SETTLEMENT SYSTEMS IN THE EVOLUTION

by

Denise Pumain


The authors of a recent bibliographical essay (Pumain and Robic, 1996) stressed the degree to which theoretical approaches to urban places are static. When such theories do include time, it is usually only implicitly. The focus of explanation tends to be on structures, and much more rarely on the forms taken by change. No theory of urban systems, especially those expressed in geometrical or mathematical models, has so far satisfactorily integrated the urban phenomenon’s temporal dimension, in the sense of change over historical time. However effective static interpretations may appear, as explanatory tools they completely overlook what is an essential characteristic of towns and cities, namely their capacity to change and adapt over time. The basic question to be addressed is that of how and why the man in the street is not mistaken to go on describing as “town” or “city” objects whose form, content and meaning have changed continuously over the centuries.

It is important to be clear that an evolutionary theory is not the same as a historical theory. The aim is not to consider the whole sweep of urban history, nor to reconstruct the particular contours of urban genesis. The approach is not that of historical study in the sense of exploring the past to seek an “explanation” in a detailed account of how a place has developed. The objective is not the descriptive one of historical geography, but is instead that of producing a theoretical formalization of urban evolution as the process of the historical transformation of specific settlements. The aim is to demonstrate that a geographical object can be interpreted as the particular outcome, from among a set of possible outcomes, of a general dynamic process. In making the simplifications necessary for modelling, the notions of uniqueness and unrepeatability that characterise historical objects are inevitably lost, but it does take into account the non-reversibility of their particular history.

Difficulties in static explanations

As their structure is evolving slowly, it is not surprising that explanations for settlement systems have been initially searched in static theories, mainly in the framework of economics. But there is no evidence that an adequate explanation of urban systems could be derived from the principles of general economic theories. Economic theories are missing those spatial and evolutionary features which seem indispensable for a real understanding of urban systems.

Agglomeration economies and central places

Most urban theorists who have based their work on micro-economic theory explain the spatial clustering of economic agents, assumed to be utility maximisers, by the existence of “agglomeration economies” or “economies of scale external to the firm” (Catin, 1994, p. 105). An initial difficulty with this is that although these “agglomeration economies” are frequently invoked they are never actually measured. It can also be pointed out that for decisions about localisation to be based on agglomeration economies assumes that the town or city already exists, and the theory gives no explanation of how it developed. Another criticism that can be made of these theories is that they usually assume that the town or city exists in isolation: the agents cluster according to their individual interests without there being any necessary relations with other urban places. In order to explain urban systems and to account for their hierarchical structure, the micro-economic theories make the assumption that returns increase with the size of urban centre. For example, Fujita et al. (1994) work on the hypothesis that the system optimises its operation by establishing an equilibrium between supply and demand for services at the non-scalar level of the city, while individuals optimise their localisation by maximising their utility. This means that the towns and cities which offer the widest range of services will be more attractive, and the influx of migrants will in turn cause an increase in the range of serv-
The problem is that behind all the static explanations for the size and spatial arrangement of urban settlements can be discerned the notion of equi-
librium, governed by a determining factor in the form of some force, an "invisible hand", a statisti-
cal constraint. The form of urban networks has thus been "explained" by the action of contradictory forces working for unification and diversification (Zipp, 1949), by the micro-economic theory of market agents (where competition is the spatial expression of a principle according to Berry (1967)), by a prin-
ciple of maximising social interaction (Claval, 1982) or by a statistical process in which a structure is deduced from a mechanism controlled by a lim-
iting or optimising principle (entropy maximising, Curry, 1964). All these explanations in fact derive from the same hypothesis, which can be summed up thus: the configurations observed at any given moment are the result of the optimisation of a con-
straining factor, or are the expression of an equilib-
rium between contradictory actions. Fujita et al. (1994) consider as a "natural rule" (sic) that "at any instant in time, the urban system must be in spatial equilibrium" (p. 21). In this recent version, they try to introduce a dynamic element to their model by starting with a static equilibrium model and imag-
nining the kinetic transfer of dynamic time. Contrary to what is claimed by the authors, such a model does not allow us to understand "the emergence of a system of cities".

The construction of the dynamics of urban networks does not supply proof that the dynamics of urban systems are governed by a macroscopic constraining factor, or that what is observed at any given time represents an equilibrium state. Or, on the contrary, what we find are large or gradual changes in the distributions, but without knowing if the forces that are supposed to have produced them have changed. A certain dynamic can thus be identified at work within a geographical distribution. This dy-
namic is the product of interactions which, if they were allowed to operate unhindered over a long enough period, might result in an equilibrium that could be modelled. In reality, of course, the form and intensity of the interactions have always changed before this equilibrium can be reached. Imagining that a system's current state corresponds to an equilibrium thus means assuming, if it is to be explained in these terms, a set of interactions which on their own cannot in fact explain the present form (Allen, 1991). We are thus forced to take the view that the configuration observed and the transforma-
tions it experiences must be explained by a single theory, with the same theory it must be possible to explain the structure of the system and the changes occurring in it.

The somewhat tautological hypothesis used by economists to justify the concentration of economic activities in urban centres, and which presupposes the existence of agglomeration economies, does not appear to me to be necessary to explain the dif-
f erences in the size and evolution of towns and cit-
es. An evolutionary theory can account for these without having to accept the idea that large cities are more productive and more efficient in economic terms than small ones in order to explain the exist-
ence of urban networks in their present form. In-
 deed, the greater economic efficiency of large ur-
ban centres, if this is actually proved to be the case (Rousseau and Prudhomme, 1993), and thus the ex-
istence of agglomeration economies, could be as much the consequence as the cause of their success.

Dynamic models of urban growth

With the creation of historical and comparative data bases (Bairoch, 1996) a large volume of ob-
servations relating to change in urban systems is now available. Early efforts to incorporate this ma-
terial in mathematical models, as in Christaller's book (1933) and the "transfer" of dynamic models from the "pure" sciences of mathematics, statistics, physics and chemistry. However simplis-
tic these models may be, they have a valuable fil-
tering role. By making possible the comparison be-
 tween urban dynamics and the dynamics of other systems they have helped to identify a number of specific aspects of urban evolution and thus con-
tributed to the development of more sophisticated methodologies.

Many models present merely a simplified de-
scription of urban systems in terms of sets of pop-
ulation centres characterised by their number of inhabitants. Although this simplification is con-
tinuedly crude it is not meaningless given that 'de-
mographic size' is the variable which best express-
 a wide range of functional properties of towns and cities. Population size is a correlate for many quantitative descriptors, such as the number of jobs, firms or homes, but also for qualitative de-
scriptors such as the diversity or complexity of economic activities within urban space. It consti-
tutes the main "dimension" of an urban system, its most important factor of differentiation. (Reiner and Parr, 1980).

Knowing the size of a city typically supplies a lot of other information about that city, and can thus be used to make a wide range of fairly accurate predictions. Size presents the addi-
tional advantage of being measured over very long periods in a reasonably straightforward and
comparable way, and of keeping a relatively con-
sistent significance in terms of the 'success' of the
system it represents. The fact of attracting and
keeping a certain number of residents was for long taken as the measure of a city's capacity to provide the means of subsistence for this population
(through local production and innovation, but also by the extraction of surplus from the countryside or other urban centres, or by imposing its terms in a system of exchanges). For present-day interna-
tional comparisons in particular, however, this in-
dicator of urban success needs to be completed by an
evaluation of living standards and levels of out-
put, as well as by indicators of quality of life and the
environment.

For the purposes of this article we will just use this description of the system in terms of city size for the formal part of the theory. This does not mean that an evolutionary theory has no further exten-
sion. It is merely the first stage, and it will in fact be seen that a coherent and formalised conceptualisa-
tion leads to including far more complex aspects of the definition of an urban system even in this sim-
plified version of the theory.

Gibrat's model

The simplest alternative to the static point of view is the approach which treats urban systems as the product of a stochastic process distributing popu-
lation growth between cities. This involves trans-
forming a statistical model for the dynamic descrip-
tion of an urban system. In his book: Les inégalités économiques (1931), Gibrat demonstrated that
when cities are growing at the same average speed but with fluctuations or inequalities of growth, the distribution of the city sizes will consistently take a particular form.

Let there be a set of localities with a certain size distribution (they can even be all the same size at
the start of the process) and whose evolution over a long period (several centuries, for example) is
modelled, using a large number of short time peri-
ds. It is assumed that in each time period (for ex-
ample, one year, or ten years) the population P_t of
each locality grows on average (though with fluctua-
tions e_t) by an amount dP which is proportional to
the population P_t (and low in relation to P_t):

\[ dP_t = r P_t + e_t \]

where the size of a city typically supplies a lot of other information about that city, and can thus be used to make a wide range of
This is the same as saying that the proportional rate of change of the population $dP/P$, (which is the measure usually expressed in percentage terms and used to describe and compare the growth of cities and regions) has the same average value for all settlements, whatever their size at the beginning of each time period. If in addition the distribution of their urban attributes is independent from one time period to the next, the growth process it defines will always result in a distribution of settlement sizes that is a lognormal distribution, highly skewed, characterised by a large number of villages and small towns and a geometric decline in the number of cities according to their size. If the rate of growth and its variations are known it is even possible to predict (at a known probability) by how much the size disparity will increase.

With this process the attractiveness of cities does not have to be assumed to increase with their size in order to obtain a highly uneven size distribution, comprising a small number of very large cities (consequently there is no need for a theory that like that of agglomeration economies). In addition, this model can be "tested" empirically against the statistical observations of urban growth rates at regular intervals over long periods. It has been confirmed for many different countries and periods (Robson, 1973; Pumain, 1982; de Vries, 1984; Guérin-Pace, 1993; Moriconi-Ebrard, 1993). A more satisfactory "explanation" for the form of the distribution of city sizes is thus obtained when we view it as resulting from a dynamic process for the distribution of urban functions, rather than as the expression of a static equilibrium. Compared with these static interpretations, Gibbard's model is in fact the only one to show why this form of organisation is unchanged over time and to deduce a gradual concentration of the population in increasingly large cities at the top of increasingly skewed distributions. The model thus provides an "explanation" for the gradual differentiation in size between cities that are involved in the same evolution, though with an uneven effect produced by the "accidental" (or random, in the sense that they are not determined by the model) repetition in some places of incremental increases that are larger or smaller than those experienced by the set of other cities. As a first approximation, then, Gibbard's model provides both a good description of growth in an urban system and an evaluation of its hierarchical structure. It can be used for predicting the evolution of city sizes over periods of up to several decades.

However, a theory of urban systems cannot be restricted to the formulation of this purely statistical model, expressed as a function that accommodate the systematic deviations from Gibbard's model that are observed: first, over the long term there is usually a slight positive correlation between urban growth rates and city size; second, they are sometimes temporally autocorrelated for periods of several decades (Robson, 1973; Pumain, 1982; Guérin-Pace, 1993). In addition, the "primary" cities, the two or three largest in the hierarchy according to the geographer finds national urban hierarchies, whose size appears "disproportionate" compared with the rest of the urban system, are in fact agglomerations which have grown systematically faster than the rest of the urban network over several decades (Moriconi-Ebrard, 1993; Pumain and Moriconi-Ebrard, 1997). The model therefore has to be corrected and completed so as to accommodate what is in relation to the simple stochastic process appear as "anomalies" but which are in fact revealing processes that are specific to the evolution of the urban system. This leads us to an explicit formalisation of the relations of interdependence between cities, and thus to a shift from a formalisation of a set of cities to that of a system of cities. In the Gibbard model, the notion of interdependence between cities remains implicit and is presented only in two ways in the sense that the growth rates of cities obey a distribution with the same average and with a given variance: the implication is thus that they belong to the same "system". This hypothesis does not appear as such in the statistical model and it is when the model is transferred to cities that it becomes logical to introduce it as expressing the information exchange and competition between cities.

Empirical observation also necessitates consideration of another process which is that of the selection of those members of the set of inhabited localities which grow into towns and cities. The numbers described by Gibbard in fact apply not only to urban places but more generally to the settlement system as a whole. Clearly, however, this dynamics tends to produce not one but two logarithmic distributions, which empirically represent the contrast between the size of their elements. From our perspective, this implies that two subsets of localities have been subject to a long-term "selection", one for a dynamic system that is the urbanisation process, and the other for the evolution of the urban system, as has been the case for the towns and rural communities in France over the last 200 years (Pumain 1995, p. 437). The distinction between towns and cities and villages thus leads to incorporating into the evolutionary theory qualitative considerations concerning the functions and historical change known as the "urban transition".

A dynamics of interdependence

Communication, exchange and interdependence between similar entities are properties which appear to be linked to a geographical definition of the urban concept. For this reason and for a long time the geographer finds it hard to accept the definition (presented as a "non-definition") put forward by H. Hegesin (1996), which emphasises the concept of density but does not satisfactorily articulate the two identifiable levels of urban organisation, the city and the city system, and above all fails to integrate the "ecological" dimension, that is, to provide an explanation of the urban object that is "vertical" rather than simply "horizontal" or "horizontal". The individual urban place cannot be dissociated from the system of urban places, and exchange of whatever type or form, including therefore long-distance relations, is integral to the urban function and are part of its specificity. According to Henri Reymond (1981), what distinguishes the town and city from other forms of settlement (village, mining centre), which exploit population from a single site, is the possibility of exploiting a situation, the resources of several different sites. This conception is consistent with the findings of historical and archaeological research regarding the genesis of towns and their evolution. For P. Riaux, "the existence of an urban centre presupposes not only an agricultural surplus but also the means of exchanging this surplus" (1995).

Reymond (personal communication, 1995) justifies his definition using the concept of limiting factor borrowed from ecology, which postulates that at any given moment the resources available to a population is established in a particular location act as a limit to its development: "The system of urban habitat is the basic ecological organism which, of all those invented by the human race, can eliminate the demand completely the set of limiting factors of a site, whatever its own resources. Towns and cities have succeeded in overcoming the uncertainties arising from local limiting factors and in controlling these local difficulties, by developing multiple links and forming ever faster networks. Sites that are linked in this way can support a population much larger than was possible on the initial site. This evolution is not finished, and the tool that these towns and cities become is continually being invented."

The existence of a pattern in the evolution of interdependent urban centres was noticed long ago, for example by Reynaud (1841) when he introduced the concept of "general city system". Berry's (1964) famous phrase "cities as systems within in cities of systems" was a use of systems theory as the framework in which to formalize the relationship between the level of organization and its evolution. The model is based on the classificatory collection of neighbourhoods, institutions and actors linked by systemic relations, and the city system, organized on the principles of central place theory. In his book published in 1977, Peed stresses the dynamic aspect of the notion of urban system by defining it in extremely general terms as: "a national or regional set of cities which are interdependent in such a way that any significant change in the economic activities, occupational structure, total income, or population of one member city will directly or indirectly bring about some modification in the economic activities, occupational structure, total income, or population of one or more set members" (p. 13).

Subsequent empirical research on the ways in which urban systems have expressed the value of interorganisational dynamics within the framework of self-organisation theory. This formalisation can be used to refine Peed's definition and in particular his notion of "significant change". Not all local transformations will affect the other cities in the system. Many of the changes that affect towns and cities over periods of several years or several decades (in urban morphology, in demographic, social and economic structures, etc.), though they could be described as 'significant', have no effect on the structure of the system, that is on the form of inter-urban differentiation, because they affect all towns and cities (Pumain and Saint Julien, 1978; Sanders, 1992). The structure of the system will also not be modified by changes that are particular to a town or city, because they do not last long enough from one period to the next and are more fluctuations which do not alter the relative position of urban centres within the system. A structural change will occur when one of these fluctuations persists and increases to the point that it modifies this position, usually of several cities at the same time, related to the development of a new specialisation, for example. In this case there is said to be a bifurcation in the history of the system. Most fluctuations have no effect if they occur while the
system is on a stable path (Pumain, 1997; Pumain et al., 1989). A bifurcation may also be the result of a perturbation originating outside the system in question. Models based on self-organisation theories Observations of the way in which interdependent cities evolve has prompted the transfer to geography of models originating in chemistry and physics. The analogy with physical systems, using the dissipative structures approach (Allen, 1978) or synergetics (Haken, 1977; Sanders, 1992), involves formalising the dynamics of systems as the product of interactions between microscopic units which, when taken together form structures that are detectable at the macroscopic level. The system thus described is "open"; in the case of a physical system this means that it receives an input of "energy" from the exterior which allows "self-organisation" to occur. When the model is transferred to human systems, this energy input is assimilated to innovation, or to a trend towards growth of the system, which acts as the "motor" of its dynamics. Application of the self-organisation concept to these systems suggests that the fact that they are structured by a persistent and thus identifiable architecture cannot be explained by the action of a network operator or by any endogenous political or economic objective, such as could be described in terms of some constraint to be respected or quantity to be optimised. The structure is in fact the spontaneous product of multiple interactions between many separate agents. Each of these has its own objectives and strategies, but do not need to be known in detail in order to predict (or simulate) what will be the system's general architecture. In this way the mass of individuals' decisions can be reduced to a small number of parameters and equations which describe the system's evolution at a macroscopic level: this is in principle possible using the probabilistic models taken from synergetics and self-organisation theory, in which a "master equation" expresses the overall effect of the interactions occurring between the particles whose individual state and position within the system cannot be known (or predicted). The notion of random events related to the existence of several levels of observation in the system: the microscopic phenomena influence the form of the macroscopic phenomenon, while the latter's law which unites this microscopic phenomena, or, in other words, subjects them to its "point of view". The particular laws of these two types of phenomena are not usually random laws, and these are merely the laws of their interaction (Haken, 1977).

Self-organising systems are not in a state of equilibrium at any given time; their transformation is the result of internal and external dynamic processes, and their structure is the product of this evolution. Changes in urban centres or urban systems take forms similar to those in the physical theories of self-organisation. This combines, on the one hand, a stability or slow transformation in the system's structure (the macro-geographical state of the urban system) reflected in the form of the size distribution, the main differentiating factors in the economic activities and social composition of cities), and on the other, very large fluctuations in the relative position of its elements: at the meso-geographical level this corresponds to the rapid shifts individual cities make between "states" of growth, stagnation, contraction; or changes of rank, within the size hierarchy over the medium term; or ten-}

From dynamics to evolutionary theory These various analyses do help to identify the characteristics of an authentic evolutionary theory of urban systems (Pumain, 1997). We will not focus here on the substantial transformations in urban form and economy, which in a legitimate way are the basis for most urban research agendas (for instance, Bourne, 1995). Our objective is to concentrate on the few features which can be considered as remaining general and specific properties of urban systems over time. An evolutionary theory thus has to accept that its object is transformed over time, based on its previous forms but acquiring new properties. The theory must set out the general pattern of these changes and identify their direction, which can then serve as the basis for forecasts. Of course such a theory may be seen as an over-simplification compared to the complexity of real processes of change. It would be premature to claim that a complete formalisation is possible, but we can indicate some guidelines for this which may be of use in its elaboration.

The notion of urban system A simple definition of an urban network or city-

some authors have also suggested incorporating in the theory not only aspects of physical dynamics but also evolutionary processes based on biological theories. Allen (1991) has argued that towns and cities belong to the sorts of systems in which new forms and functionalities are created. The occurrence of innovation in the systems is not due to some optimisation of their functioning at a given time but results from the practices and "discoveries" and "inventories" of non-averaged individuals. This is in fact a social interpretation of the notion of diversity from the biological theory of evolution. The models based on this theory require very powerful computers to simulate the endogenous emergence of innovation. However, it is not certain that social innovation, and the learning and collective processes that it implies, can be modelled in quite the same way as a biological phenomenon. Future models will probably have to include cognitive processes, both for their role in the appearance and selection of innovation, and as additional regulators in the evolution. It may then be possible to explain why bifurcations or "photic" behaviour are less common in urban systems than in other types of evolution.

SETTLEMENT SYSTEMS IN THE EVOLUTION

Denis Pumain

Some authors have also suggested incorporating in the theory not only aspects of physical dynamics but also evolutionary processes based on biological theories. Allen (1991) has argued that towns and cities belong to the sorts of systems in which new forms and functionalities are created. The occurrence of innovation in the systems is not due to some optimisation of their functioning at a given time but results from the practices and “discoveries” and “inventories” of non-averaged individuals. This is in fact a social interpretation of the notion of diversity from the biological theory of evolution. The models based on this theory require very powerful computers to simulate the endogenous emergence of innovation. However, it is not certain that social innovation, and the learning and collective processes that it implies, can be modelled in quite the same way as a biological phenomenon. Future models will probably have to include cognitive processes, both for their role in the appearance and selection of innovation, and as additional regulators in the evolution. It may then be possible to explain why bifurcations or “photic” behaviour are less common in urban systems than in other types of evolutionary systems.

From dynamics to evolutionary theory These various analyses do help to identify the characteristics of an authentic evolutionary theory of urban systems (Pumain, 1997). We will not focus here on the substantial transformations in urban form and economy, which in a legitimate way are the basis for most urban research agendas (for instance, Bourne, 1995). Our objective is to concentrate on the few features which can be considered as remaining general and specific properties of urban systems over time. An evolutionary theory thus has to accept that its object is transformed over time, based on its previous forms but acquiring new properties. The theory must set out the general pattern of these changes and identify their direction, which can then serve as the basis for forecasts. Of course such a theory may be seen as an over-simplification compared to the complexity of real processes of change. It would be premature to claim that a complete formalisation is possible, but we can indicate some guidelines for this which may be of use in its elaboration.

The notion of urban system A simple definition of an urban network or city-

Some authors have also suggested incorporating in the theory not only aspects of physical dynamics but also evolutionary processes based on biological theories. Allen (1991) has argued that towns and cities belong to the sorts of systems in which new forms and functionalities are created. The occurrence of innovation in the systems is not due to some optimisation of their functioning at a given time but results from the practices and “discoveries” and “inventories” of non-averaged individuals. This is in fact a social interpretation of the notion of diversity from the biological theory of evolution. The models based on this theory require very powerful computers to simulate the endogenous emergence of innovation. However, it is not certain that social innovation, and the learning and collective processes that it implies, can be modelled in quite the same way as a biological phenomenon. Future models will probably have to include cognitive processes, both for their role in the appearance and selection of innovation, and as additional regulators in the evolution. It may then be possible to explain why bifurcations or “photic” behaviour are less common in urban systems than in other types of evolutionary systems.

From dynamics to evolutionary theory These various analyses do help to identify the characteristics of an authentic evolutionary theory of urban systems (Pumain, 1997). We will not focus here on the substantial transformations in urban form and economy, which in a legitimate way are the basis for most urban research agendas (for instance, Bourne, 1995). Our objective is to concentrate on the few features which can be considered as remaining general and specific properties of urban systems over time. An evolutionary theory thus has to accept that its object is transformed over time, based on its previous forms but acquiring new properties. The theory must set out the general pattern of these changes and identify their direction, which can then serve as the basis for forecasts. Of course such a theory may be seen as an over-simplification compared to the complexity of real processes of change. It would be premature to claim that a complete formalisation is possible, but we can indicate some guidelines for this which may be of use in its elaboration.

The notion of urban system A simple definition of an urban network or city-

Some authors have also suggested incorporating in the theory not only aspects of physical dynamics but also evolutionary processes based on biological theories. Allen (1991) has argued that towns and cities belong to the sorts of systems in which new forms and functionalities are created. The occurrence of innovation in the systems is not due to some optimisation of their functioning at a given time but results from the practices and “discoveries” and “inventories” of non-averaged individuals. This is in fact a social interpretation of the notion of diversity from the biological theory of evolution. The models based on this theory require very powerful computers to simulate the endogenous emergence of innovation. However, it is not certain that social innovation, and the learning and collective processes that it implies, can be modelled in quite the same way as a biological phenomenon. Future models will probably have to include cognitive processes, both for their role in the appearance and selection of innovation, and as additional regulators in the evolution. It may then be possible to explain why bifurcations or “photic” behaviour are less common in urban systems than in other types of evolutionary systems.

From dynamics to evolutionary theory These various analyses do help to identify the characteristics of an authentic evolutionary theory of urban systems (Pumain, 1997). We will not focus here on the substantial transformations in urban form and economy, which in a legitimate way are the basis for most urban research agendas (for instance, Bourne, 1995). Our objective is to concentrate on the few features which can be considered as remaining general and specific properties of urban systems over time. An evolutionary theory thus has to accept that its object is transformed over time, based on its previous forms but acquiring new properties. The theory must set out the general pattern of these changes and identify their direction, which can then serve as the basis for forecasts. Of course such a theory may be seen as an over-simplification compared to the complexity of real processes of change. It would be premature to claim that a complete formalisation is possible, but we can indicate some guidelines for this which may be of use in its elaboration.

The notion of urban system A simple definition of an urban network or city-
tem is that of a subset of a settlement system. The latter is made up of all the inhabited localities of the same territory – hamlets, villages, towns and cities – as well as the relations (which can be specified later) that exist between these localities and between them and their surroundings. The fact of being part of the same territory gives a degree of unity to the system's functioning, by providing relatively homogeneous conditions of communication between its elements. The territory is defined here by political and cultural standards, as a contiguous portion of the earth's surface occupied or controlled by a group. The notion of territory also implies a degree of delineation (its "boundaries") limit the intensity of interactions between localities that belong to different territories) which justifies the identification of a "system". However, it must be noted that the development of urban systems continues regardless of changes in politically-defined territories. Furthermore, when the forms of political organization, of particular territories have remained stable over long periods, they can give rise to specific variants of the general urban system. Finally, it is clear that to varying degrees, cities, particularly those in the upper reaches of the urban hierarchy, are indifferent to territorial boundaries. Since the range of urban relations is only strongly correlated to the size of the centre, the reference territory should be proportionate to the level of the hierarchy being considered. The notion of system applied to urban settlements is thus of a relatively limited value and needs to be replaced by a more suitable metaphor. However, it is useful for the purposes of modelling, so we shall continue to use it with some reservations until something better is found.

The time-scale of evolutionary change is that in which ••complexification occurs. Adoption of the evolutionary perspective implies integrating the specific temporalities of cities and the urban system into our definitions of these objects. This is only possible if we adopt a relative concept of geographical space.

Relative space and geographical entities

Two concepts of space are found in the work of geographers. The more widely used treats space as a physical reference which acts as the setting and framework for the analysis of the human activities distributed within it (Beguin and Thiese, 1979). Space thus construed is endowed with straightforward geometric properties, even though it gives only an approximate and distorted representation of the earth’s surface, as, for example, in cartographical representations. The other concept defines space on the basis of the relations which exist between localities. Because the distribution of these relations is not homogeneous or isotropic or continuous, the representation of the “relative” space thus produced is not always easy. However, the concept of relative space is essential for understanding the development of what in terms of absolute space are referred to as “spatial structures”, that is, detectable and identifiable geographical objects.

Although the theory is still in its infancy, an evolutionary perspective would show how these geographical entities, in the form of villages, areas of jurisdictions, towns, territories, networks, are constructed through the integration over time of relations between localities. Contact and exchange are responsible for the spread and imitation of practices, for producing personal interaction and shared modes of thought and expression. These relations gradually create the similarities which underpin the formation of homogeneous landscapes and regions, as well as the memberships which define the territories and networks. They are at the origin of the construction of a new space in which increase the attractiveness of some localities. The interlocking of these constructions at different levels is what forms geographical space. A static account of this was given by Philbrick (1957), while a formal analysis has been attempted by Grasland (1994). They have yet to be incorporated into a dynamic theory.

Urban centres construct two spaces

The dynamic models in which cities are formalised as self-organising systems use the representation of an abstract space that is linked by flows of goods, people and information to form a system. Progressing from a dynamic theory to an evolutionary theory requires considering the notion of relative space, one that is defined by these relations and flows. It then becomes necessary to distinguish the two levels of observation of the urban phenomenon that are the city and the urban system. All the connections of a given city to another, which correspond to two types of territory, with different relational scales. Second, the changes in this relative space are a source of information about the origins and history of the urban systems (Painui, 1995).

At any given time, the urban system, or more generally the settlement system, is responsible for the articulation of territorial entities on different geographical levels, by establishing links between levels of organisation with contrasting space-time dimensions. The settlement system is what makes possible the coexistence of at least two main levels. The first level is that of the village or urban place in its role of territory of daily activity. At different periods, the composition and importance of the spatial structure have varied widely, depending in particular on the degree of specialisation of places for residence, work, commerce, or leisure. The average time devoted to these daily journeys, and thus time to be added to that of other activities carried out during the day, remains on the whole relatively constant in the order of approximately one hour, and sometimes between one and two hours. The spatial expansion of the territories of daily life is thus severely limited by the time necessary for internal journeys, the speed of which, particularly in the cities, is itself reduced by high density levels (for example, the average speed of road traffic in Paris today is estimated to be 18 km/h).

The main functions of the daily territory represented by each cluster in the settlement system are habitat and production of food and services. In most cases, however, these main functions are no longer regulated locally. Its evolution is in the main determined elsewhere, in the nodes of the network where decision-making occurs. Some of the power centres in these networks are clearly identified, taking the form of a political or administrative capital, the headquarters of a multinational company. More usually, the forms of the flows of goods, services, labour and ideas which are responsible for the functioning of territories of the higher level of organization, on national, regional, or even multinational levels, are determined through multiple networks of negotiation, competition, and complementarity.

These networks are able to function thanks to the urban system. The forms of circulation which exist in the city and the urban system and that between localities and organization govern the system’s spatial structure, as defined by the distances between and size of the urban elements. At this level the distance between the main centres corresponds to the distances that can be traveled in a single day. This length of time is accepted as the average needed for the stages in a journey (Reclus, 1895) or for a return trip necessary to conduct fairly frequent negotiations in localities which are not visited every day (Cauvin et al., 1989), within the networks of exchanges which can exist between headquarters and subsidiaries, between large companies and their clients, between a capital and its main administrative relays. Although it is not very precise, this measure of a single day’s journey allows a much larger spatial extension for the territories of control and co-ordination than for those of daily life. For any given transport technology, not only the length of time determining the spacing of centres in a territory, but also the speed of circulation between them, is in general much greater than that possible within each individual centre (Orléan, 1993; Dagny, 1995).

Definition of a relative space for representing the relational possibilities both within and between urban centres is at the origin of the attention given to fractal geometry for the study of these objects (Arzilhau, 1985; Frankhauser, 1993; Battu and Longley, 1994). The relational space is strongly organized by density gradients and has a fractal structure. Applied to the description of urban structures, the notion of fractality has the advantage of offering a new reference model to replace that of a homogeneous occupation of space as is used for calculating density. Instead it is the direct expression of the anthropological occupation of space, by individuals and aggregates, and the fractality of the space of centrality in the description of urban objects. The effect of this is that urban densities (buildings, population, activities, journeys, etc.) do not have a homogeneous spatial distribution but vary systematically depending on their position in relation to one or several centres, following a dilution or fragmentation going from the centre to the periphery. A fractal model is therefore more suitable than the Euclidean model for analysing the properties of urban systems and for identifying their specific features. The early experiments on urban centres and urban networks have found a good fit for the fractal models at different levels of analysis (Le Bras, 1996). Some authors have also drawn attention to the existence of a clear discontinuity between the fractality of the internal spaces of towns and cities and the external space of interaction (Le Bras, 1995), which would appear to confirm that urban centres induce space-time structuring on two levels. Fractal representation is also an antidote for the short-comings of the spatial and location analysis of the geometry of urban systems in the Christaller models, which take no account of the differentiation in density which central places create by their presence. According to M.-C. Robin (1982), this “distorting” of the spatial models of central places, which was
Towards a theory of the evolution of settlement systems

A conception of the urban system as being responsible for the articulation between two levels of space-time relations implies seeking an alternative to a static formal theory. The present structure of urban systems represents rather a state of equilibrium or an equilibrium pervaded by some constraint at a particular moment. The spatial configuration of urban centres, their size and spacing, are the product of more or less constant processes whereby towns and cities are adapted to modifications in the proportions of space by the speeds of circulation, and an adjustment of their internal structure to the forms of social, functional and technical change.

Evolutionary theory situates the present structure of the urban system in the context of the general history of settlement systems. The latter is deeply influenced by the process of the urban transition, which is the process by which the urban centres of ancient cities, transformed by the processes of population and urban activities that occur in a biological predator-prey model, inter-urban competition is not a zero-sum game and, because it encourages innovation, is continually pushing back the factors that threaten to limit urban expansion. It is the source of the process of the social division of labour which expresses this impulse and generates new urban specialisations, and which must consequently be conceptualised as being endogenous to a model.

Competition between urban centres to attract distant resources and the product of innovations leads sooner or later (depending on the speed of communications) to a process of growth and transformation experienced by all the towns and cities in the territory under review. In the context of this generalised quantitative and qualitative expansion, urbanisation is interpreted as a diffusion expansion process affecting all parts of the urban system. This process is stochastic, made up of small adjustments by which the place of urban centres in the system is continually adapted (or corrected), with each tending to retain the same relative position. This same process explains the system’s hierarchical organisation as described by Gibran’s model.

The anticipations and imitations performed by the urban agents are non linear processes, however, which can produce bifurcations, that is, modifications to the structure defined by such adaptive processes. It is for this reason that in periods when systems of settlement are being established, during upheavals in the political organisation of territories, or when particular regimes of innovations appear (rapid growth, crises, decline) can significantly modify the position of towns in the system. Modelling these non-linear interactions requires a more complex dynamic than that emplaced or corrected, with the expansion of urban systems not appearing to be a continuous process over time. In the past, the range of urban interactions was for long restricted by the very slow speed of communications, thus preventing towns and cities from developing beyond the limits set by the size of the accessible territories, which were usually limited in extent. Historians have established that even in the wealthiest pre-industrial economies the average urban population did not exceed 10 per cent of the total population (assessed by the process of urbanisation that began with the earliest towns, some two or three thousand years after the appearance of agriculture (Bairoch, 1985), accelerated significantly when the transport system opened the way for an extension of the territories and networks that urban centres were able to dominate. The urban transition, which took urbanization levels up to 80 per cent in the industrial nations in little more than a hundred years, could in fact be interpreted as the fairly brutal condensation of a dynamics of inter-urban competition which although it had started long ago had previously been held in check by slow communications. The transformation of economies from an agrarian to manufacturing and then service basis was accompanied by far-reaching modifications in the distribution of settlement, resulting in a pattern of urban clusters that is regularly spaced, and appreciably more hierarchical, contrasted and concentrated than the previous settlement system (see, for example maps of the evolution of the European between countries of the Old World and those where settlement is of more recent origin (Moriconi-EBard, 1993). As an example, the average space between two urban centres with over 10,000 inhabitants is less than 15 km in China and in Europe, compared with nearly 50 km in North America. Similarly, an index measuring the size of hierarchical contrasts between these urban populations, expressed by the slope of the ‘rank-size’ distribution, is less than 1 in Europe or only 1.2 in North America, whereas its average value is the same (1.05) in the developed and developing countries. In addition, all these systems exhibit a long-term trend to more pronounced hierarchical contrasts. This could be explained by the initial advantage of the largest centres because those innovations are diffused hierarchically (a strengthening of the hierarchy from above). But it also occurs systematically as a result of the time-space convergence, whereby the largest urban centres are able to short-circuit the clienteles of the smaller centres that occupy an intermediary position, also inducing a ‘simplification from the bottom’ of the urban hierarchy. The most frequently observed deviation from the Gibran model of urban system adaptation between urban growth and city size, can thus be incorporated into an evolutionary urban theory in which towns and cities are responsible for articulating two types of relational spaces situated on two different time scales and which take into account the historical trend, based on technological progress, to a contraction of geographical space, as measured by journey times.

The construction over time of an increasingly hierarchical settlement system, marked by growing contrasts in the size of its elements, is thus a consequence of the optimal organisation of circulation in a territory at a given moment (which disqualifies static explanatory theories such as central place theory and spatial economic equilibrium), but rather of the gradual and gradual accumulation of a system originally created for a particular speed of circulation and which has subsequently been transformed through the effect of faster speeds. In the course of this evolution, the system has also enabled societies to gradually free themselves from local ecological dependence, on the resources offered by their particular site, while reducing the hazards and uncertainties that threaten their survival and development, by considerably enlarging the scope for substitution and complementarity of resources, through the ability to draw on those of increasingly distant sites.

With this systemisation it becomes possible to combine into a unified concept the organisational duality between, first, the hierarchical organisation of the urban systems that assure the territorial control and services, and second, their formation of a network for the exercise of various specialist functions. Unlike in central place theory, we do not have to accept a priori that there are functions which are differentiated by their presence or range, and then order the urban centres into a hierarchical configuration according to their size and the distance between them. These functions are actually created in the urban centres, they are a-

Geographia Australis. 82 B (2000) 2
space in which relations are structured by a centre that can be reached in an hour, the surface that entity covered in space would be roughly double what it was in 1800, and roughly double the 1950 figure in 1990. Examination of this object, which in the relational space-time would be consistently defined as a "town" or "city", is unlikely to reveal the pattern of growth followed by decline in the central area and continuous growth in peripheral districts ever further (physically) from the centre, as proclaimed by the theorists of "counterurbanization" (for a critical review see Davies, 1994) and whose forms have been well described by the specialists of peri-urbanisation. On the contrary, the likelihood is that the town or city defined in an evolutionary way will be found to present a sustained trend to growth, compared with its peripheral. Similarly, many of the debates over the "urban exodus" to smaller local spaces could also be interpreted as extensions, in the light of the extent of the spatial range of interactions.

These notions of evolutionary spatial entities can also help geographers to improve their contribution to certain questions of forecasting: regarding the nature of the post-urban transition, that is about the future evolution of urban systems, the increasing speed of inter-urban circulation and the extension of relational networks leave no doubt about the continuing trend to concentration: in a space-time that is smaller in relation to the space-context, an urban network comprising fewer urban places is the transformed equivalent of the earlier urban network. Finally, an evolutionary perspective invites us to take a different view of what are assumed to be the optimal forms in regional planning: the same mechanisms of change may have produced slightly different forms of urban system [for example, French or Roman centralised network as compared with the networks found in the Rhineland], which constitute so many variants that are equally valid and not necessarily in contradiction with each other, with the capacity of adapting to the changes that are occurring without their structures necessarily becoming more similar. There is no optimal urban system, there are just systems which continue to adapt.

Acknowledgment
This paper is in part a translation of a paper published in French in L'Espace Géographique (1997, 2, 119–134). We thank Godfrey Rogers for his help in translation.

References
DENISE PUMAIN


