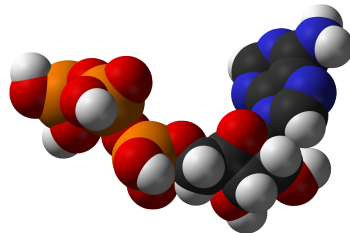


Anàlisi de xarxes metabòliques: On la biologia es troba amb l'anàlisi convexa

Joan Saldaña

Genòmica, proteòmica i metabolòmica són algunes de les anomenades ciències *òmiques* que podem trobar quan llegim alguna notícia de ciència o medicina, o consultem el pla d'estudis d'alguns graus i màsters de l'àmbit biomèdic. Sovint, gairebé sempre, són paraules que associem a l'Estadística i, més concretament, a l'anàlisi de dades ("dades òmiques" en diuen alguns) i al "big data". Totes aquestes àrees d'estudi formen part del que avui es coneix com a *Biologia de Sistemes* i, curiosament, si mireu el contingut de molts dels llibres sobre aquesta matèria, trobareu més capítols dedicats a matemàtiques (equacions diferencials, àlgebra lineal, optimització, teoria de grafs, etc.) que no pas a temes de biologia o d'estadística. Ara bé, m'atreviria a dir que, a casa nostra, una gran majoria de matemàtics desconex com d'útils són molts dels seus coneixements aplicats a allò que hom ha qualificat de biologia de l'era post-genoma. Per la meua part, espero que aquest article ajudi a difondre aquesta contribució de les Matemàtiques a (una part de) la Biologia del segle XXI.

D'aquestes ciències òmiques, ens centrarem en la metabolòmica, és a dir, en l'estudi sistemàtic a gran escala de substrats i productes que formen el metabolisme. A nivell cel·lular, el metabolisme es defineix com el conjunt de reaccions químiques que transformen els nutrients provinents de l'exterior (metabòlits externs) en les biomolècules que formen els éssers vius (metabòlits interns). La majoria d'aquestes reaccions estan catalitzades per enzims,



proteïnes controlades per la regulació genètica i que determinen les velocitats (taxes) de les reaccions.

Els conceptes que presentarem provenen de l'anàlisi convexa i, durant els darrers 25-30 anys, s'han anat generalitzant i aplicant a l'anàlisi de xarxes metabòliques per tal de definir quines rutes metabòliques fonamenten el seu funcionament. Per exemple, el concepte de vector elemental que veurem més endavant va ser introduït per Urbanczik i Wagner el 2005 en el context de fluxos en xarxes [18], però ja havia estat introduït molt abans, el 1969, per Rockafellar [11]. En termes generals, els investigadors en aquest camp estan interessats a establir quines propietats caracteritzen les rutes principals que transformen determinats substrats en determinats productes, i en construir algorismes que permetin destriar-les entre els milions de possibles rutes que defineixen les xarxes a nivell cel·lular. Pensem que, des d'un punt de vista algorímic, la naturalesa combinatoria de la formació de possibles rutes metabòliques (vg. [14]) comporta que els algorismes per a la seva enumeració inverteixin un temps de càlcul que creix de manera exponencial amb la mida de la xarxa. Per tant, avui dia encara s'estan proposant nous conceptes i nous mètodes de selecció i simplificació de rutes que ajudin a delimitar aspectes fonamentals de les xarxes metabòliques necessaris per analitzar-les a escala genòmica [2, 4].

Aquest esclat el va provocar l'allau de dades "òmiques" arran del desenvolupament del projecte genoma humà als anys 90, i la posterior seqüenciació del material genètic de molts organismes. Avui dia es tenen cartografiades les xarxes metabòliques de diversos organismes unicel·lulars a escala de tot el genoma i es publiquen reconstruccions del metabolisme humà cada cop més completes [16]. De fet, l'objectiu d'estudi en aquest àmbit passa, cada cop més, de les vies o rutes metabòliques considerades de manera aïllada, a les xarxes metabòliques que les integren i que van de les rutes lligades al que es coneix com metabolisme central del carboni, a les xarxes a escala genòmica que comprenen el conjunt de reaccions que tenen lloc a l'interior d'una cèl·lula. En biotecnologia i la indústria farmacèutica, aquest coneixement del metabolisme permet, entre altres coses, la identificació de rutes per augmentar la producció de certes molècules d'interès o el disseny de drogues específiques que permetin, per exemple, bloquejar certes reaccions cel·lulars en el tractament de malalties o en l'estudi de l'impacte de certes mutacions.

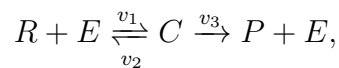
El punt de partida d'un model de metabolisme seria el següent. Conegudes les velocitats v_j de les n reaccions químiques que formen una xarxa, podem escriure la taxa de canvi de la concentració x_i del metabòlit m_i en

funció del temps com

$$\frac{dx_i(t)}{dt} = \sum_{j=1}^n S_{ij}v_j(t), \quad i = 1, \dots, m, \quad (1)$$

on S_{ij} és el coeficient estequiomètric del metabòlit m_i en la reacció j afectat de signe, de manera que $S_{ij} > 0$ si m_i és un producte de la reacció j (i, per tant, es produeixen S_{ij} molècules d' m_i), i $S_{ij} < 0$ si m_i n'és un reactiu (se'n consumeixen S_{ij} molècules).

Per exemple, les reaccions constitutives bàsiques de les rutes metabòliques són aquelles en què un substrat (o reactiu) R es combina amb una molècula d'enzim lliure E per formar, de manera reversible, un complex C , el qual acaba donant un producte P més la molècula inicial d'enzim lliure [15]. Aquestes reaccions les podem representar mitjançant el diagrama següent



on hem indicat les velocitats v_i de cada reacció. D'acord amb la definició donada de la matriu estequiomètrica S , per a aquestes reaccions tenim

$$S = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 1 \\ 1 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

on cada fila correspon a un dels quatre metabòlits de les tres les reaccions (R , E , C i P respectivament), i cada columna conté els coeficients estequiomètrics (afectats de signe) de cada una de les reaccions. Fixem-nos que en el cas de la primera reacció, que és reversible, s'ha desdoblada en dues d'irreversibles de manera que la segona columna de S és igual a la primera però amb signes canviats. Per reduir la dimensió del nostre sistema, molts cops es considera una sola velocitat per a cada reacció reversible que serà positiva o negativa segons la direcció en què tingui lloc la reacció. A més, en condicions biològiques normals, reaccions catalitzades per enzims com la de l'exemple tenen un fort caràcter unidireccional de manera que el flux net de la reacció anirà en el sentit de la producció del producte. De fet, rutes metabòliques oposades com la glicòlisi (oxidació de la glucosa) i la gliconeogènesi (formació de glucosa) no són reversibles sinó que sovint utilitzen reaccions diferents per raons termodinàmiques [5] (vg. Fig. 1).

D'altra banda, la llei d'acció de masses de la cinètica química ens diu que les velocitats de reacció (mols per unitat de volum i per unitat de temps) són proporcionals al producte de les concentracions dels *reactius* elevades al

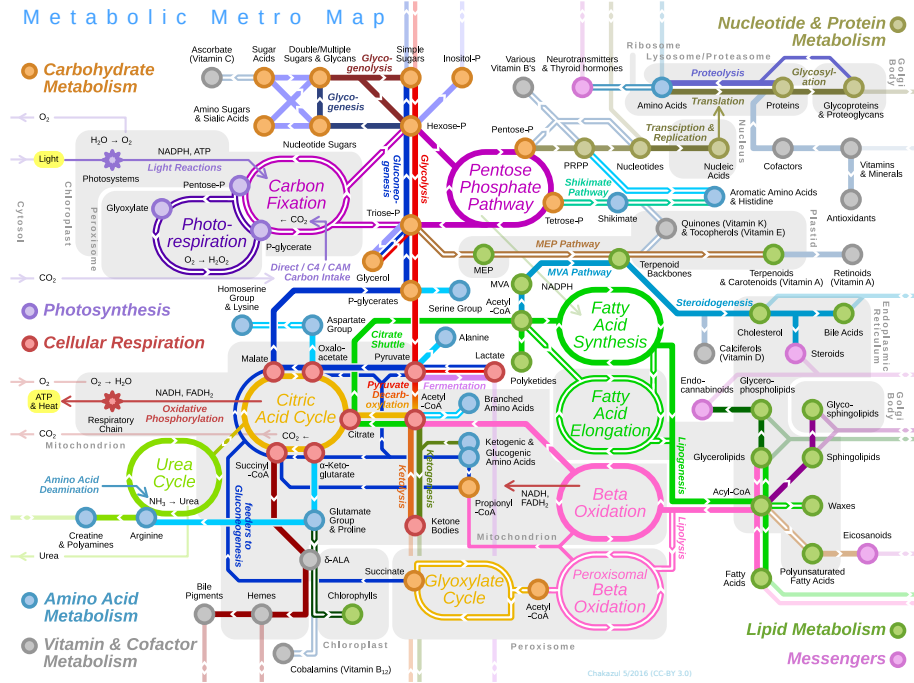


Figura 1: Esquema general del metabolisme cel·lular.

nombre de molècules amb què intervenen (com a reactius) en la reacció [15], és a dir, que la velocitat de la reacció j ve donada per

$$v_j(t) = k_j \prod_{i=1}^m (x_i(t))^{|S_{ij}|}, \quad j = 1, \dots, n, \quad (2)$$

on el coeficient de proporcionalitat k_j és l'anomenada constant de velocitat de la reacció j . A l'exemple anterior, com que tots els coeficients estequiomètrics són 1, el vector de velocitats és

$$\mathbf{v} = \begin{pmatrix} k_1 x_r x_e \\ k_2 x_c \\ k_3 x_c \end{pmatrix},$$

on x_r , x_e , i x_c són les concentracions de R , E i C , respectivament, i les $k_i > 0$. El fort caràcter unidireccional de la reacció reversible de l'exemple vindria donat per $k_1 \gg k_2$.

Així, doncs, introduint l'expressió (2) dins de l'equació (1) s'obté el sistema d'equacions diferencials no lineals polinomials per a les concentracions dels metabòlits de la xarxa. En l'exemple que estem considerant de la reacció

d'un substrat R amb un enzim E per acabar donant un producte P , el sistema per a les concentracions que en resulta, un cop multiplicada la matriu estequiomètrica S pel vector de velocitats \mathbf{v} , és

$$\begin{aligned}\frac{dx_r(t)}{dt} &= -k_1x_r(t)x_e(t) + k_2x_c(t), \\ \frac{dx_e(t)}{dt} &= -k_1x_r(t)x_e(t) + (k_2 + k_3)x_c(t), \\ \frac{dx_c(t)}{dt} &= k_1x_r(t)x_e(t) - (k_2 + k_3)x_c(t), \\ \frac{dx_p(t)}{dt} &= k_3x_c(t).\end{aligned}$$

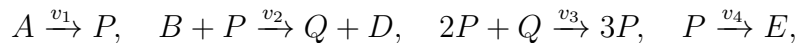
Tot i tractar-se d'un sistema de 4 equacions diferencials, el seu estudi és molt senzill de fer perquè es pot reduir a un sistema de 2 equacions aprofitant que el valor de x_p no apareix a les tres primeres equacions i, a més, que $\frac{dx_e}{dt} + \frac{dx_c}{dt} = 0$, és a dir, que la suma de les concentracions de C i E es manté constant en el temps: $x_e(t) + x_c(t) = x_e^0$ si inicialment $x_c(0) = 0$ i $x_e(0) = x_e^0$. Aquest fet ens diu que la concentració total d'enzim, la que es troba lliure (x_e) més la de l'enzim que es troba combinat formant un complex (x_c), és sempre la mateixa. Tenint en compte el que acaben de dir, el sistema al qual arribem és

$$\begin{aligned}\frac{dx_r(t)}{dt} &= k_2x_c(t) - k_1x_r(t)(x_e^0 - x_c(t)), \\ \frac{dx_c(t)}{dt} &= k_1x_r(t)(x_e^0 - x_c(t)) - (k_2 + k_3)x_c(t).\end{aligned}$$

L'únic equilibri d'aquest sistema és $(x_r^*, x_c^*) = (0, 0)$ i el comportament asimptòtic de les solucions es redueix a que tot el substrat inicial $x_r(0)$ acaba transformant-se en producte i el complex C acaba desapareixent de manera que tot l'enzim acaba estant en forma lliure. Aquest comportament, de fet, no és res que no haguéssim pogut predir des d'un bon començament, quan hem escrit les reaccions. Fixem-nos que no estem suposant ni síntesi ni degradació de l'enzim, com tampoc una entrada externa de substrat. Cap de les dues coses, però, no són certes si considerem el metabolisme en el seu conjunt i, quan es tenen en compte aquests i altres aspectes dels mecanismes de regulació, podem trobar-nos amb sistemes d'equacions diferencials les solucions dels quals tenen una dinàmica més interessant.

Aquests sistemes sorgeixen, per exemple, quan considerem les anomenades reaccions químiques *autocatalítiques*. En una d'aquestes reaccions, un o més dels productes formats és alhora un reactiu de la mateixa o d'altres

reaccions que hi estan acoblades dins d'un conjunt més gran de reaccions. Un dels sistemes en els quals es basa l'estudi d'aquest tipus de reaccions es coneix amb el nom de "Brusselator", anomenat així perquè qui el va proposar el 1968, I. Prigogine —premi Nobel de química el 1977— i el seu col·laborador R. Lefever, treballaven aleshores a la Universitat Lliure de Brusselles. Les reaccions considerades en aquest model són



on A i B són substrats que es troben en una concentració prou alta de manera que les seves concentracions no varien (x_a^0 i x_b^0 , respectivament), P i Q són reactius i productes de diferents reaccions, i D i E són els productes finals. La matriu estequiomètrica d'aquestes reaccions és

$$S = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

on cada fila correspon a una de les sis molècules implicades (A , B , P , Q , D i E , respectivament). Fixem-nos que l'element de la matriu S corresponent a P en la tercera reacció, S_{33} , ve donat pel balanç entre el nombre de molècules de P consumides i el nombre de molècules que se'n creen, és a dir, és el coeficient estequiomètric *net* de P en aquesta reacció: $S_{33} = 3 - 2 = 1$.

Procedint com hem fet a l'exemple anterior i tenint en compte que el coeficient estequiomètric de P com a reactiu en la tercera equació és 2, el vector \mathbf{v} de velocitats ve donat per

$$\mathbf{v} = \begin{pmatrix} k_1 x_a^0 \\ k_2 x_b^0 x_p \\ k_3 x_p^2 x_q \\ k_4 x_p \end{pmatrix}.$$

El sistema d'equacions diferencials per a les concentracions de P i Q que resulta del producte $S\mathbf{v}$ és

$$\frac{dx_p(t)}{dt} = \alpha_1 - \alpha_2 x_p(t) + \alpha_3 x_p^2(t) x_q(t),$$

$$\frac{dx_q(t)}{dt} = \alpha_4 x_p(t) - \alpha_3 x_p^2(t) x_q(t),$$

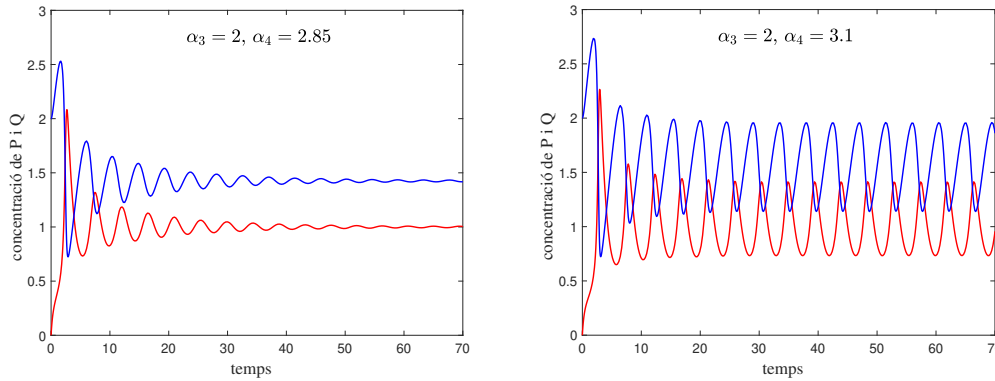


Figura 2: Solucions del Brusselator abans (esquerra) i després (dreta) de travessar la bifurcació de Hopf quan comencem sense el substrat P ($x_p(0) = 0$) i amb el substrat Q amb una concentració $x_q(0) = 2$. Passat el llindar $\alpha_4^c = 3$, apareixen les oscil·lacions al voltant de l'equilibri $(x_p^*, x_q^*) = (1, \alpha_4/\alpha_3)$.

on $\alpha_1 = k_1 x_a^0$, $\alpha_2 = k_2 x_b^0 + k_4$, $\alpha_3 = k_3$, i $\alpha_4 = k_2 x_b^0$. Aquest sistema té un únic equilibri positiu $(x_p^*, x_q^*) = \left(\frac{\alpha_1}{\alpha_2 - \alpha_4}, \frac{\alpha_4(\alpha_2 - \alpha_4)}{\alpha_1 \alpha_3} \right)$ i, a diferència de l'exemple anterior on la concentració de R no era constant, aquest equilibri existeix gràcies a l'entrada dels substrats A i B que fa que les seves concentracions es mantinguin constants i, per tant, que α_1 i α_4 siguin estrictament positives. Si, per simplificar les expressions adimensionalitzem les equacions per tenir $\alpha_1 = 1$ i $k_4 = 1$ (vg [9]), llavors $\alpha_2 = \alpha_4 + 1$ i $(x_p^*, x_q^*) = (1, \alpha_4/\alpha_3)$. En aquest cas, és fàcil veure que el sistema experimenta una bifurcació de Hopf supercrítica per a $\alpha_4^c = 1 + \alpha_3$ de manera que, per a $\alpha_4 > \alpha_4^c$, l'equilibri (x_p^*, x_q^*) es torna inestable i apareix un cicle límit estable que l'envolta. En altres paraules, d'acord amb el model del Brusselator, les concentracions de P i Q oscil·laran de manera periòdica sempre que tinguem concentracions prou altes del substrat B (vg. Fig. 2).

La primera constatació experimental de l'existència d'oscil·lacions en reaccions químiques la va fer el bioquímic rus B. Belousov construint un anàleg no biològic de la ruta metabòlica anomenada cicle de l'àcid cítric (o cicle de Krebs), en groc a la Figura 1. El caràcter oscil·latori de les reaccions es podia observar en la variació periòdica de la coloració de la solució on tenien lloc. Belousov, el 1951, va escriure un article sobre la seva troballa però, malauradament, mai no va poder ser publicat a cap revista perquè el dogma imperant a l'època era que les reaccions químiques oscil·lants no eren possibles perquè contradeien la segona llei de la termodinàmica. No va ser fins a 1959 que Belousov va publicar una breu nota en les actes d'un congrés rus de medicina, poca cosa donada la importància del que havia descobert. Més

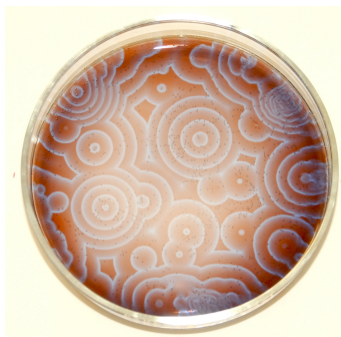


Figura 3: Patró espacial de la reacció de Belousov-Zhabotinski.

tard, el 1961, A. Zhabotinski, aleshores estudiant de doctorat, va rebre l'encàrrec del seu director de tesi de continuar l'estudi de la reacció de Belousov perquè no es coneixia encara quin era el seu mecanisme de reacció, és a dir, la seqüència detallada de totes les reaccions implicades [1]. Avui dia, el que es coneix com a reacció de Belousov-Zhabotinski ha passat a englobar tota una classe de reaccions que manifesten un comportament oscil·latori tant en el temps com a l'espai (vg. Fig. 3). Per cert, el model teòric més senzill per a la dinàmica d'aquesta reacció, publicat el 1974 per R.J. Field i R.M. Noyes, està format per un sistema de tres equacions diferencials i les oscil·lacions són degudes a l'existència d'un cicle límit atractor a \mathbb{R}^3 [9]. El model rep el nom d'"Oregonator": endevineu el perquè d'aquest nom?

1 Descripció minimal de xarxes metabòliques: el con de fluxos

Quan s'estudia el metabolisme a una escala que implica centenars o milers de metabòlits, la manca d'un coneixement quantitatiu de la cinètica de moltes de les reaccions així com el seu elevat nombre fa que una anàlisi completa de la dinàmica del sistema d'equacions diferencials que en resulta no sigui factible. Per fer-nos-en una idea, en la reconstrucció del metabolisme humà consensuada per diferents grups de recerca a partir de diferents bases de dades feta a [16], el nombre de reaccions arriba als 7440 i el de metabòlits a 5063.

Una alternativa a l'estudi detallat del sistema (1) consisteix a utilitzar la informació estequiomètrica de les reaccions que defineixen una xarxa metabòlica per estudiar els equilibris del sistema en termes del vector de velocitats $\mathbf{v} = (v_1, \dots, v_n)^T$, també anomenat *vector de fluxos*. Cadascun d'aquests

vectors representa un estat estacionari de la xarxa en el qual algunes de les reaccions romandran actives mentre que d'altres estaran inactives ($v_j = 0$) de manera que \mathbf{v} satisfà

$$S\mathbf{v} = 0, \quad (3)$$

on la matriu estequiomètrica S té dimensions $m \times n$ essent m el nombre de metabòlits en les n reaccions. Així, cada solució d'aquest sistema homogeni representa una distribució de fluxos en la xarxa que permet que cada metabòlit implicat tingui una concentració constant en el temps.

Fixem-nos que si \mathbf{v} és un vector de fluxos i totes les reaccions que el formen són reversibles, llavors $-\mathbf{v}$ també ho ha de ser. Per tant, si totes les reaccions de la xarxa fossin reversibles, els seus estats estacionaris, que venen donats pel conjunt de les solucions de (3), seria l'anomenat *espai nul* (o nucli) de la matriu S , que com sabem té estructura d'espai vectorial. Així doncs, donada una base d'aquest espai, qualsevol estat estacionari de la xarxa podria expressar-se com una combinació lineal dels seus vectors. De fet, en alguns dels primers treballs sobre xarxes metabòliques que pretenien estudiar les seves propietats estructurals apareguts a finals del anys 80 i principis dels 90, aquests vectors s'interpretaven com un conjunt de vies o rutes metabòliques, independents entre si, que caracteritzen les possibles rutes de la xarxa en estat estacionari [10]. Ara bé, el fet que hi hagi més d'una manera de descompondre un vector (les bases dels espais vectorials no són úniques), i que moltes reaccions en condicions biològiques normals tinguin un caràcter marcadament irreversible [5], fa que considerar els vectors d'una base com a definidors de rutes metabòliques fonamentals sobre les quals es construeix la xarxa no sigui l'opció més adient des d'un punt de vista bioquímic. En particular, quan es consideren reaccions irreversibles, no totes les bases de $\text{nul}(S)$ són admissibles i, per a aquelles que ho són, les combinacions lineals dels seus vectors seran admissibles si els fluxos corresponents a rutes irreversibles no apareixen amb coeficients negatius. Si això no fos així voldria dir que, per obtenir certes rutes metabòliques, caldria considerar un o més vectors de fluxos irreversibles en una direcció oposada a la que tenen determinada termodinàmicament. En el cas límit en què totes les reaccions de la xarxa fossin irreversibles, només podríem considerar combinacions lineals no negatives, també anomenades *combinacions còniques*, de vectors de fluxos.

Il·lustrem el que hem definit i discutit fins ara amb un exemple. Considerem la xarxa de reaccions de la Figura 4. La matriu estequiomètrica associada S té 3 files perquè hi ha tres metabòlits interns (A , B i C), i 6 columnes corresponents a les sis reaccions involucrades amb els coeficients estequiomètrics afectats de signe (signe $-$ per als reactius i signe $+$ per als productes). El metabòlit A s'importa des de l'exterior del nostre sistema, el

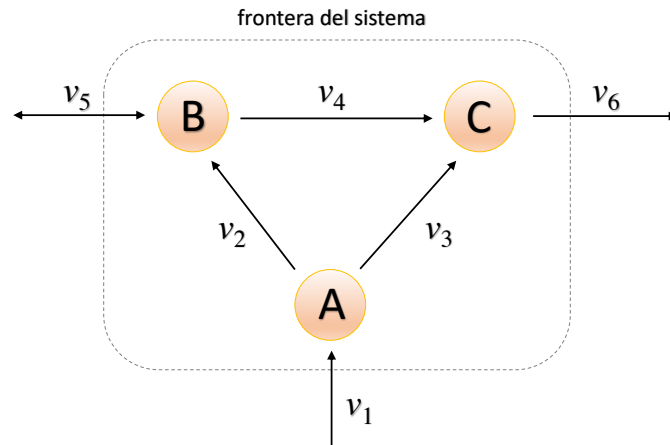


Figura 4: Xarxa de reaccions amb tres metabòlits interns A, B i C. El flux v_5 d'intercanvi del metabòlit B amb l'exterior és reversible.

metabòlit C s'exporta a l'exterior i el metabòlit B es bescanvia de manera reversible. Així, si numerem les reaccions com els fluxos de la Figura 4, i fem que la primera fila correspongui al metabòlit A , la segona a B i la tercera a C , llavors la matriu que en resulta és

$$S = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 \end{pmatrix}.$$

D'acord amb la figura, els vectors \mathbf{v} que compleixin (3) i siguin compatibles amb el sentit indicat dels fluxos tindran sis components v_i amb $v_1, v_2, v_3, v_4, v_6 \geq 0$. Pel que fa a v_5 , com que la reacció associada és reversible, pot ser positiu (B entra dins del sistema) o negatiu (en surt). Per construir S , hem pres el signe positiu ($S_{25} = 1 > 0$). Per a aquesta matriu, podem obtenir una base del seu espai nul resolent (3) tot prenent v_4, v_5 i v_6 com a paràmetres, de manera que

$$\text{nul}(S) = \langle (0, 1, -1, 1, 0, 0)^T, (1, 1, 0, 0, -1, 0)^T, (1, 0, 1, 0, 0, 1)^T \rangle.$$

Observem que el primer vector de la base no és compatible amb les direccions dels fluxos de la xarxa de la Figura 4. D'altra banda, el vector $(1, 1, 0, 1, 0, 1)$ sí que ho és i, com que s'obté com a suma del primer i el tercer vector de la base, podem intercanviar-lo amb el primer vector per construir-ne una de nova on els seus tres vectors siguin admissibles.

Així doncs, cal tenir en compte la naturalesa marcadament irreversible de certes reaccions i que, per tant, aquestes tenen un flux (net) v_i associat

no negatiu. Per fer-ho, l'espai de vectors de fluxos estacionaris admissibles (*espai de fluxos*) F que cal considerar és

$$F = \{\mathbf{v} \in \mathbb{R}^n : S\mathbf{v} = \mathbf{0}, v_i \geq 0 \forall i \in \text{Irr}\}, \quad (4)$$

on Irr és el conjunt dels índexs de les reaccions irreversibles. Per definició, doncs, l'espai de fluxos és la intersecció de $\text{nul}(S)$ amb semiespais no-negatius que passen per l'origen $\{\mathbf{v} \in \mathbb{R}^n : v_i \geq 0, i \in \text{Irr}\}$ corresponents a les reaccions irreversibles. Geomètricament, F defineix un con polièdric convex anomenat *con de fluxos*.

Aquest tipus de con, de fet, és un cas particular del que en geometria es coneix com *con polièdric general* C_g que, en la seva representació *implícita*, correspon a la intersecció de q semiespais que passen per l'origen:

$$C_g = \{\mathbf{v} \in \mathbb{R}^n | A\mathbf{v} \geq 0\},$$

essent A una matriu $q \times n$ adient. Així, el con de fluxos F es pot escriure en aquesta formulació prenent

$$A = \begin{pmatrix} S \\ -S \\ I_{\text{irr}} \end{pmatrix} \in \mathbb{R}^{q \times n}$$

on I_{irr} és una matriu $|\text{Irr}| \times n$ que conté les condicions de no negativitat de les reaccions irreversibles, $|\text{Irr}|$ és la cardinalitat del conjunt d'aquestes reaccions i $q = 2m + |\text{Irr}|$ [5]. La formulació de C_g en termes de generadors, l'anomenada representació *explícita*, comprèn les combinacions lineals no negatives (*combinacions còniques*) de vectors de fluxos $\{\mathbf{y}_j\}$ irreversibles més les combinacions lineals dels vectors de fluxos $\{\mathbf{z}_k\}$ que formen una base del subespai vectorial format per tots els vectors de fluxos que són reversibles (*espai de linealitat* del con), és a dir,

$$C_g = \left\{ \sum_{j \in J} \beta_j \mathbf{y}_j + \sum_{k \in K} \gamma_k \mathbf{z}_k \mid \beta_j \geq 0, \gamma_k \in \mathbb{R} \right\}. \quad (5)$$

En aquest cas, el conjunt $\{\mathbf{y}_j, \mathbf{z}_k\}$ de generadors de C_g no és únic, fins i tot si ens restringim a conjunts minimalis de generadors [6].

2 Anàlisi de rutes metabòliques

L'estudi de les xarxes metabòliques té com a principal objectiu la identificació de rutes que siguin rellevants des d'un punt de vista bio(tecno)lògic i, en

particular, les rutes i cicles més curts que representin estats estacionaris d'aquestes xarxes. Per entendre algunes de les definicions de rutes metabòliques mínimals que s'han proposat en la literatura, comencem amb unes nocions bàsiques d'anàlisi convexa.

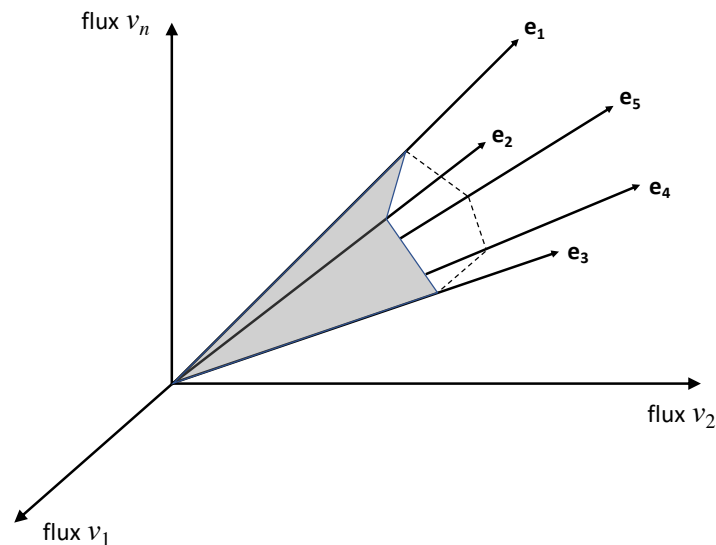


Figura 5: Espai de fluxos d'una xarxa metabòlica i els seus vectors extrems. Les solucions del sistema $S\mathbf{v} = \mathbf{0}$ amb les condicions d'irreversibilitat $v_i \geq 0$ donen lloc, en aquest exemple, a un con de fluxos apuntat no acotat. Els vectors extrems \mathbf{e}_j corresponen a les arestes del con i constitueixen el seu únic conjunt minimal de generadors. Qualsevol vector del con representa, doncs, un estat estacionari de la xarxa compatible amb les condicions d'irreversibilitat.

La propietat que defineix un con C és que, si $\mathbf{v} \in C$, llavors el raig $\{\alpha \mathbf{v} \mid \alpha \geq 0\}$ està totalment contingut a C . A més, un vector $\mathbf{e} \in C$ és *extrem* si no pot escriure's com a suma de dos vectors de C linealment independents. Per últim, recordem que un con C és *apuntat* si no conté rectes, és a dir, si $\{\alpha \mathbf{v} \mid \alpha \in \mathbb{R}\} \not\subset C$ per a cap \mathbf{v} . Per tant, C és apuntat si i només si el seu espai de linealitat és $\{\mathbf{0}\}$. Geomètricament, els raigs extrems $\{\alpha \mathbf{e} \mid \alpha \geq 0\}$ corresponen a les arestes d'un con polièdric apuntat. La propietat de ser apuntat la compleixen, per exemple, els cons de fluxos de xarxes metabòliques tals que totes les reaccions de bescanvi amb l'exterior són irreversibles perquè aquest fet no permet l'existència de vectors de fluxos reversibles. En general, tindrem cons apuntats quan hi hagi reaccions reversibles però no hi hagi cap vector de fluxos que ho sigui perquè tots ells contenen una o més reaccions irreversibles. Aquesta situació es presenta, per exemple, quan les úniques reaccions reversibles són les que comporten bescanvis de metabòlits

amb l'exterior del sistema i cada metabòlit participa, com a màxim, en una única reacció de bescanvi (vegeu l'exemple de la Figura 4) [10]. El que fa relativament senzill l'estudi dels estats estacionaris d'aquesta mena de xarxes de caràcter marcadament irreversible és el fet que els cons polièdrics convexos apuntats poden generar-se a partir de combinacions lineals no negatives dels seus vectors extrems (teorema de Minkowski) [8]. En aquest cas, l'expressió de l'espai de fluxos F donada per (5) es redueix a

$$F = \left\{ \sum_j \beta_j \mathbf{e}_j \mid \beta_j \geq 0 \right\}, \quad (6)$$

és a dir, amb $\{\mathbf{z}_k\} = \emptyset$ i essent $\{\mathbf{y}_j\} = \{\mathbf{e}_j\}$ (amb $\mathbf{e}_i \neq \lambda \mathbf{e}_j$ si $i \neq j$) l'únic conjunt minimal de generadors de F (vg. Fig. 5).

En el context de les xarxes metabòliques, els raigs extrems corresponen a les anomenades vies extremes (*extreme pathways*) per a les quals totes les reaccions internes són irreversibles. Quan hi ha vectors de fluxos que són reversibles, F ja no és un con apuntat i, per tant, no té raigs extrems o, el que és el mateix, no hi ha un únic conjunt minimal de generadors F . Aquest fet va motivar la introducció de definicions de rutes metabòliques basades en un espai de fluxos "augmentat". Aquest nou espai resulta de dividir cada reacció reversible interna en dues d'irreversibles i té, per tant, una dimensió addicional per a cada reacció desdoblada. D'aquesta manera, el con de fluxos obtingut és apuntat i torna a tenir sentit parlar de raigs extrems, tot i que ho deixin de ser quan els reduïm a l'espai de fluxos original. Els anomenats corrents extrems (*extreme currents*) i les vies extremes abans esmentades són exemples d'aquests raigs en espais de fluxos augmentats. La diferència entre aquestes dos tipus de rutes apareix a l'hora de construir el corresponent espai de fluxos augmentat. Concretament, el que les distingeix és si s'han dividit en dues o no aquelles reaccions amb l'exterior que són reversibles. Així, en el primer cas, parlem de corrents extrems mentre que, en el segon, parlem de vies extremes [6].

Fixem-nos que una conseqüència immediata de la definició de vector extrem és que aquest no es pot descompondre en vectors de fluxos més simples. Dit d'una altra manera, si $\mathbf{e} \in F$ és un vector extrem, llavors \mathbf{e} no pot expressar-se com a suma de dos vectors de F tals que, a més de tenir les mateixes components nul·les que ell, en tinguin alguna més. En termes de reaccions, un vector de fluxos no descomponible (també anomenat *irreductible* o de *suport minimal*) representa un conjunt de reaccions que formen una "unitat funcional" de la xarxa de manera que, si una d'elles no està activa ($v_i = 0$), la resta de reaccions no poden definir per elles mateixes un estat estacionari no trivial de la xarxa. En general, però, la irreductibilitat d'un

vector de fluxos no està lligada necessàriament al fet de ser un vector extrem de F perquè hi pot haver vectors de suport minimal obtinguts a partir de combinacions còniques de vectors extrems. Els vectors \mathbf{m}_i de F de suport minimal s'anomenen *modes de flux elementals* (MFE) i han jugat un paper molt important en l'anàlisi de rutes metabòliques des de la seva introducció el 1994 [13]. Per construir el conjunt de MFE associat a un con de fluxos, escollirem un representant de cada raig $\{\beta \mathbf{m}_i \mid \beta \geq 0\}$ associat a una unitat funcional diferent.

Aquesta definició de MFE implica que aquests formen un conjunt de generadors del con de fluxos però, tanmateix, no garanteix que aquest conjunt sigui minimal. Així, quan el con és apuntat, el conjunt de MFE inclou els vectors extrems i, quan no ho és, inclou una base del seu espai de linealitat [10]. Un exemple del primer cas el dóna la xarxa de la Figura 4, on totes les reaccions internes són irreversibles però n'hi ha una de bescanvi amb l'exterior que és reversible. En les Figures 6 i 7 es mostren, respectivament, els vectors extrems i els MFE d'aquesta xarxa, i es veu com els vectors extrems formen un subconjunt (cònicament independent) dels MFE. Aquest exemple mostra un fet que és general, i és que es poden obtenir conjunts minimal de generadors a partir dels MFE eliminant tots aquells modes que es poden expressar com una combinació cònica de la resta de modes [6].

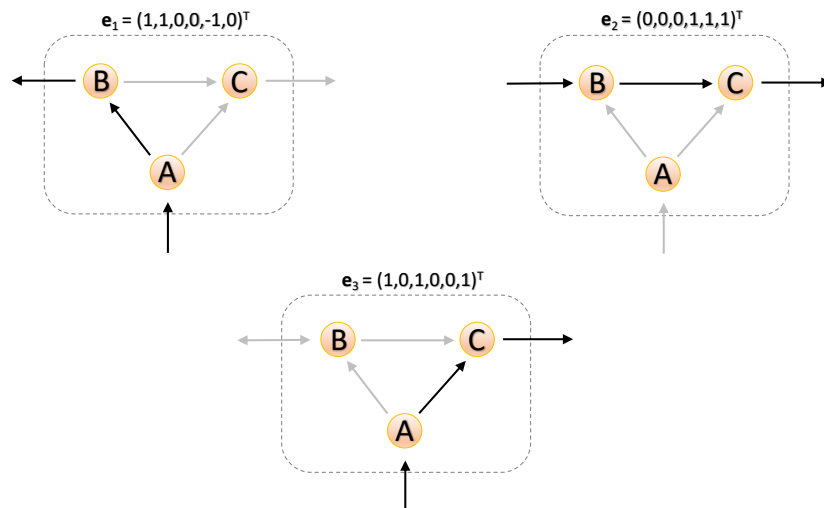


Figura 6: Vectors extrems \mathbf{e}_i associats a la xarxa de reaccions de la Figura 4.

Una definició alternativa de MFE, deslligada de la noció de suport minimal, és la que es basa en la *suma conforme de vectors* [5, 8]. De fet, aquesta definició ens permetrà determinar un conjunt de vectors de cons polièdrics

generals (i de poliedres) que tindran propietats anàlogues a les dels MFE del con de fluxos. Per entendre la idea que hi ha al darrere d'aquesta nova definició, observem que, a la xarxa de la Figura 4, \mathbf{m}_4 s'obté a partir d' \mathbf{e}_1 i \mathbf{e}_2 gràcies al fet que la cinquena component d'aquests vectors extrems tenen signes oposats i, per tant, es cancel·len l'una amb l'altra quan les sumem (vg. Fig. 7):

$$\mathbf{m}_4 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{e}_1 + \mathbf{e}_2.$$

Ara bé, des d'un punt de vista bioquímic, no té sentit construir una ruta metabòlica a partir de combinacions de vectors de fluxos on una mateixa reacció reversible tingui lloc en sentits oposats (aquests han de formar part de rutes diferents). L'absència de cancel·lacions de fluxos és, doncs, la idea que hi ha al darrere del concepte de suma conforme de vectors i, malgrat ser força natural en aquest context, és curiós que no fos tinguda en compte durant els primers anys d'estudi de rutes metabòliques [5]. Així, donats dos vectors $\mathbf{u}, \mathbf{v} \in F$, direm que la suma $\mathbf{w} = \mathbf{u} + \mathbf{v}$ és conforme si $w_i = 0$ implica $u_i = v_i = 0$, $w_i > 0$ implica $u_i, v_i \geq 0$, i $w_i < 0$ implica $u_i, v_i \leq 0$, $\forall i \in \{1, \dots, n\}$. En anàlisi convexa, la descomposició de vectors sense cancel·lacions va ser introduïda per Rockafellar en 1969, mentre que Urbanczik i Wagner la van introduir en l'anàlisi de rutes metabòliques l'any 2005.

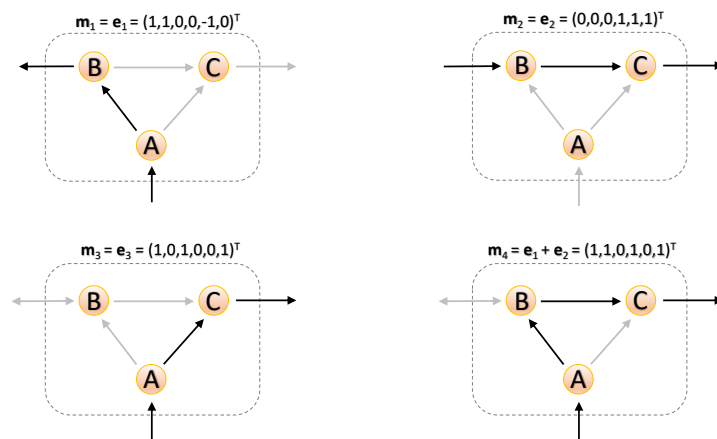


Figura 7: Modes elementals \mathbf{m}_i associats a la xarxa de reaccions de la Figura 4.

D'acord amb aquesta definició de suma conforme, els MFE són vectors

del con de fluxos que no es poden descompondre de manera conforme, és a dir, que no poden expressar-se com una suma sense cancel·lacions de dos vectors de fluxos. D'aquesta manera, arribem a un resultat fonamental en l'anàlisi de rutes metabòliques: *el conjunt de MFE constitueix l'únic conjunt minimal de generadors conformes del con de fluxos F* [5, 8].

Els MFE han esdevingut una eina fonamental per a l'estudi de xarxes que corresponen a models de l'anomenat metabolisme central, i que podem descriure com aquella part del metabolisme que comprèn les reaccions que converteixen sucres en energia i metabòlits precursors de biomolècules més complexes. Aquestes xarxes són prou petites perquè el càlcul dels MFE sigui factible. Des del punt de vista de les aplicacions, els MFE corresponen a subxarxes minimals (en el sentit de ser indescomponibles de manera conforme) que realitzen una certa funció en estat estacionari. Això ha fet que aquests hagin estat utilitzats per identificar rutes mínimes de conversió d'un cert substrat a un cert producte i, també, cicles metabòlics (rutes tancades). Ara bé, per adornar-nos de l'esforç computacional que comporta el càlcul dels MFE, fins i tot en xarxes relativament petites com aquestes, és interessant saber que, per exemple, el nombre de MFE calculats per al con de fluxos d'un model de metabolisme del bacteri *Escherichia coli* amb $m = 75$ metabòlits interns i $n = 101$ reaccions, 48 de les quals són reversibles, arriba als 169.916 modes [3].

Una de les raons de l'èxit de la noció de MFE en les aplicacions i, en especial, en els estudis de manipulacions de xarxes, rau en la següent propietat: *si eliminem una part d'una xarxa, el conjunt de MFE que en resulta està format per aquells MFE de la xarxa original que no contenen cap de les reaccions eliminades*, la qual cosa estalvia haver de calcular-los per a cada una de les subxarxes obtingudes. Una situació on això es dona és, per exemple, quan es bloqueja l'expressió de certs gens de manera que algunes reaccions deixen d'estar actives per falta de l'enzim que les catalitza o d'un substrat necessari perquè aquestes reaccions es puguin produir. Aquesta és una propietat que no tenen, per exemple, els anomenats conjunts minimals de generadors de l'espai de fluxos, que també s'han utilitzat per caracteritzar les xarxes metabòliques [6].

3 El poliedre de fluxos

A part de la condició d'irreversibilitat ($v_i > 0$) imposada a certes reaccions, sovint es té informació addicional sobre la velocitat de les reaccions que voldríem incorporar en l'anàlisi de les rutes metabòliques. Per exemple, per a algunes reaccions de bescanvi es coneixen fites superiors i/o inferiors de les

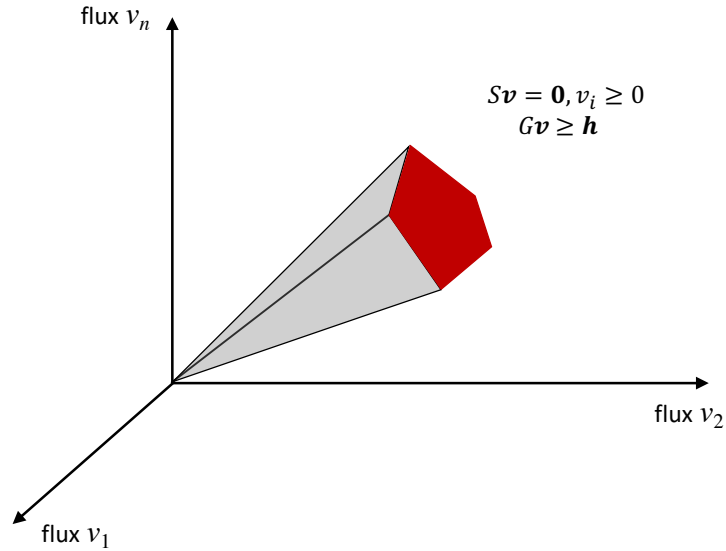


Figura 8: Poliedre de fluxos acotat. La base (en vermell) apareix per l'existència de restriccions addicionals als fluxos de la Fig. 5.

seves velocitats: $\underline{v}_i \leq v_i \leq \bar{v}_i$. Aquestes condicions addicionals les podem escriure matricialment com: $G\mathbf{v} \geq \mathbf{h}$ amb $\mathbf{h} > \mathbf{0}$ i $G \in \mathbb{R}^{r \times n}$, de manera que l'espai de fluxos ara ve donat pel conjunt

$$PF = \{\mathbf{v} \in \mathbb{R}^n \mid S\mathbf{v} = \mathbf{0}, v_i \geq 0 \forall i \in \text{Irr}, G\mathbf{v} \geq \mathbf{h}\}. \quad (7)$$

Geomètricament, PF defineix un poliedre (acotat o no) anomenat *poliedre de fluxos*.

Com passava abans amb el con de fluxos, aquest poliedre de fluxos és un cas particular de poliedre general que, en la seva representació implícita, ve donat per $P = \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} \geq \mathbf{b}\}$ amb

$$A = \begin{pmatrix} S \\ -S \\ I_{\text{irr}} \\ G \end{pmatrix} \in \mathbb{R}^{q \times n}, \quad \text{i} \quad \mathbf{b} = \begin{pmatrix} \mathbf{0} \\ \mathbf{h} \end{pmatrix} \in \mathbb{R}^q, \quad (8)$$

on $q = 2m + |\text{Irr}| + r$. Un resultat ben conegut de l'anàlisi convexa ens diu que tot poliedre P es pot expressar com un poliedre acotat, el *polítop*, i l'anomenat *con de recessió* $C_P = \{\mathbf{v} \in \mathbb{R}^n \mid A\mathbf{v} \geq \mathbf{0}\}$, un con polièdric general que conté les direccions no acotades de P . Recordem que \mathbf{p} és un *vèrtex* de P si no es pot posar com a combinació lineal convexa ($\alpha \mathbf{v}_1 + (1 - \alpha) \mathbf{v}_2$, $0 < \alpha < 1$) de cap parell de vectors $\mathbf{v}_1, \mathbf{v}_2$ de P ($\mathbf{v}_1 \neq \mathbf{v}_2$), i que un

poliedre acotat ve donat pel conjunt de combinacions lineals convexes dels seus vèrtexs. A més, direm que P és apuntat si ho és el seu con de recessió C_P . D'aquestes definicions se segueix que, si un poliedre no és apuntat, llavors no té vèrtexs ni, tampoc, un únic conjunt minimal de generadors [5]. La representació explícita de P ve donada, doncs, per

$$P = \left\{ \sum_{i \in I} \alpha_i \mathbf{p}_i + \sum_{j \in J} \beta_j \mathbf{y}_j + \sum_{k \in K} \gamma_k \mathbf{z}_k \mid \alpha_i, \beta_j \geq 0, \gamma_k \in \mathbb{R}, \sum_{i \in I} \alpha_i = 1 \right\}, \quad (9)$$

essent $\{\mathbf{p}_i\}$ el conjunt de generadors acotats (vèrtexs) del politop ($\mathbf{p}_i \in \mathbb{R}^n$), i $\{\mathbf{y}_j, \mathbf{z}_k\}$ el conjunt de generadors (no acotats) de C_P ($\mathbf{y}_j, \mathbf{z}_k \in \mathbb{R}^n$). La qualificació d'acotat o no acotat dels generadors ve donada perquè, en el primer cas, els pesos de la combinació són acotats ($\alpha_i \in [0, 1]$) mentre que, en el segon cas, no ho són [17].

Un exemple de poliedre no acotat a \mathbb{R}^2 el dona la Figura 9 i el podem escriure, en les dues representacions (implícita i explícita), com

$$\begin{aligned} P_1 &= \left\{ \mathbf{v} \in \mathbb{R}^2 \mid \begin{pmatrix} 3 & 1 \\ -3 & 3 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \geq \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\} \\ &= \left\{ \alpha_1 \begin{pmatrix} 5/6 \\ 1/2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1/6 \\ 1/2 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \beta_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \beta_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \right. \\ &\quad \left. \alpha_i, \beta_j \geq 0, \sum_i \alpha_i = 1, \right\}. \quad (10) \end{aligned}$$

En aquest cas, el politop és el triangle de vèrtexs $\mathbf{p}_1 = (5/6, 1/2)$, $\mathbf{p}_2 = (1/6, 1/2)$ i $\mathbf{p}_3 = (0, 1)$, i el con de recessió és apuntat i té per arestes $\left\{ \beta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid \beta \geq 0 \right\}$ i $\left\{ \beta \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid \beta \geq 0 \right\}$ (vg. Fig. 9).

Malgrat que les representacions de cons i poliedres en termes de generadors és útil en certs casos, el conjunt minimal de generadors no té propietats que són molt desitjables des del punt de vista de les aplicacions com sí les tenen, en canvi, els MFE en el cas dels cons de fluxos. Per exemple, ja hem comentat que el conjunt de MFE d'una xarxa metabòlica conté els conjunts de MFE de totes les possibles subxarxes i que, a més, és l'únic conjunt minimal de generadors del con de fluxos que són indescomponibles de manera conforme. En general, però, els MFE del con de recessió d'un poliedre no satisfan les noves restriccions imposades a partir de la informació addicional sobre les velocitats (d'algunes) de les reaccions.

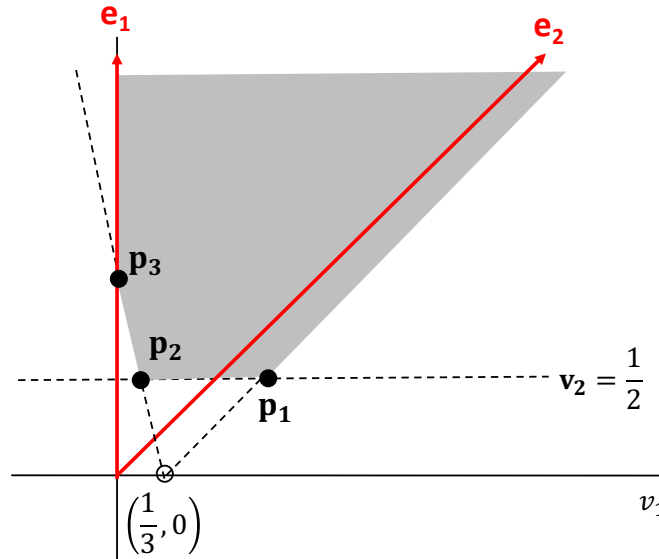


Figura 9: Poliedre no acotat (en gris) definit per (10). Els seus vèrtexs són p_1 , p_2 , i p_3 , i és no acotat en les direccions d' e_1 i e_2 que defineixen les arestes del seu con de recessió.

Anem, doncs, a definir un conjunt de vectors que jugarà el mateix paper en el cas dels poliedres que el jugat pels MFE en el cas dels cons, i ho farem utilitzant la noció de suma conforme introduïda a la secció 2.

4 Un pas més enllà dels MFE: els vectors elementals

La idea geomètrica que hi ha darrere del concepte de vector elemental es basa en el fet que la intersecció d'un poliedre P_0 amb un hiperoctant Ω_s de \mathbb{R}^n dóna lloc a un subpoliedre P_s que és sempre apuntat. En efecte, tots els elements de $P_s = P_0 \cap \Omega_s$ tenen un patró de signes $s \in \{-1, 1\}^n$ determinat per $\Omega_s = \{x \in \mathbb{R}^n \mid x_i s_i \geq 0, i = 1, \dots, n\}$ i, per tant, l'espai de linealitat del seu con de recessió C_{P_s} és el trivial. Així doncs, P_s tindrà un conjunt minimal de generadors format pels seus vèrtexs i pels vectors extrems del seu con de recessió.

Il·lustrem aquesta idea amb l'exemple del poliedre no acotat

$$P_0 = \left\{ v \in \mathbb{R}^2 \mid \begin{pmatrix} 3 & 1 \\ -3 & 3 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \geq \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}. \quad (11)$$

P_0 ve definit, doncs, per les mateixes inequacions que el poliedre P_1 donat per (10) excepte la condició de no negativitat de v_1 . Per tant, $P_1 = P_0 \cap \Omega_1$ essent $\Omega_1 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_i \geq 0, i = 1, 2\}$ el primer quadrant. La representació explícita de P_0 és

$$P_0 = \left\{ \alpha_1 \begin{pmatrix} 5/6 \\ 1/2 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1/6 \\ 1/2 \end{pmatrix} + \beta_1 \begin{pmatrix} -1 \\ 3 \end{pmatrix} + \beta_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \right. \\ \left. \alpha_i, \beta_j \geq 0, \sum_i \alpha_i = 1, \right\},$$

d'on veiem que P_0 té només dos vèrtexs, a diferència de P_1 que en té tres, i dues arestes.

El conjunt de *vectors elementals* (VE) d'un poliedre P es defineix com la unió dels conjunts minimal de generadors dels subpoliedres P_s obtinguts de les interseccions de P amb els 2^n hiperoctants Ω_s de \mathbb{R}^n . Hi haurà, doncs, dos tipus de generadors: els acotats formats pels vèrtexs dels diferents P_s , i els no acotats formats pels vectors extrems dels corresponents C_{P_s} . Com que $P = \bigcup P_s$, aquesta definició de VE comporta que aquests vectors formin un conjunt generador de P però, com passa també amb els MFE, no formin un conjunt minimal de generadors.

Així, per construcció, els VE acotats d'un poliedre P estan formats pels elements que no es poden expressar de manera conforme com a combinació lineal convexa de vectors de P i, per tant, inclouen els seus vèrtexs, però no només aquests perquè alguns d'aquests VE són vèrtexs de subpoliedres que no ho són de P . A l'exemple anterior, els VE acotats de P_0 són els tres vèrtexs del subpoliedre P_1 dels quals només dos, \mathbf{p}_1 i \mathbf{p}_2 , són vèrtexs de P_0 . Fixem-nos que \mathbf{p}_3 , malgrat no ser un vèrtex de P_0 , no es pot expressar com una combinació lineal convexa sense cancel·lacions d'elements de P_0 . A més, \mathbf{p}_3 és també vèrtex del subpoliedre P_2 obtingut per la intersecció de P_0 amb el segon quadrant $\Omega_2 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \leq 0, x_2 \geq 0\}$ (vg. Fig. 10). En aquest exemple, doncs, $P_0 = P_1 \cup P_2$.

Pel que fa als VE no acotats de P , són vectors del seu con de recessió C_P no descomponibles de manera conforme [8] i, per tant, els vectors extrems de C_P en formen part. En general, però, hi ha haurà VE no acotats que no es correspondran amb cap vector extrem de C_P malgrat ser-ho del con de recessió C_{P_s} d'algun dels subpoliedres P_s . Per exemple, el poliedre P_0 que hem definit té tres VE no acotats que són proporcionals a $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, i $\begin{pmatrix} -1 \\ 3 \end{pmatrix}$, respectivament. Els dos primers defineixen les arestes \mathbf{e}_1 i \mathbf{e}_2 de

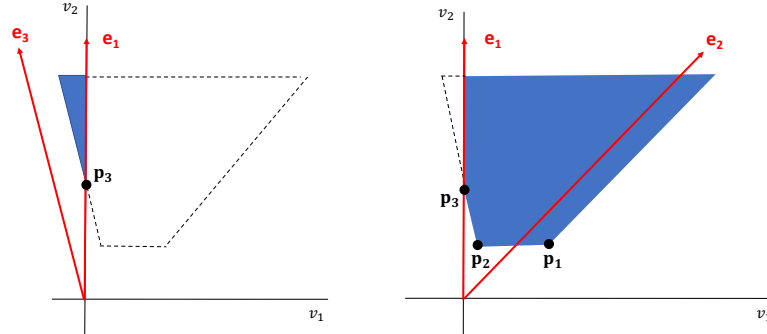


Figura 10: En blau, els poliedres P_1 (dreta) i P_2 (esquerra) resultants de la intersecció del poliedre P_0 definit per (11) amb el primer i el segon quadrants, respectivament: $P_0 = P_1 \cup P_2$. Els vectors elementals de P_0 són 6: 3 d'ells corresponen a la unió dels vèrtexs de P_1 i P_2 (\mathbf{p}_1 , \mathbf{p}_2 , \mathbf{p}_3), i els altres 3 a la unió dels vectors extrems dels seus cons de recessió (\mathbf{e}_1 , \mathbf{e}_2 , i \mathbf{e}_3).

C_{P_1} , mentre que els dos últims defineixen les arestes de C_{P_2} (vg. Fig. 10). D'aquests tres VE no acotats de P_0 , però, només el primer i el tercer són arestes de C_{P_0} . Fixem-nos que, sense cancel·lacions, no podem obtenir el segon VE a partir dels altres dos i, per tant, els tres VE formen un conjunt minimal de generadors conformes no acotats de P .

Aquest exemple ens mostra de manera intuïtiva una de les propietats que els VE tenen en comú amb els MFE: *Els vectors elementals constitueixen l'únic conjunt minimal de generadors conformes d'un poliedre* [8, 17]. La segona propietat que també tenen els VE, i que els fa ser la generalització natural dels MFE a poliedres, és que els conjunts de VE de les possibles subxarxes d'una xarxa metabòlica són subconjunts dels VE de la xarxa original.

Per calcular de manera eficient els VE d'una xarxa, ens cal una definició algebraica d'aquests vectors que permeti desenvolupar algorismes adients. De fet, la definició la dona el mateix mètode de càlcul dels VE. Conceptualment, aquest mètode es basa en la transformació del sistema no homogeni d'inequacions $\mathbf{A}\mathbf{v} \geq \mathbf{b}$ que defineix un poliedre general P de \mathbb{R}^n , en un sistema homogeni d'equacions que es pugui pensar com l'expressió d'un con de fluxos en dimensió superior. És per això que aquest mètode es coneix com el mètode d'homogenització [5, 17].

En un primer pas, amb l'ajut d'una variable auxiliar λ , escrivim el sistema $\mathbf{A}\mathbf{v} \geq \mathbf{b}$ com el sistema homogeni en $n + 1$ variables (\mathbf{v}', λ) donat per

$$(\mathbf{A} \quad -\mathbf{b}) \begin{pmatrix} \mathbf{v}' \\ \lambda \end{pmatrix} \geq 0 \quad \text{amb} \quad \lambda \geq 0. \quad (12)$$

Aquest sistema d'inequacions és l'expressió d'un con general de \mathbb{R}^{n+1} . Clarament, si \mathbf{v} satisfà $A\mathbf{v} \geq \mathbf{b}$, llavors $(\mathbf{v}, 1)$ és solució de (12). Recíprocament, si (\mathbf{v}', λ) , $\lambda > 0$, és solució de (12), llavors $\mathbf{v} = \mathbf{v}'/\lambda$ és solució del sistema no homogeni. En canvi, si $(\mathbf{v}', 0)$ és solució de (12), llavors $\mathbf{v} = \mathbf{v}'$ és solució del sistema homogeni $A\mathbf{v} \geq \mathbf{0}$ que defineix el con de recessió de P . En altres paraules, un conjunt $\{(\mathbf{v}'_k, \lambda_k)\}$ de generadors del con definit per (12) defineix un conjunt de generadors de P . Aquests últims seran acotats quan $\lambda_k > 0$, mentre que seran no acotats quan $\lambda_k = 0$. No tenim garantit, però, que aquests generadors de P siguin generadors conformes, és a dir, que siguin els seus VE.

El segon pas serà, doncs, obtenir els generadors conformes de P i això ho farem aprofitant el fet que els MFE són generadors conformes de cons de fluxos. D'aquesta manera, el càlcul dels VE es reduirà al càlcul dels MFE d'un con de fluxos auxiliar, és a dir, d'un con de la forma donada per (4) però dins d'un espai de dimensió més gran. Per fer-ho, introduïrem un vector \mathbf{y} de variables de folgança, anàlogues a les utilitzades en programació lineal per convertir desigualtats en igualtats, i escrivim (12) com (vg. [5])

$$(A \quad -I \quad -\mathbf{b}) \begin{pmatrix} \mathbf{v}' \\ \mathbf{y} \\ \lambda \end{pmatrix} = \mathbf{0} \quad \text{amb} \quad \mathbf{y} \geq \mathbf{0} \quad \text{i} \quad \lambda \geq 0. \quad (13)$$

Amb aquesta transformació, els MFE del con de fluxos representat per (13) corresponen als VE d'un poliedre P . Si P és un poliedre de fluxos, A i \mathbf{b} vénen donades per (8) i I és una matriu on les primeres $2m$ files són zero, essent m el nombre de metabòlits de la xarxa. En aquest cas, però, una transformació de les desigualtats en (7) anàloga a l'anterior dóna lloc a la següent expressió del con de fluxos auxiliar

$$\begin{pmatrix} S & \mathbf{0} & \mathbf{0} \\ G & -I & -\mathbf{h} \end{pmatrix} \begin{pmatrix} \mathbf{v}' \\ \mathbf{y} \\ \lambda \end{pmatrix} = \mathbf{0} \quad \text{amb} \quad v'_i \geq 0 \quad \forall i \in \text{Irr}, \quad \mathbf{y} \geq \mathbf{0} \quad \text{i} \quad \lambda \geq 0 \quad (14)$$

on, recordem, S és la matriu estequiomètrica de les reaccions i \mathbf{h} és el vector que conté les fites superiors i/o inferiors de les velocitats de reacció especificades per la matriu G [5]. En aquest context, hom parla a vegades de *vectors de fluxos elementals* (VFE).

El fet de poder basar el càlcul dels VE d'un poliedre en l'obtenció dels MFE d'un con de fluxos (de dimensió més gran) no és només una qüestió conceptual, sinó que té una important aplicació pràctica ja que permet aprofitar els algorismes desenvolupats recentment per augmentar l'eficiència del càlcul dels MFE de cons de fluxos de xarxes metabòliques [19]. Tot i així, avui dia l'enumeració dels MFE de models a escala genòmica continua essent impracticable [2, 4] i, per tant, també ho és el càlcul dels VE el nombre dels quals pot arribar a ser més gran que el de MFE. En l'exemple de metabolisme central d'*Escherichia coli* considerat a [5], el nombre de MFE del con de fluxos és 314.241. Si, a més, es considera una fita superior per a la taxa d'ingesta de glucosa i una fita inferior per la velocitat d'una segona reacció lligada al procés de manteniment cel·lular (anomenada ATPmaint), el con de fluxos passa a ser un poliedre de fluxos i el nombre de VFE calculats és 247.947. D'aquests, 27.864 són VFE de suport minimal, és a dir, aquest és el nombre de les subxarxes funcionals amb el mínim nombre de reaccions i que, a més, són compatibles amb les restriccions imposades a les dues velocitats de reacció anteriors. Les limitacions computacionals actuals queden paleses en el recent estudi de 17 xarxes metabòliques fet a [12] amb l'ajut d'eines com *efmtool* per al càlcul de MFE, integrada a MATLAB. El resultat és prou revelador: no es van poder calcular, per falta de memòria, els MFE de cap de les 11 xarxes amb més de 100 reaccions. Totes elles són models del metabolisme de diferents espècies de bacteris (excepte la del glòbul vermell humà) i tenen unes mides que van de 194 a 581 reaccions. Som, doncs, encara lluny de poder obtenir els MFE de xarxes de mida més gran com, per exemple, la del metabolisme humà comentada anteriorment amb les seves 7440 reaccions. D'altra banda, la possible rellevància biològica de nous conceptes introduïts per superar aquestes limitacions de càlcul, com el proposat a [12], ha de ser estudiada en futurs treballs.

Fins aquí un breu repàs de conceptes fonamentals lligats al que es coneix com anàlisi de xarxes amb restriccions. N'he deixat fora aspectes importants, mereixedors cadascun d'ells d'un nou article, com són l'optimització de fluxos [7], els algorismes del càlcul dels MFE [14], o l'aplicació de principis de la termodinàmica per reduir el conjunt inicial de MFE a aquells modes que són factibles des d'un punt de vista termodinàmic [3]. Voldria acabar, però, amb una curiositat. Cap dels articles citats en aquest treball ha estat publicat en una revista de matemàtiques, malgrat el contingut eminentment formal d'algun d'ells (vg. [8]). Això no ha estat de cap manera intencionat i podria estar relacionat amb el fet que, com es comenta en aquesta darrera referència, la descomposició conforme de cons i poliedres sembla que no ha estat estudiada en geometria. Així, de la mateixa manera que al principi

d'aquest treball dèiem que els llibres de biologia de sistemes estan farcits de matemàtiques, podem acabar dient que, cada cop més, revistes de fora de l'àmbit estrictament matemàtic acullen articles amb nous conceptes formals que ajuden a entendre i fer avançar la biologia de sistemes.

Referències

- [1] I.R. Epstein: *Anatol Zhabotinsky (1938-200)*. Nature, vol. **445**, 1053, 2008.
- [2] M.P. Gerstl, D. E. Ruckerbauer, D. Mattanovich, C. Jungreuthmayer, J. Zanghellini: *Metabolomics integrated elementary flux mode analysis in large metabolic networks*. Scientific Reports, vol. **5**, 8930, 2015.
- [3] M.P. Gerstl, C. Jungreuthmayer, S. Müller, J. Zanghellini: *Which sets of elementary flux modes form thermodynamically feasible flux distributions?* The FEBS Journal, vol. **283**, 1782–1794, 2016.
- [4] M.P. Gerstl, S. Müller, G. Regensburger, J. Zanghellini: *Flux tope analysis: studying the coordination of reaction directions in metabolic networks*. Bioinformatics, 1–8, 2018.
- [5] S. Klamt, G. Regensburger, M.P. Gerstl, C. Jungreuthmayer, S. Schuster, R. Mahadevan, J. Zanghellini, S. Müller: *From elementary flux modes to elementary flux vectors: Metabolic pathway analysis with arbitrary linear flux constraints*. PLOS Computational Biology, vol. **13** (4), 1–22, 2017.
- [6] F. Llaneras, J. Picó: *Which metabolic pathways generate and characterize the flux space? A comparison among elementary modes, extreme pathways and minimal generators*. Journal of Biomedicine and Biotechnology, vol. 2010, Article ID 753904, 1–3, 2010.
- [7] T.R. Maarleveld, R.A. Khandelwal, B.G. Olivier, B. Teusink, F.J. Bruggeman: *Basic concepts and principles of stoichiometric modeling of metabolic networks*. Biotechnology Journal, vol. **8**, 997–1008, 2013.
- [8] S. Müller, G. Regensburger: *Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis*. Frontiers in Genetics, vol. **7**, article 90, 1–11, 2016.
- [9] J.D. Murray: *Mathematical Biology: I. An Introduction*, 3a edició, Springer, 2007.

- [10] J.A. Papin, N.D. Price, S.J. Wiback, D.A. Feli, B.O. Palsson: *Metabolic pathways in the post-genome era*. TRENDS in Biochemical Sciences, vol. **28**, 250–258, 2003.
- [11] R.T. Rockafellar: *The elementary vectors of a subspace of RN* . En la monografia “Combinatorial mathematics and its applications”, pàg. 104–127, University of North Carolina Press, 1969. Reeditat el 2011. ISBN: 978-0-8078-7820-0.
- [12] A. Röhl, A. Bockmayr: *Finding MEMo: minimum sets of elementary flux modes*. Journal of Mathematical Biology vol. **79**, 1749–1777, 2019.
- [13] S. Schuster, C. Hilgetag: *On elementary flux modes in biochemical reactions systems at steady state*. Journal of Biological Systems, vol. **2**, 165–182, 2002.
- [14] S. Schuster, C. Hilgetag, J.H.Woods, D.A. Fell: *Reaction routes in biochemical reaction systems: Algebraic properties, validated calculation procedure and example from nucleotide metabolism*. Journal of Mathematical Biology, vol. **45**, 153–181, 2002.
- [15] E.D. Sontag: *Lecture notes on mathematical systems biology*, 2015. Llibre periòdicament revisat i actualitzat disponible a Internet a l’adreça <https://www.math.rutgers.edu/docman-lister/math-main/academics/undergraduate/interdisciplinary-majors/biomathematics-interdisciplinary-major/2198-lecture-notes-on-mathematical-systems-biology/file>
- [16] I. Thiele i molts altres: *A community-driven global reconstruction of human metabolism*. Nature Biotechnology, vol. **31** (5), 419–425, 2013
- [17] R. Urbanczik: *Enumerating constrained elementary flux vectors of metabolic networks*. IET Systems Biology vol. **1**, 274–279, 2007.
- [18] R. Urbanczik, C. Wagner: *Functional stoichiometric analysis of metabolic networks*. Bioinformatics, vol. **21**, 4176–4180, 2005.
- [19] J. Zanghellini, D.E. Ruckerbauer, M. Hanscho, C. Jungreuthmayer: *Elementary flux modes in a nutshell: Properties, calculation and applications*. Biotechnology Journal, vol. **8**, 1009–1016, 2013.



Departament d'informàtica,
matemàtica aplicada i estadística
Universitat de Girona
joan.saldana@udg.edu

Publicat el 2 de desembre de 2019