Dynamics of Transportation Infrastructures

N. Nakicenovic

NAE Workshop on The Evolution of Infrastructures

PREFACE

This paper was prepeared for the National Academy of Engineering Workshop on "The Evolution of Infrastructures", Woods Hole, U.S.A., 11-13 August 1986. It is based on research conducted together with Arnulf Grübler within the scope of the Dynamics of Technology Project, at the International Institute for Applied Systems Analysis (IIASA), Austria.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

Dynamics of Transportation Infrastructures

1 INTRODUCTION

The objective of this paper is to assess, on the basis of a number of indicative examples, the time needed to build new and replace old transportation infrastructures and energy transport systems. It will be shown that both the growth and senescence of transport infrastructures evolve as regular processes that can be described by S-shaped (logistic) curves. Not all growth and senescence phenomena can be described by simple logistic functions, sometimes more complex patterns can be observed. However, these more complex patterns can often be described by envelopes which can be decomposed into a number of S-shaped growth or senescence phases. Two typical cases include different growth pulses that follow each other with a saturation and a period of change in between and also simultaneous substitution of more than two competing technologies. In the first case S-shaped pulses that follow each other usually represent successive improvement in performance such as the aircraft speed records where the first pulse is associated with piston engines and the second with the new technology, the jet engine. In the second case simultaneous substitution of a number of competing technologies would usually be described by increasing market shares of new technologies and decreasing market shares of old technologies. We will show that the evolution of transport infrastructures can be described both in terms of performance or productivity of competing technologies and in terms of their market shares.

In both cases we will show that the development of transport networks is subjected to regular patterns of change that can in principle be used for forecasting future trends. Most of the examples are taken from the United States, since the United States has been one of the few countries to experience most of the technological changes that occurred during the last 200 years. However, because most of these technological changes were subsequently disseminated throughout the world, the examples also indicate the dynamics of these processes elsewhere.

Within the scope of this paper, the use of the term infrastructure is rather narrow, referring only to transportation and energy grids and networks, and other components of these two systems. These two systems are interesting because they played a crucial role in the economic and technological development process, are very capital intensive and, in general, have long lifetimes. The analysis of the historical development of these two systems will include a quantitative description of performance improvement, the general evolution of a particular infrastructure, and the replacement of old by new technologies and infrastructures in terms of their relative market shares. We use the term performance as a multidimensional concept (i.e., as a vector rather than a scalar indicator) and, where appropriate, measure the size of an infrastructure as a function of time.

The first section describes the evolution of transport systems and their related infrastructures. The analysis starts, somewhat unconventionally, with the youngest technologies, aircraft and airways, and ends with the oldest transport networks – canals and waterways. The second section describes the evolution of energy consumption and pipelines as an example of dedicated transport infrastructures. In the Appendix we briefly present the methods used for our analysis and the statistical tests.

2 TRANSPORTATION

2.1 Aircraft

Aircraft are the most successful of the advanced modes of transportation. Other concepts of rapid transport such as very high speed trains have shown some limited success but are not as universally used as aircraft. In fact, the rapid expansion of air travel during recent decades has its roots in the developments achieved in aerodynamics and other sciences many decades ago, and especially in the engineering achievements between the two World Wars. The DC-3 Airliner is often given as the example of the first "modern" passenger transport because in many ways it denotes the beginning of the "aircraft age".

The use of aircraft for transport has increased ever since and their performance has improved by about two orders of magnitude. Figure 2.1 shows the increase of air transport in the world measured in billion passenger-kilometers per year and refers to all carrier operations including those of the planned economies. The logistic function has been fitted to the actual data and it indicates that the inflection point in the growth of air carrier operations occurred about ten years ago (i.e. about 1977). This means that after a very rapid, exponential growth period there is less than one doubling left until the estimated saturation level is achieved after the year 2000. The growth rate has been declining for about ten years and if our projection is correct it will continue to do so until the total volume of all operations levels off after the year 2000 at about a 40 percent higher level than at present.

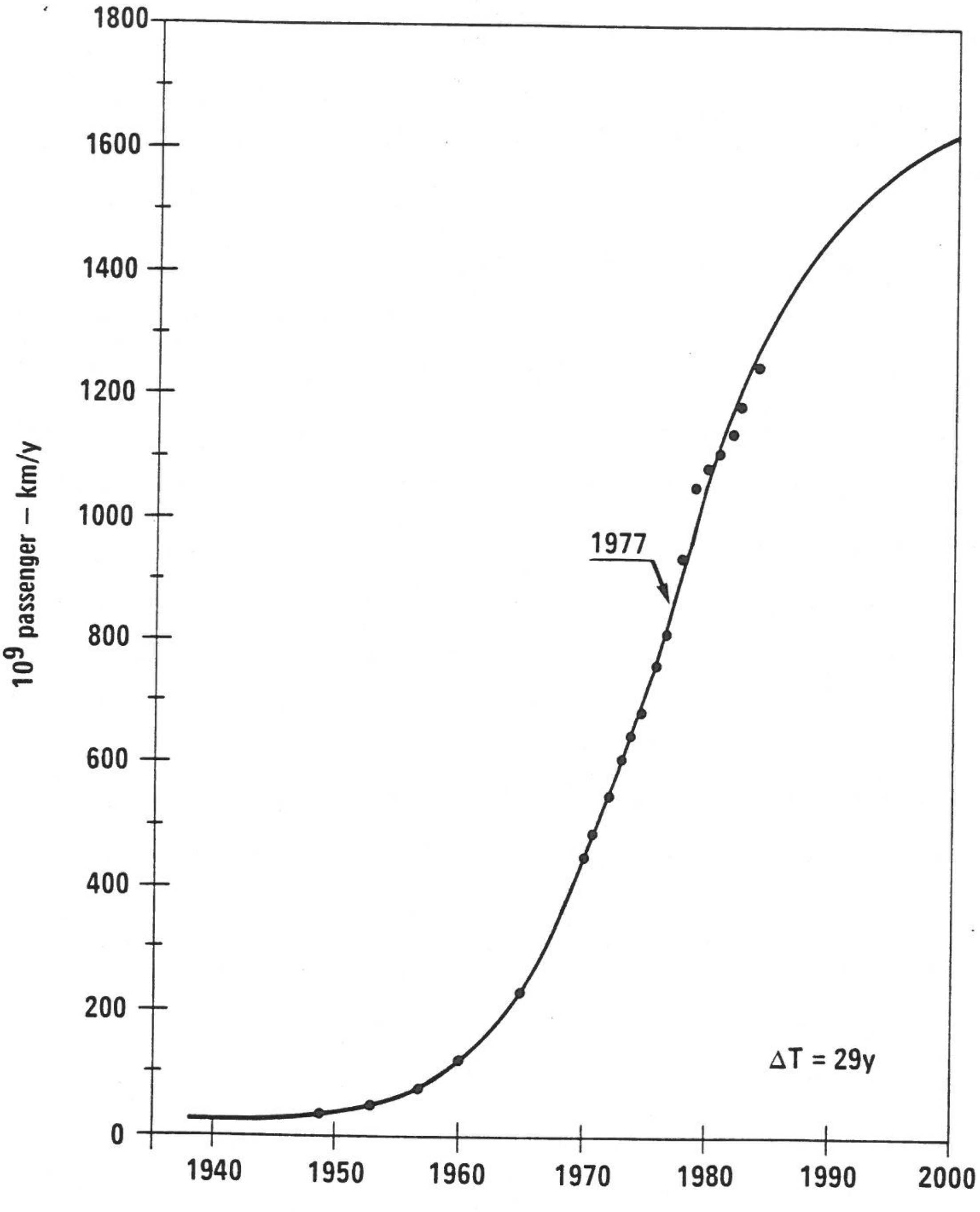

Figure 2.1 World Air Transport – All Operations.

Figure 2.2 shows the same data and fitted logistic curve transformed as $x/(\kappa-x)$, where x denotes the actual volume of all operations in a given year (from Figure 2.1) and κ is the estimated saturation level. The data and the estimated

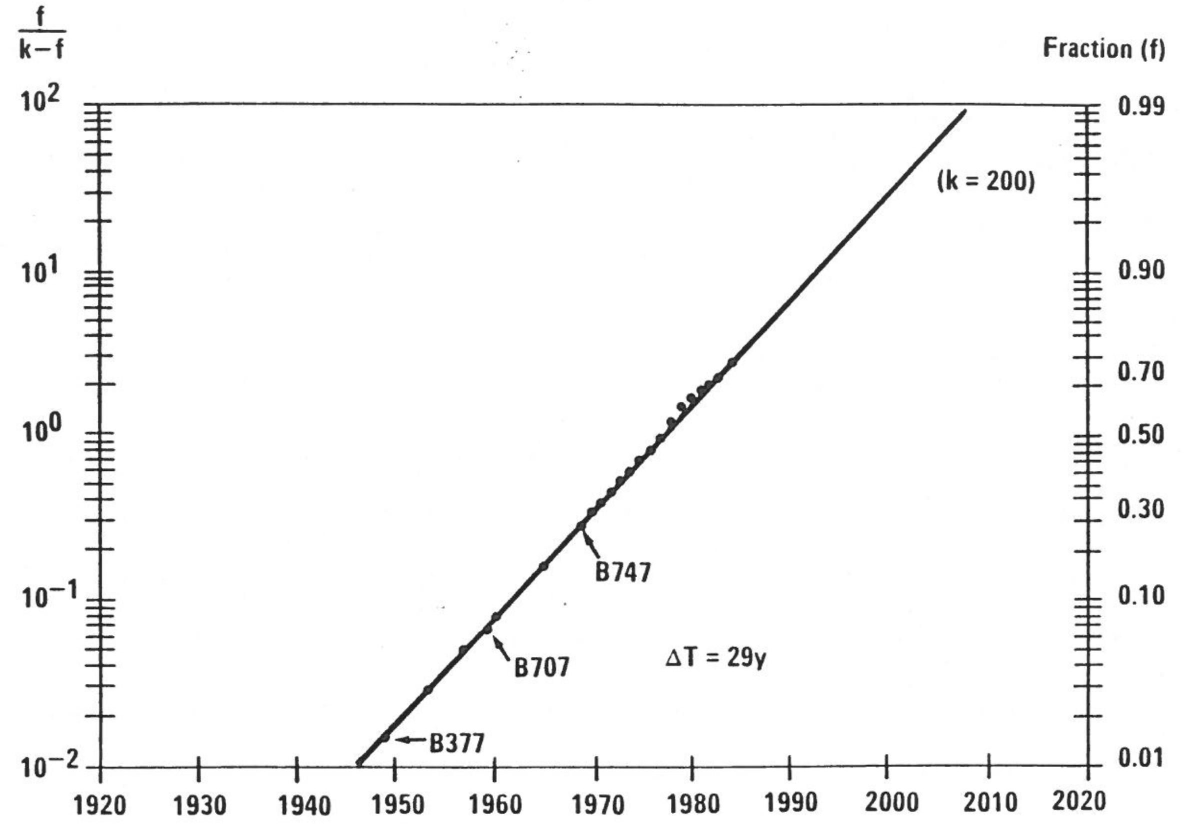

Figure 2.2 World Air Transport – Logistic Plot.

logistic trend line are plotted in Figure 2.2 as fractional shares of the saturation level, $f=x/\kappa$, which simplifies the transformation to $f/(1-f)$. Transformed in this way, the data appear to be on a straight line, which is the estimated logistic function. [1]

Perhaps the most interesting result is that it took about 30 years until world air transport growth reached the inflection point (i.e. or about half of the estimated

1 One general finding of a large number of studies is that many growth processes follow characteristic S-shaped curves. Logistic function is one of the most widely applied S-shaped growth curves and is given by: $x/(\kappa-x)=\exp(\alpha t+\beta)$, where t is the independent variable usually representing some unit of time, α, β and κ are constants, x is the actual level of growth achieved, while $\kappa-x$ is the amount of growth still to be achieved before the (usually unknown) saturation level κ is reached. Taking logarithms of both sides gives the left side of the equation to be expressed as a linear function of time so that the secular trend of a logistic growth process appears as a straight line when plotted in this way. Substituting $f=x/\kappa$ in the equation, expresses the growth process in terms of fractional share f of the asymptotic level κ reached, i.e. the equation becomes:
$f/(1-f)=\exp(\alpha t+\beta)$.

saturation level) and that within two decades saturation level would be reached. The crucial question implied by this result is what will happen after such a possible saturation? Can we expect another growth pulse, a decline, or instability of changing periods of growth and decline? We will try to show that it is most likely that a new period of growth associated with new technologies will follow the projected saturation.

To understand the implications of a possible saturation in world air transport operations and possible future developments thereafter, it is necessary to look at the transport system in general, comparing aviation with other modes of transport, and to analyze the various components of the air transport system itself. The aircraft, airports and ground services obviously represent the most important components of the air transport infrastructure. Commercial aviation infrastructure is different from that of other competing transport systems such as roads or railways. Airports are distributed and not connected as are pipelines or roads, although aircraft represent an infrastructure of complex elements in the same way hospitals or schools do.

There are a number of possible ways to describe the global fleet of commercial air transport and how it developed. An obvious descriptor of the fleet would be the number of aircraft in operation worldwide. In fact, this number increased from about 3,000 in the 1950s to almost 10,000 in the 1980s. However, during the same time the performance or carrying capacity and speed of aircraft increased by about two orders of magnitude. Thus, the size of the fleet is not the most important descriptor, since much of the traffic is allocated to the most productive aircraft operating between the large hub-airports while other aircraft constitute the feeder and distribution system to destinations with lower traffic volume. The analogy between electricity grids or road systems is very close – large aircraft correspond to high-voltage transmission lines or primary roads.

Figure 2.3 shows the improvement over time of one important performance indicator for commercial passenger aircraft, i.e., the increase of carrying capacity and speed (often called productivity) measured in passenger-kilometers per hour. Each point on the graph indicates the performance of a given aircraft when used in commercial operations for the first time. For example, the DC-3 was introduced in 1935 with a performance of about 7,400 passenger-kilometers per hour (about 21 passengers at about 350 km/h) whereas the B747 was introduced in 1969 with a performance of about 500,000 passenger-kilometers per hour (about 500 passengers at about 1,000 km/h). The largest, planned B747s can carry almost 700 passengers and are therefore about 100 times as productive as the DC-3 50 years ago.

The upper curve represents a kind of "performance feasibility" frontier for passenger aircraft, since the performance of all commercial air transports at the time they were introduced is either on or below the curve. Furthermore, long-range aircraft on the performance feasibility curve were commercially successful, while all other planes whether they were successful or not are below the curve. Thus, at any given time there appears to be only one appropriate (best) productivity specification for long-range passenger planes. Since the more recent jet transports all fly at about the same speed, the appropriate design for a new, hypothetical long-range transport should allow for a capacity of more than 700 passengers. According to the estimated curve, the asymptotic capacity for the largest aircraft is about 1,200 passengers at subsonic speed (i.e. 1.2 million passenger kilometers per hour or ten billion passenger kilometers per year). This implies that the next pulse in the aircraft performance, if it occurs, would take place after the saturation phase and would start at about 1.2 million passenger

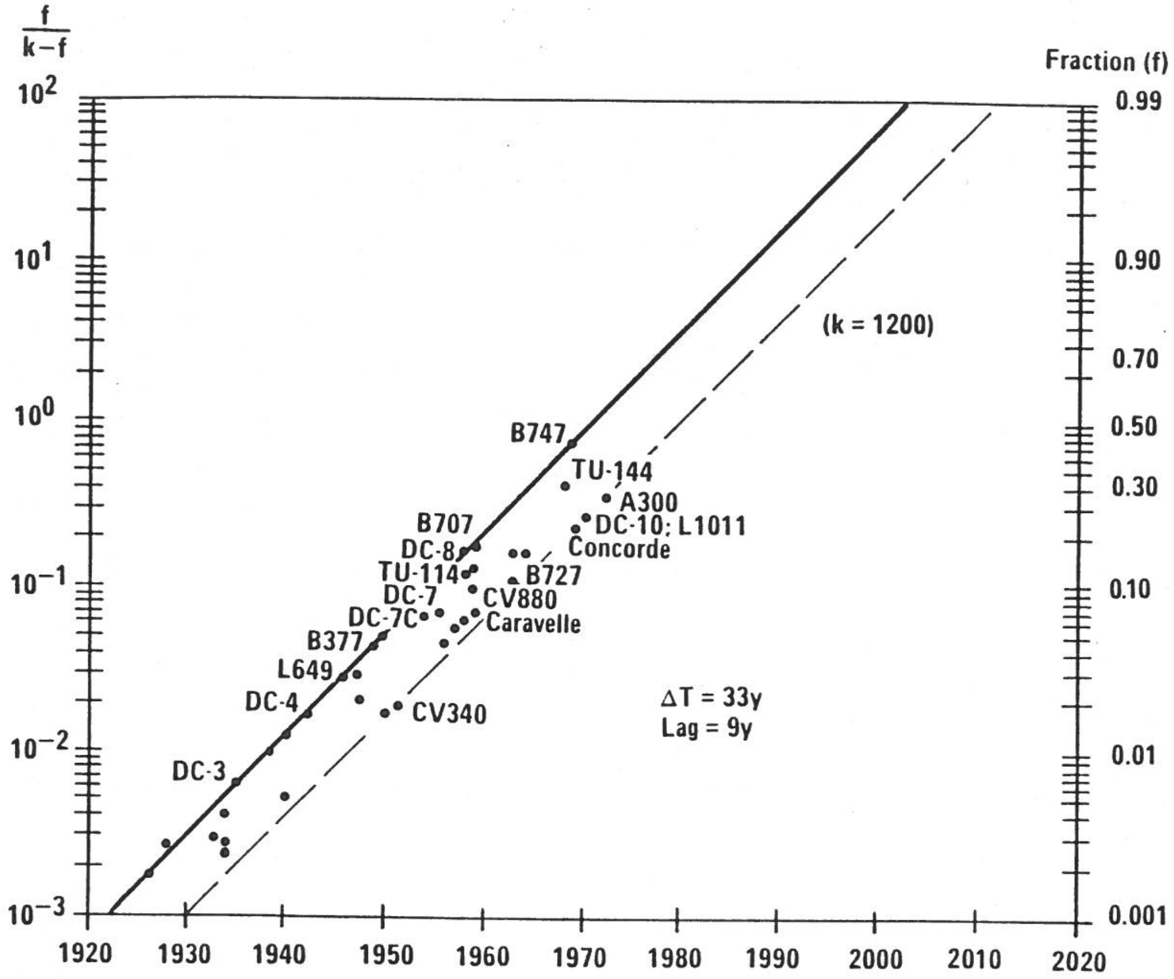

Figure 2.3 Passenger Aircraft Productivity.

kilometers per hour and grow to much higher levels by at least one if not two orders of magnitude. This indicates the possibility that the long-range aircraft after the year 2000 will be larger and faster. Another interesting feature in Figure 2.3 is that the productivity of all passenger aircraft is confined within a rather narrow band between the performance feasibility curve and a "parallel" logistic curve with a lag of about nine years. This "parallel" logistic curve represents the growth of the world air transport from Figure 2.2.

It took about 30 years for the performance of the most productive aircraft to increase from about one percent of the estimated asymptotic performance to about half that performance[2] (e.g. the DC-3 represents the one-percent achievement and the B747 roughly the 50-percent mark). In many ways, the achievement of the 50 percent level represents a structural change in the development of the whole passenger aircraft industry and airlines. Since we are dealing with *S*-shaped growth (i.e. a logistic curve), the growth process is exponential until the inflection point (i.e. the 50 percent level) is achieved. This means that at the beginning there are many doublings in the productivity of aircraft within periods of only a few years. However, once the inflection point (the 50-percent level) is reached, there is only one doubling left until the saturation level is reached. In the example from Figure 2.3, this development phase occurred in 1969 with the introduction of the B747. Since the B747 could in principle be stretched by about a factor of two, it could also remain to be the largest long-range aircraft during the next two decades. Thereafter, as was mentioned above, a new phase of growth is conceivable with either larger or faster long-range aircraft and in the more distant future possibly both.

Before the inflection point, stretching does not help for more than a few years due to rather frequent doublings in productivity so that principally new solutions are necessary. This also means that a wrong model on the market can be a crucial but nevertheless reversible mistake provided that the manufacturer has large enough resources at his disposal to launch a new, improved performance model on the market after a few years. However, after the inflection point this is no longer a possible strategy since those aircraft that are successful can be stretched to meet

2 We define Δt as the time elapsed between the achievement of one and 90 percent of the saturation level κ, i.e. in this example $\Delta t = 30.8$ years. Due to the symmetry of the logistic function, the same time is required for the increase from 50 to 99 percent of the saturation level. An alternative definition of Δt is the time elapsed between the achievement of 10 and 90 percent level. In this case Δt would be slightly different from the other definition, but for all practical applications both definitions can be used interchangeably.

the market demand through saturation. Therefore, toward the late 1960s designing new long-range transport suddenly became riskier than before since a second chance of launching a new model after a few years was no longer possible. The B747, for example, had the appropriate productivity, while two other competitors, the DC-10 and L1011 were too small. Launching a 700 to 800 passenger aircraft in the near future could constitute a great risk since a stretched B747 can do the same job. A smaller long-range aircraft would most likely constitute a failure since it would fall short of the primary market requirements. The only routes open for smaller longer-range transport such as the MD 11 or Airbus 340 would be those for which the B747 is too large. Since in the long-run most of the long-range traffic will be between large and larger hubs and less in the form of direct traffic between smaller airports, the market niche for long-range aircraft with less than 500 passengers would be very limited. A feasible alternative would be to design the MD 11 and the Airbus 340 as wide-body shorter-range transports that would be suitable for frequent cycles. The design of a cruse supersonic aircraft with a capacity for 300 to 400 passengers, or a hypersonic transport with say 200 to 300 passengers, may therefore be a better strategy for the first years of the next century. To some extent this also explains why Concorde cannot be a commercial success – with a capacity of 100 passengers it was 150 passengers short of being a serious competitor to the B747. How probable is the development of a large cruse supersonic or hypersonic transport? S-pulses do not occur alone, but usually in pairs – thus at saturation, structural change will occur leading to a new growth pulse (probably S-shaped) – leading to new productivity requirements and therefore also to supersonic or hypersonic transport. The alternative is that air transport would saturate permanently over the next few decades and the wide-bodied families of subsonic transports would actually constitute the asymptotic technology achievement. We will show below that this alternative is unlikely because air transport may become the dominant mode of passenger (long-distance) travel in the next century.

Figure 2.4 shows the performance improvement of civil aircraft engines since the beginning of aviation. The first piston engine visible on the plot is the French Antoinette engine rated at 50 hp in 1906 and the last is the American Wright Turbo Compound rated at 3400 hp in 1950. These represent an improvement in power of almost two orders of magnitude during a period of 44 years, and about 90 percent of the estimated saturation level for piston engines at about 3800 hp. A parallel development in the maximal thrust of jet engines follows with a lag of about 30 years: starting in 1944 with the German Junkers Juno 004 (rated at 900 kg) and ending with the American Pratt and Whitney JT9D in the early 1980s with a 90 percent estimated thrust saturation level of about 29000 kg. Both pulses in the improvement of aircraft engines are characterized with a Δt of about 30 years. The midpoints of the two pulses occurred in 1936 and 1967 (i.e., the respective inflection points) and coincide with the introduction of the DC-3 and B747. In fact, the Pratt and Whitney Twin Wasp, introduced in 1930, became the power plant for the DC-3. It also served as the basis-engine for subsequent and more powerful derivatives such as the Wasp Major introduced in 1945 toward the end of the aircraft piston engine era. Thus, certain parallels are visible to the dynamics of passenger aircraft development beyond the obvious similarity in the time constant (Δt of about 30 years). Namely, those engines introduced slightly before the inflection point, such as the Twin Wasp, are "stretched" by doubling the cylinder rows to increase the power. The Wright Turbo Compound represents the last refinement in aircraft piston engines, the major additional feature being that turbo-charging was used to derive shaft-power, otherwise the engine was a direct derivative from the Wright Cyclone series introduced in the early 1930s with Cyclone 9 (which originally powered the DC-2).

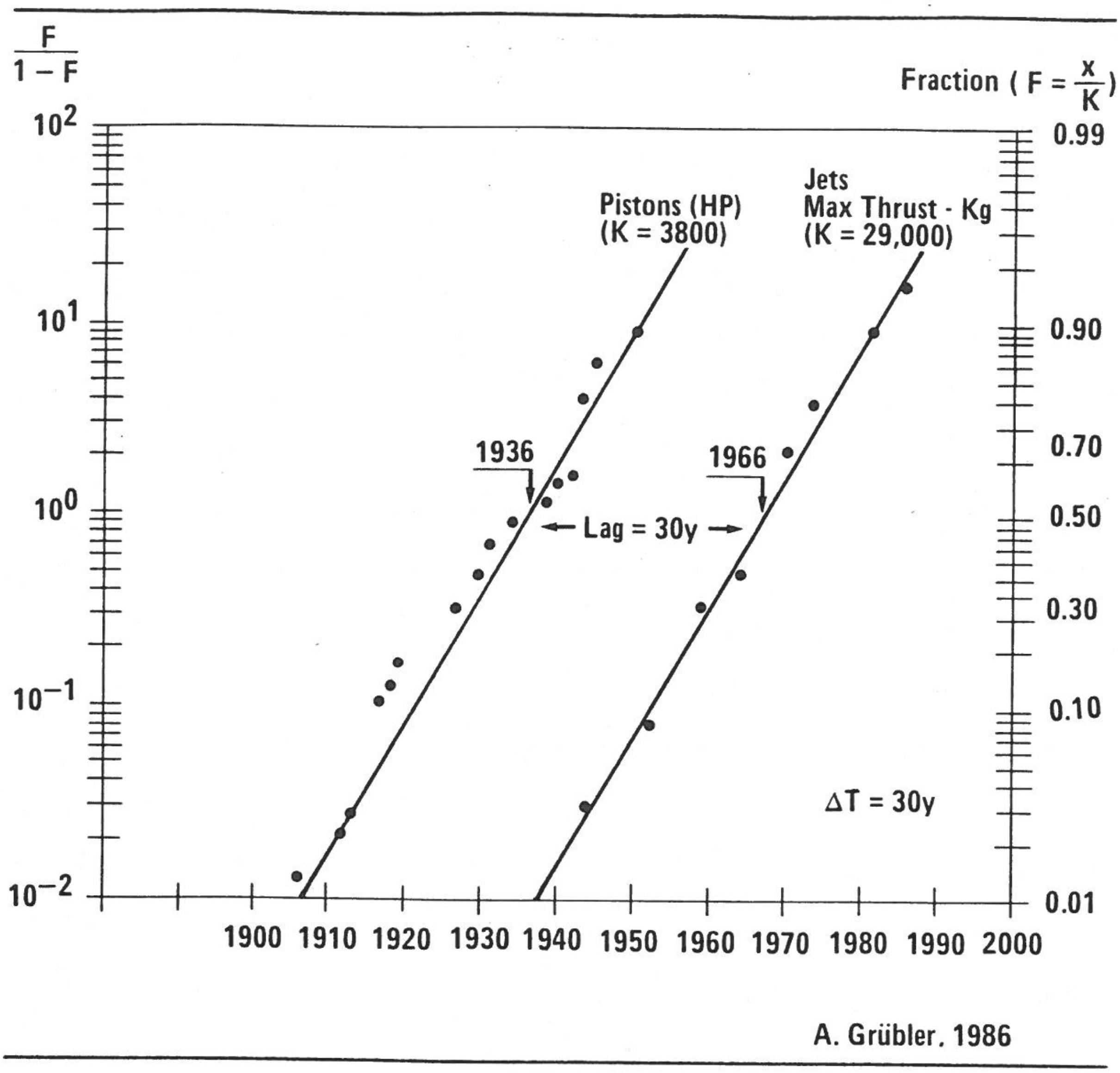

Figure 2.4 Performance of Aircraft Engines.

This result indicates that the time constant for the development of passenger aircraft as one of the most modern transportation modes appears to be about 30 years. As mentioned above, the standard industry design emerged during the depression years (i.e., the 1930s). Thirty years later the B747, the first wide-body jet to enter service, represented one of the most significant improvements in commercial transport and its productivity represented half of the estimated industry saturation level that may be approached toward the end of this century. Thus, within a period of 60 years the life cycle is completed, from standardization and subsequent rapid growth characterized by numerous improvements, through the inflection point when the emphasis changes to competition characterized by cost reductions and rationalization. These characteristics of the industry, including steady performance improvement, are well illustrated by the development of the B747 family.

The emergence of a new aircraft engine at the beginning of a new phase of growth is very likely toward the end of the century after the projected saturation has been reached. Furthermore, the characteristics of the long-range aircraft in this hypothetical growth pulse are implicit in our analysis. The saturation in the thrust of a turbo-jet and the turbo-fan engines indicates a new generation of engines emerging over the next few decades with much higher thrust ratings and growth potential. In our opinion the only realistic candidates are the ram-jet or scram engines with efficient off-design flight engine characteristics or perhaps HOTOL type air-breathing rocket engines. Thus, if a new pulse in the improvements of aircraft engines should occur, it will most likely be associated with some variant of variable-cycle air breathing engine with a very high Mach number and perhaps even balistic flight potential. This trend would single out methane as the fuel of choice between Mach numbers of about 3,5 to about 6, with a small overlap of hydrogen starting at about Mach 5. Thus cryogenic methane powering cruse supersonic aircraft may become the mode of transport with the highest productivity after the year 2000. Next we will investigate the development of an older system that was a strong competitor for the airlines in long-distance passenger transport – the automobile.

2.2 Automobile

At the beginning of the century, very few proponents of the automobile envisaged that it would develop and disseminate throughout the world so rapidly during the last 100 years. But the first *horseless carriages* already posed an alternative to the horse-drawn buggies and wagons. As a commercial and recreational vehicle, the motor car offered many potential advantages compared to other modes of transportation, especially over animal-drawn vehicles. Perhaps the most important advantage was the possibility to increase the radius of local transport; being faster the automobile allowed many entrepreneurs to expand their circle of customers and offered a more flexible mode of business and leisure transport.

At the beginning the railroads were not challenged by the automobile, but rather helped their expansion since they offered an efficient form of long-distance transport that combined well with the use of motor vehicles for local, urban and rural road transport. Within a few decades, however, the automobile became an important form of transport for both local and long-distance passenger travel in the United States. Since the 1930s up to the present, the total mileage traveled by automobiles, and motor vehicles in general, was divided almost equally between rural and urban travel.

The automobile had a relatively late start in the United States compared with European countries (e.g. France, Germany and the United Kingdom). In 1894 four motor vehicles were recorded to be in use in the United States. However, the expansion of the automobile fleet thereafter was impressive – 16 vehicles were in use in 1896, 90 in 1897, 8,000 in 1900, almost half a million ten years later and more than one million after another two years. Thus, both in terms of production and number of vehicles in use, the United States quickly surpassed European countries.

Figure 2.5 shows the rapid increase in the number of cars used in the United States. By the 1920s more than ten million automobiles were in use on American roads and the 100 million mark was surpassed in 1970. Figure 2.5 also shows that the expansion of the automobile fleet is characterized by two distinct secular trends with an inflection in the 1930s followed by less rapid growth rates. Since the two secular trends of the curve appear to be roughly linear on the logarithmic scale in Figure 2.5, the automobile fleet evolved through two exponential *pulses*. Thus in this example we can observe that the growth of the automobile fleet did not follow a simple, single S-shaped growth pulse.

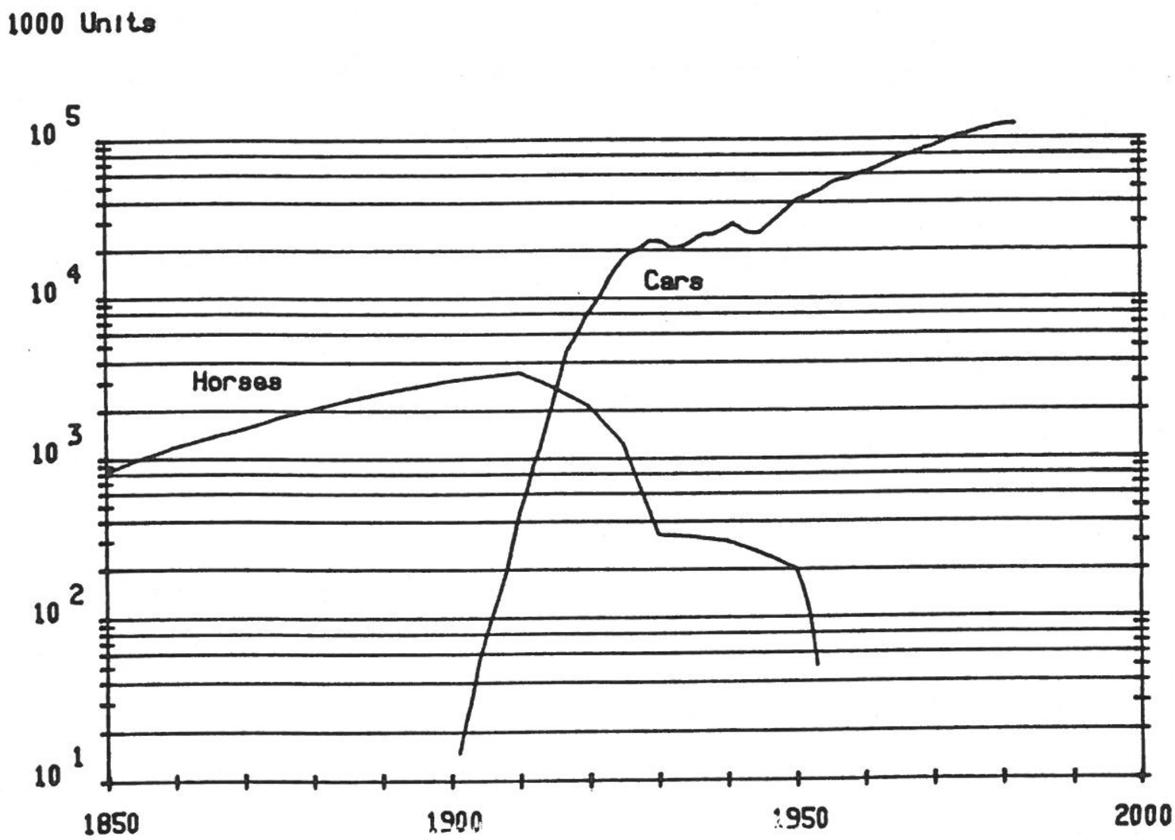

Figure 2.5 Number of Automobiles and Road Horses (and Mules) in US.

Our working hypothesis is that the two trends indicate two different phases of dissemination of motor vehicles in the United States. The first characterizes the substitution of horse-drawn road vehicles, and the second the actual growth of road transport after animal-drawn vehicles essentially disappeared from American roads. Only after the completion of this substitution process did the automobile emerge as an important competitor to the railroads for the long-distance movement of people and goods, and perhaps, also as a competitor to urban transportation modes, such as the tram, subway or local train. Thus, the first expansion phase is more rapid since it represents a "market takeover" or expansion in a special niche, whereas the second represents the actual growth of the road vehicle fleets and their associated infrastructures such as highway systems.

Due to the obvious problems associated with the lack of historical records about the exact number of horse-drawn vehicles in the United States during the first decades after the introduction of the automobile in 1895, we can only approximately describe the assumed substitution of the horse by the motor car during the first, more rapid, expansion phase of the motor vehicle fleets. As a rough approximation of this substitution process, we use the number of draft animals (road horses and mules) and automobiles given in Figure 2.5. Sometimes horse and saddle were used as a "road vehicle", but often more than one horse was used to pull buggies and wagons. City omnibuses required about 15 horses to be in use the whole day and a stage coach probably even more, since the horses were replaced at each station. Thus, Figure 2.5 may overemphasize the number of horse-drawn vehicles if the number of draft animals is used as a proxy for the number of vehicles actually in use. On the other hand, we have not included farm work animals in Figure 2.5, although certainly farm horses were also used as a means of transport, especially in the rural areas.

Thus, the disadvantage of this rough comparison of the number of animal-drawn vehicles and motor cars is that the estimates of non-farm horses and mules are certainly not very accurate and are unevenly spaced in time. Despite this disadvantage, Figure 2.6 indicates that the automobile replaced horse- and mule-drawn road vehicles during a relatively short period and that the substitution process proceeded along a logistic path. [3] Motor vehicles achieved a one-percent share in road vehicles shortly after 1900, and a 50 percent share in 1916. A complete takeover of the "market" occurred in 1930 with 23 million cars over 0.3 million road horses and mules compared to almost two million ten years earlier. This result indicates that the inflection point in the growth of the automobile fleet from Figure 2.5 actually coincides with the end of the substitution of animal-drawn road vehicles and explains the apparent "saturation" in the growth of motor vehicles perceived by many analysts during the late 1920s and early 1930s. This perceived saturation marks the beginning of a new phase in the motorization of America, with growth rates comparable to those of the expansion of horse-drawn vehicles before the automobile age. Seen from this perspective, the growth in the number of *all road vehicles* is a continuous growth process over the entire period from 1870 to date. Figure 2.7 shows the secular increase of all road vehicles, horse-drawn and motor powered, as a *logistic growth process* with an apparent saturation level of about 350 million vehicles after the year 2030 and Δt of about 100 years.

Thus, we are dealing with two simultaneous processes, the growth of road transport in general and the substitution of horse carriages and wagons by automobiles. Because these two secular developments overlap in time, together they have the effect of producing two growth trends in the automobile fleet with an

3 One general finding of a large number of studies is that substitution of an old technology by a new one, expressed in fractional terms, follows characteristic *S*-shaped curves. Fisher and Pry (1971) formulated a simple but powerful model of technological substitution by postulating that the replacement of an old by a new technology proceeds along the logistic growth curve: $f/(1-f)=\exp(\alpha t+\beta)$, where t is the independent variable usually representing some unit of time, α and β are constants, f is the fractional market share of the new competitor, while $1-f$ is that of the old one.

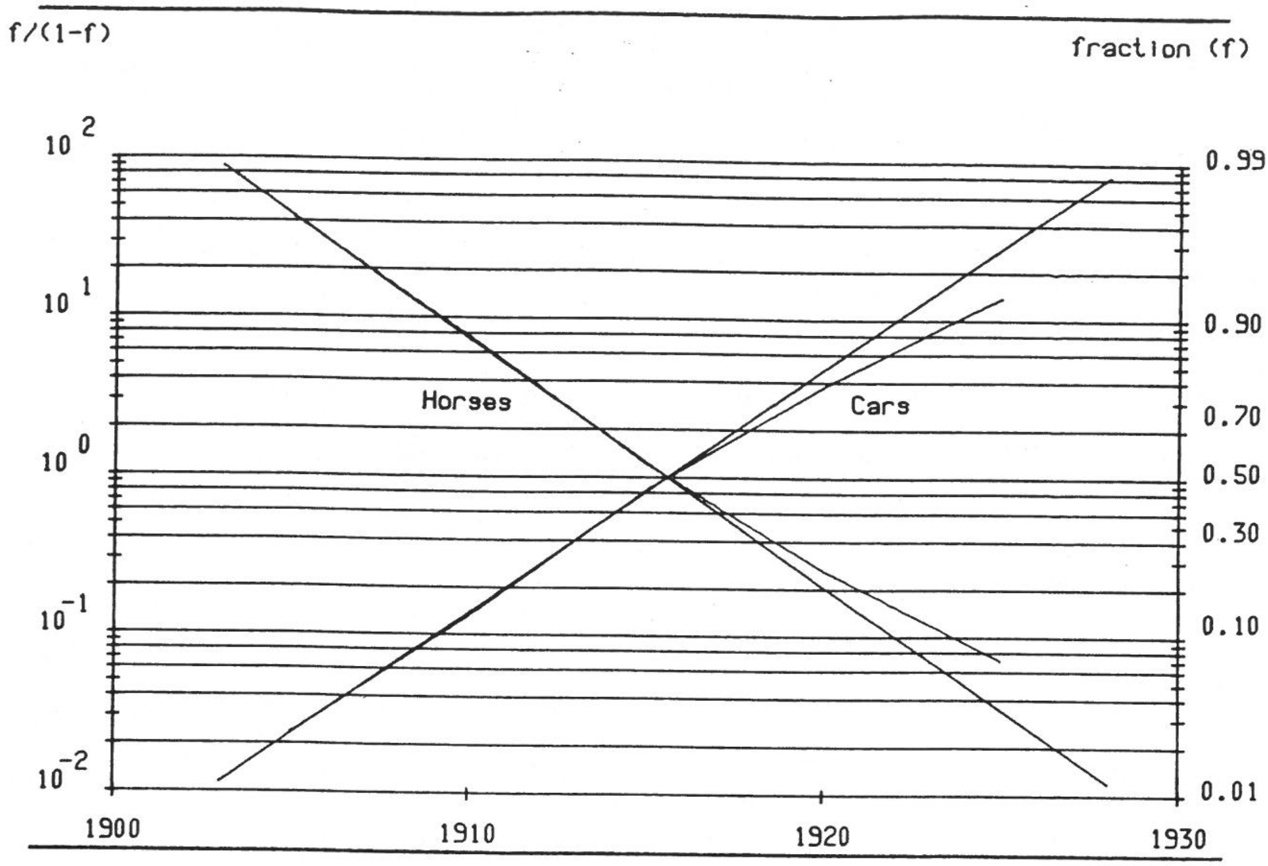

Figure 2.6 Substitution of Horses by Cars, US.

inflection point in the 1930s marking the structural change in the composition of the road vehicle fleets. However, the growth of *all* road vehicles (e.g. horses and cars) results in one single logistic trend. Thus, like aircraft, road vehicles constitute a kind of distributed infrastructure that expands in time as one connected system.

In fact, both air and road transport systems require elaborate and sophisticated infrastructures. Perhaps the most obvious examples are airports and supporting ground systems for aircraft, and roads and service infrastructure for road vehicles. In the case of trains and aircraft it is self evident that airports and railroads were constructed with the single purpose of providing the infrastructure for these transport modes. However, in the case of the automobile it is not so clear whether the roads provided the niche into which the motor car expanded or if good roads had to be developed because the car fleets were expanding. Whether the

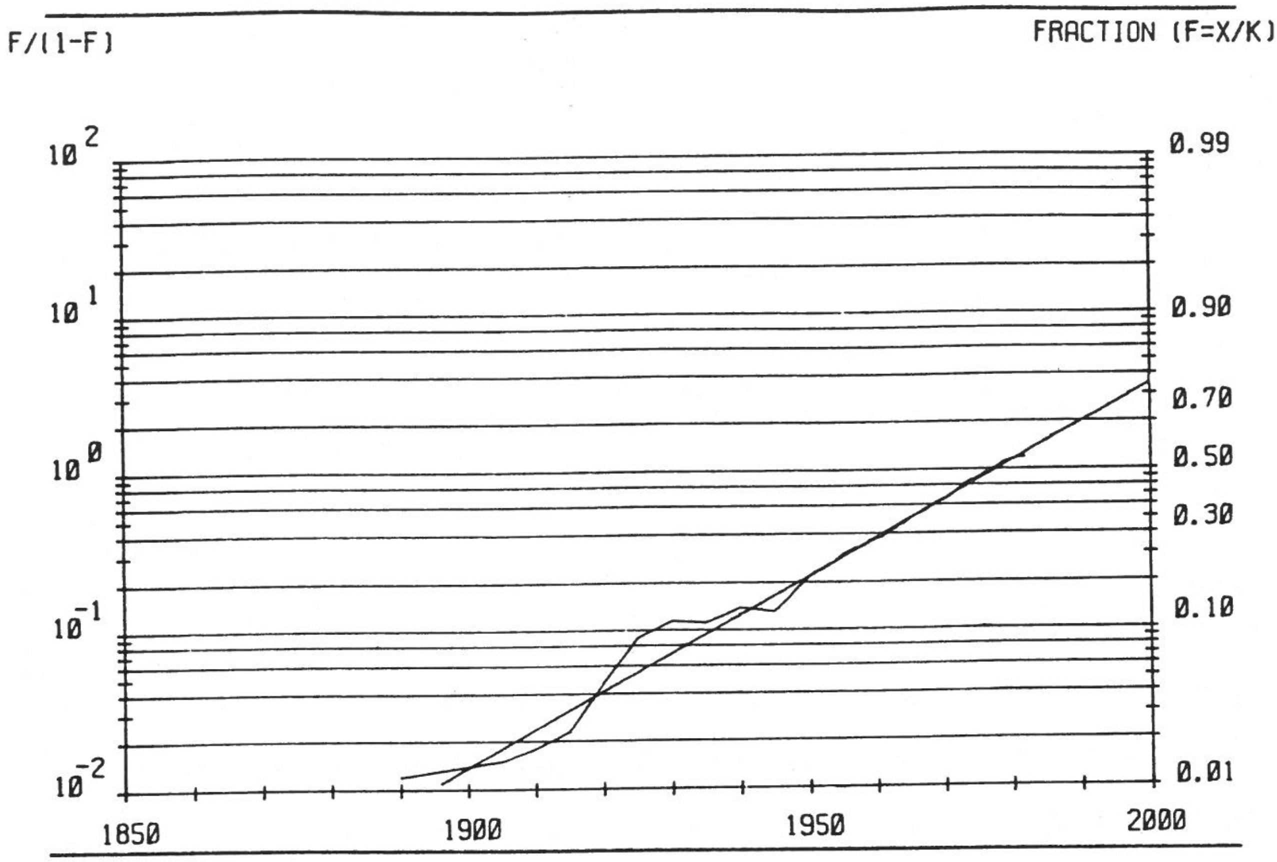

Figure 2.7 All Road Vehicles in Use, US.

automobile caused the need for good roads, or the construction of good roads caused the development of the automobile industry (an argument already debated in the 1930s, see e.g. Epstein, 1928), the expansion of road vehicle fleets is paralleled by the mileage growth of surfaced roads, although the total mileage of *all roads* increased very slowly from 3.16 million miles in 1921 to 3.85 million miles in 1981. The length of all roads has remained almost unchanged during 60 years and represents one of many indicators that *do not* portray logistic growth trends. However, it is possible that the mileage of *all roads* increased during earlier times, e.g. colonial period, and was saturated 100 years ago or even earlier.

Figure 2.8 shows the total road mileage in the United States and the mileage of urban streets (earlier defined as municipal streets), rural roads and all urban and rural surfaced roads (bituminous penetration, asphalt, concrete, wood, stone and

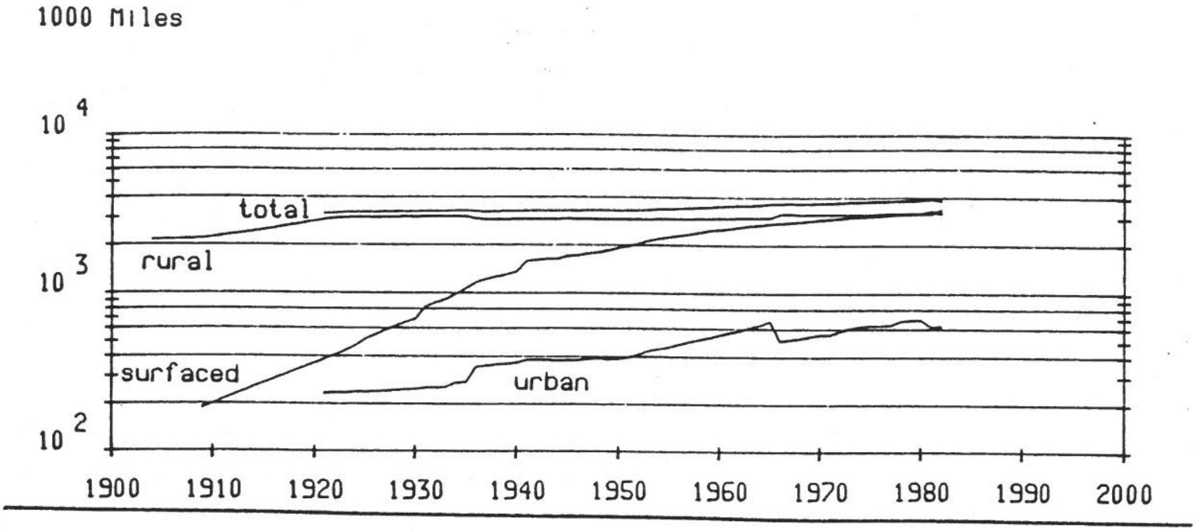

Figure 2.8 Mileage of all Roads, US.

other). The figure illustrates that the growth of surfaced roads paralleled the growth of all road vehicle fleets while the mileage of all roads remained almost unchanged. However, the expansion of surfaced roads preceded the expansion of motor vehicles. In 1905 eight percent of all roads were surfaced, but less than 80 thousand motor vehicles were in use compared to about 3.3 million road horses and mules (in addition to the 22 million draft animals used for farming). Thus, the early surfaced roads were developed for horses and not automobiles, but motor vehicles quickly expanded into the growing infrastructure.

Figure 2.9 shows the substitution of unsurfaced by surfaced roads. In 1910 about 10 percent of all roads were surfaced, during the 1940s about one half, and today about 90 percent are surfaced, so that the substitution process lasted (i.e. Δt is) about 73 years. Therefore, we can conclude that the introduction of surfaced roads preceded the introduction of motor vehicles in the United States, but that the first rapid-growth phase of motor vehicle fleets occurred while less than one half of American roads were suitable for their use. It is also interesting to note that the substitution process does not reflect the vigorous road construction effort after

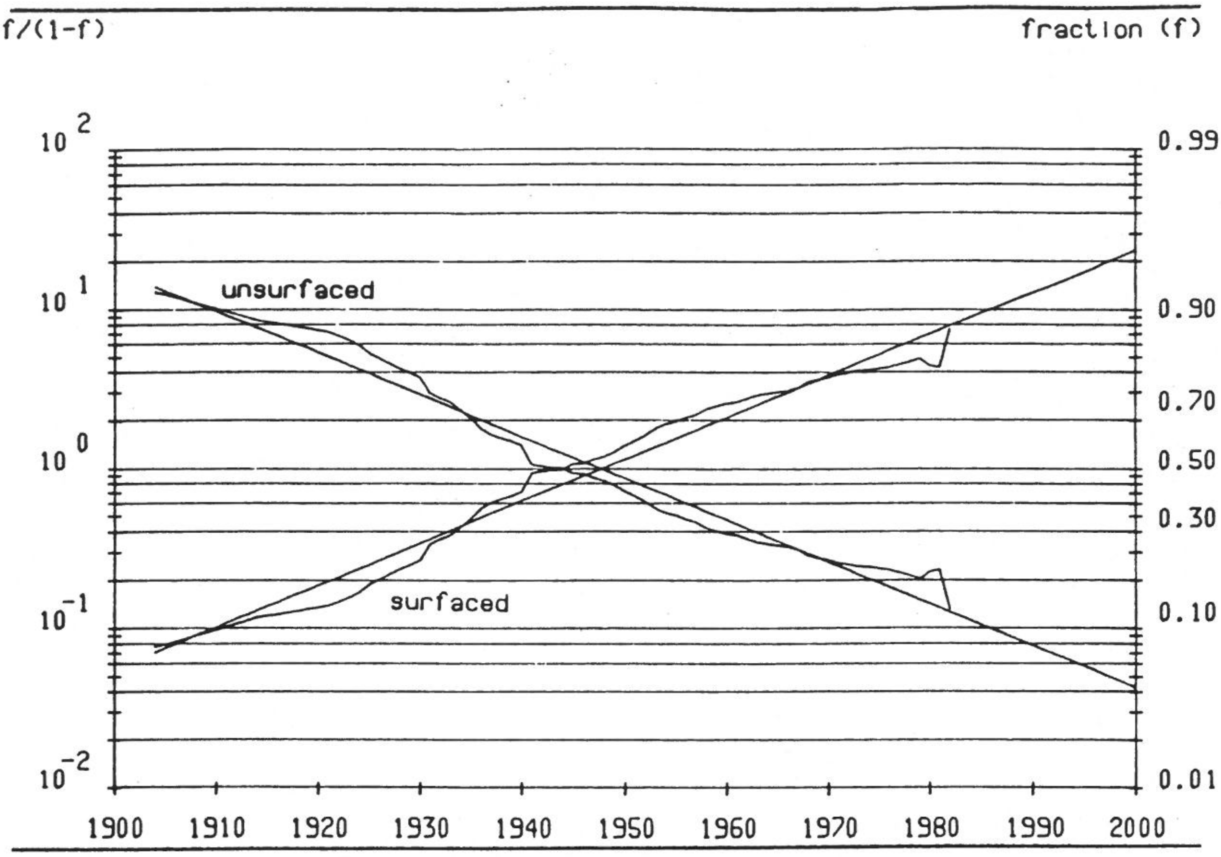

Figure 2.9 Substitution of Unsurfaced by Surfaced Roads, US.

the depression years in the United States, but rather indicates a lack of such effort during the 1910s and 1920s because the actual expansion of surfaced mileage is somewhat below the long-term trend during these two decades. A similar underexpansion occurred during the early 1970s, but appears to have been reabsorbed during the last few years. This is a reassuring result since it confirms the observation that adequate infrastructure is a prerequisite for the expansion of transport systems. Railroads and airports had to be constructed for trains and aircraft and, in the same way, surfaced roads were required for the widespread use of motor vehicles. Figure 2.10 shows the increase of surfaced road mileage as a logistic growth process with a saturation level of about 3.5 million miles (almost reached with 3.4 million miles today) that parallels the substitution process of unsurfaced by surfaced roads from Figure 2.9. The dashed logistic function (below

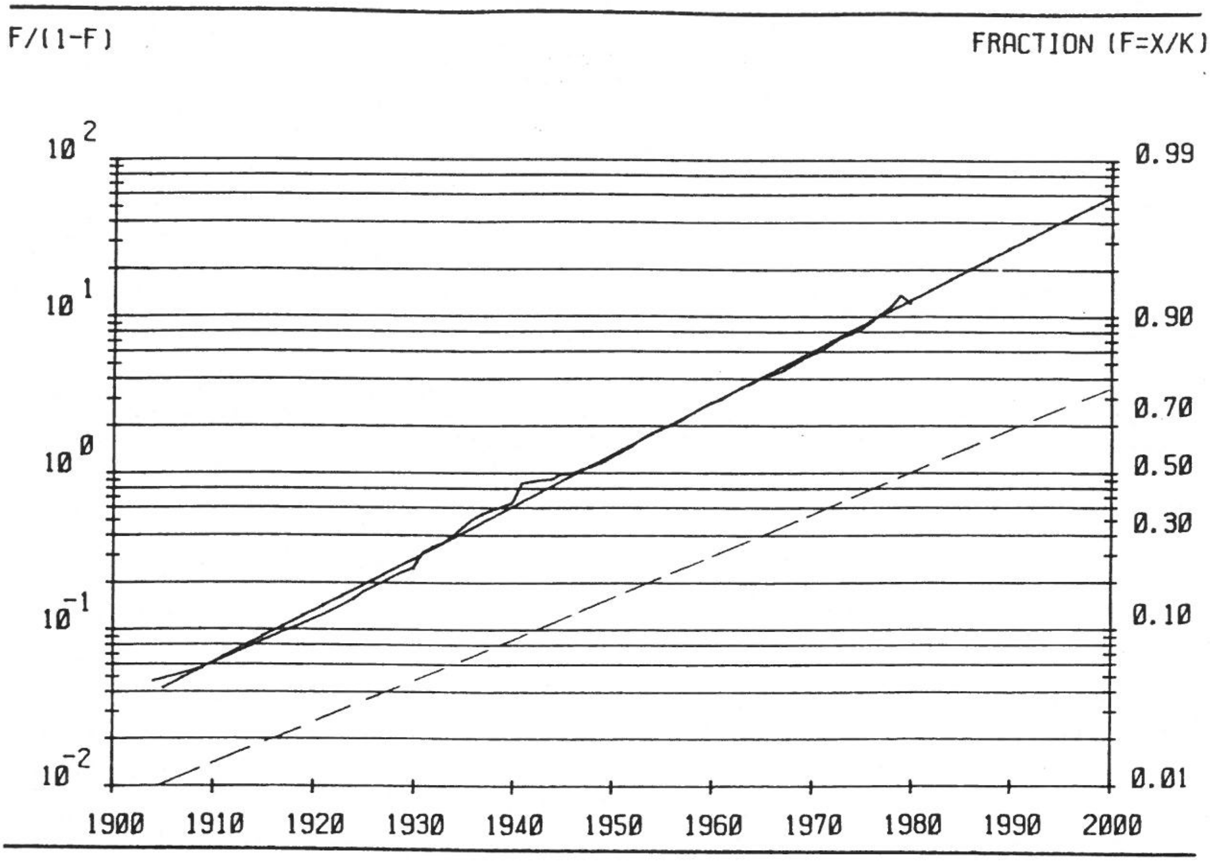

Figure 2.10 Mileage of Surfaced Roads, US.

the line for the growth of surfaced roads) represents the growth of *all* road vehicles from Figure 2.7. Thus, surfaced roads, as an important infrastructure for road transport, appear to be in saturation today, while the automobile fleet is still growing and should reach its estimated saturation phase in about 50 years.

2.3 Railroads

We have seen that air transport and the automobile are still expanding modes of passenger travel. Railroads, on the other hand, are now in the post-saturation development phase. In terms of intercity passenger traffic in the United States, their position *vis-a-vie* aircraft and the automobile is being eroded. It is symbolic of this decay process that the transcontinental railway service has been discontinued in the U.S.

While widespread air travel is only about 50 years old and the first automobile perhaps about 100 years, the first railroads of significant length were introduced 50 years earlier, or more than 150 years ago. Although railroads were also developed largely in Europe, like the automobile they had their most dramatic growth in the United States. By 1840, or more than ten years after the first commercial lines went into service, the United States had almost 4,500 km of railroad compared to Europe's 3,000 km (see Taylor, 1962).

The Baltimore and Ohio Railroad, probably the first commercial railroad in the United States, was chartered in 1829; two years later it had 13 miles of track in operation. Many projects soon followed. In 1833 the Charleston and Hamburg Railroad along the Savannah River route was the longest railroad in the world under single management with 126 miles of track. By 1835, three Boston railroads were in operation, one to Lowell, one to Worcester and the third to Providence.

To an extent, the early railway lines were feeder lines for canal and waterway transport systems in much the same way as the early commercial motor vehicles were for railways. Tramways (early dedicated tracks with animal-drawn or steam trains) in the Pennsylvania coal fields augmented canal traffic, and some of the first railroads connecting to the Erie Canal originally served as branch lines for that canal. In fact the early railroads were not allowed to compete with the Erie Canal. But even these earliest railroads were largely independent transport agents and proved to be sturdy rivals for the older canals. Thus railways were both feeder lines for canals where they could not compete with the canals, but also provided alternative traffic routes, and consequently direct competition, to canals and turnpikes. For example, the Boston and Worcester Railway was a competitor of the Blackstone Canal, and when completed, the Providence and Worcester Railroad put this canal out of business. The Baltimore and Ohio Railroad took business away from the Chesapeake and Ohio Canal, and the Charleston and Hamburg Railroad was designed to, and did to some extent, divert traffic from the lower Savannah River. This is again analogous to the competition between motor vehicles and railroads, and the eventual replacement of trains by motor vehicles in some market segments, and the symbiotic development of early motor vehicles as a feeder system to the railroad network.

Starting in the 1830s, the expansion of the railroads in the United States was very rapid for almost 100 years. This growth process is illustrated in Figure 2.11, which shows the increase in operated mileage of first and other main tracks in the United States since 1835. The mileage increased as a single logistic pulse reaching saturation in 1929 with a Δt of 54 years.

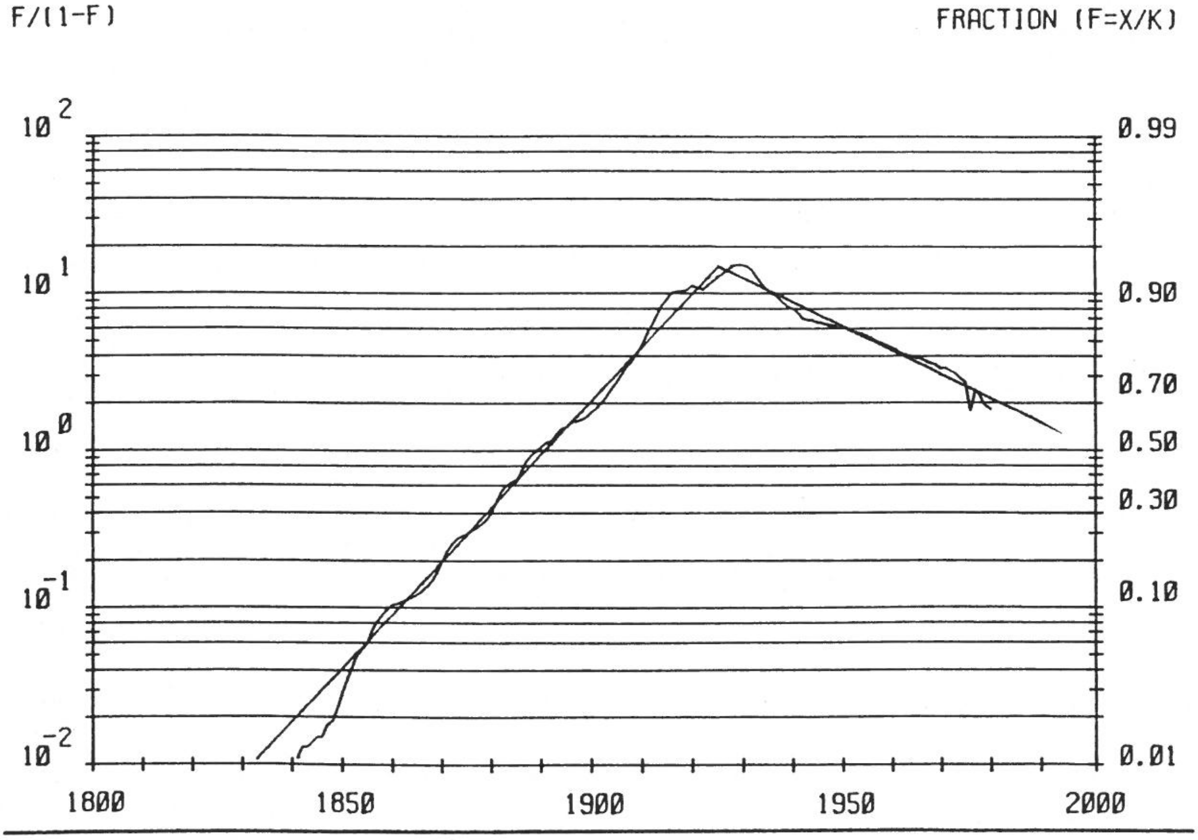

Figure 2.11 Growth Pulse in Main Rail Track Operated, US.

It declined logistically thereafter with a Δt of about 118 years. The increase in terms of the actual length of operated main track infrastructure, was from about 4,500 km (less than 3,000 miles) in 1840 to over 480,000 km (about 300,000 miles) in 1929; about two orders of magnitude in 90 years. In comparison, the decline was a rather slow process: by the 1980s the length of main track actually in operation decreased by about one third to some 320,000 km (about 200,000 miles). Thus, the phase of decline (about 50 years up to the present) appears to be slower than the growth phase probably because older technologies and infrastructures can enter new niches after they saturate and by doing so may avoid complete displacement by new competitors. For example, sailing ships, wood fire and horse riding have all become favorite leisure time and sporting activities although in their original markets they were replaced by new technologies long ago. Thus, in this sense the railways may enter a new era of use although they have virtually lost any significance as a mode of intercity *passenger travel*. Local passenger and tourist traffic (perhaps similar to ocean cruises) may be alternative uses for this large infrastructure in the future. Some canals are already serving such a purpose although they have not been profitable as a means of passenger transport since the hay days of railroads. But both canals and railroads still offer *efficient transport* for low-value goods (per unit weight and/or volume).

Many technological improvements were introduced in the United States during the first decades of the railroad expansion. Most of the early American railroads adopted wooden rails often 20 to 25 feet long capped with an iron strap or bar. In 1831 Robert I. Stevens ordered iron T-rails from England and installed them on the New Jersey Railroad (see Taylor, 1962), but it was not until the 1860s that T-rails became a prominent design feature on American railroads. In 1847 they became mandatory in the state of New York, although wooden rails were still in use on the western routes then under construction. Iron rails were a significant improvement since they made the use of heavier locomotives and loads possible and thereby increased the efficiency of rail transport. Later, the iron rails were substituted by steel. Thus, the basic construction materials in railroad infrastructure were shifted from wood to larger shares of iron and later steel. At the same time, the energy sources also changed from the draft animals that were used in the first tramways to steam locomotives fired by wood, which then remained the principal fuel used by railroads until about 1870 (Schurr and Netschert, 1960). Wood was then replaced by coal and later steam locomotives in general were replaced by diesel ones. Figure 2.12 shows the substitution of steam by diesel locomotives.

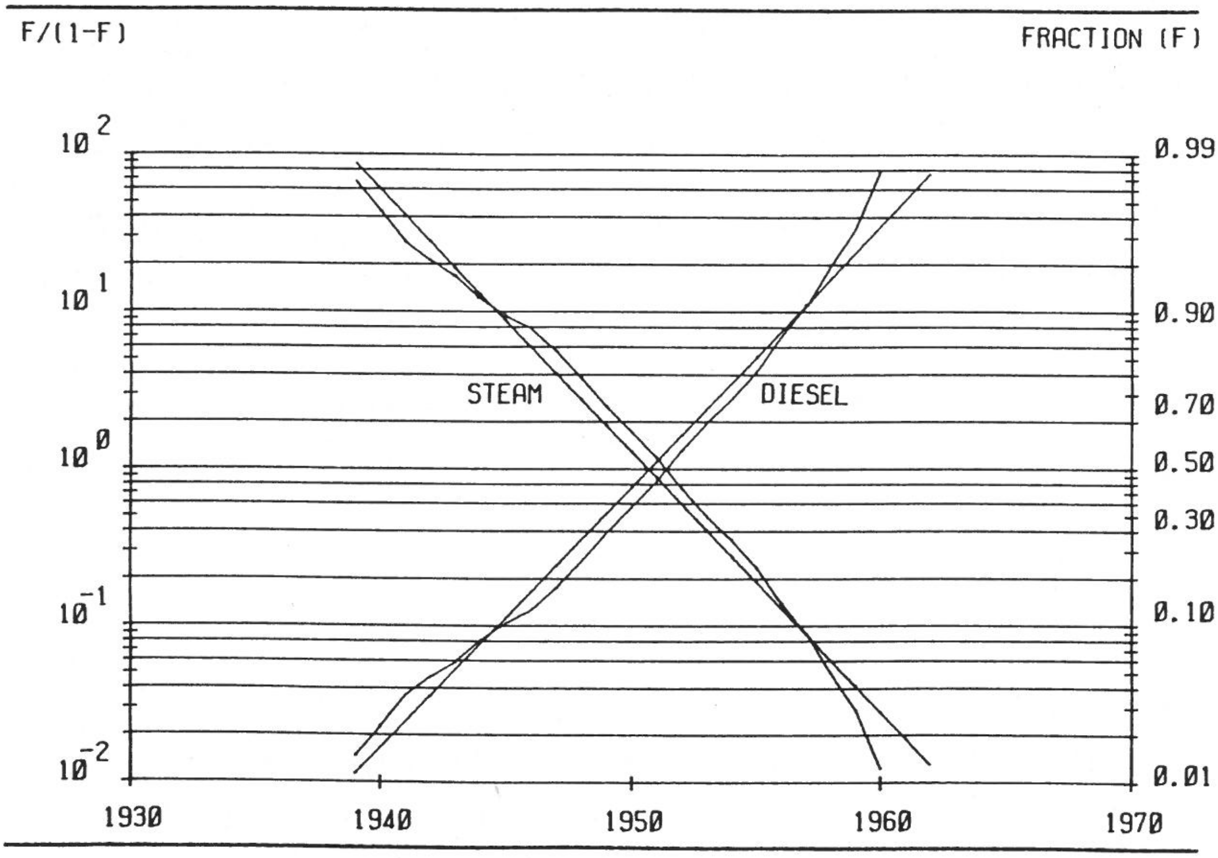

Figure 2.12 Substitution of Steam by Diesel Locomotives, US.

The first diesel locomotive was introduced in 1925 during the time when the total number of locomotives peaked at about 69,000 and a few years before the length of main track in operation reached its maximum and started to decline. Some electric locomotives were introduced earlier but they never gained any importance in the United States and always stayed below a two percent share in the total locomotive fleet. The replacement of steam by diesel locomotives was a swift process that lasted slightly longer than 20 years (a little shorter than the time taken for the substitution of horses by automobiles that lasted less than 30 years). In 1938 diesel locomotives reached a one-precent share in total locomotives and by the 1960s more than a 99 percent share. Thus, while the substitution of steam by diesel locomotives was a fast technological change in the composition of the fleet, the continuous increase in performance of the railroad transport system is characterized with longer secular trends. Since the first designs at the beginning of the last century, the traction of locomotives increased through improvements in the efficiency of energy conversion and subsequent increase of horsepower, and through increased weight (only possible because of the improvements in tracks and materials). In terms of total installed horsepower of all locomotives, the absolute peak was achieved in 1929 with more than 100 million horsepower of about 60,000 (mostly steam) locomotives. This peak coincides with the maximal length of main tracks also achieved in 1929. Although the total installed horsepower decreased after 1929, the average tractive effort of all locomotives continued to increase through the 1970s because the total number of locomotives also declined since 1925.

A number of other indicators of the railroad system in the United States show that the 1920s represented the culmination of railways. While the capacity continued to increase in an analogous way to the tractive effort of locomotives, the number of passenger-train cars in service peaked in 1924 at more than 57,000 and freight-train cars in service peaked in 1925 at more than 2.4 million. The number of passengers carried and passenger-miles also peaked during the same decade (passengers at over 1.2 billion and passenger-miles at more than 47 billion both in 1920). All of these indicators declined subsequently by somewhere between 30 and

40 percent of the maximal values reached during the 1920s.

A certain parallel in the development of railroads and automobiles can be detected. In a more superficial way both systems imitated the design of the transport modes that they were competing with and eventually substituted. First motor vehicles were literally ***horseless carriages***: they were almost identical except for the difference in the prime mover. Analogously, early railroad passenger cars were similar to stagecoaches both in design and appearance. Nevertheless, both of the new transportation technologies represented radical improvements and changes with respect to their predecessors. Schumpeter illustrated this point by saying that you can "add as many mail-coaches as you please, you will never set a railroad by doing so" (Schumpeter, 1935).

More fundamentally, the expansion of main track was only slightly faster than the expansion of surfaced roads in the United States. Both growth processes are characterized with a Δt of roughly 50 years (i.e. railroads of about 54 and roads about 58), but are separated in time by about 50 years. The inflection points of the two growth pulses are separated by 56 years; the main tracks reached half the saturation level in 1890 and surfaced roads in 1948. Thus, the growth of these two infrastructures is characterized by a "time constant" of about 50 years, while the decline of railroad infrastructure appears to be a much slower process.

Another fundamental similarity in the evolution of the two transportation systems is that important changes in road transport occurred during the years when most performance indicators of railways were saturating. Railways saturated during the 1920s and during the same decade the substitution of horses by automobiles was completed. Other fundamental innovations were also introduced in the rapidly expanding automotive industry during this period such as mass production and closed car bodies. Thus, while the railroads were saturating, the automobile industry consolidated and introduced important changes that generated subsequent growth.

It is useful therefore to distinguish two different aspects in the evolution of the two transport infrastructures. While the growth of main tracks and surfaced roads lasted on the order of 50 years, the technological changes and substitution of old by new equipment is a much faster process of about 10 to 20 years at the level of locomotive or automotive fleets. These processes continue even after the saturation phase is reached as the substitution of steam by diesel locomotives indicated. Thus, technological changes are important in both growing and declining industries. In growing industries new technologies are introduced through new additions and replacement, while in the declining industries new technologies are introduced exclusively through partial replacement of old technologies.

2.4 Canals

It appears that the evolution of rail and road transport systems portrays similar features and time constants (Δt) in their performance improvement, technological substitution, and in the increase in the size of their supporting infrastructures. The important events in the evolution of the two systems, however, are displaced in time by about 50 years. Railroads saturated during the 1920s when some of the most important technological changes and improvements in road transport were initiated. Thus, this similarity is perhaps indicative of an invariance in the development pattern of transport systems and their underlying infrastructures.

A serious problem arises, however, when comparing these two transportation modes with those that do not depend exclusively on the rigid, man-made links between them. Airways and waterways are examples of transport systems that rely less on the man-made links between the nodes because they use the natural environment (i.e. air, rivers and coastal waters) but nevertheless require an elaborate infrastructure such as airports, harbors and canals. Consequently, it is difficult to assess the total length of the implicit air and waterway routes that would be equivalent to the length of main railroad tracks or surfaced roads. In both cases, however, there are abstract concepts that would, in principle, correspond to the length of the grid: the network of certified route carriers or federal airways in air transport, and the total length of continental waterways and canals. Unfortunately, the annual increase in actual operated mileage length of all air carrier routes and the mileage of used continental waterways and canals is not very accurately documented in historical records. Thus, here we can rely only on sparse accounts and probably inaccurate estimates.

The first decades of the 19th century in the United States marked the beginning of large roads, canals, tramways and later railway construction projects. The years from about 1800 to 1830 have been called the "turnpike era" because during this period a number of the roads designed for travel between the larger towns or to the west across the mountains were completed. Turnpikes, however, were already being abandoned during the 1820s because of a lack of financial success (see Taylor, 1962). During the same period attention was given to canal construction in an attempt to develop a more effective means of internal transportation to complement coastal merchant transport while the construction of turnpikes declined. The "canal era" lasted until the railways became the main mode of long-distance transport a few decades later. From this point of view, the 1830s were very turbulent years: many turnpikes were abandoned, canal construction was reaching its peak, and some early railway projects were already completed. Thus, since the 1830s there have been at least three different transport modes that to some extent provided complementary services but were also competing directly in many market segments. In contrast to turnpikes and railways, canals connected the various natural links of the inland waterway system and did not represent a transport infrastructure in themselves. Instead, they are rather comparable to bridges and tunnels as the connecting links of road or rail transport networks. Thus, canals made the inland waterways into an integrated transport system by connecting the lakes in the north with the rivers in the south-east and mid-west. However, the great waterway across the United States never materialized.

The first canals were built during the 1780s and the Richmond Falls was the first canal to exceed a length of 7 miles when completed in 1793. From then on the pace of canal construction accelerated: by 1800 the total length of canals exceeded 50 miles and by 1825 more than 1,000 miles were in operation. Rapid construction continued for another 20 years or so, reaching 3.600 miles in 1850 and leveling off at about 4,000 miles 20 years later. Thereafter the canal business was so seriously eroded by competition from the railroads that many important canals were decommissioned. Subsequently, the length of all canals in use proceeded to decline. In the United States, the rise and fall of operated canals parallels the growth and decline of the main railway tracks but is displaced in time by about 50 years. In both cases the growth in mileage of the operated infrastructure increased rapidly, but after saturation the decline was less rapid by far.

Figure 2.13 shows the increase in the length of canals in the United States as a logistic growth pulse with a saturation level of about 4,000 miles after the 1860s. The inflection point, or the maximum rate of growth, was reached in 1832 and the Δt is about 30 years. Thus, canals have a time constant comparable to that of the airways but shorter than railways and roads. A possible explanation for this difference is that both canals and airways represent only one important component

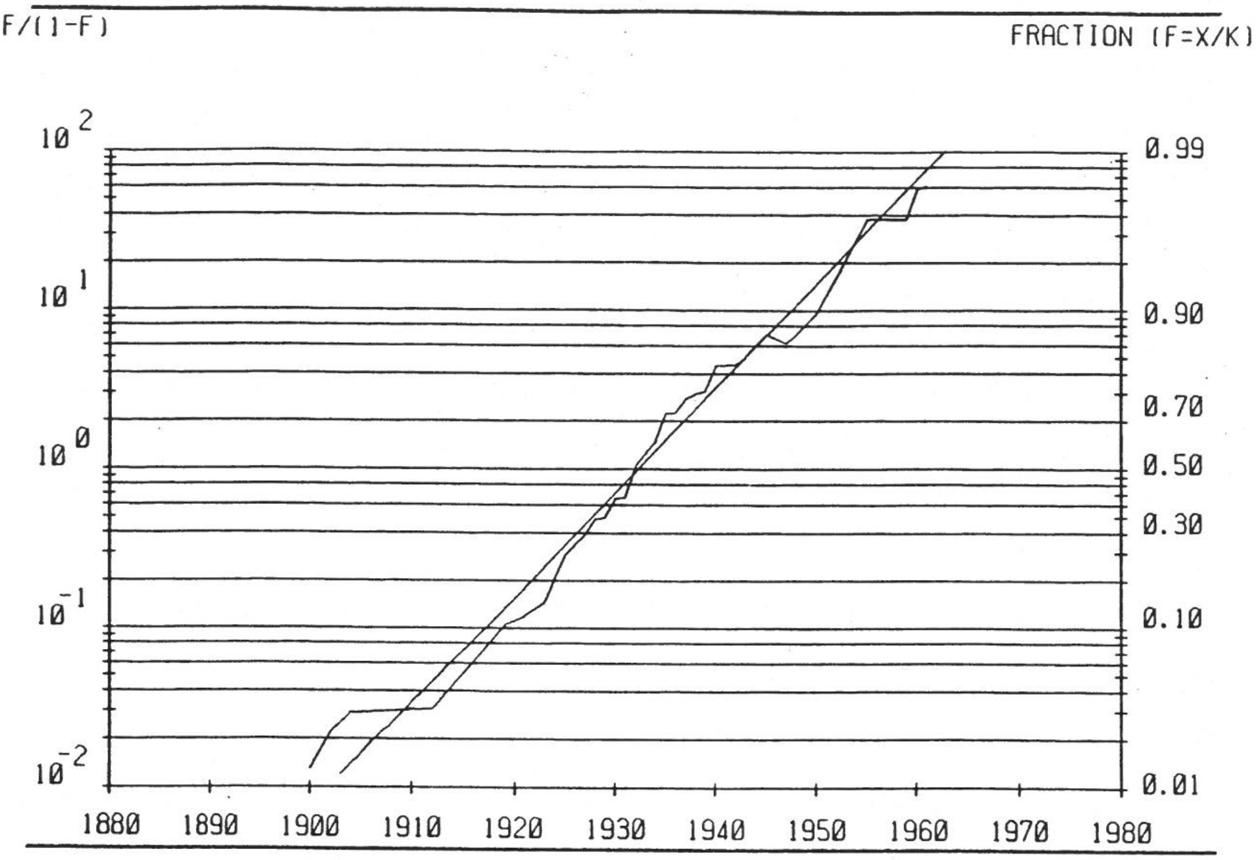

Figure 2.13 Length of Canals, US.

of the actual transport infrastructures, which for the former includes all waterways in addition to some man-made canals, and for the latter a large infrastructure of supporting systems, such as the airports and traffic control. On the other hand, railroads and surfaced roads in themselves constitute a large connected infrastructure that is essentially all built for a specific purpose, whereas to a large extent the water and airways utilize the natural environment with additional supporting infrastructures. Thus, this result may simply illustrate that it takes longer to construct large infrastructures such as roads and tracks compared with constructing selected links (canals or more abstractly air corridors) and nodes (airports and harbors) in order to expand a new transport system such as water and airways into an already available natural environment.

2.5 Transport Infrastructures

The difference in the time constants between air and inland water transport systems on the one hand and rail and road transport on the other, indicate that at least at this level of comparison those transport systems with more extensive infrastructural requirements may take longer to expand and, possibly, to complete the whole life cycle from growth to saturation and later senescence. In order to assess whether the time constants are really different, Figure 2.14 shows the substitution of different transport systems during the last 180 years in the United States in terms of the operated mileage of the competing infrastructures – canals, main railway tracks, surfaced roads and federal airway route miles. Thus, the growth of the transport networks is seen here as competition of the four transport modes in terms of the total length of operated mileage. Seen from this perspective the substitution of the four systems over time appears as a regular process.[4] A number of invariant features are inherent in this substitution process. At any give time, there are at least three important transport infrastructures competing in terms of shares of the total route length of all transport modes. The longest transport infrastructure is always longer than at least half the total length.

---

4 The fractional shares (*f*) are not plotted directly but as the linear transportation of the logistic curve, i.e. $f/(1-f)$ – in this more general case, as the ratio of the market share taken by a given energy source over the sum of the market shares of all other competing transport infrastructures. This form of presentation reveals the logistic substitution path as an almost linear secular trend with small annual perturbations. Thus, the presence of some linear trends in Figure 2.14 indicates where the fractional substitution of transport infrastructures follows a logistic curve.
In dealing with more than two competing technologies, we must generalize the Fisher and Pry model, since in such cases logistic substitution cannot be preserved in all phases of the substitution process. Every competitor undergoes three distinct substitution phases: growth, saturation and decline. This is illustrated by the substitution path of rail tracks, which curves through a maximum from increasing to declining market shares (see Figure 2.14). In the model of the substitution process, we assume that only one competitor is in the saturation phase at any given time, that declining technologies fade away steadily at logistic rates and that new competitors enter the market and grow at logistic rates. As a result, the saturating technology is left with the residual market shares (i.e., the difference between 1 and the sum of fractional market shares of all other competitors) and is forced to follow a nonlogistic path that joins its period of growth to its subsequent period of decline. After the current, saturating competitor has reached a logistic rate of decline, the next oldest competitor enters its saturation phase and the process is repeated until all but the most recent competitor are in decline. A more comprehensive description of the model and assumptions is given in Nakicenovic (1979).

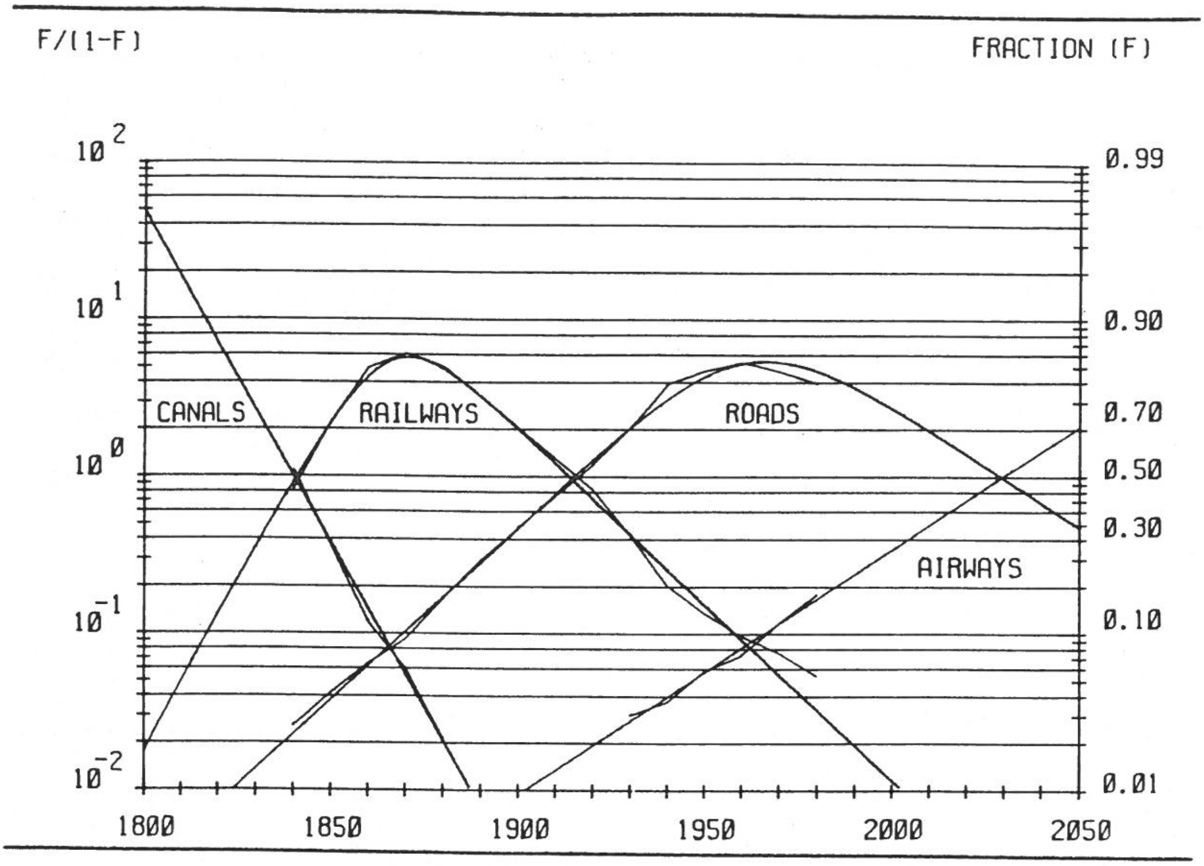

Figure 2.14 Substitution of Transport Infrastructures, US.

Consequently, the second and third longest transport infrastructures sum to less than half of the total length. This symmetry and dominating role of the longest infrastructure has been complemented by another regular feature during the last 180 years. The time constant (Δt) increases from less than 50 years for the growth of railway track shares to almost 90 years for surfaced roads and more than 130 years for airway routes as shares of total length. The time constant increased by about 40 years for successive infrastructures. The distance among the saturation periods (time of maximal market shares) of rail tracks and surfaced roads is about 100 years.

This result may appear to contradict our earlier observation that the total length of rail tracks and surfaced roads took longer to construct than water and airways. In fact, the substitution dynamics of infrastructure lengths offers a surprisingly consistent time table compared with the duration of growth pulses of the four transport modes during the past 180 years. The apparent inconsistency is due to the different ways of measuring the growth rates and life cycles of the respective infrastructures. In the case of ***market shares*** we are analyzing the increase of a particular transport infrastructure in terms of the length of all networks together. Thus, even the very rapid growth rate of airway route mileage is translated into a comparatively long time constant because at the same time the total length of all transport networks is also growing rapidly. Due to these rapid growth rates, the share of surfaced roads has been declining since the 1970s whereas the total length of surfaced roads is still growing toward the saturation level (see Figure 2.10).

Thus, the total length of transport infrastructure can still be growing and even be decades away from ultimate saturation and final senescence while the shares of its length in the length of all transport infrastructures are declining. This was the case with canals, railroads and surfaced roads. Thus, the saturation and decline of market shares precedes saturation in absolute growth in an increasing market. This is so by definition, but it means that the eventual saturation of any competing technology can be anticipated in terms of the substitution dynamics in a growing market. The logistic growth phase of rail tracks spans the period from 1840 to say 1916 (see Figure 2.11) with saturation in the 1920s and senescence thereafter. The dominance of rail tracks in terms of market shares also lasted form 1840 to 1915. In this particular case this means that the beginning of the growth pulse coincides with the beginning of market dominance and that saturation immediately follows the loss of market dominance. The saturation in market shares, however, occurred in the 1870s, indicating that saturation in the growth pulse would occur in the future, although at that time it was more than four decades away.

Although the substitution of older by younger transportation modes is a regular process when measured in terms of total mileage of the four transport infrastructures, it is nevertheless only a proxy for the real dynamics of transportation systems, which should be measured in some common performance unit. However, since transport systems provide a whole range of services, it is difficult to define such a common descriptor. Two obvious choices would be to use ton or person kilometers per year as a common unit. Nevertheless, this is still inappropriate since such a common unit does not distinguish between freight and passenger transport or between short and long-distances. Thus, there is no obvious short-cut and it appears necessary to analyze both passenger and freight separately for both short and long distances (e.g. intra and intercity travel). Fortunately, historical data are available that make it possible to reconstruct the dynamics of these substitution processes for at least some countries. For example, Figure 2.15 shows the substitution of different transportation modes in intercity (long-distance) passenger travel in the United States since 1950.

By excluding urban and metropolitan transport because explicit statistics are generally lacking, the competition for intercity passenger traffic is reduced to four significant transport modes. Figure 2.15 shows the substitution of railways, busses, cars and airways as competing transportation systems for long-distance travel. The shares of each mode of transport are calculated in terms of passenger-miles traveled in a given year. According to the market substitution analysis in Figure 2.15, the automobile is still the main choice for the majority of Americans and will remain dominantly so well into the next century in spite of the market share increases gained by air travel. The losers are the railways and busses. The substitution dynamics indicate that during the next two decades this trend will expand the share of airways in intercity travel to almost 40 percent, reduce the share of automobiles to little more than 60 percent, and virtually eliminate railways

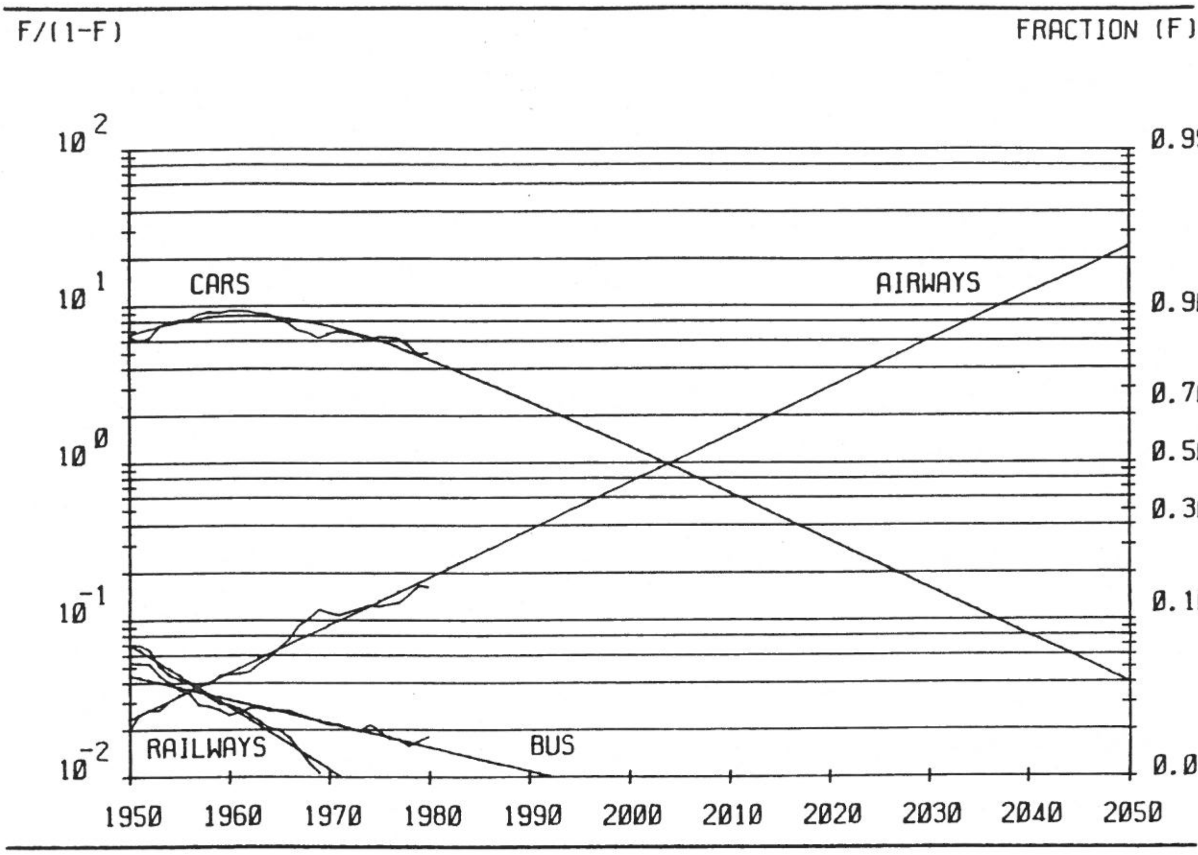

Figure 2.15 Intercity Passenger Traffic Substitution, US.

and busses. Thus, automobiles will remain the most important, although declining, form of long-distance travel throughout this century in the United States, while aircraft will continue their rapid expansion by increasing the market shares they currently control.

Unfortunately, Figure 2.15 does not indicate the full drama of this replacement process because data are not available for the period before 1950. Nevertheless, the substitution process is consistent with the mileage substitution of the respective transportation infrastructures from Figure 2.14. For example, the share of main railway tracks decreases down to less than a one-percent share by the end of the century, and the share of railways in intercity travel phases-out even earlier during the 1970s. During the same decade the shares of both automobiles in intercity travel and of surfaced road mileage are saturating. In both cases the shares of airways are increasing and are expected to reach a position of dominance (with control of more than half the "market") during the first decades of the next century. This result suggests that a new growth pulse in commercial aviation may be due in about two decades. In conjunction with our analysis of the future trends in the aircraft productivity (see), we conclude that cruse supersonic and possibly also air-breathing hypersonic transports could be the only technology that is likely to provide the growth potential to be achieved in the next expansion phase of passenger travel.

All of these examples have indicated that the structural changes in transport systems and their respective infrastructures take place over long time periods. There is also strong evidence that some features of these changes are rather invariant in time. Each of the transport systems developed (with logistic improvement of performance and increases in the length of their respective networks), saturated, and then proceeded into a period of senescence. Exactly the same pattern is repeated in the secular changes of market shares as these systems compete with each other.

The above is obviously only a partial description of these complex systems and there are certainly many other, perhaps better, ways of describing the dynamics of these systems. The intriguing aspect of the logistic description of the development of transportation systems is that they appear to be interwoven with regular features and a pattern of substitution dynamics over a period of about two centuries. In order to indicate that this is not a singular feature unique to the evolution of transportation systems, we will now show that a similar regularity exists in the evolution of energy systems. Furthermore, the future developments of transport infrastructures are related to likely energy system changes, especially with respect to energy transport and propulsion for prime moves.

3 PRIMARY ENERGY

3.1 Energy Consumption

At the beginning of the 19th century, fuel wood, agricultural wastes and mechanical wind and water power supplied most of the inanimate energy in addition to animal and human muscle power. This, for present standards, poor energy menu, already represented a sophisticated energy system compared to earlier practices. We have seen that a considerable infrastructure of roads (turnpikes) and canals was in place in the 19th century for timber and later coal transport. Mining, manufacture and irrigation were also usually associated with elaborate systems of dams and water-wheels. This indicates that in the early industrial development phase energy use depended on the transport system, and that energy was one of the more important components of goods transported on canals, waterways, turnpikes and roads.

In primary energy terms, fuel wood represented most of the energy inputs. Figure 3.1 gives the annual consumption of fuel wood, fossil energy sources, direct uses of mechanical water power and hydroelectric power in the United States since 1800. Data are plotted on a logarithmic scale and show the exponential growth phases in consumption of the most important sources of primary energy during the last 180 years by piece-wise linear secular trends.

In the United States the consumption of fuel wood, once the most important source of energy, has declined since the beginning of the century, although its use is still widespread, especially in the developing parts of the world. With the expansion of railroads, the steel industry and the application of steam in general,

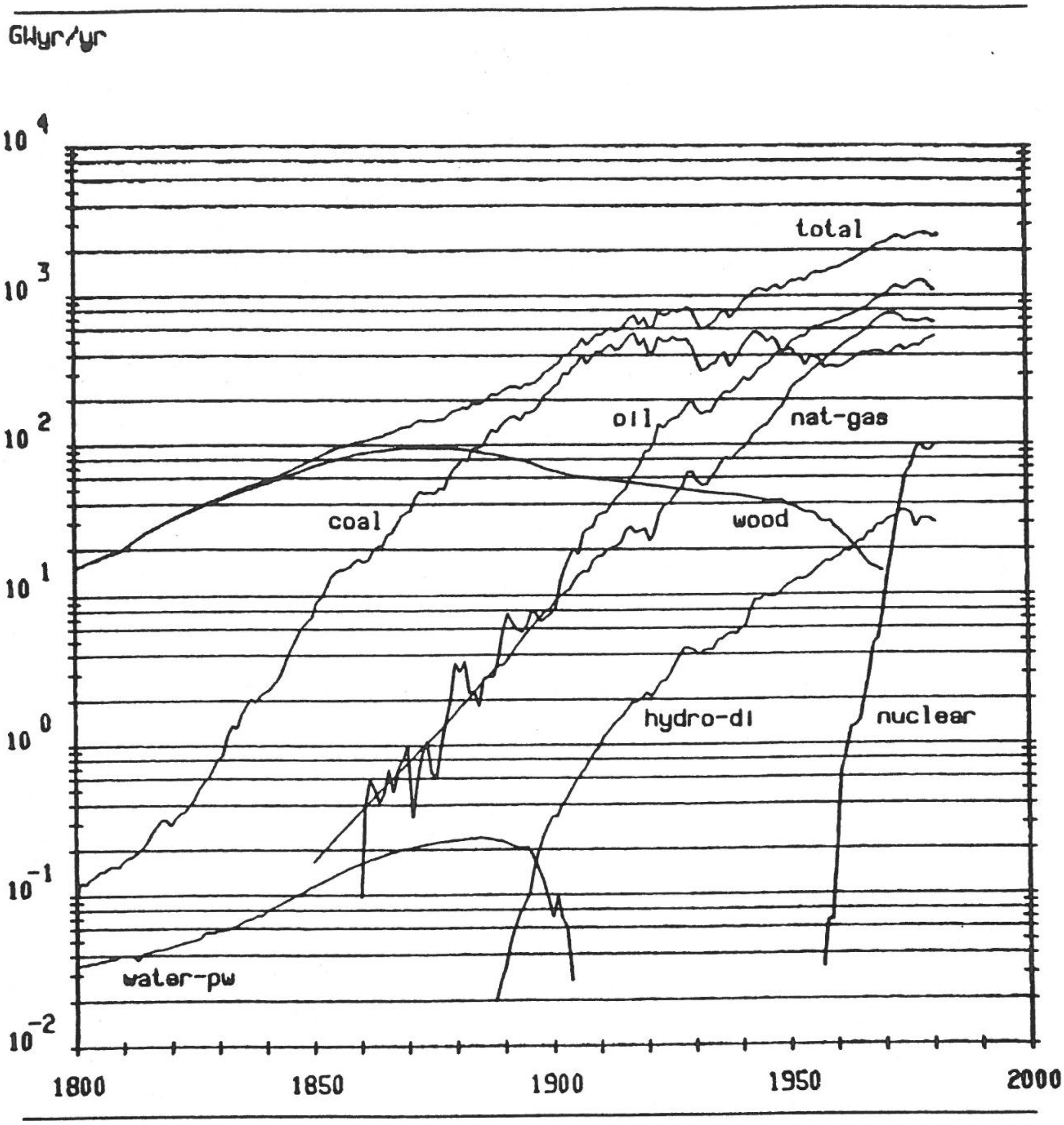

Figure 3.1 Primary Energy Consumption, U.S.

coal use increased exponentially until the 1910s and has oscillated ever since. Both oil and natural gas were introduced during the 1870s and their consumption has increased with even more rapid growth rates ever since. In fact, oil and natural gas curves have the same shape and almost identical growth rates; they are just shifted in time by about 10 to 15 years. The use of oil and natural gas grew in parallel with the petrochemical industry, electricity and electrical industry, internal combustion and electric prime movers. Nuclear energy is still in its early phase of development, therefore the steep growth prevailing over the last decade may not be indicative of its future role. During the last years, however, the growth of nuclear energy in the United States and worldwide has declined to more moderate rates.

3.2 Energy Substitution

Thus, during this period of almost two centuries, energy consumption did not draw equally from all sources, nor did the use of all energy sources increase equally. Yet, primary energy consumption (including fuel wood) increased exponentially at an average growth rate of about three percent per year. The decline of older energy sources was more than compensated for by the rapid growth of the new ones. Therefore, it is evident that the older forms of energy have been substituted by the newer ones. Figure 3.2 shows primary energy substitution in the United States. [5]

Mechanical water power (mostly hydro and some wind mills) and hydroelectric power are not observable in the figure due to their low contribution to total energy supply: they barely exceeded the one-percent level during very short periods and are otherwise under that critical level. Thus, before the 1820s fuel wood provided for virtually all the energy needs of the United States. Coal entered the competition process in 1817 at the one-percent level and up to the 1880s it was essentially a two technology market – whatever gains coal made were translated into losses for fuel wood.

5 As in Figure 2.14, the fractional shares (f) are not plotted directly but as the linear transformation of the logistic curve, i.e. $f/(1-f)$ – as the ratio of the market share taken by a given energy source over the sum of the market shares of all other competing energy sources. Also in this figure, this form of presentation reveals the logistic substitution path as an almost linear secular trend with small annual perturbations. Thus, the presence of some linear trends in Figure 3.2 indicates where the fractional substitution of energy sources follows a logistic curve.

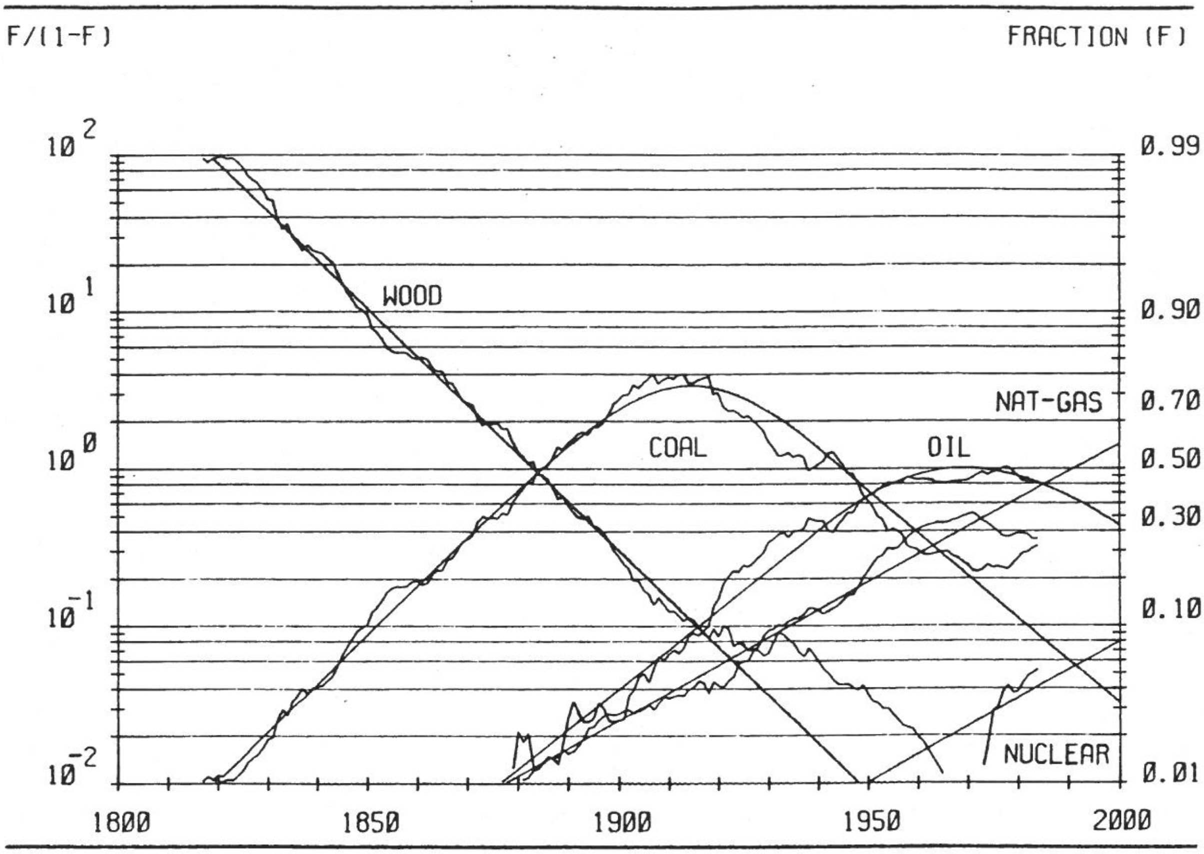

Figure 3.2 Primary Energy Substitution, US.

Nevertheless, wood remained an important construction material and source of heat and power well into the last decades of the 19th century so that the steam age began in an economy based on wood use. The first steam-boats and locomotives were fired with wood, which remained the principal fuel used by railroads until about 1870 (Schurr and Netschert, 1960). The iron industry was another large wood consumer. Around 1850, more than half of all the iron produced was still smelted with charcoal. In spite of these large uses of wood, the total amount consumed in manufacturing and transportation was small compared to the huge quantities used in households. By 1880, the industrial use of wood declined so that domestic use accounted for more than 96 percent of fuel wood consumed (Schurr and Netschert, 1960). At the same time, coal already supplied almost one-half of all energy needs, most of it being used by emerging industries. In 1880, coal supplied almost 90 percent of the fuel used for smelting iron. Thus iron was used in turn for great construction prospects including larger railroad infrastructures which by that time used coal as the fuel of choice. Thus, in this respect, the end of the last century marks the beginning of the industrial development period in the United States.

The first use of crude oil and natural gas in the United States dates back to the beginning of the 19th century and both reached the one-percent market share during the 1880s. From then on the use of crude oil expanded somewhat faster as time progressed and by 1950 crude oil consumption surpassed that of coal. The use of natural gas surpassed coal nine years later. It should be noted that even as late as the 1920s the use of crude oil was not much larger than the consumption of fuel wood. When compared with earlier periods, it is remarkable that the structure of energy consumption changed more during the period of oil dominance than ever before.

The 1950s, when oil became the dominant source of energy, represent the beginning of more intense competition between various energy sources both in the United States and in the world. Over the period of 150 years the energy source that dominated the energy supply at the same time also contributed more than one-half of all primary energy consumption – from 1800 to 1880 fuel wood and from 1880 to 1950 coal. This is similar to the dominance at any given time of one transport infrastructure over all other with more than half of all mileage that was observed in the substitution of canals, rail tracks, surfaced roads and airways (see Figure 2.14). During the 1970s, crude oil was close to achieving a 50-percent share, but before actually surpassing this mark proceeded to decline. Thus, during the last three decades three important sources of energy shared the market without a single source having a pronounced dominance, which is contrary to the pattern observed during earlier periods and the substitution of transport infrastructures.

Figure 3.2 shows natural gas as the dominating energy source after the 1980s, although crude oil still maintains about a 30 percent market share by the end of the century. The future potential competitors of natural gas, such as nuclear or solar energy, have not yet captured sufficient market shares to allow estimation of their future penetration rates. The starting point for market penetration of nuclear energy can be dated back to the 1960s when nuclear power acquired slightly less than a one-percent share in primary energy. Allowing for further cancellations of already ordered power plants and possible decommission of those in operation and construction, we have assumed that nuclear energy could at most double its current market share to about four percent by the year 2000 and have used this assumed penetration rate for the projection in Figure 3.2. This leaves natural gas with the lion's share in primary energy advancing its position to the dominant energy source after this century.

Thus, a prominent feature of the projection of primary energy substitution dynamics into the future is the emergence of natural gas as the dominant energy source during the next decades. According to Figure 3.2. more than half of all the primary energy consumed during the first decades of the next century will be natural gas. This result illustrates not only that the natural gas bubble will be absorbed in a few years, but also that methane technologies will develop in the future creating new growth sectors. One of the important used may in fact be aviation in the more distant future.

Although, this result is unexpected in terms of the numerous energy debates of the last 10 to 15 years, it is perhaps reassuring that we may not have to rely on nuclear or alternative energy sources for another 50 years or so after all. Instead, the possible future that emerges would require less radical changes, but still leave challenging tasks to develop new technologies and improve the performance of already employed technologies, such as deep drilling, pipelines, and methane conversion into other energy carriers (e.g., electricity and methanol). Despite the current difficulties in expanding the use of natural gas in a time of worldwide economic slowdown and low energy (i.e., crude oil) prices, our scenario paints a different picture, albeit in the long run. In order to gain a better understanding of how such changes may come about to allow methane to emerge as the dominant energy source of the future, we will briefly analyze the evolution of energy transport infrastructures.

3.3 Oil and Gas Transport

In contrast to oil that is traded globally and gas that is transported over continental distances, fuel wood was primarily consumed locally, close to the source of timber. Some fuel wood was transported over longer distances mostly by river flotation and distributed by waterways or roads. Coal on the other hand, represents a more concentrated form of energy than fuel wood and coal mines a more concentrated source than forests (since coal has a higher heat content per unit weight than wood) so that it was generally transported over longer distances than fuel wood. The extreme case is the modest overseas coal transport, but more usually coal was transported (nationwide) by barges, trains or trucks (earlier horse wagons). Thus, the shift from a wood to coal economy was accompanied by the expansion of energy transport over longer distances and to increasing number of transport modes. The widespread use of crude oil brought another transport mode in addition to tankers, trains and trucks – oil pipelines. Pipelines are becoming an important freight transport mode with market shares in total ton-kilometers per year comparable to those of train and truck transport. They are also comparable to railways in terms of the total length of the infrastructure or grid: the total length of main track in the United States is, about 200,000 miles, slightly shorter than crude oil pipelines with about 230,000 miles. It is interesting to note that in terms of ton-kilometers, train car loads and revenue, coal transport represents by far the largest commodity group in rail transport. Thus, although the total length of the rail and oil pipeline grids is equivalent, the big difference between the two infrastructures is that since the 1920s the railroad system has been declining while oil pipelines have been expanding. Figure 3.3 shows the rapid increase in the pipeline length for crude oil and petroleum products in the United States.

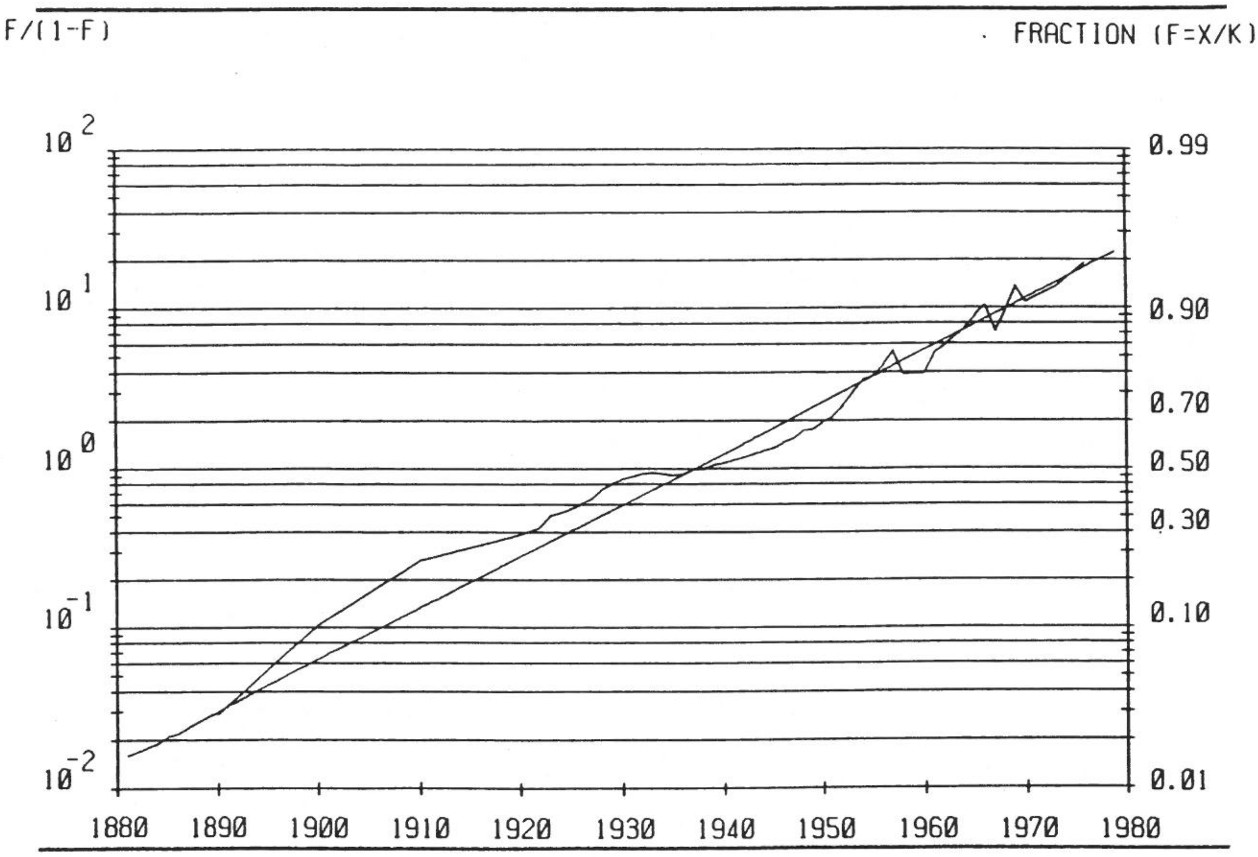

Figure 3.3 Crude and Products Oil Pipelines Length, U.S.

The expansion of the oil pipeline grid parallels the increase of crude oil shares in total primary energy consumption. Oil reached a one-percent share in energy during the 1880s and at the same time the rapid increase in oil pipeline mileage started and followed exponential trends until the 1930s (the inflection point occurred in 1937). As if by coincidence, oil's largest competitor, coal, reached maximal shares in primary energy during the same decade. Thus, by 1937 about half of the current length of the oil pipeline network was already in place and the growth rates declined slowly. During the 1980s the length should reach the asymptotic level about the same time as crude oil shares in total primary energy saturate.

The time constant (Δt) of the expansion of oil pipelines is 59 years or about the same as for the expansion of rail tracks and surfaced roads (about 54 and 59 years, respectively). However, in the analysis of the expansion of oil pipelines we encountered numerical instability in the values of the estimated parameters with respect to different estimation algorithms used in our analysis. In particular, the result reported in Figure 3.3 is based on an (estimated) saturation level, κ, of 238.8 thousand miles with the Δt of 59.2 years, whereas an alternative estimation algorithm (used to test the stability of estimated parameters with respect to changes in the assumptions about the weighing of the observations) gave different estimates, namely, 246.5 thousand miles and 78.7 years, respectively. In all other examples, the variation of the parameter values was within the estimated uncertainty ranges and in the as shown in the Appendix. Only in the case of oil pipelines the value for Δt was outside this range.

Although oil is still the most important energy source, in terms of primary energy consumption it is slowly being replaced by natural gas. In fact, the amount of associated natural gas is decreasing in total natural gas production and, therefore, higher shares of natural gas transport and end-use are based on gas and less on oil technologies (see, Nakicenovic and Grübler, 1987). This process is also reflected in the increase of natural gas transport and distribution pipelines when compared with oil pipelines from Figure 3.3. According to Schurr, et.al. (1960) the earliest recorded commercial use of natural gas in America dates back to 1821 (at the time when coal supplied just one-percent of primary energy and fuel wood and draft animals the rest), when it was used as lighting fuel in Fredonia, New York. Natural gas continued to be used sporadically throughout the 19th century. The first pipeline was constructed from Murrysville to Pittsburgh (Pennsylvania) in 1883 after the discovery of a large well in 1878.

Despite such pioneering projects by the emerging oil and gas industry, methane was in general considered a waste product. By 1878, both crude oil and natural gas surpassed a one-percent share in primary energy consumption but at that time most of the natural gas consumed was used in the vicinity of the oil fields.

The rapid expansion of the natural gas pipeline network started during the 1890s, or about 20 years after the growth of oil pipelines was initiated. Figure 3.4 shows that this 20 year shift in time persists through most of the growth cycle of the natural gas transport system and distribution infrastructure. The inflection point, with about half the eventual saturation level achieved, occurred in 1965 or almost 30 years after the inflection in the growth of oil pipelines. Since the time constant is about 60 years, and therefore shorter but still comparable to that of oil pipelines (59 years), saturation should also occur more than 20 years later during the 2020s. Again this is symmetrical to the relationship between the growth phases of oil pipelines and oil penetration in primary energy. The growth pulse started when oil achieved a one-precent share of primary energy, inflection occurred during the time when oil became the second largest energy source (bypassing fuel wood), and saturation of pipeline length was synchronous with the saturation of market shares. Exactly the same pattern can be observed during the growth pulse of natural gas pipelines by comparing Figure 3.4 and Figure 3.2. The growth started towards the end of the last century when natural gas achieved a one-percent share in primary energy. The inflection point was reached in 1965 when natural gas became the

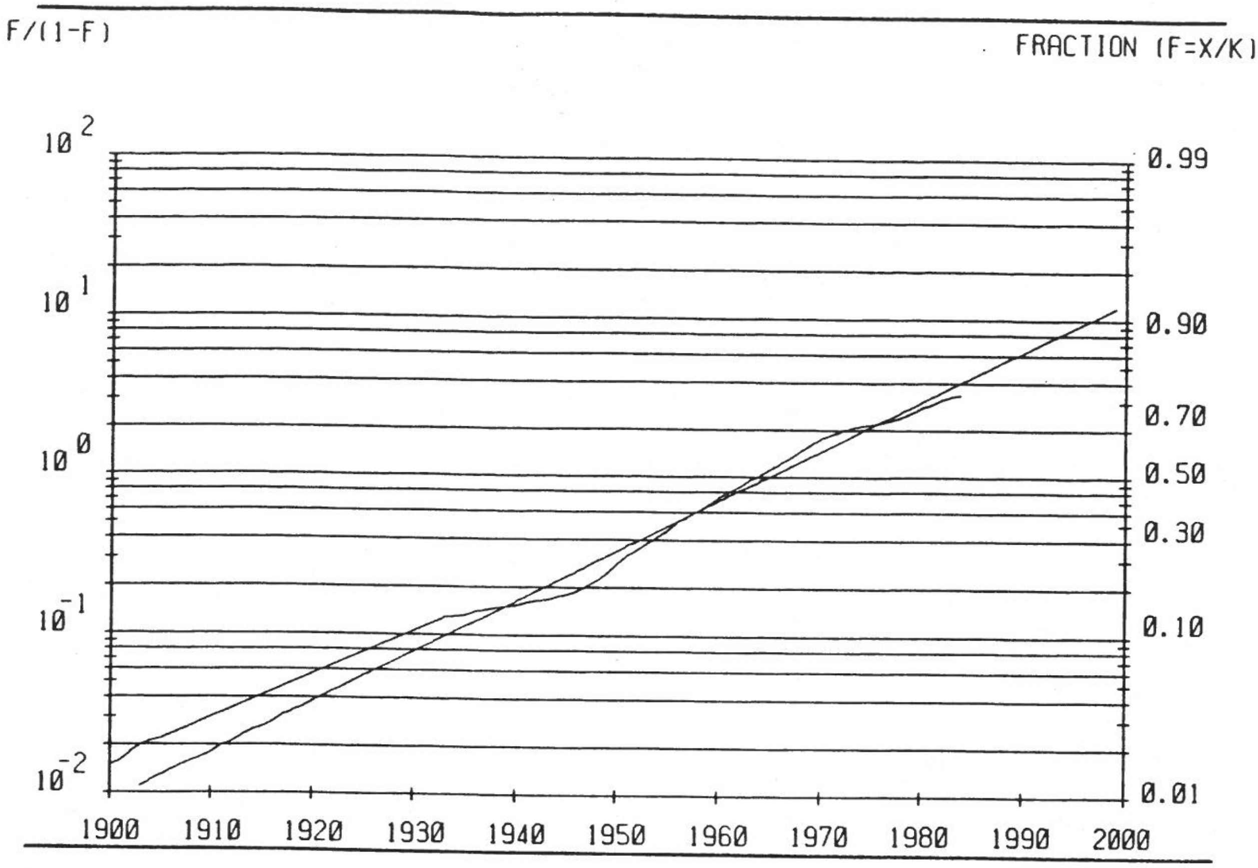

Figure 3.4 Natural Gas Pipeline Length, U.S.

second largest energy source (by passing coal), and saturation of both natural gas market shares and length of pipeline should be achieved during the 2020s.

A large difference between the growth pulses of oil and natural gas pipelines is in the length of the respective transport and distribution networks. Figure 3.3 gives a saturation level estimate for oil pipeline length of about 240,000 miles (or about the current length of railroad tracks), whereas the asymptotic level for the length of natural gas pipelines is estimated at more than 1,300,000 miles (more than five times higher). For the time being natural gas is transported almost exclusively through the pipeline grid. Oil and petroleum products however are also shipped by tankers, trains, trucks and for some military use even by aircraft. Besides some smaller quantities of liquefied natural gas and liquid natural gas products, most natural gas reaches the consumer either in a gaseous form or as electricity. The pipeline network for natural gas transport and distribution is therefore also much longer than that for crude oil and petroleum products. This of course poses the question of whether we can expect natural gas to continue to be almost exclusively transported by pipelines in the future, especially if its projected use expands as dramatically as illustrated in Figure 3.2. For liquid natural gas products especially, it is likely that other transport modes will also be used, conceivably even aircraft. From the technical point of view there are no principal obstacles to using this transport mode for energy, the only question is whether it would be economical and competitive to do so.

This cannot be resolved here, but we mention this alternative for the future because similar solutions were found in the past to meet the ever increasing need to transport more energy over longer distances. Both transport and energy systems must meet stringer economic and environmental requirements over longer distances and in shorter time. A faster means of transport, (i.e. over larger area) and denser and cleaner energy forms were necessary technological measures in the past to improve the performance of these systems. We can therefore expect further improvements in the near future and these could be fulfilled by a stronger reliance on natural gas and aircraft until better solutions and the development of completely new systems are devised during the next century. During the next four to five decades we will rely on improved versions of the current energy and transport systems and infrastructures, since our results indicate that the time constants involved in installing new infrastructures are in the order of 50 years.

4 CONCLUSIONS

In a number of examples analyzed in this paper, we have shown that the growth and senescence of transport infrastructures can be described as a regular process with logistic secular trends. We have also observed that the time constants (Δt) of these growth processes clustered around 50 years with a range of between 30 to 59 years. Considering that the quality of historical data concerning the growth of the transport infrastructures leaves much to be desired, it is remarkable that the four network or grid transport systems cluster even more densely between 54 and 59 years. This includes the time constants for railway tracks, surfaced roads, oil and natural gas pipelines. The time constants for canals and performance improvements in aircraft transport were shorter at about 30 years. In the case of canals and railway tracks senescence lasts much longer than the growth process. The total length of canals, and railways 50 years later, declined with a much longer time constant (of almost 150 years) so that canals are still in use despite a century long decline.

This perhaps explains the long time constants in the substitution of infrastructures. Although the *growth* in length of *all* network or grid transportation infrastructures was a very fast process, from about 4,000 miles of canals to about 4 million miles of railways, roads and airways (a growth of three orders of magnitude in about 150 years), the *substitution* of transport infrastructures (see Figure 2.14) lasted longer than the actual growth pulses for each of the four individual infrastructures. The time constant Δt of the substitution process increased from 50 years for railway shares to almost 90 years for surfaced roads and more than 130 years for airways, compared with a time constant of 30 years for the *growth pulse* of canals and about 50 years for railways and roads. This also illustrates the permanence of infrastructures and their sites.

The decline of infrastructures is a very slow process. New infrastructures substitute the older ones more through rapid growth of the whole transport system, rather than through ***physical*** destruction or replacement of older infrastructures.

Despite a remarkable regularity of these growth and substitution processes *in toto*, the variation of the parameter values with respect to different estimation algorithms was outside the uncertainty range in the case of the growth of oil pipelines. This case illustrates the fact that those growth processes that follow the simplest possible pattern, namely a single *S*-shaped path, can be difficult to measure even when they are almost complete because the saturation levels are *a priori* unknown and consequently have to be estimated from the data. Technological substitution processes are inherently more complex because market shares of all important (competing) technologies have to be measured. In part also due to this higher degree of complexity, they are more stable from the statistical point of view. The most sensitive parameter of the logistic function, the asymptotic saturation level κ, is *a priori* known in this case. By definition, it can never exceed 100 percent of a given market.

A more consistent analysis of the evolution of transport systems and their infrastructures would require a comparison of the performance and services provided by the different modes over long periods in addition to the analysis of the length of different transport grids. For the time being it was not possible to reconstruct such an indicator for the overall transport system, but Figure 2.15 shows a good proxy for such a comparison of intercity passenger travel during the last 37 years. The substitution of the four major modes of long-distance travel in the United States is a regular process that singles out aircraft as the transport mode with the highest growth potential. In terms of the dynamics and general implications, this result is consistent with the time constants observed for the growth pulses and substitution of the four transport modes – canals, railways, roads and airways.

The primary energy *substitution* dynamics (from Figure 3.2) indicates that technological substitution can be characterized by very regular time constants provided that the historical data is accurate enough to provide the information required for the analysis. This is to an extent possible in the case of primary energy consumption because different energy sources can be measured in common (physical, energy) units and because their use is relatively well documented. Furthermore, the technological changes in the energy system, i.e. the shift from older to newer energy sources, is also reflected in a parallel evolution of the energy transport infrastructures. Both fuel wood and coal were transported fundamentally by canals and railroads, in the same way as most other goods were. Oil and natural gas, however, are also transported by dedicated pipeline infrastructures. It was shown that the evolution of these energy transport infrastructures parallels the changes in the energy system, but that at the same time their growth patterns are no different from the expansion of other transport infrastructures.

A general conclusion is that technological changes, such as the substitution of road vehicles or locomotives, last a few decades, while changes in the transport system and evolution of infrastructures are longer processes with time constants of between three and seven decades. This means the transport infrastructures themselves will not change drastically during the rest of the century, but concurrent technological changes in vehicles, aircraft and necessary equipment will improve both the performance and quality of the transport services provided by the current systems.

More specifically, road vehicles will remain the dominant form of long-distance transport during the next 20 years, but the importance of airways should increase in time. Both transport and energy systems must meet stringer economic and environmental requirements over *longer distances and in a shorter time*. Both transportation and energy systems in the United States appear to be still in the development phase because we were not able to detect any significant signs of saturation, but rather we observed that older, still dominant technologies are slowly being displaced by newer competitors which in principle offer substantial growth potential.

Thus, our results of transport infrastructure evolution and energy substitution indicate that the next generation of aircraft and natural gas technologies represent evolutionary rather than revolutionary technologies that could meet these more stringent requirements. Both aircraft and natural gas technologies could meet these requirements through refinements and improvements in current designs and practices during next two decades, but new solutions must be developed during this period for the decades thereafter. In particular, the performance of transport aircraft is estimated to saturate (in the current growth pulse) at about 1.2 million passenger-kilometers per hour which is reachable with current aircraft technology and some improvements. However, thereafter (in a new, hypothetical growth pulse) aircraft would be required to exceed these levels implying either large subsonic liners (with perhaps a few thousand passengers), supersonic cruise or air-breathing hypersonic transport with a more "modest" capacity of a few hundred passengers. The first alternative appears unlikely since it would require very powerful turbofan engines and their rating appears to be saturating in unison with the aircraft performance (see Figure 2.4). It would also imply that a very large number of passengers would have to be processed at terminals over very short intervals, fundamentally creating new peak load management problems for air control, airports and many service systems. Thus, a new engine (perhaps turbofan/rampet or ram/scram) could power a supersonic hypersonic aircraft of the next century with a performance of a few thousand passenger-kilometers per hour. The natural gas economy also provides the necessary liquid methane or other endothermic hydrocarbon fuel and later also cryogenic hydrogen.

APPENDIX

I. Estimation Methods for the Logistic Growth Function

The analysis of the development of transport infrastructures and systems was based on the hypothesis that technological growth and substitution processes can be described by the logistic function. In the simplest case the logistic growth function describes the technological life cycle from the early development phase, to the rapid growth and expansion phase, all the way to the eventual saturation phase. This is very similar to the use of the function in biology to describe growth processes. Examples include the growth of organisms and populations in constrained environments, but also range to more complex processes such as the growth of a child's vocabulary or bio-assay.

By analogy, we have used the logistic growth curve to analyze the expansion of air travel, aircraft productivity, length of surfaced roads, rail tracks, canals and oil and gas pipelines. This was the simplest case of technology (or infrastructure) growth because in all of these examples one single logistic curve described the whole technological life cycle. Unfortunately, although this is the simplest class of examples when compared with the more complex pattern of technological substitution, the statistical problems of estimating the parameters of the logistic function from the empirical data are sometimes substantial, especially in those cases where the ultimate saturation level has not been reached.

In brief, the problem is to estimate the three parameters of the logistic function:

$$x(t) = \frac{\kappa}{1+\exp(-\alpha t - \beta)}, \tag{A.1}$$

where t is the independent variable usually representing some unit of time and α, β and κ are the unknown parameters. Alternatively, the logistic growth curve can be expressed as a linear function of time by taking the (usually) unknown parameter κ to the left side of the equation and by taking logarithms of both sides:

$$\ln\frac{x(t)}{\kappa - x(t)} = \alpha t + \beta, \tag{A.2}$$

which is a very convenient form for showing the logistic growth process on the semi-logarithmic paper because the historical data indicate a linear secular trend (assuming that the parameter κ can be estimated from the data or that it is *a priori* known).

The three parameters have the following interpretation in terms of the underlying growth process. α denotes the rate of growth, β is the location parameter (it shifts the function in time, but does not affect its shape) and κ is the asymptote that bounds the function and therefore specifies the level at which the growth process saturates (as t tends to infinity $x(t)$ approaches κ). Thus, all three parameters have clear physical interpretations, although α and β values are not necessarily intuitively clear. The logistic function is a symmetric S-shaped growth function, which means that it has a point of inflection, t_0, where the growth rate reaches a maximum ($\dot{x}(t_0)=-\alpha\kappa/4$). Symmetry implies that the value of the function is half the asymptote at the point of inflection ($x(t_0)=\kappa/2$).

Thus, we can define the location parameter of the function as the point of inflection or alternatively, at the time when 50 percent of the saturation level, κ, is reached. In terms of the parameters, the point of inflection is given as $t_0 = -\beta/\alpha$. The growth rate of the function can be defined alternatively as the length of the time interval needed to grow from 10 to 90 percent of the saturation level, κ. We call the length of this interval $\Delta t = (\ln 81)/\alpha$. This second set of three estimated parameters, κ, Δt and t_0, also specifies the logistic growth curve (in the same way as κ, α and β do), but in addition these alternative parameters all have very clear and obvious intuitive interpretations.

We used the method of non-linear least square regression to estimate the three unknown parameters of the logistic function from the empirical observations. In order to test the sensitivity of the estimated values of the parameters with respect to different assumptions about the errors and weighing of observations, we also used alternative estimation algorithms.[6] The estimation methods used in the analysis are reported in Grübler, Posch and Nakicenovic (1987).

The values of the estimated parameters for the logistic growth processes are given in the Table 1 with the correlation coefficient R^2 and uncertainty ranges for the parameters. The estimation of the uncertainty ranges is based on a Monte Carlo simulation approach by Debecker and Modis (1986) for the three-parameter logistic function. A total of 33693 different S-curves were generated and subsequently fitted by the logistic function providing values for κ, Δt and t_0, with known and varying distribution of statistical fluctuations. They concluded that the value of the estimated asymptote, κ, is the most sensitive (it varies the most) depending on the accuracy of the data and amount of data available. In general, their results indicate that as a rule-of-thumb the uncertainty of the parameter κ will be less than 20

6 In this particular application, the difference was basically in the assumptions about weighing of the observations in the estimation procedure, e.g. whether unit or data dependent weights are used.

Table 1 Estimates of Three-Parameter Logistic Functions.

Figure	Units	κ	κ-range	Δt	Δt-range	t_0	R^2
2.1	10^9 pass-km/yr	1747.2	±91.9	30.0	±0.8	1977	0.996
2.2	10^6 pass-km/hr	199.5	±10.5	30.0	±0.8	1977	0.996
2.3	10^3 pass-km/hr	1224.4	±197.5	30.8	±1.0	1971	0.994
2.4	pistons hp	3858.4	±137.3	29.7	±0.7	1936	0.977
2.4	jets kg	28380.0	±910.0	25.6	±0.6	1967	0.968
2.7	10^6 vehicles	358.8	±76.3	95.5	±3.7	1994	0.997
2.10	10^3 miles	3485.8	±111.8	58.6	±1.3	1948	0.996
2.11	10^3 miles	319.5	±10.3	54.0	±1.2	1890	0.998
2.11	(decline)	319.5	(none)	-117.8	±4.8	2001	0.952
2.13	miles	4003.9	±128.1	28.9	±0.6	1832	0.991
3.5	10^3 miles	238.8	(none)	59.2	(none)	1937	0.960
3.5	(alternative)	246.5	±8.5	78.7	±1.8	1937	0.983
3.6	10^3 miles	1428.8	±56.9	60.2	±1.5	1965	0.993

percent within a 95 percent confidence level, provided that at least half of the data are available (i.e. at least up to the point of inflection) with precision better than ten percent. A more comprehensive treatment of the uncertainty ranges and the estimation methods is given in Posch, Grübler and Nakicenovic (1987).

In all cases, except the example from Figure 3.4, the values of the estimated parameters were within the specified uncertainty ranges. Even alternative algorithms used to estimate the parameters provided values within the specified ranges.

II. The Technological Substitution Models and Parameter Estimates

In general, the examples of technological substitution are inherently more complex than the determination of single logistic growth pulses. However, from the statistical point of view, the estimation of substitution processes is much simpler. Setting $\kappa=1$ we obtain a two-parameter logistic function with a *known* (unit) asymptote. In other words, the three-parameter logistic function $x(t)$ can be normalized by setting $f(t)=x(t)/\kappa$ and for given values of κ reduces to a function with two unknown parameters. In the examples of technological substitution this known asymptote specifies the size of the "market" in which old technologies are substituted by new ones. In the simplest case, there are only two technologies. The market shares are $f(t)$ for the new technology and $1-f(t)$ for the old one (see Fisher and Pry, 1971). Thus, only the values of t_0 and Δt must be determined from the observations.

Fisher and Pry used the two-parameter logistic function to describe a large number of technological substitution processes. The basic assumption postulated by Fisher and Pry is that once substitution of the new for the old has progressed as far as a few percent, it will proceed to completion along a logistic substitution curve

$$\ln\frac{f(t)}{1-f(t)} = \alpha t + \beta, \tag{A.3}$$

We have used ordinary least squares to estimate parameters Δt and t_0 in all cases where one old technology is substituted by a new one. Table 2 gives the estimated parameter values, the estimation period and the correlation coefficients for the three examples given in the paper.

In dealing with more than two competing technologies, we have had to generalize the Fisher-Pry model since logistic substitution cannot be preserved in all phases of the substitution process. Every given technology undergoes three distinct substitution phases: growth, saturation, and decline. The growth phase is similar to the Fisher-Pry substitution, but it usually ends before complete market takeover is reached. It is followed by the saturation phase, which is not logistic but which encompasses the slowing of growth and the beginning of decline. After the saturation phase of a technology, its market share declines logistically, (e.g., see the path of coal substitution in Figure 3.2 or railways and roads substitution in

Figure 2.14).

We assume that only one technology saturates the market at any given time, that declining technologies fade away steadily at logistic rates "uninfluenced" by competition from new technologies, and that new technologies enter the market and grow at logistic rates. The current saturating technology is then left with the residual market share and is forced to follow a nonlogistic path that curves from growth to decline and connects its period of logistic growth to its subsequent period of logistic decline. After the current saturating technology has reached a logistic rate of decline, the next oldest technology enters its saturation phase, and the process is repeated until all but the most recent technology are in decline.

Let us assume that there are n competing technologies ordered chronologically in the sequence of their appearance in the market, technology 1 being the oldest and technology n the youngest, i.e., $i=1,2,\ldots,n$. This means that all technologies with indices k where $k<j$ will saturate before the technology with index j, and technologies l where $l>j$ will saturate after technology j.

Let us denote the historical time series by $x(t)$ where the indices $i=1,2,\ldots,n$ represent the competing technologies and t the time points of the historical period for which the data are available, i.e., year, month, etc. The fractional market shares of competing technologies, $f_i(t)$, are obtained by normalizing the sum of the absolute shares to one:

$$f_i(t) = \frac{x_i(t)}{\sum_j x_j(t)}. \qquad \text{(A.4)}$$

By applying the linear transform of the logistic function to the fractional market shares,

$$y_i(t) = \ln\left[\frac{f_i(t)}{1-f_i(t)}\right], \tag{A.5}$$

we have n transformed time series with piece-wise linear secular trends. In fact, there are only three distinct possibilities - either a decreasing or an increasing linear trend or a phase of linear increase connected by a nonlinear saturation phase with a phase of linear decline. The oldest technology ($i=1$) always displays a declining linear trend and the youngest technology ($i=n$) an increasing linear trend (see, Figure Figure 3.2). These linear trends can be estimated, including the increasing linear trends of technologies that enter the saturation phase during the historical period.

These linear trends are estimated for each competing technology by ordinary least squares. Table 2 gives the estimated parameter values for the three multiple substitution processes given in the paper. It also gives the estimation period (historical time interval for which the parameters were estimated) and the correlation coefficient. Parameter values are not given for cars in Figure 2.15 because they have been saturating during the whole historical period and consequently their substitution path is specified in the model.

Thus, we have n equations with estimated parameters Δt and t_0. These n estimated linear equations can be transformed into n logistic functions in coefficients α and β:

$$f_i(t) = \frac{1}{1+\exp(-\alpha_i t - \beta_i)}, \tag{A.6}$$

where $f_i(t)$ are now the estimated fractional market shares of technology i. Due to

Table 2 Estimates of Two-Parameter Logistic Function

Figure	Units	Estimation Period	Δt	t_0	R^2
2.6	10^6 vehicles				
	horses	1900-1919	-12.3	1916	0.999
	cars	1900-1919	12.3	1916	0.999
2.9	10^3 miles				
	unsurfaced	1904-1982	-72.7	1948	0.967
	surfaced	1904-1982	72.7	1948	0.967
2.12	10^3 locomotives				
	steam	1939-1957	-12.6	1951	0.992
	diesel	1939-1957	12.6	1951	0.992
2.14	10^3 miles				
	canals	1840-1900	-59.9	1841	0.985
	railways	1840-1860	55.0	1841	0.999
	roads	1890-1940	86.5	1915	0.996
	airways	1930-1980	127.5	2030	0.990
2.15	10^9 pass-km/yr				
	railways	1950-1980	-47.9	1921	0.982
	bus	1950-1980	-125.2	1862	0.885
	cars	(saturation)			
	airways	1950-1980	63.3	2004	0.969
3.2	GWyr/yr				
	wood	1860-1900	-61.9	1883	0.992
	coal	1830-1895	69.4	1885	0.989
	oil	1890-1950	70.6	1886	0.960
	nat-gas	1890-1965	99.0	1985	0.975

the fact that these estimated logistic functions do not capture the saturation phases and represent only growing or declining logistic trends, for some t their sum will exceed 1. Although they do not necessarily sum to 1 for all t, it is obvious that this condition must be satisfied all the time. This condition holds by definition for the observed market shares. To avoid this problem, we leave $n-1$ estimated logistic equations in their original form (i.e., as specified by coefficients α_i and β_i) and define the n-th equation as the residual, as the difference between the sum of other n-1 estimated market shares $f_i(t)$ and 1. The residual equation represents the *oldest still growing* technology and we call this technology j, such that $\alpha_j \geq 0$ where $\alpha_{j-1} < 0$ and $j > 1$. This chosen technology cannot be the oldest technology, i.e., $j \neq 1$, since the oldest technology is substituted by the newer technologies and, consequently, its market shares decline logistically from the start, i.e., $\alpha_1 < 0$.

Given the estimated coefficients α_i and β_i, and the sequence of the appearance of the technologies on the market $i=1,2,\ldots,n$, the estimated fractional market shares are defined by

$$f_j(t) = 1 - \sum_{i \neq j} \frac{1}{1+\exp(-\alpha_i t - \beta_i)} \tag{A.7}$$

and

$$f_i(t) = \frac{1}{1+\exp(-\alpha_i t - \beta_i)}, \tag{A.8}$$

where $j = \{k=1,2,\ldots,n \mid \alpha_{k-1}<0, \alpha_k \geq 0\}$ and $i \neq j$ (i.e., $i=1,\ldots,j-1,j+1,\ldots,n$), for technology j in its transition period at time t, and where $i \neq j$, for all other technologies i. Thus, at the beginning (at time t_0) we have $n-1$ technologies denoted by indices $i \neq j$ which follow logistic substitution paths, at least one technology $i=1$ which leaves the market logistically, and only one technology j which forms the residual of the market, i.e. the complement of the sum of other technologies and 1. Let us denote the point in time when technology j is defined as a residual t_j. When we apply the linear transform of the logistic function to the market shares of technology j, defined above, we obtain a nonlinear function:

$$y_j(t) = \ln\left[\frac{f_j(t)}{1-f_j(t)}\right]. \tag{A.9}$$

This function has negative curvature, it increases then passes through a maximum where technology j has its greatest market penetration, and finally decreases. After the slope becomes negative, the curvature diminishes for a time, indicating that $f_j(t)$ is approaching the logistic form, but then, unless technology j is shifted into its period of logistic decline, the curvature can begin to increase as newer technologies acquire larger market shares. Phenomenological evidence from a number of substitutions suggests that the end of the saturation phase should be identified with the time at which the curvature of $y_i(t)$, relative to its slope, reaches its minimum value. We take this criterion as the final constraint in our generalization of the substitution model, and from it we determine the coefficients for the j-th technology in its logistic decline.

Thus, we search for the point in time where the rate of decrease of $y_j(t)$ approximates a constant. From this point on, we set the rate of change equal to this constant and thus define a new logistic function. We approximate this point of constant slope by requiring that the relative change of slope is minimal,

$$\frac{\ddot{y}_j(t)}{\dot{y}_j(t)} = minimum \tag{A.10}$$

for $t_j \le t < t_e$, $\ddot{y}_j(t) < 0$ and $\dot{y}_j(t) < 0$.

If this condition is satisfied, let us denote the time point when this occurs $t_{j+1} > t_j$, we determine the new coefficients for technology j

$$\alpha_j = \dot{y}_j(t_{j+1}) \tag{A.11}$$

and

$$\beta_j = y_j(t_{j+1}) - \dot{y}_j(t_{j+1})t_{j+1}. \qquad \text{(A.12)}$$

After time point t_j, technology $j+1$ enters its residual phase, and the process is repeated either until the last technology n enters the saturation phase, or the end of the time interval (t_e) is encountered.

These expressions determine the temporal relationships between the competing technologies. They have been formulated in algorithmic form. Figure 1 shows the flowchart of the algorithm that describes the logistic substitution process. A more detailed description of this procedure and the software package for the generalized logistic substitution model is given in Nakicenovic (1979 and 1984).

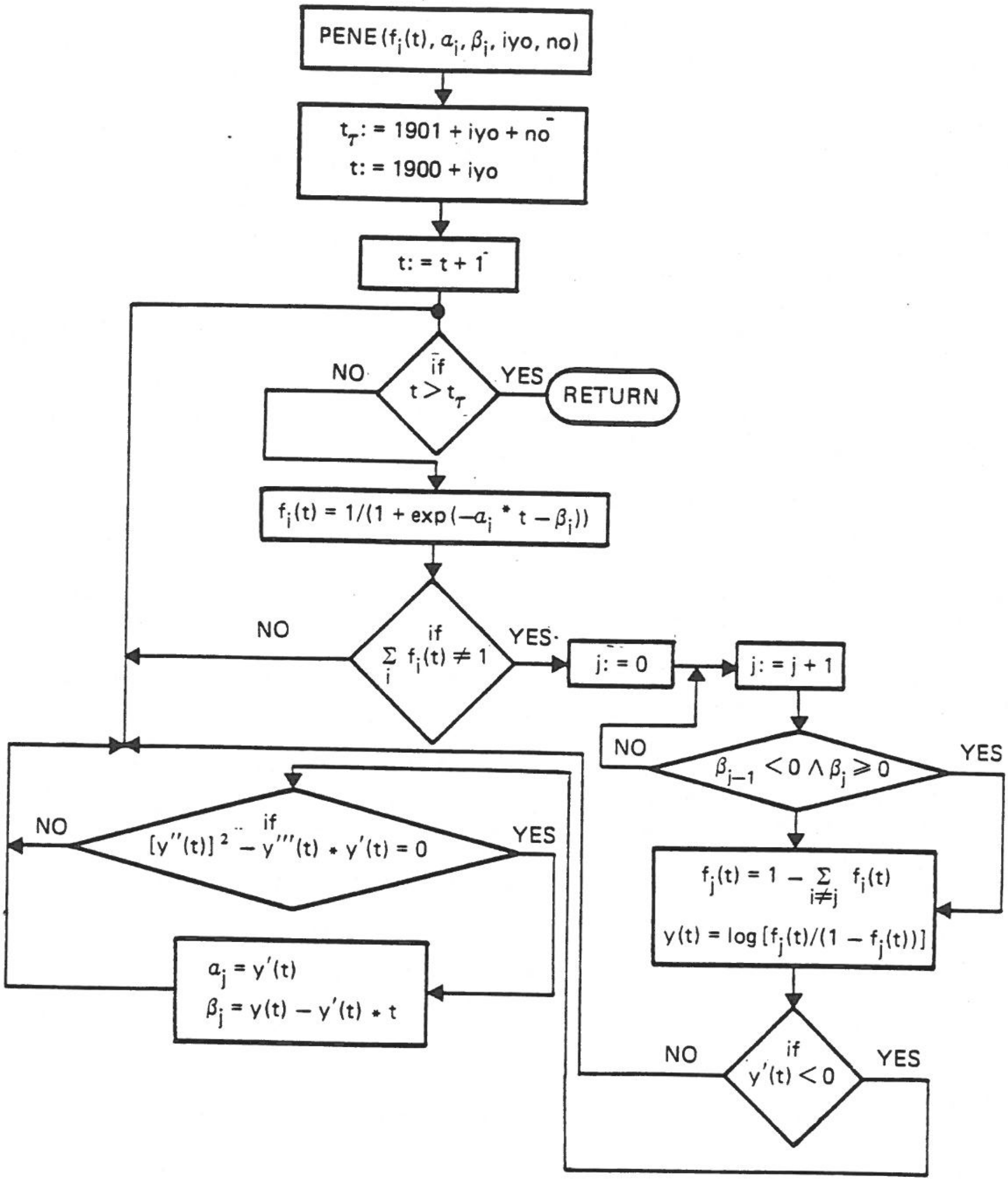

Figure 1 Flowchart of the logistic substitution algorithm.

BIBLIOGRAPHY

Abernathy, W. J., 1978, The Productivity Dilemma, Roadblock to Innovation in the Automobile Industry, Baltimore and London: The John Hopkins University Press.

Aerospace Facts and Figures, 1962, Washington, D. C.: Aerospace Industries Association of America, Inc.

Aerospace Facts and Figures, 1964, Washington, D. C.: Aerospace Industries Association of America, Inc.

Aerospace Facts and Figures, 1966, Washington, D. C.: Aerospace Industries Association of America, Inc.

Aerospace Facts and Figures, 1984, Washington, D. C.: Aerospace Industries Association of America, Inc.

Aerospace Facts and Figures, 1985, Washington, D. C.: Aerospace Industries Association of America, Inc.

Automobile Facts and Figures, 1971, Detroit: Motor Vehicle Manufacturers Association of the United States, Inc.

Casti, J. and N. Nakicenovic, 1983, The World Automotive Industry in Transition: A Framework for Projection into the 21st Century, WP-83-2, March 1983, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Chilton's Automotive Industries, 1985, October, Volume 165, Nr. 10, p. 23.

Epstein, R. C., 1928, The Automobile Industry, Its Economic and Commercial Development, Chicago & New York: A. W. Shaw Company.

Fisher, J. C., and R. H. Pry, 1971, A Simple Substitution Model of Technological Change, Technological Forecasting and Social Change, 3:75-88.

Fisher, J. C., 1974, Energy Crises in Perspective, New York, John Wiley and Sons.

Fisher, J. C., 1985, personal communication.

Forest Service Report, 1946, A Reappraisal of the Forest Situation, Potential Requirements for Timber Products in the United States, Forest Service Report No. 2.

Goodrich, C., 1960, Government Promotion of American Canals and Railroads, 1800-1890, New York: Columbia University Press.

Goodrich, C., J. Rubin, H. J. Cranmer and H. H. Segal, 1972, Canals and American Economic Development, Port Washington: Ira J. Friedman Division, Kennikat Press.

Grübler, A., 1986, Forthcoming, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Grübler, A., and N. Nakicenovic, 1987, The Dynamic Evolution of Methane Technologies, WP-87-2, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Grübler, A., N. Nakicenovic and M. Posch, 1987, Algorithms and Software Package for Estimating *S*-Shaped Curves, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Hadfield, C., 1986, World Canals, Inland Navigation Past and Present, London: David & Charles.

Harcourt, L. F. V., 1896, Rivers and Canals, Vol. I and II, Oxford: At the Clarendon Press.

Jane's, 1919, All the World's Aircraft 1919, London: Jane's Publishing Company.

Jane's, 1960, All the World's Aircraft 1959-1960, London: Jane's Publishing Company.

Jane's, 1970, All the World's Aircraft 1969-1970, London: Jane's Publishing Company.

Jane's, 1980, All the World's Aircraft 1979-1980, London: Jane's Publishing Company.

Jane's, 1984, All the World's Aircraft 1984-1985, London: Jane's Publishing Company.

Marchetti, C., and N. Nakicenovic, 1979, The Dynamics of Energy Systems and the Logistic Substitution Model, RR-79-13, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Marchetti, C., 1979, Energy Systems - The Broader Context, Technological Forecasting and Social Change, 14:191-203.

Marchetti, C., 1983, The Automobile in a System Context, The Past 80 Years and the Next 20 Years, Technological Forecasting and Social Change, 23:3-23.

Marchetti, C., 1986, Fifty-Year Pulsation in Human Affairs, Analysis of Some Physical Indicators, Futures, June 1986:376-388.

Martino, J. P., 1983, Technological Forecasting for Decision Making, Second Edition, New York: North-Holland.

Nakicenovic, N., 1979, Software Package for the Logistic Substitution Model, RR-79-12, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Nakicenovic, N., 1984, Growth to Limits, Long Waves and the Dynamics of Technology, University of Vienna.

Nakicenovic, N., 1986, The Automobile Road to Technological Change, Diffusion of the Automobile as a Process of Technological Substitution, Technological Forecasting and Social Change 29:309-340.

Poor, H. V., 1860, History of the Railroads and Canals of the United States of America, Vol. I, II, III, New York: John H. Schults & Co.

Putnam, P. C., 1953, Energy in the Future, New York: D. Van Nostrand Co.

Reynolds, R. and A. Pierson, 1942, Fuel Wood Used in the United States, 1630-1930, USDA Forest Service, Circular No. 641.

Schurr, S. and B. Netschert, 1960, Energy in the American Economy, 1850-1975, An Economic Study of its History and Prospects, Baltimore, Published for Resources for the Future, Inc. by The Johns Hopkins Press.

Schumpeter, J.A., 1935, The Analysis of Economic Change, Review of Economic Statistics, Volume 7:2-10.

Taylor, G. R., 1962, The Transportation Revolution, 1815-1860, Vol. IV, The Economic History of the United States, New York: Holt, Rinehart and Winston.

U. S. Department of Commerce, 1970, Historical Statistics of the United States, Colonial Times to 1970, Vol. I and II, Bureau of the Census, Washington D. C.

U. S. Department of Commerce, 1975, Statistical Abstract of the United States 1975, 96th Annual Edition, Bureau of the Census, Washington D. C.

U. S. Department of Commerce, 1980, Statistical Abstract of the United States 1980, 101st Edition, Bureau of the Census, Washington D. C.

U. S. Department of Commerce, 1981, Statistical Abstract of the United States 1981, 102d Edition, Bureau of the Census, Washington D. C.

U. S. Department of Commerce, 1983, Statistical Abstract of the United States 1983-83, 103d Edition, Bureau of the Census, Washington D. C.

U. S. Department of Commerce, 1984, Statistical Abstract of the United States 1984, 104th Edition, Bureau of the Census, Washington D. C.

U. S. Department of Energy, 1980, State Energy Data Report, Statistical Tables and Technical Documentation, 1960 Through 1978, National Technical Information Services, PB80-191224, Washington D. C.

U. S. Department of Energy, 1982, Monthly Energy Review, August 1982, Energy Information Administration, DOE/EIA-0035(82/08), Washington D. C.

Ward's Automotive Yearbook, 1974, Thirty-Sixth Edition, Detroit: Ward's Communications Inc.

Ward's Automotive Yearbook, 1984, Forty-Sixth Edition, Detroit: Ward's Communications Inc.