión incontrovertible, determinada exclusiva mente por la evidencia de los llamados "hechos experimentales"; ya que no resulta impensable, que con una mayor actividad científica en el programa lorentziano, éste pudiera llegar también a explicar esos mismos hechos.

Por lo tanto, si la decisión de aceptación o abandono de su programa no puede estar basada solamente en lo que ha dado por llamarse la "evidencia experimental", esto nos conduce a preguntarnos si la comunidad científica utiliza criterios metodológicos bien definidos para tomar dicha decisión, o si por el contrario, el desarrollo de la actividad científica es un proceso anárquico, carente de tales criterios.

Trataremos ahora de dar respuesta a estas preguntas incorporando conceptos provenientes de la historia de la ciencia.

A primera vista, es posible reconocer que en la historia de la física, al igual que en la historia política, han existido épocas de tranquilidad y etapas revolucionarias[33)]. En las épocas de tranquilidad, la mayor parte de la actividad de la comunidad científica se encuentra concentrada en el desarrollo de un determinado conjunto de programas de investigación no-competitivos entre si. Es entonces, cuando los objetivos de la actividad científica consisten en el fortalecimiento y ampliación de dichos programas de investigación, con el fin de lograr la mejor articulación posible entre programas distintos, para poder ir conformando asi, una estructura teórica coherente capaz de explicar el comportamiento de la naturaleza como un todo.

A esta estructura de programas de investigación, considerados por la gran mayoría de la Comunidad Científica como un conjunto de conocimientos no-cuestionables y articulados a lo largo de directrices aceptadas también por los miembros de esta misma comunidad, es a lo que llamaremos, preservando el espíritu con que Kuhn[33)] introdujo el término, paradigma.

Es importante hacer notar, que un paradigma no es una estructura estática, sino que muy por el contrario, es una estructura dinámica que puede comenzar a construirse a través de un programa de investigación empujado por un pequeño sector de la Comunidad Científica, el cual logra extenderlo a otra áreas del conocimiento e incorpora, al mismo tiempo, a otros sectores de la comunidad a participar de dicha actividad. Ahora bien, sólo cuando la mayor parte de la comunidad, activa en esas áreas de la ciencia, ya han aceptado sin mas cuestionamiento, el desarrollar el mismo conjunto de programas y el seguir las mismas directrices de evolución, se dice entonces que se ha establecido un paradigma.

Por lo tanto, la amplitud y solidez de un paradigma depende de su grado de desarrollo. De esta manera podemos hablar del paradigma newtoniano para referirnos, a la aceptación de las leyes, que sobre el movimiento de los cuerpos, propone Newton en los Principia [34)], a todas sus subsecuentes reformulaciones, como las de Lagrange y Hamilton[35)], a sus múltiples aplicaciones a la descripción del movimiento de partículas puntuales, de cuerpos rígidos y de cuerpos deformables, a su arti culación con diversos programas de investigación en mecánica celeste, elasticidad, hidrodinámica, óptica, electricidad, magnetismo, teoría del calor, teoría atómica de la materia, teoría cinética de los gases, mecánica estadística, para constituir así una estructura de conocimientos a la que se le conoce bajo el nombre de: visión mecánica del mundo físico.

La construcción de este paradigma consumió la actividad de la comunidad científica por mas de dos siglos y definió una de las épocas clásicas en la historia de la física: el florecimiento de la física newtoniana.

Es importante hacer notar, que dentro de la evolución estructural de un paradigma no todos los programas adquieren la misma relevancia, ya que es fácil mostrar que la actividad científica siempre se encuentra concentráda con mayor intensidad en el desarrollo de algunos programas específicos.

Es muy claro que durante el siglo XVIII, dentro del paradigma newtoniano, el desarrollo de los programas de mecánica celeste, elasticidad, hidrodinámica y óptica consumieron la mayor parte de la actividad científica en comparación, por ejemplo, a los programas de electricidad y magnetismo. Y aún menor fue la actividad dedicada a los programas de teoría atómica y mecánica estadística, los cuales adquirieron una gran importancia sólo hasta finales del siglo XIX.

Otro ejemplo de paradigma, podría ser el paradigma cuántico, que se refiere a la aceptación de las leyes de la mecánica ondulatoria[36)] que describen el movimiento del mundo microscópico, a todas sus diversas reformulaciones matemáticas; a su articulación con programas de investigación en electrodinámica, química, física atómica y molecular, física del estado sólido, astrofísica, física nuclear, física de altas energías, teoría de muchos cuerpos, biología molecular, etcétera. A esta estructura de conocimientos se le conoce bajo el nombre de: visión cuántica del mundo microscópico, y ha sido construída por la comunidad científica dentro de los últimos cincuenta años.

Con respecto a la importancia relativa en el desarrollo de los distintos programas del paradigma cuántico, cabe señalar que los programas de física nuclear y física del estado sólido adquieren una importancia singular después de la Segunda Guerra Mundial.

Ahora bien, durante las épocas de florecimiento de un paradigma, la aparición de discordancias entre teoría y experimento son consideradas "anomalías", que obviamente no cuestionan ni los fundamentos, ni la estructura del paradigma. El objetivo de la actividad científica se reduce entonces a mostrar que las "anomalías" representan discordancias aparentes, ya que de acuerdo al paradigma siempre será posible encontrar una explicación causal de los resultados experimentales utilizando, única y exclusivamente, los elementos conceptuales y metodológicos que brinda el propio paradigma. Estos elementos pueden consistir en la introducción de hipótesis auxiliares de conceptos nuevos, en la inclusión de factores que no habían sido considerados como relevantes o en una re-interpretación adecuada de los datos experimentales, todo esto sin afectar, claro está, las bases fundamentales del paradigma. A la actividad científica que se desarrolla en estas condiciones es a lo que se ha llamado: ciencia normal[33], y su propósito es la solución de los enigmas originados por la aparición de "anomalías".

Un ejemplo actual de ciencia normal es la actividad realizada dentro del programa de investigación en física del estado sólido, el cual forma parte del paradigma cuántico y del paradigma maxwelliano. Del primero ya hemos hecho referencia con anterioridad y sobre el segundo, podemos señalar, que se fundamenta en el reconocimiento de que la dinámica del llamado campo electromagnético se encuentra descrita por las ecuaciones de Maxwell[37]. Con respecto a los principios estructurales que guían al programa, podríamos citar al atomismo, que concibe a los sólidos como sistemas compuestos por multitud de partículas microscópicas,

Ahora bien, dado que la ciencia normal se encuentra enmarcada dentro de la articulación de uno o varios paradigmas, los frutos de la investigación científica normal servirán unicamente para fortalecer y estructurar los fundamentos y principios en que se encuentran sustentados los paradigmas, pero jamás podrán ser capaces de originar descubrimientos que signifiquen cambios radicales en la visión que se tiene del mundo físico.

Así, el descubrimiento experimental de la superconductividad, realizado a principios del presente siglo[38], representó una de las "anomalías" mas espectaculares en el comportamien to eléctrico de materiales conductores y uno de los enigmas mas intrigantes que tuvo que arrastrar el programa de investigación en física del estado sólido, por mas de cincuenta años.

A pesar de todo, la casi totalidad de la comunidad científica, dedicada a la solución de este enigma, nunca se apartó de los lineamientos del programa de física de sólidos y siempre mantuvo la confianza de poder resolverlo, sin necesidad de poner en duda los fundamentos de los paradigmas que enmarcan al programa. Finalmente en 1957, aparece en la literatura científica, una solución apropiada a este enigma, basada esencialmente en el manejo adecuado de un nuevo concepto: el apareamiento de electrones[39)]. Por lo tanto, la formulación de una teoría, como la teoría de la superconductividad, es una actividad que cae dentro de lo que hemos llamado, ciencia normal, término que hemos utilizado para contraponerlo posteriormente a lo que llamaremos, ciencia revoucionaria.

Sin embargo, es un hecho, que los grandes avances de la física no podían haberse realizado dentro del marco de la ciencia normal, ya que éstos significaron transformaciones muy profundas en los fundamentos que sustentaban la visión que tenía la comunidad científica sobre algún aspecto del mundo físico. En otras palabras, estos grandes avances, representaron el derrumbamiento de un paradigma y la articulación de uno nuevo que lo reemplazase, actividad que cae fuera del alcance de la ciencia normal.

Pero esto nos conduce ahora a preguntarnos, ¿Cómo es posible que pueda darse el derrumbamiento de un paradigma, cuando la comunidad científica ha dedicado la mayor parte de sus esfuerzos a fortalecerlo y estructurarlo?

Para contestar esta pregunta, proseguiremos ahora el desarrollo de la analogía que habíamos establecido con la historia política, para darnos cuenta, que aún en las épocas de mayor tranquilidad, los sistemas políticos no constituyen una estructura monolítica, ni el control ideológico del Estado llega a ser un control absoluto. Siempre existen sectores sociales descontentos, con demandas cuya satisfacción requiere de un cambio radical en la estructura y composición del poder político. Estos sectores manifiestan su inconformidad con diversos actos de rebeldía al orden establecido, los cuales, en su etapa inicial, son fuertemente reprimidos y finalmente sofocados. No es sino hasta que los sectores disidentes comienzan a mostrar una organización combativa y una cohesión ideológica capaz de infundir confianza a otros grupos sociales, cuando es posible desencadenar una crisis en sectores considerados claves a la estabilidad del sistema. Son estas circunstancias, aunadas a una situación favorable en el panorama internacional, los indicadores, de lo que podríamos llamar, la etapa revolucionaria.

El objetivo fundamental de la revolución es la transformación radical de las estructuras de poder y el establecimiento de un nuevo orden social.

De la misma manera, aún en las épocas de mayor auge en el establecimiento de un determinado paradigma, y apesar de la intensa actividad dedicada a la ciencia normal, nunca se logra alcanzar un consenso absoluto. No sólo existe actividad en programas de investigación que impulsan la evolución del paradigma en direcciones distintas; sino que también, la persistencia contínua de resultados experimentales "anómalos", de enigmas no resueltos, de hipótesis auxiliares no justificadas, de contradicciones internas no explicadas, pueden llegar a provocar una crisis en el desarrollo de algunos programas de investigación. Ahora bien, si la crisis aparece en programas considerados por la comunidad, como los de mayor relevancia, se le presta entonces una atención particular.

El surgimiento de crisis, es capaz de generar, en pequeños sectores de la comunidad, un profundó sentimiento de insatisfacción que los puede llevar al extremo de cuestionar los conceptos y principios que sustentan al propio paradigma.

En sus inicios, estos pequeños sectores, trabajando fuera del paradigma establecido, comienzan a construir teorías que logran interpretar ciertos resultados "anómalos" aislados. La reacción de la comunidad científica ante este tipo de teorías es, en general, o de rechazo o, simplemente, de olvido.

Baste recordar el fuerte rechazo inicial a las hipótesis, que sobre la estructura atómica, enuncia Bohr[40)] en los comienzos de la mecánica cuántica. Bohr supone, sin mayor justificación y en contra de los principios fundamentales de la mecá nica newtoniana, que los electrones se pueden mover alrededor del núcleo atómico sólo en un cierto número de órbitas cuantiza das, y que además, durante su movimiento orbital, no son capaces de radiar energía electromagnética, otra hipótesis que se o pone a la teoría de la radiación del paradigma maxwelliano. Sin embargo, al suponer que los electrones radían sólo cuando pasan de una órbita a otra, aparece una explicación al misterioso e-nigma de las líneas espectrales en los espectros de radiación a tómica.

La simplicidad de esta explicación, aunada al hecho de que el modelo aparece en una época, en que la comunidad científica otorga singular atención a la crisis en que se encontraba la teoría atómica, hace que el modelo atómico de Bohr capte el interés de algunos sectores de la comunidad.

La etapa revolucionaria, propiamente dicha, se inicia sólo cuando las teorías construidas por los grupos disidentes, comienzan a inspirar un cierto grado de confianza a sectores mas amplios de la comunidad. Esto se puede deber, a que dichas teorías, además de ofrecer una interpretación coherente a un con junto de resultados considerados "anómalos", brindan un marco teórico promisorio para la superación de la crisis de algún pro grama considerado importante.

Pero la adopción de conceptos y principios ajenos al paradigma establecido, genera también un tipo de problema completamente nuevo y diferente. Por otra parte, muchos de los problemas y programas de investigación articulados en el paradigma establecido, pueden carecer completamente de sentido cuando se analizan en términos de los nuevos principios. En otras palabras, el programa revolucionario va generando en su desarrollo, sus propios problemas, sus propias contradicciones, sus propias aplicaciones, sus propias "anomalías" y sus propias prioridades. El verdadero propósito de una revolución científica no es el de generar una explicación alternativa a los problemas y contradicciones del paradigma establecido, ni la de resolver <u>todos</u> sus enigmas y ni la de comprender <u>todas</u> sus "anomalías", sino el de generar una visión radicalmente diferente del mundo físico con su respectiva problemática.

Así, el modelo atómico de Bohr, además de proponer una posible solución al enigma de las líneas espectrales, plantea un problema completamente nuevo: la justificación de los postulados de cuantización. Esto, aunado a las complicaciones requeridas por el modelo de Bohr para explicar cuantitativamente, los espectros de átomos complejos, dió lugar al desarrollo de teorías mas sofisticadas, como la mecánica matricial[41)] y la mecánica ondulatoria[36)]. A su vez, estas teorías generaron sus propios problemas conceptuales, cuya discusión se concentró en el debate sobre origen físico de las relaciones de incertidumbre de Heisenberg[42)] y lo que dió por llamarse, el problema de la dualidad onda-corpúsculo[43)]. Estas discusiones se alargaron por varios años, se adentraron en el terreno filosófico y ocuparon la actividad de las mentes mas brillantes de la época.

Por otra parte, uno de los problemas mas importantes en los programas sobre la estructura atómica, enclavados en el paradigma newtoniano y en el paradigma maxwelliano, era el problema de la estabilidad atómica. De acuerdo a la teoría de la radiación, basada en las ecuaciones de Maxwell, cualquier partícula cargada en movimiento orbital, radía energía electromagnética. Por lo tanto, si se considera al átomo como un pequeño sistema planetario, con el núcleo jugando el papel del sol y a los electrones, el papel de los planetas, estos últimos eventualmente perderían, por radiación, toda su energía de movimiento y terminarían en reposo junto al núcleo. Predicción que entra en desacuerdo con la existencia misma de átomos estables.

Este problema, por ejemplo, carece completamente de sentido dentro del esquema propuesto por el programa cuántico, ya que de acuerdo a este programa, además de que no es posible adjudicar una trayectoria a los electrones, se supone, sin mayor justificación, que en su estado base, los electrones se comportan como una distribución <u>estática</u> de carga, incapaz, por lo tanto, de radiar energía electromagnética.

Resulta entonces, que problemas que aparecen como fundamentales para un cierto programa de investigación, no son considerados ni siquiera como problemas cuando se les examina bajo la luz de principios radicalmente distintos.

De la misma manera, todos los problemas del programa lorentziano acerca de la estructura del éter, de su descripción dinámica alrededor de cuerpos en movimiento, de su relevancia en los fenómenos gravitatorios y el comportamiento de partículas atómicas, carecen de sentido cuando son considerados desde el punto de vista del programa einsteniano ó el programa cuántico, los cuales se basan en la suposición de que el éter, simplemente, no existe.

Es importante hacer notar, que el programa revolucionario genera también, nuevos tipos de aplicaciones que se integran en nuevos programas de investigación, no contemplados siquiera en el viejo paradigma. Por ejemplo, el programa cuántico genera una intensa actividad científica en programas de física molecular, física nuclear y física del estado sólido. La articulación de estos programas, bajo las leyes de mecánica cuántica, conduce a la estructuración del nuevo paradigma.

Queremos señalar, que aunque se asegure[44] que es posible "deducir" las leyes de la mecánica newtoniana como límite apropiado de las leyes de la mecánica cuántica, esto no implica que el paradigma cuántico sea una extensión del paradigma newtoniano; ya que, como hemos visto, cada paradigma posee su propia problemática. Lo que esto significa es, simplemente, que aquellos programas del paradigma newtoniano, relacionados con el comportamiento de cuerpos macroscópicos, no se vieron afectados tan gravemente por el programa revolucionario, y aunque su incorporación al paradigma cuántico no altere la formulación matemática de sus problemas, ésta si requiere de todo un proceso de re-interpretación.

Con esto queremos decir, que el paradigma cuántico y el paradigma newtoniano, representan dos visiones incompatibles del mundo físico; y que, por consiguiente, la adopción de un nuevo paradigma no representa un proceso acumulativo de conocimientos.

Hasta ahora hemos descrito únicamente como, dentro de un paradigma, una serie de anomalías o contradicciones internas puede dar lugar a una crisis, como una crisis, a su vez, puede generar la formulación de un programa de investigación revolucionario y como, finalmente, este programa se puede convertir en un nuevo paradigma. Además hemos ilustrado este proceso, con ejemplos tomados de la historia de la física.

Lo que no nos hemos preguntado es, ¿Porqué ciertas "anomalías" son juzgadas como mas importantes que otras, hasta el punto de provocar situaciones consideradas críticas?, ¿Porqué, en una determinada época, la actividad científica se encuentra concentrada en el desarrollo de ciertos programas de investigación? y ¿Porqué, en otras épocas, la comunidad dirige sus intereses al desarrollo de otros programas?

Creemos que la respuesta a estas preguntas es indispensable si queremos pasar de un nivel puramente descriptivo al análisis del proceso de formación y destrucción del Consenso; es decir, al análisis de las razones por las cuales la comunidad científica abandona un determinado conjunto de programas y principios rectores, para adoptar otro radicalmente distinto.

Es claro, que uno de los factores importantes en el proceso de formación de un consenso, es el impacto causado por los primeros éxitos logrados por las nuevas teorías, al explicar algunos de los resultados considerados como "anómalos". Sin embargo, la magnitud de este impacto depende, en gran parte, de la importancia que la comunidad científica le otorga al programa que se ha visto entorpecido por la aparición de dichos resultados "anómalos".

Por ejemplo, en el siglo XVIII no existía modelo alguno que pudiera explicar, cuantitativamente muchas de las propiedades conocidas del comportamiento de los gases. En este sentido, podríamos decir que los gases obedecían un comportamiento "anómalo". No obstante, en 1738, Daniel Bernoulli[45)] logra deducir la Ley de Boyle, suponiendo que los gases estaban compuestos por pequeñas partículas esféricas que chocaban elásticamente unas conotras y con las paredes del recipiente. Estas ideas, a pesar de que fueron retomadas con éxito cien años mas tarde, no tuvieron ningún impacto en una época en que la comunidad científica consideraba que el desarrollo de otros programas de investigación era de mayor relevancia.

Esto nos indica claramente, que son los programas considerados importantes, los que tienen mejores posibilidades de desarrollo y en donde las crisis recibirán una mayor atención y vigilancia.

Ahora bien, la importancia que la comunidad científica le confiere a los distintos programas de investigación, depende en si, de la combinación de muchos factores. Estos factores pueden frenar el desarrollo de un determinado programa, haciendo que la comunidad pierda interés en él, o, por el contrario, pueden estimular su progreso, atrayendo la atención de otros sectores de la comunidad.

Aunque es difícil separar cada uno de estos factores, creemos que es posible distinguir algunos de ellos. Primeramente mencionaremos el factor tecnológico, ya que muchas veces, la realización de experimentos considerados significativos para el desarrollo de un programa, se ve frenada por la carencia de una tecnología adecuada.

Baste mencionar el experimento requerido por las predicciones opuestas de la teoría corpuscular y la teoría ondulatoria, sobre la magnitud relativa de la velocidad de la luz en el agua y en el aire, hechas a principios del siglo XVII. Como es bien sabido, el experimento no fue realizado sino hasta 1850 por Fizeau[46)] y Foucault[47)], debido esencialmente a problemas de carácter tecnológico.

Por otra parte, los avances tecnológicos actuales, que han hecho posible la construcción de los grandes radiotelescopios, así como el lanzamiento de satélites artificiales y laboratorios espaciales, han dado un nuevo impulso al desarrollo de diversos programas de investigación en astronomía.

Otro factor que influye en la importancia dada por la comunidad a los diversos programas de investigación es la aplicación de la ciencia. Después de todo, el objetivo que busca la comunidad científica al construir una estructura de conocimientos que le ayuda a comprender el mundo físico, no es sólo la de satisfacer una curiosidad, sino que dicha estructura le es útil para resolver una amplia gama de problemas.

Muchos de esos problemas son generados por la aparición de contradicciones internas o de resultados "anómalos" y muchos otros provienen de las necesidades de la estructura social donde la comunidad científica se encuentra inmersa.

Por ejemplo, podríamos decir que el problema sobre la naturaleza ondulatoria o corpuscular de la luz, aparece en el Siglo XVII, debido a las contradicciones internas de ambos modelos al tratar de explicar el conjunto de fenómenos ópticos conocidos hasta entonces. En cambio, en este mismo siglo, los problemas sobre el comportamiento de la luz a través de lentes, adquiere una importancia particular, debido a sus aplicaciones en la construcción de instrumental óptico.

Sin embargo, cuando las relaciones entre la comunidad científica y los sectores encargados de la aplicación de la ciencia, se realizan en el seno de sistemas sociales con un alto grado de organización, resultará entonces muy difícil descubrir las verdaderas causas que dieron origen a un determinado problema.

Por ejemplo, el problema de la naturaleza cuántica de la luz, que se manifiesta al tratar de explicar el espectro de radiación del cuerpo negro, no se puede desligar de la importancia que tenía el estudio sobre los espectros de radiación, para la industria química europea del siglo XIX[48)].

Creemos también importante considerar el factor cultural, en el cual incluímos todas las influencias que puedan ejercer sobre la comunidad científica, las diversas corrientes filosóficas, religiosas o políticas, para la selección y elaboración de un programa de investigación.

Por ejemplo, se suele justificar la inclinación de la comunidad en favor de alguna teoría, utilizando argumentos de simplicidad, de generalidad o de belleza[49)].

Es también posible incluir, lo que podría llamarse, el factor psicológico, si creemos que la comunidad se ha enfrentado, en el curso de su actividad científica, a diversos obstáculos de carácter epistemológico que han impedido el progreso de ciertos programas de investigación. Por ejemplo, se ha llegado a pensar[50)] que estos obstáculos están presentes en el proceso mismo de conocimiento y que sólo la búsqueda de la abstracción hace posible que el espíritu científico los logre superar.

El problema reside ahora en analizar cuidadosamente cual es la influencia que cada uno de estos factores ha ejercido sobre la comunidad científica, para que ésta intensifique su actividad en el desarrollo de determinados programas de investigación. El objeto de este análisis sería el de determinar el proceso mediante el cual, la compleja combinación de todos estos factores hacen posible la instauración de un consenso conceptual y metodológico en el seno de la comunidad científica.

Es evidente que todos estos factores son generados, de una u otra manera, por el sistema social donde la comunidad científica se encuentra inmersa. Creemos, por lo tanto, que cualquier análisis que no considere a la comunidad científica como un grupo social estructurado dentro de un sistema social mas amplio, puede caer fácilmente en el peligro de no poder encontrar racionalidad alguna en el comportamiento de la comunidad científica en el ejercicio de su actividad[51)].

Por otra parte, tenemos que ser conscientes, que no va a ser posible efectuar el análisis de las relaciones de la comunidad científica con la estructura social que la contiene, sin echar mano de una teoría social. Es muy claro, que para poder analizar las relaciones estructurales entre diversos grupos sociales, se requiere de una teoría sobre la estructura de los sistemas sociales, que brinde un marco conceptual y metodológico que sustente la elaboración de dicho análisis.

Esto nos conduce ahora a la necesidad de seleccionar alguna teoría sobre la estructura de los sistemas sociales para poder llevar a cabo nuestro análisis; pero esto nos enfrenta, a su vez, al problema metodólogico de encontrar los criterios que nos indiquen qué teoría social es la mas adecuada y el tipo de validez que tendría entonces un análisis asi realizado.

Creemos importante señalar, que el no ser conscientes de este problema, puede dar origen al peligro de conceder un grado de validez demasiado amplio, a análisis fundamentados sobre alguna teoría social, seleccionada en base a una decisión no justificada desde un punto de vista estrictamente metodológico. Lo mismo puede decirse de los análisis sobre la evolución psíquica de la mentalidad científica, hechos a la luz de alguna teoría psicoanalítica.

Como vemos, nos encontramos ahora ante un problema metodológico mucho mas complicado que el problema que dió inicio al presente trabajo. Con esto queremos decir, que hemos comenzado analizando los problemas metodológicos de la física, cuyo objeto de estudio es el mundo inanimado, hemos concluído, que para resolver dichos problemas, necesitamos analizar ahora los problemas metodológicos de la ciencias sociales, cuyo objeto de estudio es la sociedad en su conjunto.

Esto nos indica claramente, que la separación que usualmente se hace de la metodología de las ciencias experimentales en contraposición a la metodología de las ciencias sociales, es una separación artificiosa; ya que los problemas metodológicos de ambas, forman parte de una misma problemática.

6. Conclusiones

En el presente trabajo hemos analizado los problemas estrictamente metodológicos del método experimental, tomando como ejemplo su aplicación en la física.

Comunmente se cree que la objetividad del conocimiento en la física, recae sobre la posibilidad de confrontar teoría con experimento; lo cual hace, que sea la naturaleza misma, a través del experimento, la que decida sobre la validez o falsedad de las teorías.

Sin embargo, hemos mostrado que la confrontación teoría-experimento no es suficiente para refutar una teoría, ya que para esto, es absolutamente indispensable aceptar un conjunto de conocimientos como no problemáticos. No obstante, dado que esta aceptación siempre se realiza, de una u otra manera, injustificadamente, esto nos condujo a concluir que es estrictamente imposible decidir si una teoría ha sido finalmente refutada por el experimento. La conclusión fue esencialmente la misma aún ampliando el concepto de teoría al concepto dinámico de programa de investigación.

Con esto mostramos que no existe ninguna posibilidad de rechazar un programa de investigación si nos basamos exclusivamente en la confrontación teoría-experimento. Fue, por tanto, necesario recurrir al análisis histórico para examinar los criterios que, de hecho, ha utilizado la comunidad científica para decidir sobre el rechazo o aceptación de un determinado programa de investigación. Nuestra conclusión fue, que dichos criterios son esencialmente los de consenso; por lo que nos avocamos a examinar los factores que determinan la formación y la destrucción de dicho consenso.

Finalmente señalamos, que no es posible analizar la influencia de factores generados por la estructura social donde la comunidad científica se encuentra inmersa, sin tener, a su vez, una teoría sobre dicha estructura social. Pero la selección de dicha teoría representa, nuevamente, un problema metodológico en si.

Por lo tanto, creemos haber mostrado, en que sentido, los problemas metodológicos de las ciencias experimentales forman parte del problema metodológico de las ciencias sociales y las ciencias del hombre. Creemos también haber apuntado los peligros a que conduce el no reconocer la magnitud de este problema.

Agradecimientos

Los autores desean manifestar que el presente trabajo se originó a consecuencia de las estimulantes discusiones que sostuvimos con Marco A. Martínez, José Marquina, Ignacio Campos y Leonardo Sánchez a quienes agradecemos sus valiosos comentarios y agudas críticas. Agradecemos también las sugerencias de Elaine Reynoso durante la elaboración del trabajo. Queremos también agradecer la cooperación de Pilar González, Violeta Castillo y Josefina Ramos en la preparación tipográfica.

Bibliografía

1. K.R. Popper, Conjectures and Refulations, Rotledge, 1963; K.R. Popper, The Logic of Scientific Discovery, Harper, 1965.
2. J.C. Maxwell, A Treatise on Electricity and Magnetism, Dover Publications, NY, 1954.
3. Ver, por ejemplo, el artículo de revisión titulado "On Dirac magnetic poles" en Old and New Problems in Elementary Particles, ed. G. Puppi, Academic Press, NY (1968); ver también los artículos de P.G.H. Sandars, Contemporary Physics 7 (1966) 419 ; R.H. Carrigan, Nuovo Cimento 38 (1965) 638; R.L. Fleischer, H. R. Hart, I.S. Jacobs, P.B. Price, W.M. Schwarz y F. Aumento, Phys, Rev. 184 (1969) 1393; R.L. Fleischer, P.B. Price, R.T. Woods, Phys. Rev. 184 (1969) 1398.
4. P.A.M. Dirac, Proc. Roy. Soc. A133 (1931) 60; Phys. Rev. 74 (1948) 817; J. Schwinger, Science 165 (1969) 757; A.S. Goldhaber Phys. Rev. 140 (1965) B1407.
5. Esta Teoría se encuentra expuesta en las obras de R. Descartes intituladas Dioptrique, publicada en 1637, Météores en 1638 y los Principia Philosophiae en 1644. Para una revisión escueta de esta teoría se puede ver E. Whittaker, A History of the Theories of Aether and Electricity, Harper and Brothers N.Y. 1960, Vol. I pp 4-12.
6. Sir Isaac Newton, Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World, Univ. of California Press, Berkeley 1974.
7. A. Einstein y M. Growssmann, Zeitschrift fur Math. u. Phys. 62 (1964) 215. A. Einstein, Berlin Sitzungsberichte (1915) pp. 778, 799, 831, 844.
8. A.A. Michelson y E.W. Morley, Amer. Jour, Sci. 34 (1887) 333; Phil. Mag. 24 (1887) 449. Ver también el artículo intitulado "The Michelson-Morley experiment" de R.S. Shankland, Am.J. Phys. 32 (1964) 16.
9. Ver por ejemplo, A.S. Kompaneyets, Theoretical Physics, Gordon and Breach, N.Y., 1962 p. 191. y el artículo intitulado "Einstein, Michelson and the crucial experimet" de G. Holton, Isis, 60 (1969) pp 133-197, en especial p. 160.
10. A. Einstein, Annalen der Physik, 17 (1905) 891. Ver también las discusiones en J.M. Levy-Lebond, Am.H. Phys. 44 (1976) 271; L.C Baird, Am.J.Phys. 44 (1976) 167.
11. P.B. Price, E.K. Shick, W.Z. Osborne y L.S. Pinsky, Phys. Rev. Lett. 35 (1975) 487.
12. Ver por ejemplo, A.F. Alexander, The Planet Uranus, Faber and Faber, London 1965;
13. U.J. Leverrier, Ann. Obs. París, 1859, Vol. 5 p. 109.
14. Ver por ejemplo R.L. Waterfield, A Hundred Years of Astronomy, The Macmillan Co. 1938. p 30.

15. U.J. Leverrier, Ann. Obs. París, 1859, Vol 5, p. 109. también vease, por ejemplo, F. Hoyle, Astronomy and Cosmology, W.H. Freeman and Co., 1975. p. 476.
16. P.B. Price, E.K. Shirk, W.Z. Osborne y L.S. Pinsky, Phys. Rev. D 18 (1978) 1382; M.W. Friedlander, Phys. Rev. Lett 35 (1975) 1167; L. Alvarez, LBL Report No. 4260, 1975 (no publicado); P.H. Fowler, Proceedings of the Fourteenth International Conference on Cosmic Rays, Munich, 1975, Ed. por Klans Pinkau, Max-Planck-Institute, München, 1975. Vol. 12 p. 4049; R.L Fleischer y R.M. Walker, Phys. Rev. Lett. 35 (1975) 1412.
17. Ver por ejemplo, F. Hoyle, Astronomy and Cosmology, W.H. Freeman and Co., 1975 p. 440; M. Grosser, The Discovery of Neptune, Dover, N.Y. 1979.
18. Ver por ejemplo, B.C. Murray y E. Burgess, Flight to Mercury, Columbia Univ. Press, 1976.
19. S. Newcomb, Suppl, Am. Ephem. Naut. Alm. 1897. Ver además, como referencias generales de algunos intentos de explicación:
J.C. Brandt y P.W. Hodge, Solar System Astrophysics, Mc Graw Hill Co., NY. 1964; A.J. Meadows, Early Solar Physics, Pergamon Press, Oxford, 1970.
20. Ver por ejemplo, B. Russell, An Outline of Philosophy, the World Publishing Co., 1960. Cap. 24.
21. I. Lakatos y A. Musgrave, Criticism and the Growth of Knowledge, Cambridge 1970. Ver también el artículo intitulado "Criticism and the Methodology of Scientific Research Programmes" de I. Lakatos, Proceeding of the Aristotelian Society (1968) Vol. 69.
22. H.A. Lorentz, The Theory of Electrons, B.G. Teubner, Leipzig, 1909.
23. H.A. Lorentz, Proc. Amst. Acad. 2 (1900) 559; la traducción francesa aparece en Arch. Néerl. 7 (1902) 325.
24. E. Zaher, Brit. J. Philos. Sci. 24 (1973) 95, 233.
25. K. Feyerabend, Brit. J. Philos. Sci. 25 (1974) 25.
26. A. Einstein, Berlin Sitzungsberichte (1915) 831
27. J. Soldner, Berliner Astr. Jahrb. 1804. p. 161, reimpreso en Ann. d. Phys. 65 (1921) 593. Sobre las observaciones del fenómeno se puede consultar G. van Biesbroek, Ast. S. 60 (1950) 49; E. Finlay-Freundlich y W. Ledermann, Mon. Not. R.A.S. 104 (1944) 40.
28. M. Brouwer, Am. J. Phys. 48 (1980) 425.
29. H.A. Lorentz, Versl. Kon. Akad. Wet. (Amsterdam) 24 (1916) 1759. Ver también la discusión de J. Mehra, The Physicis's Conception of Nature, Reidel, Boston, 1973. pp 135, 136.
30. H.A. Lorentz, Collected Works, Vol 9, Nyhoff, Amsterdam, 1934. pp 274, 275.
31. E. Whittaker, A History of the Theories of Aether and Electrycity, Harper and Brothers, N.Y. 1960. Vol I.
32. T.A. Brody, A.M. Cetto y L. de la Peña, Rev. Mex. Fís. 26 (1979) 59.

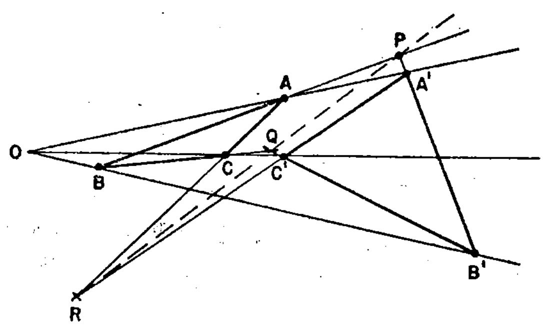
33. T.S. Kuhn, *La Estructura de las Revoluciones Científicas*, Fondo de Cultura Económica, México, 1971. Ver también T.S. Kuhn, *The Copernican Revolution*, Modern Library, 1959.
34. Sir Isaac Newton, *Philosophiae Naturalis Principia Mathematica*, presentado ante la Real Sociedad, en tres partes, entre 1686 y 1687.
35. Ver por ejemplo, H. Goldstein, *Classical Mechanics*, Addison-Wesley Pub. Co., Reading, 1959. Cap. 2.
36. E. Schrödinger, Annalen der Physik 79 (1926) 361, 489, 734. Ver también L. de Broglie, Le Journal de Physique et le Radium, 7 (1926) 321.
37. Ver por ejemplo, L.D. Landau y E.M. Lifshitz, *Teoría Clásica de Campos*, Ed. Reverté, Barcelona, 1967. Cap. 4.
38. H.K. Onnes, Leiden Comm. 124c (1911).
39. J. Bardeen, L.N. Cooper y J.R. Schrieffer, Phys. Rev. 108 (1957) 1175.
40. N. Bohr, Nature, 121 (1928) 580; Phys. Rev., 48 (1935) 696. Ver también la discusión de N. Bohr, *Atomic Physics*, Blackie and Son Ltd, London, 1962. Cap4.
41. W. Heisenberg, Zeitschrift für Physik, 35 (1925) 557. La conexión formal entre la mecánica matricial y la mecánica ondulatoria aparece en E. Schrödinger, Annalen der Physik, 79 (1926) 734 y C. Eckart, Phys. Rev. 28 (1926) 711.
42. W. Heisenberg, Zeitschrift fur Physik 43 (1927) 172. Ver también el artículo de revisión de J.W.M. Du Mond y E.R. Cohen, Rev. Mod. Phys. 25 (1953) 691.
43. Ver por ejemplo, L. de la Peña, *Introducción a la Mecánica Cuántica*, C.E.C.S.A. México, 1979. p. 62.
44. Ver por ejemplo, E. Merzbacker, *Quantum Mechanics*, John Wiley and Sons. N.Y. 1961. p. 42.
45. D. Bernoulli, *Hydrodynamica*, Strasbourg, 1738. Sec. 10. Ver también R. Hooke, *De Potentia Restitutiva*, London 1678.
46. H.L. Fizeau, Comptes Rendus, 30 (1850) 569.
47. L. Foucault, Comptes Rendus, 30 (1850) 551.
48. Como referencias generales de investigaciones en esta dirección se pueden consultar: H. Ross y S. Ross, *Economía Política de la Ciencia*, Ed. Nueva Imágen, México, 1979; *La radicalización de la Ciencia*, Ed. Nueva Imágen, México, 1979; J.M. Levy-Leblond, *Autocrítica de la Ciencia*, Ed. Nueva Imágen, México, 1980; G. Ciccotti, M. Cinni, M.A. de Moria, G. Jona-Lasino, *L' ape e L'architetto*; *Paradigmi Cientifici e Materialismo Storico*, Feltrinelli Economica S.T.A., Milano, 1977.
49. Ver por ejemplo, B. Easlea, *La Liberación Social y los Objetivos de la Ciencia*, Siglo XXI, España, 1973.
50. G. Bachelard, *La Formación del Espíritu Científico*, Siglo XXI, México, 1979; *El Compromiso Racionalista*, Siglo XXI, México, 1980.
51. Ver por ejemplo, Paul Feyerabend, *Against Method*, Verso, London, 1978.

EL TEOREMA DE DESARGUES

Guadalupe Lucio

Uno de los teoremas fundamentales de geometría proyectiva es el famoso teorema sobre triángulos de Desargues:

Si en un plano, dos triángulos ABC y A'B'C' son tales que las rectas que unen AA', BB' y CC' son concurrentes en un punto 0, los lados correspondientes se cortan en tres puntos colineales.



A pesar de la sencillez de la figura, constituída sólo por rectas, la demostración no es trivial; más aún, si se piensa en los teoremas de la geometría euclideana, se verá que este teorema no se parece a la mayoría de ellos y que la demostración del teorema en términos de Euclides, aunque es posible, sería algo latosa.

Nos remitiremos, entonces, a tratar de averiguar cómo es que Desargues se interesa por este resultado.

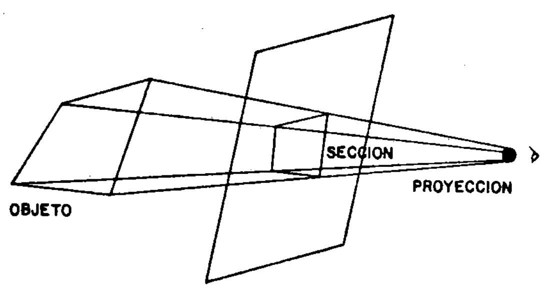
Podemos decir que el estudio de la perspectiva, o sea el arte de representar el espacio sobre una superficie plana, fue el motor del trabajo de Desargues.

Este estudio está profundamente ligado con el desarrollo de la pintura durante el Renacimiento. Los artistas renacentistas, que trataban de pintar de manera realista, tuvieron que resolver el siguiente problema técnico: ¿Cómo se puede hacer un dibujo bidimensional de un objeto tridimensional de tal manera que el dibujo se vea como el objeto?

Durante la Edad Media se pasó por alto este problema, como podemos ver en las figuras pertenecientes a esta época. Estas figuras están aplanadas sobre la tela, o así nos lo parecen.

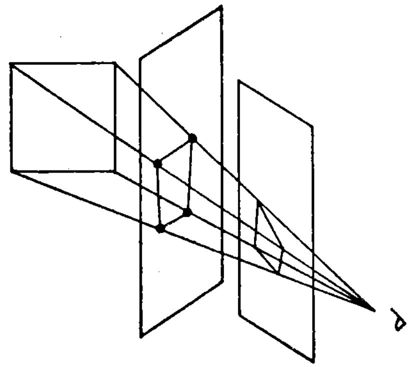
Los pintores renacentistas lucharon por más de cien años para encontrar un esquema que les permitiera representar un mundo tridimensional en una tela bidimensional.

Ya que la imagen de un objeto depende de la posición del observador, una de las primeras suposiciones que se hicieron para resolver el problema fue la de considerar al observador en un punto fijo. Para simplificar más el problema se suponía también que se usaba solamente un ojo. Después, el pintor se imaginaba que de cada punto del objeto partía un rayo de luz hacia el ojo. A esta colección de rayos de luz, o de rectas concurrentes, les llamaban proyección.



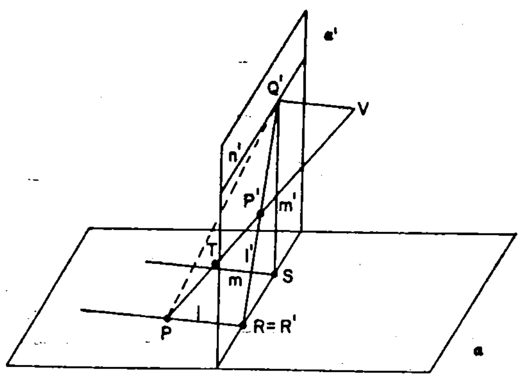
Si se supone, como lo hicieron los pintores del siglo XIV, que se interpone una ventana transparente entre el objeto y el ojo, ésta intersecta a la proyección en una colección de puntos a los que llamaban sección. La sección obtenida sobre la ventana transparente es una representación bidimensional de un objeto tridimensional. Así, el problema del pintor se reduce al problema geométrico de determinar los puntos de la sección obtenida sobre la ventana.

Hay que observar que la sección obtenida depende no sólo de donde se coloque el observador, sino de la posición de la ventana. Así se pueden obtener diferentes secciones de la proyección de un objeto, dependiendo de donde se coloque ésta.



Examinemos ahora la perspectiva desde un punto de vista geométrico.

Supongamos que el punto V representa el ojo del artista, el plano α es el plano que deseamos representar y el plano α' la tela.



A todo punto P en el plano α lo podemos representar por un punto P′ en el plano α', de la siguiente manera: Trazamos la recta VP y a P le asociamos el punto P′ donde VP corta al plano α'. Al punto P′ le llamamos la proyección de P en el plano α' desde V y al plano α' la proyección de α desde V.

A la recta de intersección de α y α' le llamamos el eje de proyección y vemos que todos los puntos sobre el eje se proyectan en sí mismos.

Tomemos ahora una recta ℓ en α que corte al eje de proyección en un punto R. Si proyectamos todos los puntos de ℓ en α', desde V, obtenemos una recta ℓ' que pasa por R, ya que, la intersección del plano que contiene a V y ℓ con el plano α' es una recta que contiene a R. Así, la proyección de una recta ℓ en α, que corta al eje de proyección en R, es una recta ℓ' en α' que también pasa por R.

Si tomamos ahora una recta n′ en α' tal que el plano que contiene a V y n′ es paralelo al plano α, las rectas VQ′, para Q′ en n′, no intersectan al plano α. Esto es, n′ no es la proyección de recta alguna en α. Por esta razón llamaremos a n′ la línea de fuga de la proyección.

Tomemos ahora una recta m en α paralela a ℓ, su imagen es una recta m′ en α'. Ahora, si ℓ' corta a n′ en Q′, veremos que m′ corta a n′ en el mismo punto.

Tomemos el plano que contiene a V, Q′ y S el punto donde m corta al eje de proyección. Este plano contiene a m, por ser m paralela a ℓ y ℓ paralela a VQ′, por lo tanto corta al plano α en la recta m y al plano α' en la recta SQ′; esto es, la proyección de la recta m es la recta SQ′.

Es decir, rectas paralelas en α deben representarse como rectas concurrentes en un punto de la línea de fuga en α' y a direcciones distintas en α corresponden puntos distintos en la línea de fuga

La afirmación anterior corresponde al desarrollo de la perspectiva durante un largo período.

Supongamos ahora que tomamos un punto T en la recta VP, tal que T no está en α.

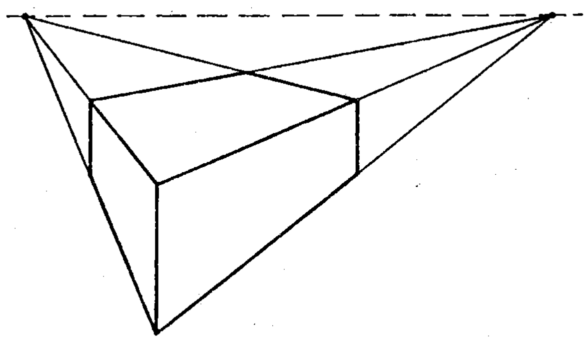
La proyección de T en α' es el mismo punto P′ considerado antes. La recta Q′P es una recta en el espacio que se proyecta en Q′R en el plano α'. De hecho, cualquier recta en el plano definido por P, Q′, R estará representada por la recta Q′R. Esto es, cuando proyectamos el espacio en un plano, perdemos la correspondencia biunívoca que teníamos en el plano (excepto por la línea de fuga); ahora, a una recta le puede corresponder un punto y a un plano una recta.

Pero podemos seguir afirmando que cualquier recta en el espacio que sea paralela a la recta VQ′ estará representada por una recta que pase por Q′.

Sea n una recta, en el espacio, paralela a VQ′. Tomemos el plano que contiene a VQ′ y a n (esto se puede hacer por ser paralelas), la intersección de este plano con

α' es una recta que contiene a Q'; por lo tanto, si unimos V con los puntos de n obtenemos en α' una recta que pasa por Q'. Esto es, existe tan sólo un punto sobre la línea de fuga en el plano α' que corresponde a un haz de paralelas en el espacio.

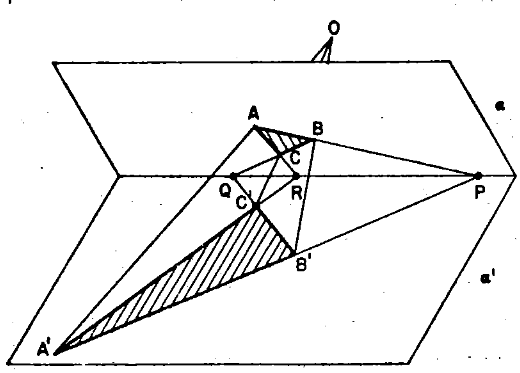
Esto es la justificación del dibujo en perspectiva de un cubo, que ya se hacía en el siglo XV.



Es esta problemática, la del estudio de la perspectiva, la que impulsa el trabajo de Desargues. Su motivación fundamental era combinar los teoremas sobre perspectiva de forma que les fueran de utilidad a los artistas, ingenieros, etc.

Volvamos ahora, a la luz de esta problemática, al teorema que nos preocupa.

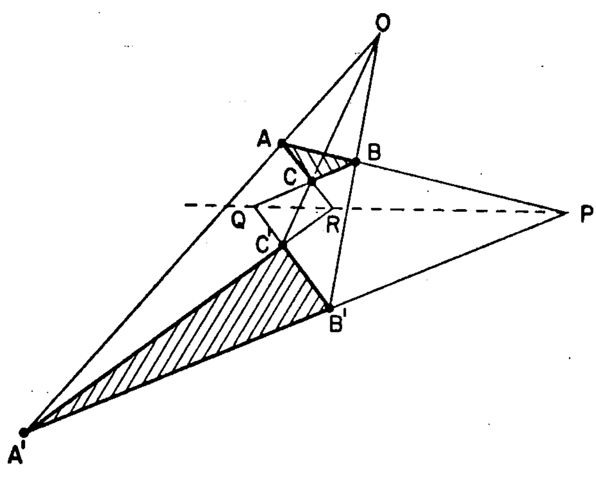
Supongamos ahora, que los triángulos ABC y A'B'C' están en planos distintos α y α' y que AA', BB' y CC' son concurrentes en un punto 0. Veremos que las intersecciones de lados correspondientes son colineales.



Este teorema lo podemos interpretar de la siguiente manera: Proyectamos el triángulo ABC en el plano α, desde 0, en el plano α'. Ya que OAA', OBB' y OCC' son colineales, la proyección en α' es el triángulo A'B'C'. Luego, la proyección de AB sobre α' es A'B', la de BC es B'C' y la CA es C'A'.

Sea P la intersección de AB con el eje de proyección, por lo tanto A'B' debe pasar por P. ASí, la intersección de AB y A'B' es un punto sobre el eje de proyección. De manera análoga se puede probar que la intersección de BC con B'C' y la CA con C'A' están sobre el eje de proyección. De aquí que efectivamente las intersecciones de lados correspondientes sea colineales.

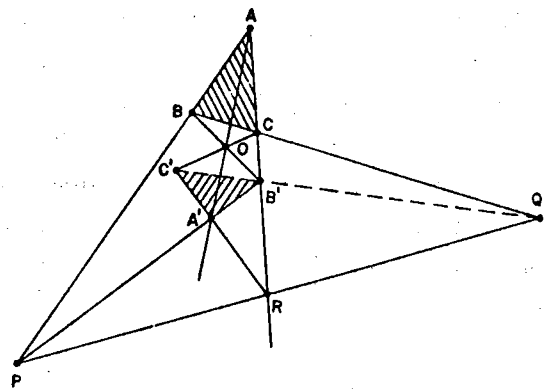
Si tomamos ahora una fotografía de la configuración anterior, es decir, si proyectamos adecuadamente sobre otro plano, obtenemos la figura para el teorema de Desargues en el plano.



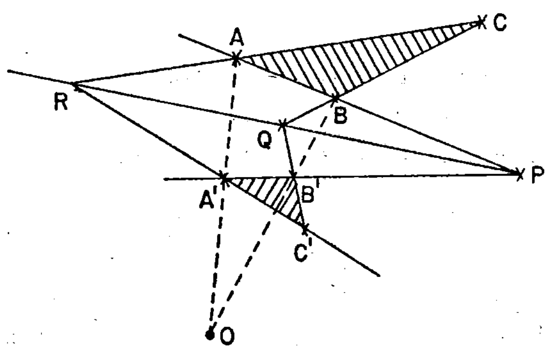
Ya que la proyección preserva colinealidad, los puntos P, Q y R deben ser colineales.

Al punto 0 le llamamos el centro de perspectiva de los triángulos ABC y A'B'C' y a la recta PQR, el eje de perspectiva.

Esta configuración tiene una admirable simetría. Se puede ver que cualquier punto de la configuración es equivalente. Esto es, cualquier punto de la configuración puede ser el centro de perspectiva y se pueden renombrar los demás puntos para que el teorema siga siendo válido.



Supongamos ahora que tenemos dos triángulos ABC y A'B'C' en el plano y que las intersecciones AB ∩ A'B' = P, BC ∩ B'C' = Q, CA ∩ C'A' = R, son colineales.



Veremos que las rectas AA', BB' y CC' son concurrentes. Esto se llama, a veces, el recíproco del teorema de Desargues y, como veremos, puede probarse a partir del teorema de Desargues.

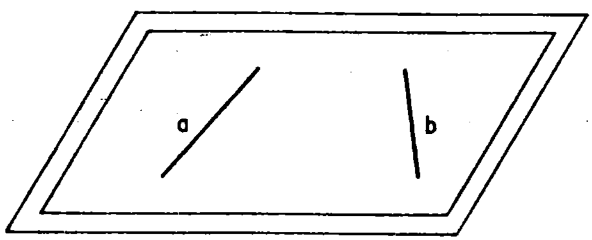
Sea O = AA' ∩ BB'. Consideremos los triángulos AA'R y BB'Q. Las rectas AB, A'B' y RQ son concurrenres en P, por lo tanto, por el teorema de Desargues las intersecciones de lados correspondientes son colinales. Estas intersecciones son:

$$AA' \cap BB' = 0, \quad A'R \cap B'Q = C', \quad RA \cap QB = C,$$

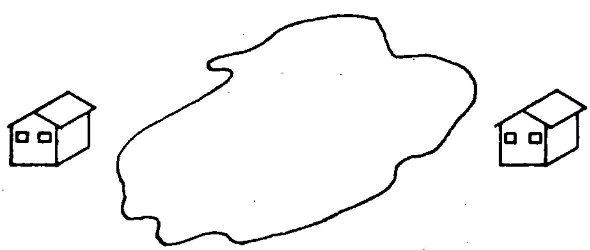
por lo tanto 0, C y C' están alineados y AA', BB' y CC' son concurrentes.

El teorema de Desargues nos puede ser de gran utilidad para resolver problemas de construcción relacionados con puntos y rectas inaccesibles.

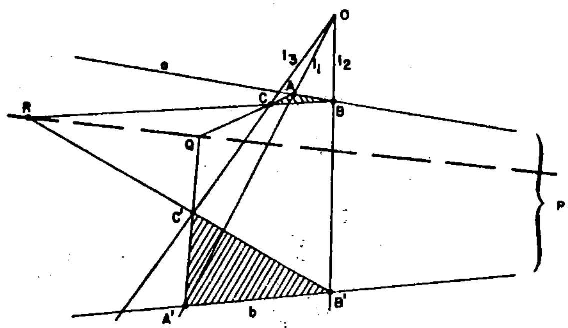
Se dice que un punto dado es un punto inaccesible si está determinado por dos rectas dadas que se intersectan en tal forma que, en el desarrollo del problema dado, el punto de intersección no se puede usar para la construcción. Por ejemplo, si un dibujante tiene sobre la mesa de dibujo las rectas a y b que se cortan fuera de los límites de la misma.



Se dice que una recta es inaccesible si está determinada por dos puntos dados en tal forma que no se puede usar para la construcción. Por ejemplo, en topografía es frecuente que no se puedan alinear dos objetos porque lo impide un obstáculo.



Veamos primero, como trazar una recta que pase por un punto Q y un punto inaccesible P dado por la intersección de dos rectas a y b.



Tomemos un punto arbitrario O que no esté en las rectas a y b. Tracemos las rectas ℓ_1, ℓ_2, ℓ_3 que pasen por O y que corten a a y b. Llamemos:

$$A = a \cap \ell_1 \qquad A' = b \cap \ell_1$$
$$B = a \cap \ell_2 \qquad B' = b \cap \ell_2$$

Tracemos AQ y A'Q. Sea $C = AQ \cap \ell_3$, $C' = A'Q \cap \ell_3$. Sea $BC \cap B'C' = R$.

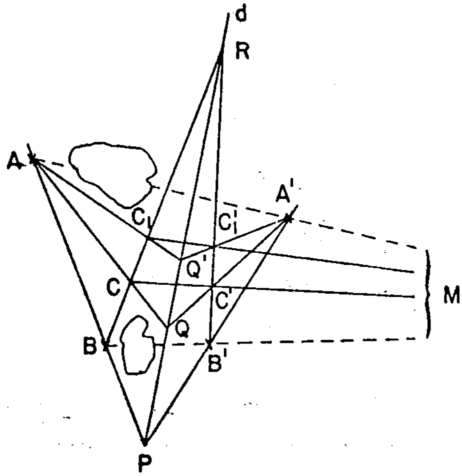
Considerando los triángulos ABC y A'B'C', por el teorema de Desargues, P, Q y R son colineales. Por lo tanto la recta buscada es QR.

Supongamos ahora que queremos construir el punto de intersección M de dos rectas inaccesibles AA' y BB'.

Sea $P = AB \cap A'B'$. Sea d una recta arbitraria que pasa por P y sean Q y R dos puntos (diferentes de P) en d. Sea $C = AQ \cap BR$ y $C' = A'Q \cap B'R$. Se tiene entonces que los triángulos ABC y A'B'C' son tales que:

$$AB \cap A'B' = P, \quad BC \cap B'C' = R, \quad CA \cap C'A' = Q$$

y que P, Q y R son colineales por construcción. Por el recíproco del teorema de Desargues se tiene que las rectas AA', BB' y CC' son concurrentes; esto es CC' pasa por M.



Cambiando la posición de la recta d o bien la posición de cualquiera de los puntos Q y R, podemos construir en la misma forma otra pareja de puntos C_1 y C_1', tales que la recta que pasa por ellos, también pase por el punto M. Entonces podemos construir el punto M como el punto de intersección de las rectas CC' y C_1C_1'.

LOS APROXIMANTES DE PADÉ

Francisco Gutiérrez Santos

Resumen

En ciertas aplicaciones de las matemáticas, aparecen una o más series infinitas. Desde el advenimiento del cálculo, con la aparición de la serie de Taylor, el valor de una serie se ha prestado a ciertas controversias. Este pequeño opúsculo introductorio, tratará de ilustrar el método de Padé para aproximar el valor de una serie, cuando el argumento es menor que la unidad, en particular para el caso de la función conocida cos x.

Sin embargo, la utilidad del método de Padé y de ciertas generalizaciones que se han hecho recientemente, es el de extraer el máximo de información de una serie cuya representación analítica, es desconocida en términos de funciones conocidas. Uno de los aspectos menos difundidos en relación a este método, es el de que permite obtener información parcial, aún en aquellos casos donde la serie no sea convergente, como es el caso de cierta clase de series divergentes de Stieltjes.

El método aproximante de Padé a sido útil en la obtención de información cuantitativa respecto de la solución de múltiples problemas

que han surgido en la aplicación de este método a la Física como a la Química (ver la Bibliografía). La información se extrae de la expansión, en términos de perturbaciones (lo que en general es más fácil de obtener que una respuesta exacta) y del conocimiento cualitativo que se tenga de la solución.

El aproximante de Padé tiene la forma de un cociente de polinomios. El numerador tiene grado m y el denominador grado n, es decir, el aproximante de Padé se denota así, $[m,n]$.

El método de Padé, lo estableció originalmente Frobenius, al ofrecer una representación aproximada, de una función representada en términos de una serie de potencias, a partir de fracciones racionales.

Sin embargo, no fue hasta recientemente, con el advenimiento de las computadoras, que se hizo posible la generalización de la aplicación de este método. Diremos que el método de Padé consiste en poder aproximar, cualquier serie de potencias;

$$(1) \qquad f(x) = \sum_{i=0}^{\infty} c_i x^i, \qquad x \in \mathbb{R}$$

que converja o no, en términos de fracciones racionales. Lo que se da en llamar, tabla de Padé de una función, f(x), se construye a partir del cociente

$$(2) \qquad [m,n] = \frac{P_{mn}(x)}{Q_{mn}(x)} = \frac{a_0+a_1x+\ldots+a_mx^m}{b_0+b_1x+\ldots+b_nx^n} = \frac{\sum_{i=0}^{m} a_i x^i}{\sum_{i=0}^{n} b_i x^i}$$

donde m y n pueden ser números enteros tan grandes como se quiera, según el grado de aproximación que se desee.

El problema consiste entonces en hallar los coeficientes a_i , b_i de cada uno de los polinomios P_{mn} y Q_{mn}. Para evitar dividir entre cero cuando $x = 0$, definimos $b_0 \equiv 1$.

Dado que buscamos una proximación a la serie (1) a partir de (2), es necesario encontrar una relación que conecte nuestra información, es decir, la serie que conocemos con los polinomios $P_{nm}(x)$ y $Q_{mn}(x)$. En otras palabras conectar los coeficientes de (1), las c_i con los coeficientes de los polinomios, P_{nm} y Q_{mn}, las a_i y b_i. Para relacionar cada a_i y b_i con las c_i procederemos de la siguiente manera.

Restemos (2) de (1)

$$\sum_{i=0}^{\infty} c_i x^i - \frac{\sum_{i=0}^{m} a_i x^i}{1 + \sum_{i=1}^{n} b_i x^i} = g(x) \tag{3}$$

Donde $g(x)$ es una serie bien particular cuya forma nos interesa determinar. Multipliquemos a (3) por

$$i + \sum_{i=1}^{n} b_i x^i = Q_{mn}(x) \tag{4}$$

tendremos que

$$\left[1+\sum_{i=1}^{n} b_i x^i\right]\left[\sum_{i=0}^{\infty} c_i x^i\right]-\sum_{i=0}^{m} a_i x^i=\left[1+\sum_{i=1}^{n} b_i x^i\right] g(x)=$$

$$=\left[1+b_1 x^1+b_2 x^2+\ldots+b_n x^n\right]\left[c_0+c_1 x^1+c_2 x^2+\ldots\right]-\left[a_0+a_1 x+a_2 x^2+\ldots+a_m x^m\right]$$

$$=\left[1+b_1 x^1+b_2 x^2+\ldots+b_n x^n\right] g(x)$$

$$=(c_0-a_0)+(c_0 b_1+c_1-a_1) x+(c_0 b_2+b_1 c_1+c_2-a_2) x^2+\ldots=$$

$$(5) \qquad \left[1+\sum_{i=1}^{n} b_i x^i\right]\left[\sum_{i=0}^{\infty} c_i x^i\right]-\sum_{i=0}^{m} a_i x^i=\left[1+\sum_{i=1}^{n} b_i x^i\right] g(x)$$

Para tener una mejor idea del método que vamos a seguir para determinar a los coeficinetes a_i y b_i de los polinomios $P_{mn}(x)$ y $Q_{mn}(x)$, veremos varios casos concretos.

CASO 1. Sea $P_{mn}(x) = a_0 + a_1 x$, $Q_{mn}(x) = 1 + b_1 x$ es decir, $m = 1$, $n = 1$. Entonces

$$(1+b_1 x)\left(\sum_{i=0}^{\infty} c_i x^i\right)-(a_0+a_1 x)=(1+b_1 x) g(x)=$$

$$=(1+b_1 x)(c_0+c_1 x+c_2 x^2+\ldots)-(a_0+a_1 x)$$

$$=(c_0+b_1 c_0 x+c_1 x+b_1 c_1 x^2+c_2 x^2+\ldots)-(a_0+a_1 x)$$

$$=(c_0-a_0)+(b_1 c_0+c_1-a_1) x+(b_1 c_1+c_2) x^2+(1+b_1 x) \sum_{i=3}^{\infty} c_i x^i .$$

Los términos de orden mayor o igual a 3 son los que restan y no quedaron especificados pero son iguales a

$$\sum_{i=3}^{\infty} c_i x^i$$

es decir,

$$(6) \qquad (c_0 - a_0) + (b_1 c_0 + c_1 - a_1)x + (b_1 c_1 + c_2)x^2 + (1 + b_1 x) \sum_{i=3}^{\infty} c_i x^i =$$

$$= (1 + b_1 x) g(x)$$

de donde se sigue, que para que resulte una función en potencias de x, donde la potencia menor sea 3 se deberá cumplir que

$$(7) \qquad \begin{cases} c_0 - a_0 = 0 \\ b_1 c_0 + c_1 - a_1 = 0 \\ b_1 c_1 + c_2 = 0 \end{cases}$$

que son tres ecuaciones con tres incognitas a_0, a_1 y b_1. Como vemos $(1 + b_1 x) g(x)$, es una serie cuya potencia más pequeña es 3, ya que de (6) se obtiene: $(1 + b_1 x) \sum_{i=3}^{\infty} c_i x^i = (1 + b_1 x) g(x)$ es decir

$$(8) \qquad g(x) = \sum_{i=3}^{\infty} c_i x^i$$

CASO 2. Sea $P_{mn}(x) = a_0 + a_1x + a_2x^2$; $Q_{mn}(x) = 1 + b_1x$ aquí $m = 2$ y $n = 1$. De (5) vemos que

$$(1+b_1x)\left(\sum_{i=0}^{\infty} c_i x^i\right) - (a_0 + a_1x + a_2x^2) = (1+b_1x)g(x)$$

Desarrollando el lado izquierdo de la igualdad

$$(1+b_1x)\left[c_0 + c_1x + c_2x^2 + \ldots\right] - a_0 - a_1x - a_2x^2 =$$

$$(c_0 - a_0) + (c_1 - a_1 + b_1c_0)x + (c_2 - a_2 + b_1c_1)x^2 +$$

$$(c_3 + b_1c_2)x^3 + (1+b_1x)\sum_{i=4}^{\infty} c_i x^i = (1+b_1x)g(x) \tag{9}$$

Para poder determinar los coeficientes de $P_{21}(x)$ y $Q_{21}(x)$, en términos de las c_i; $i = 0,1,2,3$ bastará con establecer las siguientes relaciones:

$$\begin{aligned} c_0 - a_0 &= 0 \\ c_1 - a_1 + b_1c_0 &= 0 \\ c_2 - a_2 + b_1c_1 &= 0 \\ c_3 + b_1c_2 &= 0 \end{aligned}$$

despejando apropiadamente tendremos que:

$$(10) \qquad \begin{cases} a_0 = c_0 \\ a_1 = \dfrac{c_1c_2 - c_0c_3}{c_2} \\ a_2 = \dfrac{c_2^2 - c_1c_3}{c_2} \\ b_1 = \dfrac{-c_3}{c_2}\,, \end{cases}$$

por la relación anterior se reduce a

$$(1+b_1x)\sum_{i=4}^{\infty} c_ix^i = (1+b_1x)g(x)$$

es decir,

$$g(x) = \sum_{i=4}^{\infty} c_ix^i \; .$$

En general se puede ver que $g(x)$ está dado por

$$(11) \qquad g(x) = \sum_{i=m+n+1}^{\infty} c_ix^i$$

y que cada coeficientes a_i y b_i $(b_0 = 1)$ puede determinarse a partir de las relaciones

$$(12) \qquad AP_{mn}(x) = \begin{vmatrix} c_{m-n+1} & c_{m-n+2} & \cdots & c_{m+1} \\ \vdots & \vdots & \cdots & \vdots \\ c_m & c_{m+1} & \cdots & c_{m+n} \\ \sum_{i=0}^{m} c_{i-n}x^i & \sum_{i=0}^{m} c_{i-n+1}x^i & \cdots & \sum_{i=0}^{m} c_i x^i \end{vmatrix}$$

de tal forma que cuando $i - n < 0$, $c_{i-n} = 0$

$$(13) \qquad AQ_{mn}(x) = \begin{vmatrix} c_{m-n+1} & c_{m-n+2} & \cdots & c_{m+1} \\ \vdots & \vdots & & \\ c_m & c_{m+1} & \cdots & c_{m+n} \\ x^n & x^{n-1} & \cdots & 1 \end{vmatrix}$$

donde

$$(14) \qquad A = \begin{vmatrix} c_{m-n+1} & c_{m-n+2} & \cdots & c_m \\ \vdots & \vdots & & \vdots \\ c_m & c_{m+1} & \cdots & c_{m+n-1} \end{vmatrix}$$

cuando $n \neq 1$, $m \neq 1$. Si $m = 1$ y $n \neq 0$, $A = c_1$. Si $n = 1$ y $m \neq 1$ entonces $A = c_m$. Nótese que el determinante A es de orden $n-1$.

Apliquemos las relaciones antes mencionadas a los casos (1) y (2) que tratamos anteriormente.

CASO (1-a). Supongamos que P_{nm} y Q_{nm} están dados por

(15) $$P_{11}(x) = a_0 + a_1 x, \quad Q_{11}(x) = 1 + b_1 x$$

con la serie, $\sum_{i=0}^{\infty} c_i x^i$, que se quiere aproximar; utilizando (12), (10) y (11) tendremos que

(16) $$A = c_1 \quad \text{aquí} \quad m = n = 1$$

$$c_1 P_{11}(x) = \begin{vmatrix} c_1 & c_2 \\ c_{-1} + c_0 x & c_0 + c_1 x \end{vmatrix} =$$

$$= \begin{vmatrix} c_1 & c_2 \\ c_0 x & c_0 + c_1 x \end{vmatrix} = c_1(c_0 + c_1 x) - c_0 c_2 x$$

es decir,

(17) $$c_1 P_{11}(x) = c_1 c_0 + \left[c_1^2 - c_0 c_2\right] x ,$$

por otro lado,

$$c_1Q_{11}(x) = \begin{vmatrix} c_1 & c_2 \\ x & 1 \end{vmatrix} = c_1 - c_2x$$

entonces

(18) $$c_1Q_{11}(x) = c_1 - c_2x$$

Observando el lado derecho de las relaciones (15) y (16), nos damos cuenta que, están bien determinadas a partir de la serie (1). Es el lado izquierdo de esas mismas relaciones lo que queda por determinar.

Utilizando (13) en (15) y (16) tendremos

$$c_1(a_0 + a_1x) = c_1c_0 + \left[c_1^2 - c_0c_2\right]x$$

$$c_1\left[1 + b_1x\right] = c_1 - c_2x$$

para que se satisfagan es necesario que

$$c_1a_0 = c_1c_0 , \qquad c_1a_1 = c_1^2 - c_0c_2$$

y

$$c_1 = c_1 , \qquad c_1b_1 = -c_2$$

de donde a_0, a_1, y b_1 están dados por

$$a_0 - c_0 = 0$$
$$c_1 a_1 - c_1^2 - c_0 c_2 = 0$$
$$c_1 b_1 + c_2 = 0$$

que es el mismo resultado obtenido más lentamente en (7), ya que a través de cada conjunto de relaciones obtenemos

$$(19) \qquad \left\{ \begin{aligned} a_0 &= c_0 \\ a_1 &= \frac{c_1^2 + c_0 c_2}{c_1} \\ b_1 &= -\frac{c_2}{c_1} \end{aligned} \right. \qquad\qquad c_1 \neq 0$$

CASO (2-a). Nosotros inicialmente (como en el caso 2), consideramos que

$$P_{mn}(x) = P_{21}(x) = a_0 + a_1 x + a_2 x^2$$
$$Q_{mn}(x) = Q_{21}(x) = 1 + b_1 x \qquad \therefore \begin{array}{l} m = 2 \\ n = 1 \end{array}$$

Aplicando a (14), dado que $n = 1$ y $m = 2$ entonces

$$(20) \qquad A = c_2$$

$$AP_{21} = c_2P_{21} = \begin{vmatrix} c_2 & c_3 \\ c_{-1}+c_0x+c_1x^2 & c_0+c_1x+c_2x^2 \end{vmatrix} =$$

$$= \begin{vmatrix} c_2 & c_3 \\ c_0x+c_1x^2 & c_0+c_1x+c_2x^2 \end{vmatrix} =$$

$$= c_0c_2 + (c_1c_2 - c_0c_3)x + (c_2^2 - c_1c_3)x^2$$

es decir,

$$(21) \qquad c_2\left[a_0 + a_1x + a_2x^2\right] = c_0c_2 + (c_1c_2 - c_0c_3)x + (c_2^2 - c_1c_3)x^2,$$

por otro lado, aplicando (13), tendremos:

$$AQ_{21} = \begin{vmatrix} c_2 & c_3 \\ x & 1 \end{vmatrix} = c_2 - c_3x$$

por lo tanto

$$(22) \qquad c_2(1+b_1x) = c_2 - c_3x.$$

A partir de (21) y (22) nosotros podemos determinar las a_i y las b_i; con el resultado siguiente:

$$a_0 = c_0$$

$$a_1 = \frac{c_1 c_2 - c_0 c_3}{c_2}$$

$$a_2 = \frac{c_2^2 - c_1 c_3}{c_2}$$

y

$$b_1 = \frac{-c_3}{c_2}$$

comparando los anteriores resultados con los obtenidos en (10) vemos que son los mismos.

Apliquemos lo anteriormente visto a un caso concreto. Supongamos que

$$\sum_{i=0}^{\infty} c_i x^i = \cos x \quad \text{con} \quad c_i = 0 \quad \text{si} \quad i = 1,3,5,\ldots \tag{23}$$

es decir, tratemos de aproximar con el método de Padé la serie, $\cos x$, que vamos a escribir como

$$\cos x = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} (x^2)^i = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} \omega^i \tag{24}$$

con $\omega \equiv x^2$.

Antes de aproximar el valor de $\cos x$, es relevante informar que si consideramos el caso $m = n$, es decir, la diagonal de la tabla de Padé, entonces nuestra aproximación resulta más eficiente, por razones que mencionaremos más adelante.

Cálculo de algunos términos diagonales, $m = n$, de la tabla de Padé, para el caso del $\cos x$,

CASO 1. $m = n = 1$

(25) $$P_{11}(\omega) = a_0 + a_1\omega = a_0 + a_1x^2$$

(26) $$Q_{11}(\omega) = 1 + b_1\omega = 1 + b_1x^2$$

de donde, aplicando (12), (13) y (14) obtenemos

$$A = c_1$$

$$AP_{11}(x) = c_1P_{11}(x) = \begin{vmatrix} c_1 & c_2 \\ c_0x^2 & c_0 + c_1x^2 \end{vmatrix} =$$

$$= c_1\left[a_0 + a_1x^2\right] = c_1\left[c_0 + c_1x^2\right] - c_0c_2x^2$$

es decir,

$$(c_1a_0 - c_1c_0) + \left[a_1c_1 - c_1^2 + c_0c_2\right]x^2 = 0$$

$$c_1a_0 - c_1c_0 = 0$$

$$a_1c_1 - c_1^2 + c_0c_2 = 0$$

entonces

(27) $$a_0 = c_0$$

(28) $$a_1 = \frac{c_1^2 - c_0c_2}{c_1}$$

por otro lado,

$$AQ_{11}(x) = c_1Q_{11}(x) = \begin{vmatrix} c_1 & c_2 \\ x^2 & 1 \end{vmatrix}$$

$$c_1\left[1 + b_1x^2\right] = c_1 - c_2x^2$$

es decir,

(29) $$b_1 = \frac{-c_2}{c_1}$$

para calcular ahora c_0, c_1 y c_2 con el objeto de determinar a_0, a_1 y b_1 basta con escribir los primeros tres términos de la serie

$\cos x$, utilizando a (23) y (24)

(30) $$\sum_{i=0}^{\infty} c_i \omega^i = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} \omega^i \quad \text{con} \quad \omega = x^2$$

$$c_0 + c_1 + c_2\omega^2 + \sum_{i=3}^{\infty} c_i\omega^i = 1 - \frac{\omega}{2!} + \frac{\omega^2}{4!} + \sum_{i=3} \frac{(-1)^i}{(2i)!} \omega^i$$

de donde

(31) $$c_0 = 1$$

(32) $$c_1 = -\frac{1}{2!} = -\frac{1}{2}$$

(33) $$c_2 = \frac{1}{4!} = \frac{1}{24}$$

A partir de las anteriores relaciones y de (27), (28) y (29) obtenemos

(35) $$a_0 = 1$$

(34) $$a_1 = \frac{1/4 - 1/24}{-1/2} = \frac{1}{12} - \frac{1}{2} = \frac{-5}{12}$$

(36) $$b_1 = \frac{2!}{4!} = \frac{1}{12}$$

En estos momentos podemos escribir a

$$P_{11}(x) = a_0 + a_1 x^2 \quad y \quad Q_{11}(x) = 1 + b_1 x^2$$

como

(37) $$P_{11}(x) = 1 - \frac{5}{12} x^2$$

(38) $$Q_{11}(x) = 1 + \frac{1}{12} x^2$$

de donde

(39) $$\frac{P_{11}(x)}{Q_{11}(x)} = \frac{12 - 5x^2}{12 + x^2}$$

A nosotros nos gustaría tener una idea razonable del error que aparece en cada aproximación de Padé, a la función $f(x)$. Por ejemplo, ¿qué tan buena es la aproximación $m = n = 1$, que se tiene en (39), a la función $\cos x$?

Escribamos de nuevo (5)

$$\left[1 + \sum_{i=1}^{n} b_i x^i\right]\left[\sum_{i=0}^{\infty} c_i x^i\right] - \sum_{i=0}^{m} a_i x^i = \left[1 + \sum_{i=0}^{n} b_i x^i\right] g(x)$$

ahí vimos que para determinar los coeficientes de P_{mn} y Q_{mn} es decir, las a_i y las b_i a partir de las c_i se necesitó igualar a cero todos los coeficientes de $x^{i'}$ con $i' = m+n+1$. En el caso de que nos interesen los elementos diagonales $(m=n)$ de la tabla de Padé, entonces $i' = 2m+1$.

Lo anterior se ve más claramente desarrollando la relación (5):

$$\sum_{i=0}^{m+n+1} c_i x^i + \left[\sum_{i=1}^{n} b_i x^i\right]\left[\sum_{i=0}^{\infty} c_i x^i\right] - \sum_{i=0}^{m} a_i x^i + \sum_{i=m+n+1}^{\infty} c_i x^i =$$

$$= x^{m+n+1} \sum_{i=0}^{\infty} \gamma_i x^i$$

es decir,

$$Q_{mn}(x) \sum_{i=0}^{\infty} c_i x^i - P_{mn}(x) = x^{m+n+1} \sum_{i=0}^{\infty} \gamma_i x^i \tag{40}$$

En general sucede que cuando consideramos la diagonal, $(m = n)$, de la tabla de Padé, los coeficientes γ_i, decrecen tan rápidamente, que basta con considerar el término

$$\left| \frac{\gamma_0 x^{2n+1}}{Q_{nn}(x)} \right| > 0 \tag{41}$$

para obtener una buena aproximación, del valor absoluto del error, al que se incurre si uno posee a

$$\frac{P_{nn}(x)}{Q_{nn}(x)} \tag{42}$$

en lugar de

$$f(x) = \sum_{i=0}^{\infty} c_i x^i$$

Lo anterior es particularmente el caso cuando nos restringimos a valores de x tales que $|x| < 1$, ya que entonces $Q_{nn}(x)$ difiere poco de la unidad ya que los coeficientes b_i decresen muy rapidamente. Por lo anterior, cuando, $|x| < 1$, entonces tendremos una buena medida del error incurrido, al utilizar a (42) si calculamos únicamente a γ_0.

Se puede ver que

$$\gamma_0 = \frac{\begin{vmatrix} c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_{n+1} & c_{n+2} & \cdots & c_{2n+1} \end{vmatrix}}{\begin{vmatrix} c_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & & \vdots \\ c_n & c_{n+1} & & c_{2n} \end{vmatrix}} \tag{43}$$

Evaluemos a γ_0 para el caso que acabamos de ver, es decir, cuando $f(x) = \cos x$. Aplicando la realción anterior, donde, $m = n = 1$.

$$\gamma_0 = \frac{\begin{vmatrix} c_1 & c_2 \\ c_2 & c_3 \end{vmatrix}}{c_1} = \frac{c_1 c_3 - c_2^2}{c_1}$$

Dado que

$$c_1 = -\frac{1}{2!} = -\frac{1}{2}$$

$$c_2 = \frac{1}{4!} = \frac{1}{24}$$

$$c_3 = -\frac{1}{6!} = \frac{-1}{720}$$

de donde

$$\gamma_0 = \frac{\left(-\frac{1}{2!}\right)\left(-\frac{1}{6!}\right) - \left(\frac{1}{4!}\right)^2}{-1/2!}$$

$$\gamma_0 = \frac{3}{2}\,\frac{1}{6!} \sim 0.00208 = 2.08 \times 10^{-3}$$

será la medida aproximada del error incurrido.

El lector interesado encontrará ilustrativo calcular el siguiente término diagonal de la tabla de Padé, $m = n = 2$, para el caso

$$f(x) = \cos x$$

en donde se obtiene

$$\frac{P_{22}(x)}{Q_{22}(x)} = \frac{15120 - 6900x^2 + 313x^4}{15120 + 660x^2 - 13x^4}$$

$$|x| < 1$$

y la medida de la aproximación en este caso es igual a

$$|\gamma_0| = \left| -\frac{59}{42}\ \frac{1}{10!} \right| \sim 3.871 \times 10^{-7}$$

lo cual nos da una muy buena aproximación al valor de $\cos x$, como podemos verificar inmediatamente.

Sea $x = \pi/4 < 1$, efectuando los cálculos correspondientes obtenemos

Descripción	Función	Resultado	Porcentaje de error
Valor Exacto	$\cos \pi/4$	0.707106	
$\frac{P_{22}(\pi/4)}{Q_{22}(\pi/4)}$	$\frac{15120-6900\left(\frac{\pi}{4}\right)^2+313\left(\frac{\pi}{4}\right)^4}{15120+660\left(\frac{\pi}{4}\right)^2+13\left(\frac{\pi}{4}\right)^4}$	0.707107	1.41×10^{-4}
Los primeros 5 términos de la serie	$1-\frac{(\pi/4)^2}{2!}+\frac{(\pi/4)^4}{4!}-\frac{(\pi/4)^6}{6!}+\frac{(\pi/4)^8}{8!}$	0.707229	1.74×10^{-2}

TABLA 1

En la anterior tabla se encuentaran los resultados al calcular el valor exacto y aproximado de $\cos x$ hasta 6 cifras decimales. Dado que el cálculo de P_{22} y Q_{22} involucra la utilización de, cinco $(m+n+1 = 2+2+1)$, términos de la serie, el cálculo aproximado de $\cos x$, sólo contiene los primeros cinco términos de la serie. Nótese que, además de la ventaja del método de Padé en relación al cálculo aproximado de $\cos x$, tomando los primeros cinco términos, de la serie no sólo está en ahorrarnos algunos cálculos, en la obtención de la aproximación, sino lo que es aún más importante, en que es un método de aproximación que resulta ser más exacto para el caso que tratamos.

En realidad la utilidad práctica del método de Padé radica en los casos donde es desconocida la función a la que es igual una serie y donde sólo se conoce información parcial de la misma. Para los lectores interesados véase la bibliografía.

BIBLIOGRAFIA

1. T.J. Stieltjes, Research sur les fractions continués, Ann. Fac. Sci. Toulouse 8, J, 1-122 (1894); Ann. Fac. Sci. Toulouse 9, A, 1-47 (19895).

2. H. Padé, Sur la représentation approchée d'une fonction par des fractiones rationelles, Thesis, Ann. Ecole. Nor. 9, 1-93 (1892).

3. Un libro que contiene los anteriores trabajos (1) y (2) y ciertos desarrollos de la teoría de Padé es "Continued Fractions" de H.S. Wall, Van Nostrand, Princeton, New Jersey (1948).

4. Dos libros más recientes que contienen información útil relativa a la aplicación en Física Teórica de los Aproximantes de Padé son: G.A. Baker, "The theory and applications of the Padé approximant method", Advances in Theoretical Physics, Vol. 1, K.A. Brueckner, ed., Academic Press, N.Y. (1965).
G.A. Baker y J.L. Gammel, "The Padé approximant in theoretical physics", Academic Press, New York, (1970).

INVESTIGACION EN EDUCACION MATEMATICA

Fernando Hitt ★

La investigación en educación la podemos dividir, en primera instancia, en dos grandes bloques: Pedagogía* y Didáctica*.

En la antigüedad no había diferencia entre ellos. Por ejemplo, en el siglo XVII, Comenius escribió su obra maestra "La Grande Didactique", y su contenido general trata sobre consejos pedagógicos. Sin embargo, la pedagogía se fue desarrollando haciéndose cada vez más exacta y provocando investigaciones científicas que culminaron en una ciencia bien estructurada.

Sin tratar de ahondar en este tema, que es conocido y del que existe una gran cantidad de bibliografía, sólo expondré los principales objetivos de la investigación pedagógica [11]. Existen cinco objetivos generales, cuya obtención no es problema exclusivo de los pedagogos. La colaboración de los especialistas de otras ciencias humanas y matemáticas es absolutamente necesaria para tal fin. Así, la comunicación entre médicos, biólogos, sicólogos, sociólogos, antropólogos, economistas, historiadores, filósofos, matemáticos, etc., es preponderante. Los objetivos generales son los siguientes:

★ Sección de Matemática Educativa del Centro de Investigación y Estudios Avanzados del IPN.

---

*Al final del escrito, antes de la bibliografía, se encuentran las definiciones de todos los términos que estén señalados con el asterisco.

I. Conocer al alumno, como niño (o adulto) y como aprendiz.

II. Conocer a los educadores, la enseñanza y el aprendizaje.

III. Conocer las materias a enseñar.

IV. Conocer el sistema educativo.

V. Conocer las soluciones aportadas por otros científicos.

Partiendo de una clasificación de las ciencias humanas, se distinguen cuatro grupos en la investigación pedagógica:

1. La investigación nomotética*
2. La investigación histórica
3. La investigación normativa
4. La investigación filosófica.

Cada una de ellas está bien definida y estructurada, a la persona que quiera ahondar en el tema se le sugiere [11] y [17], por ejemplo.

Como se había dicho en un principio, la didáctica y la pedagogía se confundían, y se siguen confundiendo, como una misma ciencia. En realidad, la didáctica trata de llegar al mismo estatus como ciencia que tiene la pedagogía.

Puesto que nuestro interés está centrado en la enseñanza de las matemáticas, nos restringiremos a Didáctica de las Matemáticas.

A continuación mostraré algunos tópicos (sin pretender ser exhaustivo) relacionados a la investigación en didáctica de las matemáticas:

Didáctica de las Matemáticas

- Evaluación* y Docimología*
- Obstáculos didácticos* y epistemológicos.
- Heurística*
- Teorías del Aprendizaje y Taxonomías (clasificaciones).
- Historia de las ideas matemáticas
- Desarrollo Curricular
- Nuevos métodos de investigación

1. EVALUACION Y DOCIMOLOGIA

La evaluación trata sobre pronósticos, juicios o diagnósticos, determinados por un objeto inicial; el objeto en cuestión puede, por ejemplo, ser un sujeto de examen, desencadenando de la parte del examinado una reacción e inmediatamente después, una respuesta.

Se pueden distinguir cuatro tipos de evaluaciones:

Evaluación

- Formativa (I)
- Sumativa (II)
- Comparativa (III)
- Sobre criterios (IV)

i) Para el tipo (I), la evaluación es hecha en el transcurso de un aprendizaje, con el objeto de influir sobre el desarrollo del aprendizaje. Así, la evaluación formativa tiene como único fin el de reconocer dónde y en qué un alumno tuvo una dificultad e informarle sobre ésta. Tal evaluación no se traduce en calificación.

ii) Para el tipo (II), la evaluación se hace al final del aprendizaje, o bien, fuera de él. La evaluación sumativa toma de la formativa el caráter de diagnóstico. Para el educador este carácter de diagnóstico individual reviste una gran importancia, pues la evaluación se traduce en una calificación para el alumno.

iii) Evaluación comparativa. Ésta también tiene carácter de diagnóstico, pero no individual. En este caso, el diagnóstico nos sirve para la comparación entre individuos o grupos de individuos.

iv) Evaluación sobre criterios. Este tipo de evaluación puede ser como I, II o III; lo que la caracteriza es el hecho de que para la realización de un examen sobre criterios se tiene primeramente una fase de elaboración y una de criterios, después de una examinación y, finalmente, una de corrección.

Ejemplo:

Indicar, subrayando la respuesta deseada, si los siguientes números son decimales:

a)	0.747474 ...	sí	no
b)	0.333333 ...	sí	no
c)	7.21	sí	no
d)	1/3	sí	no
e)	5	sí	no
f)	0.4747	sí	no

El criterio que se podría tomar es el siguiente: 6 puntos a la respuesta "no, no, sí, no, sí, sí" ; 3 puntos a la respuesta "no, no, sí, no, no, sí" ; 0 puntos en cualquier otro caso. El criterio siempre es formulado antes de conocer las respuestas de los alumnos.

Sobre Docimología, sólo diré que es la ciencia que estudia los exámenes. Ella se interesa en el comportamiento del examinador y del examinado, así como en los sistemas de calificación.

2. OBSTACULOS DIDACTICOS Y EPISTEMOLOGICOS [1].

El ejemplo está relacionado con los números decimales. Sabemos que los chinos tenían un sistema de medida decimal en el siglo XIII antes de Cristo. Los babilonios tenían la numeración de posición, y los pitagóricos el conocimiento de las fracciones. Así, se tuvo que esperar a Al Kashi, nacido en 1427, e independientemente a S. Stevin en 1585, para que los decimales aparecieran. Claro está que hubo otras personas antes que ellos que utilizaban fracciones, como Bonfils de Tarascón, nacido en 1350; Regiomontanus, en 1563, etc... Pero es Stevin quien recibe el mérito por proponer la substitución de fracciones decimales a las fracciones racionales de manera que se pudieran utilizar los cálculos y reglas ya conocidos de los números naturales. La "vulgarización" de los decimales se tornó entonces en un problema que lo llamaré de tipo didáctico. Fue necesario esperar dos siglos para que se empezará a llevar a cabo tal vulgarización, Gobain hizo una obra dedicada a los comerciantes; y D'Alambert, en 1779, presentó en una enciclopedia un artículo sobre los decimales en forma matemática.

En 1784, el abad Boussut presentó los decimales a la manera "naturalista"; es decir, como enteros con un punto. Así, se anunció una ruptura entre las fracciones decimales y los "decimales populares". Los esfuerzos de vulgarización fueron facilitados por la elección del sistema decimal. Por tanto, la enseñanza de los decimales se topa con problemas técnicos y socio-económicos; así, un niño piensa como si todos los números fueran enteros relativos. Por consiguiente, las nociones topológicas serán perturbadas durante mucho tiempo y, por ejemplo, el niño no encontrará un decimal entre 3.25 y 3.26, a cierta edad, y/o multiplicará de la siguiente forma dos decimales:

$$0.2 \times 0.3 = 0.6$$

Con todo esto, sacamos como conclusión que es así como se formó un obstáculo didáctico, y el estudio de este fenomeno atañe a la didáctica de las matemáticas. Para finalizar, les diré que tal obstáculo no ha sido superado y se sigue estudiando. (Cabe aclarar que la noción de obstáculo, en general, atañe a psicólogos y filósofos).

3. HEURISTICA* .

La Heurística entra en el dominio de la Didáctica, y la pregunta más importante es: ¿Cuándo sabremos si al analizar la resolución de un problema se siguió una conducta programada o un proceso heurístico ?

Con el siguiente ejemplo, trataremos de clarificar lo anterior; para tal efecto, usaremos un ejercicio clasificado entre los automatismos* (éste, fue resuelto por un alumno de primero de preparatoria).

El problema trata sobre la resolución de una desigualdad. Primero se muestra el trabajo realizado por el alumno, e inmediatamente después mi interpretación en el tiempo de realización (aunque el alumno no llegó a la solución $[4, +\infty[$, el alumno demuestra una <u>habilidad matemática</u>: la sensibilidad a la contradicción matemática).

Pregunta 3. Estudio de la desigualdad:

$$(0.2)[0.4x + 15] - 0.8x \leq 0.12$$

a) ¿Para qué valores de x, la desigualdad se satisface?

$$\cancel{0.08x + 0.30 - 0.8x \leq 0.12}$$

$$\cancel{0.00 \leq 0.12} \quad -0.72x \leq 0.12 - 3$$

$$0.72x \geq 2.88$$

$$x \geq \frac{2.88}{0.72} = 0.4$$

la desigualdad se satisface para $x \geq 0.4$ $[0.4; +\infty[$

b) Verificar que la desigualdad se satisface para $x = 10$.

$$(0.2)[0.4 \times 10 + 15] - 0.8 \times 10 \leq 0.12$$

$$0.2\,[4 + 15] - 8 \leq 0.12$$

$$0.8 + 0.30 - 8 \leq 0.12$$

$$1.10 - 8 \leq 0.12$$

$$-6.8 \leq 0.12$$

Pregunta 3. Estudio de la desigualdad:

$$(0.2)[0.4x + 15] - 0.8x \leq 0.12$$

a) ¿Para qué valores de x, la desigualdad se satisface?

$$\cancel{0.8x} + 0.30 - \cancel{0.8x} \leq 0.12$$
$$0.30 \leq 0.12$$
$$\cancel{x} = 0.12$$

b) Verificar que la desigualdad se satisface para $x = 10$.

Pregunta 3. Estudio de la desigualdad:

$$(0.2)[0.4x + 15] - 0.8x \leq 0.12$$

a) ¿Para qué valores de x, la desigualdad se satisface?

$$\cancel{0.8x} + 0.30 - \cancel{0.8x} \leq 0.12$$
$$0.30 \leq 0.12$$
$$\cancel{x} = 0.12$$

b) Verificar que la desigualdad se satisface para $x = 10$.

$$(0.2)[0.4 \times 10 + 15] - 0.8 \times 10 \leq 0.12$$
$$0.2 [4 + 15] - 8 \leq 0.12$$
$$0.8 + 0.30 - 8 \leq 0.12$$
$$1.10 - 8 \leq 0.12$$
$$-6.8 \leq 0.12$$

Pregunta 3. Estudio de la desigualdad:

$$(0.2)[0.4x+15]-0.8x \le 0.12$$

a) ¿Para qué valores de x, la desigualdad se satisface?

$\cancel{0.08x + 0.30 - 0.8x} \le 0.12$

$\cancel{0.60 \le 0.12}$

$\cancel{[illegible]}$

b) Verificar que la desigualdad se satisface para $x=10$.

$(0.2)[0.4\times 10+15] - 0.8\times 10 \le 0.12$

$0.2[4+15]-8 \le 0.12$

$0.8+0.30-8 \le 0.12$

$1.10-8 \le 0.12$

$-6.8 \le 0.12$

Pregunta 3. Estudio de la desigualdad:

$$(0.2)[0.4x+15]-0.8x \le 0.12$$

a) ¿Para qué valores de x, la desigualdad se satisface?

$\cancel{0.08x + 0.30 - 0.8x} \le 0.12$

$\cancel{0.60 \le 0.12}$ $-0.72x \le 0.12-3$

$0.72x \ge 2.88$

$x \ge \frac{2.88}{0.72} = 0.4$

la desigualdad se satisface para $x \ge 0.4$ $\cancel{[illegible]}$

$[0.4;+\infty[$

b) Verificar que la desigualdad se satisface para $x=10$.

$(0.2)[0.4\times 10+15] - 0.8\times 10 \le 0.12$

$0.2[4+15]-8 \le 0.12$

$0.8+0.30-8 \le 0.12$

$1.10-8 \le 0.12$

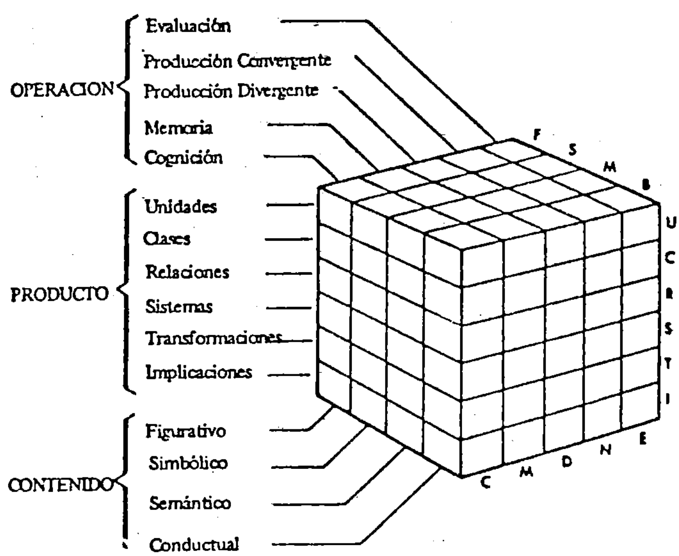
$-6.8 \le 0.12$

4. TEORIAS DEL APRENDIZAJE Y TAXONOMIAS* (clasificación).

Las diferentes corrientes en teorías del aprendizaje producen diferencias en modelos que tratan de explicar el desarrollo de la inteligencia por medio de las habilidades intelectuales, teniendo como producto concepciones de clasificaciones muy variadas.

Un modelo de la estructura intelectual dada por J.P. Guilford, es el siguiente (él descartó la idea de un esquema jerárquico):

Modelo de la estructura intelectual
(Guilford)



Otra clasificación o taxonomía es el modelo propuesto por Robert M. Gagné, que contempla cinco categorías, sin llegar a dar un orden de complejidad y un modelo sobre aprendizaje y memoria:

Información verbal

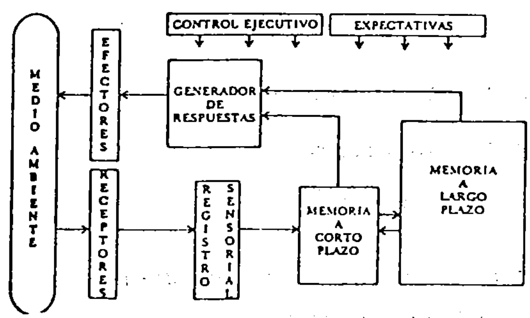
Habilidades intelectuales

Estrategias cognoscitivas

Actitudes

Habilidades motoras

El modelo básico del aprendizaje y la memoria sobre el que se fundamentan las teorías modernas del "procesamiento de la información"



Para Piaget y seguidores, la evolución de la inteligencia consiste en una transformación continua de su estructura por medio de la asimilación, la acomodación y la adaptación.

La práctica de estas teorías en el nivel educativo y sobre todo para la enseñanza de las matemáticas, ha sido y es un problema abierto.

Existen otras teorías que tienen como meta el conocer la estructura de las habilidades matemáticas. Por ejemplo, la escuela soviética se centró en la investigación de problemas detectores de habilidades matemáticas, detectando las siguientes clases: Generalización, Abreviación en el razonamiento, e

Inversión de un proceso, razonamiento u orden de operaciones. Para una lectura más precisa de estas teorías, al final se enlista una serie de lecturas.

El enorme impulso que ha tenido la teoría conductista en Estados Unidos, tiene su aplicación en matemáticas bajo la taxonomía NLSMA (1969); ésta es producto del trabajo de B. Bloom y seguidores. Ella comprende cuatro bloques:

A.0 Computación

B.0 Comprensión

C.0 Aplicación

D.0 Análisis

Cada bloque a su vez es dividido. Si bien, "el mercado educativo" está saturado de ideas relacionadas con esta teoría, en la práctica la enseñanza de tópicos matemáticos se queda en los primeros niveles. En seguida se da un ejemplo sobre la ecuación de 2do. grado, en relación con esta taxonomía.

Clasificación NLSMA de la ecuación de 2do. grado.

NIVEL A: <u>Conocimiento de hechos específicos memorizados. (Computación).</u>

A1: Conocimiento de hechos específicos.

1. Entre los siguientes polinomios, ¿cuáles son de 2do. grado?

	verdadero	falso
$2+2x+2x^4$	☐	☐
$2-2x^2+x$	☐	☐
$2x^5+x^4+2x^3-x^2+2x+2$	☐	☐
x^2-1	☐	☐

<u>Nota</u>. El único hecho específico que se requiere para resolver el ejercicio, es el siguiente: "Se llama grado de un polinomio reducido, al grado del monomio de más alto grado del polinomio".

2. Entre los polinomios siguientes:

a) x^2-1 b) x^3 c) $4-4x^2$ d) $x+2x^2+36$

- Todos son de 2do. grado. ☐
- Ninguno es de 2do. grado. ☐
- Solamente a) y c) son de 2do. grado. ☐
- Otro caso diferente a los anteriores. ☐

A2: Conocimiento de terminología.

Cinco es raíz de la ecuación $2x^2-4x-30=0$.

Es decir:

- Substituyendo $x=5$ en la ecuación, se obtiene $0=0$ ☐
- Substituyendo $x=5$ en la ecuación, se obtiene $x=0$ ☐
- 5 es solución de la ecuación $2x^2-4x-30=5$ ☐
- 5 es solución de la ecuación $2x^2-4x-25=0$ ☐
- Otro caso diferente a los anteriores. ☐

A3: Habilidad para efectuar algoritmos.

1. Las raíces o soluciones de la ecuación $7x^2+x-30=0$ son:

☐ $x=2$ y $x=15/7$

☐ $x=-2$ y $x=30/14$

☐ $x=7$ y $x=30$

☐ No hay solución.

☐ Ninguno de los anteriores.

2. Las soluciones de la ecuación $x^2-x-12=0$ son $x=4$ y $x=-3$, entonces la ecuación la podemos escribir de la forma:

☐ $(x+4)(x-3)=0$

☐ $(x-4)(x+3)=0$

☐ $(x-4)(x-3)=0$

☐ $(x+4)(x+3)=0$

☐ Otro caso diferente a los anteriores.

3. Si tenemos la ecuación $(x+1)(2x-7/3)=0$, el conjunto solución es:

- $\{+1; -\frac{7}{6}\}$ ☐
- $\{+1; +\frac{7}{6}\}$ ☐
- $\{-1; -\frac{7}{6}\}$ ☐

- No hay solución. ☐
- Ninguno de los anteriores. ☐

NIVEL B: Conocimiento y utilización de conceptos memorizados. (Comprensión).

B1: Conocimiento de conceptos.

1. La ecuación $ax^2 + bx + c = 0$, $a \neq 0$, $a, b, c, \in \mathbb{R}$ puede tener
   - Tres soluciones reales ☐
   - Dos soluciones reales o complejas ☐
   - Una solución real y otra compleja ☐
   - Una solución doble real ☐

2. Entre las siguientes expresiones, indicar si es o no un polinomio de segundo grado:

	verdero	falso
$2 + 3x + 2x^4 - x$	☐	☐
$x + 2x^2 + x^3$	☐	☐
$2x^3 + x - x^3 + x^2 - x^3 + 1$	☐	☐
$5x + 3x^2 - 1$	☐	☐
$x^2 - 1/x + 1$	☐	☐
$1/x + x - 3$	☐	☐
$4x - 4x^2 + 4 - 4x$	☐	☐
$x^2 + \sqrt{x} + x + 6$	☐	☐

3. Se tienen los números $-3/2$ y 5
   ¿Es posible encontrar una ecuación de segundo grado tal que $-3/2$ y 5 sean sus soluciones? Sí ☐ No ☐
   (Si la respuesta es sí, escribir la ecuación) ____________

B2: Conocimiento de principios, reglas y generalizaciones.

1. Sea el polinomio $p(x) = dx^2 + x + 2$, con d un número real:
   - el polinomio es de 2do. grado ☐
   - el polinomio no es de 2do. grado ☐
   - el grado del polinomio depende del valor de d ☐
   - Otro caso diferente de los anteriores. ☐

B3: Conocimiento de la estructura matemática.

1. Se sabe que $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ nos da las dos soluciones de la ecuación $ax^2 + bx + c = 0$, $a \neq 0$ $a, b, c \in \mathbb{R}$. Si $b^2 - 4ac = 0$, entonces:
   - ☐ No hay solución.
   - ☐ Se obtiene una solución doble.
   - ☐ Se obtienen dos soluciones diferentes.
   - ☐ Otro caso diferente a los anteriores.

2. La misma pregunta que 1., pero ahora con $b^2 - 4ac > 0$.

3. La misma pregunta que 1., pero ahora con $b^2 - 4ac < 0$.

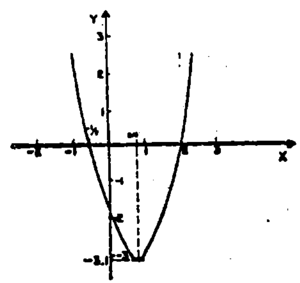
S4. <u>Habilidad para transformar elementos de un problema de una forma a otra.</u>

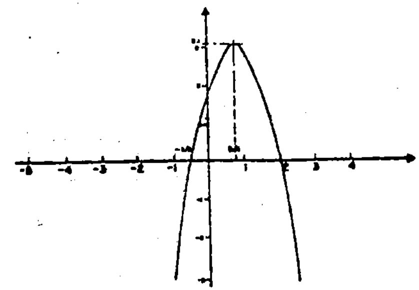
1. Escribir las ecuaciones necesaria para resolver el problema siguiente, y decir el significado de cada letra utilizada (es necesario llegar a una ecuación de segundo grado, pero no se pide resolverla).

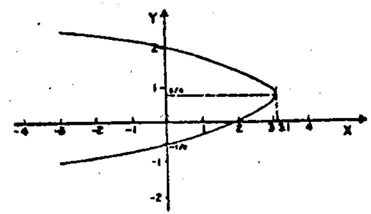
   Un electricista paga \$144 por un tramo de cable. A la siguiente vez, él trata de pagar la misma suma de dinero por la misma cantidad de cable que la primera. Pero, el dependiente le dice que en vista de un aumento de precios de \$1 por metro de cable, él tendrá dos metros de cable menos. ¿Cuántos metros de cable recibió la primera vez?

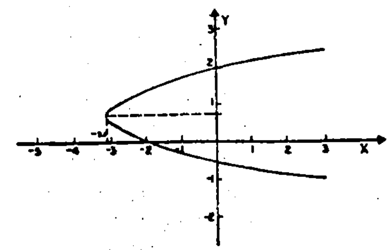
2. Si en el problema anterior se cambian metros por yardas, ¿el número obtenido en la solución anterior, cambia?

3. Se tiene el polinomio $p(x) = 2x^2 - 3x - 2$, él se anula para $x = -1/2$ y $x = 2$. ¿Cuál es entonces su gráfica?









Otro caso diferente a los anteriores.

B5: <u>Habilidad para seguir un razonamiento.</u>

1. Se tiene la ecuación de segundo grado $ax^2+bx=0$, $a\neq 0$ $a,b \in \mathbb{R}$. Entonces:

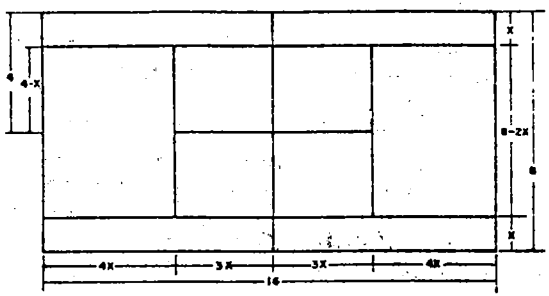
   Se puede dividir entre x y obtener $ax+b=0$, de donde $x=-\frac{b}{a}$ es la única solución.

   No es posible dividir entre x y por tanto no se sabe nada acerca de las soluciones.

   Se puede escribir $x(ax+b)=0$, y por tanto $x=0$ y $x=-\frac{b}{a}$ son soluciones.

B6: <u>Habilidad para leer e interpretar un enunciado.</u>

Se tiene un terreno rectangular de 8 por 16 metros, y se quiere hacer una cancha de tenis como lo indica la siguiente figura:



NOTA: Este problema se puede resolver de manera inmediata al resolver $4x+3x+3x+4x=16$.

Matematización del problema

¿Cuál es el valor que puede tomar x?	$0\leq x\leq 4$
Se tienen 4 rectángulos de misma área, ¿cuál es el área de cada uno de ellos en función de x?	$(3x)(4-x)$
Se tienen 2 rectángulos de misma área, ¿cuál es el área de cada uno de ellos en función de x?	$(4x)(8-2x)$
Se tienen otros 4 rectángulos de misma área, ¿cuál es el área de cada uno de ellos en función de x?	$(8)(x)$
¿Cuál es el área total en función de x?	$4(3x)(4-x)+2(4x)(8-2x)+4(8)(x)$
Esta última es por hipótesis igual a 8×16, es decir	$12x(4-x)+8x(8-2x)+32x=128$
Reduciendo	$7x^2-36x+32=0$

resolviendo la ecuación $7x^2-36x+32=0$, se obtienen las soluciones, $x=4$ y $x=8/7$. Pero, para nuestro problema la única solución es $x=\frac{8}{7}$.

NIVEL C: <u>Las aplicaciones.</u>

C1: <u>Habilidad para resolver problemas rutinarios.</u>

1. Calcular las soluciones de la ecuación:

$$(3x+2)(4x-1)+(2x-1)(3x+2)=0$$

2. ¿Cuál es la gráfica de la función $p(x)=6x^2+4x$?

C2: Habilidad para hacer comparaciones.

1. Se tienen dos rectángulos: uno tiene 4x de base, $2(x+1)$ de altura (x es la incógnita), y $48\,m^2$ de área; el otro tiene 2x de base, y $(x+1)$ de altura. Entonces:

   El área del rectángulo de base 4x es dos veces más grande que el área del otro.

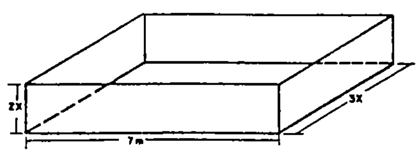
   El área del rectángulo de base 4x es cuatro veces más grande que el área del otro.

   No se puede dar una respuesta, pues hacen falta más datos.

   Otro caso diferente a los anteriores.

C3: Habilidad para interpretar y analizar datos.

1. Un salón de clases tiene la forma siguiente:



la superficie de los cuatro muros es de $104\,m^2$. ¿Cuál es el valor de x? ¿Cuál es la superficie del techo?

Nota. Los hechos específicos que se requieren para solucionar este problema son:

El área de un rectangulo de largo l y altura h es $A = l \times h$.

Si las longitudes se expresan en metros, el área se expresará en metros cuadrados.

Relación entre la superficie de los cuatro muros y el área de los cuatro rectángulos:

$$(7)(2x) + (7)(2x) + (2x)(3x) + (2x)(3x) = 104$$

Ecuación reducida y ordenada $3x^2 + 7x - 26 = 0$.

Las soluciones de una ecuación de 2do. grado son todos los valores de x que verifican: $ax^2 + bx + c = 0$ a,b,c números reales y $a \neq 0$.

Las soluciones en $\mathbb{R}$ de la ecuación de 2do. grado son:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \text{con} \quad b^2 - 4ac \geq 0$$

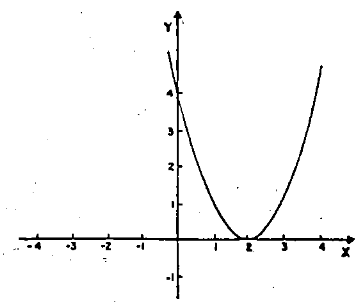
entonces $x = 2$ y $x = -26/6$ son las ecuaciones de $3x^2+7x-26=0$, pero, $x = -26/6$ aunque es solución de la ecuación, ella no es solución del problema.

$x = 2$ es solución del problema.

La superficie del techo es de $42\,m^2$.

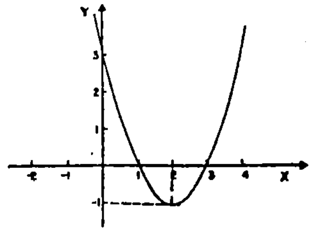
C4: Habilidad para reconocer patrones, isomorfismos y simetrías.

1. Se tiene la función $p(x) = (x-2)^2$, su gráfica es de la forma siguiente:

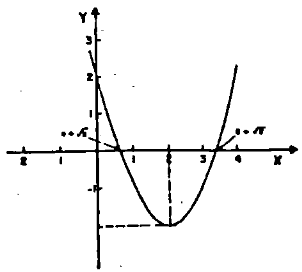


Nosotros podemos definir las funciones:

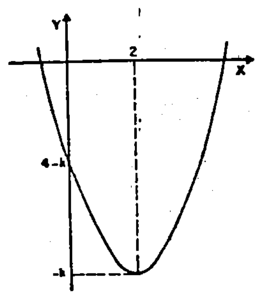
$p_1(x) = p(x) - 1$, es decir, $p_1(x) = x^2 - 4x + 3$ que tiene como gráfica



$p_2(x) = p(x) - 2$, es decir, $p_2(x) = x^2 - 4x + 2$ que tiene como gráfica



¿La gráfica de la función $p_k(x) = p(x) - k$, $k \in \mathbb{N}$ es la que se muestra en la figura de abajo?



Es necesario validar tu respuesta:

NIVEL D: <u>El descubrimiento. (Análisis)</u>.

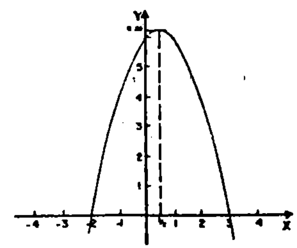
D1: <u>Habilidad para resolver problemas no rutinarios.</u>

1. ¿Se puede encontrar un entero relativo x tal que

$$x^2 + 8x + 15 < 0 \,?$$

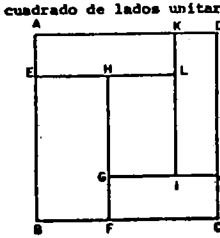
D2: <u>Habilidad para descubrir relaciones.</u>

1. Encontrar una ecuación de 2do. grado tal que -2 y 3 sean sus soluciones, y un polinomio de 2do. grado tal que su gráfica sea de la forma:



D3: <u>Habilidad para hacer demostraciones.</u>

1. A B C D: cuadrado de lados unitarios:

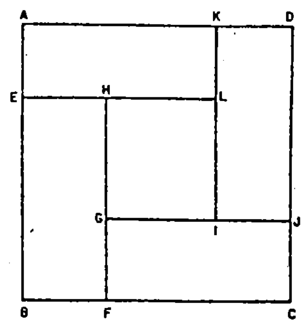


¿Es posible dividir el cuadrado en 5 rectángulos de diferentes dimensiones, pero de igual área?

Si la respuesta es sí, ¿cuáles son las dimensiones de los 5 rectángulos?

D4: Habilidad para criticar la validez de un razonamiento.

1. A B C D: cuadrado de lados unitarios.



¿Es posible dividir el cuadrado en 5 rectángulos de diferentes dimensiones, pero de igual área?

Se pone $EB = x$ $(x \neq \frac{1}{2}$ y $\frac{1}{5} < x < \frac{4}{5})$.

Se busca la expresión de las dimensiones en función de x.

Se encuentra cada vez:

$BF = \frac{1}{5x}$ $\qquad$ $FC = 1 - BF = \frac{5x-1}{5x}$

$FG = \frac{1}{5FC} = \frac{x}{5x-1}$ $\qquad$ $GH = x - FG = \frac{x(5x-2)}{5x-1}$

$GI = \frac{1}{5GH} = \frac{5x-1}{5x(5x-2)}$ $\qquad$ $IJ = FC - GI = \frac{(5x-1)(5x-3)}{5x(5x-2)}$

$JD = \frac{1}{5JI} = \frac{x(5x-2)}{(5x-3)(5x-1)}$ $\qquad$ $JD = 1 - FG = \frac{4x-1}{5x-1}$

Las dos expresiones encontradas de JD deben ser compatibles, de donde:

$$5x^2 - 5x + 1 = 0$$

Conclusión: Si hubiera una solución, deberían tenerse más de dos, lo que es imposible. ¿Todo el razonamiento anterior es exacto y completo?

D5: Habilidad para formular y validar generalizaciones.

1. Se tiene un polinomio de 2do. grado $p(x) = ax^2 + bx + c$ $a \neq 0$, $a, b, c \in \mathbb{R}$.

Si $a < 0$, ¿se puede encontar un valor de x para el cual p(x) sea máximo?

Si $a > 0$, ¿se puede encontar un valor de x para el cual p(x) sea mínimo?

es necesario validar la respuesta.

5. HISTORIA DE LAS IDEAS MATEMATICAS .

Los obstáculos didácticos y epistemológicos no son fácilmente detectables; una herramienta poderosa para tal efecto es la historia de las ideas matemáticas. Bajo este punto de vista, queda desechada la idea ingenua de que el seguir el proceso de transformación de una idea matemática, en el tiempo, sólo proporciona cultura matemática.

Lo que se muestra a continuación es un ejemplo del punto de vista arriba citado.

Génesis del concepto de convergencia y convergencia uniforme.

Los conceptos de función, continuidad, continuidad uniforme, convergencia y convergencia uniforme, se desarrollaron progresivamente, según las necesidades prácticas que existían en la resolución y esclarecimiento de problemas relacionados con la física y la matemática.

DESARROLLO HISTORICO.

Entre los sabios del siglo XIV, se encuentra Thomas Bradwardine que, en su tratado (Tractus de Proportionibus, 1328), aborda el concepto de función de potencia. También, Nicole Oresme (1323-1382) en su Algorismus Proportionum, estudia las reglas para manipular las funciones potencia, extendiéndose a potencias fraccionarias. Oresme, utilizando las expresiones de "latitud" y "longitud", en la representación de las trayectorias de los astros, fue conducido al señalamiento gráfico. Se considera comunmente que Oresme con las palabras "latitudo formarum", introdujo el germen de la idea de función.

No es sino hasta el siglo XVII, fundamentalmente, que se inicia un estudio de las curvas bajo el concepto de función.

Descartes muestra el camino con sus aplicaciones de métodos algebraicos en geometría. Newton, en su "Methodus fluxionum" (escrito en 1671 y publicado en 1736), usa una expresión que se aproxima al sentido de la palabra de función, "relata quantitas". Leibniz emplea la expresión "functionem faciens" o "functio", para designar medidas cuyas variaciones están bajo una ley.

Jean Bernoulli (1667-1748), en su memoria de 1718, da una definición de función, utilizando la letra griega φ para designar la función, escribe: "se llama función de una medida variable, una cantidad que se compone en cierta forma de esta medida y de constantes".

Es en los trabajos de Leonhard Euler (1707-1783), donde encontramos una idea precisa de función continua y discontinua. Euler, en su primer volumen de su "Introductio in analysis infinitorum" (1748), definió una función como cualquier expresión analítica: "función es cualquier expresión analítica de cantidades variables y de números, o con cantidades constantes.

En el mismo escrito, Euler impulsa principalmente la teoría algebraica de funciones, además, denota las constantes como a, b, c ; y las variables como x, y, z , e introduce su clasificación de las funciones como algebraicas o trascendentes, explícitas o implícitas, uniformes o multiformes (es decir, simples o con varios valores). Pero, el análisis de las cuerdas vibrantes dio lugar a la primera extensión de la noción de función y condujo a Euler, en su segundo volumen, a hacer una distinción entre una curva "continua" o "discontinua". Así, una curva que podía representarse por una ecuación algebraica o trascendente (por una misma ley) se le llamaba curva continua; y curva discontinua (mixta o irregular) a aquella que exigía diferentes ecuaciones para la representación de sus diversas partes constitutivas. Por tanto, en el sentido Euleriano, las funciones determinadas por una función analítica para todo el dominio de la variable independiente, se le llamaba continua. Entonces, por un lado se tenían las funciones "verdaderas" y por el otro las funciones discontinuas o arbitrarias (representación de líneas trazadas por un movimiento libre de la mano) que no eran funciones "verdaderas". Es decir, Euler consideraba como continuas las curvas que obedecían a una ley constante; y como discontinuas, las curvas cuyas diversas funciones tienen como expresión las diferentes funciones de x, a estas últimas, él las consideraba como diferenciables por pedazos.

Una definición precisa no se tenía hasta ese momento (1741), pero al tratar D'Alambert, entre otros, de restringir la selección de las funciones, provocó objeciones de otros matemáticos de la época, entre ellas la de Euler. Este último, no satisfecho con su definición de función, formuló otra en su Institutiones calculi differentialis (1755), esta nueva concepción consiste en dar la definición de función como una relación arbitraria entre las cantidades variables de cualquier dependencia, es decir: "las cantidades dependiendo de otras, de manera que si las otras cambian, sus cantidades cambian también. Es frecuente nombrarlas funciones, metonimia, que es evidente y que contiene en ella misma las formas con las que una cantidad pueda ser determinada por las otras. Entonces, si yo pudiera nombrar la variable de todas las cantidades que dependen en cierta forma de x o que pueden ser determinadas, la llamaría función de esta variable".

El nuevo concepto de función está desligado en ese momento de expresiones analíticas. La definición de Euler fue utilizada por otros matemáticos, por ejemplo, Condorcet, Lagrange y otros. En 1796, en su "Théorie des fonctions analytiques...", Lagrange da la definición siguiente: "Se llama función de una o de varias cantidades, toda expresión de cálculo en la cual sus cantidades entran de una manera cualquiera, mezcladas o no con otras cantidades que se miran como poseyendo valores dados e invariantes en tanto que las cantidades de la función pueden recibir todos los valores posibles".

A partir de una nota de Euler en relación con las series trigonométricas, que dice:

Si $f(x) = a_1 \operatorname{sen} x + a_2 \operatorname{sen} 2x + \cdots + \frac{1}{2} b_0 + b_1 \cos x + b_2 \cos 2x + \cdots$

los coeficientes son determinados por las fórmulas:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(x) \operatorname{sen} nx\, dx \quad ; \quad b_n = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(x) \cos nx\, dx$$

Joseph Fourier (1768-1830) estaba convencido de que esta determinación sigue siendo aplicable cuando la función $f(x)$ es dada arbitrariamente (Memoire sur la Propagation de la Chaleur, 1807)*. La impresión fue profunda, en algunos matemáticos de la época, cuando Fourier afirmó que toda función arbitraria, sea simple o compuesta de parte de funciones, puede ser repre-

*Por esta memoria Fourier recibió el gran premio de matemáticas en 1812 y fue hasta 1822 cuando se publicó su escrito dentro de su libro "Théorie de la Chaleur".

sentada por una serie trigonométrica*.

En esa misma memoria, Fourier afirma que: "la serie de funciones continuas

$$\cos x - \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x - \cdots$$

- Converge en todo punto.
- Su función límite se compone de líneas rectas separadas, cada una de las cuales es paralela al eje de las x e igual a la circunferencia. Estas paralelas están alternativamente situadas por encima y por debajo del eje, con una distancia de $\pi/4$ entre dos de ellas, estando <u>unidas por perpendiculares</u> que forman a su vez parte de la línea".

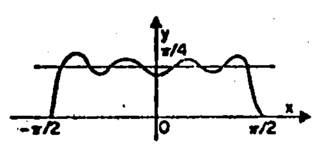
Fourier consideraba que la función límite era continua, es decir**:

---

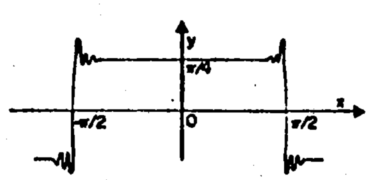
*Dirichlet posteriormente probó que la función $X(x)$ igual a 1 para x racional y X igual a 0 para x irracional, admite la representación siguiente:

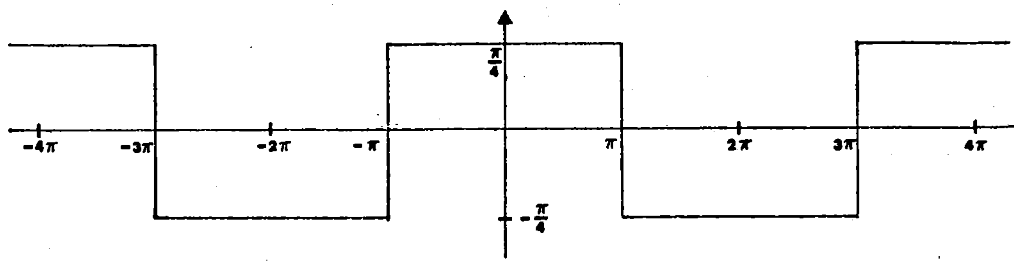
$$X(x) = \lim_{m}[\lim_{n}(\cos^2 m!\,\pi x)^{2n}];$$

**Para $y = \cos x - \frac{1}{3}\cos 3x + \cdots - \frac{1}{4m-1}\cos(4m-1)x$ se tiene



Cuando m crece, se tiene:





Seguramente Fourier no quería ir en contra del "axioma", según el cual "lo que es verdadero hasta el límite es verdadero en el límite" (principio de continuidad de Leibniz, 1687). Además, los científicos del siglo XVIII estaban influenciados por el viejo adagio latín "natura non facit saltus"; de hecho, algunos matemáticos consideraban que la noción de discontinuidad estaba excluida de la ciencia. Ellos pensaban que tal noción encierra una idea enfermiza (curvas discontinuas o irregulares según Euler, ilegítimas según Fourier, etc.)*.

Cauchy y el Problema de la Convergencia Uniforme (dos enfoques):

La idea de función y de continuidad estaban latentes a principios del siglo XIX. Una definición precisa del concepto de continuidad se encuentra en los trabajos de B. Bolzano, de 1817:

> "... Según una definición correcta de lo que se entiende por la expresión que una función f(x) para todos los valores de x, que se encuentran, al interior, o fuera de ciertos límites, cambia, según la

*Existe un paso obscuro en la interpretación del trabajo de Fourier, ya que el término "discontinuo" (aproximadamente, en el sentido de Cauchy) aparece en algunos manuscritos no publicados de Poinsot (1807) y de Fourier (1809) (ver Lakatos [10] pág. 151, o [6] pág. 153.

> ley de continuidad, si cualquier x toma un tal valor, la diferencia $f(x+\omega) - f(x)$ puede llegar a ser más pequeña que toda cantidad dada, cuando se pueda aceptar ω, tan pequeña como se quiera".

Esta definición de Bolzano indica un concepto local de continuidad. En el mismo tratado, él critica, como incorrecta, la definición de la continuidad, según la cual cuando $f(x) = a$, $f(y) = b$, $a \neq b$, $f(z)$ toma todos los valores entre a y b si $x \leq z \leq y$*.

Augustin-Luis Cauchy (1789-1857) en su Cours d'Analyse à l'Ecole Polytechnique, en 1821, da la definición siguiente para la continuidad: "... Sea $f(x)$, una función de la variable x, y supongamos que esta función admite constantemente un valor único y finito. Si, partiendo de un valor de x comprendido entre esos dos límites, se le atribuye a la variable x un incremento infinitamente pequeño α, la función misma recibirá por incremento la diferencia

$$f(x+\alpha) - f(x)$$

que dependerá, al mismo tiempo, de la nueva variable α y del

---

*Esta definición persiste algunos años después (ver G. Glaesser [4]), por ejemplo, en el Cours d'Analyse de l'Ecole Politechnique (1841) de Jean-Marie Duhamel:

> "Una función se dice c o n t i n u a cuando se hace variar de una manera continua las cantidades donde ella depende, ella (la función) es constantemente real y no puede pasar de un valor a otro sin pasar por todos los intermediarios".

La función $f(x) = \begin{cases} \operatorname{sen} 1/x & \text{para } x \neq 0 \\ 1/2 & \text{para } x = 0 \end{cases}$ cumple con tal definición y no es continua en cero.

En este mismo escrito, se encuentra la definición de límite, a saber:

> "Se dice límite de una cantidad variable, una cantidad fija, la cual se aproxima a ella indefinidamente".

La sucesión $U_n = \frac{1}{n}$ no tiende a -10, sin embargo ella se le aproxima indefinidamente.

valor x. Así, la función $f(x)$ será, entre los dos límites asignados a la variable x, una función continua de esta variable, si para cada valor de x intermediario entre esos límites, el valor numérico de la diferencia

$$f(x+\alpha) - f(x)$$

decrece infinitamente con α. En otros términos, la función $f(x)$ permanecerá continua con respecto a x entre los límites dados, si, entre esos límites un incremento infinitamente pequeño de la variable siempre produce un incremento infinitamente pequeño de la función misma. Se dice nuevamente que la función $f(x)$ es en la vecindad de un valor particular atribuido a la variable x, función continua de esta variable, siempre que sea continua entre dos límites de x, aun muy próximos, que encierran el valor en cuestión".

En la definición de Cauchy no se sabe concretamente si la continuidad es local, global o uniforme. Aún en su Mémoire sur les fonctions continues (1844) su definición de continuidad aparece en los mismos términos que utilizó en su memoria de 1821. En su memoria de 1844 encontramos también notas interesantes sobre la historia del concepto de función continua, él escribe: "En las obras de Euler y de Lagrange, una función es llamada c o n t i n u a o d i s c o n t i n u a, si los diversos valores de esta función, correspondientes a diversos valores de la variable, son o no sujetos a una misma ley, son o no proporcionados por una misma ecuación. En esos términos, la continuidad de las funciones se encontraba definida por estos ilustres geómetras, cuando ellos decían que «las funciones arbitrarias, introducidas por la integración de ecuaciones que involucran derivadas parciales, pueden ser funciones continuas o discontinuas». Con todo, la definición que hemos presentado está lejos de ofrecer una precisión matemática; porque, si los diversos valores de una función, correspondientes a los diversos valores de una variable, dependiendo de dos o más ecuaciones diferentes, no impedirá el disminuir el número de esas ecuaciones y aun de reemplazarlas por una única ecuación, cuya descomposición proporcionará todas las otras. Aún más: las leyes analíticas a las que las funciones pueden ser ajustadas se encuentran generalmente expresadas por fórmulas algebraicas o trascendentes, y puede suceder que diversas fórmulas representen, para ciertos valores de una variable x, la misma función; además, para otros valores de x, funciones diferentes. Por consiguiente, si se considera la definición de Euler y la de Lagrange como aplicables a todas las especies de funciones, sean algebraicas o trascendentes, un simple cambio de notación bastará frecuentemente para transformar una función continua en una discontinua, y recíprocamente. Así, por ejemplo, si x designa una variable real, una función que se reducirá, tanto a $+x$ como a $-x$, según que la variable x sea positiva o negativa, deberá, por ese motivo, ser colocada en la clase de las funciones discontinuas, y sin embargo, la misma función podrá ser mirada como continua, cuando se le representa por la integral indefinida

$$\frac{2}{\pi}\int_0^\infty \frac{x^2\,dt}{t^2+x^2}$$

o aun por el radical

$$\sqrt{x^2}$$

que es el valor particular de la función continua

$$\sqrt{x^2 + t^2}$$

correspondiente a un valor nulo de t, el carácter de continuidad en las funciones, mirado desde el punto de vista que en un principio analizaron los geómetras, es un carácter vago e indeterminado".

Teorema de Cauchy (Cours d'Analyse, 1821).

"En tanto que los diferentes términos de la serie

(1) $u_0,\ u_1,\ u_2,\ \ldots,\ u_n,\ u_{n+1},\ \ldots$

son funciones de una misma variable x, continua con respecto a esta variable en la vecindad de un valor particular en la cual la serie es convergente, la suma s de la serie es también, en la vecindad de este valor particular, función continua de x".

La aserción es que si todas las funciones $u_n(x)$ son continuas en un intervalo $a < x < b$ (el cual incluye un x_0 particular) y si la serie (1) converge en el intervalo a la suma $s(x)$ entonces $s(x)$ también es continua. Cauchy, en su demostración, introduce las sumas parciales

$$s_n(x) = u_0(x) + u_1(x) + \cdots + u_n(x)$$

y los residuos

$$r_n(x) = s(x) - s_n(x)$$

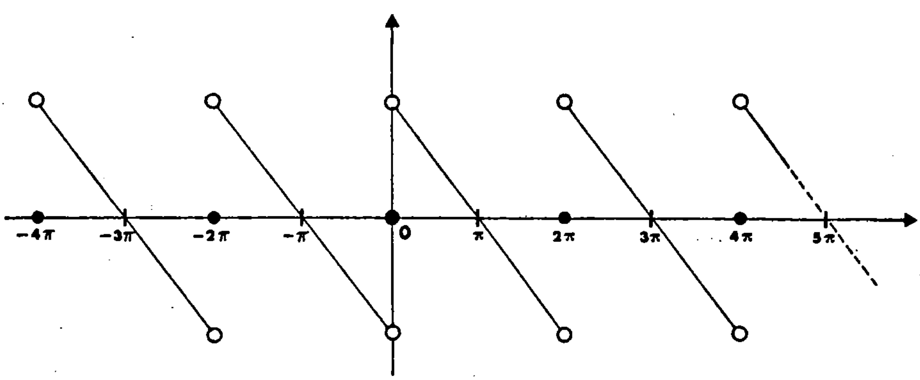
y argumenta como sigue:

"Así puesto, consideremos los incrementos que reciben esas tres funciones (es decir, r_n, s_n y s), en tanto se hace crecer x en una cantidad infinitamente pequeña. El incremento de s_n será, para todos los valores posibles de r_n una cantidad infinitamente pequeña; y aquella de r_n permanecerá insensible al mismo tiempo que r_n, si se le atribuye a n un valor muy considerable. Por consiguiente, el incremento de la función x no podrá ser más que una cantidad infinitamente pequeña".

Abel (1802-1829) en su artículo sobre las series binomiales (1826) plantea lo siguiente: "Me parece que hay algunas excepciones al teorema de Cauchy"; y pone como ejemplo

$$\operatorname{sen}\phi - \frac{1}{2}\operatorname{sen}\phi + \frac{1}{3}\operatorname{sen}3\phi - \cdots$$

cuya gráfica sería:



Abel en ese momento comenzó a tratar de responderse la siguiente pregunta: "¿Cuál es el dominio seguro del teorema de Cauchy?"

He aquí su respuesta: "El dominio de validez de los teoremas del análisis en general y de los teoremas acerca de la continuidad de la función límite en particular está restringido a series de potencias. Todas las excepciones conocidas de este principio de restringir el análisis al interior de las seguras fronteras de las series de potencias, dejando así fuera las queridas series trigonométricas de Fourier como si fuesen una jungla incontrolable, en las que las excepciones son la norma, y los éxitos un milagro".

El método de exclusión de excepciones seguido por Abel es más patente en una carta a Hansteen (29 de marzo de 1826), en donde caracterizaba la «miserable inducción euleriana» como un método que conduce a generalizaciones falsas y sin fundamento, preguntándose la razón por la cual tales procedimientos han llevado de hecho a tan p o c a s calamidades. Su respuesta es: "A mi modo de ver, la razón estriba en que en análisis se ocupa uno en gran medida de funciones que se puedan representar mediante series de potencias. Tan pronto como aparecen otras funciones (cosa que sólo ocurre rara vez), entonces [la inducción] no funciona ya y de esas conclusiones falsas surge un infinito número de teoremas incorrectos, llevando uno a los demás. He investigado varios de ellos y he tenido la suerte de resolver el problema..." *.

En realidad, Fourier en su "Mèmoire sur la propagation de la Chaleur" (1807), proporcionó un ejemplo que con el tiempo contradeciría el teorema de Cauchy de 1821. Fourier escribía: "la serie

$$\cos x - \frac{1}{3}\cos 3x + \frac{1}{5}\cos 5x - \cdots$$

converge a una función continua". Él tenía dudas al respecto y constató que la convergencia era lenta: "La convergencia no

*De hecho, Abel pensaba que la restricción del dominio (a series de potencia) de las series era absolutamente necesario, desechando así las series divergentes. Abel escribía al respecto: "Las series divergentes son obra del demonio. Lo único que producen son calamidades y paradojas".

es lo suficientemente rápida como para producir una aproximación fácil, aunque es suficiente para la verdad de la ecuación".

Hasta 1847, no se había modificado ni el teorema de Cauchy ni su demostración. Seidel descubrió la hipótesis oculta en la prueba de Cauchy (esto desde el punto de vista del continuo de Weierstrass):

Veamos la prueba de Cauchy y la dada por Seidel (la primera también desde el punto de vista del continuo de Weierstrass):

PRUEBA DE CAUCHY.

Dado $\varepsilon > 0$:

(1) Hay un δ tal que, para cualquier α, si $|\alpha| < \delta$, entonces $|s_n(x+\alpha) - s_n(x)| < \varepsilon$ (existe tal δ, debido a la continuidad de $s_n(x)$).

(2) Hay un N tal que $|r_n(x)| < \varepsilon$, para todo $n \geq N$ (existe tal N, debido a la convergencia de $\Sigma u_n(x)$).

(3) Hay un N_1 tal que $|r_n(x+\alpha)| < \varepsilon$, para todo $n \geq N_1$ (existe tal N_1, debido a la convergencia de $\Sigma u_n(x+\alpha)$)

(4) Conclusión:

$$\begin{aligned} |u(x+\alpha) - u(x)| &= |u_n(x+\alpha) + r_n(x+\alpha) - s_n(x) - r_n(x)| \\ &\leq |s_n(x+\alpha) - s_n(x)| + |r_n(x)| + |r_n(x+\alpha)| \\ &< 3\varepsilon, \text{ para todo } |\alpha| < \delta . \end{aligned}$$

PRUEBA DE SEIDEL.

(1) $|s_n(x+\alpha) - s_n(x)| < \delta$, si $|\alpha| < \delta(\delta, x, n)$.

(2) $|r_n(x)| < \delta$, si $n > N(\delta, x)$.

(3) $|r_n(x+\alpha)| < \delta$, si $n > N(\delta, x+\alpha)$.

(4) Conclusión:

$$
\begin{aligned}
|u(x+\alpha) - u(x)| &= |u_n(x+\alpha) + r_n(x+\alpha) - s_n(x) - r_n(x)| \\
&\leq |s_n(x+\alpha) - s_n(x)| + |r_n(x)| + |r_n(x+\alpha)| \\
&< 3\delta, \quad \text{si} \quad n > \max_z N(\delta, z) \\
&\quad \text{y} \quad |\alpha| < \delta(\delta, x, n).
\end{aligned}
$$

El lema oculto en la demostración de Cauchy es que $\max_z N(\delta, z)$ debe existir para cualquier δ fijo. Este requisito llegó a denominarse **convergencia uniforme.**

De esa época hasta mediados del siglo XX, se consideró errónea la demostración de Cauchy.

La interpretación anterior sufrió un cambio fundamental desde el punto de vista del Análisis no-Estándar. Robinson, en el capítulo X de su Nonstandard Analysis (1966), revisó la historia del cálculo desde el punto de vista no-estándar*, y en particular el teorema sobre la convergencia de una serie de funciones continuas.

---

*Sistema de números Hiperreales (ver [6] pág. 25, [16], o [3] pág. 28).

Definición. Una estructura S es un sistema de números Hiperreales si tiene las tres propiedades siguientes:

1. S contiene al sistema de números reales. Por esto no sólo entendemos que todos los números reales están en S, sino también que cada función y relación definida sobre los reales está también definida sobre los números de S.
2. S contiene un infinitesimal. Es decir, existe un número 0 en S tal que 0 > 0 y además 0 < r para cada número real positivo.
3. Las mismas afirmaciones de L son verdaderas en ambos S y R. Si B es cualquier afirmación de L, entonces B es verdad en S si y sólo si B es verdad en R.

El teorema de Cauchy resulta cierto en Análisis no-Estándar, sin embargo, la demostración dada por Cauchy desde este punto de vista es errónea*(ver Robinson [16], pág. 272).

La primera impresión de Lakatos con respecto al mismo teorema fue la descrita en las páginas anteriores. Él tenía dudas acerca de esa interpretación, y no quiso publicarla en su tesis doctoral (Essays in the Logic of Mathematical Discovery, 1961)** antes de hacer un análisis desde el punto de vista del Análisis no-Estándar (ver [9], pág. 158).

Es sorprendente la manera en que Lakatos [9] analiza el teorema y la prueba de Cauchy. En su artículo "The Significance of Non-standard Analysis for the History and Philosophy of Mathematics", presentado en el Coloquio Internacional de Lógica (Hannover, 1966):

> "De Hecho, el teorema de Cauchy es cierto y y su demostración tan correcta como lo puede ser una demostración informal. Siguiendo a Robinson podemos probar que el argumento de Cauchy, si no es interpretado como un argumento pre-Weierstrassiano, sino como uno Leibniz-Cauchysiano genuino, es el siguiente:
>
> Sea $\lim s_n(x) = s(x)$, donde los $s_n(x)$ son continuos. Entonces, en principio, para probar que $s(x)$ es continua para algún x_1, tenemos que probar que

*Lakatos la considera correcta desde un punto de vista mucho menos formal que Robinson.

**Los escritos de Lakatos fueron publicados por sus colegas John Worrall y Elie Zahar, dando lugar al libro Proofs and Refutations (Pruebas y Refutaciones, ed., Alianza Editorial, 1978), de la ed., Cambridge University Press, 1976.

$s(x_1+\alpha) - s(x_1)$ es un infinitésimo para todo infinitesimal α. (Aquí se utiliza el concepto de continuidad, que sería equivalente al concepto de Weierstrass sólo si cualquier proposición que es verdad para toda cantidad infinitesimal es también verdad para cantidades finitas suficientemente pequeñas, y viceversa).

Ahora

$$|s(x_1+\alpha) - s(x_1)| = |s_n(x_1+\alpha) - s_n(x_1) + r_n(x_1+\alpha) - r_n(x_1)|$$

$$\leq |s_n(x_1+\alpha) - s_n(x_1)| + |r_n(x_1+\alpha)| + |r_n(x_1)|,$$

donde los r_n son los residuos. Cauchy pensó que el lado izquierdo de la expresión anterior era infinitesimal para todo α infinitesimal dado que $|s_n(x_1+\alpha) - s_n(x_1)|$ es infinitesimal para todo n según la definición de Cauchy de **continuidad**; y $|r_n(x_1+\alpha)|$ y $|r_n(x_1)|$ son también infinitesimales para todo n infinitamente grande de acuerdo a la definición de Cauchy de **límite***: $a_n \longrightarrow 0$ si a_n es infinitesimal para n infinitamente grande.

Por supuesto, este argumento implica que $s_n(x)$ debería ser definido, continuo, y que converja no solamente en puntos estándar en el continuo de Weierstrass, sino en cada punto del continuo "denso" de Cauchy... Esta interpretación difunde completamente una nueva luz sobre el famoso "error" de Cauchy: **Cauchy no tuvo ningún error en absoluto, él sólo probó un teorema completamente diferente,**

*Cours d'Analyse de Cauchy (1821): "En tanto que los valores numéricos sucesivos de una misma variable decrecen indefinidamente, de manera que queda por debajo de todo número dado, esta variable llega a ser lo que se dice un **infinitamente pequeño** o una cantidad **infinitamente pequeña**. Una variable de esta especie tiene a cero por límite".

acerca de sucesiones de funciones transfinitas que Cauchy-convergen sobre el continuo de Leibniz".

Hasta este punto, se podría pensar que Cauchy es un precursor de Robinson (o del Análisis no-Estándar[3]). Al analizar el teorema de Cauchy desde el punto de vista del Análisis no-Estándar vemos que es verdadero, pero la demostración de Cauchy a este teorema es errónea desde el punto de vista estructural de la teoría de Robinson.

De acuerdo a Robinson, Cauchy aproximadamente en 1953 cambió de idea, ya que en su "Note sur les séries convergentes dont les divers temes son des fonctions continues d'une variable réelle ou imaginaire, entre des limites données", dice lo siguiente:

"Al establecer, en mi Analyse Algébrique, las reglas generales relativas a la convergencia de las series, yo he, además, enunciado el teorema siguiente:

En tanto que los diferentes términos de la serie

(1) $u_0, u_1, u_2, \ldots, u_n, u_{n+1}, \ldots$

son funciones de una misma variable x, continua con respecto a esta variable, en la vecindad de un valor particular para el cual la serie es convergente, la suma s de la serie es también, en la vecindad de este valor particular, función continua de x.

Como lo han remarcado M.M. Bouquet y Briot, este teorema se verifica por las series ordenadas siguiendo las potencias ascendentes de una variable. Pero, para otras series, no serían admitidas sin restricción. Así, por ejemplo, es verdad que la serie

$$(2)\quad \operatorname{sen} x\,,\ \frac{\operatorname{sen} 2x}{2}\,,\ \frac{\operatorname{sen} 3x}{3}\,,\ \ldots\,,$$

siempre convergente para valores reales de x, tiene por suma una función de x que permanece continua, en tanto que x, supuesta real, varía, en la vecindad de un valor diferente de un múltiplo $\pm 2n\pi$ de la circunferencia 2π, y que se reduce, en particular, a $\frac{\pi - x}{2}$, entre los límites $x = 0$, $x = 2\pi$. Pero, en esos límites, la suma s de la serie (2) llega a ser discontinua, y esta suma, considerada como función de la variable real x, adquiere, en el lugar del valor

$$+\frac{\pi}{2} \quad \text{o} \quad -\frac{\pi}{2}$$

dado por la fórmula

$$s = \frac{\pi - x}{2}\,,$$

el valor singular $s = 0$, que reaparece nuevamente cuando se supone

$$x = \pm\, 2n\pi$$

n siendo un número entero cualquiera.

Por lo demás, es fácil de ver cómo se debe modificar el enunciado del teorema, para que no haya lugar a ninguna excepción. Eso es lo que voy a explicar en pocas palabras.

De la definición propuesta en mi Analyse Algébrique, y generalmente adoptada en el presente, una función u de la variable real x será **c o n t i n u a** entre los dos límites dados de x, si, esta función admitiendo para cada valor intermediario de x un valor único y finito, un incremento infinitamente pequeño atribuido a la variable produce siempre, entre los límites que trata, un incremento infinitamente pequeño de la función misma. Así, concibiendo que la serie (1) permanece convergente, y que sus diversos términos sean funciones continuas de una variable real x, para todos los valores de x entre ciertos límites. Sean entonces

s la suma de la serie;

s_n la suma de sus n primeros términos;

$r_n = s - s_n = u_n + u_{n+1} + \cdots$ el resto de la serie prolongada indefinidamente a partir del término general u_n.

Si se nombra a n' un número entero superior a n, el resto r_n no será otra cosa que el límite hacia el que convergerá, para valores crecientes de n', la diferencia

$$s_{n'} - s_n = u_n + u_{n+1} + \cdots + u_{n'-1} \tag{3}$$

Concibiendo, ahora, que al atribuir a n un valor suficientemente grande puede hacerse, para todos los valores de x comprendidos entre los límites dados, el módulo de la expresión (3) [cualquiera que sea n'], y, por consiguiente, el módulo de r_n, inferiores a un número δ tan pequeño como se quiera. Como un incremento atribuido a x podrá nuevamente ser supuesto tan próximo a cero para que el incremento correspondiente de s_n tenga un módulo inferior a un número tan pequeño como se desee, es claro que bastará atribuir al número n un valor infinitamente grande, y al incremento de x un valor infinitamente pequeño para demostrar, entre los límites dados, la continuidad de la función

$$s = s_n + r_n$$

Pero esta demostración supone evidentemente que la expresión (3) cumple la condición arriba enunciada, es decir, que esta expresión llega a ser infinitamente pequeña para un valor infinitamente grande atribuido a un número entero n. Por otra parte, si esta condición se cumple, la serie (1) será evidentemente convergente. En consecuencia, se puede enunciar el teorema siguiente:

Teorema I. Si los diferentes términos de la serie

$$(1) \qquad u_0,\ u_1,\ u_2,\ \ldots,\ u_n,\ u_{n+1},\ \ldots$$

son funciones de la variable real x, continuas, con respecto a esta variable, entre los límites dados; si, por otra parte, la suma

$$(3) \qquad u_n + u_{n+1} + \cdots + u_{n'-1}$$

llega a ser infinitamente pequeña para valores infinitamente grandes de números enteros n y $n' > n$, la serie (1) será convergente, y la suma s de la serie (1) será, entre los límites dados, una función continua de la variable x.

Si a la serie (1) se substituye la serie (2), la expresión

(3), reducida a la suma

$$(4) \qquad \frac{\operatorname{sen}(n+1)x}{n+1} + \frac{\operatorname{sen}(n+2)x}{n+2} + \cdots + \frac{\operatorname{sen} n' x}{n'}$$

se anulará para $x = 0$; pero, para los valores de x muy cercanos a cero, por ejemplo $x = 1/n$, n siendo un número muy grande, ella [la suma] podrá diferir notablemente de cero; y si, al atribuir a n un valor muy grande, uno pone no solamente $x = 1/n$, sino además $n' = \infty$, la suma (4), o, lo que es lo mismo, el residuo r_n de la serie (2) se reducirá sensiblemente a la integral

$$\int_1^\infty \frac{\operatorname{sen} x}{x} dx = \frac{\pi}{2} - 1 + \frac{1}{1\cdot 2\cdot 3}\,\frac{1}{3} - \frac{1}{1\cdot 2\cdot 3\cdot 4\cdot 5}\,\frac{1}{5} + \cdots = 0.6244\cdots$$

Añadimos que, para un valor positivo de x, pero muy cercano a cero, la suma s de la serie (2) se reducirá sensiblemente a la integral

$$\int_0^\infty \frac{\operatorname{sen} x}{x}\, dx = \frac{\pi}{2} = 1.570796\cdots$$

Sea ahora

$$z = x + yi$$

una variable imaginaria. Esta variable podrá..."

En este pasaje, Cauchy menciona a MM. Bouquet y Briot, en relación a las excepciones al teorema*. Esto quiere decir que Cauchy en lugar de pensar como Abel (restringir el dominio

*No menciona, en relación al teorema, a Fourier ni a Abel.

del teorema a series de potencia solamente) añade la condición de convergencia uniforme.

Desde el punto de vista del Análisis no-Estándar, no había necesidad de añadir ninguna condición, ni de analizar la serie $\sum_{r=1}^{\infty} \frac{\operatorname{sen} r x}{r}$ para asegurar o no su convergencia. ¿Creyó Cauchy que había cometido un error en su teorema de 1821?, o, ¿cambió sus ideas sobre los infinitésimos para acercarse a la teoría del continuo de Weierstrass?

Lakatos pone especial cuidado a estas preguntas, él cree que, si primeramente se juzgó a Cauchy desde el punto de vista del continuo de Weierstrass, caeríamos en un error al tratar de juzgar los teoremas y demostraciones de Cauchy desde el punto de vista formal del Análisis no-Estándar. Además, sería un error el esperar que, ahora, la historia del cálculo se dirija hacia Robinson y no hacia Weierstrass como comunmente los historiadores lo han supuesto.

El posible cambio de Cauchy acerca de los infinitésimos es más palpable en un trabajo de 1844, él escribe: "Para descartar completamente la idea de que las fórmulas empleadas en el cálculo diferencial son fórmulas aproximativas, y no fórmulas rigurosamente exactas, me parece importante el considerar las diferenciales como cantidades finitas, distinguiéndolas cuidadosamente de los incrementos infinitamente pequeños de las variables. La consideración de esos últimos incrementos puede y debe ser empleada como medio de descubrimiento o de demostración en la investigación de las fórmulas o en el estable cimiento de teoremas. Pero entonces el calculador se sirve de los infinitamente pequeños como intermediarios que deben conducirle al conocimiento de las relaciones que subsisten entre las cantidades finitas; y jamás, a mi parecer, de las cantidades infinitamente pequeñas no deben ser admitidas en las ecuaciones finales, donde su presencia llegaría a ser sin objeto ni utilidad..."

Es importante este pasaje pues Cauchy por un lado desecha las cantidades infinitamente pequeñas, y por otro afirma que las diferenciales son cantidades finitas y en 1829 decía que d_x = cte., es arbitrario y podría aun ser infinitesimal. La inclinación de la balanza es más hacia Weierstrass que hacia Robinson. Sin embargo, al analizar la génesis de los conceptos de continuidad, continuidad uniforme, convergencia y convergencia uniforme desde el punto de vista de Robinson, nos muestra diferencias que antes no veíamos con las solas ideas de Weierstrass. Se ha afinado el instrumento para analizar la historia del cálculo, en particular los conceptos arriba mencionados, y posiblemente no es el definitivo.

6. DESARROLLO CURRICULAR.

Ésta forma parte de la investigación pedagógica y, principalmente, describe los rasgos esenciales de un proyecto educativo. El desarrollo curricular en matemáticas no deja de lado los puntos mencionados anteriormente, y en base a ellos comprende los siguientes temas:

a) Selección de objetivos.

b) Análisis matemático del contenido.

c) Análisis genético del contenido.

d) Análisis epistemológico del contenido.

e) Descripción detallada del curriculum.

f) Experimentación del proyecto.

g) Formación y actualización de los profesores de matemáticas.

h) Evaluación y perfeccionamiento del proyecto.

7. NUEVOS METODOS DE INVESTIGACION Y ENSEÑANZA.

Las nuevas teorías de la comunicación como la tecnología existente, han producido cambios en la investigación y la enseñanza de las matemáticas. Así, han surgido los métodos audiovisuales y la computación aplicados a la educación matemática.

El análisis anterior ha sido con el fin de dar un panorama de la investigación en educación matemática; algunas de las apreciaciones no son completas y por ignorancia u olvido pude haber omitido otras que sean fundamentales. Con todo, espero que el escrito muestre que la educación matemática no está dividida en dos campos, la educación y las matemáticas, sino que hay una relación muy estrecha que las une coherentemente.

DEFINICIONES

Didáctica—Estudio de fenómenos ligados a la enseñanza e, indirectamente, al aprendizaje.

Docimología—Ciencia de los exámenes. Ella se interesa en el comportamiento del examinador y del examinado, e igualmente en los sistemas de calificación.

Ejercicio de tipo automatismo—Es aquel cuyo enunciado indica un mecanismo o procedimiento a seguir por el alumno para la resolución del ejercicio.

Evaluación—Juicio, pronóstico o diagnóstico determinado por un objeto inicial.

Heurística—Ciencia que trata sobre la investigación y resolución de problemas.

Investigación nomotética—Es aquella que tiende a estructurar y a generalizar las leyes que gobiernan una ciencia.

Obstáculo didáctico—Mecanismo que conduce a tratar una nueva noción matemática con la estructura anteriormente adquirida, pudiendo estar el individuo en situación contradictoria.

Pedagodía—Conjunto de proyectos, intervenciones, situaciones y técnicas para transmitir el conocimiento.

Taxonomía—Clasificación.

BIBLIOGRAFIA

[1] Brousseau, Guy—Les obstacles épistémologiques et les problems en mathématiques. CIEAEM, Bélgica, 1976.

[2] Cauchy, Augustin—Oeuvres Complètes.

[3] Cleave, John P.—Cauchy, Convergence and Continuity. British Journal for Philosophy of Science, 22 (1971).

[4] Glaeser, Georges—La transmission des connaissances mathématiques hier, aujord'hui, demain. L'Enseignement Mathématique (Revue Internationale), II serie, tomo XVIII, fascículos 3-4, julio-diciembre, 1972.

[5] Glaeser, Georges—The Curriculum Development. Curso de 3er. Ciclo de Didáctica de las Matemáticas. Universidad Luis Pasteur, Estrasburgo, Francia, 1976.

[6] Henle, James M. - Kleinberg, Eugene M.—Infinitesimal Calculus, MIT Press, 1979.

[7] Hitt, Fernando—Análisis en componentes principales y Análisis de las correpondencias. Revista Matemáticas y Enseñanza, No. 14.

[8] Hitt, Fernando—Rapport sur la classification NLSMA de l'equation du 2ème. degré. Rapports et diplômes élabores pour le D.E.A. de Didactique des mathématiques, Universidad Louis Pasteur de Estrasburgo, Francia, 1978-1979.

[9] Lakatos, Imre—Cauchy and the Continum (The Significance of Non-standard Analysis for the History and Philosophy of Mathematics; Mathematical Intelligencer.

[10] Lakatos, Imre—Pruebas y Refutaciones (la lógica del descubrimiento matemático). Alianza Editorial, 1978.

[11] Landsheere, Gilbert—Evaluation continue et examens precis de docimologie. Ed. Labor, Nathon, 1976.

[12] Landsheere, Gilbert—Introduction à la recherche en educación. Armand Colin-Bourrelier, París, 1976.

[13] Manheim, Jerom H.—The genesis of point set topology. Pergamon Press Inc., 1964.

[14] Phili, Christine—"Le Developpement du Concept de Function". Seminario de Filosofía y de Matemáticas de la Escuela Normal Superior, París.

[15] Pluvinage, François—Evaluation et docimologie. L'ouvert No. 12 (Revista de Alsacia y del IREM de Estrasburgo), junio, 1977.

[16] Robinson—Non-Standard Analysis. Exclusivamente capítulo X (Concerning the History of the Calculus). Amsterdam, North Holland, 1974 .

[17] Van Dalen, Mayer—Manual de técnica de la investigación educacional. Paidós, Buenos Aires, 1978.

LAS MATEMATICAS APLICADAS COMO UNA MANERA DE VER LA NATURALEZA

Antonmaría Minzoni

Cuando se habla de matemáticas aplicadas se entienden cosas muy variadas. Con frecuencia se piensa que las matemáticas aplicadas son la utilización de herramientas (que han descubierto, mas o menos por azar, los matemáticos "puros" de otros tiempos) para resolver ecuaciones de "relevancia" como en ingeniería, medicina, y economía. En otras ocasiones se piensa que el trabajo del matemático aplicado empieza con las computadoras y que su función es la de manejar y utilizar estas herramientas.* También se piensa que las llamadas matemáticas aplicadas son de alguna manera útiles en el mismo sentido que la medicina. Este tipo de ideas creo se ha tratado muy ampliamente y no pienso referirme a ellas en esta plática más que de manera tangencial. De lo que yo quiero platicar con ustedes es de un aspecto que yo considero la escencia misma de las matemáticas aplicadas que es su estilo peculiar de plantear problemas sobre la naturaleza que nos rodea, de aquí el título de la plática.

Para ilustrar el pensamiento y resumir lo que yo considero sus partes fundamentales haremos un poco de historia, que desde luego, por la limitación del tiempo y de mis conocimientos no puede ser muy larga.

Podemos considerar, en mi opinión, el origen de la matemática aplicada moderna con Galileo quien se interesó toda su vida por dar una formulación cuantitativa de ciertos fenómenos naturales y encontrar leyes generales que los controlan en muchos casos. Fue, en mi opinión, la línea de pensamiento de Newton la primera que reúne y de hecho hizo conocer las ideas y finalidades de lo que considero matemática aplicada. En primer lugar para describir el movimiento de los planetas, los aproximó por puntos; este es el primer proceso que se hace en matemáticas aplicadas, el de aproximar de una manera muy drástica los fenómenos con el fin de entender ciertos aspec-

* Originalmente, que debe de resolver ecuaciones que otro le da sin averiguar más.

tos y la aproximación debe ser de tal naturaleza que no se pierdan las propiedades básicas del objeto que se desea estudiar. La segunda parte del proceso fue la de (inventar) encontrar la formulación adecuada, para esto inventó el cálculo y las ecuaciones diferenciales, que en la actualidad son herramientas básicas de las ciencias naturales cuando éstas se hacen cuantitativas. Una vez obtenida una formulación satisfactoria ésta puede resultar intratable, de hecho es el problema de muchos cuerpos en el caso del sistema planetario. Aquí nuevamente Newton recurre a aproximaciones que reducen el problema a uno tratable, el problema de la tierra girando alrededor del sol. Finalmente, a pesar de las fuertes aproximaciones, se explican en base al modelo de las leyes de Kepler. Esto da enorme confianza en las aproximaciones y en la capacidad de la herramienta inventada y el pensamiento involucrado para resolver problemas del mismo tipo.

En etapas posteriores el modelo se refina si desean incluir más efectos, más partículas y de manera natural surge la idea de utilizar las mismas ideas para estudiar otros fenómenos naturales. Poco después de Newton, Hamilton quiere estudiar la óptica aproximando un rayo de luz por partículas y explicar de manera coherente todos los fenómenos que no involucran difracción. Al mismo tiempo Lagrange llevó las herramientas del cálculo dadas por Newton a una mayor perfección que permite formular con confianza problemas más complejos y tener una base más firme para las aproximaciones.* Es este proceso el que da origen al cálculo y a los sistemas de ecuaciones ordinarias así como todos los problemas asociados a estas disciplinas y desde luego no voy a discutir los éxitos de este pensamiento que son bien conocidas por todos ustedes.

La segunda etapa en la evolución de las matemáticas aplicadas es, a mi manera de ver, cuando se da el paso del estudio de partículas al estudio de sistemas macroscópicos como el agua, los cuerpos elásticos, los gases, etc. Aquí nuevamente se hace una aproximación muy drástica que es el de tratar estos materiales como un continuo sin tomar en cuenta los detalles del movimiento fino de sus componentes. Esto da origen al estudio de las ecuaciones parciales y a la formulación de las primeras de ellas. Nuevamente se tiene un proceso parecido al de Newton sólo que con más agentes involucrados. En mi opinión, el trabajo fundamental que relaciona el pensa-

* De hecho, formular problemas muy complicados.

miento, aproximando la creación de matemáticos y modelos para entender fenómenos complejos, es la obra de Lord Rayleigh que aún hoy, después de más de 100 años, tiene enorme valor. En este caso también se extienden los modelos, se hacen mejoras sustanciales en las técnicas del cálculo y se aplican estas ideas a una enorme variedad de fenómenos principalmente en agua y medios elásticos; excelentes ejemplos de esto son las obras monumentales de Lamb, Love, Stoker, Ainj. A principios de este siglo se adquiere una gran confianza y conocimiento de las ecuaciones básicas y fundamentalmente de la aproximación lineal y con muy contados esfuerzos en la dirección de comprender fenómenos no lineales. En esta etapa el pensamiento analógico es también sumamente fructuoso. Maxwell utilizó una analogía con materiales elásticos para formular cuantitativamente las leyes de Faraday, prediciendo la propagación de ondas electromagnéticas. Así también, utilizó todos los antecedentes en líquidos y medios elásticos para predecir muchos fenómenos electromagnéticos y explicar otros conocidos. Por otra parte la experiencia con membranas vibrantes motiva por analogía de Broglie y a Schrödinger a formular las ecuaciones que modelan algunos fenómenos atómicos y microscópicos; pero a esto volveremos más adelante. Vemos que en estas dos etapas los esquemas de pensamiento son parecidos y si bien no ha evolucionado en las ideas centrales de aproximación y modelos simplificados resueltos con precisión, si se ha hecho más sofisticado y por sus éxitos se tiene una mayor confianza en él. Dejamos así la época moderna y el resto de la plática lo dedicaré a describir la época contemporánea y, los últimos minutos, a especular sobre la evolución futura.

Podemos considerar que la época contemporánea trae dos ingredientes nuevos al pensamiento en matemáticas aplicadas. Se inicia esto con el trabajo de Poincaré. Fue Poincaré quien introdujo dos nuevas actitudes en forma explícita: Es él quien introduce la idea del estudio cualitativo sistemático de las ecuaciones de movimiento, desarrollando para ello técnicas de geometría y se puede decir, inventando la topología. La segunda idea es la de proceder de manera sistemática en las aproximaciones y estimar los errores involucrados en ellos. Es el primero que comprende, de una manera sistemática (con el ánimo de clasificar y separar efectos), el estudio de problemas no lineales. Por otra parte, es Sommerfeld quien introduce y pone en práctica la idea de encontrar aproximaciones, de manera sistemática, a soluciones de problemas de ecuaciones en derivadas parciales, motivado en gran parte por la necesidad de entender las ondas de radio. Esto aumenta la comprensión de muchos fenómenos y explica de manera coherente los fenómenos de difracción que escaparon al análisis durante muchos años. Si fue con Poincaré que se empezó a estudiar las ecuaciones ordinarias no lineales y los efectos nuevos que estos son capaces de describir, fue con Prandtl, a principios de este siglo, cuando se emprendió el mismo camino para las ecuaciones parciales con el fin de entender flujos alrededor de olas. Estas ideas y las de Poincaré unifican las ideas de aproximación de soluciones y simplificación de modelos y estas ideas apoyadas en la experiencia previa son el centro del pensamiento en matemáticas aplicadas de hoy en día que se conocen con el nombre de Teoría de Perturbaciones.

Estas ideas fueron importadas a América e impulsadas por R. Courant desde los años 30 hasta la fecha. Originalmente estas ideas fueron aplicadas a los problemas de donde nacieron fundamentalmente la óptica y la mecánica de fluidos. Sin embargo, su éxito las hizo extenderse al estudio de problemas de reacciones químicas, propagación de epidemias, vida de poblaciones, estudio de sistemas celulares simples, modelos económicos, estudio de sistemas complejos como la física de los plasmas, etc. Desde luego no quiero entrar en detalles sobre estas aplicaciones específicas y si alguien está interesado tendré mucho gusto en pasarle referencias.

Antes de concluir quiero discutir el papel de la computadora en el desarrollo del pensamiento de las matemáticas aplicadas. Originalmente se pensó en ellas como sólo una ayuda para completar cálculos que no se podrían hacer más que con mucho trabajo rutinario sujeto a errores. Sin embargo, en los últimos 15-20 años, la filosofía ha cambiado y se concibe la computadora como una herramienta para experimentar sobre el comportamiento cualitativo de ecuaciones que se han planteado para mejorar así la instrucción sobre su comportamiento y adecuación al fenómeno que se estudie. Sólo en fecha muy reciente se está acoplando este enfoque con la teoría de perturbaciones y soluciones exactas y aún no es claro cuales serán los productos de esta nueva idea.

No quisiera concluir esta plática sin antes mencionar algunas de las limitaciones de este pensamiento. En primer lugar no pude reproducir detalles finos e incluir todos los efectos, en este sentido es parecido al del fisiólogo que da explicaciones macroscópicas sin describir la bioquímica detallada. No es clara su aplicabilidad a las ciencias sociales y economía, de hecho no se ha producido aún mucho en esos campos con estas técnicas. Apenas ahora se está experimentando su uso en formulaciones cuantitativas de problemas biológicos y si bien unos resultados son alentadores la mayoría de los problemas parece, en mi opinión, aún escapar al análisis y es por esto que creo que este es uno de los campos donde uno puede esperar cosas nuevas.

Para concluir quiero decir que en mi opinión, las matemáticas aplicadas son una manifestación de la idea de ver a la naturaleza en forma aproximada, con modelos simples cuyo comportamiento se puede entender construyendo las ideas y herramientas apropiadas. Lejos de estar muerto son estas las ideas que impulsan y permiten el avance de este pensamiento, las otras concepciones son en mi opinión la descripción de meras manifestaciones externas de estas ideas.

Anexo a la conferencia.

Quiero añadir a la conferencia algunos de los comentarios del auditorio al terminar ésta, así como respuestas y algunas reflexiones que fueron motivadas por este intercambio de ideas después de la plática.

La primera observación por parte del público fue que omití algunas contribuciones importantes. Concretamente se me preguntó por qué no había yo mencionado el trabajo de Maupertius y en general el planteamiento de los problemas de mecánica a través del cálculo de variaciones.

La razón por lo cual no mencioné el trabajo relacionado con el cálculo de variaciones es la siguiente:

Las leyes fundamentales de la mecánica y su utilización para resolver problemas originales son independientes del cálculo de variaciones. Este último provee un esquema que unifica métodos de solución e ideas pero es un desarrollo posterior y alternativo al que se planteó en la conferencia. De hecho, el formalismo canónico en mecánica no ha ayudado, hasta una fecha muy reciente, 1970's, a resolver problemas que no sean solubles por otro método. Es por esto que considero que el cálculo de variaciones no ha jugado un papel esencial en el desarrollo del pensamiento en matemáticas aplicadas.

Al plantear al auditorio esta idea sobre el cálculo de variaciones, otra persona apuntó que el cálculo de variaciones había sido fundamental en la comprensión de los fenómenos cuánticos.

Desde luego contesté que esta idea no es la correcta, desde el punto de vista histórico, ya que las leyes fundamentales de los fenómenos cuánticos fueron descubiertos independientemente del cálculo de variaciones. Fueron más bien las ideas de propagación de ondas y óptica las que llevaron a su comprensión. Desde luego el cálculo de variaciones y el formalismo canónico de la mecánica clásica jugaron un papel importante para esquematizar la teoría. Este es, en mi opinión, nuevamente un ejemplo donde el cálculo de variaciones interviene para aclarar el pensamiento establecido previamente de otras maneras.

Al notar tanto interés en la audiencia sobre el papel del cálculo de variaciones en la evolución del pensamiento en matemáticas aplicadas, creo que debo añadir algunos comentarios al respecto para que esta presentación no quede truncada.

Es por esto que aprovecho este escrito para exponer algunas opiniones, muy personales, sobre el uso y la evolución del cálculo de variaciones en conexión con las matemáticas aplicadas.

En primer lugar en mecánica clásica el cálculo de variaciones permite plantear las ecuaciones de sistemas complicados de una manera sistemática, evitando el problema de balancear fuerzas substituyendolo por el de balancear energías y encontrar coordenadas apropiadas para la formulación del problema. Es también una ayuda valiosa, a veces, para encontrar constantes de movimiento y reducir así el número de ecuaciones por resolverse.

Una situación muy análoga ocurre en los problemas de ecuaciones en derivadas parciales nuevamente es un gran auxiliar para plantear ecuaciones y encontrar cantidades conservadas. En este contexto provee la base y las ideas para los métodos de elementos finitos, o de Galerkin para encontrar soluciones aproximadas a problemas de equilibrio. Este tipo de pensamiento ha tenido gran impacto en la ingeniería moderna.

Hay dos usos novedosos del cálculo de variaciones y el formalismo canónico asociado con él en matemáticas aplicadas que pueden desmentir la opinión sobre el cálculo de variaciones que he sostenido en esta plática; los expondré brevemente.

En primer lugar se han descubierto técnicas de aproximación asintótica para problemas no lineales que tienen como base ideas de cálculo de variaciones y su uso es natural en ese contexto. Estas ideas han dado luz sobre la naturaleza de estos problemas y es necesario más trabajo en esta área para obtener una perspectiva histórica de estas contribuciones.

Por otra parte hay ciertas ecuaciones parciales no lineales que aparecen como límites naturales de sistemas mas complejos que son sistemas Hamiltonianos de dimensión infinita. Aunque no se sabe decidir cuando estos sistemas son completamente integrables es uno de los éxitos más grandes de las matemáticas aplicadas en los últimos 10 años el haber descubierto como integrar explícitamente varios casos particulares. En estos casos el formalismo de la mecánica clásica extendido a sistemas de dimensión infinita ha iluminado nuestro concepto de interacción y evolución de ondas no lineales que aparecen en varios contextos físicos. No es aún claro en este campo, actualmente en desarrollo, cuál será la contribución del cálculo de variaciones y los conceptos de él derivados. El tiempo pondrá en perspectiva valorar la importancia relativa de las varias ideas involucradas.