

Elementary growth modes/vectors and minimal autocatalytic sets for kinetic/constraint-based models of cellular growth

Stefan Müller

February 24, 2021

Abstract

Elementary vectors are fundamental objects in polyhedral geometry. In metabolic pathway analysis, elementary vectors range from elementary flux modes (of the flux cone) and elementary flux vectors (of a flux polyhedron) via elementary conversion modes (of the conversion cone) to minimal cut sets (of a dual polyhedron) in computational strain design.

To better understand cellular phenotypes with optimal (or suboptimal) growth rate, we introduce and analyze classes of elementary vectors for models of cellular growth. *Growth modes* (GMs) only depend on stoichiometry, but not on growth rate or concentrations; they are elements of the growth cone. *Elementary growth modes* (EGMs) are conformally non-decomposable GMs; unlike elementary flux modes, they are not support-minimal, in general. Most importantly, every GM can be written as a conformal sum of EGMs. *Growth vectors* (GVs) and *elementary growth vectors* (EGVs) also depend on growth rate, concentrations, and linear constraints; they are elements of a growth polyhedron. Again, every GV can be written as a conformal sum of EGVs. To relate the new concepts to other branches of theory, we define *autocatalytic* GMs and the corresponding (*minimal autocatalytic sets* of reactions).

As a case study, we consider whole cell models (simple kinetic models of self-fabrication). First, we use EGMs to derive an upper bound for growth rate that only depends on enzyme kinetics. Next, we study growth rate maximization (via control parameters for ribosome kinetics). In particular, we analyze *growth states* (GSs) and *elementary growth states* (EGSs) as introduced in [de Groot et al, 2020]. Unlike EGMs, EGSs depend on (metabolite) concentrations and growth rate. Most importantly, (i) we show that EGSs are support-minimal, (ii) we give a simple proof for the fact that maximum growth rate is attained at an EGS, and (iii) we show that, at every optimal EGS, the ribosome capacity constraint is active. Finally, we determine the dependence of EGSs on growth rate, and we study the relation between EGSs and minimal autocatalytic sets, EGMs, and elementary flux modes. Along the way, we point out (and resolve) mathematical issues in [de Groot et al, 2020].

Keywords: cellular self-fabrication, minimal pathways, growth rate, polyhedral geometry, elementary vectors, conformal generators

S. Müller (st.mueller@univie.ac.at), Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

1 Introduction

Cell proliferation involves cell growth and cell division. In more abstract terms, cellular *self-fabrication* involves cellular *self-maintenance*, including processes such as signaling, transport, metabolism, and gene regulation, and cellular *self-replication*, including macromolecular synthesis and cell cycle. Indeed, during one cycle, a cell self-fabricates all its constituents (metabolites, enzymes, lipids, DNA, ...); it grows. By cell division, this leads to growth on the population level (microbial growth) or the tissue level (multicellular development).

The processes contributing to cellular growth can be summarized in a stoichiometric matrix N with rows corresponding to constituents (molecular species) and columns corresponding to processes (chemical reactions). In the deterministic setting, one has the dynamical system $\frac{dX}{dt} = NV(X)$ for the copy numbers X of molecules (extensive variables), as determined by N and the reaction rates $V(X)$. After introducing concentrations x (intensive variables), one obtains $\frac{dx}{dt} = Nv(x) - \mu x$, cf. [12]. The second term represents “dilution by growth”, where growth rate μ is given (as a function of x) by a linear constraint arising from dry weight or cell volume. The dynamical systems apply to individual cells or, as averages, to cell populations. Steady state for the intensive variables, that is, $x(t) = \text{const}$ or $Nv(x) = \mu x$, corresponds to balanced growth for the extensive variables, that is, $X(t) = X(0)e^{\mu t}$. On the population level, one has $\frac{dn}{dt} = \mu n$ for the copy number n of cells, cf. [9, 22].

Particular models of cellular growth as well as model classes can be categorized by several dimensions: (i) the set of cellular processes modeled explicitly, (ii) the model structure – from detailed (elementary reaction steps as one extreme) to coarse-grained (one reaction for every process as another extreme), (iii) dynamic vs. steady-state, and (iv) kinetic (with nonlinear dependence $v(x)$ of reaction rates on concentrations) vs. constraint-based (with linear constraints in x and v). For a recent review on mathematical models of cellular growth, see e.g. [6].

Traditional growth models often involve an approximative “biomass reaction”, which specifies macromolecular synthesis in terms of precursors. In terms of categories (i) and (ii), such models consider metabolism in detail and macromolecular synthesis as a coarse-grained process (the biomass reaction). In terms of (iii) and (iv), they often assume steady state and are constraint-based. In detail, concentrations x are fixed, and steady-state reaction rates (fluxes) v are considered as independent variables. The stoichiometric matrix has rows corresponding to metabolites and columns corresponding to metabolic reactions and the biomass reaction. Further, there are irreversibility constraints. Altogether, one considers $Nv = 0$ and $v_{\mathcal{I}} \geq 0$, defining the *flux cone* (mathematically, a linear subspace with nonnegativity constraints, that is, an s-cone [19]).

Metabolic pathway analysis aims to identify biologically meaningful routes in a metabolic network, in particular, minimal routes. The (biologically and mathematically) fundamental abstractions of minimal metabolic pathways are *elementary flux modes* (EFMs) [24, 25]. Formally, EFMs are the support-minimal vectors of the flux cone. Most importantly, every element of the flux cone can be written as a conormal sum of EFMs (a sum without cancellations) [19]. In convex analysis, *elementary vectors* (of linear subspaces) have been introduced 25 years before their use in metabolic pathway analysis [23]. Recently, the concept

of elementary vectors and the corresponding conformal sum theorems have been extended from linear subspaces to polyhedral cones and polyhedra [19]. Indeed, elementary vectors form *unique* sets of *conformal generators* for linear subspaces, polyhedral cones, and polyhedra.

Given the flux cone, flux balance analysis (FBA) adds linear constraints (e.g. flux bounds) and maximizes (biomass) flux [8]. Altogether, FBA considers a linear program on $Nv = 0$, $v_{\mathcal{I}} \geq 0$, and e.g. $v^{\text{lb}} \leq v \leq v^{\text{ub}}$, defining a *flux polyhedron*. Its elementary vectors are called *elementary flux vectors* (EFVs) [27] and are not support-minimal, in general. As a consequence, maximum flux is attained at EFVs. In other words, EFVs are the missing link between EFM analysis, describing all (stoichiometrically) *feasible* solutions, and FBA, identifying *optimal* solutions [16]. Interestingly, in kinetic models (without flux bounds and with one enzyme capacity constraint), maximum flux is attained at EFMs [20, 29].

Finally, two more classes of elementary vectors have been introduced in metabolic pathway analysis: *elementary conversion modes* (ECMs) of the conversion cone (the projection of the flux cone onto the exchange reactions) [28] and *minimal cut sets* (MCSs) of a dual polyhedron (minimal sets of gene knockouts) [2]. By using ECMs, EFMs need not be enumerated (which is computationally infeasible for genome-scale metabolic models [30]) in order to determine all minimal conversions of substrates into products, and by using MCSs, EFMs need not be enumerated in order to determine all minimal interventions in computational strain design.

Recently, more refined models of cellular growth have been studied, where individual synthesis reactions for macromolecules replace the traditional biomass reaction. In the category of constraint-based models, prominent examples are resource balance analysis (RBA) [11] and metabolism and macromolecular expression (ME) models [17]. In particular, RBA considers linear *capacity constraints*: given the concentration of a catalyst, kinetics implies an upper bound for the flux of the corresponding catalytic reaction. (E.g. for the flux v of a metabolic reaction catalyzed by enzyme E , one has $v \leq k^{\text{cat}}x_E$.) In addition, ME models consider the *genotype-phenotype map* of macromolecular synthesis in detail (and the resulting inequality constraints for the fluxes involved). Finally, in the category of kinetic models, whole cell models (also known as self-replicator or self-fabrication models) aim to give a complete picture of cellular processes, however, on a coarse-grained level [5, 18]. In fact, such models often are *hybrid* (kinetic and constraint-based), involving control parameters for optimization and the corresponding constraints.

In this work, we introduce and analyze classes of elementary vectors for *general* models of cellular growth: *Elementary growth modes* (EGMs) are elementary vectors of the *growth cone*; they only depend on stoichiometry and hence apply to general growth models (kinetic or constraint-based). *Elementary growth vectors* are elementary vectors of a *growth polyhedron*; they also depend on growth rate, concentrations, and fluxes and hence apply to constraint-based models. Unlike EFMs (but like ECMs), EGMs are elementary vectors of a general polyhedral cone, and like EFVs, EGVs are elementary vectors of a polyhedron. To demonstrate the relevance of the new concepts, we relate them to the theory of *autocatalytic sets*.

As a case study, we consider whole cell models (simple kinetic models of self-fabrication). To characterize cellular phenotypes with maximum growth rate, we use *elementary growth states* (EGSs) as introduced in [5]. Unlike EGMs, EGSs are not defined for general growth models and also depend on concentrations and growth rate. Still, every EGS can be written as a conformal sum of EGMs. We obtain several new results on EGSs (and hence on growth rate maximization) and simple proofs of existing results.

Outline and main results. In **Section 2**, we provide new results on elementary vectors in polyhedral geometry, cf. Propositions 1 and 2. In **Section 3**, we introduce *elementary growth modes* and *vectors* for kinetic and constraint-based models of cellular growth, we provide the corresponding conformal sum theorems, cf. Theorems 9, 15 and Corollary 10, we define *minimal autocatalytic sets* of reactions, and we illustrate all new concepts in models of a minimal network, cf. Examples 13 and 16. In **Section 4**, we consider simple kinetic models of self-fabrication and related constraint-based models, we instantiate general definitions and results, and we derive an upper bound for growth rate that only depends on enzyme kinetics (and the set of active enzymes), cf. Theorem 21. In **Section 5**, we study growth rate maximization (via control parameters for ribosome kinetics). In particular, we analyze *elementary growth states* (EGSs) as introduced in [5]. Most importantly, we show that (i) EGSs are support-minimal, (ii) maximum growth rate is attained at an EGS, and (iii), at every optimal EGS, the ribosome capacity constraint is active, cf. Theorem 28, Proposition 29, Theorem 30, and Corollary 31. Moreover, we find that (E)GSs correspond to (minimal) autocatalytic sets, cf. Proposition 33, and we illustrate the relation between EGSs and EGMs in models of small networks, cf. Examples 34, 35, and 36. Under the assumption of proportional synthesis, we determine the dependence of EGSs on growth rate, cf. Theorem 37, and we study the relation between EGSs and elementary flux modes, cf. Theorems 38 and 39. Finally, in **Section 6**, we discuss the terminology of elementary vectors in metabolic pathway analysis.

In Supplement A, we summarize definitions and results on elementary vectors in polyhedral geometry. In Supplement B, we provide a minimal derivation of the dynamic model of cellular growth. Finally, in Supplement C, we list mathematical issues in [5].

Mathematical notation. We denote the positive real numbers by $\mathbb{R}_{>}$ and the nonnegative real numbers by \mathbb{R}_{\geq} . For $x \in \mathbb{R}^n$, we write $x > 0$ if $x \in \mathbb{R}_{>}^n$, $x \geq 0$ if $x \in \mathbb{R}_{\geq}^n$, and we denote its *support* by $\text{supp}(x) = \{i \mid x_i \neq 0\}$. Recall that a nonzero vector $x \in X \subseteq \mathbb{R}^n$ is *support-minimal* if, for all nonzero $x' \in X$, $\text{supp}(x') \subseteq \text{supp}(x)$ implies $\text{supp}(x') = \text{supp}(x)$. For $x \in \mathbb{R}^n$, we define its *sign vector* $\text{sign}(x) \in \{-, 0, +\}^n$ by applying the sign function component-wise, that is, $\text{sign}(x)_i = \text{sign}(x_i)$ for $i = 1, \dots, n$. The relations $0 < -$ and $0 < +$ on $\{-, 0, +\}^n$ induce a partial order on $\{-, 0, +\}^n$: for $X, Y \in \{-, 0, +\}^n$, we write $X \leq Y$ if the inequality holds component-wise. For $x, y \in \mathbb{R}^n$, we denote the component-wise product by $x \circ y \in \mathbb{R}^n$, that is, $(x \circ y)_i = x_i y_i$. For $n \in \mathbb{N}$, we write $[n] = \{1, \dots, n\}$. For $x \in \mathbb{R}^n$ and index set $I \subset [n]$, we write $x_I \in \mathbb{R}^I$ for the corresponding subvector. For $A \in \mathbb{R}^{m \times n}$ and index sets $I \subset [m]$, $J \subset [n]$, we write $A_{I,J} \in \mathbb{R}^{I \times J}$ for the corresponding submatrix; if $I = [m]$, we just write $A_{*,J} \in \mathbb{R}^{m \times J}$.

2 Elementary vectors

For the objects of polyhedral geometry (subspaces, cones, polyhedra), there is no *unique* minimal set of generators, in general. However, *elementary vectors* (EVs) form unique sets of *conformal* generators [19, Section 3.4]. For linear subspaces and s-cones (arising from linear subspaces and nonnegativity constraints), elementary vectors are the support-minimal (SM) vectors; for general cones, they are the conformally non-decomposable (cND) vectors; and for general polyhedra, they are the convex-conformally non-decomposable (ccND) vectors plus the cND vectors of the recession cone [19].

In Supplement A, we summarize basic definitions and results for s-cones, general polyhedral cones, and polyhedra; cf. Theorems 43, 44, and 45. Below, we provide new results on special polyhedra (arising from affine subspaces and non-negativity constraints). In fact, Propositions 1 and 2 will be crucial in the study of elementary growth states; we state them here to complete the mathematical preliminaries before introducing growth models.

2.1 Special polyhedra

Recall the definitions of linear and affine subspaces: In a linear subspace $S \subseteq \mathbb{R}^n$, we have that $x, y \in S$ and $\lambda, \mu \in \mathbb{R}$ imply $\lambda x + \mu y \in S$. In an affine subspace $T \subseteq \mathbb{R}^n$, we have that $x, y \in T$ and $\lambda \in \mathbb{R}$ imply $\lambda x + (1 - \lambda)y \in T$. We call an affine subspace *genuine* if it is not also a linear subspace.

Proposition 1. *Let $T \subseteq \mathbb{R}^n$ be a genuine affine subspace. If two SM vectors have the same support, then they are identical.*

Proof. Let $e, e' \in T$ be distinct SM vectors that have the same support. Since T is not a linear subspace, they are not scalar multiples. Consider $x = \lambda e + (1 - \lambda)e' \in T$ with $\lambda \in \mathbb{R}$. Indeed, choose λ such that $x_i = \lambda e_i + (1 - \lambda)e'_i = 0$ for some $i \in \text{supp}(e)$ (with $e_i \neq e'_i$). That is, $\text{supp}(x) \subset \text{supp}(e)$, and e is not SM, a contradiction. \square

The corresponding result for linear subspaces/s-cones is Corollary 42 in Supplement A.1.

Proposition 2. *Let $T \subseteq \mathbb{R}^n$ be a genuine affine subspace and $O \subseteq \mathbb{R}^n$ be a closed orthant. The vertices of the polyhedron $P = T \cap O$ are its SM vectors.*

Proof. It suffices to show the statement in the notationally simplest case $O = \mathbb{R}_{\geq}^n$. (Then, $x \in O$ is equivalent to $x \geq 0$.)

(VE \Rightarrow SM). Assume $x \in P$ is not SM, that is, there exists $x^1 \in P$ with $\text{supp}(x^1) \subset \text{supp}(x)$, and consider $x^2 = \lambda x + (1 - \lambda)x^1 \in T$. Clearly, there exists $\lambda > 1$ such that $x^2 \geq 0$ (that is, $x^2 \in O$) and hence $x^2 \in P$. Now, $x = (1 - \frac{1}{\lambda})x^1 + \frac{1}{\lambda}x^2 = (1 - \lambda')x^1 + \lambda'x^2$ with $0 < \lambda' < 1$, that is, x is not a vertex.

(SM \Rightarrow VE). Assume $x \in P$ is not a vertex, that is, there exist $x^1, x^2 \in P$ with $x^1 \neq x^2$ and $0 < \lambda < 1$ such that $x = \lambda x^1 + (1 - \lambda)x^2$. If $\text{supp}(x^1) =$

$\text{supp}(x^2) = \text{supp}(x)$, consider $x' = \lambda'x^1 + (1 - \lambda')x^2 \in T$. Clearly, there exists a largest $\lambda' > 1$ such that $x' \geq 0$ (that is, $x' \in O$) and hence $x' \in P$. For this λ , $\text{supp}(x') \subset \text{supp}(x)$, that is, x is not SM. \square

3 Growth models

Notation. We denote fundamental objects and quantities as follows:

Mol	set of molecular species	
Rxn	set of chemical reactions	
$N \in \mathbb{R}^{\text{Mol} \times \text{Rxn}}$	stoichiometric matrix	(unit: 1)
$\rho \in \mathbb{R}_{>}^{\text{Mol}}$	molar masses	(unit: g/mol)
$x \in \mathbb{R}_{\geq}^{\text{Mol}}$	concentrations	(unit: mol/g)
$v(x) \in \mathbb{R}^{\text{Rxn}}$	reaction rates	(unit: mol/g/s)

In Supplement B, we provide a minimal derivation of the dynamic model of cellular growth.

A *dynamic* growth model is given by a dynamical system (involving chemical reactions and growth) and a linear constraint representing dry weight (or, alternatively, cell volume):

	$\frac{dx}{dt} = Nv(x) - \mu x$	(1a)
and	$\rho^T x = 1.$	(1b)

In particular, growth rate is given by

$$\mu = \rho^T N v. \quad (2)$$

Note. At steady state, system (1a) and (2) is equivalent to system (1). However, if (2) is used as a definition of μ in (1a), then the mathematical treatment often becomes less transparent.

Remark 3 (Conservation laws). In a growth model, there can be no conservation laws. In mathematical terms, $\ker N^T \cap \mathbb{R}_{\geq}^{\text{Mol}} = \{0\}$.

To see this, assume $c^T N = 0$ with $0 \neq c \geq 0$, for example, assume $c_1 = c_2 = 1$ and $c_i = 0$, otherwise. Then, $\frac{d(c^T x)}{dt} = \frac{d(x_1 + x_2)}{dt} = -\mu(x_1 + x_2) \leq 0$, and $\mu > 0$ implies $x_1 = x_2 = 0$ at steady state.

Remark 4 (Dependent concentrations). Even if there are no conservation relations, there can be dependent concentrations. (However, biochemically, this is a degenerate case.) In mathematical terms, $\ker N^T \neq \{0\}$.

To illustrate this, assume $c^T N = 0$ with $0 \neq c$, for example, assume $c_1 = 1$, $c_2 = -1$, and $c_i = 0$, otherwise. Then, $\frac{d(c^T x)}{dt} = \frac{d(x_1 - x_2)}{dt} = -\mu(x_1 - x_2)$, and $\mu > 0$ implies $x_1 = x_2$ at steady state.

In the following, we assume that all dependent variables have been eliminated. In particular, $\ker N^T = \{0\}$, that is, $\text{im } N = \mathbb{R}^{\text{Mol}}$.

Elementary vectors

At steady state,

$$Nv = \mu x \geq 0.$$

Further, $v_{\mathcal{I}} \geq 0$, where $\mathcal{I} \subseteq \text{Rxn}$ denotes the set of irreversible reactions.

Note. In general, all reactions are reversible, but in a given setting, reactions may have a given direction, as determined by thermodynamics.

Definition 5. *Growth modes* (GMs) for the dynamic growth model (1) are elements of the *growth cone*

$$C_g = \{v \in \mathbb{R}^{\text{Rxn}} \mid Nv \geq 0, v_{\mathcal{I}} \geq 0\}.$$

A GM $v \in C_g$ has an *associated* growth rate $\mu(v) = \rho^T Nv \geq 0$ and, if $\mu(v) > 0$, an *associated* concentration vector $x(v) = Nv/\mu(v) \in \mathbb{R}_{\geq}^{\text{Mol}}$.

Elementary growth modes (EGMs) are conformally non-decomposable GMs.

Note. GMs only depend on stoichiometry; in particular, they do not depend on concentrations or growth rate. (But they have an associated growth rate and an associated concentration vector.)

Note. The growth cone C_g is a general polyhedral cone (not an s-cone like the flux cone); its elementary vectors are conformally non-decomposable (but not support-minimal, in general); see e.g. Example 13.

Remark 6 (Computation). For the computation of EGMs, the system of inequalities defining C_g is transformed to a system of equalities by the introduction of slack variables, and SM vectors of the resulting higher-dimensional s-cone are computed by (variants of) the double description method, see e.g. [16].

The definition of GMs immediately implies the scale invariance of associated concentrations.

Proposition 7. For a GM $v \in C_g$ with associated concentration $x(v)$ and $\lambda > 0$, it holds that $x(\lambda v) = x(v)$.

Further, the assumption $\ker N^T \cap \mathbb{R}_{\geq}^{\text{Mol}} = \{0\}$ (no conservation laws) allows to characterize zero growth.

Proposition 8. For a GM $v \in C_g$, $Nv = 0$ is equivalent to $\mu(v) = 0$.

GMs v with $\mu(v) = 0$ (“zero growth modes”) are flux modes (FMs), that is, elements of the flux cone

$$C_f = \{v \in \mathbb{R}^{\text{Rxn}} \mid Nv = 0, v_{\mathcal{I}} \geq 0\}.$$

(In general, this is not the flux cone arising from traditional growth models, involving an approximate “biomass reaction”.) EGMs v with $\mu(v) = 0$ are elementary flux modes (EFMs), that is, SM elements of C_f . They are not SM elements of C_g , in general; see e.g. Example 13.

Finally, we apply the general theory of elementary vectors and state the main result of this section.

Theorem 9. *Every nonzero GM is a conformal sum of EGMs.*

Proof. By Theorem 44 in Supplement A.2: The growth cone C_g is a general polyhedral cone, and its elementary vectors are the conformally non-decomposable vectors, that is, the EGMs. \square

In fact, we can be more specific.

Corollary 10. *Let v be a nonzero GM with associated growth rate $\mu(v) =: \mu$. Then, there exist (possibly empty) finite sets E_0 and E_μ of EGMs with associated growth rates 0 and μ , respectively, such that*

$$v = \sum_{e \in E_0} e + \sum_{e \in E_\mu} \lambda_e e \quad \text{with } \text{sign}(e) \leq \text{sign}(v),$$

$\lambda_e \geq 0$, and $\sum_{e \in E_\mu} \lambda_e = 1$. Moreover, if $\mu > 0$, then $x(v) = \sum_{e \in E_\mu} \lambda_e x(e)$.

Proof. By Theorem 9, there exist (possibly empty) finite sets E_0 and $E_{>}$ of EGMs (with associated growth rates 0 and >0 , respectively) such that

$$v = \sum_{e \in E_0} e + \sum_{e \in E_{>}} e \quad \text{with } \text{sign}(e) \leq \text{sign}(v).$$

In particular, $\mu = \mu(v) = \sum_{e \in E_{>}} \mu(e)$. Now,

$$\sum_{e \in E_{>}} e = \sum_{e \in E_{>}} \frac{\mu(e)}{\mu} \frac{\mu}{\mu(e)} e = \sum_{e' \in E_\mu} \lambda_{e'} e',$$

where $e' = \frac{\mu}{\mu(e)} e$ with $\mu(e') = \mu$ and $\lambda_{e'} = \frac{\mu(e)}{\mu}$ with $\sum_{e' \in E_\mu} \lambda_{e'} = 1$.

Finally, if $\mu > 0$, then $x(v) = Nv/\mu = \sum_{e \in E_\mu} \lambda_e Ne/\mu = \sum_{e \in E_\mu} \lambda_e x(e)$. \square

Note. In Corollary 10, we actually fix growth rate which turns the growth cone into a polyhedron. Hence, the result is also an instance of Theorem 45 in Supplement A.3.

Autocatalysis

Cellular growth is autocatalytic in the sense that the cell fabricates itself (thereby exchanging substrates/products with the environment). One needs to distinguish this notion of “network autocatalysis” from “autocatalytic subnetworks” (technically: autocatalytic cycles/cores) [3, 4].

A main requirement for network autocatalysis is the existence of a growth mode where *all* species involved in active reactions have nonzero associated concentrations.

Definition 11. A GM $v \in C_g$ is *strict* if, for every $r \in \text{supp}(v) \subseteq \text{Rxn}$ and $s \in \text{Mol}$, $N_{sr} \neq 0$ implies $(Nv)_s > 0$.

Clearly, in the overall reaction corresponding to a strict GM, all species appear on the product side, in particular, with a larger stoichiometric coefficient than on the educt side. If a species appears also on the educt side, then it is *formally autocatalytic* (cf. [1]), and one may call the GM itself autocatalytic. In fact, there are several competing notions of autocatalytic species and subnetworks like formally/exclusively autocatalytic and autocatalytic cycles/cores (cf. [1, 3, 4]). Before we state possible definitions for network autocatalysis, we distinguish two modeling approaches.

- **Detailed models** (without individual catalytic reactions)

In this approach, catalysis occurs on the level of (small) subnetworks. In particular, individual reactions are not catalytic. For example, a simple catalytic mechanism (involving enzyme E, substrate S, and product P) is given by $E + S \leftrightarrow ES \leftrightarrow EP \leftrightarrow E + P$.

- **Coarse-grained models** (with individual catalytic reactions)

In this approach, catalysis occurs on the level of individual reactions. For example, the catalytic mechanism above is written as $E + S \leftrightarrow E + P$ or $S \xrightarrow{E} P$. Due to coarse-graining, catalysis cannot be identified from the stoichiometric matrix. Hence, for every catalytic reaction, the corresponding catalyst is specified explicitly.

For **detailed models**, one may call a GM $v \in C_g$ *autocatalytic* if (i) it is strict and (ii) it contains an autocatalytic species or subnetwork. For example, one may require the existence of a formally/exclusively autocatalytic species or a catalytic/autocatalytic subnetwork. Formal definitions and their comparison are subject of future work. In this work, all examples (and the case study in Sections 4 and 5) are specified as coarse-grained models. For **coarse-grained models**, we give a formal definition of network autocatalysis.

Definition 12. For coarse-grained models, a GM $v \in C_g$ is *basically catalytic* (BC) if there is a catalytic reaction $r \in \text{supp}(v)$. Further, a GM $v \in C_g$ is *catalytically closed* (CC) if, for every catalytic reaction $r \in \text{supp}(v)$, it holds that $(Nv)_c > 0$ for the corresponding catalyst $c \in \text{Mol}$. Finally, a GM $v \in C_g$ is *autocatalytic* (AC) if it is strict, BC, and CC.

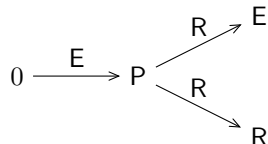
A subset of reactions $S \subseteq \text{Rxn}$ is *autocatalytic* (AC) if there exists an autocatalytic GM $v \in C_g$ with $S = \text{supp}(v)$. A nonempty subset of reactions is *minimally autocatalytic* (MAC) if it is AC and inclusion-minimal.

Note. A closure condition is also crucial in classical definitions of “reflexive autocatalysis” [13, 15, 26] and “chemical organizations” [7, 10, 14]. (For detailed models, a closure condition is not required. In that approach, closure is implied by strictness.)

Note. One may define an *elementary* autocatalytic GM as an autocatalytic GM with minimal support. However, the definition of autocatalysis goes beyond polyhedral geometry, and hence the term *elementary (vector)* is not appropriate. See also the discussion of terminology in Section 6.

Clearly, network autocatalysis as in Definition 12 implies formal autocatalysis for all catalytic species.

Example 13. Consider the following “minimal network”, involving precursor P, “enzyme” E, and ribosome R:



“Metabolism” (the production of the precursor P) is catalyzed by the enzyme E, and the synthesis of E and R is catalyzed by the ribosome R. The corresponding stoichiometric matrix amounts to

$$N = \begin{array}{c} \text{P} \\ \text{E} \\ \text{R} \end{array} \begin{array}{c} \xrightarrow{\text{E}} \quad \rightarrow \text{E} \quad \rightarrow \text{R} \\ \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array} .$$

Further, $\rho = (\rho_{\text{P}}, \rho_{\text{E}}, \rho_{\text{R}})^T = \rho_{\text{P}} (1, 1, 1)^T$, by mass conservation. (Normalized) EGMs with associated growth rates and mass fractions are given by

$$\begin{aligned}
 e^1 &= u \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & e^2 &= u \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, & e^3 &= u \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \\
 \mu(e^1) &= u \rho_{\text{P}}, & \mu(e^2) &= u \rho_{\text{P}}, & \mu(e^3) &= u \rho_{\text{P}}, \\
 \rho \circ x(e^1) &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & \rho \circ x(e^2) &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, & \rho \circ x(e^3) &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.
 \end{aligned}$$

To obtain correct units, we introduced the factor $u = 1 \text{ mol/g/s}$.

In a (biologically meaningful) kinetic model of the minimal network, all steady-state reaction rates (fluxes) and all (associated) concentrations are nonzero. However, in all EGMs, at least one flux is zero, and only one concentration is nonzero. By definition, EGMs only depend on stoichiometry and hence do not reflect (constraints arising from) kinetics. Still, every flux vector of a kinetic model can be written as a conformal sum of EGMs. See also Examples 16 and 34, where the minimal network is studied in the contexts of constraint-based and kinetic models, respectively.

All EGMs involve catalytic reactions, and hence they are BC. However, for all EGMs, at least one active catalyst has zero (associated) concentration, and hence no EGM is CC. Still, every convex combination $v = \lambda_1 e^1 + \lambda_2 e^2 + \lambda_3 e^3$ of EGMs with $\lambda_1 \geq 0$, $\lambda_2, \lambda_3 > 0$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$ (and hence $\mu(v) = u \rho_{\text{P}}$) has nonzero catalyst concentrations, and hence v is CC. If further $\lambda_1 > 0$, then also the precursor concentration is nonzero, that is, v is strict and ultimately AC. Clearly, $S = \text{supp}(v) = \text{Rxn}$ (the set of all reactions) is the unique (M)AC set of reactions.

3.1 Constraint-based models

For many systems, kinetic models are not yet available, and constraint-based models are used. Steady-state reaction rates (fluxes) v are considered as independent variables, that is, the non-linear dependence of the kinetics on the concentrations x is neglected. Most importantly, catalytic processes (in the kinetic model) imply linear capacity constraints for x and v , and additional constraints can be formulated for processes that are not catalytic (in the given model), e.g. lower bounds for concentrations or fluxes. Most compactly, the linear constraints can be written as

$$Ax + Bv \geq b$$

with $A \in \mathbb{R}^{m \times \text{Mol}}$, $B \in \mathbb{R}^{m \times \text{Rxn}}$, and $b \in \mathbb{R}^m$.

Altogether, *constraint-based* growth models involve steady-state (with irreversibility), dry weight, and additional linear constraints,

$$\begin{aligned} Nv = \mu x \geq 0 \quad (\text{with } v_{\mathcal{I}} \geq 0), \\ \rho^T x = 1, \\ Ax + Bv \geq b. \end{aligned} \tag{3}$$

Note. For given growth rate $\mu \geq 0$, one may consider the polyhedron

$$P(\mu) = \left\{ (x, v) \in \mathbb{R}^{\text{Mol}} \times \mathbb{R}^{\text{Rxn}} \mid \begin{aligned} Nv = \mu x, v_{\mathcal{I}} \geq 0, \\ \rho^T x = 1, \\ Ax + Bv \geq b \end{aligned} \right\}.$$

However, for $\mu > 0$ and given fluxes v , the concentration vector $x = Nv/\mu$ is not an independent variable.

Elementary vectors

Definition 14. Let $\mu > 0$. *Growth vectors* (GVs) for the constraint-based growth model (3) are elements of the *growth polyhedron*

$$P_g(\mu) = \left\{ v \in \mathbb{R}^{\text{Rxn}} \mid \begin{aligned} Nv \geq 0, v_{\mathcal{I}} \geq 0, \\ \rho^T Nv = \mu, \\ (AN + \mu B)v \geq \mu b \end{aligned} \right\}.$$

A GV $v \in P_g(\mu)$ has an *associated* concentration vector $x(v) = Nv/\mu \in \mathbb{R}_{\geq}^{\text{Mol}}$.

Elementary growth vectors (EGVs) are convex-conformally non-decomposable GVs and conformally non-decomposable elements of the recession cone

$$R = \left\{ v \in \mathbb{R}^{\text{Rxn}} \mid Nv = 0, v_{\mathcal{I}} \geq 0, Bv \geq 0 \right\}.$$

Again, we can apply the general theory of elementary vectors and state our main result.

Theorem 15. *Let v be a GV for growth rate $\mu > 0$. Then, there exist finite sets $E_0 \subseteq R$ and $E_\mu \subseteq P_g(\mu)$ of EGVs such that*

$$v = \sum_{e \in E_0} e + \sum_{e \in E_\mu} \lambda_e e \quad \text{with } \text{sign}(e) \leq \text{sign}(v),$$

$\lambda_e \geq 0$, and $\sum_{e \in E_\mu} \lambda_e = 1$. Moreover, $x(v) = \sum_{e \in E_\mu} \lambda_e x(e)$.

Proof. By Theorem 45 in Supplement A.3: The growth polyhedron $P_g(\mu)$ is a general polyhedron, and its elementary vectors are the convex-conformally non-decomposable vectors and the conformally non-decomposable vectors of its recession cone. \square

Note. Every GV is also a GM, and hence every nonzero GV is also a conformal sum of EGMs.

Example 16. Consider the “minimal network” introduced in Example 13 together with the kinetics

$$\begin{aligned} v_P &:= v_{0 \rightarrow P}(x) = \kappa(x_P) x_E, \\ v_E &:= v_{P \rightarrow E}(x) = \tau_E(x_P) \alpha_E(x_P) x_R, \\ v_R &:= v_{P \rightarrow R}(x) = \tau_R(x_P) \alpha_R(x_P) x_R, \end{aligned}$$

where $\alpha_E + \alpha_R \leq 1$ due to limited ribosome capacity. The reaction rates are given by particular kinetic functions (depending on the precursor concentration x_P) times the concentrations of the catalyzing molecules (enzyme and ribosome concentrations, x_E and x_R , respectively).

Now, $\kappa(x_P) \leq k^{\text{cat}}$ and $\tau_E(x_P) \leq k_E^{\text{tl}}$, $\tau_R(x_P) \leq k_R^{\text{tl}}$, (plus $\alpha_E + \alpha_R \leq 1$) imply the capacity constraints

$$\begin{aligned} \frac{v_P}{k^{\text{cat}}} &\leq x_E, \\ \frac{v_E}{k_E^{\text{tl}}} + \frac{v_R}{k_R^{\text{tl}}} &\leq x_R. \end{aligned}$$

At steady state, $Nv = \mu x$, that is,

$$\begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} v_P \\ v_E \\ v_R \end{pmatrix} = \mu \begin{pmatrix} x_P \\ x_E \\ x_R \end{pmatrix}.$$

In particular, $x_E = v_E/\mu$, $x_R = v_R/\mu$, and the capacity constraints can be rewritten in terms of fluxes,

$$\begin{aligned} -\frac{\mu}{k^{\text{cat}}} v_P + v_E &\geq 0, \\ -\frac{\mu}{k_E^{\text{tl}}} v_E + \left(1 - \frac{\mu}{k_R^{\text{tl}}}\right) v_R &\geq 0. \end{aligned} \tag{4}$$

The resulting growth polyhedron amounts to

$$\begin{aligned} P_g(\mu) &= \{v \in \mathbb{R}^{\{P,E,R\}} \mid Nv \geq 0, v \geq 0, \\ &\quad \rho_P v_P = \mu, \\ &\quad \text{and (4)}\}. \end{aligned}$$

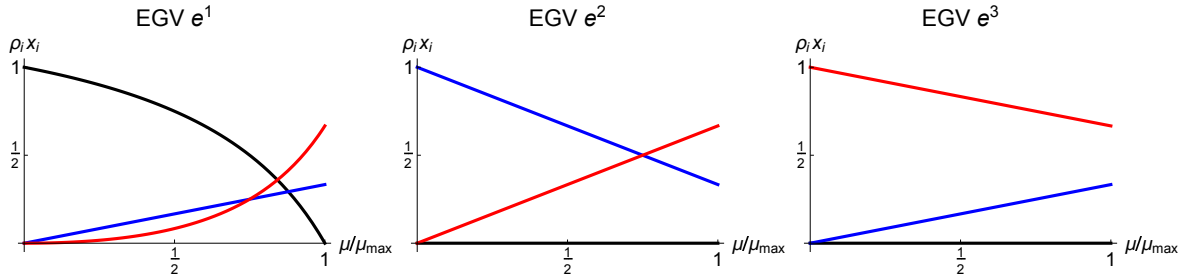


Figure 1: Mass fractions as functions of growth rate for the EGVs e^1 , e^2 , e^3 of the minimal network. Black/blue/red lines show mass fractions of precursor P/enzyme E/ribosome R.

For a quantitative analysis, we set $\frac{k^{\text{cat}}}{k_{\text{R}}^{\text{tl}}} = 2$, $\frac{k_{\text{E}}^{\text{tl}}}{k_{\text{R}}^{\text{tl}}} = 1$, introduce $\hat{\mu} = \frac{\mu}{k_{\text{R}}^{\text{tl}}}$, and find $\hat{\mu} \leq \hat{\mu}_{\text{max}} = \frac{2}{3}$. The EGVs and the associated mass fractions are given by

$$e^1 = \frac{\mu}{\rho_{\text{P}}} \begin{pmatrix} 1 \\ \frac{\hat{\mu}}{2} \\ \frac{\hat{\mu}^2}{2(1-\hat{\mu})} \end{pmatrix}, \quad e^2 = \frac{\mu}{\rho_{\text{P}}} \begin{pmatrix} 1 \\ 1 - \hat{\mu} \\ \hat{\mu} \end{pmatrix}, \quad e^3 = \frac{\mu}{\rho_{\text{P}}} \begin{pmatrix} 1 \\ \frac{\hat{\mu}}{2} \\ 1 - \frac{\hat{\mu}}{2} \end{pmatrix},$$

$$\rho \circ x(e^1) = \begin{pmatrix} 1 - \frac{\hat{\mu}}{2} - \frac{\hat{\mu}^2}{2(1-\hat{\mu})} \\ \frac{\hat{\mu}}{2} \\ \frac{\hat{\mu}^2}{2(1-\hat{\mu})} \end{pmatrix}, \quad \rho \circ x(e^2) = \begin{pmatrix} 0 \\ 1 - \hat{\mu} \\ \hat{\mu} \end{pmatrix}, \quad \rho \circ x(e^3) = \begin{pmatrix} 0 \\ \frac{\hat{\mu}}{2} \\ 1 - \frac{\hat{\mu}}{2} \end{pmatrix}.$$

For illustration, the associated mass fractions are shown as functions of growth rate in Figure 1.

The defining difference between the EGVs concerns the (in)activity of the inequality constraints $x_{\text{P}} \geq 0$, $x_{\text{E}} \geq \frac{v_{\text{P}}}{k^{\text{cat}}}$, and $x_{\text{R}} \geq \frac{v_{\text{E}}}{k_{\text{E}}^{\text{tl}}} + \frac{v_{\text{R}}}{k_{\text{R}}^{\text{tl}}}$. For $\mu < \mu_{\text{max}}$, all EGVs have one inactive constraint,

$$e^1: x_{\text{P}} > 0, \quad e^2: x_{\text{E}} > \frac{v_{\text{P}}}{k^{\text{cat}}}, \quad e^3: x_{\text{R}} > \frac{v_{\text{E}}}{k_{\text{E}}^{\text{tl}}} + \frac{v_{\text{R}}}{k_{\text{R}}^{\text{tl}}}.$$

For $\mu = \mu_{\text{max}}$, all EGVs are identical, $e^1 = e^2 = e^3$, and all inequality constraints are active, $x_{\text{P}} = 0$, $x_{\text{E}} = \frac{v_{\text{P}}}{k^{\text{cat}}}$, and $x_{\text{R}} = \frac{v_{\text{E}}}{k_{\text{E}}^{\text{tl}}} + \frac{v_{\text{R}}}{k_{\text{R}}^{\text{tl}}}$. The associated mass fractions are given by $\rho \circ x = (0, 1 - \hat{\mu}_{\text{max}}, \hat{\mu}_{\text{max}})^T = (0, \frac{1}{3}, \frac{2}{3})^T$.

For all EGVs, all fluxes are nonzero. For $\mu < \mu_{\text{max}}$, also all concentrations are nonzero for e^1 , whereas $x_{\text{P}} = 0$ and $x_{\text{E}}, x_{\text{R}} > 0$ for e^2 and e^3 . Obviously, the capacity constraints do not *fully* reflect kinetics.

All EGVs are BC (since they involve catalytic reactions) and CC (since the active catalysts have nonzero concentrations). If $\mu < \mu_{\text{max}}$, then e^1 also has nonzero precursor concentration, that is, e^1 is strict and ultimately AC. As stated in Example 13, $S = \text{supp}(e^1) = \text{Rxn}$ is the unique (M)AC set of reactions.

4 Simple kinetic models of self-fabrication

Simple kinetic models of self-fabrication as studied in [5, 18] consider metabolism as well as enzyme and ribosome synthesis. In such models, one distinguishes the set of *metabolites* Met and the set of (catalytic) *macromolecules* Mac (namely, enzymes and the ribosome). Correspondingly, one considers the set of *metabolic reactions* Rmet (one reaction per enzyme) and the set of *synthesis reactions* Rsyn (one reaction per macromolecule). To summarize, the sets of molecules Mol and reactions Rxn are given by

$$\begin{aligned} \text{Mol} &= \text{Met} \cup \text{Mac} \quad \text{with} \quad \text{Mac} = \text{Enz} \cup \{\text{rib}\}, \\ \text{and} \quad \text{Rxn} &= \text{Rmet} \cup \text{Rsyn}. \end{aligned}$$

The dynamic growth model (1) for the concentrations $x \in \mathbb{R}_{\geq}^{\text{Mol}}$ takes the form

$$\frac{d}{dt} \begin{pmatrix} x_{\text{Met}} \\ x_{\text{Mac}} \end{pmatrix} = \begin{pmatrix} M & -S \\ 0 & I \end{pmatrix} \begin{pmatrix} v_{\text{Rmet}}(x) \\ v_{\text{Rsyn}}(x) \end{pmatrix} - \mu \begin{pmatrix} x_{\text{Met}} \\ x_{\text{Mac}} \end{pmatrix} \quad (5a)$$

with the dry weight constraint

$$\rho_{\text{Met}}^T x_{\text{Met}} + \rho_{\text{Mac}}^T x_{\text{Mac}} = 1. \quad (5b)$$

Thereby, we write $M \in \mathbb{R}^{\text{Met} \times \text{Rmet}}$ for the stoichiometric matrix of metabolism (without a ‘‘biomass reaction’’) and $S \in \mathbb{R}^{\text{Met} \times \text{Rsyn}}$ for the stoichiometric matrix of synthesis, specifying the stoichiometric coefficients of the metabolites (precursors, cofactors, ...) in the synthesis reactions (for enzymes and ribosome). Further, we write $v_{\text{Rmet}}(x) \in \mathbb{R}^{\text{Rmet}}$ for the rates of the metabolic reactions and $v_{\text{Rsyn}}(x) \in \mathbb{R}_{\geq}^{\text{Rsyn}}$ for the synthesis rates; that is, $v_{\text{Rsyn}}(x) \geq 0$.

Remark 17. Following Remark 4, we assume $\ker M^T = \{0\}$, that is, $\text{im } M = \mathbb{R}^{\text{Met}}$ in the following. Further, we assume that S has nonnegative/nonpositive rows, and every column has at least one positive entry.

In simplified notation, the dynamic growth model (5) takes the form:

$$\frac{d}{dt} \begin{pmatrix} x \\ c \end{pmatrix} = \begin{pmatrix} M & -S \\ 0 & I \end{pmatrix} \begin{pmatrix} v(x, c) \\ w(x, c) \end{pmatrix} - \mu \begin{pmatrix} x \\ c \end{pmatrix} \quad (6a)$$

with

$$\rho_x^T x + \rho_c^T c = 1. \quad (6b)$$

By abuse of the symbols x and v , we write $x = x_{\text{Met}}$ (for the metabolite concentrations), $v = v_{\text{Rmet}}$ (for the metabolic reactions), and $\rho_x = \rho_{\text{Met}}$ (for the molar masses of the metabolites). Analogously, we write $c = x_{\text{Mac}}$ (for the enzyme/ribosome concentrations), $w = v_{\text{Rsyn}}$ (for the synthesis reactions), and $\rho_c = \rho_{\text{Mac}}$ (for the molar masses of enzymes/ribosome).

Elementary vectors

At steady state,

$$Mv - Sw = \mu x \geq 0, \quad (7a)$$

$$w = \mu c \geq 0. \quad (7b)$$

Further, $v_{\bar{\mathcal{I}}} \geq 0$, where $\bar{\mathcal{I}} \subseteq \text{Rmet}$ denotes the set of irreversible metabolic reactions. (Implicitly, all synthesis reactions Rsyn are assumed to be irreversible.)

In the following, we identify Rmet with Enz (the metabolic reactions with the catalyzing enzymes) and Rsyn with Mac (the synthesis reactions with the synthesized macromolecules).

We instantiate Definition 5 and Theorem 9.

Definition 18. *Growth modes* (GMs) for the dynamic growth model (6) are elements of the *growth cone*

$$\bar{C}_g = \{(v, w) \in \mathbb{R}^{\text{Enz}} \times \mathbb{R}^{\text{Mac}} \mid Mv - Sw \geq 0, v_{\bar{\mathcal{I}}} \geq 0, w \geq 0\}.$$

A GM $(v, w) \in \bar{C}_g$ has an *associated* growth rate $\mu(v, w) = \rho_x^T(Mv - Sw) + \rho_c^T w$ and, if $\mu(v, w) > 0$, *associated* concentration vectors $x(v, w) = (Mv - Sw)/\mu(v, w)$ and $c(v, w) = w/\mu(v, w)$.

Elementary growth modes (EGMs) are the conformally non-decomposable vectors of \bar{C}_g .

Theorem 19. *Every nonzero GM $(v, w) \in \bar{C}_g$ is a conformal sum of EGMs.*

The properties of the matrix S imply the following result.

Proposition 20. *Let $(v, w) \in \bar{C}_g$ be a GM. If $Mv = 0$, then $w = 0$.*

Below, we use Theorem 19 and Proposition 20 to derive an upper bound for growth rate that only depends on enzyme kinetics.

Enzyme kinetics

Often, one considers enzyme kinetics of the form

$$v(x, c) = \kappa(x) \circ c_{\text{Enz}} \quad \text{with} \quad -k^{\text{cat},-} \leq \kappa(x) \leq k^{\text{cat},+}. \quad (8)$$

The rate of the metabolic reaction catalyzed by enzyme $i \in \text{Enz}$ is given by $v_i(x, c_i) = \kappa_i(x) c_i$, that is, by a particular kinetics times the enzyme concentration.

Enzyme kinetics alone (without specifying ribosome kinetics) implies an upper bound for growth rate (which depends on the set of active enzymes).

Theorem 21. *Assume steady state (7) and let enzyme kinetics $v(x, c)$ be given by (8). Then, for every x , there is an upper bound for μ , depending on $\text{supp}(v)$.*

Proof. By assumption, $Mv - Sw = \mu x$, $w = \mu c$, $v_{\mathcal{I}} \geq 0$, and $v(x, c) = \kappa(x) \circ c_{\text{Enz}}$. Obviously, (v, w) is a GM. By Theorem 19, (v, w) is a conformal sum of EGMs. That is, there exists a set E of representative EGMs $e = (v^e, w^e)$ conforming to (v, w) (with one representative EGM on each ray of EGMs) such that

$$(v, w) = \sum_{e \in E} \lambda_e (v^e, w^e) \quad \text{with } \lambda_e \geq 0.$$

Let $s = \text{supp}(v)$. If $i \in s$, then $v_i/\kappa_i = c_i = w_i/\mu$, that is, $\mu = w_i/(v_i/\kappa_i)$, and further

$$\mu = \frac{\sum_{i \in s} w_i}{\sum_{i \in s} v_i/\kappa_i},$$

since $r = a/b = c/d$ implies $r = (a + c)/(b + d)$. Using the conformal sum,

$$\begin{aligned} \mu &= \frac{\sum_{i \in s} \sum_{e \in E} \lambda_e w_i^e}{\sum_{i \in s} \sum_{e \in E} \lambda_e v_i^e/\kappa_i} \\ &= \frac{\sum_{e \in E} \lambda_e \sum_{i \in s} w_i^e}{\sum_{e \in E} \lambda_e \sum_{i \in s} v_i^e/\kappa_i} \\ &= \frac{\sum_{e \in E} \lambda_e b_e}{\sum_{e \in E} \lambda_e a_e} \end{aligned}$$

with $a_e = \sum_{i \in s} v_i^e/\kappa_i \geq 0$ and $b_e = \sum_{i \in s} w_i^e \geq 0$.

By Proposition 20, $v^e = 0$ implies $w^e = 0$. Equivalently, $a_e = 0$ implies $b_e = 0$, that is, $b_e > 0$ implies $a_e > 0$. Hence, μ as a ratio of linear functions (in the variables λ_e) is bounded. In fact,

$$\mu \leq \max_{e \in E: a_e > 0} \frac{b_e}{a_e}.$$

□

Ribosome kinetics (with control parameters)

Often, one considers ribosome kinetics of the form

$$w(x, c) = \tau(x) \circ \alpha \cdot c_{\text{rib}} \quad \text{with } \tau(x) \leq k^{\text{tl}}, \quad (9a)$$

involving the control parameters $\alpha \in \mathbb{R}_{\geq}^{\text{Mac}}$ with

$$\sum_{i \in \text{Mac}} \alpha_i \leq \bar{\alpha} \leq 1. \quad (9b)$$

The rate of the synthesis reaction for macromolecule $i \in \text{Mac} = \text{Enz} \cup \{\text{rib}\}$ is given by $w_i(x, c) = \tau_i(x) \alpha_i c_r$, that is by a particular kinetics times a fraction of the ribosome concentration. Most importantly, the control parameters (ribosome fractions) α are used to study growth rate maximization in Section 5.

Steady state (7), enzyme kinetics (8), and ribosome kinetics (9a) imply basic results regarding the case $\mu = 0$.

Proposition 22. *Let $x \in \mathbb{R}_{\geq}^{\text{Met}}$ be nonzero. The following statements hold:*

- $\mu = 0 \implies w = 0 \implies Mv = 0.$
- $c_{\text{rib}} = 0 \implies w = 0 \implies \mu = 0 \vee c_{\text{Enz}} = 0.$
 $c_{\text{Enz}} = 0 \implies v = 0 \wedge w_{\text{Enz}} = 0 \implies w_{\text{rib}} = 0 \implies \mu = 0.$
Hence, $c_{\text{rib}} = 0 \implies \mu = 0.$

Conversely, $\mu = 0 \implies c_{\text{rib}} = 0$ does not hold if $\tau(x) \circ \alpha = 0$ (a degenerate case). In the following, we often assume $\mu > 0$ (and hence $c_{\text{rib}}, w_{\text{rib}}, \tau_{\text{rib}}, \alpha_{\text{rib}} > 0$).

4.1 Related constraint-based models

In the rest of this section, we specify a constraint-based model corresponding to the simple kinetic model of self-fabrication.

First, enzyme kinetics (8) implies the capacity constraints

$$-k^{\text{cat},-} \circ c_{\text{Enz}} \leq v \leq k^{\text{cat},+} \circ c_{\text{Enz}},$$

which can be written as

$$\begin{aligned} c_{\text{Enz}} + \text{diag}(k^{\text{cat},-})^{-1} v &\geq 0, \\ c_{\text{Enz}} - \text{diag}(k^{\text{cat},+})^{-1} v &\geq 0. \end{aligned}$$

Second, ribosome kinetics (9a) and the capacity constraint (9b) imply

$$\sum_{i \in \text{Mac}} \frac{w_i}{k_i^{\text{tl}}} \leq c_{\text{rib}},$$

that is,

$$c_{\text{rib}} - ((k^{\text{tl}})^{-1})^T w \geq 0.$$

Altogether,

$$\begin{pmatrix} 0 & \text{I} & 0 \\ 0 & \text{I} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ c_{\text{Enz}} \\ c_{\text{rib}} \end{pmatrix} + \begin{pmatrix} \text{diag}(k^{\text{cat},-})^{-1} & 0 \\ -\text{diag}(k^{\text{cat},+})^{-1} & 0 \\ 0 & -((k^{\text{tl}})^{-1})^T \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} \geq 0, \quad (10)$$

which is of the general form $A \begin{pmatrix} x \\ c \end{pmatrix} + B \begin{pmatrix} v \\ w \end{pmatrix} \geq b$. Here, the metabolite concentrations x do not contribute to the constraints, and the right-hand side b is zero.

In general, constraint-based growth models (3) involve steady state (with irreversibility), dry weight, and additional linear constraints. Here, the model involves steady state (7) (with irreversibility $v_{\bar{I}} \geq 0$), dry weight (6b), and the additional linear constraints (10).

After eliminating enzyme/ribosome concentrations c via steady state (7b), that is, $c = w/\mu$, the additional constraints only involve the reaction rates (fluxes) v and w ,

$$\begin{aligned} w_{\text{Enz}} + \mu \text{diag}(k^{\text{cat},-})^{-1} v &\geq 0, \\ w_{\text{Enz}} - \mu \text{diag}(k^{\text{cat},+})^{-1} v &\geq 0, \\ w_{\text{rib}} - \mu ((k^{\text{tl}})^{-1})^T w &\geq 0. \end{aligned} \quad (11)$$

By Definition 14, the growth polyhedron for the constraint-based model is given by

$$\bar{P}_g(\mu) = \left\{ (v, w) \in \mathbb{R}^{\text{Enz}} \times \mathbb{R}^{\text{Mac}} \mid \begin{aligned} &Mv - Sw \geq 0, v_{\bar{r}} \geq 0, w \geq 0, \\ &\rho_x^T Mv = \mu, \\ &\text{and (11)} \end{aligned} \right\},$$

and GVs and EGVs are defined accordingly.

5 Growth rate maximization

We consider the simple kinetic model of self-fabrication and study the problem of growth rate maximization. In fact, we maximize growth rate at steady state (7) and for given enzyme kinetics (8), ribosome kinetics (9a) with ribosome capacity (9b), and dry weight (6b).

Thereby, we further simplify notation and introduce the abbreviations $E = \text{Enz}$ and $r = \text{rib}$ and hence $\text{Mac} = E \cup \{r\}$.

Problem 23.

$$\begin{aligned} &\max_{\substack{x \geq 0, c \geq 0, \\ v; w, \alpha \geq 0, \\ \mu > 0}} \mu \\ &\text{s.t. } \quad Mv - Sw = \mu x, \quad (\text{mss}) \\ &\quad \quad w = \mu c, \quad (\text{ess/rss}) \\ &\quad \quad v = \kappa(x) \circ c_E, \quad (\text{ek}) \\ &\quad \quad w = \tau(x) \circ \alpha \cdot c_r, \quad (\text{rk}) \\ &\quad \quad \rho_x^T x + \rho_c^T c = 1, \quad (\text{dw}) \\ &\quad \quad \sum_{i \in E \cup \{r\}} \alpha_i \leq \bar{\alpha}. \quad (\text{rc}) \end{aligned}$$

Legend. mss/ess/rss = metabolite/enzyme/ribosome steady state, ek/rk = enzyme/ribosome kinetics, dw = dry weight, rc = ribosome capacity

Remark 24. We write $x \geq 0$ in the optimization problem. In fact, we assume $x \in D$, where D is a compact (bounded and closed) set. As a consequence, the maximum is attained.

In the following, we eliminate the enzyme/ribosome concentrations c and the fluxes v, w as independent optimization variables. That is, we only keep the control parameters (ribosome fractions) α as well as the metabolite concentrations x and growth rate μ .

Enzyme/ribosome steady state (7b) and ribosome kinetics (9a) imply

$$\mu c = w = \tau \circ \alpha \cdot c_r,$$

that is,

$$\begin{aligned}\mu c_E &= \tau_E \circ \alpha_E \cdot c_r, \\ \mu &= \tau_r \alpha_r,\end{aligned}$$

which yields c_E in terms of α_E (and x , μ).

Now, dry weight (6b) implies

$$\rho_x^T x + \rho_c^T (\tau \circ \alpha) c_r / \mu = 1,$$

that is,

$$(\rho_c \circ \tau)^T \alpha \cdot c_r = \mu (1 - \rho_x^T x),$$

which yields c_r in terms of α (and x , μ).

As a consequence, all enzyme concentrations, all fluxes, and also growth rate are multiples of the ribosome concentration c_r ,

$$\begin{aligned}v &= \kappa \circ c_E = \kappa \circ \tau_E \circ \alpha_E \cdot c_r / \mu, \\ w &= \tau \circ \alpha \cdot c_r, \\ \mu &= \frac{(\rho_c \circ \tau)^T \alpha}{1 - \rho_x^T x} c_r,\end{aligned}$$

and we divide metabolite steady state by $c_r > 0$. Altogether, we formulate growth rate maximization in terms of α (and x , μ).

Problem 25 (in terms of α).

$$\begin{aligned}\max_{\substack{x \geq 0, \mu > 0, \\ \alpha \geq 0}} \quad & \mu \\ \text{s.t.} \quad & M \hat{v} - S \hat{w} = \hat{\mu} x, & (\text{mss}) \\ & \mu = \tau_r(x) \alpha_r, & (\text{rss, rk}) \\ & \sum_{i \in \text{EU}\{r\}} \alpha_i \leq \bar{\alpha}, & (\text{rc}) \\ \text{where} \quad & \hat{v} = \kappa(x) \circ \tau_E(x) \circ \alpha_E / \mu, & (\text{ess, ek, rk}) \\ & \hat{w} = \tau(x) \circ \alpha, & (\text{rk}) \\ & \hat{\mu} = \frac{(\rho_c \circ \tau(x))^T \alpha}{1 - \rho_x^T x}. & (\text{ess/rss, rk, dw})\end{aligned}$$

Legend. Labels denote the dependence of the equations/inequalities on the original constraints in Problem 23.

Note. The last three equations are not constraints, but definitions of \hat{v} , \hat{w} , $\hat{\mu}$ in terms of α (and x , μ).

Elementary vectors

We follow the definition in [5], but write it in a more transparent way which allows to recognize the origin of the constraints and the resulting affine subspace at the same time.

Definition 26. Let $\mu > 0$. *Growth states* (GSs) for Problem 25 are elements of the polyhedron

$$\begin{aligned} \bar{P}(x, \mu) = \{ \alpha \geq 0 \mid & M\hat{v} - S\hat{w} = \hat{\mu}x \text{ and } \tau_r(x)\alpha_r = \mu, \text{ where} \\ & \hat{v} = \kappa(x) \circ \tau_E(x) \circ \alpha_E/\mu, \\ & \hat{w} = \tau(x) \circ \alpha, \\ & \hat{\mu} = \frac{(\rho_c \circ \tau(x))^T \alpha}{1 - \rho_x^T x} \}. \end{aligned}$$

Elementary growth states (EGSs) are the vertices of \bar{P} .

Note. $\bar{P}(x, \mu)$ is the intersection of an affine subspace with the nonnegative orthant and hence unbounded, in general; that is, \bar{P} is a polyhedron; see also Example 35. The elementary vectors of a polyhedron are its convex-conformally non-decomposable (ccND) vectors and the conformally non-decomposable vectors of its recession cone; see Supplement A.3. We define EGSs as ccND vectors which agree with the vertices due to nonnegativity.

By the definition of EGSs (as vertices of a polyhedron), we have the following (necessarily conformal) sum theorem.

Proposition 27. *For given polyhedron $\bar{P}(x, \mu)$, every convex combination of EGSs is a GS, but not vice versa, in general.*

By Theorem 21, for every x , there is an upper bound for μ , depending on $\text{supp}(v)$. (And ribosome capacity further limits growth rate.) Clearly, $\bar{P}(x, \mu)$ is empty for μ above the global upper bound. Moreover, different EGSs have different upper bounds, in general, and the number of vertices of $\bar{P}(x, \mu)$ depends on μ ; see also Example 36.

The following result is stated (with an incorrect proof) in [5]. We give a simple proof which uses a polyhedral geometry argument and does not require the implicit function theorem (IFT).

Theorem 28. *The maximum of Problem 25 is attained at an EGS.*

Proof. Let $H_{rc} = \{ \alpha \mid \sum_{i \in E \cup \{r\}} \alpha_i \leq \bar{\alpha} \}$ be the halfspace given by ribosome capacity. For given x and μ , in particular, for their optimal values, the set of feasible α is given by the polytope $P = \bar{P}(x, \mu) \cap H_{rc}$. If $P \neq \emptyset$ (as for optimal x and μ), then at least one vertex of P is a vertex of \bar{P} , that is, an EGS. \square

Note. Let x and μ be optimal. By the proof above, the maximum is attained at every EGS $\alpha \in \bar{P}(x, \mu)$ that fulfills $\sum_{i \in E \cup \{r\}} \alpha_i \leq \bar{\alpha}$.

Next, we show that EGSs (vertices of $\bar{P}(x, \mu)$) are support-minimal, a crucial result that is missing in [5].

Proposition 29. *EGSs are the support-minimal vectors of the polyhedron $\bar{P}(x, \mu)$.*

Proof. The polyhedron $\bar{P}(x, \mu)$ is given by the linear equations $M\hat{v} - S\hat{w} = \hat{\mu}x$ (homogeneous in α), the inhomogeneous equation $\tau_r(x)\alpha_r = \mu$, and nonnegativity $\alpha \geq 0$, that is, it is the intersection of a genuine affine subspace with the nonnegative orthant. By Proposition 2, the vertices of \bar{P} are its support-minimal vectors. \square

Using Proposition 29, we show that EGSs with the same support also exist in a neighbourhood (of x and μ).

Theorem 30. *Let $\alpha \in \bar{P}(x, \mu)$ be an EGS with $s = \text{supp}(x)$. Then, there exists an open neighborhood $U \subset \mathbb{R}_{>}^s \times \mathbb{R}_{>}$ of (x_s, μ) such that, for all $(\xi, \mu') \in U$, α can be continuously extended to an EGS $\alpha' \in \bar{P}(x', \mu')$ with $\text{supp}(x') = s$, $x'_s = \xi$, and $\text{supp}(\alpha') = \text{supp}(\alpha)$.*

Proof. Altogether, the polyhedron $\bar{P}(x, \mu)$ is given by inhomogeneous linear equations

$$A(x, \mu) \alpha = b(x, \mu) \quad (12)$$

and $\alpha \geq 0$, that is, it is the intersection of a genuine affine subspace with the non-negative orthant. By Proposition 29, the EGS $\alpha \in \bar{P}(x, \mu)$ is support-minimal. By Proposition 1, α is the unique solution of (12) with support $S = \text{supp}(\alpha)$. Hence, there is an invertible square submatrix $A_{R,S}(x, \mu)$ of A (with $|R| = |S|$) and a subvector $b_R(x, \mu)$ of b such that the subvector $\alpha_S > 0$ of α fulfills

$$F(x, \mu, \alpha_S) = A_{R,S}(x, \mu) \alpha_S - b_R(x, \mu) = 0.$$

Clearly, $\frac{\partial F}{\partial \alpha_S} = A_{R,S}(x, \mu)$ is invertible.

Now, consider x_s , the nonzero part of x . By the IFT, there exists an open neighborhood $U \subset \mathbb{R}_{>}^s \times \mathbb{R}_{>}$ of (x_s, μ) such that, for all $(\xi, \mu') \in U$, α_S can be continuously extended to $\alpha'_S > 0$. With $x' \in \mathbb{R}_{\geq}^{\text{Met}}$ given by $\text{supp}(x') = s$ and $x'_s = \xi$, the EGS α can be continuously extended to the EGS $\alpha' \in \bar{P}(x', \mu')$ with $\text{supp}(\alpha') = S$. \square

A weaker (and incorrect) version of Theorem 30 is stated in [5]. In fact, it has an unnecessary assumption on the support of α , and it lacks an assumption on the support of x .

Finally, using Theorem 30, we can further characterize optimal solutions (of growth rate maximization).

Corollary 31. *At every optimal EGS of Problem 25, the ribosome capacity constraint is active.*

Proof. Let $\alpha \in \bar{P}(x, \mu)$ be an EGS for maximum μ (and corresponding optimal x). By Theorem 30, there is $\mu' > \mu$ such that α can be continuously extended to an EGS $\alpha' \in \bar{P}(x, \mu')$ with $\text{supp}(\alpha') = \text{supp}(\alpha)$. Now, assume $\sum_{i \in \text{EU}\{r\}} \alpha_i < \bar{\alpha}$. By continuity (of α as a function of μ), also $\sum_{i \in \text{EU}\{r\}} \alpha'_i < \bar{\alpha}$. Hence, (x, μ', α') is a feasible solution of Problem 25 which contradicts the maximality of μ . \square

In the main text of [5], it is (incorrectly) claimed that, at maximum growth rate, the ribosome capacity constraint need not be active (but some ribosome fraction may become zero).

5.1 EGSs and minimal autocatalytic sets

By Definitions 26 and 18, every GS has a corresponding GM.

Proposition 32. *Let $\alpha \in \bar{P}(x, \mu)$ be a GS, and let $v = \kappa \circ \tau_E \circ \alpha_E \cdot c_r / \mu$ and $w = \tau \circ \alpha \cdot c_r$ with $c_r = \mu \frac{1 - \rho_x^T x}{(\rho_c \circ \tau)^T \alpha}$ be the corresponding (metabolic and synthesis) fluxes. Then, (v, w) is a GM with associated growth rate $\mu(v, w) = \mu$.*

Moreover, by Definition 12, the GM corresponding to a GS is autocatalytic. Thereby, we make the (biochemically meaningful) assumption on (enzyme and ribosome) kinetics that, if a species s is consumed in a reaction r (short: $s \rightarrow 0$), then $v_r > 0$ implies $x_s > 0$.

Proposition 33. *Let $\alpha \in \bar{P}(x, \mu)$ be a GS, and let (v, w) be the corresponding GM. Then, (v, w) is AC, and hence $\text{supp}(v, w)$ is an AC subset of reactions.*

Proof. First, we show that (v, w) is strict. Assume that $i \in \text{supp}(v) \subseteq E$ and $M_{si} \neq 0$ for some species $s \in \text{Met}$. The terms $M_{si}v_i$ and $-S_{sj}w_j$ in the sum $(Mv - Sw)_s \geq 0$ can be positive, negative, or zero, and there is a positive term. If there is a negative term, then there is a reaction that consumes s and hence $x_s > 0$. In any case, $(Mv - Sw)_s = \mu x_s > 0$. Alternatively, assume that $j \in \text{supp}(w) \subseteq E \cup \{r\}$. Regarding species $s \in \text{Met}$, the proof is analogous. Regarding species $s \in E \cup \{r\}$, obviously $I_{sj} \neq 0$ if and only if $s = j$ and hence $(Iw)_s = w_j > 0$.

Since all reactions are catalytic, (v, w) is BC. It remains to show that (v, w) is CC. Assume that $i \in \text{supp}(v) \subseteq E$. Then, $\tau_i, \alpha_i, c_r > 0$ and hence $(Iw)_i = w_i > 0$. Finally, $\tau_r, \alpha_r > 0$ (by the definition of \bar{P}) and hence $(Iw)_r = w_r > 0$. \square

We have shown that GSs correspond to autocatalytic (AC) sets of reactions. One may conjecture that EGSs correspond to minimal autocatalytic (MAC) sets of reactions. However, the formalization of this statement is beyond the scope of this work. Recall that the definition of EGSs depends on enzyme and ribosome kinetics (with control parameters), that is, ultimately on metabolite concentrations, whereas the definition of MAC sets of reactions only depends on stoichiometry.

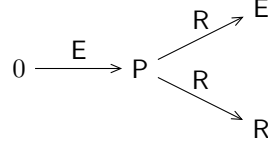
5.2 EGSs and EGMs

In three examples, we determine the EGSs and write them as conformal sums of EGMs.

First, we consider the minimal network studied in Examples 13 and 16, and then we extend it in two different ways: we assume that “metabolism” consists of (i) two anti-parallel pathways or (ii) two parallel (alternative) pathways. Thereby, we demonstrate two properties of the polyhedron of growth states (GSs): (i) it can be unbounded, and (ii) the number of its vertices (EGSs) can change with growth rate.

As in Examples 13 and 16, we use E to denote an individual enzyme (whereas $E = \text{Enz}$ is used as an abbreviation for the set of all enzymes in all other parts of Section 5) and R to denote the ribosome (instead of $r = \text{rib}$).

Example 34. Consider the minimal network studied in Examples 13 and 16:



The corresponding matrices of metabolism and synthesis amount to

$$(M \quad -S) = \begin{array}{c} \text{E} \\ \text{P} \end{array} \left(\begin{array}{c|cc} \text{E} & \rightarrow\text{E} & \rightarrow\text{R} \\ 1 & -1 & -1 \end{array} \right).$$

Further, $\rho_x = \rho_{\text{P}} =: \rho$ and $\rho_c = (\rho_{\text{E}}, \rho_{\text{R}})^T = \rho(1, 1)^T$, by mass conservation. Analogously, we assume $\tau(x) = (\tau_{\text{E}}(x), \tau_{\text{R}}(x))^T = \bar{\tau}(x)(1, 1)^T$.

(Normalized) EGMs are given by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

multiplied by a factor $u = 1$ mol/g/s. All EGMs have associated growth rate $\mu = u\rho$. By Definition 18, the first EGM has associated concentrations $x = x_{\text{P}} = 1$ and $c = (c_{\text{E}}, c_{\text{R}})^T = 0$, and the other two have $x = 0$ and $c \neq 0$.

Let $\alpha \in \bar{P}(x, \mu)$ be a GS, and let

$$\begin{aligned}
 \hat{v} &= v/c_{\text{R}} = \kappa \circ \tau_{\text{E}} \circ \alpha_{\text{E}} / \mu = \kappa \bar{\tau} \alpha_{\text{E}} / \mu && \in \mathbb{R}^{\text{P}}, \\
 \hat{w} &= w/c_{\text{R}} = \tau \circ \alpha = \bar{\tau} \alpha && \in \mathbb{R}^{\{\text{E}, \text{R}\}}, \\
 \hat{\mu} &= \mu/c_{\text{R}} = \frac{(\rho_c \circ \tau(x))^T \alpha}{1 - \rho_x^T x} = \frac{\rho \bar{\tau} (\alpha_{\text{E}} + \alpha_{\text{R}})}{1 - \rho x} > 0.
 \end{aligned}$$

By Definition 26, $M\hat{v} - S\hat{w} = \hat{\mu}x$ and $\tau_{\text{R}}\alpha_{\text{R}} = \mu$. Explicitly,

$$\frac{\kappa \alpha_{\text{E}}}{\mu} - \alpha_{\text{E}} - \alpha_{\text{R}} = \frac{\rho x (\alpha_{\text{E}} + \alpha_{\text{R}})}{1 - \rho x} \quad \text{and} \quad \bar{\tau} \alpha_{\text{R}} = \mu,$$

which has the unique solution

$$\alpha_{\text{E}} = \frac{\mu^2}{\bar{\tau}(\kappa(1 - \rho x) - \mu)} \quad \text{and} \quad \alpha_{\text{R}} = \frac{\mu}{\bar{\tau}},$$

thereby assuming $\kappa(x), \bar{\tau}(x) > 0$. That is, in the minimal network, there is a unique GS and hence a unique EGS. Moreover,

$$\mu < \mu_{\text{ub}}(x) = \kappa(x)(1 - \rho x).$$

In fact, the ribosome capacity constraint $\alpha_{\text{E}} + \alpha_{\text{R}} \leq 1$ (which is not part of Definition 26) further limits growth rate,

$$\mu \leq \mu_{\text{max}}(x) = \frac{\mu_{\text{ub}}(x) \bar{\tau}(x)}{\mu_{\text{ub}}(x) + \bar{\tau}(x)} < \mu_{\text{ub}}(x).$$

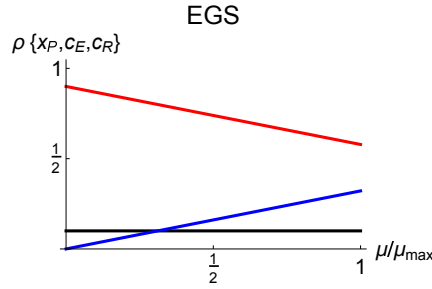


Figure 2: Mass fractions as functions of growth rate for the unique EGS of the minimal network (for fixed precursor concentration). Black/blue/red lines show mass fractions of precursor P/enzyme E/ribosome R.

By Proposition 32, the dimensionless vector $\frac{1}{\bar{\tau}}(\hat{v}, \hat{w})$ is the corresponding (scaled) GM, and further, by Theorem 19, it is a conormal sum of (scaled) EGMs,

$$\frac{1}{\bar{\tau}} \begin{pmatrix} \hat{v} \\ \hat{w} \end{pmatrix} = \begin{pmatrix} \frac{\kappa \alpha_E}{\mu} \\ \alpha_E \\ \alpha_R \end{pmatrix} = \frac{\rho x (\alpha_E + \alpha_R)}{1 - \rho x} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_E \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_R \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

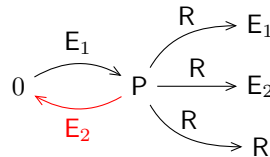
The corresponding mass fractions are given by

$$\rho \begin{pmatrix} x \\ c_E \\ c_R \end{pmatrix} = \begin{pmatrix} \rho x \\ (1 - \rho x) \frac{\alpha_E}{\alpha_E + \alpha_R} \\ (1 - \rho x) \frac{\alpha_R}{\alpha_E + \alpha_R} \end{pmatrix}.$$

For a quantitative analysis, we fix $x = x_P$; in particular, we set $\frac{\kappa(x)}{\bar{\tau}(x)} = 2$, $\rho x = \frac{1}{10}$ and find $\mu_{\text{ub}} = \frac{9}{5} \bar{\tau}$ and $\mu_{\text{max}} = \frac{9}{14} \bar{\tau}$. The resulting mass fractions are shown as functions of growth rate in Figure 2. Interestingly, they are linear. For $\mu = \mu_{\text{max}}$, $\rho(x, c_E, c_R)^T = (\frac{1}{10}, \frac{9}{28}, \frac{81}{140})^T$. Compare with Figure 1 and recall that EGVs do not depend on the precursor concentration x_P , but have an associated x_P .

By Proposition 33, the GM (v, w) corresponding to the unique (E)GS is AC. As stated in Example 13, $S = \text{supp}(v, w) = \text{Rxn}$ is the unique (M)AC set of reactions.

Example 35. “Metabolism” (the production and consumption of the precursor P) is catalyzed by the “enzymes” E_1 and E_2 , respectively, and the synthesis of E_1 , E_2 , and R is catalyzed by the ribosome R:



The corresponding matrices of metabolism and synthesis amount to

$$(M \quad -S) = \begin{array}{c} \text{E}_1 \quad \text{E}_2 \\ \text{P} \end{array} \left(\begin{array}{cc|ccc} \text{E}_1 & \text{E}_2 & \rightarrow \text{E}_1 & \rightarrow \text{E}_2 & \rightarrow \text{R} \\ 1 & -1 & -1 & -1 & -1 \end{array} \right).$$

Further, $\rho_x = \rho_P =: \rho$ and $\rho_c = (\rho_{E_1}, \rho_{E_2}, \rho_R)^T = \rho(1, 1, 1)^T$, by mass conservation. Analogously, we assume $\tau(x) = (\tau_{E_1}(x), \tau_{E_2}(x), \tau_R(x))^T = \bar{\tau}(x)(1, 1, 1)^T$.

(Normalized) EGMs are given by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

multiplied by a factor $u = 1$ mol/g/s. All EGMs except the last one have associated growth rate $\mu = u\rho$, and the last one has $\mu = 0$. By Definition 18, the first EGM has associated concentrations $x = 1$ and $c = 0$, and the second group of EGMs has $x = 0$ and $c \neq 0$.

There is only one EGS $\alpha \in \bar{P}(x, \mu)$, and it has $\alpha_{E_2} = 0$. This reduces the problem to the minimal network (with $E = E_1$), and the unique solution is given by $\alpha_{E_1} = \frac{\mu^2}{\bar{\tau}(\kappa_{E_1}(1-\rho x) - \mu)}$ and $\alpha_R = \frac{\mu}{\bar{\tau}}$.

By Proposition 33, the GM (v, w) corresponding to the unique EGS is AC. In fact, $S = \text{supp}(v, w) = \{\xrightarrow{E_1}, \rightarrow E_1, \rightarrow R\}$ is the unique MAC set of reactions.

In general, a GS $\alpha \in \bar{P}(x, \mu)$ is given by

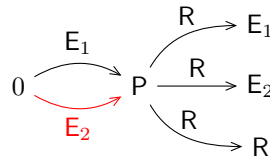
$$\frac{\kappa_{E_1} \alpha_{E_1}}{\mu} - \frac{\kappa_{E_2} \alpha_{E_2}}{\mu} - \alpha_{E_1} - \alpha_{E_2} - \alpha_R = \frac{\rho x (\alpha_{E_1} + \alpha_{E_2} + \alpha_R)}{1 - \rho x} \quad \text{and} \quad \bar{\tau} \alpha_R = \mu,$$

and the corresponding (scaled) GM can be specified as a conformal sum of (scaled) EGMs,

$$\frac{1}{\bar{\tau}} \begin{pmatrix} \hat{v} \\ \hat{w} \end{pmatrix} = \begin{pmatrix} \frac{\kappa_{E_1} \alpha_{E_1}}{\mu} \\ \frac{\kappa_{E_2} \alpha_{E_2}}{\mu} \\ \alpha_{E_1} \\ \alpha_{E_2} \\ \alpha_R \end{pmatrix} = \frac{\rho x (\alpha_{E_1} + \alpha_{E_2} + \alpha_R)}{1 - \rho x} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \alpha_{E_1} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_{E_2} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_R \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

with $\alpha_R = \frac{\mu}{\bar{\tau}}$ and $\beta \geq 0$. Clearly, for the unique EGS, $\alpha_{E_2} = \beta = 0$, and for unbounded GSs, $\beta = \frac{\kappa_{E_2} \alpha_{E_2}}{\mu} \rightarrow \infty$.

Example 36 (Alternative pathways). “Metabolism” (the production of the precursor P) is catalyzed alternatively by the “enzymes” E_1 and E_2 , and the synthesis of E_1 , E_2 , and R is catalyzed by the ribosome R:



The corresponding matrices of metabolism and synthesis amount to

$$(M \quad -S) = \begin{matrix} & \begin{matrix} E_1 & E_2 \end{matrix} & & & & & \\ \begin{matrix} P \end{matrix} & \left(\begin{array}{cc|ccc} 1 & 1 & -1 & -1 & -1 \end{array} \right) \end{matrix}$$

Further, $\rho_x = \rho_P =: \rho$ and $\rho_c = (\rho_{E_1}, \rho_{E_2}, \rho_R)^T = \rho(1, 1, 1)^T$, by mass conservation. Analogously, we assume $\tau(x) = (\tau_{E_1}(x), \tau_{E_2}(x), \tau_R(x))^T = \bar{\tau}(x)(1, 1, 1)^T$.

(Normalized) EGMs are given by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

multiplied by a factor $u = 1 \text{ mol/g/s}$. All EGMs have associated growth rate $\mu = u\rho$. By Definition 18, the first group has associated concentrations $x = x_P = 1$ and $c = (c_{E_1}, c_{E_2}, c_R)^T = 0$ and the other two groups have $x = 0$ and $c \neq 0$.

Let $\alpha \in \bar{P}(x, \mu)$ be the EGS with $\alpha_{E_2} = 0$. This reduces the problem to the minimal network (with $E = E_1$), and the unique solution is given by $\alpha_{E_1} = \frac{\mu^2}{\bar{\tau}(\kappa_{E_1}(1-\rho x) - \mu)}$ and $\alpha_R = \frac{\mu}{\bar{\tau}}$. Moreover, $\mu < \mu_{\text{ub},1}(x) = \kappa_{E_1}(x)(1 - \rho x)$.

Analogously, $\mu < \mu_{\text{ub},2}(x) = \kappa_{E_2}(x)(1 - \rho x)$ for the EGS α with $\alpha_{E_1} = 0$. Hence, depending on μ (and x), there are either two, one, or zero EGSs.

By Proposition 33, the GMs (v, w) corresponding to the EGS α with $\alpha_{E_2} = 0$ is AC. In fact, $S = \text{supp}(v, w) = \{\xrightarrow{E_1}, \rightarrow E_1, \rightarrow R\}$ is a MAC set of reactions. Analogously, $S = \{\xrightarrow{E_2}, \rightarrow E_2, \rightarrow R\}$ is a MAC set of reactions.

Proportional synthesis

The final results of this work concern the dependence of EGSs on growth rate and the relation between EGSs and EFM.

For simplicity, we assume that the reaction vectors of the synthesis reactions (for enzymes and ribosome) are proportional. Let $s^i \in \mathbb{R}^{\text{Met}}$ be the column of $S \in \mathbb{R}^{\text{Met} \times (E \cup \{r\})}$ corresponding to $i \in E \cup \{r\}$ (and note that $\rho_{c,i} = \rho_x^T s^i$). Now, let $i^* \in E$ be a representative enzyme with corresponding column $s := s^{i^*}$ and molar mass $\rho := \rho_{c,i^*}$. Then,

$$s^i = \frac{\rho_{c,i}}{\rho_{c,i^*}} s^{i^*} = \frac{\rho_{c,i}}{\rho} s.$$

In matrix notation,

$$S = \frac{1}{\rho} s \rho_c^T,$$

involving a dyadic product. At steady state (7),

$$Sw = \mu Sc = \frac{\mu}{\rho} (\rho_c^T c) s = \frac{\mu}{\rho} (1 - \rho_x^T x) s \quad (13a)$$

and further

$$Mv = Sw + \mu x = \frac{\mu}{\rho} ((1 - \rho_x^T x) s + \rho x) = \frac{\mu}{\rho} n \quad (13b)$$

with the dimensionless vector

$$n = (1 - \rho_x^T x) s + \rho x. \quad (13c)$$

Note. The vector n summarizes the effects of both synthesis reactions and dilution by growth.

5.3 EGSs and growth rate

Under the assumption of proportional synthesis, growth states (GSs) are elements of the (simplified) polyhedron

$$\begin{aligned} \bar{P}(x, \mu) = \{ \alpha \geq 0 \mid M\hat{v} = \frac{\hat{\mu}}{\rho} n \text{ and } \tau_r(x) \alpha_r = \mu, \text{ where} \\ \hat{v} = \kappa(x) \circ \tau_E(x) \circ \alpha_E / \mu, \\ \hat{\mu} = \frac{(\rho_c \circ \tau(x))^T \alpha}{1 - \rho_x^T x} \}, \end{aligned}$$

and elementary growth states (EGSs) are the vertices of \bar{P} . For fixed x , the dependence of EGSs on μ can be stated explicitly.

Theorem 37. *Assume proportional synthesis, and let $\alpha^* \in \bar{P}(x, \mu^*)$ be an EGS. Then, for every EGS $\alpha \in \bar{P}(x, \mu)$ with $\text{supp}(\alpha) = \text{supp}(\alpha^*)$, it holds that*

$$\alpha_E = \frac{(1 - \epsilon^*) \left(\frac{\mu}{\mu^*} \right)^2}{1 - \epsilon^* \frac{\mu}{\mu^*}} \alpha_E^* \quad \text{and} \quad \alpha_r = \frac{\mu}{\mu^*} \alpha_r^*$$

with

$$0 < \epsilon^*(x) = \frac{(\rho_{c,E} \circ \tau_E(x))^T \alpha_E^*}{(\rho_c \circ \tau(x))^T \alpha^*} < 1.$$

In particular, $\mu < \mu_{\text{ub}}(x) = \frac{\mu^*}{\epsilon^*(x)}$. The aggregated mass fractions of metabolites, enzymes, and ribosome are given by

$$\begin{pmatrix} \rho_x^T x \\ \rho_{c,E}^T c_E \\ \rho_{c,r}^T c_r \end{pmatrix} = \begin{pmatrix} \rho_x^T x \\ (1 - \rho_x^T x) \epsilon^* \frac{\mu}{\mu^*} \\ (1 - \rho_x^T x) (1 - \epsilon^* \frac{\mu}{\mu^*}) \end{pmatrix}.$$

Proof. The definition of $\bar{P}(x, \mu)$ implies $\tau_r(x) \alpha_r^* = \mu^*$, $\tau_r(x) \alpha_r = \mu$, and hence $\alpha_r = \gamma \alpha_r^*$ with $\gamma = \frac{\mu}{\mu^*}$. Analogously, assume $\alpha_E = \beta \alpha_E^*$ with $\beta > 0$. Then, $\hat{v} = \frac{\beta}{\gamma} \hat{v}^*$,

$$\begin{aligned} (\rho_c \circ \tau(x))^T \alpha &= (\rho_{c,E} \circ \tau_E(x))^T \alpha_E + \rho_{c,r} \tau_r(x) \alpha_r \\ &= \beta (\rho_{c,E} \circ \tau_E(x))^T \alpha_E^* + \gamma \rho_{c,r} \tau_r(x) \alpha_r^* \\ &= (\beta \epsilon^* + \gamma (1 - \epsilon^*)) (\rho_c \circ \tau(x))^T \alpha^*, \end{aligned}$$

and hence also $\hat{\mu} = (\beta \epsilon^* + \gamma (1 - \epsilon^*)) \hat{\mu}^*$. Now, consider $M\hat{v}^* = \frac{\hat{\mu}^*}{\rho} n$ and $M\hat{v} = \frac{\hat{\mu}}{\rho} n$. In the transition from μ^* to μ , the left- and right-hand sides scale with $\frac{\beta}{\gamma}$ and $\beta \epsilon^* + \gamma (1 - \epsilon^*)$, respectively. Hence, the assumption above is consistent if $\frac{\beta}{\gamma} = \beta \epsilon^* + \gamma (1 - \epsilon^*)$, that is, $\beta = \frac{(1 - \epsilon^*) \gamma^2}{1 - \epsilon^* \gamma}$.

To determine the mass fractions, observe that

$$\rho_{c,E} \circ c_E = \rho_{c,E} \circ \tau_E(x) \circ \alpha_E \cdot \frac{1}{\hat{\mu}}$$

scales with $\frac{\beta}{\beta \epsilon^* + \gamma (1 - \epsilon^*)} = \gamma$. Clearly, $\rho_{c,E}^T c_E^* + \rho_{c,r}^T c_r^* = 1 - \rho_x^T x$, and finally,

$$\frac{\rho_{c,E}^T c_E^*}{\rho_{c,E}^T c_E^* + \rho_{c,r}^T c_r^*} = \frac{(\rho_{c,E} \circ \tau_E(x))^T \alpha_E^*}{(\rho_c \circ \tau(x))^T \alpha^*} = \epsilon^*. \quad \square$$

Note. All control parameters (ribosome fractions) scale with the same nonlinear factor (as a function of growth rate), whereas ribosome synthesis scales linearly. Interestingly, enzyme mass fractions scale linearly (and hence also ribosome mass fraction). See also Figure 2 for the EGS of the minimal network.

5.4 EGSs and EFMs

Under the assumption of proportional synthesis, we introduce a “biomass reaction” **bm** with reaction vector $n_{\text{bm}} = -n$ and reaction rate $v_{\text{bm}} = \frac{\mu}{\rho}$ (unit: mol/g/s) corresponding to growth rate μ (unit: 1/s).

Note. For the definition of reaction rate, a scaling factor (unit: g/mol) is required; for simplicity, we choose the molar mass ρ of the representative protein.

At steady state, $Mv = \frac{\mu}{\rho} n$, cf. Equation (13b), that is,

$$(M \quad -n) \begin{pmatrix} v \\ v_{\text{bm}} \end{pmatrix} = 0.$$

Indeed, $(M \quad -n) \in \mathbb{R}^{\text{Met} \times (\text{EU}\{\text{bm}\})}$ is the stoichiometric matrix of traditional growth models, that is, the stoichiometric matrix of metabolism extended by the biomass reaction vector.

Flux modes (FMs) are elements of the corresponding flux cone

$$\bar{C}(x) = \{(v, v_{\text{bm}}) \in \mathbb{R}^{\text{EU}\{\text{bm}\}} \mid Mv - v_{\text{bm}} n = 0, v_{\bar{I}} \geq 0\},$$

and elementary flux modes (EFMs) are the support-minimal vectors of \bar{C} . The flux cone is an s-cone, and by Theorem 43 in Supplement A.1, every FM is a conormal sum of EFMs.

We characterize when the FM corresponding to an EGS is an EFM itself.

Theorem 38. *Assume proportional synthesis, and let $\alpha \in \bar{P}(x, \mu)$ be an EGS with $s = \text{supp}(\alpha_{\text{E}})$. Further, let $v \in \mathbb{R}^{\text{E}}$ be the corresponding metabolic fluxes and $v_{\text{bm}} = \mu/\rho$. Then, $(v, v_{\text{bm}}) \in \bar{C}(x)$ is an EFM if and only if $M_{*,s}$ has rank $|s|$.*

Proof. By assumption, $M_{*,s}v_s - v_{\text{bm}} n = 0$. By Proposition 41 in Supplement A.1, (v, v_{bm}) is an EFM if and only if $\ker(M_{*,s} \quad -n)$ is one-dimensional. Equivalently, $\ker M_{*,s} = \{0\}$, that is, $M_{*,s}$ has rank $|s|$. \square

Alternatively, we introduce a biomass reaction that considers only the synthesis reactions, but not dilution by growth. Indeed, $Sw = \frac{\mu}{\rho} n'$ with $n' = (1 - \rho_x^T x) s$, cf. Equation (13a), and we define $v_{\text{bm}} = \frac{\mu}{\rho}$ and $n_{\text{bm}} = -n'$.

At steady state, $Mv - Sw = \mu x$, cf. Equation (7), that is,

$$(M \quad -n') \begin{pmatrix} v \\ v_{\text{bm}} \end{pmatrix} = \mu x,$$

and the corresponding flux cone is given by

$$\bar{C}'(x) = \{(v, v_{\text{bm}}) \in \mathbb{R}^{\text{EU}\{\text{bm}\}} \mid Mv - v_{\text{bm}} n' = 0, v_{\bar{I}} \geq 0\}.$$

In traditional growth models, this definition of the flux cone arises from the assumption $Mv \approx Sw$, that is, $\mu x \ll |Mv|, |Sw|$ (with absolute values taken component-wise). This assumption has been made in FBA models, but is not always justified (in particular, if $(Sw)_i = 0$ and hence $(Mv)_i = \mu x_i$).

Again, we characterize when the FM corresponding to an EGS is an EFM itself.

Theorem 39. *Assume proportional synthesis, and let $\alpha \in \bar{P}(x, \mu)$ be an EGS with $s = \text{supp}(\alpha_E)$. Further, let $v \in \mathbb{R}^E$ be the corresponding metabolic fluxes and $v_{\text{bm}} = \mu/\rho$. Then, the following statements are equivalent:*

- (i) $v = v^1 + v^2$, where $(v^1, v_{\text{bm}}) \in \bar{C}'(x)$ is an EFM and $Mv^2 = \mu x$.
- (ii) $M_{*,s}$ has rank $|s|$.

Proof. See Theorem 38. □

Remark 40. In Theorem 39, both fluxes are proportional to growth rate, $Mv^1 = \frac{\mu}{\rho} n'$ and $Mv^2 = \mu x$. If the contribution of metabolites to dry weight is negligible, then the vector v^2 can be seen as a perturbation of the EFM v^1 . Indeed, $\rho_x^T x \ll 1$ implies $\mu \approx \rho_x^T Mv^1 \gg \rho_x^T Mv^2$.

6 Terminology

Elementary flux modes (EFMs) are defined as support-minimal vectors of the flux cone, for general (growth or non-growth) models, not just for traditional growth models (involving an approximative biomass reaction). In particular, EFMs only depend on stoichiometry. Similarly, elementary flux *vectors* (EFVs) are defined for general models with additional linear constraints.

In this work, we consider general *growth* models, and we define *elementary growth modes* (EGMs) as conformally non-decomposable vectors of the growth cone. Importantly, EGMs only depend on stoichiometry. Similarly, we introduce elementary growth *vectors* (EGVs) for general growth models with additional linear constraints (such as FBA and ME models).

For simple kinetic models of self-fabrication (with control parameters), *elementary growth states* (EGSSs) were introduced as vertices of a polyhedron (in the space of ribosome fractions) in [5]. In this work, we have shown that EGSSs are support-minimal. Hence, it makes sense to give a name to the supports of EGSSs. Unfortunately, these sets were called “elementary growth modes (EGMs)” in [5], without defining “growth modes (GMs)”. For several reasons, this terminology is not appropriate: Unlike EFMs, (i) “EGMs” are not defined for general growth models, but just for simple kinetic models of self-fabrication, (ii) “EGMs” do not depend on stoichiometry only, but also on growth rate and concentrations, in particular, on enzyme and ribosome kinetics, and (iii) “EGMs” are not elementary (vectors) in the sense of polyhedral geometry, but supports (sets). In this work, we have shown that supports of (E)GSSs correspond to (*minimal*) *autocatalytic sets* of reactions, defined for coarse-grained (but otherwise general) growth models. In a follow-up paper, we elaborate on the relation between detailed growth models and autocatalysis.

Acknowledgements

We thank Ralf Steuer for discussions on kinetic/constraint-based models of cellular growth (see also the review [21]) and Christoph Flamm for discussions on autocatalytic sets. SM was supported by the Austrian Science Fund (FWF), project P33218.

References

- [1] Jakob L. Andersen, Christoph Flamm, Daniel Merkle, and Peter F. Stadler. Defining Autocatalysis in Chemical Reaction Networks. *J. Syst. Chem.*, 8:121–133, 2020.
- [2] Kathrin Ballerstein, Axel von Kamp, Steffen Klamt, and Utz-Uwe Haus. Minimal cut sets in a metabolic network are elementary modes in a dual network. *Bioinformatics*, 28(3):381–387, December 2011.
- [3] Uri Barenholz, Dan Davidi, Ed Reznik, Yinon Bar-On, Niv Antonovsky, Elad Noor, and Ron Milo. Design principles of autocatalytic cycles constrain enzyme kinetics and force low substrate saturation at flux branch points. *eLife*, 6, February 2017.
- [4] Alex Blokhuis, David Lacoste, and Philippe Nghe. Universal motifs and the diversity of autocatalytic systems. *Proceedings of the National Academy of Sciences*, 117(41):25230–25236, September 2020.
- [5] Daan H. de Groot, Josephus Hulshof, Bas Teusink, Frank J. Bruggeman, and Robert Planqué. Elementary Growth Modes provide a molecular description of cellular self-fabrication. *PLoS Comput Biol*, 16(1):e1007559, 2020.
- [6] Hidde de Jong, Stefano Casagrandi, Nils Giordano, Eugenio Cinquemani, Delphine Ropers, Johannes Geiselmann, and Jean-Luc Gouzé. Mathematical modelling of microbes: metabolism, gene expression and growth. *Journal of The Royal Society Interface*, 14(136):20170502, November 2017.
- [7] Peter Dittrich and Pietro Speroni di Fenizio. Chemical organisation theory. *Bulletin of Mathematical Biology*, 69(4):1199–1231, April 2007.
- [8] Jeremy S. Edwards, Markus Covert, and Bernhard Palsson. Metabolic modelling of microbes: the flux-balance approach. *Environmental Microbiology*, 4(3):133–140, March 2002.
- [9] Itzhak Fishov, Arieh Zaritsky, and N. B. Grover. On microbial states of growth†. *Molecular Microbiology*, 15(5):789–794, March 1995.
- [10] Walter Fontana and Leo W. Buss. “The arrival of the fittest”: Toward a theory of biological organization. *Bulletin of Mathematical Biology*, 56(1):1–64, January 1994.
- [11] A. Goelzer and V. Fromion. Bacterial growth rate reflects a bottleneck in resource allocation. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1810(10):978–988, October 2011.

- [12] Reinhart Heinrich and Stefan Schuster. *The Regulation of Cellular Systems*. Springer US, 1996.
- [13] Wim Hordijk and Mike Steel. Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *Journal of Theoretical Biology*, 227(4):451–461, April 2004.
- [14] Christoph Kaleta, Florian Centler, and Peter Dittrich. Analyzing molecular reaction networks: From pathways to chemical organizations. *Molecular Biotechnology*, 34(2):117–124, 2006.
- [15] Stuart A. Kauffman. Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119(1):1–24, March 1986.
- [16] Steffen Klamt, Georg Regensburger, Matthias P. Gerstl, Christian Jungreuthmayer, Stefan Schuster, R. Mahadevan, Jürgen Zanghellini, and Stefan Müller. From elementary flux modes to elementary flux vectors: Metabolic pathway analysis with arbitrary linear flux constraints. *PLoS Comput. Biol.*, 13(4):e1005409, 2017.
- [17] Joshua A. Lerman, Daniel R. Hyduke, Haythem Latif, Vasilii A. Portnoy, Nathan E. Lewis, Jeffrey D. Orth, Alexandra C. Schrimpe-Rutledge, Richard D. Smith, Joshua N. Adkins, Karsten Zengler, and Bernhard O. Palsson. In silico method for modelling metabolism and gene product expression at genome scale. *Nature Communications*, 3(1), January 2012.
- [18] Douwe Molenaar, Rogier van Berlo, Dick de Ridder, and Bas Teusink. Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology*, 5(1):323, January 2009.
- [19] Stefan Müller and Georg Regensburger. Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis. *Front. Genet.*, 7(90):1–11, 2016.
- [20] Stefan Müller, Georg Regensburger, and Ralf Steuer. Enzyme allocation problems in kinetic metabolic networks: Optimal solutions are elementary flux modes. *J. Theoret. Biol.*, 347:182–190, 2014.
- [21] Stefan Müller and Ralf Steuer. Computational models of microbial growth. 2021. In preparation.
- [22] F. C. Neidhardt. Bacterial growth: constant obsession with dN/dt . *J. Bacteriol.*, 181(24):7405–7408, Dec 1999.
- [23] Ralph T. Rockafellar. The elementary vectors of a subspace of R^N . In *Combinatorial Mathematics and its Applications (Proc. Conf., Univ. North Carolina, Chapel Hill, N.C., 1967)*, pages 104–127. Univ. North Carolina Press, Chapel Hill, N.C., 1969.
- [24] Stefan Schuster and Claus Hilgetag. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.*, 2:165–182, 1994.

- [25] Stefan Schuster, Claus Hilgetag, John H. Woods, and David A. Fell. Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J. Math. Biol.*, 45(2):153–181, 2002.
- [26] M. Steel. The emergence of a self-catalysing structure in abstract origin-of-life models. *Applied Mathematics Letters*, 13(3):91–95, April 2000.
- [27] Robert Urbanczik. Enumerating constrained elementary flux vectors of metabolic networks. *IET Syst Biol*, 1(5):274–279, Sep 2007.
- [28] Robert Urbanczik and Clemens Wagner. Functional stoichiometric analysis of metabolic networks. *Bioinformatics*, 21(22):4176–4180, 2005.
- [29] Meike T. Wortel, Han Peters, Josephus Hulshof, Bas Teusink, and Frank J. Bruggeman. Metabolic states with maximal specific rate carry flux through an elementary flux mode. *FEBS Journal*, 281(6):1547–1555, February 2014.
- [30] Jürgen Zanghellini, Matthias P. Gerstl, Michael Hanscho, Govind Nair, Georg Regensburger, Stefan Müller, and Christian Jungreuthmayer. Toward genome-scale metabolic pathway analysis. In *Industrial Biotechnology: Microorganisms*, Advanced Biotechnology Book Series, pages 111–123. Wiley, 2017.

Supplementary material

A Elementary vectors

Below, we summarize basic definitions and results for s-cones, general polyhedral cones, and polyhedra.

A.1 S-cones

Given a linear subspace $S \subseteq \mathbb{R}^n$ and an index set $I \subseteq [n]$, an *s-cone* (special cone, subspace cone) is given by $C(S, I) = \{x \in \mathbb{R}^n \mid x \in S, x_I \geq 0\}$. Note that a linear subspace is an s-cone, $S = C(S, \emptyset)$.

Before we state the fundamental result for s-cones, we give an alternative definition of SM vectors, in order to demonstrate the underlying proof techniques.

Proposition 41. *Let $C(S, I)$ be an s-cone. A nonzero vector $x \in C(S, I)$ with $s = \text{supp}(x)$ is SM if and only if the linear subspace $S_s = \{x \in S \mid \text{supp}(x) \subseteq s\}$ has dimension one.*

Proof. Obviously, $x \in S_s$. (\Leftarrow) If $\dim S_s = 1$, then every element of S_s is a scalar multiple of x , and hence x is SM. (\Rightarrow) If x is SM, assume $\dim S_s > 1$. Then, there exists $x' \in S_s$ (not necessarily $x' \in C(S, I)$) which is not a scalar multiple of x . Consider $x^* = x - \lambda x'$ with $\lambda \in \mathbb{R}$. Note that $x^* \neq 0$. Indeed, choose λ such that $\text{sign}(x^*) \leq \text{sign}(x)$ (no sign change of x^* relative to x) and $x_i^* = x_i - \lambda x'_i = 0$ for some $i \in S_s$. Then, $x_i^* \geq 0$ and $\text{supp}(x^*) \subset \text{supp}(x)$. To summarize, $0 \neq x^* \in C(S, I)$, and x is not SM, a contradiction. \square

Corollary 42. *Let $C(S, I)$ be an s-cone. If two SM vectors have the same support, then they are scalar multiples.*

A vector $x \in C(S, I)$ is *elementary* if it is SM. (For linear subspaces, the definition of elementary vectors (EVs) as SM vectors was given in [23].)

The following result is fundamental. See [19, Theorem 3] based on [23, Theorem 1].

Theorem 43. *Let $C(S, I)$ be an s-cone. Every nonzero vector $x \in C(S, I)$ is a conformal sum of EVs. That is, there exists a finite set E of EVs such that*

$$x = \sum_{e \in E} e \quad \text{with } \text{sign}(e) \leq \text{sign}(x).$$

The set E can be chosen such that $|E| \leq \dim(S)$ and $|E| \leq |\text{supp}(x)|$.

A.2 General polyhedral cones

Let C be a polyhedral cone, that is, $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$ for some $A \in \mathbb{R}^{m \times n}$. A nonzero vector $x \in C$ is *conformally non-decomposable* (cND) if, for all nonzero $x^1, x^2 \in C$ with $\text{sign}(x^1), \text{sign}(x^2) \leq \text{sign}(x)$, the decomposition $x = x^1 + x^2$ implies $x^1 = \lambda x^2$ with $\lambda > 0$. A vector $x \in C$ is *elementary* if it is cND.

By defining EVs as cND vectors (instead of SM vectors), Theorem 43 can be extended to general polyhedral cones. See [19, Theorem 8].

Theorem 44. *Let $C = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$ be a polyhedral cone. Every nonzero vector $x \in C$ is a conformal sum of EVs. That is, there exists a finite set E of EVs such that*

$$x = \sum_{e \in E} e \quad \text{with } \text{sign}(e) \leq \text{sign}(x).$$

The set E can be chosen such that $|E| \leq \dim(C)$ and $|E| \leq |\text{supp}(x)| + |\text{supp}(Ax)|$.

A.3 Polyhedra

Let P be a polyhedron, that is, $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ for some $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. A vector $x \in P$ is *convex-conformally non-decomposable* (ccND) if for all $x^1, x^2 \in P$ with $\text{sign}(x^1), \text{sign}(x^2) \leq \text{sign}(x)$ and $0 < \lambda < 1$, the decomposition $x = \lambda x^1 + (1 - \lambda)x^2$ implies $x^1 = x^2$.

Let $R = \{x \in \mathbb{R}^n \mid Ax \geq 0\}$ be the recession cone of P . A vector $e \in P \cup R$ is *elementary* (an EV of P) if $e \in P$ is ccND or $e \in R$ is cND.

Ultimately, Theorem 43 can be extended to general polyhedral cones. See [19, Theorem 13].

Theorem 45. *Let $P = \{x \mid Ax \geq b\}$ be a polyhedron and $R = \{x \mid Ax \geq 0\}$ its recession cone. Every vector $x \in P$ is a conformal sum of EVs. That is, there exist finite sets $E_0 \subseteq R$ and $E_1 \subseteq P$ of EVs such that*

$$x = \sum_{e \in E_0} e + \sum_{e \in E_1} \lambda_e e \quad \text{with } \text{sign}(e) \leq \text{sign}(x),$$

$\lambda_e \geq 0$, and $\sum_{e \in E_1} \lambda_e = 1$. (Hence, $|E_1| \geq 1$.)

The set $E = E_0 \cup E_1$ can be chosen such that $|E| \leq \dim(P) + 1$ and $|E| \leq |\text{supp}(x)| + |\text{supp}(Ax)| + 1$.

B A minimal derivation of the dynamic growth model

We denote fundamental objects and quantities as follows:

Mol	set of molecular species	
Rxn	set of chemical reactions	
$N \in \mathbb{R}^{\text{Mol} \times \text{Rxn}}$	stoichiometric matrix	(unit: 1)
$X \in \mathbb{R}_{\geq}^{\text{Mol}}$	amounts of substance	(unit: mol)
$R(X) \in \mathbb{R}^{\text{Rxn}}$	reaction rates (extensive)	(unit: mol/s)
$\rho \in \mathbb{R}_{>}^{\text{Mol}}$	molar masses	(unit: g/mol)

The chemical reactions induce the dynamical system

$$\frac{dX}{dt} = NR. \tag{14}$$

We define dry weight (correctly: dry mass),

$$M = \sum_i \rho_i X_i = \rho^T X, \quad (\text{unit: g}) \quad (15a)$$

the intensive quantities

$$x = \frac{X}{M} \in \mathbb{R}_{\geq}^{\text{Mol}}, \quad (\text{unit: mol/g}) \quad (15b)$$

$$v = \frac{R}{M} \in \mathbb{R}^{\text{Rxn}}, \quad (\text{unit: mol/g/s}) \quad (15c)$$

and growth rate

$$\mu = \frac{1}{M} \frac{dM}{dt}. \quad (\text{unit: 1/s}) \quad (15d)$$

Note. We use mass instead of volume to define the “concentrations” x , the (intensive) reaction rates v , and growth rate μ . Thereby, we avoid discussions about “molecular volume”, osmotic pressure, etc. at this point. Moreover, in practice, cellular composition is often given in the unit mol/g (dry weight).

Finally, we recall the chain rule (of differentiation),

$$\frac{d}{dt} \frac{X}{M} = \frac{1}{M} \frac{dX}{dt} - \frac{X}{M^2} \frac{dM}{dt}.$$

Now, Equations (14), (15), and the chain rule yield the dynamic model of growth.

and

$$\frac{dx}{dt} = Nv(x) - \mu x$$

$$\rho^T x = 1.$$

Note. The derivation even holds for unbalanced growth.

C Mathematical issues in [de Groot et al, 2020]

The paper [5] is inspirational from the modeling perspective. Unfortunately, the mathematical treatment is not rigorous. In particular, Definition 1 is flawed, and a crucial intermediate result (regarding the support minimality of EGSs, cf. Proposition 29 in this work) is missing. As a consequence, almost every statement and/or its proof is imprecise or incomplete.

Incorrect statements

- Definition 1: $\mathcal{P}_{x,\mu}$ is not a polytope, in general, but a polyhedron.
- Theorem 1: As a consequence of Definition 1, not every element of $\mathcal{P}_{x,\mu}$ is a convex combination its vertices (EGSs), in general.
- Theorem 3: A condition on the support of x_0 is missing. Obviously, if $(x_0)_i = 0$ for some i , then there is no neighbourhood of (x_0, μ_0) such that the EGS α can be continuously extended.

Moreover, the statement has an unnecessary assumption, namely, that the support of the EGS equals its feasible basis.

- Theorem 5: As another consequence of Definition 1, even if the polyhedron $\mathcal{P}_{x,\mu}$ intersects the polyhedron P_{cons} determined by more than one inequality constraint, the polytope P_{EGSs} generated by the vertices (EGSs) of $\mathcal{P}_{x,\mu}$ does not intersect P_{cons} , in general. Hence, feasible solutions (in particular, optimal solutions) are not sums of EGSs, in general.

- Theorem 6: The statement requires the assumption that P_D has full rank. Under the assumption that $(P_D - \phi)$ has full rank, row reduction applied to $(P_D - \phi)$ does not yield $(I - \hat{\phi})$

Moreover, certain dimensions do not match. On the one hand, $D \subseteq \{1, \dots, n+1\}$ with $|D| = m+1$ is a feasible basis (of $B \in \mathbb{R}^{(m+1) \times (n+1)}$). On the other hand, $P \in \mathbb{R}^{m \times n}$ is the stoichiometric matrix. Hence, P_D is ill-defined.

Finally, the notation $\mathcal{O}(\mu)$ is misleading. In the sum $v = V + \mathcal{O}(\mu)$, also the “biomass flux mode” V is $\mathcal{O}(\mu)$.

- Theorems 7 and 8 are based on the incorrect Theorem 6.

Incorrect proof

- Theorem 4: Theorem 3 is used without checking the (unnecessary) assumption that the support of the EGS equals its feasible basis.

Moreover, the proof is too complicated, in particular, the implicit function theorem is not needed.