Journal of Mathematical Chemistry 12(1993)319–363

Autocatalytic and other general networks for chemical mechanisms, pathways, and cycles: Their systematic and topological generation

Oktay Sinanoğlu

Department of Chemistry and Department of Molecular Biochemistry/Biophysics, Yale University, 225 Prospect Street, P.O. Box 6666, New Haven, CT 06511, USA

Methods to generate a priori all the finite number of possible mechanisms of chemical reactions and/or synthetic pathways or thermodynamic cycles, which we represent by general networks, are given for any number of reaction steps or total number of species (reactants, products, catalysts, and intermediates). General networks do not place limitations on the types of species, e.g. intermediates can be short-lived, thereby participating in at most two elementary reaction steps, or longer-lived, participating in more than one way. Step stoichiometric coefficients can be more than unity. Reactants or products may also act as catalysts or inhibitors. Species vertices and the general networks themselves in which they occur are classified topologically. Topological invariants of the networks with respect to the number of reaction steps are found. Mechanisms with desired features, e.g. containing certain numbers of generalized catalysts, chains, autocatalysts, etc., are generated using the invariants, from the simplest prototypes, for successively larger numbers of reaction steps. Special emphasis is given to autocatalytic networks due to their role in chemical oscillations, dynamical instabilities and in self-replicating reactions. Examples given include the malic acid cycle, oscillatory cycles in glycolysis, Lotka–Volterra–Prigogine–Glansdorff models, and others. Oscillating and/or self-replicating cycles that have been invoked in various contexts are shown to have a common topological feature. The methods are useful also in the many autocatalytic processes of chemical engineering importance.

1. Introduction

Recently, it has become possible to introduce a good amount of systematics into the finding of reaction mechanisms [1]. In the past, experimentalists have proposed various mechanisms for specific observed reactions, but any systematic way of arriving at the possible mechanisms had been missing. In the first paper (I) [1], it was noted that for a given overall reaction type such as $A + B \Rightarrow C + D$, etc. with any A, B, C, D, as long as the equation obeys stoichiometry there are only a fairly small and finite number of possible mechanisms once the number ρ of elementary steps, or alternatively, the number σ of chemical species-moles to be admitted, are chosen. In fact, ρ and σ largely fix each other within narrow limits

© J.C. Baltzer AG, Science Publishers

(the $\sigma(\rho)$-plot) [1]. Mechanisms were denoted by simple *networks* drawn and their more compact *skeletons* [1].

Although skeletons, classified topologically in paper I [1] with respect to the numbers of catalytic, chain reaction, feedback, etc. loops they contain, were general in obtaining all the possible mechanisms/networks for various overall reactions, it was found to be convenient to deal with two broad types of mechanisms/networks: (i) *laminar* networks, and (ii) *turbulent* networks [1].

"*Laminar mechanisms*" [1] are the ones in which: (a) the stoichiometric coefficients ν_i of any species i in the *elementary reaction steps* are $|\nu_i| = 1$ or 0; (b) any *internal species* occurs in at most two elementary reaction steps. Otherwise, the network $\mathcal{N}$ (and its mechanisms $\{\mathcal{M}\}$) are "*turbulent*".

Paper I [1] gave, with chemical and biochemical examples, some of the general features of mechanisms/networks and, as a related problem, ways of generating a priori *pathways* to synthesize a target organic or bio-molecule from a pool of common reagents. The latter methodology [1] is useful in "*computer-assisted synthesis design*".

Also in paper I [1], all a priori mechanisms are classified, starting with just the numbers ρ or σ. This leads to a number of possible types of *overall reactions* ($\Rightarrow$), (OVR), depending on the network classes.

In subsequent papers (IIa, IIb) [2], we dealt with the more restricted question: how to find all the a priori mechanisms for a *given type of OVR*. Paper IIa [2] obtained all the laminar $\mathcal{M}$ and $\mathcal{N}$ for OVR types $A + B \Rightarrow C$ and $A \Rightarrow B$; paper IIb [2] obtained the OVR types $A + B \Rightarrow C + D$, giving many chemical examples.

In another paper [3] (which is not prerequisite to the present one), further details of one method to generate the a priori laminar mechanisms were presented. All laminar $\{\mathcal{M}\}$ with all the OVR types they lead to were obtained for $\rho = 2$, including the additional OVR type $2A + B \Rightarrow 2C + D$ and this OVR type also with $\rho = 3$ (cf. ref. [4]), again giving chemical examples.

The present paper deals with turbulent and general networks of any type: how to generate them and how to classify them. Autocatalytic steps [5], i.e. where some external products may act as catalysts, are one source of turbulent network aspects. Such steps are thought to be crucial as a cause of dynamic instability of steady states [6,7], in biological oscillating states [7,8], as well as in models of self-replicating systems. They are of course also important in chemical engineering.

The set of papers mentioned above have dealt with reacting mixtures mostly in batch (in this sense "closed") systems. For many applications, it was necessary to treat "open" or flow reactor-type systems, e.g. in the study of possible steady states. This has recently been done admirably by Poland [9]. Other works include the use of a "kinetics complexity index" (based, however, on the rate law expressions and for "linear mechanisms" only) by Bonchev et al. [10], the earlier algebraic approach of Sellers [11,12], the stability methods of Clarke [13], and the work of Barone et al. [14], and others (see refs. [10,11] and references cited therein).

In a different type of chemical application of the network theory, the complicated case of finding the a priori possible mechanisms for the *pyrolysis of ethane* [15] is treated in more detail in another paper [16].

More of the mechanisms or pathways invoked in chemistry or biochemistry are laminar. Turbulent ones arise when some intermediates are fairly long-lived, so that they have an opportunity to participate in a number of different reaction steps. As mentioned, "turbulence" arises also when an external species acts as a catalyst or is involved in self-regulation of the system. Another cause is the presence of homolytic steps such as ($A \rightarrow 2B$), still another is the occurrence of several OVRs with some product(s) in common. The latter is easily dealt with, such OVRs decoupled (cf. below) and this type of fairly trivial "turbulence" eliminated. In fact, such networks were therefore considered "laminar" but not "strictly laminar" (cf. paper I [1]).

Due to their greater complexity, most unsolved mechanisms are probably turbulent. Some reactions having turbulent features are:

(1) decomposition of acetone;

(2) decomposition of dimethyl ether;

(3) polymerization of diborane;

(4) decomposition of acetaldehyde;

(5) hydrolysis of esters, etc.

In the latter case, the reaction is catalyzed by the hydrogen ions. If one of the products is an acid, acid catalysis can be enhanced by the product. If we take this effect into account, the network is a turbulent one. If we disregard the dissociation of the acid, the reaction can be treated as if it were laminar. However, the turbulent mechanism is more appropriate for the interpretation of kinetic features under a wider range of conditions. Turbulent reactions are also found in enzyme reactions when an enzyme has more than one reaction site. Turbulence also occurs frequently in the organic and inorganic oxidation and reduction reactions.

In what follows, the different types of turbulent networks will be treated with regard to the types of chemical species vertices that can arise. The topological relations derived will lead to their systematic generation for increasing numbers of reaction steps. Special emphasis will be given to autocatalytic mechanisms, as well as not necessarily kinetic but, for example, metabolic or thermodynamic cycles or pathways. The methods described in this paper will also apply to *general networks*, without making a distinction between laminar and turbulent. Various examples, important in relation to dynamic instabilities in networks, chemical oscillations and biochemical models, will be given. The networks derived are involved also in chemical industrial problems such as fermentation reactions with micro-organisms on organic matter, oxidation of rubber, formation of red lead (Pb_3O_4) from powdered lead by oxidation [5], and many other examples, not to mention the ones in physical organic chemistry.

2. Laminar versus turbulent mechanisms and the species vertex as a source or sink

In a *strictly laminar network* [1], $\mathcal{N}_{\ell}$, one *species-mole line* (—) occurs for each species. The species-mole (sm) lines come together only at reaction (wiggle) lines (~~~). All *sm* lines that are thereby connected constitute a *lineblock* which is compressed into a *dot-point* in the *skeleton* $\mathcal{S}$ of that network $\mathcal{N}$. (For further definitions and the theory up to here, the reader is referred to paper I [1].)

In a *strictly laminar network*, the number of sm lines n equals the number of distinct chemical species σ,

$$n = \sigma, \tag{1ℓ}$$

where [1]

$$\sigma = \sigma_{\text{int}} + \sigma_{\text{ext}}. \tag{2}$$

Further, at each strictly laminar lineblock i, similarly

$$n_i = \overline{\omega}_i, \tag{3ℓ}$$

where $\overline{\omega}_i \equiv$ number of species at i, and of course

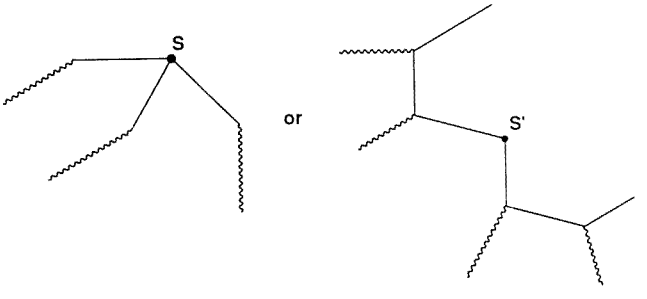
$$\sum_i n_i = n \tag{4}$$

and

$$\sum_i \overline{\omega}_i = \sigma \tag{5}$$

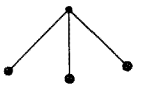
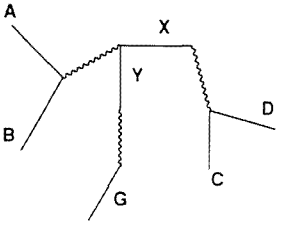
for the $\mathcal{N}_{\ell}$. Recall that in the $\mathcal{S}$ of $\mathcal{N}_{\ell}$, the dot-points $\{i\}$ get "*weights*" $\{\overline{\omega}_i\}$ depending on the $\mathcal{N}_{\ell} \subset \mathcal{S}$.

In *turbulent networks* $\mathcal{N}_{\text{t}}$, there may be several moles n_s occurring for a given species s. If so, the sm lines are joined to a species vertex [1] s, e.g.

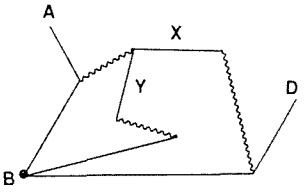


(6)

Note that starting from a laminar network, this may also have the effect of merging two or more otherwise separate lineblocks into one. This would change the skeleton; for example,

$\mathcal{S}$:  of $\mathcal{N}_\ell$: 

but if B = G = C, $\mathcal{S}$ becomes

$\mathcal{S}'$:  of $\mathcal{N}_t$: 

However, it would be awkward to start with the $\{\mathcal{N}_\ell\}$ already obtained and continue to merge sm lines into s-vertices to obtain some $\{\mathcal{N}_t\}$. Rather, below we still take one $\mathcal{S}$ at a time and derive methods to obtain all the turbulent as well as laminar $\mathcal{N}$s of that $\mathcal{S}$ without altering the $\mathcal{S}$.

For turbulent networks, eq. (1) no longer holds. Instead, we have

$$n > \sigma, \tag{1t}$$

or generally, for any $\mathcal{N}$(t or ℓ),

$$\boxed{n \geq \sigma}\,, \tag{1'}$$

where the equality holds for the laminar cases.

Similarly, if a lineblock i displays species vertices

$$n_i > \overline{\omega}_i, \tag{3t}$$

or for a general i (ℓ or t),

$$\boxed{n_i \geq \overline{\omega}_i}\,. \tag{3'}$$

For laminar networks, their skeletal dot-points $\{i\}$ were labeled [1–3] by weights $\{\overline{\omega}_i\}$.

For *general networks* (t or ℓ), we see that skeletal dot-points $\{i\}$ will have to have "*double weights*" ($\overline{\omega}_i$, n_i) satisfying eqs. (3′) and (1′).

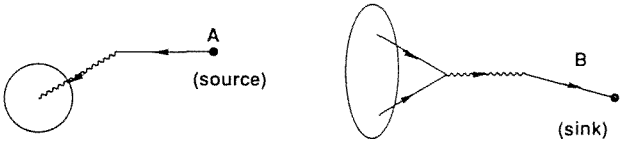
The relations between ρ, σ, γ, σ_{int}, σ_{ext} were given in [1] for general networks (ℓ or t). Rules for obtaining the possible weights $\overline{\omega}_i$ to an $\mathcal{S}$ to obtain the $\{\mathcal{N}_\ell\}$ out

of that $\mathcal{S}$ were derived and used in [1–3]. In the following, the double weights (ω_i, n_i) will be derived for the $\mathcal{S}$ to obtain the additional $\mathcal{N}_t$ possibilities out of that $\mathcal{S}$.

The skeletons possible with elementary reaction step numbers $\rho = 1, 2, 3$ and 4 were displayed in tables 1 and 2 of [1]. For a given $\mathcal{S}$, of the possible (ω_i, n_i) double weight labelings of the dot-points, those double weights which show $\omega_i = n_i$ correspond, of course, to the usual [2,3] single weight (ω_i) laminar subset of $\{\mathcal{N}\}_{\mathcal{S}}$. The others give the turbelent $\{\mathcal{N}_t\}_{\mathcal{S}}$.

In laminar networks, it was shown [1–4] that one does not need to show the reaction direction arrows. The direction of material flow is implicit in the network itself for each sm line and reaction step line, up to a reversal of all arrows.

In strictly laminar networks, the only "*sources*" or "*sinks*" of molecules are the free tips of external species mole-lines.



(7)

In turbulent networks, *all* species vertices $\{s\}$ are sources or sinks or both. They are sorted out as follows.

3. Types of species vertices

Consider a species vertex s of a species $s \in \{A, B, \ldots; X, Y, \ldots\}$ with n_s species-mole lines. The n_s is also

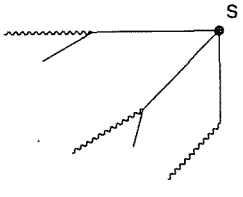
$$n_s = \sum_{J\geq 1}^{\rho} (|\nu_{s,L}^J| + |\nu_{s,R}^J|), \tag{8}$$

i.e. the absolute value sums of stoichiometric coefficients of s in the left (L) and right (R)-hand sides of each elementary step J. Also

$$\sum_{s\geq 1}^{\sigma} n_s = n \tag{9}$$

for the full network.

If n_s is odd, s must be a source or a sink; hence, s is an external species. Since each line has at

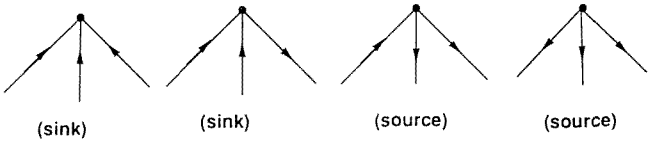


(10)

its other end a (~~~) line, when *consistency arrows* [1,3] are placed on the $\mathcal{N}$, each sm line acquires an arrow to or from s. Thus,

$$n_s = \vec{n}_s + \overleftarrow{n}_s. \tag{11}$$

With eq. (10), the possibilities are



(12)

Thus, for n_s = odd:

$$\begin{cases} \vec{n}_s > \overleftarrow{n}_s & \text{(sink)}, \\ \vec{n}_s > \overleftarrow{n}_s & \text{(source)} \end{cases} \tag{13a}$$

with n_s even:

$$\begin{cases} \vec{n}_s > \overleftarrow{n}_s & \text{(sink) } (\therefore s = \text{ext}), \\ \vec{n}_s < \overleftarrow{n}_s & \text{(source) } (\therefore s = \text{ext}), \\ \vec{n}_s = \overleftarrow{n}_s & \text{(source and sink cancelled)} \\ & (\therefore s = \text{ internal species}) \end{cases} \tag{13b}$$

which proves the following theorem.

THEOREM 1

A species s with s-vertex s in $\mathcal{N}$ with degree n_s *odd* is an external species. If n_s is even, s can be an *internal* or *external* species.

From eqs. (11)–(13), we also have:

THEOREM 2

Without regard to the rest of $\mathcal{N}$, the possible types of an s of degree n_s are given by the partitions of n_s into two integers $(\vec{n}_s, \overleftarrow{n}_s)$.

For example, in eq. (12) with $n_s = 3$,

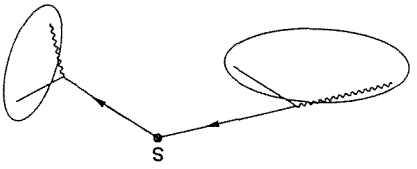
$3 = 3 + 0,\ 2 + 1,\ 1 + 2,\ 0 + 3.$

Which partition $(\vec{n}_s, \overleftarrow{n}_s)$ the s-vertex has, depends on the flows implied by the $\mathcal{N}$ consistent with the desired OVR [1–3].

With n_s even, the case of $n_s = 2$ deserves special mention due to its relation to laminar $\mathcal{N}$s,

$$n_s = 2 = 2 + 0,\ 1 + 1,\ 0 + 2.$$

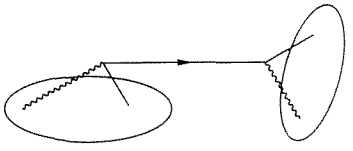
The $n_s = 2$ and $(\vec{n}_s, \bar{n}_s) = (1, 1)$ gives the simplest type of internal species:



(14)

Since the *same mole* is going in and out of s, in (14) the s can be removed, the line "unkinks", and this section of $\mathcal{N}$ becomes laminar.

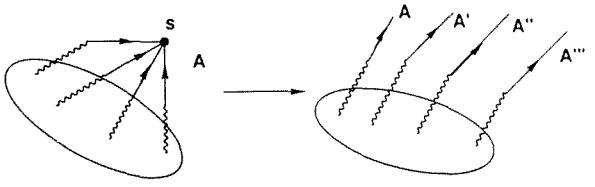
A laminar network is one in which all internal species are of the type $n_s = 2$; $(\vec{n}_s, \bar{n}_s) = (1, 1)$. Equation (14) and



(14′)

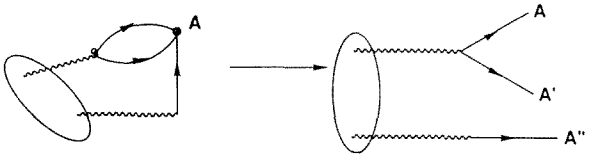
are topologically equivalent as stated in [1]. The use of (14′) instead of (14) leads to eqs. (1ℓ) and (3ℓ).

Further, some external species vertices ($\vec{n}_s \neq \bar{n}_s$) can also be removed, allowing some $\mathcal{N}$s to be treated as "*laminar*" although they are not "strictly laminar" (recall the definitions of these two in terms of the $\{\nu_i\}$ of the steps and of the OVR in [1]). These are the s with the n_s partitions $(\vec{n}_s, \bar{n}_s) = (n_s, 0)$ or $(0, n_s)$. For example,



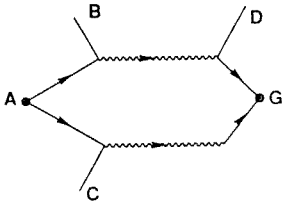
(15)

or, for example,



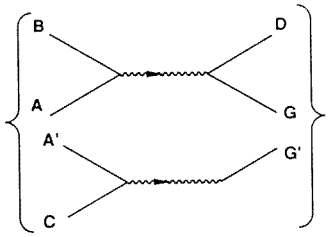
Such decoupling of s joints of lineblocks with the resulting change in the skeleton (e.g. ⬭ → ⋖) will not cause any difficulty in the stoichiometric aspects of the $\mathcal{N}$, nor in the rate laws aspects, as long as the $\mathcal{N}$ is used in connection with a definite OVR (the same assignment of "consistency arrows") [1,3].

If all the species vertices of $n_s > 1$ are of the full-external, i.e. (n_s, 0) or (0, n_s) type, that "quasi-turbulent" $\mathcal{N}$ (i.e. "laminar" but not "strictly laminar") may split into several disconnected laminar $\mathcal{N}$s, several independent OVRs. For example,



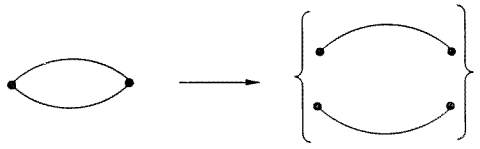
$$\text{for} \begin{cases} A + B \rightarrow D + G \\ A + C \rightarrow G \\ \hline \text{OVR:}\ \ 2A + B + C \Rightarrow D + 2G \end{cases} \qquad (16)$$

becomes



$$\begin{aligned} &\text{OVR}_1\text{: } A + B \Rightarrow D + G, \\ &\text{OVR}_2\text{: } A + C \Rightarrow G \end{aligned} \qquad (17)$$

with the resulting disconnection in the S:



(18)

(Note that eq. (18) does not occur for the cyclic assignment of reaction arrows, which corresponds, however, to a different OVR: B ⇒ D + C and also violates our rule regarding the type of the A and G vertices.)

In summary, the following types of "turbulent networks" are noted vis-a-vis the types of species vertices they contain (see section 4).

4. Types of networks with respect to their turbulent nature

Assume networks are initially drawn in full generality, i.e. displaying a species vertex s for each of the σ species {s}. Then the {$\mathcal{N}$} types are:

(I) Each external species has a species vertex of $n_s = 1$ with the partition $(\vec{n}_s, \overleftarrow{n}_s) = (1, 0)$ or $(0, 1)$.

Each internal species has a species vertex of $n_s = 2$ with the partition (1, 1). (Then they may be "unkinked"; eqs. (14) (14').) These $\{\mathcal{N}\}_{\mathrm{I}}$ are "*strictly laminar*" [1].

(II) Each external s-vertex has $n_s \geq 1$, but with the partition $(\vec{n}_s, \overleftarrow{n}_s) = (n_s, 0)$ or $(0, n_s)$.

Each internal s-vertex has $n_s = 2$ with (1, 1) (so they may be "unkinked"). These $\{\mathcal{N}\}_{\mathrm{II}}$ are "quasi-turbulent", using the terminology of [1]: "laminar" but not "strictly laminar".

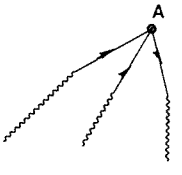
The external s-vertices of $\{\mathcal{N}\}_{\mathrm{II}}$ can be split; then $\{\mathcal{N}\}_{\mathrm{II}} \rightarrow \{\mathcal{N}_l\}$. This process splits some dot-points of the initial $\mathcal{S}$, then (IIa) either a single $\mathcal{S}'$ (one connected piece) still results, but with ring [1] number $r' < r$, or (IIb) $\mathcal{S}$ breaks up into several disconnected smaller skeletons such that $\mathcal{S} = \mathcal{S}'_1 \oplus \mathcal{S}'_2 + \ldots$. For (IIb), the initial $\mathrm{OVR} = \sum_i (\mathrm{OVR})_i$.

(III) All the external s-vertices are "laminar" as in cases (I) or (II) above, i.e. s_{ext}: $(n_s, 0)$ or $(0, n_s)$.

In general, any internal s has of course to have the partition $(\frac{1}{2} n_s^{\mathrm{int}}, \frac{1}{2} n_s^{\mathrm{int}})$, but here in type III, some $n_s^{\mathrm{int}} > 2$. Then the only turbulence of $\{\mathcal{N}\}_{\mathrm{III}}$ is in the internal species. Examples of such turbulent internal species can be seen in fig. 1.

We may call the $\{\mathcal{N}\}_{\mathrm{III}}$ "internally turbulent" networks.

(IV) Any network containing one or more species vertices such that $n_s > 2$ with an unequal partition $(\vec{n}_s \neq \overleftarrow{n}_s)$ and $\vec{n}_s, \overleftarrow{n}_s \neq 0$. Such a vertex is an external species one; however, one that also partly acts as internal. If furthermore $\vec{n}_s > \overleftarrow{n}_s$ (i.e. more s is generated than consumed at each pass through $\mathcal{N}$), then the network is "*autocatalytic*", $\{\mathcal{N}\}_{\mathrm{IVa}}$. For example,



(19)

If $\vec{n}_s < \overleftarrow{n}_s$ for all {s} with $(\vec{n}_s \neq \overleftarrow{n}_s)$, we shall put that $\mathcal{N}$ in type $\{\mathcal{N}\}_{\mathrm{IVb}}$.

Each s of the autocatalytic type can be classified further, depending on the topology of the full lineblock it occurs in (cf. the section on network piece generation below). Further, in type $\{\mathcal{N}\}_{\mathrm{IV}}$ the informal species may be laminar or turbulent.

Table 1 lists the possible types of species vertices and the resulting $\{\mathcal{N}\}$ according to types of $\{n_s\}$ values and partitions $\{(\vec{n}_s, \overleftarrow{n}_s)\}$.

k	$(\omega_{i(k)}, n_{i(k)})$	$\mathcal{N}$-lineblock	S-dot point
1	(1, 1)		i(1)
	(1,2)		
	(2, 2)		
2	(1, 2)		i(2)
	(1, 3)		
	(1, 4)		

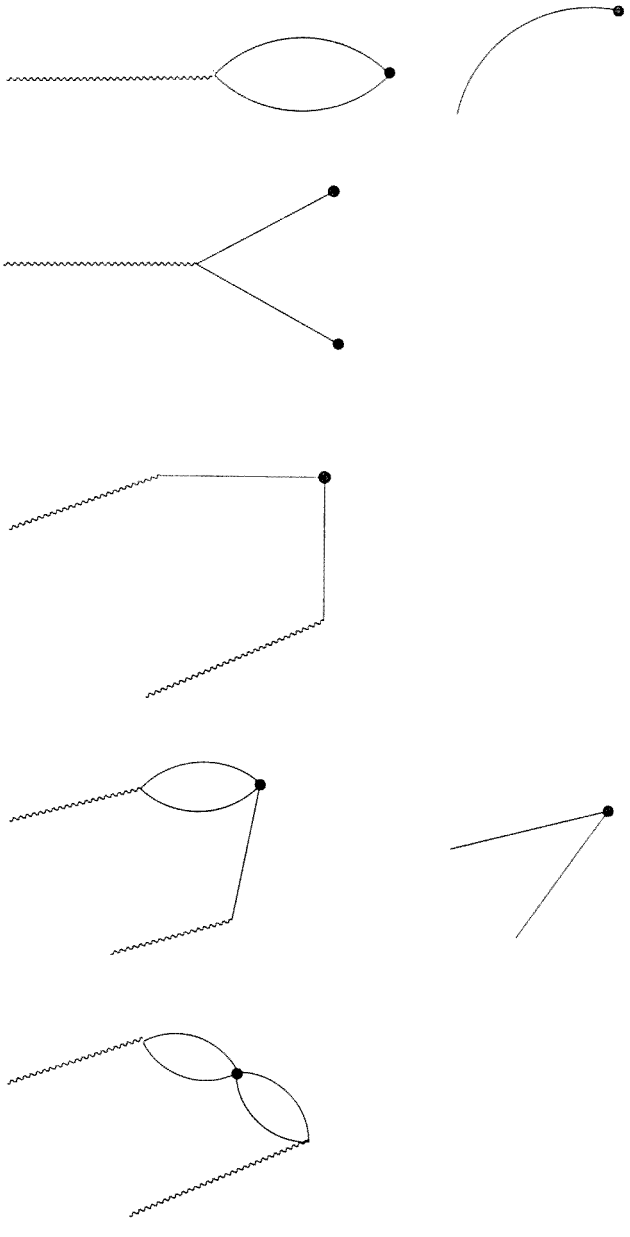


Fig. 1 (continues)

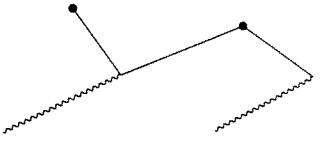
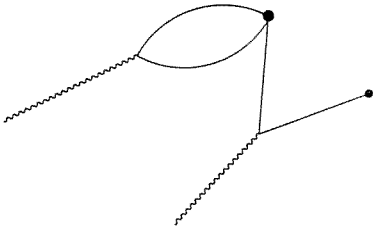
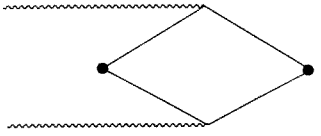
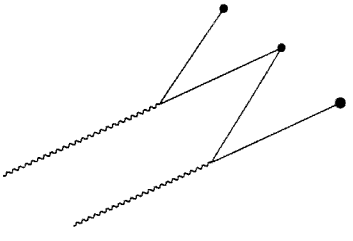
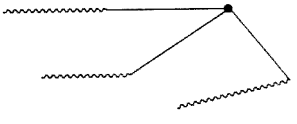
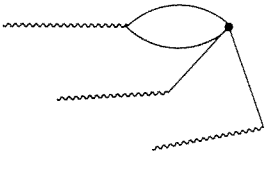
k	$(\omega_{i(k)}, n_{i(k)})$	$\mathcal{N}$-lineblock	$\mathcal{S}$-dot point
2 (cont'd)	(2, 3)		
	(2,4)		
	(3, 4)		
3	(1, 3)		 i(3)
	(1, 4)		

Fig. 1 (continues)

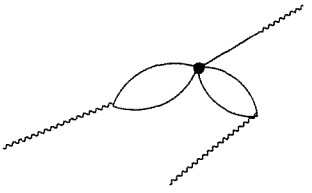
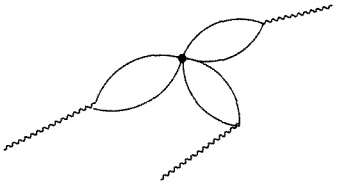
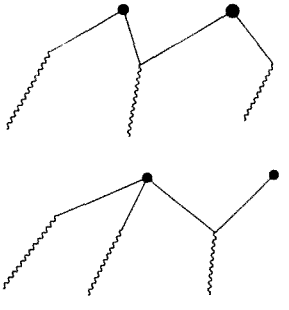
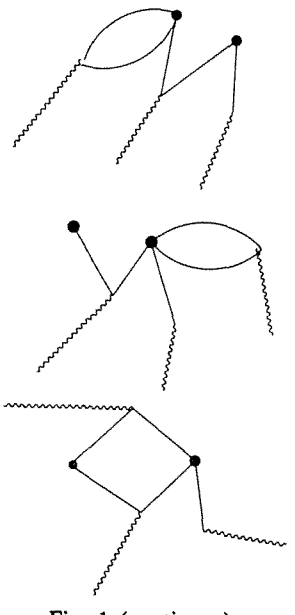
k	$(\omega_{i(k)}, n_{i(k)})$	$\mathcal{N}$-lineblock	$\mathcal{S}$-dot point
3 (cont'd)	(1, 5)		
	(1,6)		
	(2, 4)		
	(2, 5)		

Fig. 1 (continues)

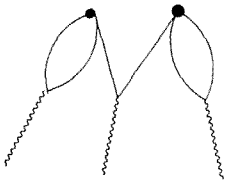
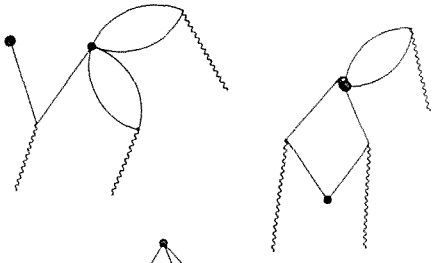
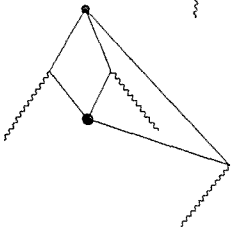
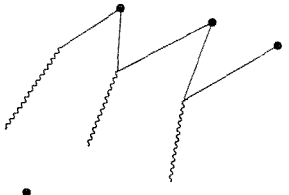
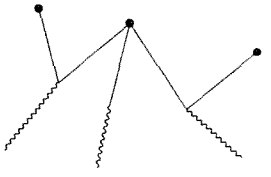
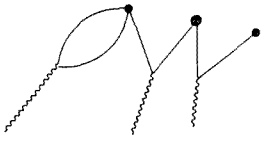
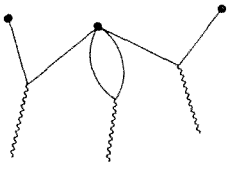
k	$(\omega_{i(k)}, n_{i(k)})$	$\mathcal{N}$-lineblock	$\mathcal{S}$-dot point
3 (cont'd)			
	(2, 6)		
			
	(3, 5)		
			
			
	(3, 6)		

Fig. 1 (continues)

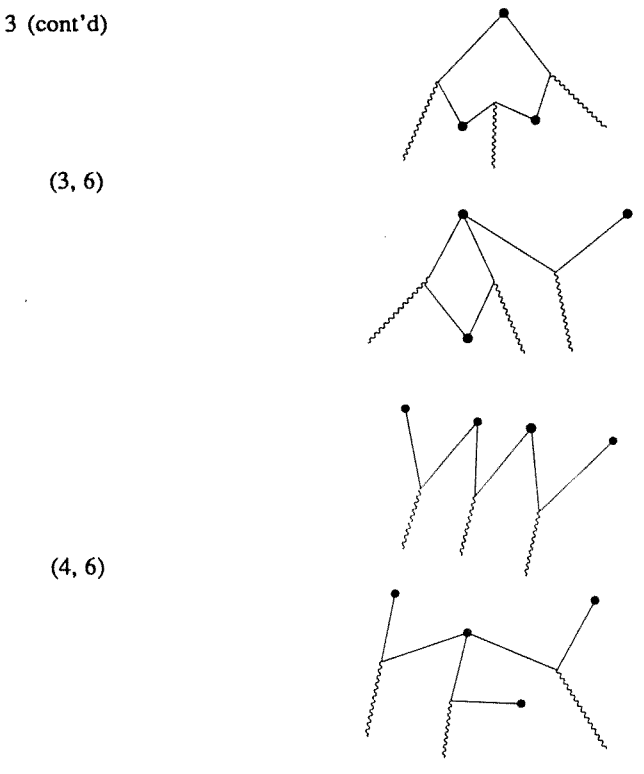


Fig. 1. Lineblock types possible for each double weighted dot-point i(k); $(\omega_{i(k)}, n_{i(k)})$.

Table 1

The possible types of species vertices and the resulting $\{\mathcal{N}\}$ network types.

Species vertex types in $\mathcal{N}$		Network type
each external species: $n_s = 1$; $(\vec{n}_s, \overleftarrow{n}_s) = (1, 0)$ or $(0, 1)$ each internal species: $n_s = 2$; $(\vec{n}_s, \overleftarrow{n}_s) = (1, 1)$		$\{\mathcal{N}\}_{\mathrm{I}}$; "strictly laminar"
each external species: $n_s \geq 1$; $(\vec{n}_s, \overleftarrow{n}_s) = (n_s, 0)$ or $(0, n_s)$ each internal species: $n_s = 2$; $(\vec{n}_s, \overleftarrow{n}_s) = (1, 1)$		$\{\mathcal{N}\}_{\mathrm{II}}$; "laminar"
each external species: $n_s \geq 1$; $(\vec{n}_s, \overleftarrow{n}_s) = (n_s, 0)$ or $(0, n_s)$ some internal species: $n_s > 2$; $(\vec{n}_s, \overleftarrow{n}_s) = (\frac{1}{2}n_s, \frac{1}{2}n_s)$		$\{\mathcal{N}\}_{\mathrm{III}}$; "internally turbulent"
one or more species vertices:		
$n_s > 2$, $\vec{n}_s \neq \overleftarrow{n}_s$, $\vec{n}_s, \overleftarrow{n}_s \neq 0$	$\vec{n}_s > \overleftarrow{n}_s$	$\{\mathcal{N}\}_{\mathrm{IVA}}$; autocatalytic
	$\vec{n}_s < \overleftarrow{n}_s$	$\{\mathcal{N}\}_{\mathrm{IVb}}$; self-inhibitory

5. Generation of all a priori turbulent networks/mechanisms

To generate all networks, turbulent and/or laminar, one starts with a look at the $\rho(\sigma)$ relations [1] to select the reasonable range of possible ρ and/or σ values. Then take a ρ. Obtain all the skeletons $\{S\}_\rho$ for this ρ. The $\rho = 1, 2, 3, 4$ skeletons are already tabulated in [1]. From here on, the procedure differs depending on whether one is treating just the laminar cases (already given in [1–4]) or the turbulent and laminar cases both.

For laminar only, eqs. (1ℓ), (3ℓ) apply. Thus, each dot-point i of S acquires a weight ω_i ($=n_i$). The possible weights $\{\omega_i\}$ for all the dot-points of S are found. For just laminar, the possible [1,2] $\omega_i \in \{k-1, k, \bar{k}, k+1\}$, where k is the star-value ("degree") of that dot-point i in the particular S.

Possible weight combinations are assigned to the S, producing a number of *weighted skeletons* $\{S_\omega\}$. For each S_ω, each dot-point i of weight ω_i is "opened up" into a lineblock which gives that piece of the network. The pieces are joined, preserving the S.

Thus,

laminar only: $\boxed{\rho \rightarrow S \rightarrow S_\omega \rightarrow \mathcal{N}_\ell}$ (20ℓ)

For general $\{\mathcal{N}\}$, turbulent and laminar allowed, we have eqs. (1′) and (3′). When turbulent, a lineblock has more species-mole (sm) lines n_i than the number of species ω_i.

Thus, now the dot-points of S have double weights (ω_i, n_i) which, when "opened up", give the turbulent network pieces.

6. Doubly weighted skeletons for general networks (turbulent and laminar both)

We shall develop the method for general networks (ℓ and t) in a way that covers all the types of table 1. The laminar method [2,3] is then recovered as a special case.

For the general method, the networks are considered drawn with a species vertex for each species, even for the internal laminar ones. That is, $n_s = 2$; $(\vec{n}_s, \bar{n}_s) = 1, 1$ vertices are drawn before any laminar "unkinking" (eqs. (14), (14′)). Then:

(1) Each dot-point i of star value k of a skeleton has a set of possible double weights $\{(\omega_{i(k)}, n_{i(k)})\}$.

(2) Each double-weight assignment gives rise to one or several types of general lineblock (a turbulent or laminar network piece).

(3) For turbulent lineblocks as well as for the "kinked laminar" ones, $\omega_i < n_i$. After the kinked laminar species vertices are identified (by their $n_s = 2$; $(\vec{n}_s, \bar{n}_s) = (1, 1)$

nature) and are unkinked (eq. (14) → (14′)), if the network piece does not have any turbulence left, one recovers the usual laminar case [2,3] with $\omega_i = n_i$.

(4) The possible double weights of $i(k)$ are given by theorems 3(a) and 3(b).

THEOREM 3(a)

The double weights $\{(\omega_{i(k)}, n_{i(k)})\}$ of a skeletal dot-point i of star value ("degree") k, denoted i(k) are sets of integers satisfying

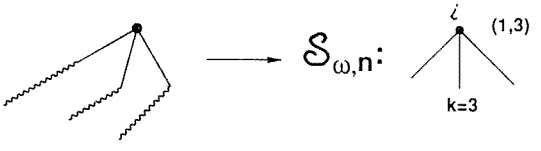
$$\begin{cases} 1 \le \omega_{i(k)} \le (k+1), & (21a) \\ k \le n_{i(k)} \le 2k & (21b) \end{cases}$$

and

$$\begin{cases} \omega_{i(k)} < n_{i(k)} \quad \text{for } k \ge 1, & (21c) \\ \omega_{i(k)} = n_{i(k)} \quad \text{for } k = 1. & (21c') \end{cases}$$

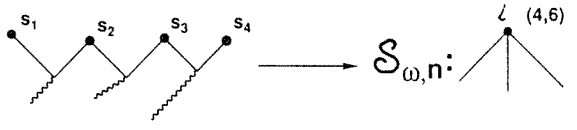
Proof

Since reaction steps were uni- or bi- only, either one or two sm lines come out of each of the k-wiggle lines. Hence, eq. (21b). A minimum of k sm lines are needed to obtain a connected lineblock. This min $n_{i(k)}$ is for $\omega_{i(k)} = 1$, which is the min ω.



(22)

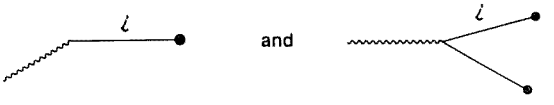
Max $n_{i(k)}$ is $2k$. The max $\omega_{i(k)}$ for $n_{i(k)} = 2k$ is obtained for the kinked laminar case (as noted in [1]):



(23)

$\{[\omega_{i(k)}]_{max} = k + 1;\ [n_{i(k)}]_{max} = 2k\}$. Any merging of species vertices starting from the kinked laminar decrease ω_i and/or n_i.

Equation (21c′) holds only for $k = 1$ for the lineblocks



$n_{i(1)} = 1$ $\quad$ $n_{i(1)} = 2$

$\omega_{i(1)} = 1$ $\quad$ $\omega_{i(1)} = 2$

otherwise all other lineblocks including $k = 1$; $(\omega, n) = (1, 2)$ and $k > 1$ turbulent ones and kinked laminar ones obey eq. (21c).

Comment 1. Equations (21) reduce to the inequalities of the usual ("unkinked") laminar ones [1–3] after the unkinking and $\omega_{i(k)} = n_{i(k)}$. $(k - 1 \leq n_{i(k)} \leq k + 1)$.

Comment 2. Theorem 3(a) is necessary but not sufficient for a set (ω, n) to be a double weight, i.e. to correspond to a connected lineblock. There is a further restrictive relation between ω and n. Hence, we need:

THEOREM 3(b)

A pair of integers satisfying eqs. (21) is a double weight if also

$$\boxed{n_{i(k)} - \omega_{i(k)} - k + 1 \geq 0.} \tag{24}$$

Proof

Lineblock i is a connected graph of $(\omega_{i(k)} + k)$ vertices and $n_{i(k)}$ lines. They are general, may be planar or non-planar. Therefore, they satisfy a topological relation like the one for skeletons $(\gamma = \rho - r + 1)$ proved in that context in [1]. The corresponding ring number, here $R = n_{i(k)} - \omega_{i(k)} - k + 1$ (eq. (24)), is non-negative. □

Comment 3. For an arbitrary 3D graph $R \in \{0, 1, 2, \ldots, n_{i(k)}\}$; however, R may not reach the higher values if the graph is a lineblock due to eqs. (21) limiting that type of graph. The k vertices can be only of degree 1 (k_1 of them) or 2 (k_2 of them). Thus, $k = k_1 + k_2$; $n_{i(k)} = k_1 + 2k_2$. Hence, $k_1 = 2k - n_{i(k)}$ and $k_2 = n_{i(k)} - k$. Then

$$R = k_2 - \omega_{i(k)} + 1, \tag{25a}$$

$$k \geq k_2 \geq 0 \tag{25b}$$

for a lineblock.

Comment 4. We see that eqs. (21c) and (21c′) also follow from (24). We could have omitted them, but wrote them to make the finding of double weights more convenient.

PRACTICAL COROLLARY TO THEOREMS 3(a) AND 3(b)

Skeletal double weights corresponding to lineblocks are the integer solutions of eqs. (21a), (21b) and (24).

Double weights for skeletal dot-points $\{i(k)\}$ are shown in table 2 for $k \in \{1, 2, 3\}$.

The lineblocks (turbulent and "kinked" representation of laminar) these double weights give rise to are shown in fig. 1.

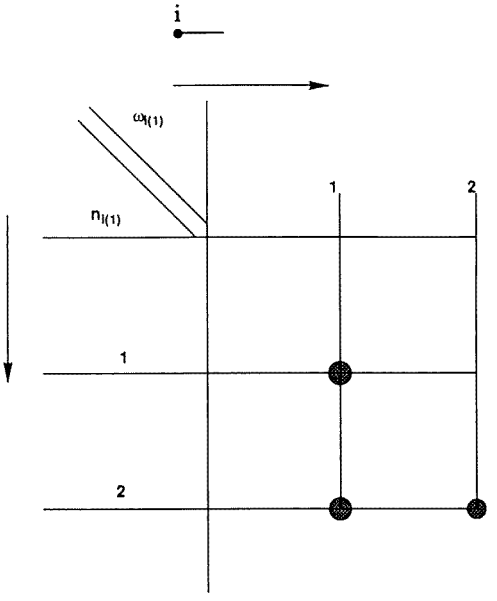
Table 2

Double weights for each skeletal dot-point i(k); $k \in \{1, 2, 3\}$.

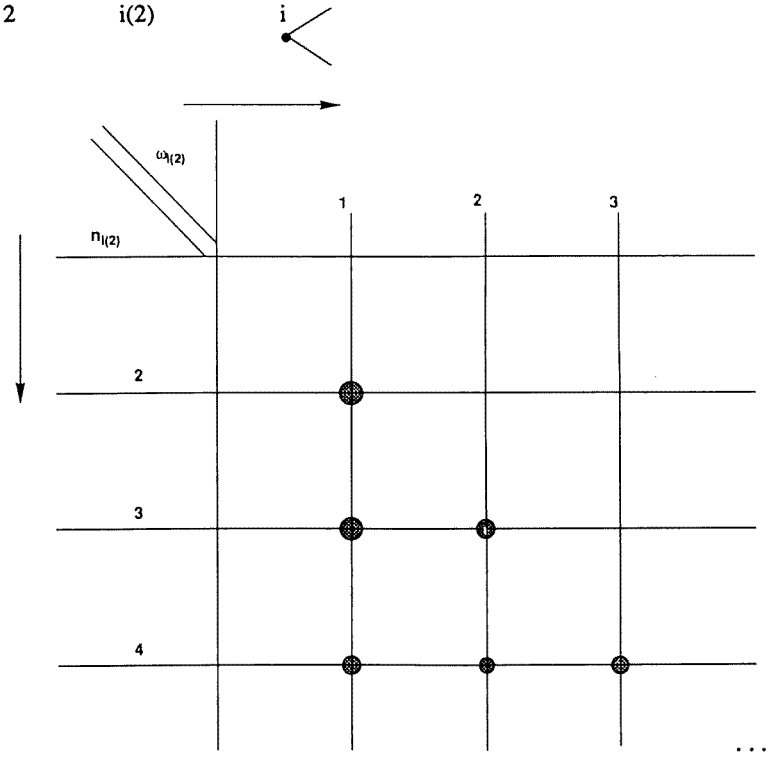
Skeletal dot-point: i(1)

Double weights $(\omega_{i(1)}, n_{i(1)})$

$k = 1$

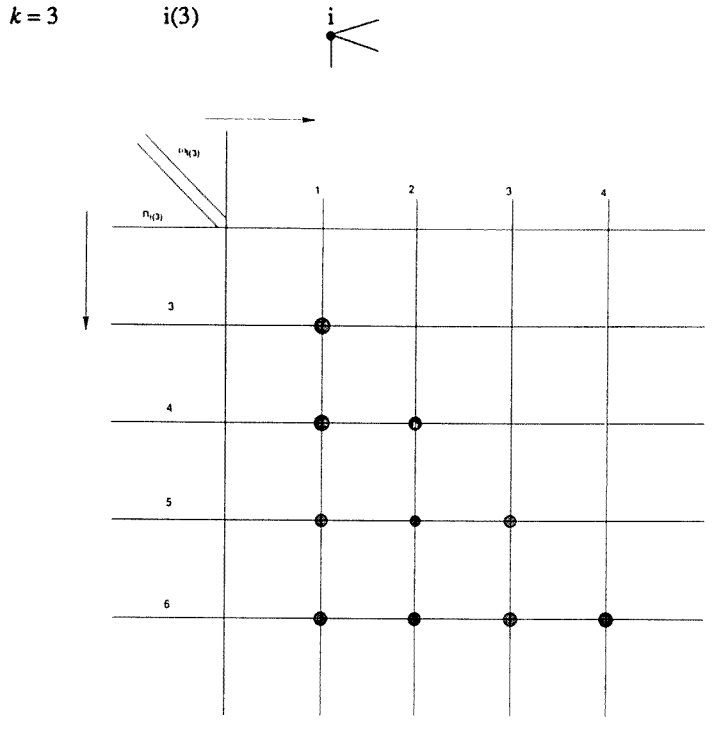


$k = 2$ i(2)



. . . continues

Table 2 (continued)



7. Topological classification of general lineblocks

A topological invariant[1] for a connected lineblock is its number of rings R given by eq. (25a). Since given the k value of i(k), the double weight $(\omega_{i(k)}, n_{i(k)})$ determines the R value, the lineblocks of an i(k) arising out of the same double weight are in the same R class; they have the same number of rings.

This is sufficient, but not necessary. The necessary and sufficient condition for arbitrary lineblocks to be in the same R class is that they have the same value of $((n_{i(k)} - \omega_{i(k)} - k)$ (theorem 3).

The following sets of numbers specify the same set of lineblocks in fig. 1:

$$\left.\begin{array}{c}(k,\ \omega_i,\ n_i)\\ \text{or}\\ (k_1,\ k_2,\ \omega_i)\\ \text{or}\\ (k,\ R,\ \omega_i)\end{array}\right\} \tag{26}$$

[1] In the sense of invariance under elementary subdivisions of the lines of a graph.

However, there are usually still several lineblocks in each such set (cf. e.g. fig. 1; $k = 3$; $(\omega_i, n_i) = (2, 5)$). Each lineblock within such a set is further specified down to a unique one, by a sequential list of the vertex degrees; for example, the ones in fig. 1 in the lineblocks set $\{(k, \omega, n) = (3, 3, 6)\}$ are given by the lists $(2, \overline{3}, 2, \overline{2}, 2, \overline{1})$, $(\overline{1}, 2, \overline{4}, 2, 2, \overline{1})$ $(2, \overline{2}, 2, \overline{2}, 2, \overline{2})$, and $(2, \overline{2}, 2, \overline{3}, 2, \overline{1})$, where the barred ones are the species vertex sm line degree (or "turbulence index" of s) [1] n_s values. The non-barred ones are the "uni–bi degree" of each reaction vertex. Each sequence is composed of two R-consistent partitions of the integer $n_{i(k)}$.

We shall not go into the combinatorial aspects further here, but assume that the n_s values of each species s can be read off on the lineblocks drawn for given $(k, \omega_{i(k)}, n_{i(k)})$ or (k, R, ω) as in fig. 1.

8. Combining the lineblocks according to a given skeleton S to obtain its full networks

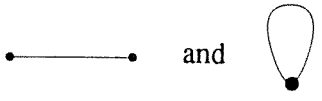
The lineblocks in fig. 1 (for $k = 1, 2, 3$) and similarly higher k ones can now be combined to obtain the general networks of ρ steps and a skeleton S of ρ, noting that [1]

$$\sum_{i\geq 1}^{\gamma} k_i = 2\rho \tag{27}$$

for γ lineblocks. The wiggle lines of the same or different lineblocks are spliced according to the S_ρ.

(A) THE $\rho = 1$, ALL ONE-STEP NETWORK TYPES

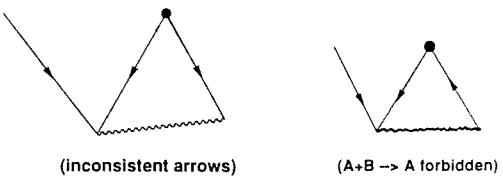
For $\rho = 1$, the $\{S\}$ are [1]



(a) (b) (28)

The $k = 1$ lineblocks, $m = 3$ of them, give $m(m + 1)/2$, i.e. six networks for S_{1a}. These are as the elementary reaction types [3], including the turbulent versions (fig. 2(a)). Each $\mathcal{N}$ implies also the reverse reaction. The only turbulent lineblocks from S_{1a} are of the homolytic type.

Since $\gamma = 1$ and $k = 2$ in the S_{1b}, the other $\rho = 1$ networks are given from the splicing of the $k = 2$ lineblocks of fig. 1. Thus, one has the ones in fig. 2(b). The first is the simplest case of autocatalysis, the second of catalysis. Other lineblocks of $k = 2$ lead to either trivial (e.g. A $\rightarrow$ A), or inconsistent [1,3] or forbidden networks such as

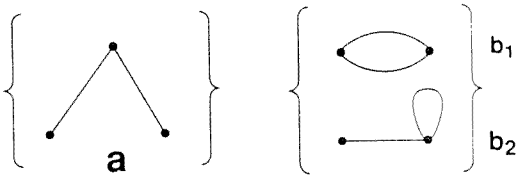


(29)

Trivial or forbidden ones arise when there are none or only one external species vertex in the network; hence, they are easily recognized.

(B) THE $\rho = 2$ GENERAL NETWORKS

The $\rho = 2$ skeletons are of ring numbers [1] $r = 0$, 1 and 2. All laminar networks and mechanisms possible with $\rho = 2$ were derived earlier [2,3]. From the general lineblocks in fig. 1, we can obtain all the general ones, including turbulent, for the $r = 0$ and 1 skeletons



(30)

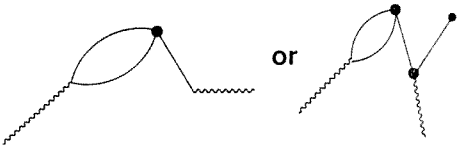
For the one of $r = 2$,  , we would also need the $k = 4$ lineblocks.

(i) *Networks of* S_{2a},  .

Six pairs of $k = 1$ lineblocks are combined, one for each end of eq. (30a), with one of the $k = 2$ cases of fig. 2.

The $k = 2$, double weights $(\omega_i, n_i) \in \{(1, 2), (2, 3), (3, 4)\}$, together with any of the $k = 1$ ones give the laminar networks already listed in refs. [2,3].

The $k = 2$, $\{(1, 4), (2, 4)_2\}$ yield either essentially a laminar mechanism [2] (for the species vertices internal, i.e. $\rightarrow 2A \rightarrow$ or $\rightarrow A + B \rightarrow$), or ones that split into two $\rho = 1$ laminar reactions at the $k = 2$ turbulent species vertex (for $\{s\}$ external, i.e. ($\rightarrow 2A$ and $\rightarrow 2B$) or ($\rightarrow A + B$ and $\rightarrow A + B$). The remaining ones are $k = 2$; $(\omega_i, n_i) \in \{(1, 3), (2, 4)_1\}$. These



(31)

are inserted into any one of the S_{1a} networks of fig. 2. For example,

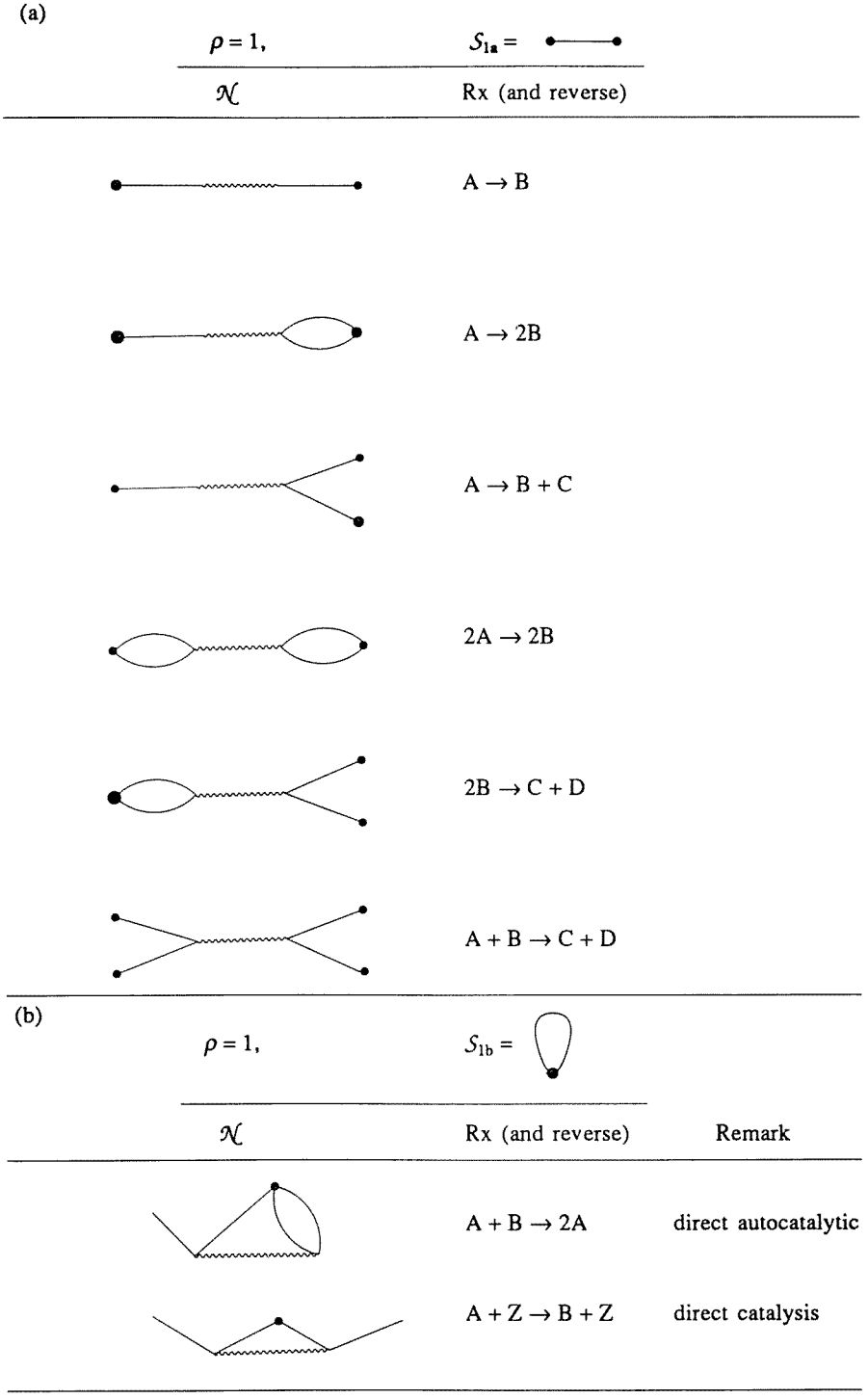
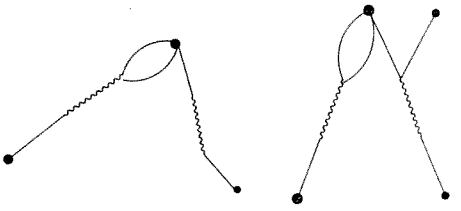


Fig. 2. One-step general network types.



(32)

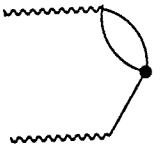
$$\begin{cases} B \rightarrow 2A \\ A \rightarrow C \end{cases} \quad \text{or} \quad \begin{cases} B \rightarrow 2A \\ A + C \rightarrow D \end{cases}$$

If the A-vertex in these were internal, i.e. $(\vec{n}_s, \overleftarrow{n}_s) = (3, 0)$ or $(0, 3)$, eqs. (32) too would split into two $\rho = 1$ reactions each.

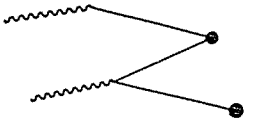
(ii) *Networks of* $\mathcal{S}_{2b_1}$, 

For S =  , the $m = 7$ lineblocks of $k = 2$ would splice into more than $m(m + 1)/2 = 49$ networks; however, most of these are already obtained or are trivial (cf. below). Very few new ones arise.

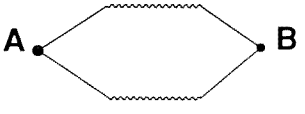
An asymmetric lineblock like



can combine two ways with another asymmetric lineblock like



However, many of these $\mathcal{N}$ would split into separate OVRs or be essentially the laminar ones already given [2,3]. For example

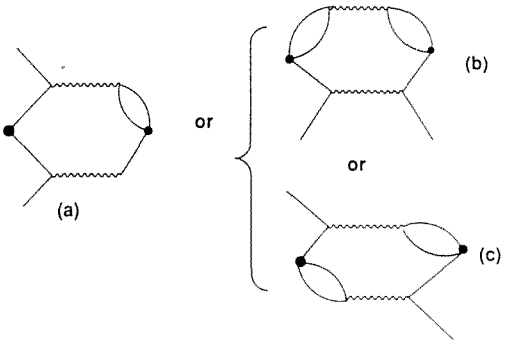


(31′)

is simply $2(A \rightarrow B)$ or the cycle $A \rightleftarrows B$ $(A \rightarrow B \rightarrow A)$.

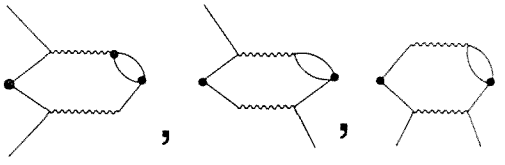
The $k = 2$, $\{(1, 2), (2, 3), (2, 4)_2, (3, 4)\}$ give the laminars [2,3], the $k = 2$ (1, 4) homolytic variant of $(2, 4)_2$.

The only new $\{\mathcal{N}\}$ are again combinations of (1, 3) or $(2, 4)_1$ with any of the seven $k = 2$ lineblocks. So again there are very few new network types (fig. 3) with $\rho = 2$, S_{2b_1}, after the inconsistent or forbidden ones too are eliminated. The new ones are all of the types



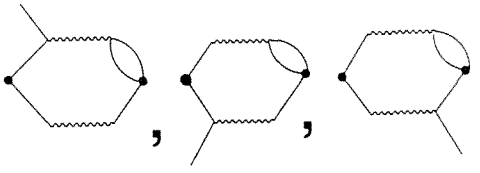
(33)

and with the single external lines removed one at a time to obtain additional $\mathcal{N}$s which correspond to different overall reactions, OVR. For example, from eq. (33a) one also obtains



(33a′)

and



(33a″)

The last external line cannot be removed since there must be at least two external species in a full network.

Clearly, not all the possibilities in deriving all networks need be listed; the maximal external lines $\mathcal{N}$s like in eqs. (33a)–(33c) suffice. If an OVR type is specified, then again just the more stripped variants of the "*complete* $\mathcal{N}$", i.e. the $\mathcal{N}$ with all its reaction vertices saturated as bimolecular, consistent with the OVR need be found.

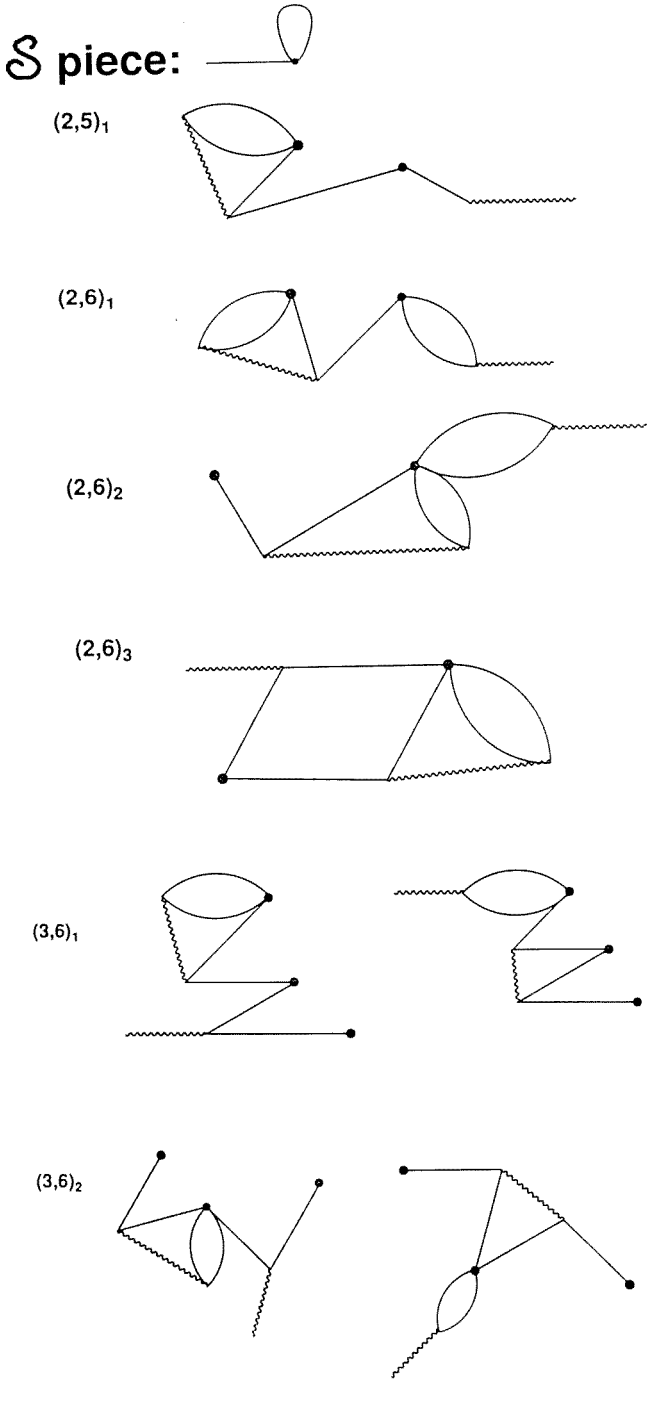


Fig. 3. The $k = 3$, turbulent network pieces.

The mechanisms of eqs. (33) are

$$\left.\begin{array}{l} A+B \rightarrow C+X \\ \underline{X+D \rightarrow \ 2A} \end{array}\right\} \qquad (\mathcal{M}33a)$$

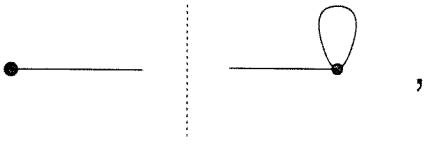
$$\left.\begin{array}{l} A+B \rightarrow C+D \\ \underline{2D \ \rightarrow \ 2A} \end{array}\right\} \qquad (\mathcal{M}33b)$$

$$\left.\begin{array}{l} A+B \rightarrow 2C \\ \underline{C+D \rightarrow 2A} \end{array}\right\} \qquad (\mathcal{M}33c)$$

Others result from the omission of some of the species of $n_s = 1$, $(\vec{n}_s, \overleftarrow{n}_s) = (1, 0)$ or $(0, 1)$ from eqs. $(\mathcal{M}33a)-(\mathcal{M}33c)$.

(iii) *Networks of* S_{2b_2}, 

These can be obtaines from the $k = 3$ lineblocks of fig. 1 by first splicing two wiggle lines at a time to each other to form the skeletal piece

 , (34)

then joining the remaining one to a $k = 1$ one from fig. 1.

Those $k = 3$ lineblocks containing at most $n_s = 2$ vertices (s) would again give the laminar $\mathcal{N}$ [2,3].

The  type network pieces containing the triangular cycle  with only one s-vertex on it or attached to it, i.e. $(\omega_i, n_i) \in \{(1, 4), (1, 5), (1, 6), (2, 4), (2, 5)_2, (2, 5)_3, (2, 6)_4\}$ are trivial or inconsistent or forbidden.

The remaining not fully laminar ones are shown in fig. 3.

We see in fig. 3 that the only network pieces that do not split up, or reduce to laminar, or become invalid or trivial, are the ones containing the autocatalytic species vertex s.

Table 3

The $\rho = 2$, $S =$ turbulent mechanisms[a] (cf. text).

$(2, 5)_1$:	$\left\{\begin{array}{r} A \rightarrow X \\ X + B \rightarrow 2B \\ \hline OVR : A \Rightarrow B \end{array}\right.$
$(2, 6)_1$:	$\left\{\begin{array}{r} A \rightarrow 2B \\ B + C \rightarrow 2C \\ \hline OVR : A \Rightarrow B + C \end{array}\right.$
$(2, 6)_2$:	$\left\{\begin{array}{r} C + B \rightarrow 2B \\ 2B \rightarrow A \\ \hline OVR : C + B \Rightarrow A \end{array}\right.$
$(2, 6)_3$:	$\left\{\begin{array}{r} A \rightarrow B + C \\ B + C \rightarrow 2B \\ \hline OVR : A \Rightarrow 2B \end{array}\right.$
$(3, 6)_1$:	$\left\{\begin{array}{r} A \rightarrow B + C \\ B + D \rightarrow 2D \\ \hline OVR : A \Rightarrow C + D \end{array}\right.$; $\left\{\begin{array}{r} A \rightarrow 2B \\ B + C \rightarrow C + D \\ \hline OVR : A \Rightarrow B + D \end{array}\right.$
$(3, 6)_2$:	$\left\{\begin{array}{r} A \rightarrow B + C \\ D + B \rightarrow 2B \\ \hline OVR : A + D \Rightarrow 2B + C \end{array}\right.$; $\left\{\begin{array}{r} A \rightarrow 2B \\ B + C \rightarrow B + D \\ \hline OVR : A + C \Rightarrow 2B + D \end{array}\right.$

[a] Additional ones are obtained by adding another external species, say G to A, in all of the above (corresponding to a $k = 1$; (2, 2) piece instead of $k = 1$; (1, 1) piece).

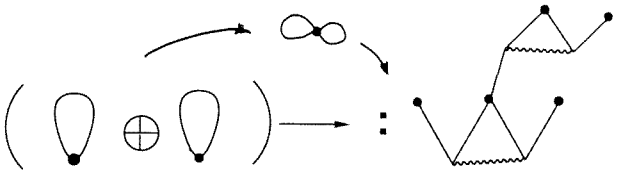
Splicing the ones of fig. 3 with the $k = 1$ lineblocks of fig. 1, we obtain the new $\rho = 2$, $S =$  networks. The maximal "complete networks" are with $k = 1$, (2, 2); the minimal with $k = 1$, (1, 1). The nine such mechanisms that result from fig. 3 are shown in table 3.

The mechanisms in table 3 lead to a number of OVR reaction types, i.e. $A \Rightarrow B$, $A \Rightarrow B + C$, $A \Rightarrow 2B$, $A + D \Rightarrow C + 2B$, and if some steps are reversed, also $A + C \Rightarrow 3B$ and $A + B \Rightarrow 2C$ (and reverse OVRs of all of these). We can also add another species , say G to A, to obtain additional ones in table 3 ($k = 1$, (2, 2)).

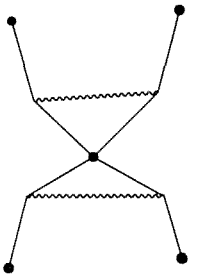
Combining the $\rho = 2$ results of the previous sections with the laminar ones obtained earlier [2,3], we have the general $\rho = 2$ mechanisms including homolytic, autocatalytic and the catalytic ones.

(iv) *The* S: ∞

Here, we skip the $\rho = r$, $S =$ ∞ turbulent ones, which can be worked out in a similar manner if desired. The systematic way to be sure to obtain all of the $\{\mathcal{N}\}$ of ∞ would be by splicing the $k = 4$ ends of each ( ) lineblock in $\binom{4}{2} = 6$ ways. A quicker way would seem to be by merging the lineblock of subskeletons, e.g. (⊕) = { ∞ }. Since there are only two lineblock types for in fig. 2, those mergers are easily worked out. They include (i): the turbulent ones from two catalyses:

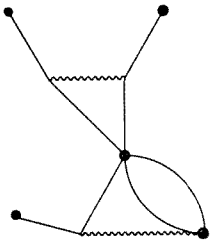


(35a)

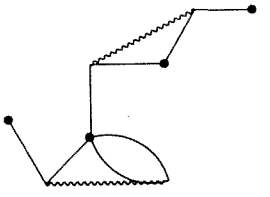


(35b)

(ii): from a catalysis and an autocatalysis:

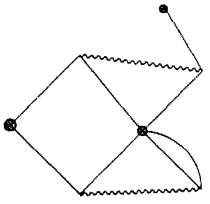


(36a)



(36b)

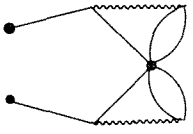
(36c)



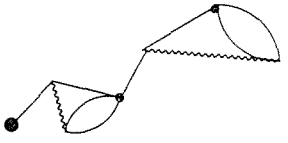
(36d)

(iii): from two autocatalyses:

(37a)



(37b)

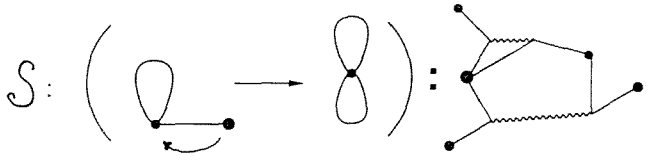


(37c)

In such mergers, the final $\mathcal{N}$ must have at least two external vertices (e.g. of $n_s = 1$, $n_s = 3, \ldots$ etc.)

The above $\{\mathcal{N}\}$ are easily interpreted; for example, in eq. (36b), the autocatalyst of a step is also reactant for another catalytic step, etc.

However, the ($\oplus$) merger does not generate all of the $\mathcal{S}=$ networks. For example,



$$\begin{array}{l} A + B \rightarrow C + D \\ D + B \rightarrow G + B \\ \hline OVR : A + B \Rightarrow C \end{array} \qquad (38)$$

results from $\mathcal{S}=$ by merging the end lineblock into the $k = 3$ one.

By such methods, i.e. splicings of k-lines or merging of sub-skeleton lineblocks, one can obtain the $\mathcal{N}$s. We shall not belabor these points further here since, if one

has a desired OVR type and/or number of catalytic, etc. rings r, the generation of $\{\mathcal{N}\}$ is much simplified to a lot fewer possibilities. The above systematic ways, however, could be used to generate an atlas of all the general networks/mechanisms of $\rho \in \{1, 2, 3, 4, \ldots\}$ ($r = 0, 1, 2, \ldots, \rho$) and of the possible OVR types. Such an atlas would serve as a handy source of the a priori possibilities for mechanistic studies.

9. Generating topologically equivalent networks of increasing number of reaction steps, ρ

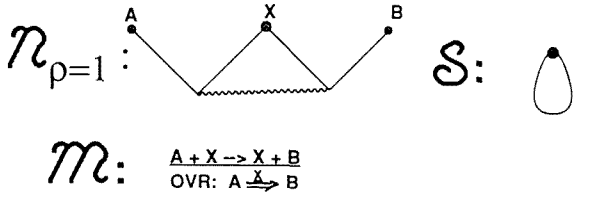
A chemically important characterization of a network/mechanism is by the types of external and of internal species vertices (section 3; n_s; $(\vec{n}_s, \overleftarrow{n}_s)$) it has and by the number of catalytic and/or autocatalytic cycles it contains. Topologically, the number of such cycles is related to the number of rings r in the skeleton and the number of homolytic or autocatalytic rings R in the lineblocks.

As we have noted, r is a topological invariant of skeletons under elementary subdivisions of (~~~) reaction step (rx) lines [1]. One may also carry out proper subdivisions (cf. below) of species-mole (sm) lines of the network $\mathcal{N}$. A theorem is stated and proved elsewhere [11] that the composite ring number $(r + R)$ is a topological invariant of the network under either or both typed of line (rx and sm) subdivisions.

Based on this result, and starting from a basic, smallest ρ, network with the desired features (types of species vertices, number of catalytic, autocatalytic cycles, . . .), successively larger networks of $\rho + 1$, $\rho + 2$, . . . are easily generated, retaining the desired features. This also allows the accommodation of richer overall reaction types (OVR) when the smallest ρ network cannot retain the desired features (cf. below).

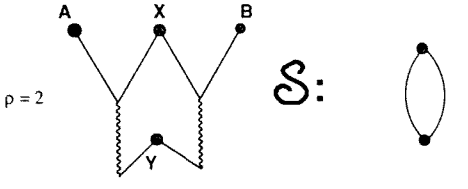
(A) SUCCESSIVELY LARGER CATALYTIC NETWORKS

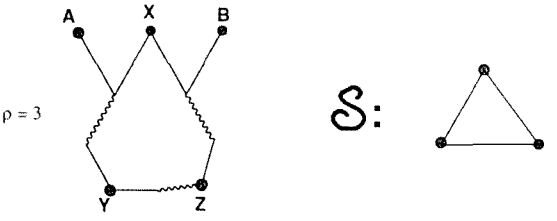
Take the simplest network of a single catalytic cycle:



(39)

(i) *Successive rx line subdivisions preserve* $r = 1$





(40b)

and so on for $\rho = 4, 5, \ldots$.

On the $\mathcal{N}$, rx line elementary subdivision give:

(41a)

corresponding to subdivisions on the $\mathcal{S}$:

(41b)

We see in eqs. (40) that two-step, three-step, ... single chain reactions are topologically, as well as in their overall chemical action, equivalent to a single direct catalyst, eq. (39$\mathcal{M}$). The action of chain steps mimicking a catalyst was noted in [1] and led to the notion of "generalized catalyst".

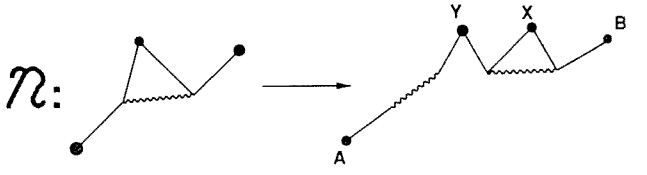
(ii) *sm line subdivision on the* $\mathcal{N}$

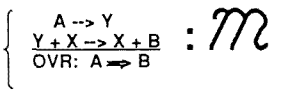
Proper sm divisions pertinent to $\mathcal{N}$ do:

(42)

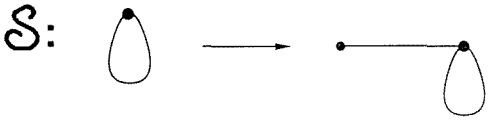
If on the composite (sm–rx) ring of an $\mathcal{N}$, sm subdivisions are equivalent to rx subdivisions; the r number is still preserved.

If on the external sm line, e.g.





(43)

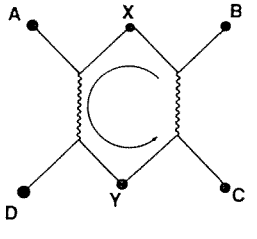


again r is preserved, but a tree is attached to the skeleton.

Thus, rx and/or sm line subdivisions generate a whole class of $\{\rho \rightarrow \rho + 1\}_m$ networks with a single "generalized catalyst" (X → one multi-step chain) [1] and with the same OVR (here, A ⇒ B).

(iii) *One generalized catalyst, larger OVRs, the "complete OVR"*

With rx steps bi- or uni-molecular the prototype single catalysis network, eq. (39) cannot accommodate an OVR larger than A ⇒ B. The successive $\{\rho \rightarrow \rho + 1\}$ generation allows one to obtain larger OVRs. Each time an $\mathcal{N}$ as in eqs. (40) is obtained, new reaction line vertices can be saturated with additional external sm lines ($n_s = 1$). To obtain all the possible OVR types, add all the $n_s = 1$ external sm lines to wherever there is room, obtaining in each case the *"complete network"* (thus defined). For example, the complete $\mathcal{N}$ of eq. (40a) is



(40a′)

The counterclockwise flow gives the "complete OVR" for this $\rho = 1$, $r = 1$, single catalytic $\mathcal{N}$:

complete OVR: A + C ⇒ D + B (40a″)

This is the largest OVR possible. Eliminating some external lines consistent with the flow gives, in the general case, the smallest OVRs possible.

Similarly, eq. (40b) would give the

$$\text{complete OVR: } A + C + F \Rightarrow G + D + B, \tag{40b}$$

but also the set of OVR types $A + C + F \Rightarrow G + D, \ldots$ down to $A \Rightarrow B$.

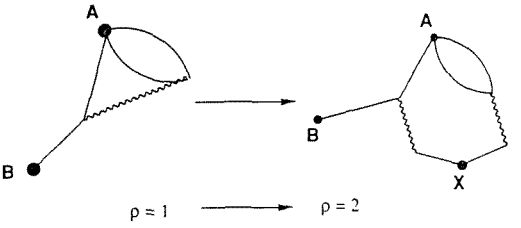
Thus, if one is attempting to generate all the $\{\mathcal{N}\}$ of a given OVR type, say $A + C \Rightarrow D + B$, and a single catalytic cycle and of varying ρ size, first the smallest ρ-$\mathcal{N}$, in this example, eq. (40a′) is found. Then, from the larger ρ *complete networks*, external sm lines are eliminated in the possible ways down to the desired OVR.

(B) GENERATION OF AUTOCATALYTIC CYCLES OF SUCCESSIVELY LARGER ρ

For one autocatalyst, the smallest, prototype network is that of $\rho = 1$, $A + B \rightarrow 2A$ (fig. 2).

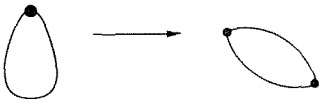
Reaction (~~) subdivisions again generate the larger ρ, the same $r = 1$ single "generalized autocatalyst" cycle mechanisms,

$\mathcal{N}$:



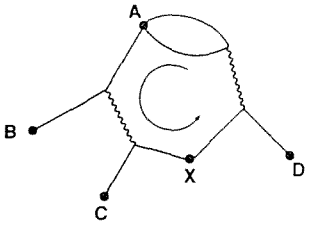
(44a)

$\mathcal{S}$:



and the complete network ($c\mathcal{N}$),

$c\mathcal{N}$:



(44b)

$c\mathcal{M}$:

$$\begin{array}{rcl} A + B & \rightarrow & X + C \\ D + X & \rightarrow & 2A \\ \hline \text{cOVR}: D + B & \overset{A}{\Rightarrow} & C + A \end{array}$$

and so on for $\rho = 3, 4, \ldots$.

Now, topologically $\mathcal{N}$-equivalent (but not necessarily $\mathcal{S}$-equivalent) larger ρ networks can also be generated by sm-line subdivisions. These will preserve the r + R number. In eq. (44), $r = 1$, $R = 1$, $r + R = 2$.

All the $\rho = 2$ complete networks $\{c\mathcal{N}\}$ arising from the $\rho = 1$ autocatalytic prototype (left-hand side of eq. (44a)) are shown in fig. 4 as obtained by sm-line subdivisions, eq. (42), as well as by rx-line subdivision, and their mechanisms $\{\mathcal{M}\}$.

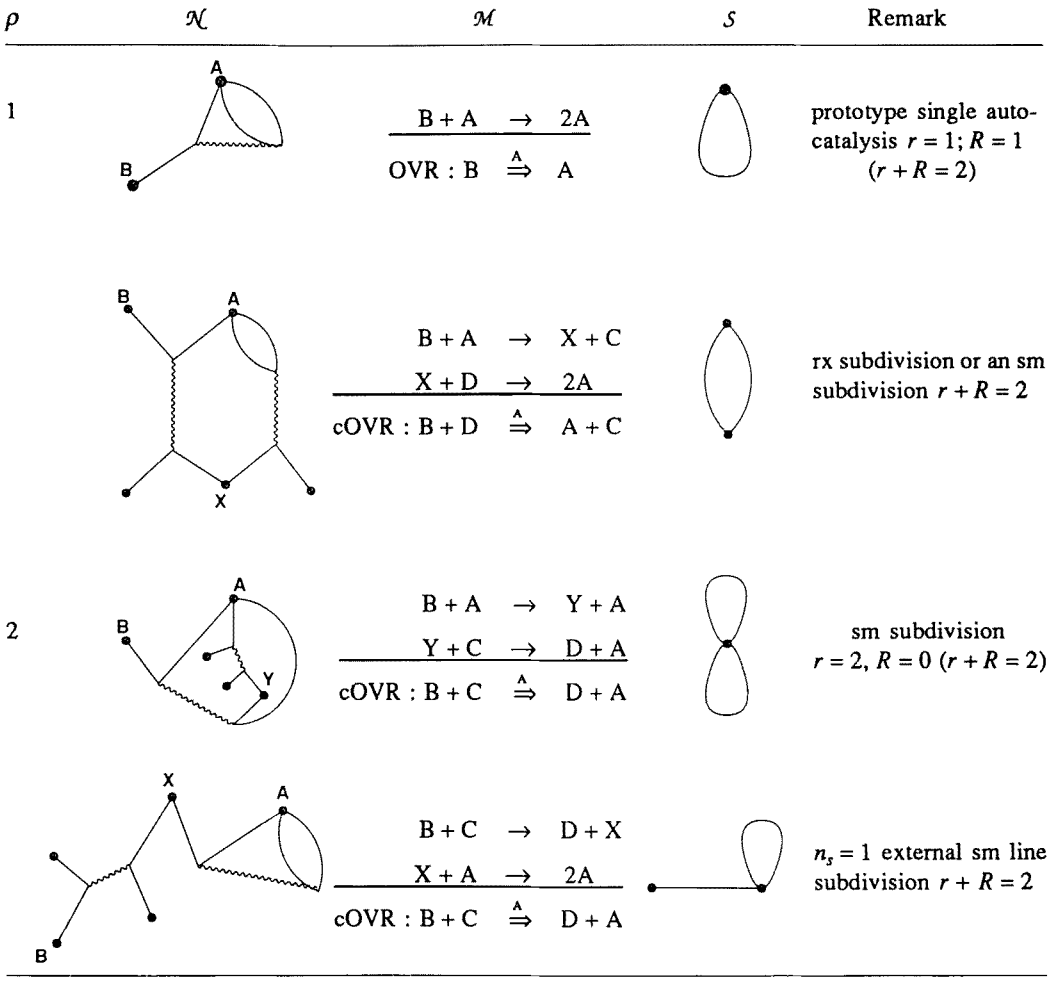


Fig. 4. Topologically $(r + R)$ equivalent complete networks $\{c\mathcal{N}\}$ generated by $\rho = 1 \rightarrow \rho + 1 = 2$ from the prototype autocatalytic reaction A + B → 2A by sm- and rx-line elementary subdivisions.

In fig. 5, we show all the one-autocatalysis networks of $\rho = 3$ derived from the B + A → 2A; OVR: $B \overset{A}{\Rightarrow} A$ prototype by sm- and/or rx-line subdivisions (topological $\mathcal{N}$-invariant $(r + R) = 2$). Unlike in fig. 4, in fig. 5 all the $\{\mathcal{M}\}$ for still

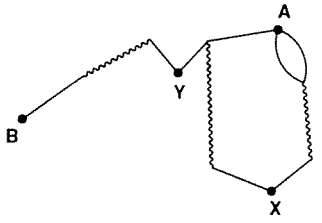
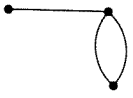
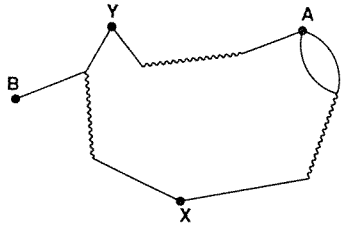
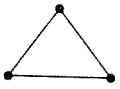
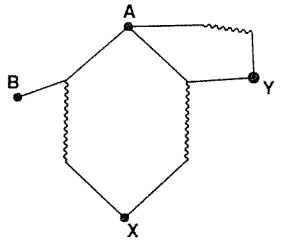
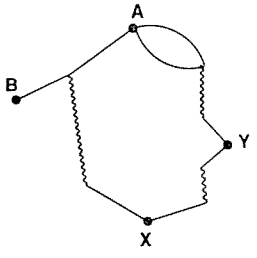
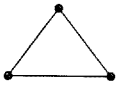
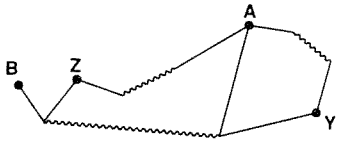
ρ	$\mathcal{N}$	$\mathcal{M}$	S
3		$B \rightarrow Y$ $A + Y \rightarrow X$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B + Y \rightarrow X$ $X \rightarrow 2A$ $A \rightarrow Y$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B + A \rightarrow X$ $X \rightarrow A + Y$ $Y \rightarrow A$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B + A \rightarrow X$ $X \rightarrow Y$ $Y \rightarrow 2A$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B + Z \rightarrow Y + A$ $Y \rightarrow A$ $A \rightarrow Z$ OVR : $B \overset{A}{\Rightarrow} A$	

Fig. 5 (continues)

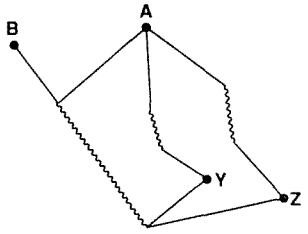
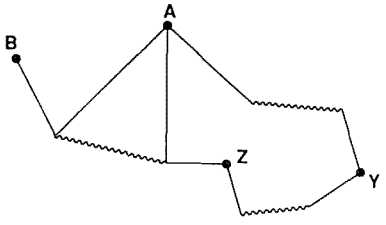
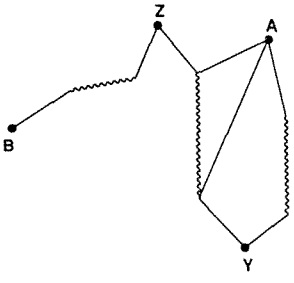
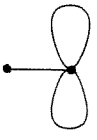
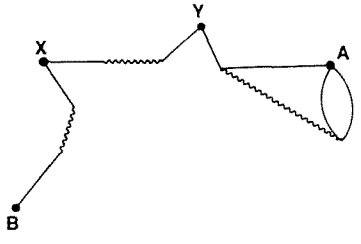
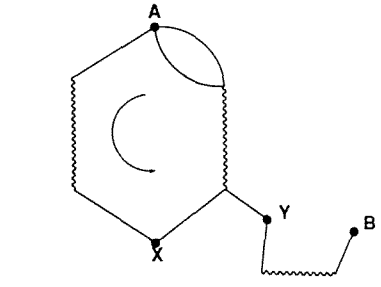
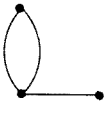
ρ	$\mathcal{N}$	$\mathcal{M}$	S
3		$B + A \rightarrow Y + Z$ $Y \rightarrow A$ $Z \rightarrow A$ OVR : $B \overset{A}{\Rightarrow} A$	
		$A + B \rightarrow A + Z$ $Z \rightarrow Y$ $Y \rightarrow A$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B \rightarrow Z$ $Z + A \rightarrow A + Y$ $Y \rightarrow A$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B \rightarrow X$ $X \rightarrow Y$ $Y + A \rightarrow 2A$ OVR : $B \overset{A}{\Rightarrow} A$	
		$B \rightarrow Y$ $Y + X \rightarrow 2A$ $A \rightarrow X$ OVR : $B \overset{A}{\Rightarrow} A$	

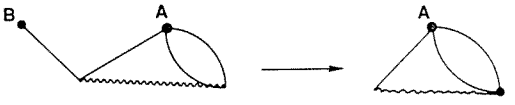
Fig. 5. Topologically $\mathcal{N}$-equivalent ($r + R = 2$ invariant) networks $\{\mathcal{N}\}$ for the OVR: $A \overset{A}{\Rightarrow} B$ generated by $\rho = 1 \rightarrow 2 \rightarrow 3$ from the prototype $A + B \rightarrow 2A$ by sm- or rx-line elementary subdivisions.

the OVR: $A \overset{A}{\Rightarrow} B$ are shown. For simplicity, each network $\mathcal{N}$ or its $\mathcal{M}$ is not completed by the addition of $n_s = 1$ external species lines to the newly obtained unsaturated (uni-) reaction vertices. For larger OVRs, those lines can be added to the $\{\mathcal{N}\}$ shown to obtain the $\{c\mathcal{N}\}$. Also, such added $n_s = 1$ lines in going from $\{\mathcal{N}\}_{\rho=1}$ to $\{\mathcal{N}\}_{\rho=2}$ to $\{c\mathcal{N}\}_{\rho=2}$ in fig. 4 can be subdivided too to obtain a few more $\rho = 3$ networks.

(C) STILL ANOTHER SIMPLE WAY TO GENERATE ALL THE POSSIBLE $(r + R)$ INVARIANT NETWORKS AND THEIR POSSIBLE OVERALL REACTION (OVR) TYPES

The above subdivision methods can be made still simpler in another systematic way, allowing one to also obtain all the possible OVR types as one proceeds.

Again, take a prototype $\mathcal{N}$, e.g. the one-autocatalysis $(r + R = 2)$ case. However, now delete the $n_s = 1$ external species line, e.g.

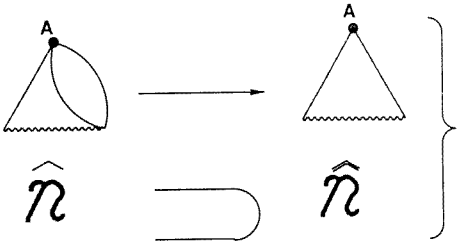


(45)

leaving the basic frame $\{\hat{\mathcal{N}}\}$ of $\mathcal{N}$

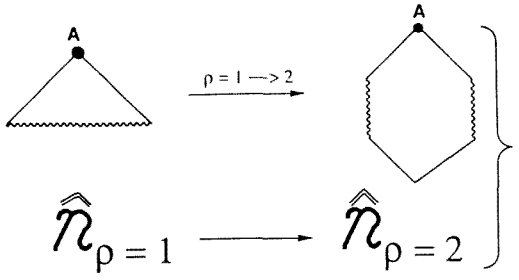
$$\mathcal{N} \supset \hat{\mathcal{N}}. \tag{45'}$$

Carry out all possible elementary subdivisions (sm and rx types) until the desired numbers $\{\rho\}$ of steps are reached. The convenience is that in general there are few resulting "expanded" $\hat{\mathcal{N}}$s due to the equivalence of many of the lines. In fact, if first the basic r frame $\hat{\mathcal{N}}$ is taken,



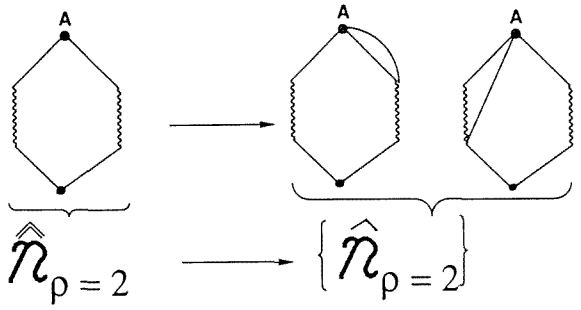
(46)

any one (sm or rx) subdivision $\rho = 1 \rightarrow 2$ yields from $\hat{\hat{\mathcal{N}}}$ only one expanded $\hat{\hat{\mathcal{N}}}$:



(47)

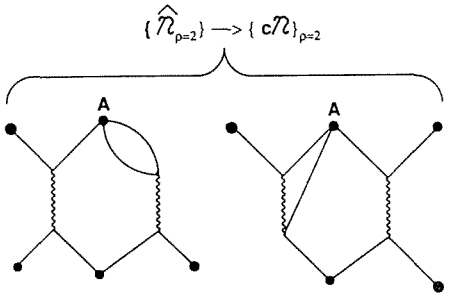
Next, the multiple A line can be added, due to symmetry, in only two possible ways:



(48)

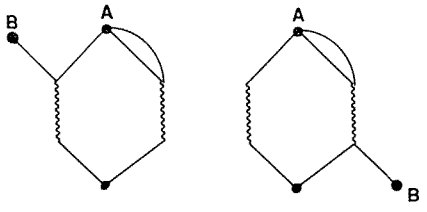
Next, all the $\{\mathcal{N}\}_{\rho=2}$ are obtained either for a specific OVR (like $B \overset{A}{\Rightarrow} A$) or for the complete cOVR, the $\{c\mathcal{N}\}_{\rho=2}$. Possible ways of adding $n_s = 1$ external species lines to $\{\hat{\mathcal{N}}\}$ or deletions from $(c\mathcal{N}\}$ give the smaller {OVR}.

From eq. (48),



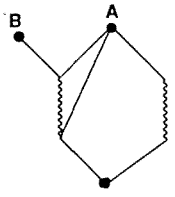
(49)

or, for example, for the OVR: $B \overset{A}{\Rightarrow} A$,



(50a)

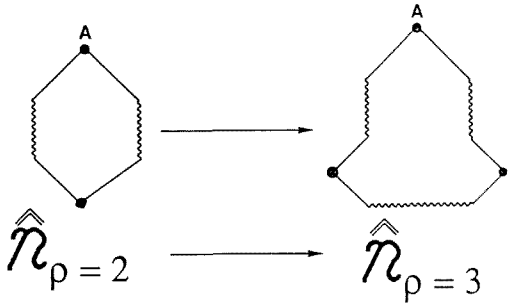
and



(50b)

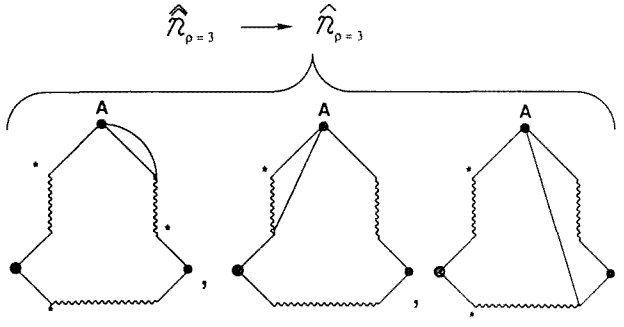
which are the only ones. Note that the direction of flow around the large ring is determined by the type, $n_s = 3$, $(\vec{n}_s, \overleftarrow{n}_s) = (2, 1)$ nature of the A vertex. Then it is clear at which points the B line can be added for the OVR: $B \overset{A}{\Rightarrow} A$.

For $\rho = 2 \rightarrow 3$, the procedure is similar. The $\hat{\hat{\mathcal{N}}}_{\rho=2}$, eq. (47) is expanded by a step,



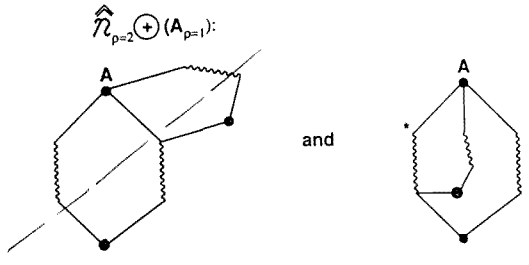
(51a)

the multi-A line added



(51b)

or just the separate multi-line of A is expanded ($\rho = 0 \rightarrow 1$) and added to $\hat{\hat{\mathcal{N}}}_{\rho=2}$ in the two possible ways, the first of which is the same as the third in eq. (51b):



(51b′)

Now the B line is easily added to (51b) and (51b′) going counter-clockwise from A. The possible positions are shown with (*) in eqs. (51b, b′) (one (*) only at a time for $B \overset{A}{\Rightarrow} A$), for consistent networks (no $A \rightarrow A$ trivial step, no $A + B \not\rightarrow A$ forbidden step). For completeness, the B line can be expanded too as in fig. 5.

10. Pathways versus mechanisms

The above results, such as the topological generation of networks of desired features, apply to synthetic pathways [1], metabolic cycles, and thermodynamic cycles [12] as well, not only to mechanisms.

The distinction between *pathways* or *cycles* and *mechanisms* is in the elementary reaction steps. In the first two, the steps are conveniently delineated chemical reactions with no *molecularity* in the kinetics sense implied. Thus, such steps are themselves like simpler OVRs, each having a mechanism of its own, but not written out. The reaction vertices of such "*pathway steps*" are therefore not restricted to being uni- or bi-molecular. The elementary reaction steps of a *mechanism*, on the other hand, of course display *molecularity*, not only stoichiometry. In the collisional sense, therefore, their rx vertices are to the largest probabilities uni- or bi- at each end, especially since steps in general are taken to imply forward and backward rates.

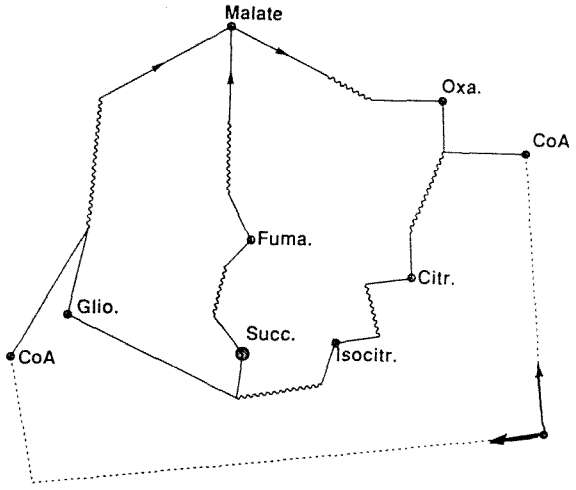
With the above methodologies, more species sm lines than two can be added at each rx vertex in generating pathways or non-kinetic cycles. Then, upper bounds in equations like (21b) are modified. The rest of the theory and methodology proceeds rather similarly, as with the kinetic mechanism problems.

11. Examples of autocatalytic and other turbulent networks

(1) *Models of self-replicating systems*: Most biological models involve some autocatalytic steps. The "chemotons" of Tibór Ganti [8] are autocatalytic cycles taken as simple ones still exhibiting self-replication (of the cycle itself as well as/or of some templates). It appears plausible that these would also exhibit stable oscillations, and might be models of *homostatic* systems [8], although Ganti's arguments on the latter are not clear, especially vis-a-vis the catalytic cycles themselves, which were discussed earlier [1].

The "chemotons" are derivable by topological equivalence from the autocatalytic prototype (eq. (45)) (cf. also fig. 5), but our procedure here would also give additional networks of similar behaviour.

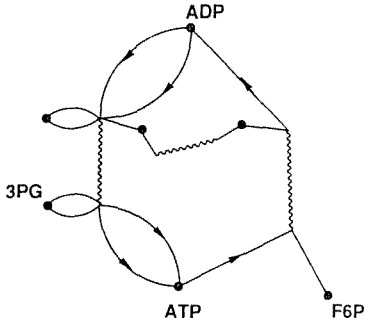
(2) *A malic acid cycle* also arises, by $[\rho \rightarrow \rho + 1]_6$, from the autocatalytic prototype (eq. (45)).



(52)

The acetyl-CoA sm lines can, but need not, be joined to an s vertex, since that vertex is a fully external [$n_s = 2$, $(\vec{n}_s, \overleftarrow{n}_s) = (0, 2)$] one and splits. Note the relation of this network (eq. (52)) to the one in eq. (51b′).

(3) *Phosphorylation in glycolysis*: The cycle generating two 3-phosphoglycerate (3PG) from fructose-6-phosphate [8] (F6P) has the network:



(53)

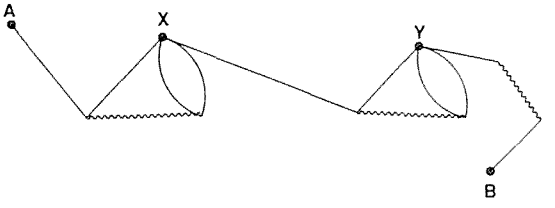
The arrows are shown for convenience. Note the relation of this network to the a priori one ($\rho \to \rho - 1$), eq. (33b).

(4) *The Lotka–Volterra*[2] *model*, discussed by Glansdorff and Prigogine [7] in connection with chemical oscillation, is

[2] The model was originally developed for ecological predator–prey kinetics, which brings us to remark that the network theory we have developed in ref. [1] and here is applicable also to other fields such as ecology and economics.

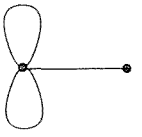
$$\begin{aligned} A + X &\rightarrow 2X \\ X + Y &\rightarrow 2Y \\ Y &\rightarrow B \\ \hline OVR: A &\Rightarrow B \end{aligned} \qquad (\mathcal{N}54)$$

with the turbulent network:



($\mathcal{N}54$)

and its skeleton



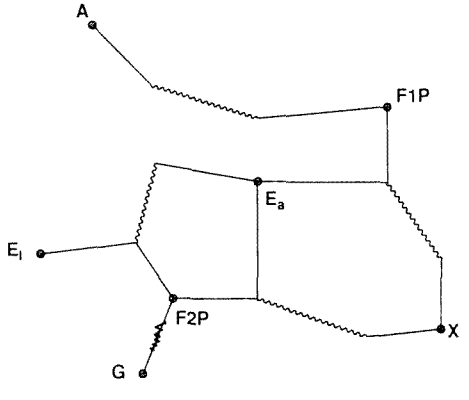
($\mathcal{S}54$)

Note the close relation of ($\mathcal{N}54$) to the network eq. (37c) we derived a priori.

(5) *Higgins model* [19] *of oscillations in the glycolytic cycle*: The mechanism is

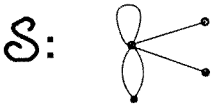
$$\begin{aligned} A &\rightarrow F1P \\ F1P + E_a &\rightarrow X \\ X &\rightarrow F2P + E_a \\ F2P + E_i &\rightarrow E_a \\ (F2P &\rightarrow G) \\ \hline OVR: A + E_1 &\Rightarrow E_a \\ +(F2P &\Rightarrow G) \end{aligned} \qquad (\mathcal{M}55)$$

with A ≡ glucose; F1P, F2P ≡ fructose-1- and-2-phosphate; E_a, E_i ≡ active and inactive forms of the enzyme; X ≡ complex; G ≡ glyceraldehyde 3P.



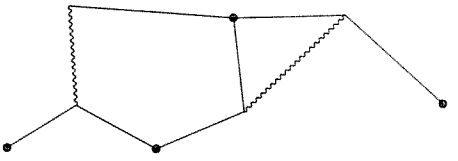
($\mathcal{N}55$)

and



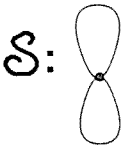
($\mathcal{S}$55)

($\mathcal{N}$55) is topologically equivalent by $[\rho \to \rho - 1]$ to



($\mathcal{N}$55′)

with the



($\mathcal{S}$55′)

Network ($\mathcal{N}$55) is one of the a priori ones in fig. 5, derived in eq. (51b). The ($\mathcal{N}$55′) is the one in eq. (38).

12. A general topological feature of models for chemical oscillations, instabilities, and self-replication with homeostasis

A number of diverse mechanisms or cycles have been proposed in connection with oscillations, instabilities, fedback, replication, etc. by various authors. We gave a few of the examples above.

Although the networks often look quite different in different models (compare, for example, ($\mathcal{N}$54) and ($\mathcal{N}$55)), we note that in each case including other examples not mentioned here, the skeletons are topologically r-equivalent with $r = 2$ and to

(56)

(cf., e.g., eqs. (53), (52) with the CoA split as it does, (54) and (55)).

On the other hand, by the method of this paper we have seen that it is possible to generate a priori the whole class of such $r = 2$ networks. Thus, other models can systematically be constructed with the given topological features.

It remains to be seen whether the class of $\{\mathcal{N}\}$ of eq. (56) is necessary and/or sufficient to obtain the particular dynamical behaviour and what, if any, other classes exist. More generally, the relation of various (r) and $(r + R)$ classes to dynamical behaviour, kinetics, and instabilities needs further examination.

Acknowledgement

The contributions to the first part of this work in its preliminary stages, in particular, vis-a-vis "laminar" networks and some of the "turbulent" ones, of Dr. L.S. Lee during our most enjoyable association some time ago, are gratefully acknowledged.

References

[1] O. Sinanoğlu, J. Am. Chem. Soc. 37(1975)2307.
[2] O. Sinanoğlu and L.S. Lee, Theor. Chim. Acta 48(1978)287; 51(1979)1.
[3] L.S. Lee and O. Sinanoğlu, Z. Phys. Chemie 125(1981)129.
[4] O. Sinanoğlu, Chim. Acta Turc. (Int. Ed.) 3(1975)155 [Trk. Ed: Türk Kimya Dergisi 3(1975)157]; Nippon Kagaku Kaishi 9(1976)1358.
[5] K. Miyazaki, J. Appl. Chem. Biotech. 23(1973)93.
[6] J.J. Tyson, J. Chem. Phys. 62(1975)1010.
[7] P. Glansdorff and I. Prigogine, *Structure, Stability, and Fluctuations* (Wiley, New York, 1971).
[8] T. Ganti, *A Theory of Biochemical Supersystems* (University Park Press, Baltimore, 1979).
[9] D. Poland (1990), unpublished; J. Chem. Phys. 95(1991)7984.
[10] D. Bonchev, D. Kamenski and O.N. Temkin, J. Math. Chem. 1(1987)345.
[11] R.B. King (ed.), *Chemical Applications of Topology and Graph Theory* (Elsevier, Amsterdam, 1983).
[12] P.H. Sellers, in: *Chemical Applications of Topology and Graph Theory*, ed. R.B. King (Elsevier, Amsterdam, 1983) pp. 420–429; Arch. Rational Mech. Ann. 44(1971)23.
[13] B.L. Clarke, in: *Chemical Applications of Topology and Graph Theory*, ed. R.B. King (Elsevier, Amsterdam, 1983) pp. 322–357; *Advances in Chemical Physics*, Vol. 43, ed. I. Prigogine and S.A. Rice (Wiley, New York, 1980) pp. 1–215.
[14] R. Barone, M. Chanon and M.L.H. Green, J. Organomet. Chem. 185(1980)86.
[15] O.K. Rice and K.F. Herzfeld, J. Am. Chem. Soc. 56(1934)284.
[16] L.S. Lee and O. Sinanoğlu, Mechanisms of thermal decomposition of ethane from the theory of chemical reaction networks, Report (1973);
L.S. Lee, Ph.D. Thesis (1973).
[17] O. Sinanoğlu, J. Math. Phys. 22(1981)1504.
[18] O. Sinanoğlu, A priori design of thermochemical cycles for hydrogen production using the theory of chemical reaction networks, Abs. 3rd Int. Congress on Quantum Chemistry, Kyoto, Japan (1979).
[19] J. Higgins, Ind. Eng. Chem. 59(1967)19.