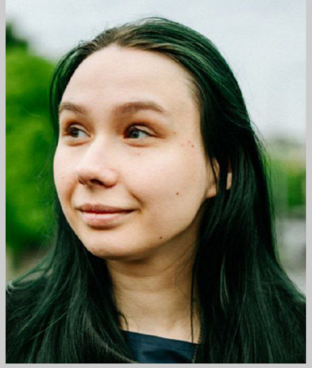
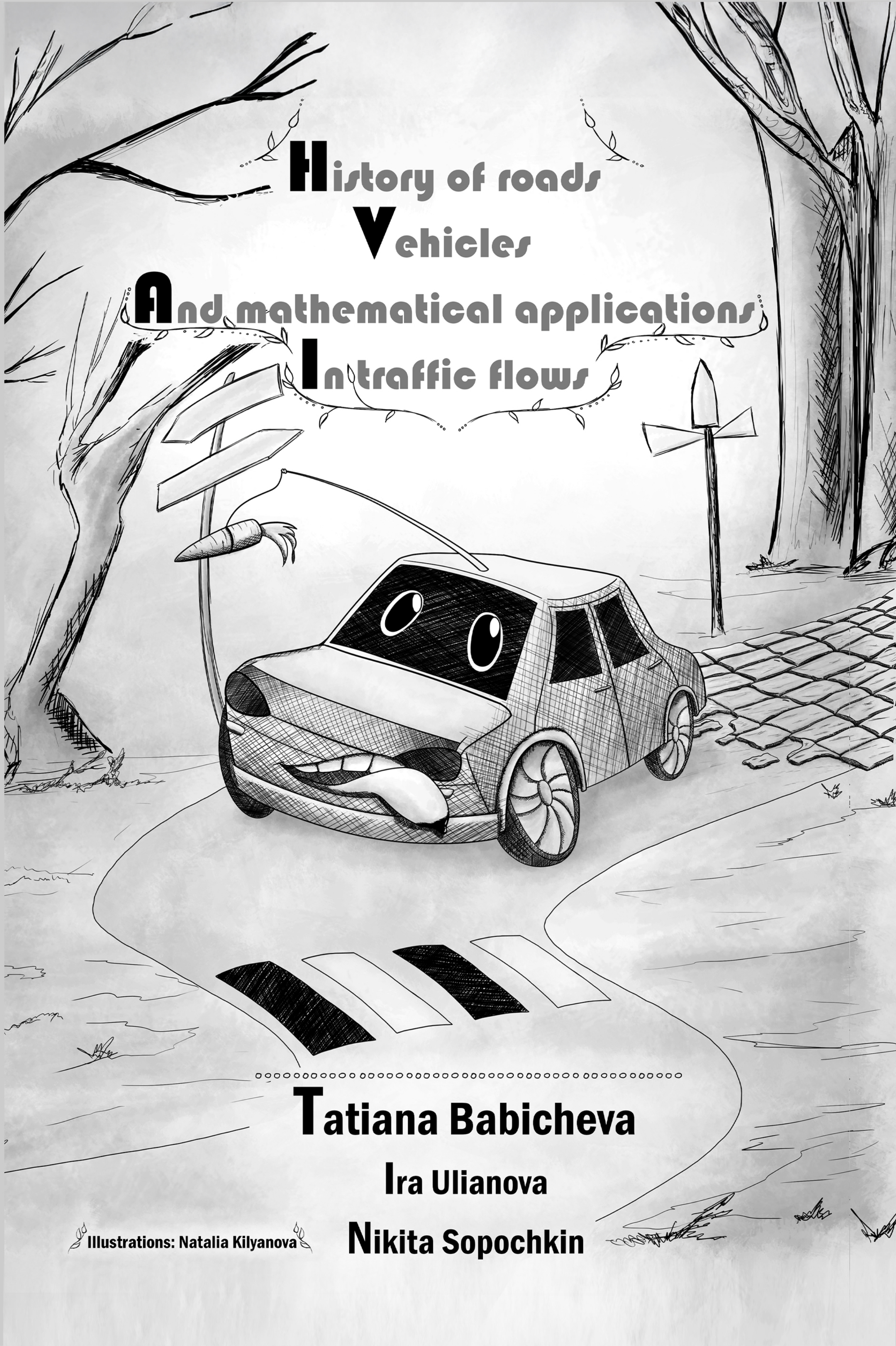
Tatiana Babicheva is an inspiring teacher who believes that teaching is not just about imparting knowledge but also about being a lifelong learner. This passion for growth and discovery led her to earn two PhDs: one in applied mathematics and its applications in Russia, and another in computer science from Université Paris-Saclay in France. She has published 10 books in Russian and French in the field of popular science, including mathematics at the competition level.

Ira Ulianova is a historian specializing in transport history and the history of education. A graduate of Herzen State University, she has devoted her research to understanding the pivotal role of transportation and education during critical periods of societal transformation. Her work primarily focuses on the Interbellum era and the Great Depression in the United States, exploring how advancements in infrastructure and policy reshaped communities and expanded opportunities.



Nikita Sopochkin is a co-writer and editor of History of Roads, Vehicles, and Mathematical Applications in Traffic Flows. Educated at the American School of Paris, Nikita has a diverse academic background with a strong interest in artificial intelligence, urban development, and sustainability. Currently studying at Babson College, Nikita continues to explore the intersection of technology and society, with a particular focus on AI-enabled solutions and smart cities.

History of Roads. Vehicles.
And mathematical Applications in Traffic flows



History of roads
Vehicles
And mathematical applications
In traffic flows

Tatiana Babicheva

Ira Ulianova

Nikita Sopochkin

Illustrations: Natalia Kilyanova

History of Roads, Vehicles, and Mathematical Applications in Traffic Flows

ISBN: 9798300307806

Cover design and illustrations: *Natalia Kilyanova*
babicheva.org.

Dedication

> “ To my family and my amazing colleagues from RATP Smart Systems
>
> —Tatiana BABICHEVA

> “ To my family, and everyone I have been fortunate enough to work with
>
> —Nikita SOPOCHKIN

> “ To everyone who goes his own way, finds and does not yield
>
> —Ira ULIANOVA

Contents

Preface from Tatiana Babicheva

My love for mathematics began in childhood. As a schoolgirl, I was the winner of the various mathematics Olympiads. Since 2006, or rather, from the first year of the Moscow Institute of Physics and Technology, I have had a keen interest in applied mathematics.

My background consists of 2 PhDs, the first in applied mathematics ("Methods of mathematical and imitational modelling of local interaction processes in transportation systems.") and the second one in informatics ("Machine Learning for the distributed and dynamic management of a fleet of autonomous taxis and shuttles"). My research interests led in areas of optimisation in complex systems, game theory, reinforcement learning. I prefer real-world applications of my knowledge starting from games and finishing with big transportation networks.

One of my passions is game theory and its popularisation, which is why I have written a book about game theory for beginners (first edition in Russian, second edition, highly changed, in French).Traffic flows are the easiest application of game theory. My appreciation of optimisation in transportation systems started with the Braess's paradox, which is linked to the conception of user equilibrium in road networks. After applying methods of mathematical and numerical optimisation during my Master's thesis and first PhD thesis, I moved to machine learning methods of optimisation in big systems of autonomous vehicles. It was an attempt to develop the artificial intelligence in such a system, simultaneously including a game theory approach (with an attempt to move towards system optimum, or to approach the Pareto front), as well as combinatorial optimisation, multiagent reinforcement learning (with different levels of granularity) etc.

After the defense of my first thesis in 2016, I received a job offer from Institut VEDECOM in France. As part of my work, I developed and implemented optimisation algorithms for a fleet of autonomous taxis (redistribution of empty vehicles, ride-sharing). Currently I am working at RATP Smart Systems as a Specialist in Operational Research, where my main goals are to develop new methods of multimodal trajectory searches in public transport networks.

In parallel with my scientific work, I have always taught. During my bachelor's and master's studies, I taught mathematics at school, during PhD studies, I taught subjects of high mathematics (Calculus and Linear Algebra) and computer science at the universities.

The love to study and to share knowledge with others led me to the fact that I am an author of several books on Olympiad mathematics (written in Russian), game theory (in Russian and in French), and I also acted as a scientific editor and expert in publishing books by other authors (in English). Also, of course, as a researcher, I am the author of many scientific articles, which you can find in google scholar.

This book is the result of the half-of-my-current-life passion to the transport, traffic flows, their optimisation and evolution. If at 2006 I have never thought that there will be real autonomous taxis, now some companies have the gouvermental permission to operate such a way, like Baidu's robotaxi at China. Soon, I will stop to regret that I have never passed my driver licence.

I cannot cover all the history and all the interesting moments of road transport in one book, but, I hope, you will enjoy its reading as much as I enjoyed its writing.

Tatiana Babicheva
Fontenay-le-Fleury, November 23, 2024

Preface from Nikita Sopochkin

The best way to examine history is through a focused lens of familiarity — looking at it in a way that allows one to compare it to what they understand, rather than through an all-encompassing view which makes it so easy to get lost. For me, as someone who with a lifelong in travel and a fascination with the logistics needed for it, this lens is that of travel.

As a result, the best way to examine this passion is by combining it with an exploration of the mathematical and specific history of travel, particularly in unexpected sections of history. It is not just a case of examining the ways we reached modern travel technology, but the strange and unexpected developments and rules that took us there. Rules established centuries ago are still used today, forming the foundations of the roads and paths that all modern travel relies on — sometimes literally, looking at the ways in which modern travel still relies on the ancient roads built so long ago. So much of what we do not examine closely or question is based on sporadic decisions or examinations made by those seeking to progress travel with the technology of their time.

As a result, while this book may not serve to provide a holistic overview of historical questions across all time, it will hopefully serve as a showcase of these past precedents in the field of travel, the work and research that informed them, the ways they were very much a product of their time, and the ways in which they have influenced us even today.

Nikita Sopochkin
Boston, November 23, 2024

Preface from Ira Ulianova

Transport history is a subject of astonishing importance. It is impossible to imagine studying the history of warfare or economic developments without considering when, where, and how people could travel. Transportation provided new possibilities but also imposed limitations that were unique to each era.

Of course, these opportunities and constraints had a profound impact on people's daily lives, the economy, and virtually every aspect related to the concept of space, everything that could be moved, and everything that could be the destination of movement.

Sometimes, economic reasons compelled the inhabitants of Imperial Russia or the United States to construct colossal transcontinental routes through desolate and dangerous lands. At other times, military ambitions led the Romans to build a road network that is still visible on the map of Europe. And sometimes, the idea that free movement is a fundamental civil right for every free individual led to the realization of global road construction projects, which involved unimaginable amounts of land movement — a feat unprecedented until the 1960s!

But what is most important to me personally is that changes in transportation possibilities alter our perception of the surrounding world and ourselves. Once, the world turned upside down with the realization that humanity could travel to unexplored lands in Asia and Americas, but even smaller details often hold significance. Imagine how a person's view of the world changes when they get behind the wheel. How does the presence of a train station or, even more so, an airport in your city affect your sense of space? Does a car driver feel more independent than a train passenger?

Not all of these questions will be answered in this book, but it will provide you with the material to contemplate how transportation shapes our lives, the world around us, and ultimately, ourselves.

Ira Ulianova
Haifa, November 23, 2024

Acknowledgments

We are deeply grateful to Grigory Simakov for his invaluable support during the creation of this book. His exceptional attention to detail, combined with his patience and professionalism, played a crucial role in refining the structure and quality of our work. Grigory's insights and assistance in navigating the challenges of editing and formatting were instrumental in bringing this project to its successful completion.

We would also like to extend our heartfelt thanks to our families for their understanding and unwavering support throughout this journey. Your encouragement and patience allowed us to dedicate ourselves fully to this work.

Finally, we owe a debt of gratitude to our colleagues and friends, whose inspiration and constructive feedback continually motivated us to strive for excellence. This book would not have been possible without each of you.

History of road construction

> Kilometres are shorter than miles. Save gas, take your next trip in kilometres.
>
> —George Carlin

1.1 History of early road construction

1.1.1 Roads of the period of primitive society and the first civilizations

In 1970, during peat excavations in the UK, traces of one of the earliest roads were discovered and named Sweet Track after its discoverer, Ray Sweet. Dendrochronology dates this road to 3807—3806 BC, making it one of the most ancient engineered roads still preserved today. The Sweet Track, found in Somerset on a swampy plain, was originally estimated to be $2,000$ km long, connecting the Neolithic pile settlement on Westhay Island with the upland village of Shapwick in Somerset.

Constructed with oak decking over beams made from young oak, ash, and linden trees, the Sweet Track was remarkable for its engineering. The materials for its construction had to be brought from elsewhere, given its swampy location. Archaeological evidence suggests that the Sweet Track was laid over an even older pathway, known as the Post-Track, which dates back to 3838 BC.

Despite the millennia, parts of the road have survived. Sections of the Sweet Track have been conserved and are on display in museums such as the British Museum and the Museum of Somerset in Taunton.

Another of the oldest roads in the UK, dating back to 3000 BC, is the Ridgeway Road. It has been preserved to this day and spans a length of 137 km. The road stretches from Wiltshire to Buckinghamshire. Initially, the track was about 400 km long, but over time, a large section fell into disrepair. It traverses the hills, which is where the name 'Ridgeway' originates, through the forested Chiltern Hills, and past the Goring Gap. Along the Ridgeway path, there are numerous historical and architectural monuments from various centuries, starting with the Neolithic era. Among them is Avebury, a mysterious structure similar to Stonehenge, consisting of tombs and sanctuaries. You can also find fortresses from the Bronze and Iron Ages, the Uffington White Horse, The Sounding Stone, and Grim's Ditch, a Bronze Age embankment that is 8 km long.

In contrast, such a developed system of roads did not exist in the island of Ireland. However, since ancient times, embankments called "toghers" were constructed there. These were piles of stones in the swamps, created over different

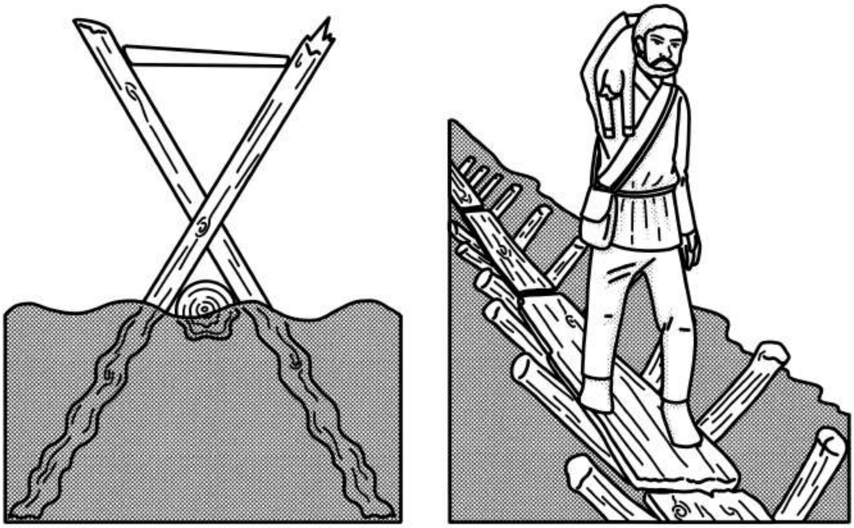


Figure 1.1: Sweet track

eras. Irish historical chronicles mention five roads leading to the Hill of Tara, the ancient capital of Ireland. The origins of these paths date back to 123 BC.

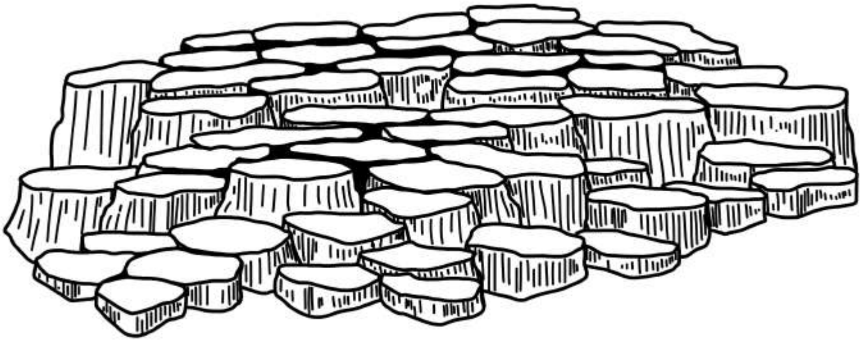


Figure 1.2: The Giants' Causeway

The Giants' Causeway, located in Northern Ireland, is of particular historical value. It consists of 40 thousand basalt columns connected to each other. These columns were formed as a result of an ancient volcanic fissure eruption, with the tallest being about 12 meters high, and the solidified lava in the cliffs reaching 28 meters (92 ft) thick in places. The road begins at the foot of the cliff, where columns with 4, 5, 7, and 8 corners form a springboard between themselves, leading into the sea. The road gained fame in the 17th century, and the first tourists visited it in the 19th century. It was declared a World Heritage Site by UNESCO in 1986 and a national nature reserve in 1987 by the Department of the Environment for Northern Ireland. Scientists explain the formation of the Giant's Causeway by the significant volcanic activity that occurred about 60 million years ago in County Antrim, Northern Ireland. Basalt, melted by high temperatures, penetrated the sediments and formed lava plateaus. The facets on the columns appeared due to the rapid cooling of the hot lava and the reduction in the volume of matter.

The development of roads was originally driven by trade and military objectives. To seize new territories and maintain control over conquered possessions, ancient states constructed networks of roadways. Roads played a crucial role in the economic and, eventually, cultural development of mankind.

In China, India, and Ancient Egypt, the network of roads was developed and the delivery of postal messages was organized, as evidenced by the traces of ancient road structures that have survived to this day. Ancient Peru had excellent roads, sometimes made of hewn stones, and on steep inclines, the roads were shaped like gentle stairs. During the construction of the Egyptian pyramids, stone blocks were transported from the banks of the Nile along specially constructed roads on skids, which were dragged along wooden beams. Herodotus reports that $100,000$ people worked for 10 years on the construction of the road from the Nile to the pyramid of Cheops. The roads in China, built in very ancient times, were made so firmly that they are still usable today. In ancient Assyria and Babylon, which conducted extensive trade with the East and West, there was a large network of roads extending to western China, India, and southward to Arabia and the country of Sheba.

The development of road infrastructure was closely linked with the evolution of transportation means. Initially, in the early stages of cultural development, humans themselves were beasts of burden, carrying loads on their backs. This primitive method of transporting goods for long distances can still be found in some parts of China, Central Africa, and the Andes in South America. With such means of transportation, people were content with narrow paths. However, when humans began to use animals (camels, horses, oxen, donkeys, deer, llamas, etc.) harnessed to logs or sledges to transport goods, narrow paths proved insufficient.

It will also be appropriate to note that, along with sledges and logs, roughly chained carts with two solid wheels appeared very early. It is impossible to determine for sure when the cart on wheels first appeared and whether it orig-

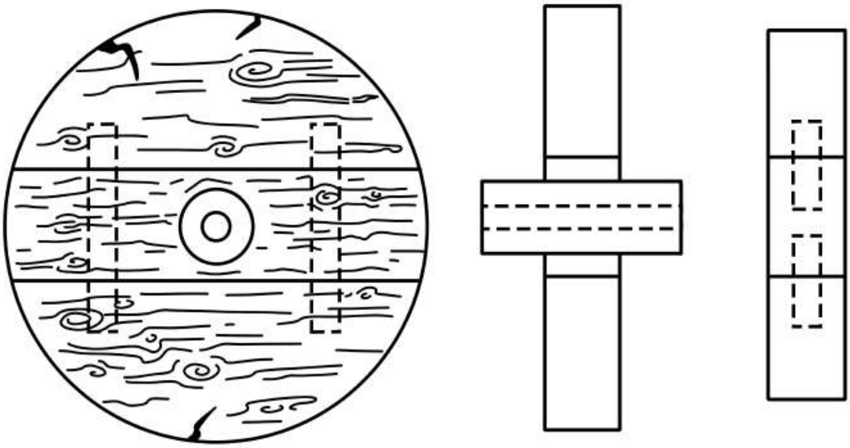


Figure 1.3: The ancient wheel reconstruction

inated from more ancient skids with rollers. In any case, the cart on wheels is very old. On many Assyrian and Egyptian bas-reliefs dated around two millennia BC, we find numerous images of carts. During the development of peat in many places in Germany, remains of wheels and carts dating back to the second millennium BC were found. Depending on the state of technology, their design was more or less perfect.

The Assyrian Empire was located in the ancient Near East, primarily in the regions corresponding to modern-day northern Iraq, northeastern Syria, and southeastern Turkey. The heartland of the Assyrian Empire was the region known as Assyria, which was situated along the Tigris River. The empire reached its peak during the 8th and 7th centuries BCE, when it was one of the dominant powers in the region, controlling vast territories across the Middle East.

It is noteworthy that the ancient Assyrians made wheels with four, six, and eight spokes, while in medieval Europe we still very often find roughly made solid wheels. In the European Bronze Age, i.e., 1700 BC, appear cast wheels with 4 or 6 spokes and strongly protruding hubs. In the Hallstatt and Late Latin eras, that is, 1000—500 years BC, there was a custom to bury warriors along with their carts. Among the most prominent discoveries are the Dejbjerg wagons from Jutland, which are two four-wheeled wagons of wood made with bronze parts.

When people began to use pack animals for transporting goods, large caravans,

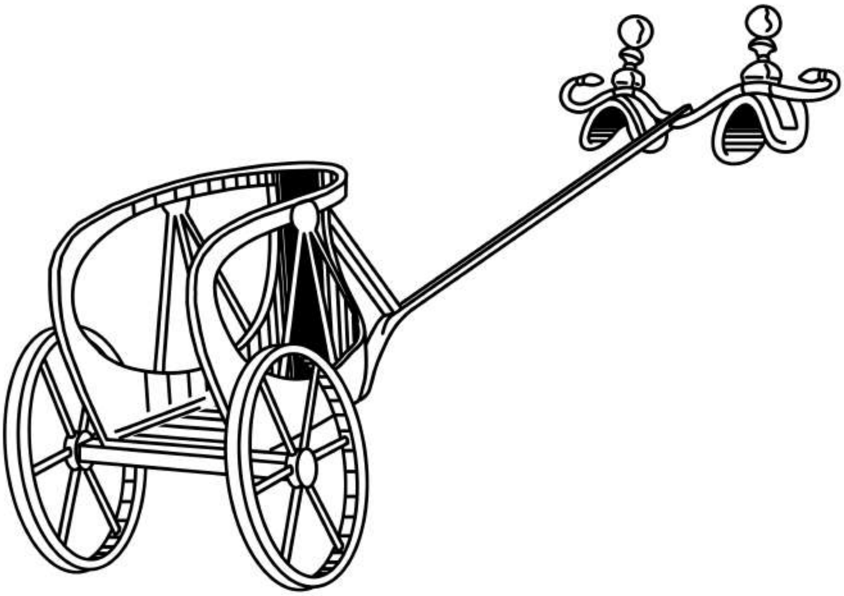


Figure 1.4: The wheel cart

often comprising 2—3 thousand camels, became common on well-trodden roads. This caravan method of trade predominated in ancient times. In Africa, this method retains its significance even now. Caravan trade necessitated ensuring water sources to quench the thirst of people and animals. These caravan routes in the East are still preserved, despite railways, as seen in Damascus, where caravans of loaded camels can be seen alongside freight trains. In ancient times, roads leading to the Chinese city Beijing were lined with trees and wide enough for a cart pulled by four animals. In India, around 300 BC, King Ashoka ordered the construction of caravanserais, milestones, and wells along roads. A caravanserai is a roadside inn along trade routes, offering shelter, food, and facilities for travelers and their animals. Roads were built by cutting down trees, leveling ground, breaking rocks, and building bridges.

The ancient Greeks, primarily a trading people and adept students of the Phoenicians, another trading civilization, paid great attention to road development, understanding the link between trade and road conditions. Greek colonies on the Black Sea served as endpoints for caravan routes from the Ganges and Indus mouths, through Bukhara, to China. The Greeks constructed sacred roads leading to temples, such as those to Delphi, Eleusis, Miletus, and Olympia. These roads featured flat ruts in rocky foundations for pilgrims and solemn chariots, resembling modern railways. Statues of gods lined these sacred roads. When constructing roads, the Greeks adapted to the terrain, preferring valleys and flat areas over mountains.

Sometimes, however, even the Greeks had to come face to face with nature. In Ancient Greece, before the Corinth Canal, ships were hauled overland across the Isthmus of Corinth to bypass the Peloponnese peninsula. Diolkos was a specially paved road with deep grooves, which connected the two sides of the Isthmus of Corinth. The ancient Greeks used to drag their ships from the Saronic Gulf to the Corinthian Gulf, specifically from the northern port of Lechaeum to the southern port of Kenchreai, thus avoiding a multi-day voyage around the Peloponnese.

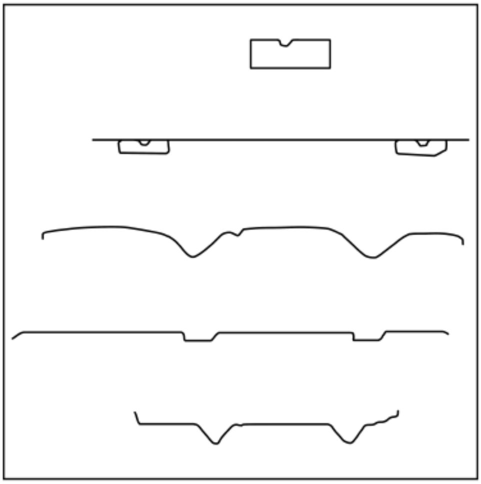


Figure 1.5: Sections of track and rails of Diolkos

The ancient Persians also became famous for building good roads. Herodotus also mentions remarkable road structures in ancient Persia, equipped with stations and buildings for travelers' lodging. Persians constructed a road from Susa to Sardis and Ephesus, spanning a length of $2,600$ km. This road, called the Royal Road, was built by the Persian ruler Darius the Great in the 5th century BCE. This route spanned from the Persian Gulf to the Mediterranean, traversing Turkey and Iran. Notable travelers like historian Herodotus, Alexander the Great, and King Midas traversed this road. Along the Persian roads, one could find many statues, bridges, temples of Zoroastrians, and distances measured in parasangs (a unit of length equal to 5.3 km) were marked with specially set stones.

One of the first transcontinental routes was the Great Silk Road, which connected the Mediterranean with East Asia. Created in the 2nd century BC, it served as a trade route, mainly transporting silk and spices. The Silk Road stretched over $7,000$ km and functioned until the 15th century. Historical records emphasize the movement of goods rather than people, as carriers often changed along the route. Camels and donkeys, numbering anywhere from three to three hundred, were used for transportation. In Europe, clothes made from oriental fabrics were fashionable, leading to mostly one-directional transportation from east to west. Since the West had little to offer China in return, settlements were made in gold and silver, resulting in a significant flow of precious metals from Europe to the East. Some states even attempted to limit purchases of goods from the East, which, however, had little effect on the course of trade.

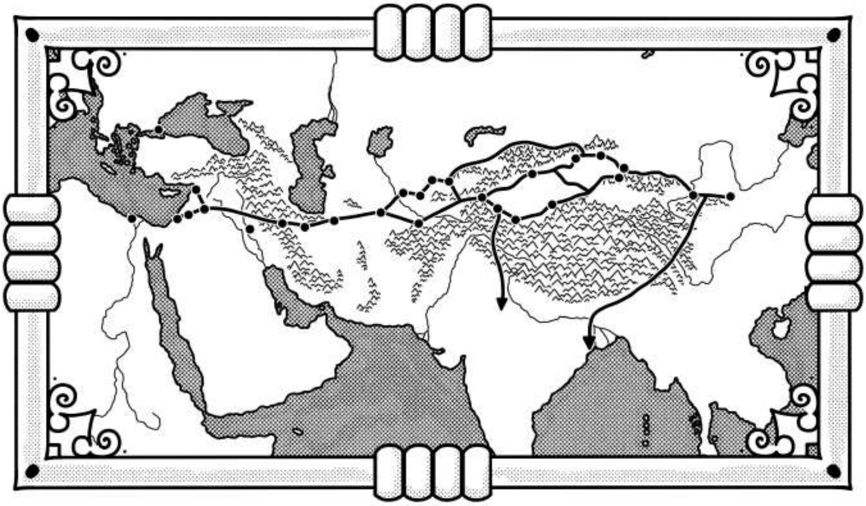


Figure 1.6: Silk road

With time and the evolution of technology, the Silk Road became a conduit for the transfer of not only goods but also new technologies. This trade route held significant economic and cultural value in the development of European and Asian countries. However, due to its length, controlling and safeguarding the Silk Road was challenging. Multiple states vied to fully own and control it. Caravans passing through were taxed, benefiting the controlling state economically. Internecine wars, driven by the struggle for ownership, often devastated settlements along the route and led to caravan robberies. This conflict over property rights persisted until the Silk Road's decline. After the fall of the Western Roman Empire in the fifth century, control over the Middle East / western part of the Silk Road remained with Byzantium. Wishing to cause economic damage to Byzantium, Persia arranged blockades on the trade route. After the Arab conquests began in the seventh century, the traffic on the Silk Road decreased significantly. To economically harm Byzantium, Persia blockaded the route. The situation worsened after the Chinese withdrawal from Central Asia in 8th century. During the Age of Discovery, the advent of sea transportation diminished the importance of caravan routes, as sea travel was quicker and ships could carry loads equivalent to thousands of animals. By the 16th century, the Silk Road ceased to function, though some local sections remained active until the 18th century.

Other significant routes included the Royal Roads. One, 225 km long, connected Thailand and Cambodia and was revered in Cambodia. Another Royal Road, the Middle East's largest trade route, linked Egypt and Syria, featuring pilgrimage sites like Mount Nebo, where Moses viewed the Promised Land. The Incense Route, another ancient trade path, connected the Mediterranean to the Arabian Peninsula, transporting goods like myrrh, incense, and African spices. In 2005, a preserved section in Israel was designated a World Heritage Site, including the ruins of Shivta, Halutsa, Avdat, and Mamshit.

In the Americas, the Inca road network, spanning several millennia, showcased advanced construction technologies. The Inca Empire, stretching from Colombia to Chile, encompassed regions of modern Ecuador, Bolivia, and Peru. Incorporating knowledge from predecessors like the Moche, Huari, Chimu, and Nazca civilizations, the Incas built an extensive network of stone-paved roads with bridges, steps, and tunnels, totaling about $30,000$ km. The network, centered in Cuzco, Peru, featured a $6,600$ km main road. Roads connected all major cities and capitals, with uniformly spaced inns and marked distances by boundary pillars called “topo”.

Excellent roads with a well-functioning courier service allowed for the management of a vast territory under unified control. Couriers, known as Chasquis, ran in a relay race fashion, passing messages to each other. They carried special horns to signal their approach to the next messenger. These Chasquis were incredibly fast and resilient, ensuring swift message delivery. Tambos, or relay stations, served as points where Chasquis could stop and transfer messages to the next runner. Each Chasqui carried a pututu, a conch shell used as a trumpet, to signal to other Chasquis when a runner was nearing, allowing them to prepare for their leg of the journey. Thus, information could be transmitted over 2000 km in just 5 days. The Incas utilized roads built by their predecessors,

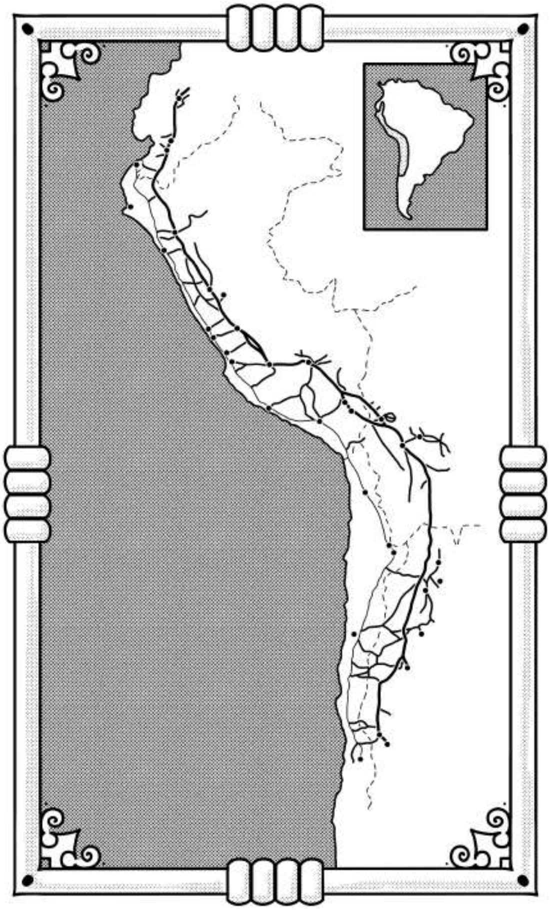
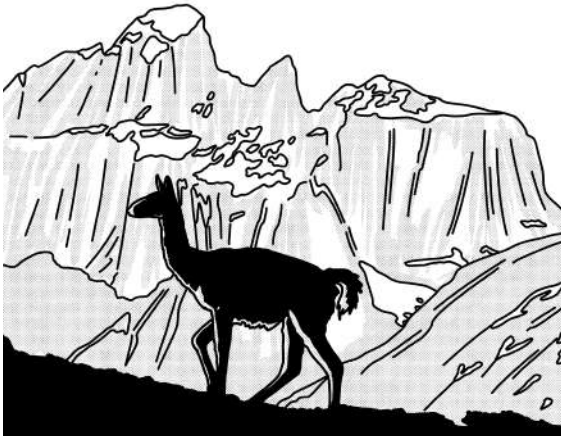


Figure 1.7: Inca road system

such as Chavín, Tiwanaku, Wari, and Chimú, and constructed approximately 16, 000 km of new roads, designed to withstand all weather conditions.

Since pre-Columbian civilizations were unfamiliar with the wheel, the Inca roads were designed for pedestrians and llama caravans. The road along the ocean coast, stretching 4, 055 km from Tumbes in the north to the Maule River in Chile, had a standard width of 7.3 m. The Andean mountain road was somewhat narrower (from 4.6 to 7.3 m) but longer, extending 5, 230 km. At least one hundred bridges were constructed on it, including wooden, stone, and even cable-stayed types; four bridges spanned the gorges of the Apurimac River. Every 7.2 km, there were distance indicators, and every 19—29 km, there were stations for travelers' rest. In addition, courier stations were located every 2.5 km. Road construction in the Inca Empire continued until the 16th century and



ceased with the Spanish conquest. The Europeans lacked the technical skills to create such a sophisticated road system, so they not only failed to continue this immense work but also neglected to maintain the existing network properly. In 1942, a Swedish expedition discovered a section of the road, which was later cleared, and now this path has become a popular tourist attraction, starting from the banks of the Rio Urubamba and ending in the Inca city of Machu Picchu.

Road networks also existed in the territory of the Slavs. Already in the III—II millennia BC in the vast areas of the middle Dnieper, the upper reaches of the Dniester and Bug, in the Carpathian region, the Black Sea region, and along the Danube, there were agricultural tribes, particularly the Trypillian tribes, which were the most developed in Europe during the pre-Scythian period. The so-called Trypillian culture, named after the initial finds near Tripillya (in the Kyiv region), is an intermediate, adjacent copper-stone culture, characterized by tribal life and a sedentary agricultural lifestyle. Some researchers [268] believe that a caravan route existed in the III-II millennia BC through the territory of the early Slavic settlements, facilitating trade relations among the peoples of Europe, Asia, and Africa.

Numerous discoveries of horse harnesses (bronze and iron bits, cheekpieces-grenzels, etc.), mostly dating to the mid-1st millennium BC, have been found across various regions of Slavic territory. In 1908, in Transcaucasia, a large wooden cart with solid wheels and a four-bull harness, dating back to the early Bronze culture around the mid-2nd millennium BC, was uncovered. Similarly, a four-wheeled wooden carriage discovered between 1937 and 1939 is also attributed to this era.

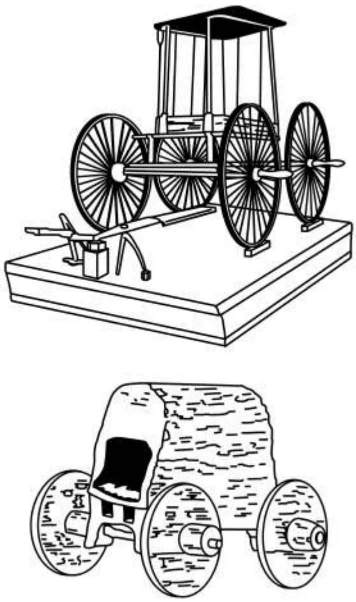


Figure 1.8: Ancient slavs cart

1.1.2 Roman roads

As you know, the Romans were ahead of all ancient peoples in the field of road construction. On the one hand, the remarkable quality, and on the other, the vast extent of the network of Roman roads, still evoke feelings of delight and surprise among our contemporaries. The primary goals of constructing these roads were to increase the state's military readiness and to facilitate long-distance trade relations. The Romans, true to their character, overcame all obstacles presented by nature in creating these communication routes. They did not hesitate to carve roads through mighty mountain ranges or to build bridges over vast plains, if, according to their design, it was necessary to connect two regions with a straight road. One of such ways was the Flaminian Way over the Apennine Mountains. In stark contrast, the Greeks adapted to the landscape of their country and took their roads along winding valleys, avoiding confrontation with nature wherever possible. Some researches think that this approach was not due to an aversion to large and challenging tasks, but rather due to the Greeks' highly developed sense of aesthetics, which often proved to be the best guide, even in technology. In our opinion, in Greece, there was a low level of centralization (or rather its absence). Building roads was economically and politically more challenging because Greece was composed of numerous city-states. Additionally, maritime communication in Greece was much more developed than, for example, in Rome due to the landscape's characteristics — the more irregular coastline created many bays, favorable for maritime activities.

The formation of the Roman road network was intimately linked with Rome's military policy, as the development of occupied territories required the establishment of an extensive road system. The Appian Way, built in 312 BC by the censor Appius Claudius Crassus, was considered the first strategic Roman road. It was the broadest paved road of its time, connecting Rome with Capua. Along this road, 6, 000 slaves who rebelled under Spartacus' leadership were crucified. The Appian Way stretched 540 km and was 7 to 8 meters wide.

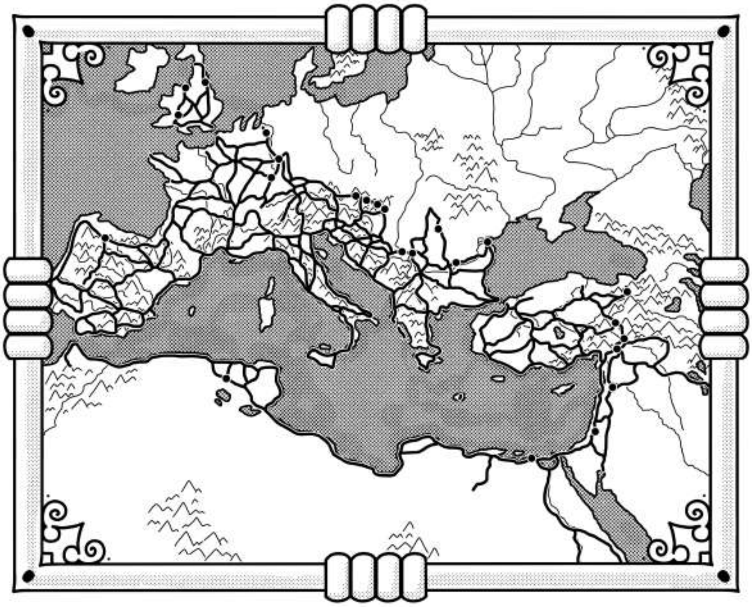


Figure 1.9: Roman road network

To clearly imagine the vast dimensions of Roman roads, it suffices to briefly mention only the main roads, adhering to the division of all Roman roads into five groups, proposed by the famous Postmaster Heinrich von Stephan. All roads originated from the center of their world, the Roman Forum. There, under the reign of Augustus, a gilded stone was erected, considered the zero-distance point for the roads. The first main road connected Rome with Africa, passing from Rome through Capua to Brindisi, and the sea route was continued from this sea port. The second main road was from Rome to Asia, stretching from Rome through Capua to Brundisium, the crossing point to Macedonia. The third main road went from Rome to Byzantium, extending from Rome to Istria, Pannonia, Thrace, and Byzantium; from here, it headed through the Bosphorus to Asia. The fourth road, from Rome to Spain, passed through Centumcellae (Trajan's famous harbor), Pisa, Genoa, and over the Pyrenees to Spain, which was interlaced with many roads. The fifth main road led from Rome to Germany and Britain.

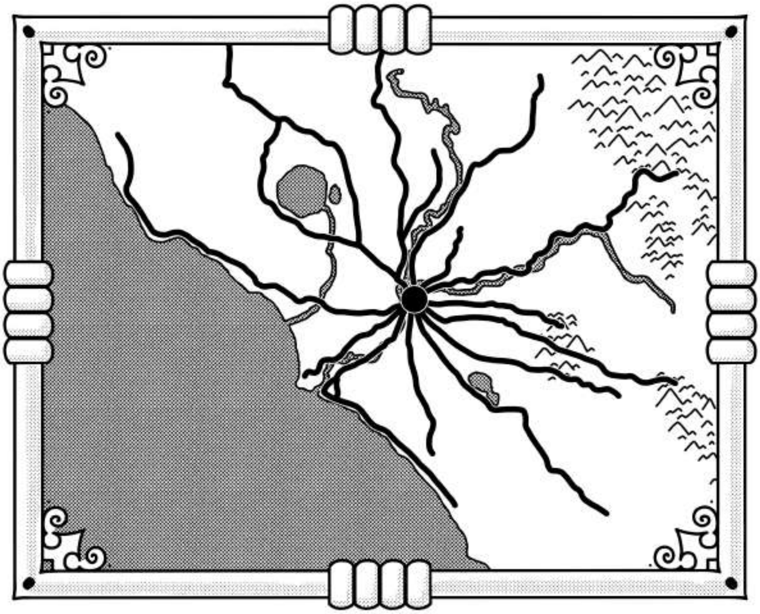


Figure 1.10: Roads went from Rome

The modern classification of Roman roads is based on the works of the Roman surveyor Siculus Flaccus and the jurist Domitius Ulpian. The Romans were among the first to employ primitive optimization methods in designing local road networks. For instance, Mommsen, in his article "Zum römischen Bodenrecht", emphasizes that the main goal of Roman land surveyors in surveying fields was to construct a road network in the fields that would provide every settler with free access from the public road to his land in the same area. These deliberately implemented orders, replacing the former "Flurgemeinschaft", ultimately prevailed in the era of the Twelve Tables laws, resulting from the conflict between the plebeians and patricians. Examining the agrimensores' data depicting the road network in the ager divisus, we observe that each limes (border) of the centuriate field was a public road of a certain width. Inside the centuries, there were no public roads, but roads "ex collatione privatorum" (made by contribution of private individuals) were established.

Roman roads varied greatly. The Romans excelled in adapting to regional conditions and always constructed roads from locally available materials. On many routes, these roads comprised solid stone massifs, comparable to lying walls. However, in the German lowlands (along the lower Rhine and Weser, particularly in Oldenburg), the Romans had to forgo such constructions. For major or significant roads, large tunnels were also built, ranging from several hundred meters to a kilometer in length. After deciding to build a road, Roman land surveyors (lat: Gromatici) began marking future routes, using various devices.

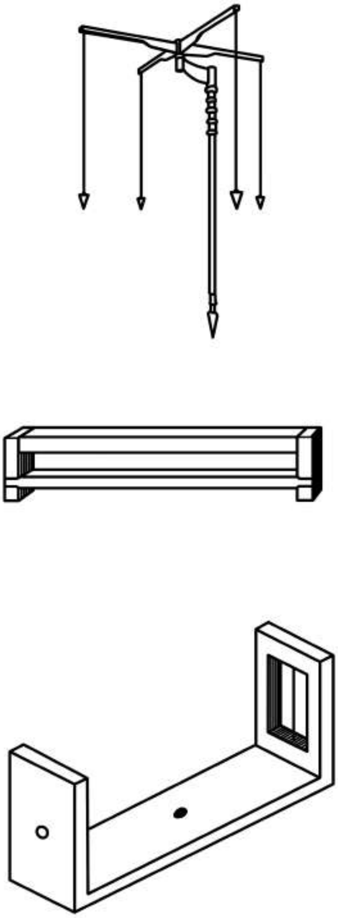


Figure 1.11: Groma, Chorobates and dioptra

- The groma, or gruma, was used to survey straight lines and right angles, hence squares or rectangles. It comprised a vertical staff with horizontal cross-pieces mounted at right angles on a bracket.
- The chorobates was used to measure horizontal planes. Similar to modern spirit levels, the chorobates consisted of a wooden beam 6 meters in length, held by two supporting legs, and equipped with two plumb lines at each end. The legs were joined to the beam by two diagonal rods with

carved notches. If the notches corresponding to the plumb lines matched on both sides, it indicated that the beam was level. On top of the beam, a groove or channel was carved. If conditions were too windy for the plumb bobs to work effectively, the surveyor could pour water into the groove and measure the plane by checking the water level.

- The dioptra was a sighting tube or, alternatively, a rod with a sight at both ends, attached to a stand. If fitted with protractors, it could be used to measure angles.

In general, Roman roads are characterized by their construction in a straight line for the maximum possible length. They typically avoided swampy areas or the immediate vicinity of rivers. When crossing water barriers, the builders preferred fords or constructed wooden or stone bridges, some of which survive to this day. In hilly terrain, the incline was minimized for safety and convenience. Sometimes, roads followed a line of equal heights before descending sharply and then continuing horizontally. At bends, roads widened to allow carriages to pass without interlocking wheels or bumping front ends. After geodetic measurements and calculations, land surveyors used milestones to mark the future road's path, clearing shrubs and other growths that hindered construction. The work of engineers was to design the road.

While individual contributors are largely unknown, records show that in 312 BC, the Roman Senate instructed the censor Appius Claudius to pave the road from Rome to Capua, a task he executed successfully.

Roads were often built simultaneously in separate sections, leading to slight directional changes, as observed by archaeologists. Soldiers frequently constructed them, particularly during peacetime, and so many roads were located near military camps. Other sections were built by slaves, local settlers, or prisoners. Contrary to popular belief, Roman roads varied in quality and maintenance, reflecting the areas they traversed.

Construction methods varied with the topography. On flat terrain, a shallow trench was dug, and after layering, the road still rose above the ground. In hilly areas, the road level was elevated using material from ditches dug alongside. Builders either leveled elevations or raised the road. In areas with settling soil, special supports were used.

After excavation, builders layered various locally sourced materials. Roman roads typically had four to five layers, including middle layers of Roman concrete. The bottom layer, about 20—30 cm thick, was stone slabs laid on a compacted subgrade and mortar, then leveled with sand. The second layer, 23 cm thick, was concrete (broken stone in mortar). The third layer, also 23 cm thick, consisted of fine gravel concrete, both compacted carefully, primarily by slaves or sometimes military units. The final layer was large stone blocks about 13 cm thick, covering an area of 0.6—0.9 square meters. For instance, the preserved parts of Appian Way have been constructed in this manner.

The following features and names of the layers of the construction of Roman roads are known nowadays. It generally consisted of the following layers:

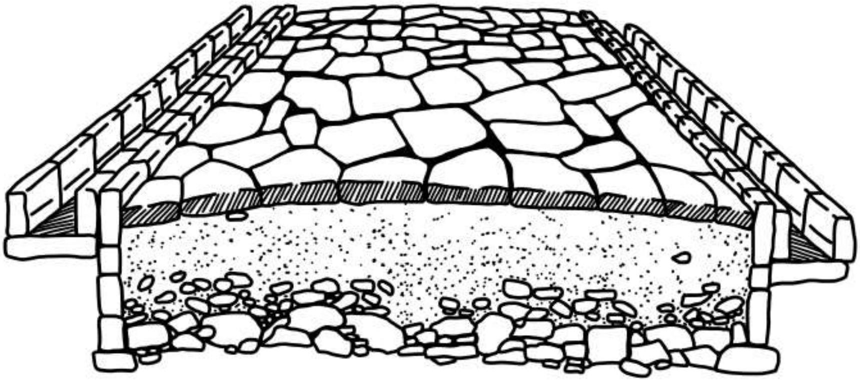


Figure 1.12: Layers of Roman roads

- Statumen (lat «backwater, abutment») — the base of the road, which was formed from large rough stone blocks. They served as the foundation of the road, and drainage was also carried out through the cracks between the slabs.
- Rudus (lat. «fragments of stones, crushed stone») or nucleus (lat. «core, hard core») — a layer of sand or a thin layer of gravel, which was placed on the statumen to level the surface. The stones used there were mostly the size of walnuts.
- Summum dorsum (lit. lat. «top surface») — the top layer of fine sand, gravel, lime, or earth. This layer had to be soft and strong at the same time. Roman roads were rarely paved, except near the cities.

The constructed road had a somewhat convex surface, allowing rainwater to drain into drainage ditches along the road. These ditches were sometimes deepened during repairs, as earth was taken from them to construct embankments. The roads, primarily used by troops, often had parallel paths for pedestrians and riders. The middle, slightly convex part was called "Agger" and was separated from the curb by a stone border. Near the road, special stones helped riders mount horses, necessary since stirrups weren't invented in Europe until the 4th century. Already in the 5th century BC, laws regulated street carriageway sizes and mandated passages between houses. Roman roads were narrow; the 12 Tables' laws specified roads to be 8 feet (2.4 m) wide, widening to 16 feet (4.8 m) on bends. Land holdings were separated by 5-foot (1.50 m) boundaries, also serving as passage paths. Roads for cattle driving had to be paved, though these requirements weren't always met.

The five main roads were central to the Roman road network, which included many smaller, less significant branches. There were large (military), provincial, communal, and hollow roads. French scientist Nicolas Bergier estimated the total road length at $10,000$ miles, but this is considered exaggerated. Roads were overseen by special commissions, with the chief warden (curator) role highly esteemed. Maintenance was publicly funded and auctioned.

The development of the post office (cursus publicus) was closely tied to road construction. During the Republic, government dispatches were sent via couriers (viatores). A well-organized system of cribs accelerated transfers. The cursus publicus, dating back to Emperor Augustus, divided the network into specific sections. Mansiones (luxuriously furnished stations) and mutationes (forage stations) marked section endpoints. Stations had about 40 horses for important persons to change horses. Roman transportation included horses, oxen, mules, and donkeys; the mail coach (rheda currens) seated 2—4 passengers. The cursus publicus administration was overseen by the Praetorian prefect. Prominent figures, like Caesar, were known for fast travel, with Caesar covering 148 km in a day, averaging about 11 km/h. The average speed was 6—8 miles per day. The cursus publicus served dignitaries, military personnel, veterans, and those with travel certificates (evectio).

The cursus publicus flourished from the 1st to the 3rd century, extending from Britain to the Tropic of Cancer and from the Pillars of Hercules to the Euphrates. The general road network was mapped (Itinerarien), including all stations. Crisis of the Third Century lead to the crisis in this movement business.

The notable Peutinger Table, named after Konrad Peutinger, is a surviving map. Comprising 12 parchment sheets, it covers lands known to the Romans in a narrow, long strip format, with about 555 city markers, 3500 names of rivers and countries, and fairly accurate distances.

The fall of Roman dominion impacted the cursus publicus. During 5th century, it was greatly restricted, eventually losing significance. In 8—9 century, the King of the Franks, Emperor of Carolingian Empire Charlemagne abolished the old postal system, retaining stations for royal commissioners. In 1464, Louis XI of France re-established a state post office.

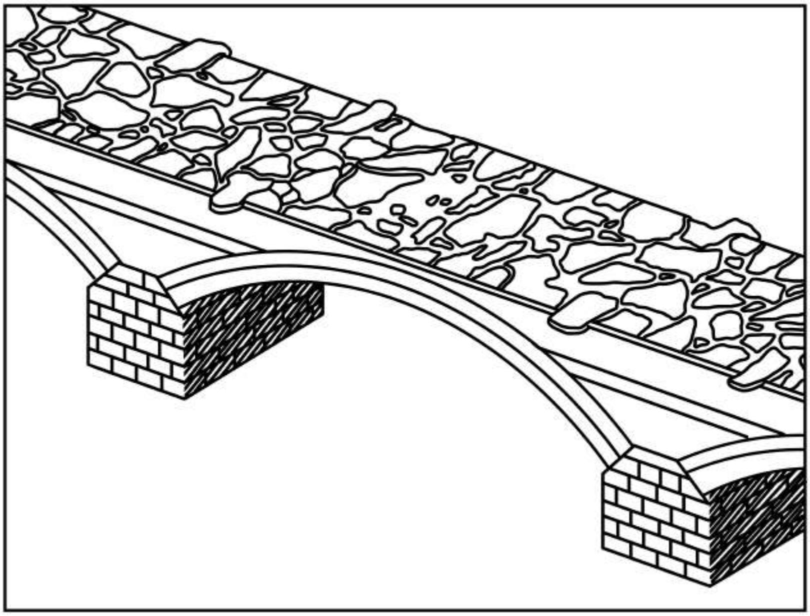


Figure 1.13: Roman bridge

As mentioned earlier, the Romans made great strides in stone bridge building. They used vaulted structures as supports and utilized cement, the secret of which was lost in the Middle Ages but later rediscovered. Bridges, more precisely aqueducts, were also used to supply cities with water. Roman historian Sextus Julius Frontinus wrote that aqueducts are the main witnesses to the greatness of the Roman Empire. Many ancient Roman bridges have survived to this day.

The Romans built bridges over relatively narrow rivers, enabling movement in any direction, even during floods. Many of these structures have survived to this day, and some are still in use. Sometimes bridges were destroyed, but the old supports were used to build new ones. Settlements often developed near these bridges. Depending on the river's width, a bridge could have one or several arches. In the latter case, each support on the upstream side of the bridge had a protrusion. These helped prevent debris carried downstream during floods from accumulating at the supports, thus protecting the bridge from destruction and the risk of being washed away. Some bridges were constructed entirely of wood and stood on piles. For greater strength, the supports were built of stone, while the superstructure was made of wood. An example of this construction is the Roman bridge at Trier, where the piers were made of stone and the decking of wood. Today, only the Roman stone pillars remain there, while the upper part was rebuilt in hewn stone later. Pontoon or floating bridges were used to cross wide rivers. At the shore, it was as if the beginning of an ordinary bridge was built, and then a pontoon bridge was attached between these supports, ensuring the bridge's stability.

How else did the Romans counter nature? In mountainous areas, it was more challenging for builders to construct straight roads, so in some cases, they cut tunnels through the rocks, allowing the road to continue unobstructed. These tunnels were often small and S-shaped, a consequence of developing adits first and then their meeting. For safety, roads were paved on the cliff side to prevent landslides, and retaining walls were used to widen the road as much as possible.

Of course, such an extensive network of roads, as well as the size of the Roman Empire, demanded efficient navigation. In addition to maps, which were available to a limited number of people and not perfectly accurate, there were local indicators. To navigate the terrain, Roman engineers erected milestones (miliaria) at regular intervals on the roadsides of viae publicae and vicinales. These cylindrical columns, 1.5 to 4 meters in height and 50 to 80 centimeters in diameter, stood on cubic bases sunk about 60—80 centimeters into the ground. Unlike modern road signs, they were not placed every mile. They indicated the distance to the nearest settlement. Atop each milestone (since travelers often rode horses or sat in carts, they could see clearly), inscriptions included the name of the emperor who ordered the road's construction or repair, his titles, a few words about the stone's origin (whether it was placed there after the construction or repair of the road), and the distance from that point to the nearest settlement, major road junction, or border. The Romans measured distances in miles. The Roman mile (lat. milia passuum) was equal to $1,000$ double steps and approximately 1.48 kilometers. On some roads, such signs were installed after the road was built, so distances were indicated in different units.

In 20 BC, Octavian Augustus became Commissioner for Roman Roads. He installed the Milliarium Aureum (lat. golden milestone) in the Roman Forum near the Temple of Saturn. All roads were supposed to start from this gilded bronze monument. It listed the largest cities of the empire and the distances to them.

The fall of the Roman Empire led to the destruction of many cities, the revival

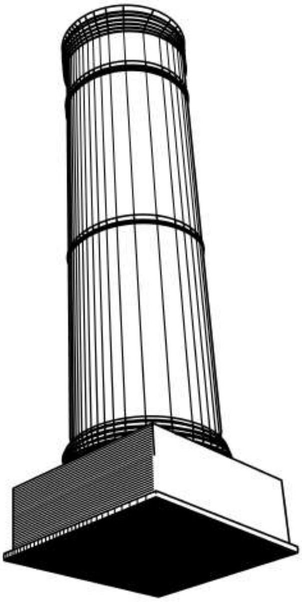


Figure 1.14: Milliarium Aureum

of which began only in the High Middle Ages. During the feudal era, road construction relied heavily on the work of the ancient Romans, but due to the expensive labor force, it was rather poorly developed.

During the Roman Empire era, in addition to ordinary separatism, rebellions occasionally erupted in distant regions. The rebels used the same roads that the conquerors were supposed to use to maintain control over these areas. This happened, for example, during the Boudiccan Rebellion in the first century AD. Later, when Germanic tribes invaded the eastern territories of the empire, they also utilized these same roads. Thus, the instrument of creation became the instrument of destruction.

Are you enjoying the book so far?

Your quick, honest review helps us immensely and only takes a minute!

Use your phone to scan the QR code and go directly to the Amazon review page.

SCAN ME

Thank you for taking the time—it truly means the world to us!

1.2 Road construction after the fall of the Roman Empire

1.2.1 Land transport in Europe in the early Middle Ages

The Western European road system fell into disrepair following the collapse of the Roman Western Empire.

In the Middle Ages, three main types of transport were used — land, river, and sea — with the majority of goods transported by sea. Land transport faced the significant obstacle of undeveloped roads. The best were the old Roman ones, with hard surfaces, lasting many centuries. Their densest network was in the territories to the west and south of the Rhine and the Danube (with key centers in Reims and Lyon); there were fewer in Spain and Britain. This network was created mainly for military-strategic purposes. Trade motives were less considered, which became apparent in the developed Middle Ages when many new trading centers emerged, and the old network increasingly conflicted with trade needs. Thus, even in Italy, the country richest in Roman roads, the 12th—13th centuries saw a "road revolution" with the emergence of many new routes aligned with trade needs. In the 13th—14th centuries, this process expanded to other Western European countries. However, the new roads were generally unpaved dirt paths, often impassable due to mud for most of the year. Travel was primarily on horseback, and goods were transported in packs. Pack transport dominated almost until the end of the developed Middle Ages, and carts, typically drawn by oxen, were only occasionally used. Consequently, transporting goods from Florence to Paris could take two or three months (a distance covered in two or three weeks under better conditions). Political circumstances were another trade barrier: insecurity on the roads, the arbitrariness of local feudal lords, and numerous customs gates (for example, there were 64 gates on the Rhine overland route in the 14th century). All this made overland trade extremely unprofitable and risky; the image of a medieval merchant often resembled that of a warrior. Only small but valuable goods were transported by land (fabrics, spices, expensive handicrafts, etc.).

The continent's main transport arteries followed rivers like the Rhine, Rhone, and Loire. However, overland routes often played a crucial role, especially in countries where rivers flowed mainly in one latitudinal direction, such as England, Spain, and Italy; here, north-south connections were primarily by roads. London was England's largest commercial and transport center. How-

Figure 1.15: Road shaped mud

ever, Coventry was the most important junction of land routes diverging to all country corners. In the Pyrenees, key meridional routes ran from Seville via Jaen to León and from Granada via Córdoba to Barcelona. In France, transport arteries concentrated in Paris; it was connected by various routes, both land and river, with almost all country regions. In Germany, the densest road network crossed the north and southwest; here were the sources of the two main river routes of the country — the Rhine and the Danube. Routes connecting Western, Eastern, and Southeastern Europe also passed through German lands. These included the route from Vienna along the Danube and its tributaries to the Balkans; from Regensburg and Prague to Krakow and further to Russia; from Pomerania through Lithuanian lands to Novgorod.

Alpine passes were crucial transport hubs in medieval Europe, divided into three groups: western, central, and eastern. In the early and developed Middle Ages, the western ones were most common — Mont Cenis, Argentière, Mont Genève, Greater and Lesser Saint Bernard, Simplon. The southernmost, Argentière, led from the Ligurian Riviera to Provence and onward to Avignon. Mont Cenis, Mont Genève, and Greater St. Bernard lay on the busiest trade route, from Lombardy and Piedmont to Champagne, Flanders, and the Rhinelands; they connected Aosta, Susa, Turin with the Lower Rhone, Grenoble, Lyon, Jura, and Lake Geneva. The most famous was Great St. Bernard; Mont Cenis was

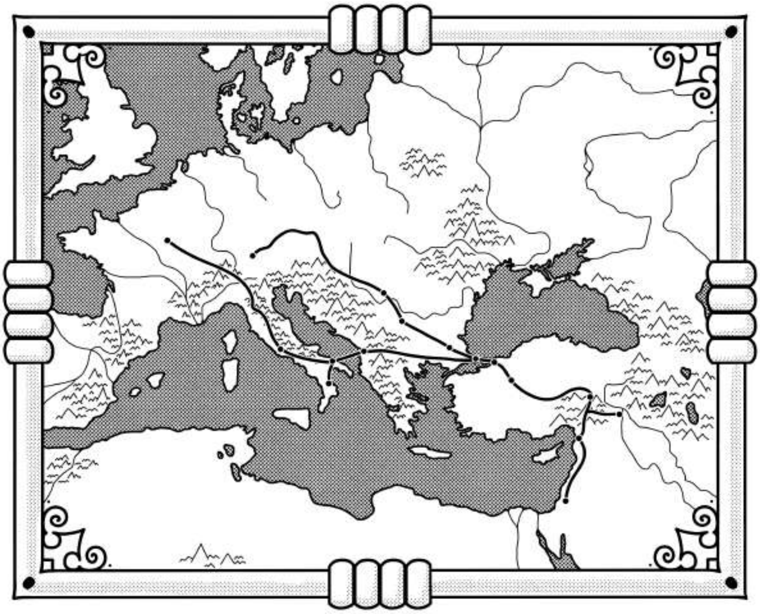


Figure 1.16: First crusade routes

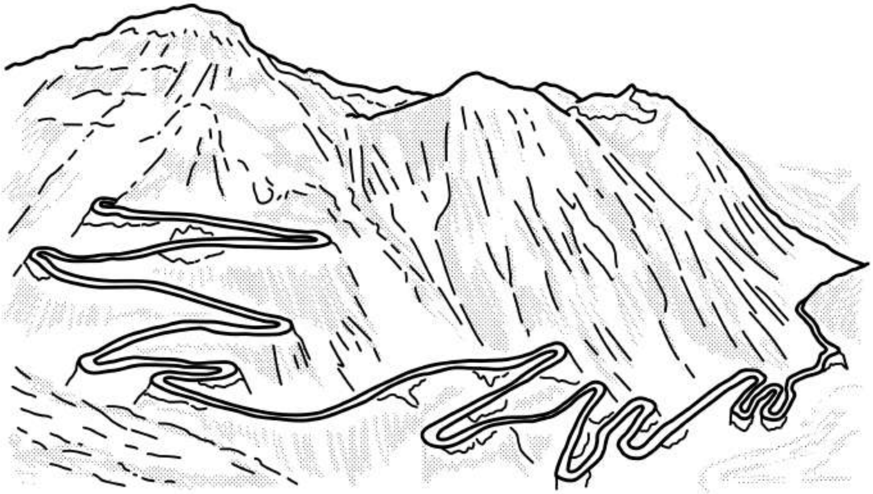


Figure 1.17: Ancient Transalpine Trade Route

most used in the early Middle Ages.

Small St. Bernard and Simplon came into use at a later time, only from the middle of the 13th century. The most important of the central Alpine passes were St. Gotthard and Septimer, lying on the shortest route between Lombardy and

Switzerland (Milan — Como — Basel) and further along the Rhine — to Germany, the Netherlands, Flanders, Champagne. At first, Septimer prevailed here, but from the 13th century, dominance passed to Saint Gotthard. In the 14th-15th centuries, due to geographical (the shortest route to the most developed trading areas of continental Europe, relative proximity to Tuscany, Venice) and political circumstances (establishment of control of the Empire over the eastern passes), the importance of the central Alpine passes increased sharply, and they became the leading ones in the communication system of Italy with the rest of Europe. However, many of them (Splügen, Lukmanier, Bernardino) were of local importance. The role of the East Alpine passes (the Brenner Pass system) was just as small. They connected Veneto with Southern Germany and further — with two water arteries, the Danube and the Oder. However, the commercial importance of this route has traditionally been relatively small; these passes played a much greater role in the political fate of the Apennine Peninsula, as it was through them that the bulk of the invasions of the lands of Italy took place both in the early and developed Middle Ages during the campaigns of the German emperors (Holy Roman Empire of the German Nation).

An important transport hub in the northern part of the continent was Jutland, located on the border of the basins of two seas - the North and the Baltic. The overland route connecting these basins passed somewhat to the south, through the lands of Schleswig-Holstein; from the beginning of the 14th century, it was replaced by the already mentioned Lübeck-Elbe canal. However, the bulk of the goods went by sea, around Jutland, through the straits of the Small and Big Belt and the Sound Strait. The latter was (until the second quarter of the 15th century) the most visited.

1.3 Roads of Eastern Europe

As for Eastern Europe, by the beginning of the 11th century, Kyivan Rus' had already reached significant development. Trade, political, and cultural ties of Kyivan Rus' with Byzantium, the West, and the East were expanding. By this time, the Kyiv princes were connected by family ties with many ruling courts.

In contrast to the previous period of the country's development, as shown by land routes during the heyday of the Kyivan state, some written and other data allow us to determine with sufficient accuracy that Kyiv was the center of all communication (and land routes in particular), having formed an extremely simplified version of the technique of road construction and the organization of road management.

Kyiv carried out the most important trade relations through five main routes, namely: the western (Lyadsky), southern (Greek), southeastern (Tmutarakan), northeastern (Murom), and northern (Novgorod and Suzdal). Kyiv also had its own way to Polesia, from where, through Pripyat, communication was carried out with Lithuania, and possibly with Sweden. Of these routes, the southern and northern routes, as well as the route to Polesia, are water-land routes. The rest of the mail routes throughout their length were overland. To characterize these routes, the western and southeastern ones are of known interest. The former was the road of relations with Western Europe, the news of which, as already noted, dates back to the 9th century. The latter was the path that connected Kyiv with the Tmutarakan principality, located at the mouth of the river Kuban on the coast of the Sea of Azov, and with Crimea; the first annalistic news about this path is from the 12th century.

The importance of land routes in the Kyivan state is also evidenced by chronicled information about the main elements of these routes (transportation, bridges, pavements, wagons, etc.). Already in the 10th century, there are references to "transportations" and a collection of various kinds of duties related to them. In the same century, there are references to carts, both four-wheeled and two-wheeled. The first chronicled mention of the use of sleds dates back to the time of Olga. Under Vladimir Monomakh, improved sleds were made.

It is possible that in Kyivan Rus' during the 9th—10th centuries (and it is absolutely indisputable in the 11th—12th centuries) there were special officials in charge of the management of roads and bridges. For example, in the Brief Russian Pravda, compiled, as historians believe, at the end of the 10th or at the beginning of the 11th century, the "lesson of bridgemen" was introduced, while the charter of Yaroslav directly refers to the "bridger", and in the Ipatiev Chronicle — "shifters" who were in charge of bridges (and, presumably, bridgers) in cities.

In the second half of the 11th century, after the death of Yaroslav, the Kyivan state began to disintegrate into a number of separate principalities, which, according to historians, was the result of the development of feudal large landownership, the political strengthening of large landowners, the economic and political growth of cities that became the new centers of certain parts of the Kyivan states. By the middle of the 12th century, and especially in the second half of it, the new political centers had become so strong and isolated that Kyiv finally ceased to be not only the capital city of a large state before that but also turned out to be far from the first among the newly formed feudal principalities. In the conditions of feudal fragmentation of the 11th—13th centuries, as well as during the heyday of the Kyivan state of the 10th—11th centuries, such factors in the development of overland routes as trade, wars, and handicraft production retained their significance, but the direction and nature of their influence were different than in the previous period. It is characteristic that during the period of feudal fragmentation, the annalistic term "path" usually meant only the "direction" along which the campaigns were going. Thus, for example, the 1127 campaign against the Krivichi followed four "paths", that is, directions — from Turov, Vladimir, Gorodok, and Klechsk.

Speaking of the development of large "straight" roads during the period of feudal fragmentation of the 11th—13th centuries, one should not think that during this period there were no such roads at all, that their direction was re-laid every time there was a new campaign, and that these directions were not relatively stable, as is confirmed by some papers. Annalistic and other data testify that even during this period, large roads, although abandoned and neglected, existed with relative stability and they were used not only for military campaigns but also for trade purposes.

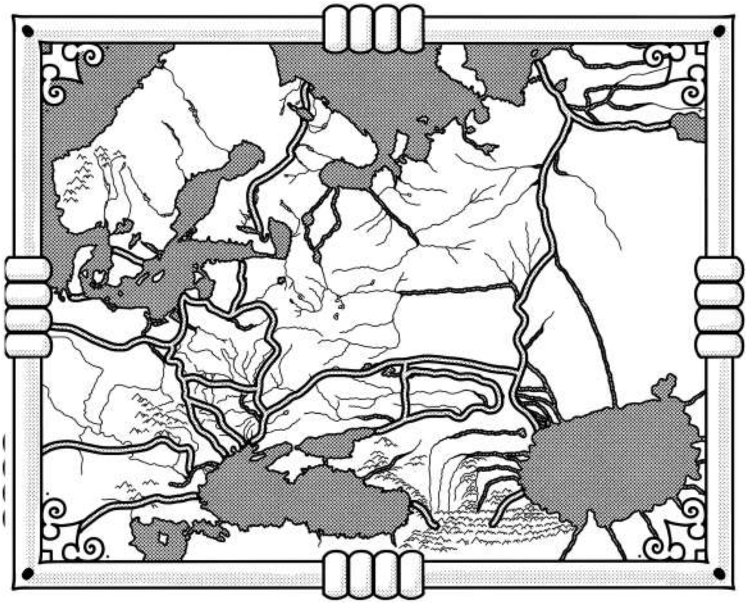


Figure 1.18: Map of Eastern Europe at 1613

As we know, the Tatar-Mongol invasion caused heavy damage to the land of Kyivan Rus' and, in particular, it had a certain effect on the development of land routes and road construction. This invasion seriously slowed down the development of the country and left a peculiar imprint on the very nature of this development. For a correct understanding of the meaning of the Mongol yoke (Mongol rule), which was mainly expressed in the system of tribute and military pressure, it is important that this yoke did not affect the organic beginnings of the life of the regular people. However, no matter how severe the wounds inflicted by the Tatar-Mongol invasion, the northeastern territory recovered relatively quickly and already from the beginning of the 14th century its political development began to be characterized by the gradual growth of the unification process, headed by the Moscow principality and which, as is known, ended at the end of the 15th and beginning of the 16th century by the formation of a Russian state.

And only in connection with the unification process, the development of crafts and trade, overland routes began to acquire a certain importance in ensuring not only internal but also external relations. Simultaneously with the strengthening of internal trade, as evidenced by both the development of their own monetary system and the peculiar customs policy of the Rus' princes and their concern for the protection of trade, a noticeable development of external trade relations of individual feudal Rus' principalities began. It is known, for example, that in the 14th century, the Novgorod-Hanse trade reached its peak.

Both the development of trade and especially the process of unification of the territories had a positive effect on the development of overland routes and the organization of communication.

According to a number of sources, it can be assumed that in the 15th century, «Novgorod was connected by roads ... with all its three hundred churchyards, pits, and cities, following the space of thousands of miles.»

It is known from the history of the craft of ancient states in Eastern Europe that in the 14th and especially in the 15th century, as before, before the Mongol invasion, carpentry was widely used in the construction and repair of roads and bridges. It is also known that in some cities of the 14th-15th centuries, not only wooden but also stone pavements were built. Furthermore, the organization of the so-called «pits» (yam), the pit service, or the pit chase, i.e., an organized communication system in which carts are prepared in advance at certain points (pit stations) awaiting the demand for transportation carried out from pit to pit with regularity. Already in 1493, the Yam (postal) service was on the road from Novgorod to Pskov. In the same '90s, there is information about potholes on the road from Moscow to Murom, as well as from Moscow to the Lithuanian border.

The runs between stations were 100 miles or more. Pit duty for the population has always been difficult in all forms. As is usually the case at the beginning, it was not strictly stipulated how many horses, fodder, carts, this or that messenger had the right to demand. Much earlier, the Tatars, who faced this circumstance, reached the point that already in the 14th century on the roads there was officially legalized robbery by passing messengers and envoys, and as a result, the population simply scattered from the roads. Therefore, the further, the more precisely it is stipulated, exactly how many means of transportation this or that messenger has the right to demand. The oldest surviving letter of travel dates back to 1489 and is addressed to the coachmen of one of the important roads of the state: Moscow to Novgorod.

1.3.1 Land transport in Europe in the Late Middle Ages.

In the late Middle Ages, the main part of traffic still fell on maritime transport, especially since new vast areas entered the area of maritime trade. As a result, shipbuilding has faced new challenges. Land transport is also evolving, although not to the same extent as maritime transport. Among draft animals, horses and mules (in southern countries) became the most common; the latter were widely used both as riding and pack animals. The wide distribution of mules (in Spain at the end of the 16th century, according to contemporaries, there were about a million of them) is a feature of the late Middle Ages. Means of transportation are also developing: iron rims appear on wheels, carriages intended only for passengers, cargo vans with a capacity of up to 2 tons, etc., appear.

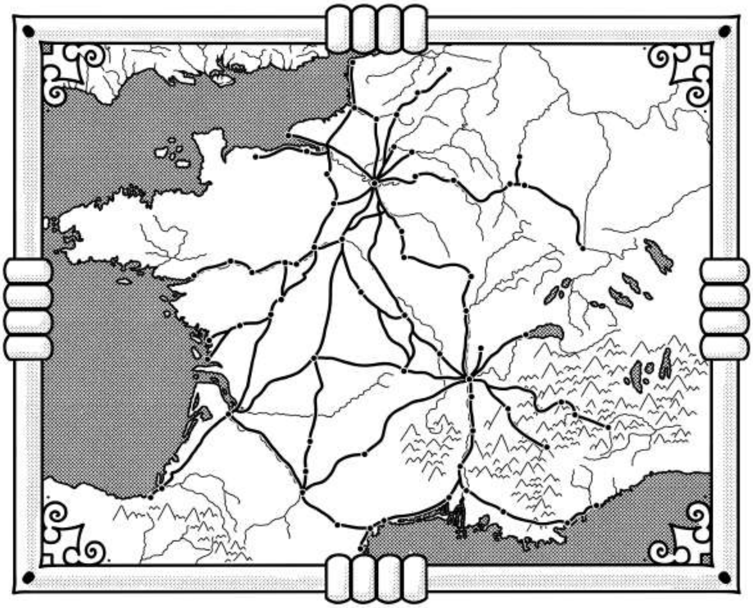


Figure 1.19: French roads around 1633

Compared to the fifteenth century there are no fundamental changes in the road network and road construction. However, control over the construction and condition of roads has increased and has become one of the important functions of the state administration; Thus, in France, under Colbert, a state road service was introduced, which was forcibly carried out by the unemployed and residents of workhouses. In the 17th century the construction of canals is also activated; if earlier in this regard, the Netherlands and Northern Italy stood out, now France joins them. As before, during the construction of the canals, two goals were pursued — draining the land and creating a transport route. Thus, the numerous canals drawn in Fennland (North-East England), in addition to draining the swamps, also provided access to the sea for the wool of eight surrounding counties. At the same time, often in the construction of canals they had in mind only their transport purpose. These are the French canals Saint-Omer — Calais and the Great Languedoc Canal du Midi (Canal royal en Languedoc), built with the latest technology of the time: over 240 km, 119 locks were built here, raising the ship to a height of about 200 m.

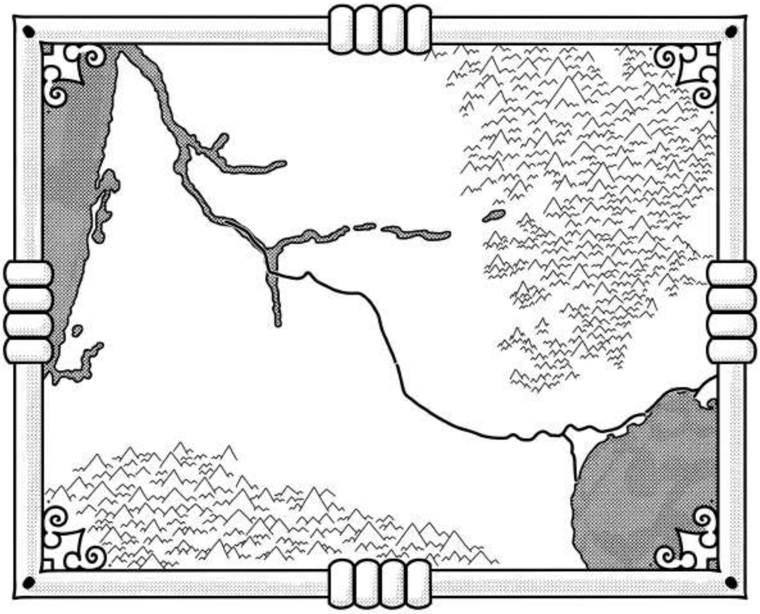


Figure 1.20: Canal du Midi

As already mentioned, the role of land transport in trade between the two regions in the 16th century gradually increased. The road network expanded more and more, with traffic on the roads becoming relatively regular and intense. For instance, in Paris, 11 tracts were created, connecting the capital with the largest cities of the country and abroad: through Orleans with Bordeaux and Toulouse, through Lyon to Avignon; roads to Saint-Malo and Rouen, Ypres, Brussels, to Cologne (via Reims and Liege), Geneva (via Dijon), Strasbourg and Frankfurt am Main (the first through Toul, the second through Verdun). In the middle of the XVII century, Paris already had a regular (at least once a week) passenger service with 43 French cities. Movement speeds have also increased. So, the couriers of Philip II reached Brussels in two weeks, the Venetian courier service created by Gabriele Taxis delivered letters from Italy to Brussels at an average speed of 130—140 km per day, that is, in 5—6 days. At sea, in some cases, speeds of 150—200 km per day were recorded. From this point of view, the timing of the news of various events is interesting.

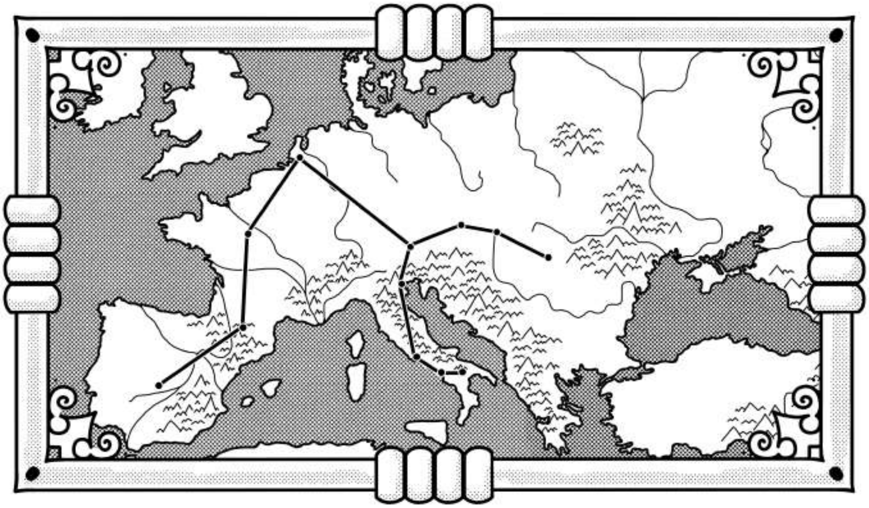


Figure 1.21: Taxis Post (1505-1516)

When first started, the Taxis postal system was for the royal court and government use only. During the early 1500s, in the documents that survive, the members of government are constantly chastised for using the post for private communication and for allowing their friends to use the post for private gain. But by the late 1500s such admonitions do not exist. We do not know when the change occurred, but we can surmise why: the growing business establishments of Europe needed a communications network. By the beginning of the seventeenth century, the post of Thurn and Taxis could be used by anyone who could pay the price.

The news of the Battle of Lepanto (off the southwestern coast of the Balkan Peninsula), which took place on October 7, 1571, arrived in Venice on October 18, in Naples on October 25, to Paris and Madrid on October 31, that is, three and a half weeks later. The message about St. Bartholomew's night went to Madrid at an average speed of 100 km per day. However, all these are extraordinary cases: the usual voyage speed of transport was still quite low. For example, merchants crossed the Mediterranean along the Tunis-Livorno line in an average of 10—11 days.

In Eastern Europe, the Moscow Principality united the lands of North-Eastern Russia around itself at the end of the 15th century. A new name for the unified state was chosen — “Russia”. The growth of the territory of Russia continued into the 16th century. By the end of the 16th century, the Volga, Urals, and

Western Siberia were included in Russia. In connection with the growth of the territory, roads in Russia have acquired special importance; on them, messengers from all the outskirts of the state delivered to Moscow news of the invasions of foreign troops, rebellions and crop failures. The central government showed particular concern for the development of the Yam post service, inherited from the Mongol Empire. During the reign of Ivan III, the first surviving travel document issued to Yuri Grek and Kulka Oksentiev, who was sent "to the Germans," dates back. In it, the sovereign ordered the ambassadors to be given "two carts to carts from pit to pit according to this letter of mine" at the entire distance from Moscow to Novgorod. In another letter of Ivan III — dated June 6, 1481 — the position of an official responsible for the condition of postal stations and roads — a Yam bailiff, was first mentioned. The pits were located at a distance of 30—50 versts. Coachmen were obliged to provide horses for all travelers with a princely letter, for their service they were exempted from tax — the sovereign tax and all duties — and, moreover, received maintenance in money and oats. Local peasants had to keep the roads in good condition under the supervision of coachmen. Under Ivan the Terrible, in 1555, a single body for managing the road business was created — the Yam hut. Already at the beginning of the XVI century. the first descriptions of large Russian roads appeared — "Russian road builder", "Perm" and "Yugorsky" road network descriptions. By the end of the XVI century, there were also "Yam books" with a description of small regional roads.

Land roads were part of the country's general transport network and were connected to waterways, being a necessary link in many transport routes. Often, land routes passed in winter along the bed of freezing rivers. Many roads ran along river communities. Finally, there was practically no land route that did not cross several rivers. At the same time, land roads still had a certain independence.

1.3.2 Road construction in Europe of XVII–XVIII centuries

In the countries of Western Europe, the resumption of road construction at first followed the path of imitation of the constructions of Roman roads. However, the economic conditions were changed. It was no longer possible to use incredible cheap slave labor for road construction, as in Ancient Rome, and there were a need to replace it with the labor of only the local population. These workers involved in road work in the order of compulsory road service or for a fee, forced to facilitate the construction of pavements on main roads, leaving local roads are practically without any improvement and maintenance.

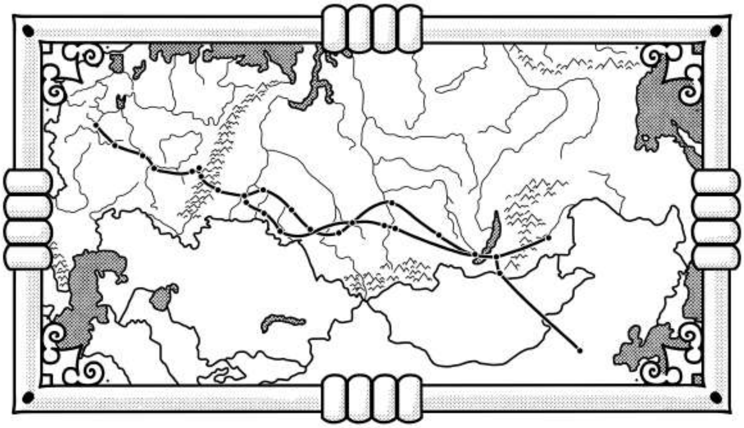


Figure 1.22: Siberian route from 1730

In the Russian state, the attention of the state authorities to the state and content of the means of communication is also increasing. This interest in the "road business" was caused both by the state need to facilitate the movement of service people and messengers, and by concern for the development of trade. Repair of roads, bridges and gates since the end of the 16th century, was assigned to the population according to the sokha (old Russian unit of land measure) division under the supervision of special leaders. A number of articles of the Council Code of 1649 placed responsibility for the safety and maintenance of land and river routes on the feudal lords. In order to eliminate obstacles to navigation, it was forbidden to build dams, and mills on many rivers.

Growing ties with Siberia, and, as, well, the signed in 1689 the Treaty if Nerchinsk, the first Russian—Chinese treaty, required the improvement of the Siberian routes. From the 1720s, the construction of the main Siberian road — the Siberian route, also known as the Great highway began. Outposts, winter huts and stations are organized on it. The government had widely involved peasant settlers for the construction and operation of this tract. Initially, it was used only for government couriers and government cargo. The decree on the construction of the longest Siberian tract in the world was issued by the Senate on March 16, 1733 in connection with the start of the second Kamchatka expedition. Initially, it was required to establish a regular, once a month, postal communication of the East Siberian cities with Tobolsk and further with St. Petersburg. For several decades, this tract was inhabited, mastered and finally formed by the 80s of 18th century.

The first attempts to improve roads in UK were described in a treatise by Thomas Procter published in London in 1607 "A Vvorthy Vvorke Profitable to this Whole Kingdome. Concerning the Mending of All High-waies, as Also for Waters and Iron Workes". The author noted: *«And heere withal is to be remembred that one principal and chiefe cause of al bad and foule waies is, that the raine water or other water doth lye and rest vpon ye highwaies (not order|lye and soundlye made) which with the working of cart wheeles & others, dooth peirce downe more deeper into the saide waies, and so more and more doth soften and rot the same, as by dayly experience is seene and knowne»* [235]. To prevent this, it was proposed to tear off a ditch 3 feet (0.9 m) deep and 4 feet (1.2 m) wide on the side of the road, distributing the excavated earth along the width of the road with an average thickness of one yard (0.91 m), and in 2 feet higher in the middle than at the edges. At the same time, the width of the road should be sufficient for the passage of two wagons. In the case of weak soils on the road, it was proposed to arrange clothes made of gravel, stone, slag, iron ore, stumps of wood or bundles of brushwood, laid in wooden frames made of logs 18 feet long and 10—14 inches in circumference, fastened together with wooden dowels. From above, this base should be covered with a layer of gravel, coarse sand or crushed stone.

There were also other pavement designs created by different architects. Construction technology has changed with almost every subsequent generation, both in connection with the accumulation of experience and changes in the requirements. At first, it was believed that artificial compaction by tamping was less effective than compaction for two to three months, but at the end of the century this opinion changed. For example, Russian engineer Gergardt wrote that roads should never suddenly rise above 4 inches, and must be filled firmly. This work must be repeated with each new filling of earth. Charles-François Exchaquet in 1787 recommended building gravel paths at least 10 inches thick in a compacted state, laying the gravel in two layers. The gravel should be *«the size of a walnut and not less than a bean, not polluted or dusty»* [202].

By the end of the XVIII century, when laying the route on the ground, some

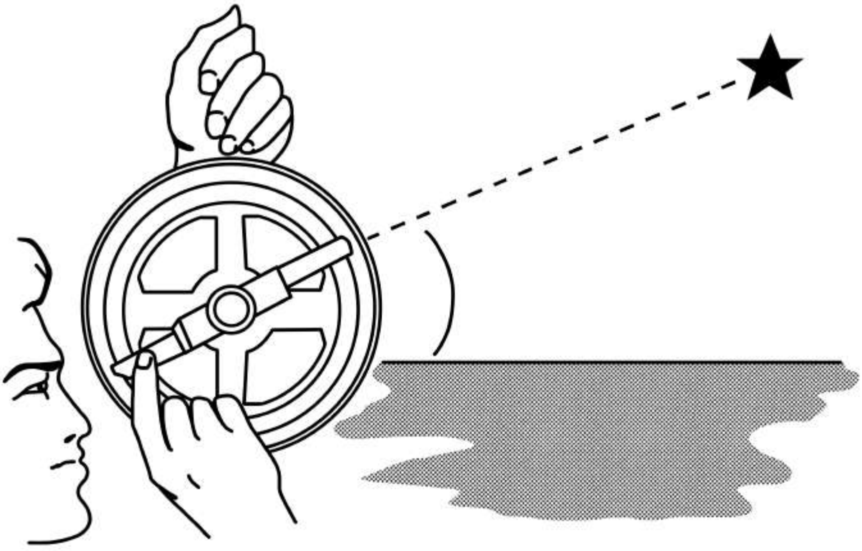


Figure 1.23: Measurement of an angle using astrolabe

geodetic tools began to come into use. The astrolabe with a compass appeared in the middle of the 16th century, the bubble level was invented by Melchisédech Thévenot in 1661. Based on it, a "y" level was proposed in 1680. In laying the tracks the inclinometers were used.

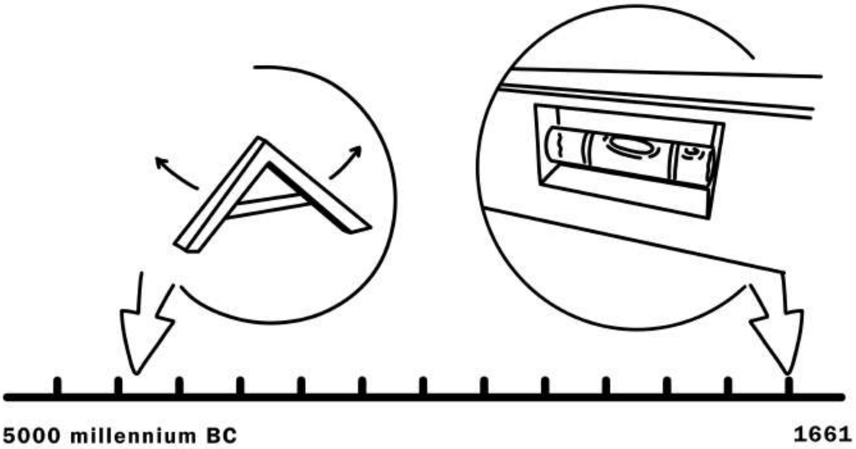


Figure 1.24: Level evolution

18th century characterized by attempts to accurately account for the properties of soils in construction. This was noted by Michael Lomonosov in a letter

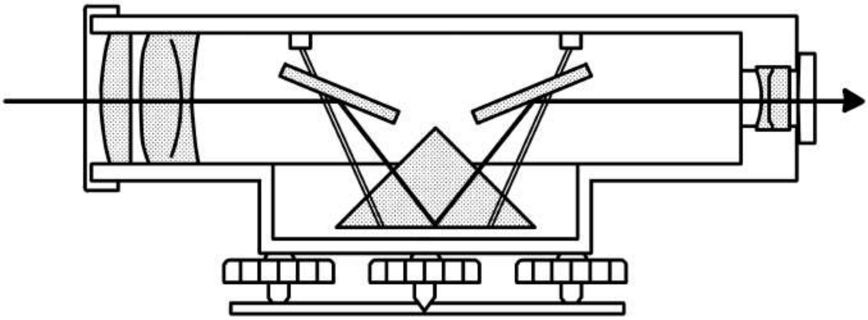


Figure 1.25: Automatic level

written in 1757—1759. The treatise was “On the layers of the earth”, indicating that *«the builder heeds the hardness of the earth in the ditches for the foundation»* [172]. He classified soils according to their composition and properties, dividing them into black earth, “clay of various kinds”, “silt or mud akin to clays”. The size of soil particles was as well taken into account in this classification.

During the period under consideration, pavements began to spread, their design almost did not differ from modern ones. There were certain requirements for their quality. The chipped cobblestone had to measure 7 to 8 inches and tapered downwards in a wedge shape. The ligation of the seams was required, *«so that in the longitudinal direction there are no matching seams that the wheels of the wagons could push apart.»* Layers of sand 6 to 8 inches thick were laid in the base, preferably fluvial and gravelly, and not quarry, which is very dusty. In the book of Christian von Lüder [174], it was indicated that when paving on both sides of the road, large stones are placed in the ground, and then smaller and smaller ones are invested. To increase the strength of the pavement, Hubert Gautier [105] proposed to arrange transverse rows of larger cobblestones 10—13 inches high through two toises (1.82 m) so that if the pavement begins to collapse, damage does not extend beyond this row. Cage paving has become ubiquitous. At the end of the 18th century, when the pace of road construction began to increase, the most widespread was pavement based on a package — stones installed with a wide side on a soil or sandy base and wedged, which later began to be replaced with rubble *“the size of a walnut made of hard rock”*, which was distributed with a layer of 8 cm. However, the pavement on packing bases did not meet the requirements of mechanized construction, and operating experience showed that they could not withstand the movement of heavy vehicles, the multiple passes of which were concentrated on a narrow rolling strip and caused longitudinal subsidence of the coatings.

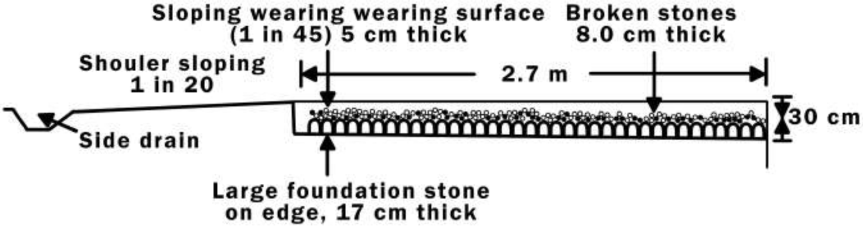


Figure 1.26: Road pavement of Tré

Pierre Trésaguet significantly reduced the thickness of the pavement, to about 24—27 cm compared with the thickness of the pavements of previously built roads, which reached 50 cm along the axis. The pavement was arranged in a trough dug in the subgrade with a convex bottom. With cohesive soils, this contributed to a partial runoff of water seeping through the pavement, and also made it possible to give the pavement a constant thickness along the entire width of the carriageway. Equally important was the convexity of the bottom of the trough for a more economical use of stone material. The bottom layer (base) of the pavement, 10 inches thick, was arranged from stones placed on the edge at the bottom of the trough, so that not one stone towered over the other. The stones were rammed with a manual rammer. On top of them, a layer 8—10 cm thick of smaller stones was laid, which were crushed on the spot and compacted by tamping. Partially penetrating into the gaps between the stones, they wedged large stones. 10 cm of gravel was laid on top.

The roads were designed for horse-drawn transport, and only at the end of the 19th century, when cars appeared, did the construction of roads using binders begin. The first of the extensive works on the use of asphalt were carried out in Paris in 1832—1835. In parallel with asphalt concrete structures, cement concrete pavements appeared in cities. The first road with cement concrete pavement was built in 1865 in England. A few years later (1876), the French workers built a concrete road in Grenoble. By the end of the XIX century, cement concrete roads appeared in many countries of Western Europe and in the USA. Asphalt and cement concrete pavements favorably differed from their predecessors in terms of technical and economic indicators. In addition, the technology of their device opened up a wide path for the mechanization of road construction work.

Before the First World War, the need for asphalt was met mainly by natural material. However, the development of cars, as well as an increase in the level of improvement of cities, caused a sharp increase in the production of binder material and a decrease in its cost. All this led to the creation and rapid expansion of the industrial production of petroleum bitumen.

1.4 "Modern" school of road construction. Road pavements

A new branch of the development of road traffic is connected with the development of capitalism. The modern school of road construction began in 1747, when the National School of Bridges and Highways was established in Paris to train engineers. The road construction methods taught at the school were based on the technologies of ancient Rome and were further improved by the engineer Pierre Trésaguet (1716—1796), who helped Napoleon I to build the main road network in France. He was the first to establish the correct principles of road construction, which have been preserved to a certain extent to the present day. The characteristic features of the road design proposed by Trésaguet are as follows:

- the convex profile of the base provides better drainage and working conditions for the structure as a whole;
- reducing the thickness of the road surface to a size justified by its strength;
- correct distribution of material by size (lower layer — package, upper layers — monotonous crushed stone and crushed stone);
- construction of a strong lateral support (curb);
- proper sealing;
- systematic repair.

The next stage in the development of road construction technology is the transition of pavement made from crushed stone, to the so-called "gravel highway", which is usually associated with the name of the Scottish road builder McAdam. The McAdam method became widespread because it was simple, cheap, and up to date. Starting in 1806 to take contracts for road construction, McAdam developed his own system for the construction and repair of roads and, having taken over in 1816 the trust of the Bristol district, the largest in England, he began to vigorously introduce this system into practice. The methods of road maintenance proposed by him turned out to be very effective and economical. The Scottish engineer John McAdam (1756—1836), who developed the principles of Trésaguet, created a school of road technicians and promoted his experiments. The highway is still called "macadam" by veteran road builders. McAdam proposed instead of a base of large stones, which significantly increased the cost of construction, to maintain a natural soil base in a dry state by creating a waterproof coating and providing drainage. In 1811, he issued an instruction on the reconstruction of old roads.

The essence of McAdam's ideas, scattered throughout the book, boils down to the following:

1. The strength of the road is determined by the soil base. Until then, it will not be possible to build roads that are not affected by seasonal and weather factors: *«until the following principles are fully realized, recognized and put into practice, namely, that the load from traffic is actually taken up by natural soil … this natural soil must be dried beforehand»* [182].

2. The role of pavement is reduced mainly to the protection of the underlying soil from soaking. *«Experience having shewn, that if water passes through a road, and fill the native soil, the road, whatever may be its thickness, loses its support, and goes to pieces»* [182]. McAdam believed that for any load, a thickness of 10 inches in a dense pavement is sufficient.

3. Pavement should rise above the ground and not be laid in a trough open in it. *«The first operation in making a road should be the reverse of digging a trench. The road should not be sunk below, but rather raised above, the ordinary level of the adjacent ground, care should at any rate be taken, that there be a sufficient fall to take off the water, so that it should always be some inches below the level of the ground upon which the road is intended to be placed: this must be done, either by making drains to lower ground, or if that be not practicable, from the nature of the country, then the soil upon which the road is proposed to be laid, must be raised by addition, so as to be some inches above the level of the water»* [182].

4. Pavement should be flat, cohesive and waterproof.

5. One-dimensional clean crushed stone or gravel should be used for pavement. *«The size of stone used on a road must be in due proportion to the space occupied by a wheel of ordinary dimensions on a smooth level surface, this point of contact will be found to be, longitudinally about an inch, and every piece of stone put into a road, which exceeds an inch in any of its dimensions, is mischievous»* [182].

6. The strength of the crushed stone bark is ensured by the mutual locking of the crushed stone. Therefore, pavement should be made of pure rubble. *«Every road is to be made of broken stone without mixture of earth, clay, chalk, or any*

other matter that will imbibe water, and be affected with frost; nothing is to be laid on the clean stone on pretence of binding; broken stone will combine by its own angles into a smooth solid surface that cannot be affected by vicissitudes of weather, or displaced by the action of wheels, which will pass over it without a jolt, and consequently without injury» [182]. *«The only way to avoid the movement of stones in the road is to use stones of the same size in it to the very bottom»* [182].

7. During the period of compaction by the movement of stone material behind the road, enhanced maintenance is required. *«After gravel is laid on the road, daily hired workers for backfilling the ruts and at the same time removing with a rake from the surface of stones that are too soft or irregular in shape, such as long flints or too large»* [182]. When compacting crushed stone placers, mainly by the passage of vehicles, McAdam noted that *«for the first sedimentation of gravel, a heavy iron roller with a diameter of* 4 *to* 5 *feet (*1.2—1.5 *m) at least can be successfully used»* [182].

8. The cross slope of the road should not be too steep. *«The road is then to be laid as flat as possible, a rise of three inches from the centre to the side is sufficient for a road thirty feet wide»* [182]. *«If the road is made flat, riders will not stick to only its middle, as is done with excessive convexity»* [182]. As a result of repeated attempts to improve the passage on the roads with a scattering of new materials, thick layers of riprap were formed on them. These layers were dismantled and replaced with crushed stone coverings, for which a large stone removed from the road was crushed away from the road. Therefore, the rebuilding of roads according to the McAdam method, which did not require a new stone, replaced the laborious and more expensive rebuilding of roads with the methods used by Thomas Telford. The work scope performed was limited to the necessary minimum, and therefore McAdam emphasized that *«on each road I was forced to change the way of work depending on local conditions, and often on funding»* [182].

These principles are uniform, however much circumstances may differ, and they must form the guide by which his judgment must be always directed [182].

1.4.1 Roll it!

The idea of crushing the clods by means of a heavy traction cylinder moved by an animal is very old, since such an agricultural implement is already mentioned by Cato the Elder and by Virgil. This agricultural use is maintained

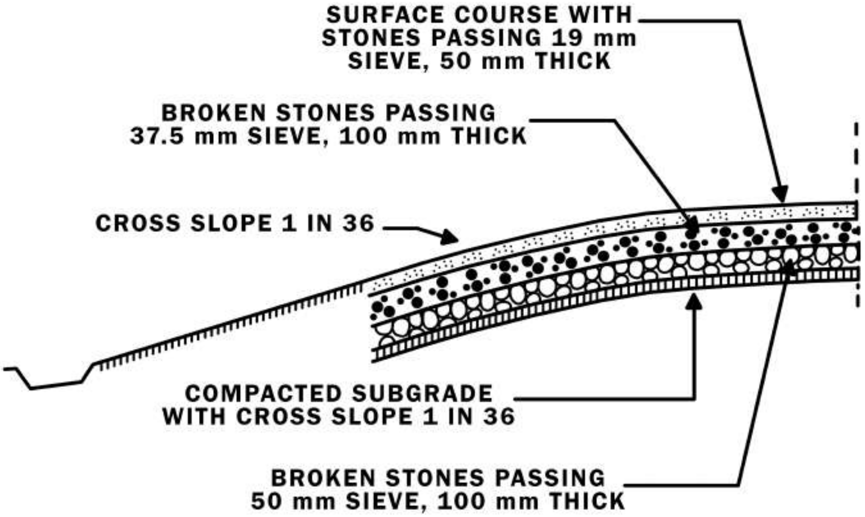


Figure 1.27: A typical cross-section of Macadam's construction

thereafter, mentioned in particular in 1690 by Furetiere; but, already in 1619, an Englishman, John Shotbolte, took out a patent for the application of the roller to roads; then in 1725, it was the turn of the German Jacob Leutpold to propose an iron one in his Theatrum Machinarium. In 1789, endorsing a well-established practice, the French Academy reported, under the term roller, "certain cylinder-shaped stones which gardeners use to level the paths in the gardens". The French engineer Louis-Alexandre de Cessart (1719—1806) was probably inspired by it when he submitted a proposal for a steamroller for roads to the Assemblee des Ponts et Chaussees on February 5, 1787. This engineer, best known for his hydraulic works, recommends a cast iron cylinder 8 feet long by 36 inches in diameter, weighing about 3500 kg. De Cessart explains: *«Certainly, by passing twenty times in succession over the gravel, each square foot would have been pressed by a rather considerable weight, and this pressure would undoubtedly be communicated throughout the entire thickness of the gravel and could form a kind of crust such as one sees on the old roads which resist the strongest cars»* [44].

But the artificial compaction of the road, in spite of its many advantages, has some difficulty in imposing itself, both in England (where it was recommended in 1822 by Paterson), and in France. This resistance is due on the one hand to the additional cost that the rolling operation entails for the project manager, on the other hand to the fact that it then remains to develop not only machines, but also effective working methods. Throughout the 19th century, a great deal of research led to decisive progress in this field. This idea of the steamroller, an idea which was then in the air, reappeared around 1820—1830 in Germany, particularly in the region of Hanover, with variable load rollers whose weight could be increased as the building progressed. settling. These machines seem quite commonly used since they were even prescribed in 1823 by the Prussian road administration.

However, the Germanic engineers practiced hardly any experimental research in this field: this driving role fell to France, where, in 1830, Lemoyne recommended the use of a cylinder of approximately one meter. Essential contributions are made by the engineer Antonin-Remy Polonceau (1778—1847). He recommends, in 1829 and especially in 1844, the use of cylinders with a large diameter (200 cm) filled with water to increase the weight (6000 kg) and (unlike McAdam who wanted very pure gravel), he proposes to spread sludge or aggregation materials on the surface to be tamped. His works and essays are widely reflected, in particular by the Annales des Ponts et Chaussees and the Allgemeine Bauzeitung. In addition, many other trials then took place in France, such as in Indre-et-Loire, in the Jura and in Maine-et-Loire. At the same time, Charles-Henry Schattenmann, director of the Bouxwiller mines (Bas-Rhin department), a polygrapher interested in various economic questions, published a brochure advocating the use of a type of steamroller derived from the one he seen in use in "Rhenish Prussia", particularly in Sarrebrück: this medium-sized machine could be gradually loaded with sand or pebbles, thanks to the boxes arranged on its chassis.

However, Germanic engineers hardly practiced Schattenmann's experiments in Alsace first of all, around 1840—1842 (which he himself describes in the most optimistic way), then in 1844 in the French capital, are widely discussed in specialized journals. In fact, the results did not seem to have been spectacular in Paris, on the Cours-la-Reine and the avenue Gabriel: during several successive tests, the Schattenmann roller, harnessed to 8 horses, had in vain passed until 42 times on the same points. The compaction was bad; the contractor attributed this phenomenon sometimes to excessive drought, sometimes to excess moisture in the aggregate materials. Despite these mediocre results, the principle of artificial settling gained ground: from the years 1839—1840, rolling was recommended to future engineers, whether they were students of the Ecole des Ponts et Chaussees or the Ecole centrale of Arts and Manufactures of Paris. On the other hand, the «Zeitschrift über das gesammte Bauwesen», published in Zürich between 1836 and 1840, the only specialist journal at that time in Switzerland, made no mention of it.

In March 1842, in the Swiss canton of Vaud, the Public Works Commission

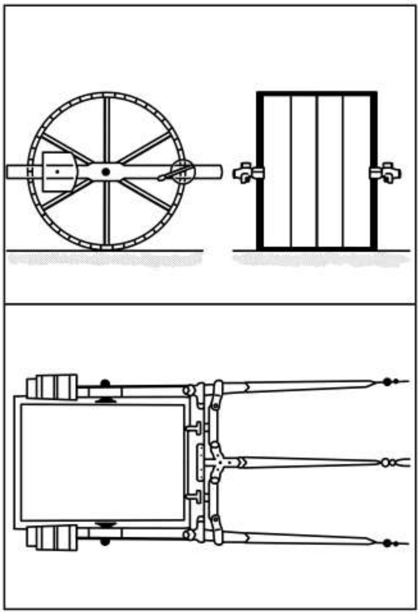


Figure 1.28: Compressor cylinder of Polonceau System, 1841

wrote to Geneva to obtain information on a steamroller arrived to the city in December of the previous year. This machine, Systeme Polonceau, came out of the Gandillot & Roy foundry in Besangon and cost 3033 francs. But the Vaud investigation was premature and it was only through a note written in May 1843 that we were informed of the tests that took place on the Carouge road. This huge roll, 150 cm wide but two meters in diameter, weighs approximately 3700 kg. To this weight must be added almost 1200 liters of water, which the cylinder is filled with. This one, thus loaded, was set in motion by four horses, on a dry and hard ground, where one had laid out a layer of broken stones of approximately 6 to 9 cm thickness, of an extent of 500 m.

The progress on this rocky surface causes vibrations which tend to unscrew the bolts of the side walls, affecting the tightness of the assembly. A companion must therefore frequently tighten these nuts. Furthermore, the width and the weight of the cylinder make it difficult to change direction. Arrived at the end of the section under construction, it is necessary to unhitch the horses, reverse the stretcher, then hitch in the other direction. Having wanted to avoid this maneuver and turn at all costs with the horses, the Geneva roadmenders irremediably damaged the edges of their cylinder from the first tests and it was emptied of its water. From now on the experimentation of the machine must therefore be done empty, but it is said to observe in spite of everything a good settling after ten to twelve passages. In fact, this first Geneva roll will be quickly abandoned... In order to be able to compare these results with those of the "Prussian roll", the Vaud administration asks the engineer Fraisse for a report. It is obviously difficult for him to judge these machines without having seen them operate, on the sole basis of the Geneva description and the dithyrambic brochure written by Schattenmann on the roller manufactured by the Dietrich company of Reichshoffen in Alsace. This last engine consists of a hollow iron cylinder, 4 feet wide and high, weighing, empty, 60 quintals. It is surmounted by a box which is loaded with stones up to a total weight of 120 quintals. Six horses are needed to drag the heavy machine and the harness is harnessed, as with the Polonceau System, sometimes forwards, sometimes backwards. According to the information available to Fraisse, the Geneva roller does not seem to have provided a surface as smooth as that described by Schattenmann; moreover, the possibility of progressively loading the "Prussian roll" leads Fraisse to recommend rather the acquisition of an Alsatian machine.

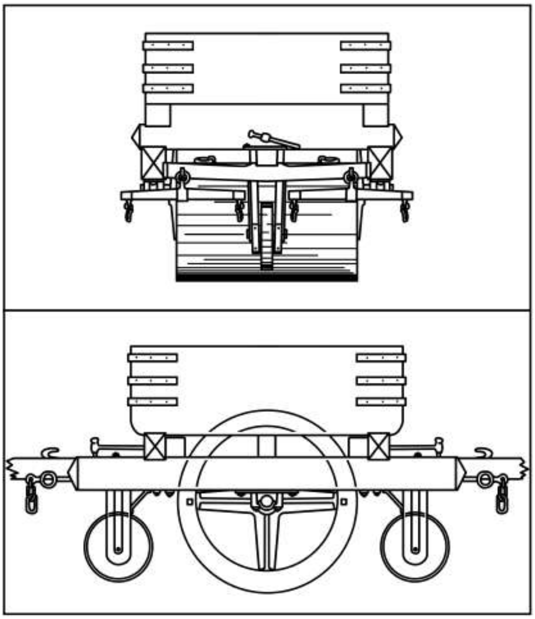


Figure 1.29: Compressor cylinder of Schattenmann System, 1842

In France, during the decade from 1840 to 1850, people were above all concerned with practical trials, notably at Autun, Orleans [230], Mäcon, and Nantes, and these experiments were published in the Annales des Ponts et Chaussees, which bears witness to the the interest they arouse. From the middle of the 19th century, certain engineers approached the question in a more rational way and sought to transcribe the operations of cylinders into mathematical formulas, taking into account the compression ratio, the nature and quantity of materials to be used, the the most favorable period and maintenance expenses. With the help of experience, one can better perceive the faults of the various rollers, generally very heavy, requiring considerable attachments, therefore very cumbersome and difficult to handle, particularly in tortuous passages. Various technical improvements will be proposed. Appear then, towards the middle of the century, the roll known as Bouilliant, easier to transport because mounted on a train of four wheels (invention of Regnault and Bouilliant in Paris), the Houyau roll, manufactured in Nantes, soon equipped with a rotating ring or a rail surrounding the whole machine, rail on which the attachment point of the hitch moves, so as to finally be able to change the direction without having to unhitch. Germany, in this field, presented as a novelty, in 1854, the cast iron roller filled with water (like the one designed by Polonceau a quarter of a century earlier), and an annular mechanism for turning the coupling was proposed by the manufacturer Ketzer of Chemnitz in 1866. In England, Amies & Bradford patented a similar system around 1868. In 1860, Louis Lemoine, a young engineer attached to the city of Bordeaux, had Daney build a first steam roller, weighing 10 tons.

This revolutionary system will know in France a development remarkable with the Ballaison rollers soon improved and built in Paris by Gellerat & Cie. These tandem rollers, with double cylinders, have the advantage of walking equally in one direction and in the other, since it suffices to reverse the steam; their only inconvenience, it was said, was the fright they can cause in horses. This fear led in 1861 the prefect of police of Paris to authorize their use only at night; but following some modifications to the apparatus, this defense was brought back three years later. England, until then, had only very moderately practiced rolling. After a first steam roller patented in 1864 by Clark and Batho, of Birmingham, Albion, however, will gain considerable ascendancy with the Aveling & Porter rollers, tricycle rollers of which one wheel, at Pavant, is the director. This company from Rochester, in 1867, was noticed by a huge steamroller of 30 tons destined for Liverpool, then by another of 20 tons ordered by Philadelphia. It was now conquering a world market, but in small steps at first. In the great adventure of steam, in fact, two concepts confront each other: the road locomotive, which implies, to be truly effective, a notable improvement in the roadways, and the railroad. If we know the success of this last means of transport during the second half of the 19th century, we often forget that this development led to a kind of stagnation for the roads.

Road locomotives have in fact had little success, except for military use, and it is not until the very end of the 19th century that the first automobile cars appear. But the political authorities had to take into account, previously, the proliferation of cyclists. Indeed, since the invention, in 1888, by John Boyd Dunlop, of the first inflatable tire for bicycles, this means of locomotion, which has become much lighter, more manageable and more comfortable, has enjoyed exceptional success. The result was a considerable revival of public interest in roads, hence a real desire to invest. In this context, animal traction rollers continue to have a certain success, and comparative studies between the two systems are multiplying. In spite of its many advantages, in fact, steam rolling only became very progressive. It is reported in 1879 that these modern rolls are still almost unknown in Germany where only the cities of Berlin and Stuttgart have them. In Switzerland, Winterthur, the first, was equipped in 1875 with an Aveling & Porter. It must be said that in our country this new technology did not seem to meet with great interest at the time: the Revue Polytechnique — Schweizerische Bauzeitung only mentioned the steam roller in 1883, and we had to wait until 1890 to find a first evocation of trials, in Bad Ems near Wiesbaden. However, following the example of Bäle and Zürich, the Council of State of Bern also bought, in 1897 (on a proposal, from 1894, of its Management Commission), a steam roller: a machine of 13 tons supplied by Fowler & Cie in Magdeburg. Contrary to what is practiced elsewhere, the Bernese roadmenders were a very thick layer of refill: 15—20 cm on the edges, 25—30 cm in the middle of the roadway. To obtain sufficient compaction of the gravel, the roller must pass 150 to 200 times over the same section!

However, with the advent of rubber tires, the condition of rolled roads worsened, because they did not roll the rubble, but, on the contrary, "pulled it out". This contributed to the emergence of "tarmak" — a coating, which included a new binder in the form of natural tar. In Switzerland, in 1721 near the city of Neuchâtel and in 1810 near the city of Seysel, deposits of asphalt rocks were discovered — limestone and sandstone impregnated with bitumen. They began to be developed for the preparation of mastic for waterproofing work. Soon it was noticed that pieces of asphalt rock that fell on the road during transportation form a solid homogeneous layer when compacted in transit. This led to the idea of building asphalt pavements. In 1829, a footpath was built in Seysel, and in the 30s the first attempts to build asphalt pavements were made.

In United States, asphalt was first used to pave streets in the 1870s. At first naturally occurring "bituminous rock" was used, such as at Ritchie Mines in Macfarlan in Ritchie County, West Virginia from 1852 to 1873. In 1876, asphalt-based paving was used to pave Pennsylvania Avenue in Washington DC, in time for the celebration of the national centennial.

Later, they began to use local asphalt rocks containing a higher percentage of bitumen, adding to them, in addition to sand, stone flour. The inventor of modern road bitumen is Professor Edward Joseph de Smedt. In 1870, a Belgian immigrant, while working at Columbia University (New York), received a patent for a new road surface he invented, which became known as French asphalt.

Already existing means of rolling and asphalt material led to the fact that the streets of rammed and rolled asphalt began to spread on the streets of large cities. In Paris in 1854 there were 800 m, in 1856 — 8 km, and in 1860 — already 230 km. In London, the first asphalt pavement appeared in 1869, in Berlin — in 1877. As the starting point of the systematic construction of improved pavements, we can consider the rapidly becoming popular laying on the streets of capital cities of pavements with “rammed asphalt” — crushed stone from natural asphalt rocks, which was heated in boilers and compacted by tamping after leveling on a solid stone base. In 1913, for the first time in Europe, rolling of the “asphalt mass” borrowed from the USA was used. The coatings are called “rolled asphalt”.

1.5 Automotive era

At the end of the XIX century, an event occurred that caused a revolutionary change in transport technology — the appearance of the automobile, a self-moving cart with an internal combustion engine. In 1885—1886 German engineer Benz installed a gasoline engine on a three-wheeled wagon, and in 1887 Daimler started serial production of cars. By 1895, the Paris-Rouen car race took place in France, with an average speed 24 km/h.

Before the invention of automobiles, limited requirements were imposed on the pavement of the roads, arising from the peculiarities of horse traction on the slopes. A horse can, working with a short-term overload, develop a traction force on the hook that is 2—3 times higher than normal, which is about 1/3 of its weight. Therefore, the steeper the climb, the shorter its length should be. The construction of roads, intended mainly for cars, made this requirement inappropriate, but gave rise to a number of requirements of a different nature.

As the number of cars increased and their dynamic qualities increased, the requirements for taking into account the peculiarities of their movement in the norms for the design of the plan and the longitudinal profile of roads increased. With the increase in traffic and the proliferation of automobiles, the problem of controlling dust on the highway has arisen. In this regard, in the early years of the XX century experiments on tarmacking were carried out.

Bad roads were not an obstacle to motorization. However, the development of automobile production in different countries set different tasks for their designers and road service. In Western Europe, where there was already a dense network of roads with solid road pavements, the task arose to take into account the requirements of high-speed vehicle traffic when maintaining roads. In countries with a sparse road network and a predominance of dirt roads — in Russia and the USA — the problem arose of ensuring travel on roads and adapting cars to the condition of these roads.

The first direction led to the development of techniques for the construction of improved pavements, the second — to the emergence of methods of mechanized construction of dirt roads as a temporary way to pass traffic of low intensity.

The mass production of automobiles gave impetus to road construction. In the USA, it unfolded in the 20s, when the number of cars exceeded 10.5 million, and roads with solid pavements accounted for only 12% of their total length.

An idea arose — to look for different ways to preserve the properties of soil during periods of waterlogging, by strengthening it — so-called "soil stabilization". Stabilization was supposed to prevent the harmful effects of increasing soil moisture, destroying cohesion in it. To increase adhesion, water-resistant binders of organic and inorganic origin or skeletal additives were introduced into the soil which increase its internal friction, such as coarse sand, gravel or crushed stone. The first successes in soil stabilization during the period of low traffic volumes led to the emergence of the slogan "soil not as a foundation, but as pavements." It was proposed "to take soil as the main material for pavement — the ubiquitous soil, but not to take it in its natural or mechanically improved state, but by turning it into a mass, elastic and strong enough for passage through various physico-chemical and technological processes and influences."

The conflict between motor transport and roads of the period of predominant horse traffic was short-lived and was an incentive for further progress in road construction technology — the mass-appearance of improved pavements based on organic binders.

Rapid progress in expanding the use of organic binders in road construction is associated with the name of the Swiss physician Ernest Guglielminetti (1862—1943). For 12 years, starting in 1902, Ernest Guglielminetti successfully used heated coal tar from a gas plant to combat dust on a 20-kilometer stretch of the Nice-Monte Carlo road with almost 1, 000 cars and a large number of horse-drawn carts. Tar was used for the annual restoration of the surface treatment of the crushed stone coating, followed by backfilling with sand.

The use of surface treatment devices showed that they not only lead to dust removal of pavements, but also significantly reduced their wear. As a result of repeated surface treatments, a kind of mat is formed on the roads — a thin layer of asphalt pavement. Initially, bottling of bitumen and tar was carried out manually from watering cans, followed by its distribution over the coating with brushes. Then boilers with a capacity of 250—350 liters appeared on carts, from which the binder flowed through holes in a horizontal tube. The cart was carried by two workers. The next step was horse-drawn distributors. They had containers for 1200—1500 liters, from which bitumen was supplied under pressure up to 8 atm and, what was important, it became possible to regulate the amount of binder supplied.

The positive effect of bitumen and tar on the strength of pavements caused the gradual emergence of new structures over the years. This was due both to the development of scientific research, and mainly to the improvement of manufactured road cars. Schematically, this process can be described as follows.

Surface treatment captured only the top layer of the pavement. The crushed stone located below was held only by the forces of the wedge and therefore, when the surface treatment was worn, the rapidly progressive destruction of the road surface resumed. The task of binding crushed stone to a greater thickness was solved by the advent of the impregnation method.

The desire to create a more uniform material for the surface layer gave rise to the idea of making pavements from stone material pre-treated with a binder. The method originated from the simplest drying of crushed stone on iron sheets and shifting with tar by hand.

Gradually, it became clear that with the appropriate selection of the mineralogical composition of stone material and bitumen or tar of low viscosity, the storage period of the “ripened” mixture can be significantly extended, and it does not stick. This made it possible to harvest processed crushed stone for future use, manufacture it at factories, including in winter, and transport it to construction over long distances, reducing the process of building a coating only to distribution over a prepared base and rolling.

One of the most famous patents for apparatus for the preparation of tarmacadam was granted in 1904 for Carlton Reid [244]. Here is the description of the proposed process: when a road is repaired, the top surface is removed down to the exposed native soil that is compacted with a heavy mechanical roller. A layer of crushed concrete is deposited onto the native soil and compacted with the roller to form the foundation of the road. A base layer of crushed rock mixed with a waste product from steel making known as slag is then laid down. A binder known as tarmacadam forms the surface layer of the road. The surface layer consists of a mixture of black sticky hydrocarbons known as asphalt, bitumen or pitch containing gravel that is coated with the hydrocarbons, normally when the asphalt is heated between 90 and 160℃. The hot mixture is then poured onto the base layer and the roller applied to produce a compacted top layer.

Edgar Purnell Hooley invented the composition for the top coat that used hot tar as the binder. He referred to a process for making ’tarmacadam’ that was abbreviated to Tarmac for a trademark. A preferred tarring mixture described by Hooley consists of the following: Tar (92.56% by weight), pitch (5.79% by weight), Portland cement (0.41% by weight), and resin (1.24% by weight). The tar is placed in a mixer and gradually raised to a temperature of about 212°F. The other ingredients are added and are thoroughly mixed with the tar.

This asphalt was prepared from stone material in the form of blast-furnace slag crushed to sand size and a minimum amount of liquid bitumen. For the construction of crushed stone coatings by the impregnation method and a surface treatment device using hot viscous bitumen, it was imperative that the crushed stone be in a dry state and that there was no dust on the crushed stones.

The next stage in the development of technology for the construction of improved road surfaces was the use of bituminous and tar emulsions and liquefied bitumens. Bituminous and tar emulsions, consisting of approximately 50% water, included 2% emulsifier and dispersed bitumen or tar. They made it possible to perform work at lower positive temperatures and wet rubble. Disintegrating upon contact with the surface of the stone material, they left a bituminous film adhering to it.

Binders thinned with volatile solvents also more easily penetrated into the spaces between the gravel, binding the dust to the surface of the stone particles. Liquid bitumen was especially widely used in the construction of gravel coatings on roads of lower categories by the displacement method on the road, since gravel materials containing a large percentage of dusty and sandy particles could be moved on the roadbed with a small number of passes of a grader or disc harrow with only low-viscosity material. Such methods were widely used in the USA and the USSR before the Second World War when creating a grassroots network of highways in the conditions of a rapid increase in motorization.

The increase in traffic intensity and the appearance of heavy vehicles on the roads required a further increase in the strength of road surfaces in comparison with crushed stone pavements treated with binders. In road construction, asphalt and cement concrete began to become widespread. Asphalt concrete arose as a development of crushed stone pavements from materials treated with binders in special installations. The fundamental difference between asphalt concrete and crushed stone treated with a binder was the mandatory presence in its composition of a fine mineral powder with a particle size of less than 0.1 mm. At the first stage of designing asphalt concrete compositions, it was assigned the role of filling the pores between sand particles, from which its original name “aggregate” was born, later replaced by the term “mineral powder”. Depending on the ratio of mineral powder and binder, the coating turned out to be too brittle or too ductile, especially in hot weather, when it left traces from the wheels and shifted when braking. Road surfaces with a low binder content quickly collapsed.

The development of motorization in countries with a dense road network with a predominance of personal cars, a sharp increase in passenger traffic and the widespread use of autotourism made it necessary to impose the same high architectural and aesthetic requirements on roads as on any engineering structures for mass use. By the beginning of the Second World War, a new direction had emerged in road design, combining landscape design, paving, and spatial smoothness. The development of high-speed automobile traffic has shown the importance of smoothly fitting the route into the landscape and for ensuring high transport and operational qualities of roads.

A century later, road equipment for asphalt laying has gone far ahead in its development. Modern asphalt pavers automatically feed, distribute, place materials and level the road.

The greatest attention is now paid to cement-concrete roads. Such surfaces are superior to asphalt in terms of strength, wear resistance and durability, and also have a number of operational and environmental benefits. High transport and operational qualities and the possibility of full mechanization of construction work put forward concrete road surfaces in first place in the world among improved coatings.

Transportation infrastructure acts as an essential component of civil infrastructure, providing vast mobility and economic benefits to the modern society. With the expanding pavement network and growing traffic, the pavement system has been deteriorating and the demand for maintenance, rehabilitation, and replacement (MR&R) activities has become ever greater and more complicated [198]. In most cases, the decision makers need to figure out which pavement sections need to be treated, when to take the MR&R actions and which maintenance treatments are appropriate. Effective MR&R strategies can greatly improve the pavement performance after maintenance, extend the service life, and maximize the social and economic benefits of the roadway infrastructure. As Figure **??** illustrates [102], Galehouse et al. (2003) found that if $1 is spent on “the right treatment on the right pavement at the right time,” it can eliminate or delay spending $6 to $10 on rehabilitation or reconstruction later.

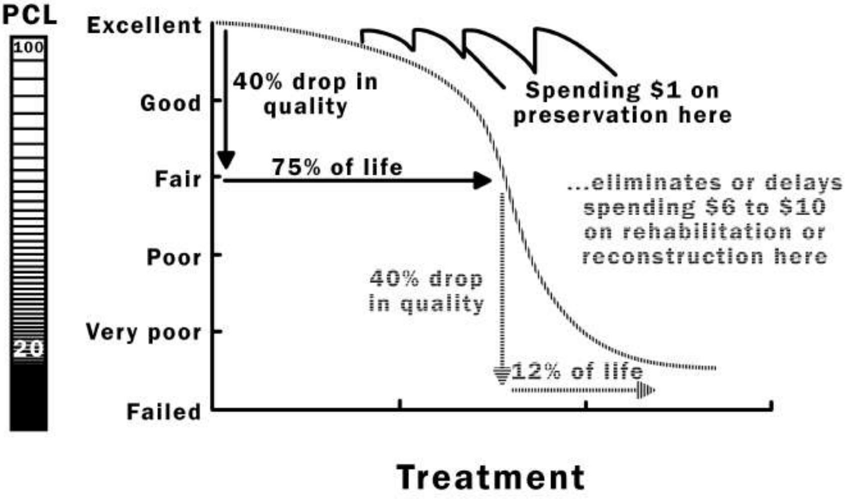


Figure 1.30: Relationship between pavement condition when treatments are applied and the treatment cost [102]

In addition to repair costs, the condition of a road surface has economic effects for road users. Rolling resistance increases on rough pavement, as does wear and tear of vehicle components. It has been estimated that poor road surfaces cost the average US driver $324 per year in vehicle repairs, or a total of $67 billion. Also, it has been estimated that small improvements in road surface conditions can decrease fuel consumption between 1.8 and 4.7% [202].

At the beginning of the 20th century, thanks to the development of transport and the increased growth of cities, the problem of improving road surfaces, organizing developments related to the management and improvement of transport networks, including the use of logical and mathematical apparatus, became acute. From this moment, the history of mathematical modeling of traffic flows begins.

The brief history of road transport

> “The car has reached complete perfection — no further improvement is required.”
>
> —Journal of the Allgemeine Automobile Zeitung (1921)

The automobile, as the realization of the human dream of self-propelled carriages for transporting people and goods, has been created and intensely improved for over two hundred years. The automobile is one of the greatest inventions of humanity, which has had and continues to have significant importance for humans. Nowadays, the significance of the automobile is felt not only in the transportation industry but also in all spheres of human life. It has become a tangible embodiment of technological progress and has transformed the face of the planet. The history of the creation of the automobile is astonishing in its fascination and unpredictability.

The predecessor of the gasoline-powered automobile was the steam carriage, more precisely a steam-powered wagon, which was crafted by the French inventor Nicolas Joseph Cugnot in 1769. This cumbersome machine could move at a speed of 2—4 km/h and carry up to three tons of cargo. To cover long distances, it was necessary to make stops every 15 minutes to reignite the furnace, as the pressure in the boiler dropped rapidly. Furthermore, the machine was poorly maneuverable and often collided with houses and fences. However, the term "automobile" began to be used precisely with the emergence of vehicles with internal combustion engines.

How did humanity progress from steam carriages to autonomous vehicles of the future? That is precisely what we will explore in this chapter.

2.1 From steam and explosion to usefulness

Attempts to design devices that could perform manual labor for a person — for example, to move something — were started a long time ago.

One of the first names worth mentioning is that of Heron of the city of Alexandria. **Heron of Alexandria** was a Greek mathematician and engineer. He is believed to have lived during the second half of the first century AD, due to the fact that he cites the lunar eclipse of March 13, 62 AD, as an example [214].

Figure 2.1: Heron of Alexandria

Heron was a prime example of the main achievements of the ancient world in the field of applied mechanics. He famously invented and described many forms of devices, with a particular example being an apparatus for measuring the length of roads, which operated on the same principle as modern taximeters. He described the apparatus of Dioptra, the great-great-grandfather of the modern theodolite. Heron first explored five types of the simplest machines: the lever, the gate, the wedge, the screw, and the block. He also laid the foundations of automation. In the work “Pneumatics,” Heron of Alexandria described a system of “magic tricks” based on the principles of using heat and pressure. People were surprised by his miracles: the doors of the temple opened themselves when a fire was lit on the altar. The scientist had invented an automatic machine for selling "holy" water, designed a ball that could be rotated by the force of a jet of steam, and invented many other systems of devices, along with automatic machines.

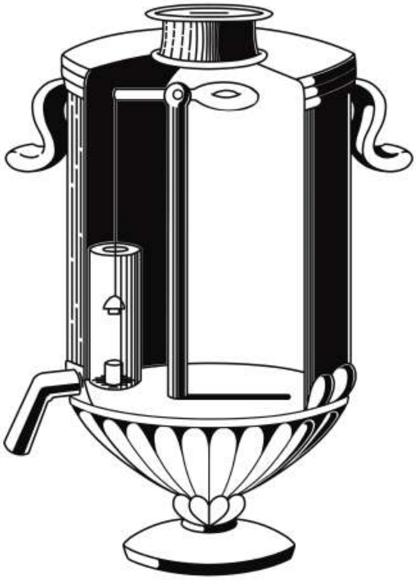


Figure 2.2: Vending Holy water machine

His major works include Metrics, Pneumatics, Automatopoetics, Mechanics (this particular work is preserved solely in Arabic), Katoptrika (the science of mirrors; survived only in the Latin translation), etc. In 1814, Heron's composition "On the Diopter" was found, which set out the rules of land surveying, based on the use of rectangular coordinates. Heron used the achievements of his predecessors in his work: Euclid, Archimedes, Straton of Lampsacus. Many of his books are irretrievably lost (such as the scrolls contained in the Alexandrian Library). A copy of his books, made in the 16th century, is contained at Oxford University.

Heron of Alexandria invented the first working steam engine and called it the "wind ball." Its design is extremely simple. A wide lead water boiler was placed over a heat source, such as burning coal. As the water boiled in two pipes, a ball in the center revolved, with steam rising. Steam jets flowed through two holes in the ball, causing it to rotate at great speed. The same principle underlies modern jet propulsion.

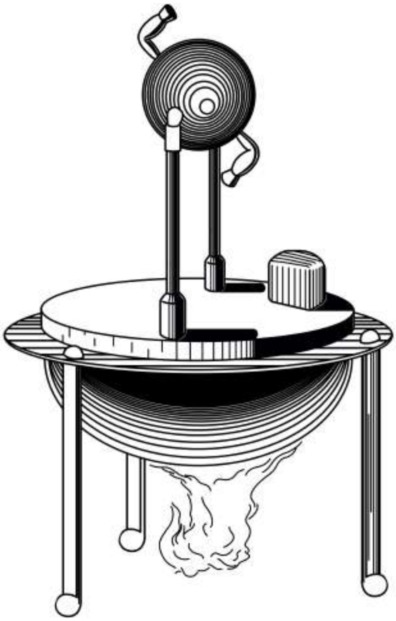


Figure 2.3: Wind ball: the first working steam engine

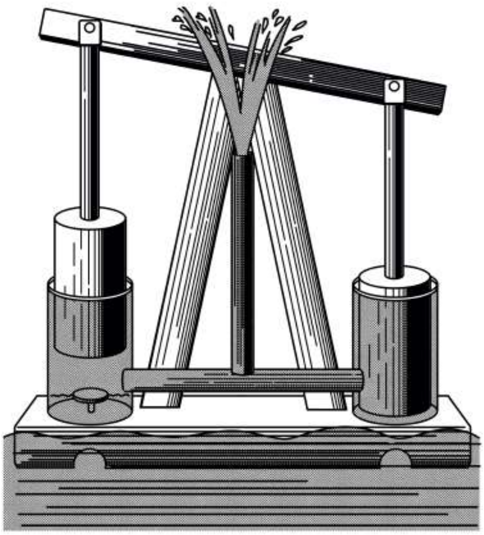


Figure 2.4: Force pump of Heron

Could his steam engine be used in practical purposes? To find the answer to this question, a specialist in Antiquity Dr. J.G. Landels of the University of Reading with the help of specialists from the engineering faculty made an accurate working model of Heron's device. He found that this model reached a high speed of rotation — at least 1500 revolutions per minute: *"The ball of Heron's device may have been the fastest-rotating object of his time."* [158].

Heron could have invented a more efficient way to use steam energy. As Landels noted, all the necessary elements for an efficient steam engine were found in the devices described by this ancient engineer. His contemporaries created cylinders and pistons with extremely high efficiency, which Heron used in the construction of a water pump. This "force pump" was widely used in the Roman world, and one of its applications was in a fire engine.

It would not have been difficult for Heron or any of his contemporaries to combine all these elements (a boiler, valves, a piston, and a cylinder) to make a functional steam engine. It has even been claimed that Heron went further in his experiments, assembling the necessary elements into an efficient steam engine, but either died during the testing phase or abandoned the idea entirely. None of these assumptions has been substantiated. Most likely, due to his many commitments, he was unable to implement the idea.

Unfortunately, Heron's invention was later forgotten. And it remained forgotten for a long time — for one and a half thousand years. Let's now return to a more recent period.

Between the late 17th and early 18th centuries, people again became interested in how to perform useful work and convert that work into movement — for example, to move heavy loads or to move themselves.

To begin with, it is worth mentioning three important names: the Frenchman Papin and the Englishmen Savery and Newcomen. Papin, though French, lived most of his life in England. He knew both Savery and Newcomen, and also worked with Hooke. **Denis Papin** was born in the French city of Blois in 1647. At the University of Angers, he studied medicine and received his doctorate but did not practice as a doctor.

Papin's career in our area of interest began after he moved to Paris, where he became an assistant to Christian Huygens, for whom he designed an air pump and other devices [247].

Figure 2.5: Denis Papin

Huygens advised his student to try using the gases produced by the combustion of gunpowder. Hot gases occupy a large space, and when cooled, they significantly decrease in volume. Heeding this advice, the young scientist built a powder engine, consisting of an iron cylinder and a piston. Papin calculated that if he ignited a pinch of gunpowder under the piston lowered to the bottom, the piston would jump up, and the space underneath would be occupied by hot gases. When they cooled down and decreased in volume, a rarefied space — a certain vacuum — would form in the cylinder, and atmospheric pressure would force the piston to lower.

Papin calculated everything correctly, but in the first experiment, the piston flew out of the cylinder despite the presence of a valve meant to hold it. Furthermore, even when the piston did not fly out, the gas formed during the explosion did not cool sufficiently, resulting in the piston only falling slightly and the machine performing minimal useful work. Papin then decided to use steam to lift the piston. Steam had suitable properties: it expanded significantly and occupied a large volume.

After leaving France for religious reasons, Papin settled in London in 1675, where he became acquainted with Robert Boyle and other members of the Royal Society. Five years later, he presented to the Society the so-called "Papin's Cauldron" — a closed vessel filled with water, beneath which was a furnace. The water boiled, and the resulting steam began to press against the walls of the vessel. To prevent the vessel from bursting, Papin installed a safety valve in the upper lid — a plug pressed down by a lever with a weight. When the pressure became too great, the steam lifted the valve and escaped; the pressure dropped, and the valve closed under the weight. Aiming to derive practical benefits from the invention, Papin turned his attention to the culinary arts. By placing the mechanism inside a small cauldron, he introduced the world's first pressure cooker to the Royal Society. "Thanks to this device, the toughest beef will become as tender and delicious as veal," the advertisement claimed.

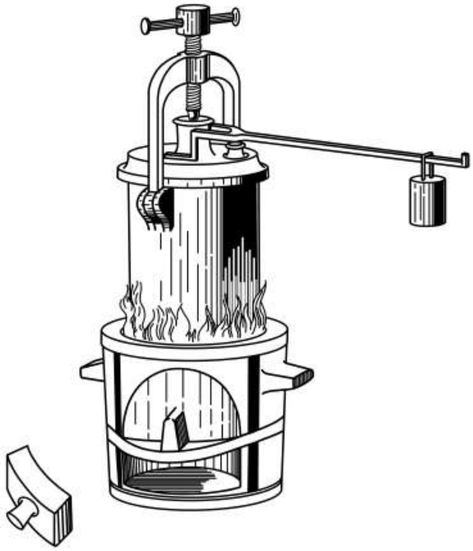


Figure 2.6: Papin's Cauldron

Later, fate took him to Venice, then back to England, and finally to Germany. Here, at the University of Marburg, he continued working on the heat engine. In his new engine, Papin used water instead of gunpowder. Water was poured under the piston, and the cylinder was heated from below. The steam lifted the piston. Then the cylinder was cooled, and all the steam inside condensed back into water. The piston fell under its own weight and atmospheric pressure. This engine worked better than the powder engine.

For practical work, it was still of little use: it was necessary to supply and remove the fire, supply cooling water, and wait for the steam to condense. These disadvantages were mainly due to the fact that the formation of the steam necessary for the operation of the engine took place in the cylinder itself. Therefore, the following idea seemed natural: what if ready-made steam, cooked, for example, in a separate boiler, were admitted into the cylinder? Then it would be enough to alternately admit steam and cooling water into the cylinder, and the engine would run at a higher speed with less fuel consumption.

The last surviving evidence of Papin's whereabouts came in a letter he wrote dated January 23, 1712. At the time, he was destitute ("I am in a sad case"), and it was believed that he died that year and was buried in an unmarked grave in London. It looks like he died in deep poverty, forgotten by everyone. Until recently, nobody even knew where his grave was.

However, a record exists for the burial of a "Denys Papin" in an 18th-century Register of Marriages & Burials, which originally came from St Bride's Church, Fleet Street, London, but is stored in the London Metropolitan Archives. The record states that "Denys Papin" was buried at St Bride's on August 26, 1713 — just a few days after his 66th birthday — and that he was buried in the Lower Ground, one of the two burial areas belonging to the church at the time. Since the discovery in 2016 of the place and date of Papin's burial in 1713, a memorial plaque has been erected in the West Entrance of St Bride's Church, Fleet Street, London, to commemorate his life and his achievements.

In general, the idea behind the engine was the same as that of Heron: we boil water in a boiler, steam is generated in large quantities, and steam under pressure presses on some turbine blades, causing work to be done.

One more name to mention is that of **Thomas Savery** (1650—1715), an English engineer and inventor [33]. He was a captain — this was the name for production managers, engineers, technicians, and foremen in those days. It is not known for certain whether he owned mines or worked as a mining engineer in Cornwall. He had a favorite pastime: he made watches and polished mirrors using a machine of his own invention. Savery was a smart and versatile person. His first invention was a rowing boat. He sent the draft to the king himself. The king approved the project and sent it to the Admiralty. Savery's long walks around the offices did not lead to anything — no one needed a boat with a rowing device. "We need something different, sir," one of the officials told the inventor. "We need a pump, a water-draining machine. Take care of this!"

In July 1698, Savery received a patent for an invention "used to raise water

Figure 2.7: Thomas Savery

and to set in motion all kinds of mills using fire, which will be of great use in pumping mines, supplying cities with water, and running mills where there is no opportunity to use water or wind."

Savery's machine was a steam pump, not an engine: it did not have a cylinder with a piston that, when moved, would set something in motion. The most important feature of this device was that the steam needed to operate the pump was prepared in a separate boiler. Steam was generated in the boiler, which was continuously heated. When a valve was opened on the pipe, steam was let into the pumping tank. Two pipes led away from the tank: one (suction) went down into the mines, and the other (discharge) went to the drainage gutter.

When steam was admitted into the tank, it began to expel the water present through the discharge pipe into the drain gutter. Then, the steam supply was stopped, and cold water was admitted through a special tube into the tank. The steam condensed, turned into water, occupying a small volume, and a vacuum was formed in the tank, i.e., reduced pressure. Water from the mine, displaced by atmospheric pressure, rushed through the suction pipe into the "empty" tank. Valves on the suction and discharge pipes allowed water to pass only from the shaft to the tank and from the tank to the gutter.

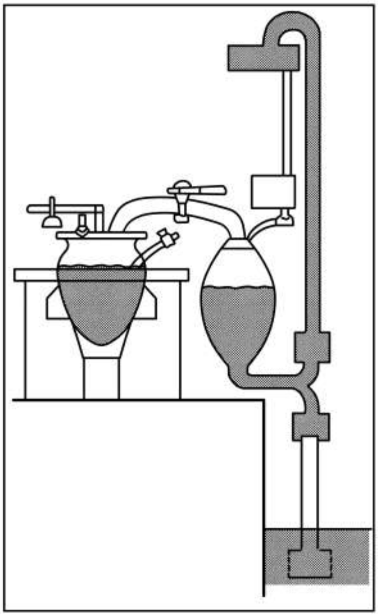


Figure 2.8: Savery's pump. 1 — bulb; 2 — boiler; 3, 6, 7, 12 — tubes; 4, 10, 11 — valves; 5, 9 — taps; 8 — tank

It was already a steam pump quite suitable for practical purposes, used to pump water only from shallow mines and mines, because atmospheric pressure cannot raise water more than ten meters. Peter the Great purchased one such pump in England in 1717 to use it for large drainage works during the construction of canals in St. Petersburg. It was the first steam engine to appear in Russia. However, it turned out to be quite weak, so Peter ordered it to be put in the Summer Garden to pump water into a tank, from where it was supplied to the fountains. Two more such pumps were installed by the merchant Trusov in his baths on the Fontanka.

The Savery pump had serious drawbacks: it was low-power, consumed a lot of fuel during operation, and worked intermittently — water was pumped out in separate portions. It could not be used as a universal motor to drive various machines and mechanisms, as most of them work continuously. Nevertheless, the Savery installation helped inventors grasp the simple idea that steam engines should use ready-made steam from a separate boiler.

Thomas Newcomen was born on February 28, 1663, in Dartmouth, and worked as a blacksmith and mechanic [53]. He had the most practical engine, perhaps because he was a very practical man, having started as a blacksmith.

Figure 2.9: Thomas Newcomen

In 1705, together with the glazier-tinker John Calley, also from Dartmouth, they built the first steam (steam-atmospheric) machine, which was distinguished by the fact that the steam was condensed by pouring water on the outside of the cylinder. In 1711, the technology was changed to injecting water inside the cylinder, which significantly accelerated the machine's movement. Among the problems they solved was the following: there is a mine that we sometimes extract something from, such as coal. It is periodically flooded with water, as the mine is located in England. How do you pump the water out of the mine? It was no longer modern to carry it in buckets, so a steam pump was invented. Steam presses on the piston, the piston rises, then this forward movement is transmitted to the pump. The piston rises to a certain level, where a valve opens, the pressure is released, cold water is pumped in, everything cools down, the piston goes back, steam raises it again, and this cyclic process repeats many times. The pump is running, the water is pumping. Everything is wonderful.

Newcomen could not obtain a patent for his invention, since the steam water elevator had already been patented by Savery in 1698, which secured any possibility of using water vapor; later they started working together.

The merit of Newcomen is that he was one of the first to implement the idea of using steam to obtain mechanical work.

Newcomen took a cylinder with a Papin piston as the basis of his machine, but the steam in his machine used for lifting the piston was obtained in a separate boiler, like Savery's.

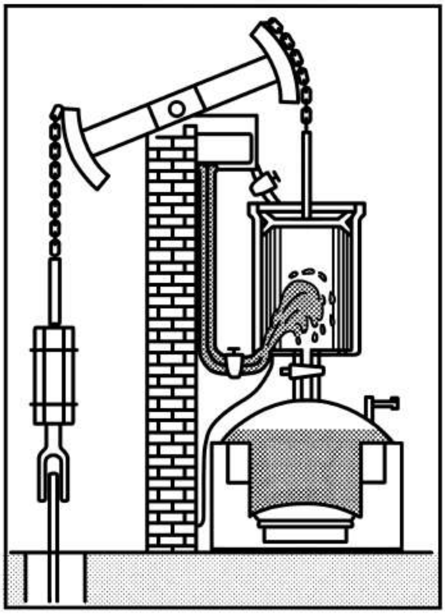


Figure 2.10: Newcomen's steam engine

The Newcomen machine worked as follows: Steam was generated in the steam boiler continuously. At a certain moment, the tap was opened, allowing steam under pressure to enter the cylinders, lifting the piston. The piston was connected to the rod of the water pump through a chain and a balancer (swing arm), which fell down during the upward stroke of the piston. Steam filled the entire cavity of the cylinder. Then, another tap was manually opened, through which cold water was injected into the cylinder. The steam would condense, and a vacuum would be created in the cylinder. Under the influence of atmospheric pressure, the piston would be lowered, pulling the balance chain with it. At the same time, the rod of the water pump went up, and the pump expelled the next portion of water. Then the process was repeated from the beginning.

Newcomen's machine worked intermittently and therefore could not drive industrial machines and mechanisms that required continuous movement for their operation. But that was not Newcomen's goal. He aimed to create a pump that could remove water from deep mines, and he succeeded.

The power of Newcomen's engine was 8 horsepower, which enabled the lifting of water from a depth of 80 meters and consumed 25 kg of coal per hour per horsepower. Newcomen began his experiments with a steam pump in 1705 and improved it for about ten years until it worked properly in 1712.

Newcomen's engine became widespread in coal and ore mines in England, as well as in France and Germany, mainly in the mining industry, though it was sometimes used to supply large cities with water. Due to its bulkiness and uneven stroke, it consumed a lot of fuel and therefore did not meet the needs of the industry, being used for highly specialized purposes, not reaching the level of a universal engine.

In the Russian Empire, work on the steam engine was carried out in the middle of the 18th century. In the city of Barnaul, based on Papin's work, an engineer named **Polzunov** decided to create a steam engine that would power bellows at a metallurgical enterprise without human intervention, replacing the water drive system that existed before. Polzunov created the first Russian engine: perhaps he owns the authorship of the idea, which was further used in all engines. He had two cylinders that worked on one shaft, assisting each other, making it a two-stroke apparatus. Unfortunately, he died of tuberculosis on May 16, 1766, just a week before his students and colleagues started the engine. The engine ran for 43 days, and during that time, it not only recouped all the costs of its creation but also brought in significant net profit, as calculated by the plant managers. By the way, the power of this engine was 40 horsepower, a substantial increase compared to the Papin and Newcomen engines, which had 5—8 horsepower.

We discussed the beginning of the invention and the first engines that had already begun to generate income. And we concluded with the fact that by the end of the 18th century, there were no such machines in widespread use until a man named James Watt appeared. Watt contributed more to the invention of engines, it is believed, than all his predecessors combined. What did Watt do? Watt created the steam engine, inventing many different nuances, details — a condenser, a parallelogram of forces. He used engines to move carts. He was the first to use steam engines for transporting goods and people.

James Watt (1736—1819) is a Scottish engineer, inventor-mechanic. He was member of the Edinburgh Royal Society, the Royal Society of London and the Paris Academy of Sciences. A unit of power is named after him — the Watt [63].

Beginning in 1763, he worked on improving Newcomen's ineffective steam-

Figure 2.11: James Watt

atmospheric machine, which was generally only suitable for pumping water. It was clear to him that the main drawback of the Newcomen machine was the alternating heating and cooling of the cylinder. How could he avoid this? Watt found the answer to this question on a sunny spring Sunday in 1765. He realized that a cylinder could remain permanently hot if the steam was drained into a separate tank through a pipeline with a valve prior to condensation. Moreover, the cylinder could stay hot and the condenser cold if the outside was covered with thermal insulation material. In addition, Watt made a number of improvements that finally turned the steam-atmospheric engine into a true steam engine. In 1768, he applied for a patent for his invention and received it, but he did not manage to build a steam engine for a long time. It was only in 1776 that Watt's steam engine was finally built and successfully passed all tests, being twice as efficient as Newcomen's machine.

In 1782, Watt created the first universal double-acting steam engine; he equipped the cylinder head with an oil seal, invented not long before, that allowed the piston rod to move freely but prevented steam from escaping the cylinder. Steam entered the cylinder alternately from one side of the piston, then from the other. As a result, the piston made both the working and the return stroke with the help of steam, which was not the case in previous machines. Since the piston rod was pulling and pushing in a double-acting steam

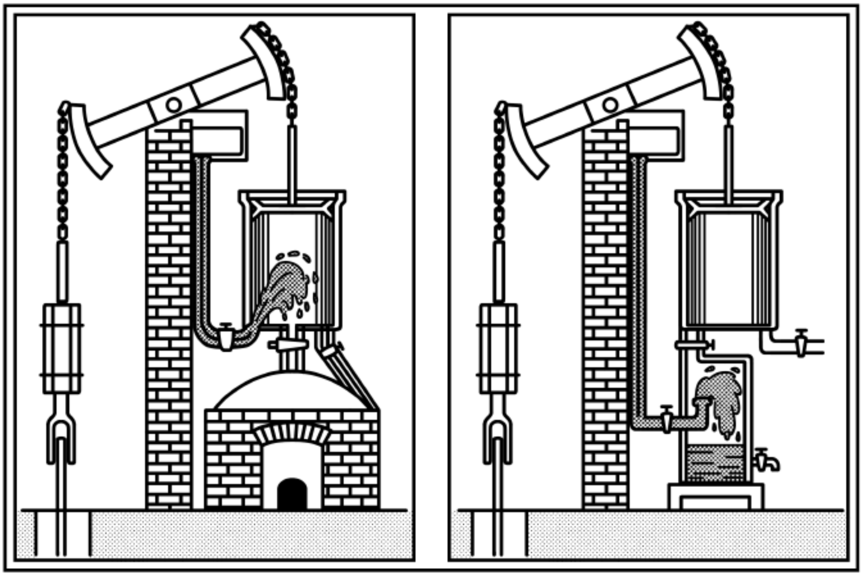


Figure 2.12: Watt's engine

engine, the old chain-and-rocker drive system, which only responded to pull, had to be redesigned.

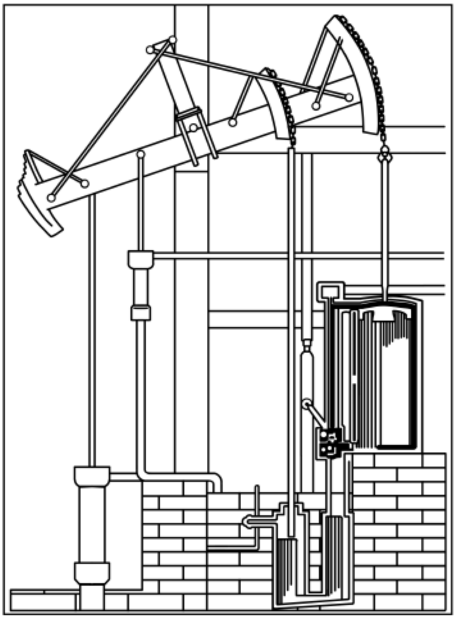


Figure 2.13: Watt's double-acting steam engine

Watt developed a coupled linkage system and used a planetary mechanism to convert the reciprocating motion of the piston rod into rotary motion, using a heavy flywheel, a centrifugal speed regulator, a disc valve, and a pressure gauge to measure vapor pressure.

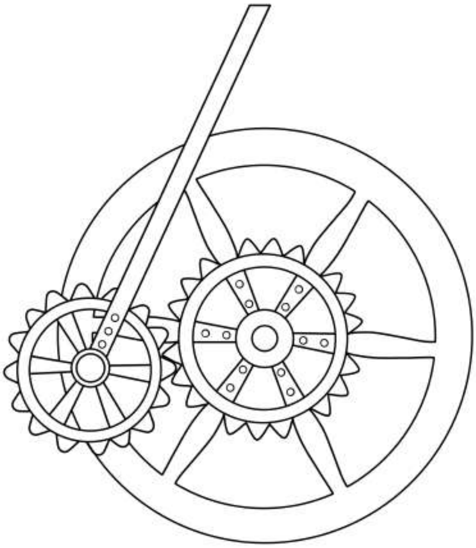


Figure 2.14: Solar and planetary gear mechanism that converts vertical motion into rotational

Watt's patented "rotary steam engine" was first widely used to power machines and lathes in spinning and weaving mills, and later in other industrial enterprises. Thus, Watt's steam engine became the invention of the century, marking the beginning of the industrial revolution. In 1785, one of the first Watt machines was installed in London at the Samuel Whitbread Brewery to grind malt. The machine did the job instead of 24 horses. Its cylinder diameter was 63 cm, the piston stroke was 1.83 m, and the flywheel diameter reached 4.27 m. The machine has been preserved to this day, and today it can be seen in action at the Powerhouse Museum in Sydney.

Science is something good; the inventions mentioned earlier bring us closer to the present day. But what about industry? Nowadays, specialized R&D departments handle the applications of these inventions. And who made James Watt's inventions both useful and profitable?

It was Matthew Boulton (1728—1809), an English industrialist and business

Figure 2.15: Matthew Boulton

partner of James Watt. In the last quarter of the 18th century, the partners produced hundreds of Boulton and Watt steam engines, a significant advancement for the time: their creation enabled the mechanization of factories [64]. Boulton introduced modern methods for minting coins, issuing millions of pieces for the UK and other countries, and supplied the Royal Mint with the most advanced equipment of the era.

Born in Birmingham, England, in 1728, Boulton was the son of a small metal fabricator who died when Boulton was 31. By then, Boulton had already established his own business for several years and later expanded it by consolidating all operations at the Soho Manufactory, which he built near Birmingham. In Soho, he introduced the latest technology into production, including the manufacture of silver plates, gilded bronze, and other decorative arts. He became associated with James Watt when Watt's business partner, John Roebuck, failed to pay off a debt to Boulton. Boulton lobbied in Parliament for an extension of Watt's patent for another 17 years, which allowed the company to market Watt's steam engine effectively. The company installed hundreds of Boulton & Watt steam engines in Britain and abroad, initially in mines and later in factories.

Boulton was a key member of the so-called Lunar Society, a group of

Figure 2.16: Lunar Society

Birmingham-area men distinguished in the arts, sciences, and theology. Its members included Watt, Erasmus Darwin, Josiah Wedgwood, and Joseph Priestley. The society met monthly on the night of the full moon. Its members are credited with developing concepts and methods in science, agriculture, manufacturing, mining, and transportation that laid the groundwork for the Industrial Revolution.

Boulton founded the Soho Mint, which soon harnessed the power of steam. He aimed to improve the dire state of Britain's coinage and, after several years of effort, was awarded a contract in 1797 to mint the first British copper coins in a quarter-century. His coins were well-designed and difficult to counterfeit; the first issue included large copper British pennies, which were minted until decimalization in 1971. He retired in 1800 but continued to oversee his mint until his death in 1809.

Watt was likely one of the first physicists to profit from his studies in physics. Incidentally, both Boulton and Watt are featured on English banknotes. Regarding physicists and profit, there is a popular internet quote from Thomas Edison: "Anything that won't sell, I don't want to invent. Its sale is proof of utility, and utility is success." Based on another popular story, Richard Feynman, known for his wit, also remarked that anyone can invent, but selling it is the challenge.

So perhaps Mr. Boulton and Mr. Watt are on the 50-pound note for a reason. In their partnership, Boulton was the entrepreneur who provided resources, and Watt was the intellectual force, and indeed, the engines from Boulton and Watt's company sold very well. But a completely amazing situation developed further — the engines worked, but there was no comprehensive theory behind them. How did they work? There was some basic theory: you heat the steam, it expands, presses on the turbine, the turbine turns, the shaft turns, and everything moves. You're right to an extent. But how do we design an engine to be more economical? How do we boil water more efficiently? Thus, humanity came to develop a theory of engines.

2.2 And what about thermodynamics?

Figure 2.17: Sadi Carnot

Many years ago, an ordinary boy named **Sadi Carnot** was born in France. However, when your father is one of the heads of state, you're probably not quite an ordinary boy.

Sadi Carnot received a good education and began to work with engines. But his life was not very successful. He wrote only one work in his entire life — it was really short — just one work, titled "Reflections on the Motive Power of Fire: A Critical Edition with the Surviving Scientific Manuscripts" [37].

In this work, Sadi Carnot formulated the basic principles of heat engines and their theory. In fact, he laid all the foundations of theoretical thermodynamics. This was applied to a heat engine and, amazingly, was based on incorrect premises.

Carnot's concept was fundamentally wrong because he relied on the theory of caloric — a kind of fluid that, during thermal contacts of bodies, could flow from one to another. He was unaware of atoms and molecules.

But his engine theory was correct. He had some kind of absolutely brilliant physical intuition. He sensed what was right and what was wrong. He understood that bodies with different temperatures were needed; otherwise, there would be complete equilibrium. So Carnot introduced the concepts of a heater and a refrigerator in an engine.

Carnot was one of those who died during a terrible cholera epidemic in Paris. The second tragedy was that the only article he published went unnoticed.

Why do we know and talk about him now? This is another amazing story. It was noticed by **Benoit Clapeyron**, who apparently did not know Carnot. He saw Carnot's work and republished it, mentioning Carnot's name everywhere, perhaps to the detriment of his own scientific interests. That turned out to be one of Clapeyron's best works, as he added several of his own contributions to Carnot's.

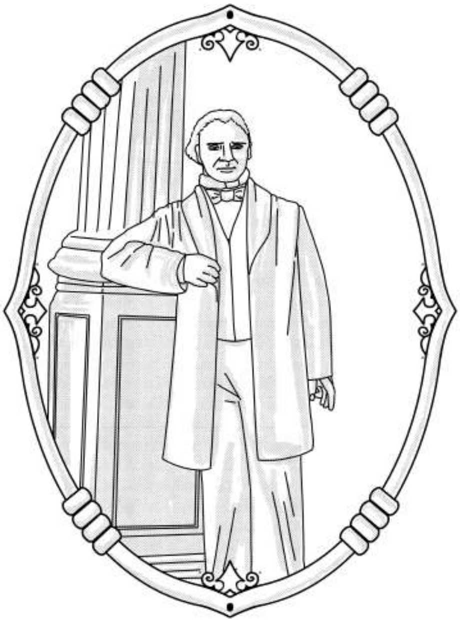


Figure 2.18: Benoit Clapeyron

Physical studies by Clapeyron [297] are devoted to heat, plasticity, and equilibrium of solids. In 1834, he gave a mathematical form to Carnot's ideas, being the first to recognize the great scientific significance of his work, which contains the formulation of the second law of thermodynamics. Based on these ideas, he introduced a graphical method into thermodynamics — indicator diagrams.

In 1834, he derived the equation of state for an ideal gas, combining the Boyle-Mariotte law, the Gay-Lussac law, and the Avogadro law, generalized in 1874 by Dmitry Mendeleev (Mendeleev-Clapeyron equation). He deduced an equation establishing the relationship between the melting and boiling points of a substance and pressure, which was thermodynamically substantiated in 1851 by Clausius (Clapeyron-Clausius equation).

The Clausius-Clapeyron equation is a thermodynamic equation relating to quasi-static (equilibrium) processes of the transition of a substance from one phase to another (evaporation, melting, sublimation, polymorphic transformation, etc.).

The equation of state of an ideal gas (sometimes called the Mendeleev-Clapeyron equation or the Clapeyron equation) is a formula that establishes the relationship between pressure, molar volume, and absolute temperature of an ideal gas. It gives the relationship

$$PV = nRT,$$

where P, V and T are the pressure, volume and temperature; n is the amount of substance; and R is the ideal gas constant.

As a result, thermodynamics was created. Of course, there were other contributors such as Clausius, who introduced the concept of entropy, Boltzmann, who provided a statistical interpretation of this concept, and Lord Kelvin.

Let's talk about them in more detail.

Rudolf Julius Emmanuel Clausius (1822—1888) was a German physicist, mechanic, and mathematician. He graduated from the University of Berlin, where he received a Ph.D. Starting in 1884, he was the rector of the University of Bonn.

Clausius's main works are dedicated to the fundamentals of thermodynamics and the kinetic theory of gases [49]. He was the first to rigorously formulate the principle of the equivalence of heat and work. In 1850, independently of William Rankine, he obtained the relationship between these quantities (the first law of thermodynamics) and developed the ideal thermodynamic cycle of a steam engine (the Rankine-Clausius cycle). That same year, simultaneously

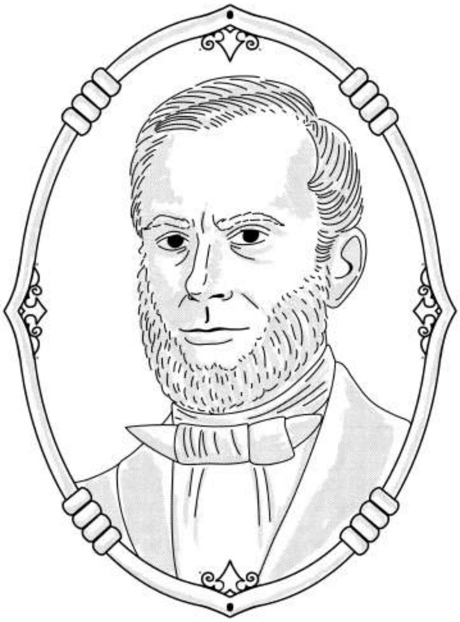


Figure 2.19: Rudolf Julius Emmanuel Clausius

with Thomson, he formulated the second law of thermodynamics: “Heat cannot by itself pass from a colder body to a warmer one.” In 1865, he introduced the concept of entropy and established its most important property: in a closed system, entropy is either constant (reversible processes) or increases (irreversible processes), thus indicating the direction in which a given process proceeds. He also proposed the idea of the “thermal death” of the Universe, extending to it the principle of increasing entropy. Boltzmann later disproved the fallacy of this idea.

Clausius contributed significantly to the development of the molecular-kinetic theory of gases. He was the first to apply a new approach — the so-called method of averages (now called statistical methods) — and explained from a unified standpoint such diverse phenomena as internal friction, thermal conductivity, and diffusion. He introduced the concept of the mean free path of molecules and in 1860 calculated its value, which later allowed the estimation of the size of molecules. He generalized the van der Waals equation of the gas state and elucidated the meaning of the equation that relates the melting (or boiling) temperature of a substance with pressure (Clapeyron-Clausius equation).

William Thomson, Baron Kelvin (1824—1907), was a British physicist and

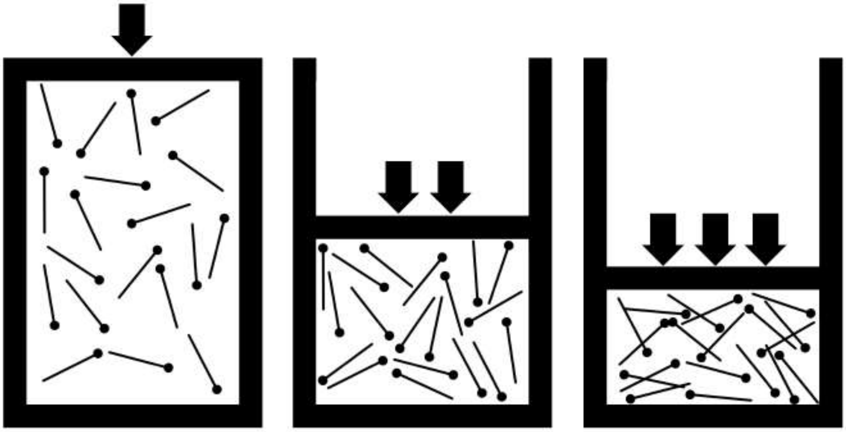


Figure 2.20: Kinetic theory of gas

mechanic known for his work in thermodynamics, mechanics, and electrodynamics. When he was seven years old, his father received a chair in mathematics at the University of Glasgow, Scotland, and moved there with his family. At the age of eight, William began attending his father's lectures, and at ten, he became a student. After graduating from Glasgow, the 17-year-old entered the University of Cambridge. During this time, his first scientific article on trigonometric series was published.

In 1845, after graduating from the university, Thomson, on the advice of his father, went to Paris to study in the laboratory of the famous French experimental physicist Henri-Victor Regnault. There, Thomson developed a method for solving electrostatic problems ("the method of electrical images"). A year later, the 22-year-old scientist returned to Glasgow, becoming a professor and head of the department of physics at the university.

In 1848, Thomson introduced the "absolute thermometric scale." He explained its name as follows: *"This scale is characterized by complete independence from the physical properties of any particular substance."* He noted that infinite cold must correspond to a finite number of degrees below zero on the air thermometer, namely, the point corresponding to the volume of air reduced to zero, which would be marked on the scale as −273℃ (more precisely, currently by international agreement, it is taken as −273.15℃).

Beginning in 1851, Thomson published a series of scientific articles under the general title "On the Dynamic Theory of Heat" [272], in which he considered

Figure 2.21: Lord Kelvin

the first and second laws of thermodynamics. At the same time, he returned to the problem of absolute temperature, noting that "the temperatures of two bodies are proportional to the amount of heat, respectively, taken and given away by the material system in two places that have these temperatures, when the system completes a full cycle of ideal reversible processes and is protected from loss or addition of heat at any other temperature."

This conclusion allowed Thomson to express the efficiency of a heat engine (Carnot cycle) using the temperatures of the heater and refrigerator.

In the same year, at the age of 27, Thomson became a member of the Royal Society of London — the English Academy of Sciences. Two years later, together with the English physicist James Joule, he found that when a gas passes adiabatically (without an influx of energy from outside) through a porous partition, its temperature decreases. This phenomenon is called the Joule-Thomson effect. Around the same time, Thomson developed the thermodynamic theory of thermoelectric phenomena.

In addition to thermodynamics, Thomson studied electromagnetic phenomena. In 1853, he published an article "On transient electric currents." Considering the change over time of the electric charge of a spherical body when it is connected by a thin conductor (wire) to the Earth, Thomson found that damped oscillations arise with certain characteristics depending on the electrical capacity of the body, the resistance of the conductor, and the electrodynamic capacitance. Subsequently, the formula reflecting the dependence of the period of free oscillations in a circuit without resistance on the indicated values was called the Thomson formula (although he himself did not derive this formula).

Figure 2.22: Thomson scattering around the sun visualises the trajectory of changed particles in visible light

Thomson was the first scientist to study electrical oscillations, and he was invited to become the chief scientific consultant in laying the first transatlantic cables designed to create a stable telegraph connection between the two continents. For his significant contribution to this work, he was elevated to nobility in 1865, and in 1892, for outstanding scientific services, he was awarded the title of Lord Kelvin (after the name of the river flowing near the university where he worked for many years).

From 1890 to 1895, Thomson held the honorary office of President of the Royal Society of London.

Sir William Thomson died at the age of 83 at Largs, near Glasgow. He left behind 25 books, 660 scientific articles, and 70 inventions.

2.3 Internal combustion engine

In this chapter, we will discuss internal combustion engines. What, exactly, is the concept behind an internal combustion engine? Why was there a need for a different design when steam engines that powered locomotives and steamships were already in use? Well, it takes time for a locomotive to start moving. You need to boil the steam in the boiler. But on the railroad, you can schedule around this boiling period. Passengers can wait until the stokers fire up their locomotive. On a steamship, as well, you can wait until everything is heated up. But if I want an engine for some purpose that needs to start immediately, it would be necessary for the heating in this engine to occur instantly.

The first such idea belonged to **Christian Huygens** [242]. He suggested injecting a small portion of gunpowder into the engine cylinder and igniting it. However, Huygens did not realize his idea. In fact, work on the internal combustion engine was delayed by 100 years. And 100 years later — at the beginning of the 19th century — the following works were associated with the names of the Niepce brothers.

The first to create the world's first working internal combustion engine was **Nicephore Niepce** (1765 — 1833). He was born more than 250 years ago into an old noble family. Just like his older **brother Claude**, he attended the College of the Oratorian Fathers. Clearly, he did not receive a natural science education there. But he also struggled with the humanities, as he failed the final exam in those subjects. At 23, he planned to study law, but a year later, in 1789, the Great French Revolution erupted. The Niepce brothers soon became officers in the revolutionary army. After dedicating three years to the cause of civil liberties, they retired and settled on their estate. From then on, they devoted themselves to invention. The French Republic was in dire need of sources of raw materials and energy. The government, in support of patriotism, offered major rewards for solving important national issues. For example, a large sum was promised for cultivating the madder plant, from which the dye indigo was extracted. The brothers grew madder, but, alas, they did not receive the prize. They invented a water-lifting machine for Versailles, but Napoleon gave the order to a major industrialist. Similarly disappointing financially were the methods they invented for extracting fibers and starch from local plants.

The brothers attempted to create an internal combustion engine. The main problem was the fuel. After all, the fuel had to burn instantly. Various options were tried: luminescent gas, coal dust, and even such exotic substances as the spores of the clubmoss plant. This is known as lycopodium and is used in theatrical productions to create the effect of a flash or explosion. The brothers Claude and Nicephore Niepce tried to use lycopodium for internal combustion engines. None of them worked in the theater.

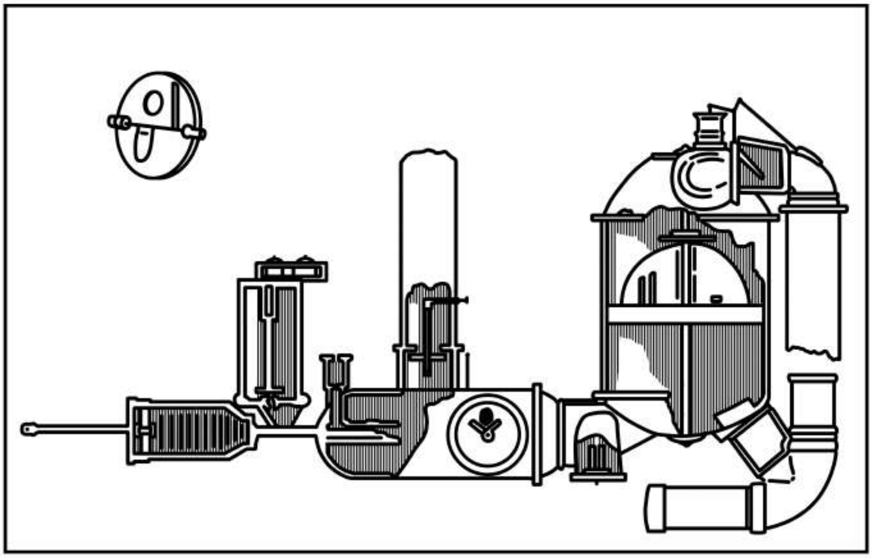


Figure 2.23: Pyreolophore

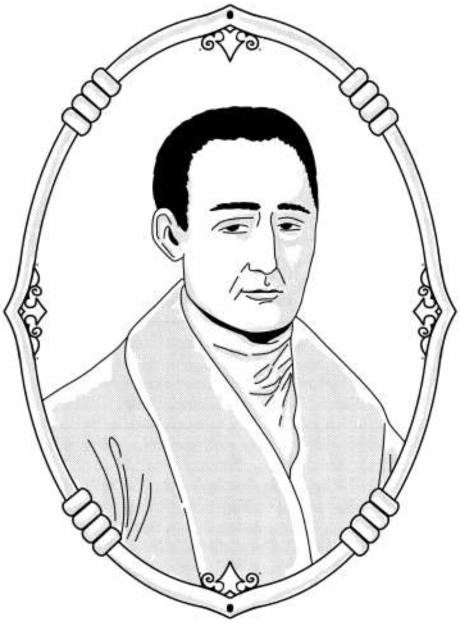


Figure 2.24: Nicephore Niepce

The brothers did not lose heart. In 1806, they submitted to the National Institute (as the French Academy of Sciences was then called) a report on a new machine that "would be comparable in strength to a steam one but would consume less fuel." The brothers called the machine "pyreolophore". It worked on coal dust, not gasoline or gas. This invention aroused great interest. Two commissioners were assigned to examine the invention. One of the commissioners was Lazare Carnot. Carnot gave a positive review, praising the machine. Although the engine had a number of shortcomings, many of them could not be eliminated at the time due to the lack of necessary technologies: the ignition of dust, for example, was carried out at atmospheric pressure, the distribution of the combustible substance inside the chamber was uneven, and the adherence of the piston to the cylinder walls required improvement. In those days, the piston of a steam engine was considered fitted to the cylinder walls if a coin could hardly pass between them.

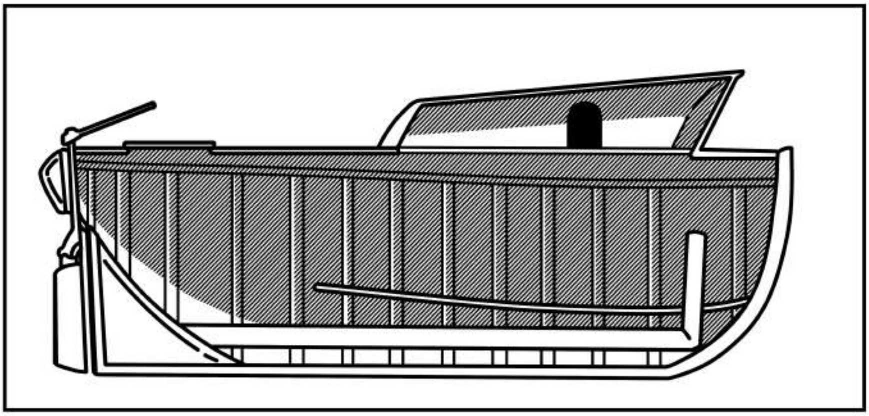


Figure 2.25: Boat of Niepce brothers

The brothers built an engine and equipped a three-meter boat with it in 1806, weighing 450 kg. The boat went up the river Seine at twice the speed of the current. The inventor brothers were looking for opportunities to enlist financial support and decided to seek an audience with Napoleon. They were preparing for the demonstration a small vessel with a two-cylinder engine. But the year was already 1811, and Napoleon did not come to Lyon, where the trials were supposed to be, as he was preparing to march on Russia. After the Russian campaign, the emperor had no time for engines.

Later, Nicephore Niepce read in the works of the chemist Lavoisier that petroleum "volatile oils" give explosive mixtures with air, and immediately appreciated this fact. But in France, the patent for the pyreolophore (engine) was about to expire, and in 1817 the elder brother Claude went to England, hoping to sell the patent there. After parting with Nicephore, Claude continued to work on the pyreolophore independently and led his inventive thought in another direction. In what direction it was difficult for Nicephore to understand: fearing the disclosure of secrets, Claude did not openly say anything in his letters but only asked endlessly for money and promised that results were imminent. It was only later that Nicephore learned from the letters that Claude was working on an engine that did not consume power, i.e., a perpetual motion machine.

By the way, Claude died soon after these efforts, and the second brother stopped working on the engine and began to experiment with photography. And we now firmly associate the name of Nicephore Niepce with the first photograph. However, Nicephore Niepce was unable to sell his engine for commercial production, and he simply did not live to see the photographic process realized in its final form.

But now, we move on to the next stages of creating an internal combustion engine. **Etienne Lenoir** [281] (1822—1900) was a French inventor of Belgian origin.

Figure 2.26: Etienne Lenoir

His father, a Belgian industrialist, died when the boy was eight years old. Planning to enter the Paris École Polytechnique — the most famous technical university in France — Lenoir was forced to give up his dream of becoming an engineer and began working as a waiter in a rather unassuming restaurant known as "The Bachelor Parisian." Owners of workshops and mechanics often met among the regulars of this institution. So, serving snacks and alcohol, the young man lived with the problems of mechanics and engineers, and a bold plan for the fundamental improvement of the engine was already beginning to form in his head. Soon, leaving his position as a garçon, Lenoir went to work in one of the many workshops, where his responsibility was to compose new enamels. About a year later, after a falling out with the owner, Lenoir became a lone mechanic who repaired everything from carriages to latrines and kitchen utensils. After working for some time, and without achieving either gratitude or money, he entered the workshop and foundry of the Italian Marinoni, which, with Lenoir's help, was transformed into an electroforming workshop. Finally, Lenoir led a comfortable life and got the opportunity for experimental invention. At that time, he created his own variations of a low-power electric motor, a dynamo regulator, and a water meter. Lenoir patented all his inventions and continued his experiments, making extensive use of the engineering experience of his predecessors. He approached the scheme of a double-acting steam engine.

The first prototype of the engine pleasantly surprised Lenoir and his sponsor Marinoni with its quietness. There were also disadvantages—it heated up too quickly during operation and required a fundamentally different cooling system. Due to a legal slip, Lenoir's machine was sealed; however, this is what prompted him to create his own company. Very soon, the company for the production of gas engines "Lenoir and Co" began to operate.

In 1860, Lenoir received a patent for his invention, and in the same year, the German engineer Otto became acquainted with the engine. Later, together with Langen, he created a company for the production of such engines. It was this company, which at first glorified Lenoir's work, and later took away his laurels.

Lenoir's machine was successfully demonstrated at the Paris Exhibition of 1862. The French magazine "Ilustration" offered the public a drawing and description of Lenoir's omnibus — a three-wheeled, eight-seater crew with this engine.

In December 1872, Lenoir's gas engine was installed on an airship, and the tests were successful. However, Lenoir's glory was short-lived — by 1878, the Germans had bypassed him. The noisy and bulky 4-stroke machine of his former colleague Otto, with a large vertical flywheel, worked with an efficiency of 16 %, while Lenoir's two-stroke engine reached only 5 %. Of course, the record was broken. But such was the time — inventions and previously unseen technologies shook the world every day, and there was still so much to come!

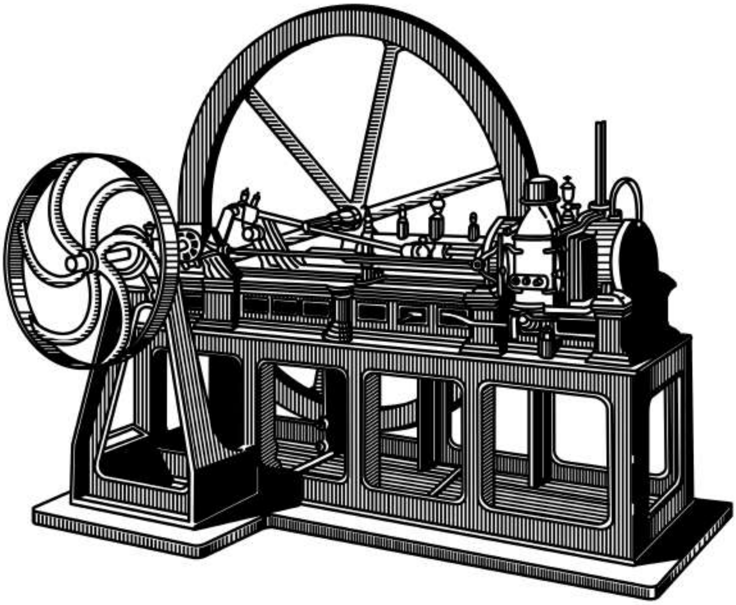


Figure 2.27: Lenoir engine

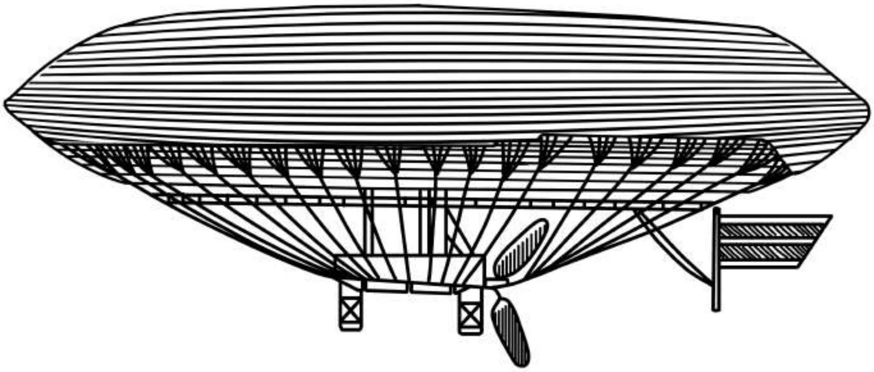


Figure 2.28: Airship of Haenlein

Jean Joseph Etienne Lenoir died on August 4, 1900. He was neither rich nor famous. He was just one of those who brought progress closer without receiving either fame or money.

Lenoir made the first few cars running on lamp gas, and he crafted both the engine and the chassis. The date of Lenoir's trip is unknown, but it is known that the first trip by car was 18 kilometers long, took 3 hours, and included 2 breakdowns. By the way, Lenoir made about 20 cars. Interestingly, one of them was sold to Russia to Emperor Alexander II. Unfortunately, there are no documents left about what happened to this car after the sale. And we do not know if the Russian emperor was the first motorist. It's likely that after the first two breakdowns, he lost interest in the car.

2.4 And then it was the Germans' turn...

One of the first, most serious people who made a very large contribution to the creation of already practical internal combustion engines that could be practically used, and who began to produce them on a commercial basis in large numbers, was Nikolaus Otto.

Figure 2.29: Nikolaus Otto

Nikolaus August Otto (1832—1891) — a German engineer and self-taught inventor, known as the inventor of the internal combustion engine. Otto's importance has been repeatedly but unsuccessfully contested, in particular by Christian Reitman, Alphonse Beau de Rochas, and Felice Matteucci.

Nikolaus Otto [210] was born in the small village of Holzhausen on the banks of the river Taunus. Within a few years, he was left without a father, and Nikolaus had to quit his studies and start working to feed his family.

However, his thirst for knowledge didn't end there, and as much as he could, he attended various courses, primarily related to technology. With a job as a traveling salesman, it was quite difficult, but he managed. Still, he had to feed his family, but the enterprising German managed to find what a few years later became a real gold mine for him.

At the end of the 1850s, Frenchman Lenoir introduced the world to a two-stroke engine of his own design. However, the engine had a large number of drawbacks, primarily a small output and a tendency towards spontaneous combustion. Nikolaus Otto sought to improve this design. As a result of his research, he concluded that the most promising option was a four-stroke engine.

In principle, experiments on creating a four-stroke engine were carried out even before Otto, but the authors faced a number of problems, primarily with the fact that the flashes of the combustible mixture in the cylinders occurred in such unexpected sequences that it was impossible to ensure an even and constant power transfer. Otto managed to find the right key to solving this problem, discovering that the problem with all previous engine designs was the mixture (proportions of fuel and oxidizer). In addition to this, he needed to solve the problem of synchronizing the fuel injection system and its combustion. It was on the development of these problems that Nikolaus Otto focused his attention. Furthermore, he managed to enlist the support of the then-prominent industrialist Otto Jogen Langen by founding Gasmotorenfabric Deiz AG. Otto focused entirely on finding the necessary solutions.

The result was not long in coming. By 1863, the first sample of an atmospheric gas engine was ready. But at that point, Otto had not yet managed to develop his famous four strokes, and he simplified the design, making it a two-stroke. However, this was already a breakthrough! For the first time, a design was created that surpassed a steam engine in terms of efficiency and was adapted for operation.

In 1866 Nikolaus Otto received a patent for his engine. Interestingly, a practically identical engine was previously patented by the Frenchman Alphonse Eugène Beau de Rochas, but since for some reason he could not create a working engine, the patent was given to Otto after he provided the commission with a working engine.

The first achievement gave Otto confidence, and he continued to work on improving his engine. Ten years later, in 1876, he finally patented the first four-stroke engine in history. An engine that worked according to the cycles we are now accustomed to: intake, compression, working stroke, exhaust. Cycles that have not changed to this day, despite changing our world.

The new engine began to sell very well. Over 15 years, $30,000$ engines were sold, which brought their creator a considerable income. However, long and

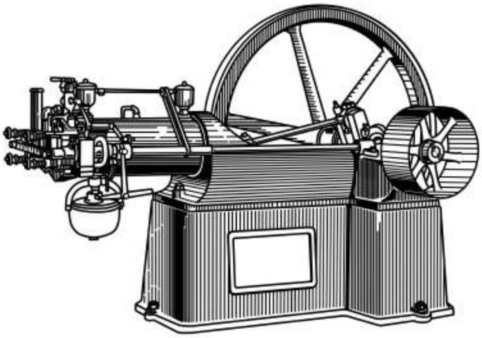


Figure 2.30: Otto's engine

hard work negatively affected Otto's health. From 1888 he began to literally fade away, and on January 26, 1891, he passed away: his heart could not stand it.

Another engineer by the name of Benz — this surname is now known to many, if not all — continued his work. Benz as in Mercedes Benz, of course. So, the engineer Benz figured out how to put Otto's engines on a cart, and make the cart's chassis turn the wheels and move. In fact, this is where the idea of the car appeared.

Karl Friedrich Michael Benz (1844—1929) — German engineer, inventor of the automobile, pioneer of the automotive industry.

Several generations of Karl Benz's ancestors lived in Pfaffenrot and were always engaged in blacksmithing. Karl's father first became a skilled blacksmith and locksmith, but later worked as a locomotive driver on the railway. Karl Benz attended secondary school in Karlsruhe and later, under the influence of his mother, entered the technical school in Karlsruhe and successfully completed it, having passed the final examinations brilliantly. While studying at the technical school, young Karl's main interest was steam locomotives and other steam-powered vehicles. A difficult period in Karl's life was the years after graduating

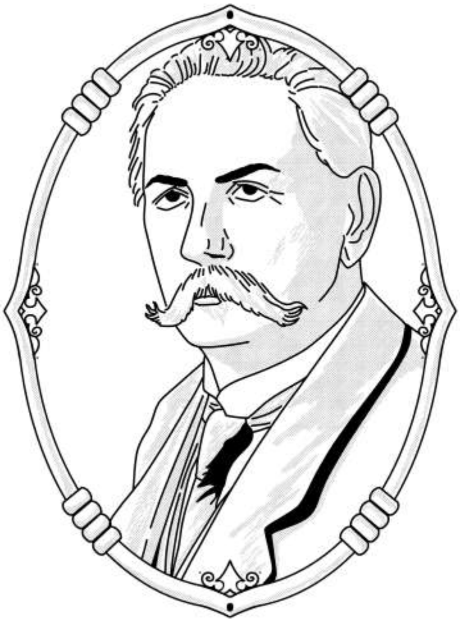


Figure 2.31: Karl Benz

from technical school. He worked as a salaried employee in many engineering enterprises but was constantly obsessed with the idea of creating a new type of engine, as Otto's naturally aspirated engines were becoming widespread at that time.

After his mother's death in 1870, Benz decided to quit his job and found his own workshop with an acquaintance where experiments could be carried out. They bought a small piece of land and started by making metal spare parts. However, as his partner resisted the idea of experimenting with engine development, Karl had to give up his dreams. Benz almost resigned himself to this.

He soon met Bertha Ringer and married her. Thanks to his wife's inheritance, he managed to buy out his leisurely partner's share and became the sole owner of the workshop. Now he could devote all his time to developing a new engine. Unfortunately, he did not pay attention to the financial condition of his enterprise, and it soon went bankrupt in 1877. All banks refused him further loans, although by this time he had developed a new internal combustion engine, and there was an urgent need to start production of a prototype model. Despite considerable difficulties, Benz managed to create a sample of a new two-stroke engine, but he could not bring it to the market, as an English company had already developed and patented a similar engine, which made it impossible to obtain an opinion on authorship. However, the Patent Office still granted a patent for the fuel system, which eventually allowed him to start producing a number of engine models. He founded a new firm that began making small two-stroke engines.

In 1885, Karl Benz and his investors founded another new firm. During the day, he worked in his workshops, and at night he experimented in a barn near his house. Perseverance, initiative, and dedication allowed Benz to overcome the initial difficulties. The result was the creation of a three-wheeled vehicle with a 4-stroke engine in his workshop. Benz himself designed and developed all the components of his car and came to the solution of many technical problems himself. In January 1886, Karl Benz received a patent for his new car, which did not generate much interest among buyers, although the engines were in high demand in the market, especially in Germany.

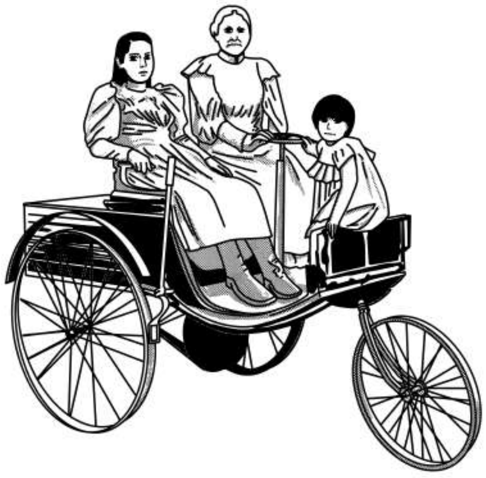


Figure 2.32: Bertha Benz

In August 1888, Bertha Benz, together with her sons of 15 and 13 years, without Karl's knowledge, made a trip in it to the neighboring city at a distance of 106 km. Along the way, she had to buy gasoline from pharmacies (it was sold there as a cleaning agent) and replace worn brake linings at a saddler's. She climbed hills several times, pushing the car. Along the way, people came running in droves to gaze at such a miracle.

The whole of Germany learned about this long-distance rally. And the press paid serious attention not only to the journey but also to Karl Benz's car itself. From that time, his path to fame and success began.

Benz cars began to be sold, financial matters improved, and the inventor began to work on new models. In 1893, the first four-wheeled vehicle was created. By the end of 1899, the 2000th car had already been released, and production figures reached 572 models per year. Thus, the Karl Benz company took first place in the world in terms of production among car manufacturers. Before World War I (1914—1918), the motor and automotive industry at Karl Benz had reached its highest development — the company's products had become famous in many countries. However, the defeat of Germany led to a complete collapse of the country's economy, including the automotive industry.

In 1889, Benz's representative in France presented his car at the Paris Motor Show. At the same time, cars of the German company "Daimler" were demonstrated there. Unfortunately, the exhibition did not bring any successful sales. This was the case until 1890, when a number of German firms became interested in the production of the Benz car. A new company was founded, exclusively producing the car.

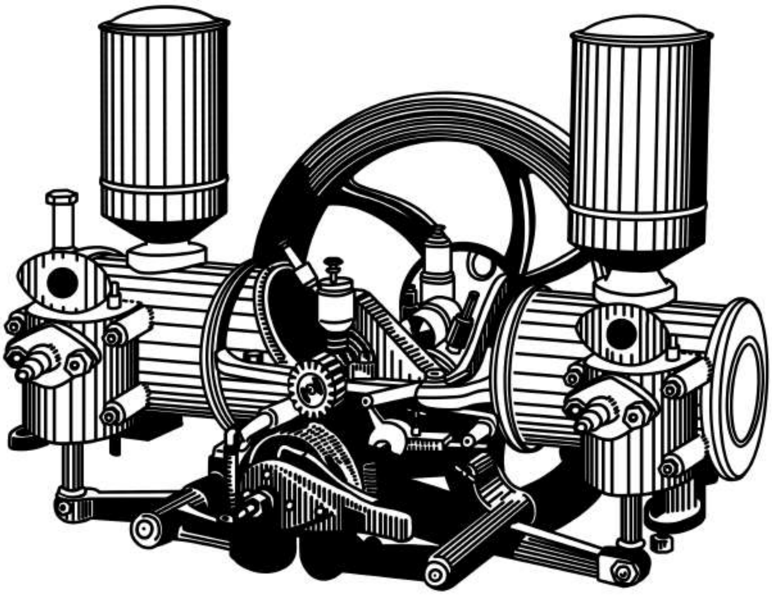


Figure 2.33: Contra-engine

In the following period, Benz continuously worked on his new project, including performing test runs of the cars. In 1897, he developed a horizontal 2-cylinder engine known as the "contra-engine." The Benz company soon achieved recognition and high popularity among buyers due to the high sporting results of its cars. Finally, after many years of failure, Karl Benz came to a more successful stage. In 1926, the company "Benz" merged with the company "Daimler," and the company "Daimler-Benz" was founded, which still exists today. Karl Benz died on April 4, 1929, at the age of 85. Unlike many other inventors, he was held in high esteem and was very wealthy at the time of his death.

But what about the fuel? Let's return to Otto. His engine ran on lamp gas. And Otto's two assistants — Maybach and Daimler — at some point realized that this was impeding the development of the engine and cars in general. In fact, the mechanism was a kind of firebox where firewood was placed. This firewood was kept at a temperature of about 200 degrees. A mixture of methane and carbon monoxide was released (this was, in fact, the luminous gas), which was the fuel of the first Otto's engines. It can be said that, as a result, they operated on wood. But at this time, the chemistry of oil was already developing. And people could already divide oil into fractions: gasoline, fuel oil, diesel fuel, and kerosene. Both Maybach and Daimler came to Otto and said, "We need to change." And Otto said, "I don't want to. My engines are selling in the tens of thousands. The profit is very good. Why should I change?" And they left Otto.

2.5 Fuel and combustion engine evolution. Or revolution?

On March 17, 1834, in a small town called Schorndorf, the future automobile genius Gottlieb Wilhelm Daimler (1834—1900) was born. This man was destined to become one of the greatest designers and engineers who would not only give the world one of the first cars, but also develop several types of engines.

Figure 2.34: Gottlieb Wilhelm Daimler

Gottlieb was born into a poor baker's family, and in 1847 he graduated from high school, after which he decided to become an apprentice to a local gunsmith. There he specialized in the manufacture of double-barreled guns. And only after another 10 years did Gottlieb enter the Polytechnic Institute of Stuttgart. After completing his studies, he worked in various enterprises in France, Belgium, and England.

His meeting with Maybach happened largely by accident but became fateful for Gottlieb. It all started with the fact that in 1863, Daimler was appointed head of the Bruderhaus plant, which was located in the city of Reutlingen. This plant was a part-time charitable institution, because disabled people, beggars, and orphans worked there. Among Gottlieb's other responsibilities were supervising the orphanage where these workers lived and trained.

One of the orphans was 19-year-old Wilhelm Maybach, who lost his parents at the age of 11, was admitted to an orphanage, and received the assignment of a designer and draftsman there. Maybach proved to be not only a talented engineer but also a reliable business partner. After they met, they quickly got along, and in the future, Wilhelm accompanied Gottlieb throughout Germany.

Their work in Karlsruhe began in 1869 when Gottlieb was assigned to the Maschinenbau-Gesellschaft Karlsruhe AG factory. At first, he was there alone, but after just 6 months, Wilhelm Maybach moved to Karlsruhe and also got a job at the factory.

The move to Cologne took place in 1879 because Daimler was appointed to the post of technical director of the enterprise, led by the inventor Nikolaus Otto. Maybach also went to this city, receiving the position of chief designer.

It is noteworthy that at the same time as them, Karl Benz created his own internal combustion engine, already a 2-stroke one. He received a patent for it in 1879.

However, Gottlieb did not have to work for long in Cologne because a year later, disagreements began with Otto. The latter was jealous of Daimler because he had a higher education, and Otto was self-taught. As a result, Gottlieb and Wilhelm left Deutz AG but continued to work.

In 1882, both inventors moved to Stuttgart. There, in a suburb called Cannstatt, they bought a house and added a workshop to it. They worked on developing a power unit, for which they planned to use petroleum. At the time, there were only three suitable fuel options: lubricating oil, petrol, and kerosene.

Their choice fell on gasoline, even though it was previously used in the domestic sphere — gasoline was used to clean clothing and was sold in pharmacies. The reason for their choice is obvious, since it has the highest degree of flammability of the three available materials.

In 1885, the designers Maybach and Daimler created their first power unit, and in the same year, they developed a carburetor. In November of 1885, they managed to assemble and patent the world's first motorcycle. It was quite primitive: a small motor and wooden wheels, placed on an ordinary frame made of wood. During the tests, Wilhelm Maybach was able to drive about three kilometers

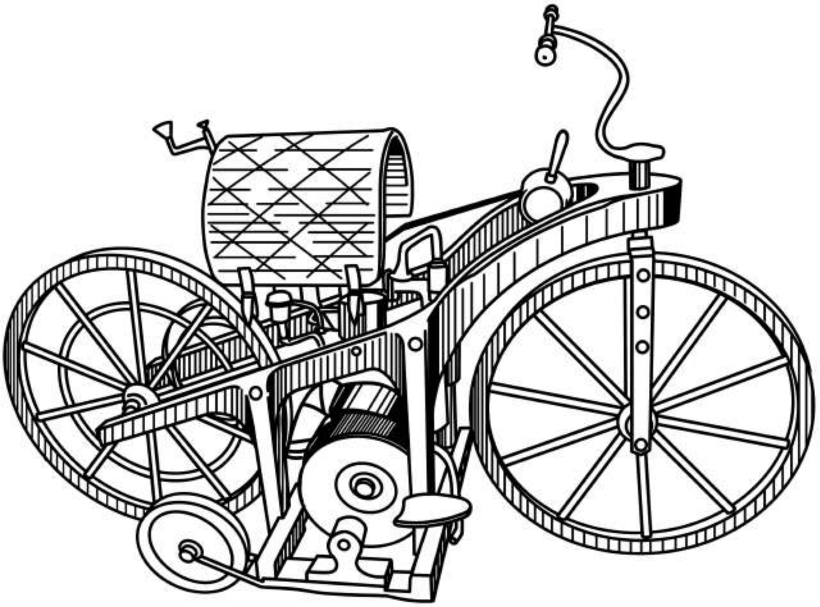


Figure 2.35: Reitwagen: world’s first motorcycle with IC engine, 1885

along the banks of the Neckar River, with his speed reaching 12 km/h.

For Emma Daimler’s birthday on March 8, 1886, these designers prepared a gift. It was a carriage, on which Wilhelm put a 1.5-horsepower engine. The torque from it to the wheels was transmitted by means of a belt. It was the first car in the world made in the form of a self-propelled crew. In addition, the engineers developed and put outboard motors on the conveyor, which for a number of years had been the source of the firm’s cash receipts.

The first car was designed in 1889 by the joint efforts of Gottlieb and Wilhelm, although it looked more like a simple carriage left without a horse. It was presented at the Paris exhibition, simultaneously with the development of Karl Benz.

Despite the death of his wife, Gottlieb did not quit his job and created in 1890 the company Daimler Motoren Gesellschaft. The company specialized not in the production of cars, but in the production of sufficiently powerful, yet compact motors that could be used on land, in air, and in water. It was this thesis that became the basis for the corporate logo of the three-pointed star, which now adorns the products of the Mercedes-Benz group. It is noteworthy that in the same year, DMG was officially reorganized, turning into a joint-stock company.

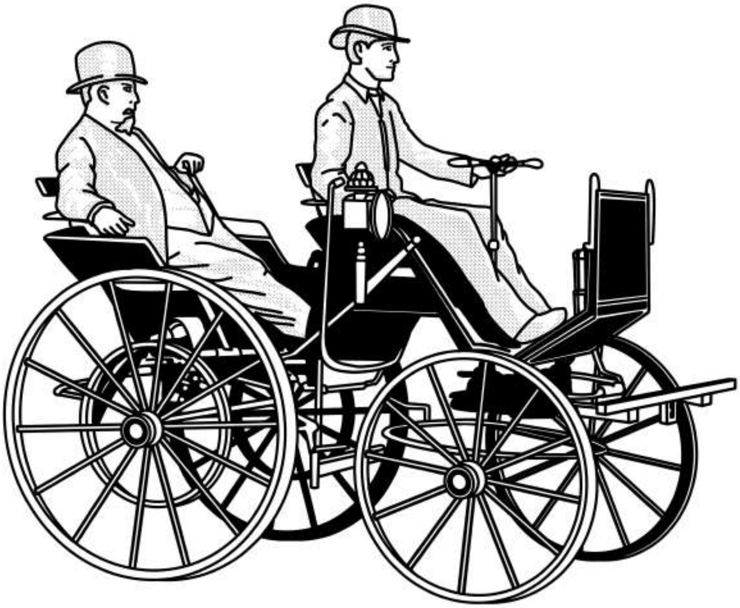


Figure 2.36: Self-driving carriage of Daimler, 1886

Wilhelm Maybach, by that time, believed in his own strength and decided to work independently, having separated from the company already in 1891. The Daimler company began producing cars, the first of which was sold in 1892.

In the winter of 1892—1893, Daimler suffered a heart attack, after which, following the advice of his doctor, he went to rest in Italy. In 1893, Daimler decided to leave Daimler Motoren Gesellschaft, although he left all his patents to the concern.

The return took place in 1894, when Daimler, together with Wilhelm Maybach and his son Paul, developed a new power unit, which was named "Phoenix." And in the same year, the English businessman Frederick Simms acquired a license for the Daimler brand and the engine of this company.

Gottlieb Daimler died at the turn of the century on March 6, 1900. His work was continued. In the future, the company had a merger with Karl Benz to look forward to, alongside a great future 666 becoming one of the world's largest automobile companies.

August Wilhelm Maybach (1846—1929) is a German auto designer and en-

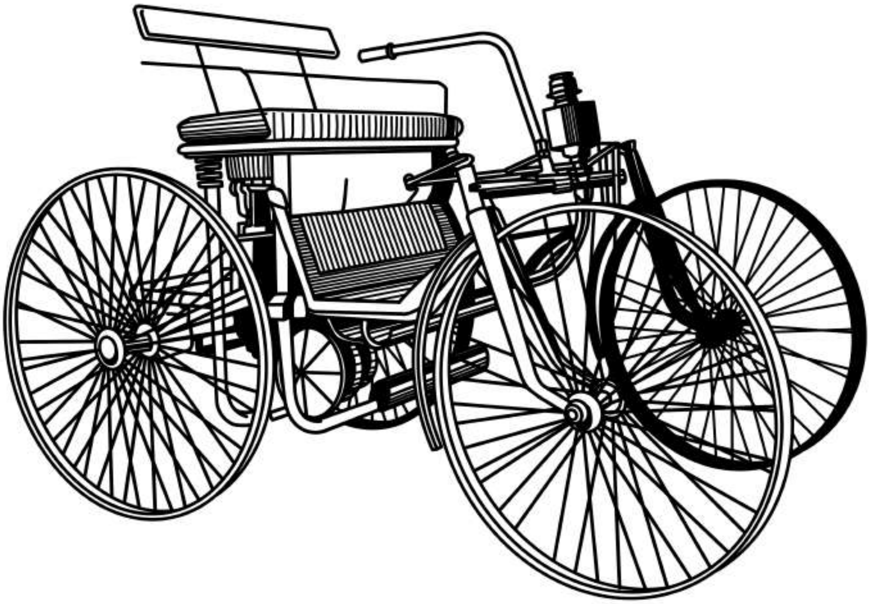


Figure 2.37: First Daimler vehicle, 1889

Figure 2.38: August Wilhelm Maybach

trepreneur, founder of the automobile companies Mercedes and Maybach.

Born in the family of a carpenter in the city of Heilbronn (Württemberg). After the sudden death of his parents in 1856, he ended up in the orphanage "Bruderhaus Reutlingen", where he worked in a workshop, the director of which at that time was Gottlieb Daimler. He turned out to be a capable student and attracted the attention of Daimler, becoming his friend and assistant.

After graduating from the orphanage school, he was sent to courses at the Reutlingen Technical College. In 1869, he entered the design office of the Maschinenbau Gesellschaft enterprise in Karlsruhe. When in 1872 Daimler received an invitation from Eugen Langen to take the position of technical director at Otto und Langen Gas Motoren Factory Deutz, which produced gas engines, he declared that he would accept it only if there was a place for his assistant Wilhelm Maybach. The contract was signed.

Maybach took up the improvement of the Otto's engine and the organization of production. In March 1882, due to disagreements that arose with the management of the enterprise, both were forced to leave the company. In April of the same year, Daimler founded the Daimler Motoren Gesellschaft company, and Maybach switched to it as a designer. A year later, the first successful test of a four-stroke engine took place, and soon its production began.

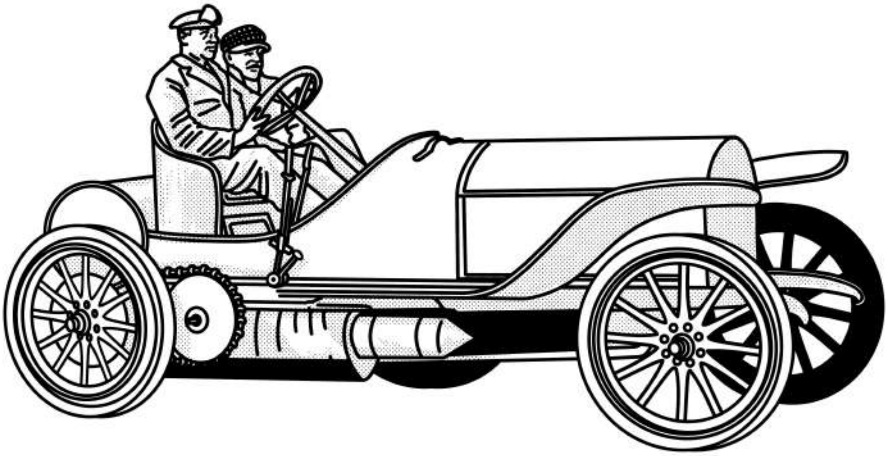


Figure 2.39: Maybach car

Maybach also participated in the work carried out by Daimler in 1885 to create the first motorcycle, which was patented a year later. In subsequent years, he was involved in the design of cars and engines for Daimler. In 1901, under his leadership, the first Mercedes passenger car was created, followed by several more models that became winners in a number of prestigious sports competitions, including the Gordon Bennett Cup. In 1906, the first Mercedes car with a 6-cylinder Maybach engine appeared.

At the beginning of 1907, Maybach left the Daimler enterprise and accepted the proposal of Count Ferdinand von Zeppelin to create a company for the manufacture of engines for airships. In 1912, this company was reorganized and moved to the city of Friedrichshafen, where the production of aircraft engines began.

After Germany lost the war and was forced to sign the Treaty of Versailles, it lost the right to produce military products at its enterprises, including engines for airplanes. Therefore, Maybach returned to the automotive industry.

In 1921, a new company appeared, bearing the name of the designer, which specialized in the manufacture of high-class cars. True, Maybach himself increasingly retired. Most of the developments introduced at the company belonged to his son Karl, who inherited his father's talent as a designer. In the spring of 1929, Wilhelm Maybach, having barely celebrated Easter with his family in Cannstatt, fell ill and died two days later. The name of the designer is immortalized in the Automotive Hall of Fame in Detroit.

Figure 2.40: Emil Jellinek

Despite the fact that it is generally accepted to associate the name of the Mercedes car with the names of Wilhelm Maybach and Gottlieb Daimler, we should also recall the German businessman Emil Jellinek, the man to whom the Mercedes automobile brand owes its beautiful name. Who was he? An idler, an idiot, a minion of fate, an eccentric rich man. Jellinek's childhood passed in Leipzig, where he constantly got involved in crazy stories, which greatly annoyed and angered his father, a respected and influential rabbi. By the age of nineteen, his father's patience snapped, and Emil was sent to France, after which, thanks to his father's connections, Emil was placed on a diplomatic mission to Morocco. There he married the daughter of a local tobacco planter, a Jewish woman named Rachel Goggman. Together with his father-in-law, they managed to establish the supply of tobacco to Europe. Then he managed to increase his capital by trading on the stock market. With that kind of money, it was easy to avoid military service, of course, for health reasons. Here he returns to Europe, just as eccentric, but now rich, where by that time (the end of the 19th century) the entire wealthy beau monde was smitten with a new expensive hobby — cars.

Jellinek had the opportunity to meet and buy new items in the automotive market. He immediately went to Stuttgart as soon as he learned about the Daimler Motoren Gesellschaft, the workshops of Daimler and Maybach. He already knew that in most races, Daimler-powered cars came first. Without hesitation, he ordered two cars with engines of 4 and 9 hp. However, the cars did not justify Jellinek's high-speed hopes. Instead of the declared 40 km/h, the Daimler Phoenix, even with nine horsepower, barely reached 25 km/h. The buyer announced to the designers that he did not want cars that could not move faster than horses. He promised to purchase four more cars, provided that their speed would reach 50 km/h. Both designers were negative about increasing the power of their offspring. But too good money was promised by an eccentric millionaire. The result pleased Emil. The cunning Jew deliberately overtook Baron Arthur Rothschild, who was on a car ride. A funny fact, but the rich of Europe at that time considered it fashionable to hide under false names. Jellinek took the pseudonym "Monsieur Mercedes". Rothschild found this "monsieur" and bought the car from him. Less than a month later, "Monsieur Mercedes" repeated his trick. Gradually, Daimler Motoren Gesellschaft gained a major league client base, and Jellinek gained serious influence over the company's owners. He could easily dictate to them his, at first glance, absurd technical solutions. The first car engines were two-cylinder and located in the rear of the car; at his whim, the next batch of cars was equipped with four-cylinders. As he reasoned, the motor in the car performs the function of a horse pulling a carriage along with it. So why doesn't the engine move to the front under the hood? After some Daimler failures in racing, a tragedy occurred — due to an accident at a famous rally, the company's best mechanic, Wilhelm Bauer, died in the hospital. Gottlieb Daimler died suddenly and unexpectedly. Maybach, like Daimler, was not a supporter of increasing speed.

Figure 2.41: Mercedes

At that time, the following legend of the appearance of the name Mercedes is most popular. When the inventors once again sat and thought, what name would they come up with for the car so that it hooked everyone, it was not an engineer, not an inventor, just a representative who entered the place where they were sitting... No, no, no, it was a representative of the Daimler company in France, Jellinek. And he told them: "What are you doing? Are you inventing a name? But what's the problem, use the name of my daughter." And so they did.

At first, the car was simply called Mercedes, without Benz. By the way, during the company's life, it was called both Daimler Benz and Daimler Mercedes. These are all the names of the engineers.

But let's return to the fuel evolution. Changes in the type of fuel in an internal combustion engine would not be possible without the following invention. Gasoline is liquid in the tank, and vaporized in the engine cylinder. Why do we need vaporized gasoline? After all, liquid gasoline does not burn. **Donát Bánki** (1859—1922), a Hungarian engineer and inventor, came up with an amazing invention that allowed gasoline to evaporate — a carburetor.

Figure 2.42: Donat Bánki

In 1893, together with another famous Hungarian engineer, János Csonka, he invented the stationary engine carburetor (also known as the Bánki-Csonka Engine). Bánki also made significant contributions to the development of the internal combustion engine compressor. He also invented the water turbine named after him.

In 1898, Donát Bánki developed a high-compression engine with a twin-diffuser carburetor, which used the evaporation method that is still used today.

The invention of the carburetor by Bánki and Csonka contributed to the development of the automotive industry, since up to this point no more efficient way had been developed to properly mix fuel and air for the engine. Some sources claim that Bánki randomly borrowed the idea for the carburetor from a flower girl. One day, returning home from the Budapest Technical University, he saw her spraying her flowers with water from her mouth. Donát Bánki also contributed to the creation of the cross-flow turbine.

For the invention of the carburetor, Bánki and Csonka received a patent on October 18, 1893, under the name "Improvement of Gasoline Engines."

The idea behind this carburetor is that if the gasoline is sufficiently finely dispersed in the air, it will be evenly distributed throughout the cylinders, and evaporation will occur in the cylinders under the action of the heat of compression. To ensure dispersion, Bánki proposed sucking gasoline into the air stream through a metering jet, and to maintain a constant composition of the mixture, he proposed maintaining a constant level of the liquid column in front of the jet. The jet was made in the form of a side hole (or holes) in a tube located perpendicular to the flow, and a small tank with a float was provided to maintain the pressure, which maintained the level at a given height (float chamber). The amount of gasoline taken in, if we neglect some factors, is proportional to the amount of air taken in. Bánki proposed his method for calculating the simplest carburetor.

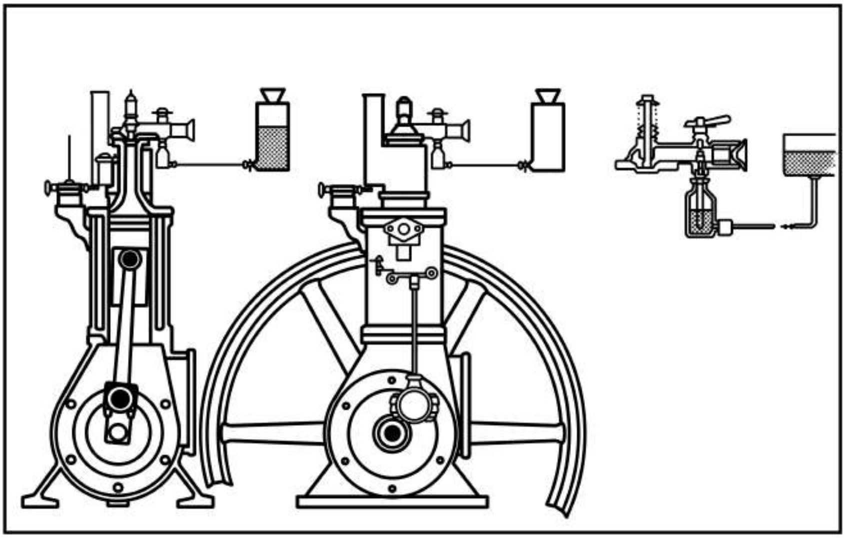


Figure 2.43: Bánki-Czonka's carburetor

The description of the Bánki-Csonka carburetor appeared long before the issuance of a patent (February 11, 1893), and on August 17, 1893, Maybach received a French patent for a similar device. Since this patent was issued before Bánki-Csonka's patent, there was a reason for long patent litigation.

Only in the 1930s, on the basis of studies confirming the priority of Bánki-Csonka, did the International Federation of Engineers recognize the carburetor as an invention of Hungarian engineers.

2.6 The birth of a modern car

Before the invention of automobiles, human thought was directed to the improvement of carriages. By the time cars were created, carriages had reached a high level both in terms of convenience for passengers and the elegance of form. It was the 19th century — the century of beauty in architecture, furniture, and clothes. The carriages were correspondingly also very beautiful. Surely you have seen them in museums or at least in photographs.

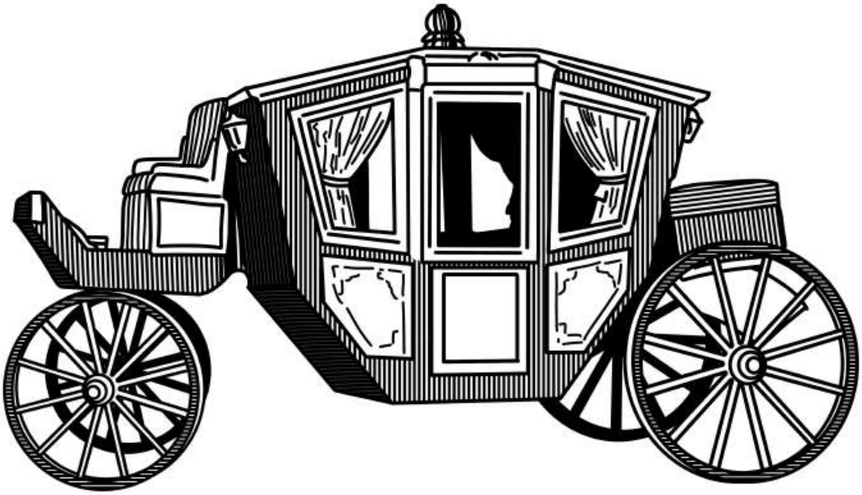


Figure 2.44: Old carriage

The first cars were created in Europe, in cities with medieval architecture where the streets were not adapted for the movement of cars. However, the first cars differed little from carriages, so they did not create any special problems. Cars were admired for being beautiful, safe, and comfortable for passengers, though they smelled a bit like gasoline. Problems appeared later as the technical power of cars increased.

For several decades, there has been fierce competition among designers who tried to make cars as powerful and fast as possible. The main indicators of the first cars exhibited in museums were power, engine size, and speed. Even today, these same parameters are showcased again at automobile exhibitions.

But how did cars evolve to have the characteristics of modern automobiles?

The engine was invented, and the fuel question was also almost solved. However, at first, cars could not climb any relatively steep hill; they lacked "strength." The fact is that internal combustion engines are capable of developing necessary operating power in a small speed range. But how then to change the torque without being limited by the engine's capabilities? There was a need for a mechanism that would transmit the engine's torque to the car's wheels. Such a mechanism was credited to Louis Renault... or was it?

Not really. It all started with the invention of Karl Benz. As mentioned earlier, in 1890, Karl's wife, Bertha, secretly set off on a long road trip to show the world her husband's car and prove that it was not a "devil's car." It was a terrible, unbearably hard journey: leather brake mechanisms wore out, fuel had to be sourced from pharmacies and shops — naphtha, a stain remover. The 0.8 HP engine did not have enough power, and Bertha and her son were constantly forced to push the car. At the end of the trip, Mr. Benz sat down with his drawings.

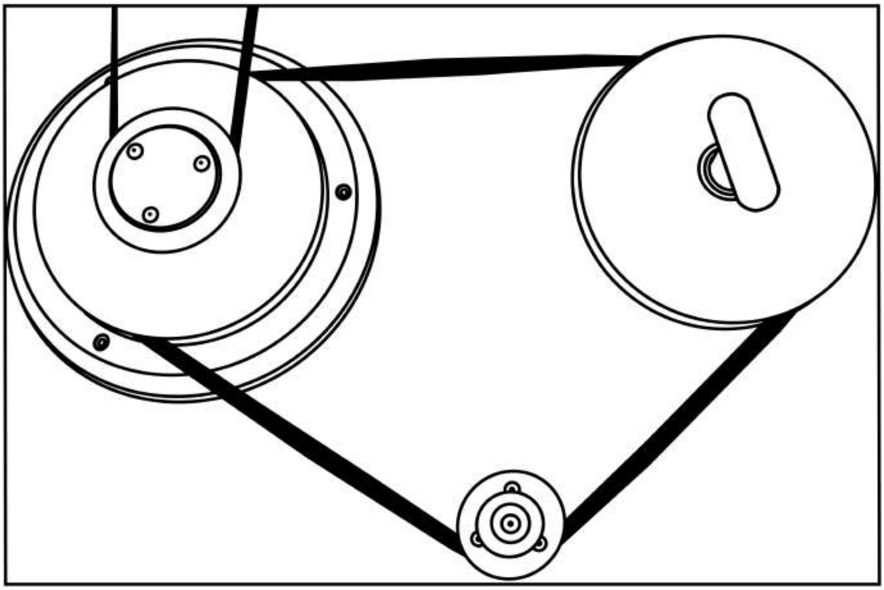


Figure 2.45: Motor shaft

The first version was primitive: two pulleys of different diameters on the drive axle, connected to the motor shaft by a belt. That's all. This design worked on the basis of simple physics, changing the torque on the drive wheels. When a pulley on the motor shaft rotated an axle pulley of a larger diameter, the torque increased, and the car started off. When the smaller pulley rotated, the angular velocity of the wheels increased and the engine force diminished. A little later, the belt was replaced with a chain, and the pulleys with sprockets.

What about Louis Renault? Error in facts? No.

A more modern gearbox was invented in 1898 and was first used in the Renault Voiturette. On December 24, 1898, Louis Renault defiantly drove his car along one of the steepest streets in Paris — without a gearbox, this would have been impossible. But front-wheel drive first appeared only in 1929 on the Cord L29 car, although it reached mass production in the post-war years.

So what did Louis Renault invent then? Essentially a driveshaft. The production model Voiturette Type A had an engine power of 1.75 HP and was equipped with the world's first gearbox with three "forward" speeds and one reverse. Direct drive with a Cardan shaft, invented by the founder of Renault, is used in rear-wheel drive cars to this day.

We are talking, of course, about a manual transmission, but an automatic transmission appeared in the USA in 1939 in Oldsmobile Custom 8 Cruiser cars.

By the way, the **steering wheel** also did not appear immediately. The first cars were equipped with tillers — control levers; there was no question of any round steering wheel. The control was obvious but ineffective at high speeds: the driver pulled the lever to the right or left, and the car turned in the indicated direction. Actually, it was the thirst for speed that gave impetus to the introduction of a round steering wheel. With the advent of automobiles, by the end of the 19th century, the first races began to appear, in which steering the tiller turned into a driver's nightmare.

In 1894, Alfred Vacheron fitted a round wheel to his Panhard 4hp and performed well at the Paris-Rouen race. By 1898, all Panhards were fitted with a steering wheel. Other automakers followed suit. The first steering wheels had a rigid unregulated steering block, which sometimes led to terrible injuries to drivers, sometimes even in not very serious accidents.

Internal combustion engine, gearbox, steering wheel. The car already looks like a modern one, doesn't it?

By the way, in 1891, Edouard Michelin created a removable pneumatic tire designed for a bicycle, and in 1895, removable pneumatic tires for cars were produced. In the same year, the tires were tested at the Paris-Bordeaux-Paris race, but the car equipped with them dropped out of the race because the tires were often punctured. Despite this, experts and motorists appreciated the smoothness of the car, and gradually all cars began to be equipped with pneumatic tires.

But what else is missing for a car to look like a modern one? And how to explain such a variety of modern cars? For example, in the race to design cars after World War II, the Americans took the lead. They aimed to make cars as big as possible, with increased power and speed, as oil was very cheap. The Europeans at that time were restoring their destroyed economy and were forced to economize. Therefore, American and European cars were noticeably different: the former were large and voracious, while the latter were small and economical. However, in the 1970s, the oil crisis broke out: oil-producing countries, united in OPEC, sharply raised oil prices. The Americans were forced to create more economical cars.

A few more interesting dates in the evolution of cars:

In 1910, the first electric horn appeared.

Air conditioning appeared in 1939 on the Packard Twelve Sedan. It was very expensive and extremely inconvenient to use.

The first audio system appeared in 1930; these were Motorola radio sets, and in 1932 the famous Blaupunkt appeared on the German Studebaker.

The wipers were invented by an American woman, Mary Anderson, who invented and patented a mechanical drive for brushes, with which drivers cleaned windshields. Power wipers were invented by Charlotte Bridgewood 14 years later, in 1917. But the mass “hanging” of wipers is the merit of Bosch.

Turn signals, as we know, appeared on the Buick Roadmaster in 1939, replacing mechanical arrows and inconvenient flashlights.

The first power steering appeared on the luxurious Chrysler Crown Imperial in 1951, replacing pneumatic systems and the power of human muscles. Three years later, power steering also reached Europe — in the French Citroen DS 19. By the way, it was the Citroen DS 19 that became the first car with disc brakes.

Anyway, it was a surprisingly advanced car for its time.

Among the first cars, there are superheroes who turned out to be the "first carriers" of several devices at once. One of these was the Cadillac Model 30 Self Starter of 1912: the model presented to the public had a starter, ignition, and the first headlights with a tungsten filament (rather than flimsy carbon). As for the starter, it replaced the cumbersome starter handle — the very handle that started the first engines. However, this handle is still relevant for some cars, for example, for some UAZs. By the way, it was the invention of the starter that partly set back the existence of electric vehicles many years ago: the use of internal combustion engines became simple, and the motivation for the development of electric motors disappeared; it was necessary to develop internal combustion engines.

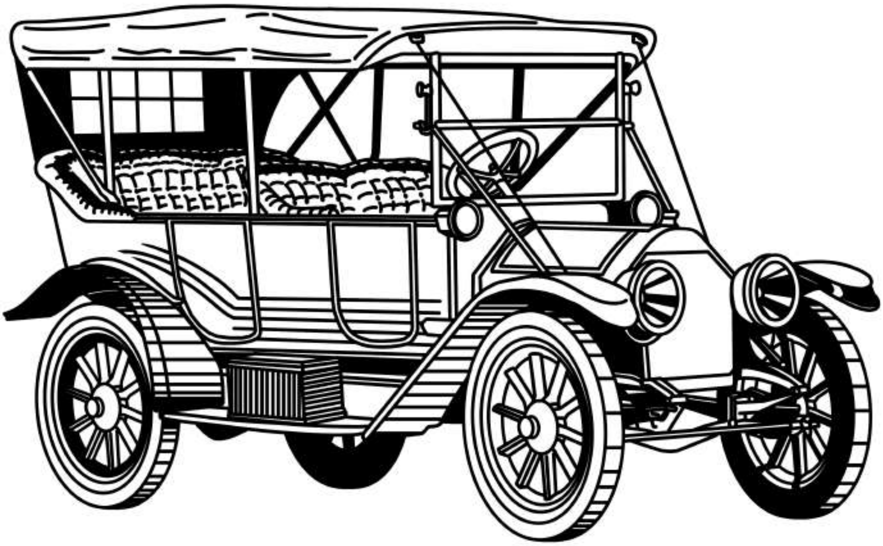


Figure 2.46: Cadillac Model 30 Self Starter

Wait, what other electric motors? The Mitsubishi i-MiEV of 2009 is considered the first mass-produced electric car... Stop! We are not interested in 2009; we are interested in 1828, when the Hungarian physicist Anyos István Jedlik invented the electric carriage. The history of electric vehicles goes back almost 200 years. The first electric cars appeared almost 50 years before the first car.

After Faraday's invention of the phenomenon of electromagnetic induction, engineers tried to put this invention into practical use. Accurate information about the time of appearance and the name of the creator of the first electric

Figure 2.47: Anyos István Jedlik

car has not been preserved.

The experiments with the ancestors of Tesla did not end there: in 1834, the blacksmith Thomas Davenport created his own version of the electric car, followed by the Dutchman Sibrand Stratingh and his assistant Christopher Becker, who equipped their invention with a battery for recharging.

All of them were heavy, moved at a speed of no more than 4 km/h, and were of little practical use. The development of electric vehicles was hampered by the lack of energy-efficient batteries.

In 1865, the Frenchman Gaston Plante presented a prototype of such a battery. It was not yet suitable for practical use, but the principles embodied in its design were adopted by other inventors. By the beginning of the 1880s, relatively light, and most importantly, sufficiently capacious and rechargeable batteries were being created. This caused a boom in electric vehicle construction.

But the most advanced and more or less “tenacious” car turned out in 1890 from William Morrison — he developed a speed of up to 22 km/h.

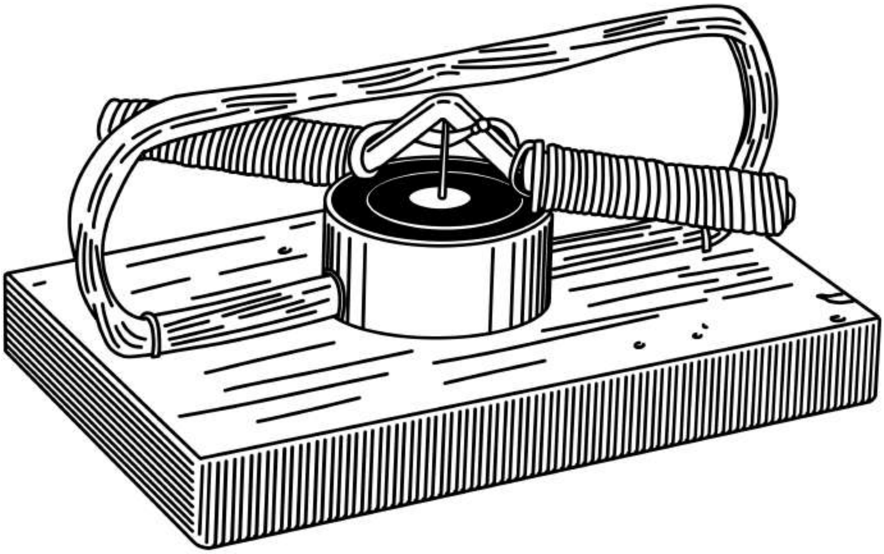


Figure 2.48: World's First Electric motor by Anyos Jedlik

The end of the 19th and the beginning of the 20th centuries can be considered the "golden age" of the electric car. At that time, few people believed in the prospects for the development of internal combustion engines. The average electric car of those years reached speeds of up to 30 km/h, and the power reserve was quite enough for trips without recharging or changing batteries during the day. At the same time, the electric motor "started" without problems in any conditions, did not require gear shifting, and worked silently. The complete opposite was represented by cars in those years. The roaring and whimsical engine, releasing fetid clouds of exhaust, the smell of gasoline and oil, the need for a manual start and gear shifting — all this deterred potential customers. At that time, only wealthy people could afford to buy a horseless carriage. And they, of course, preferred a clean, quiet, and easy-to-use electric car. Electric vehicles were so simple that they were effortlessly driven by women and elderly people. The success of the "electromobilization" of those years is also evidenced by the fact that the first speed records were set on electric vehicles. In 1895, the world's first officially registered race took place, during which the electric car of the Frenchman Charles Jeantot achieved a speed of 63 km/h. And in 1899, for the first time in history, a land vehicle exceeded the 100-kilometer speed limit. The electric car Jamais Contente, built by the Belgian Camille Jenatzy, accelerated to 105 km/h. In the first decade of the twentieth century, electric vehicles became even more widespread. They were used as taxis, fire engines, and ambulances. Their speed and range of travel without recharging increased.

Some models were equipped with a regenerative braking system. Electric vehicles were most popular in the USA — by the beginning of the 20th century, 30 percent of all cars in the USA were electric.

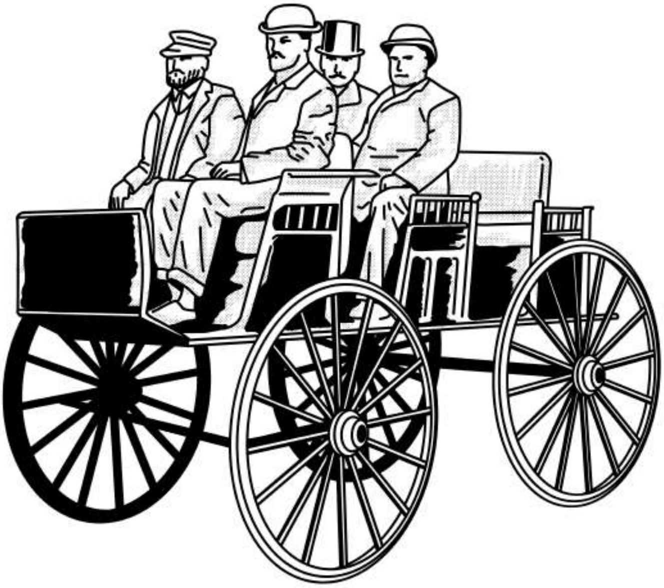


Figure 2.49: William Morrison's electric vehicle, 1890

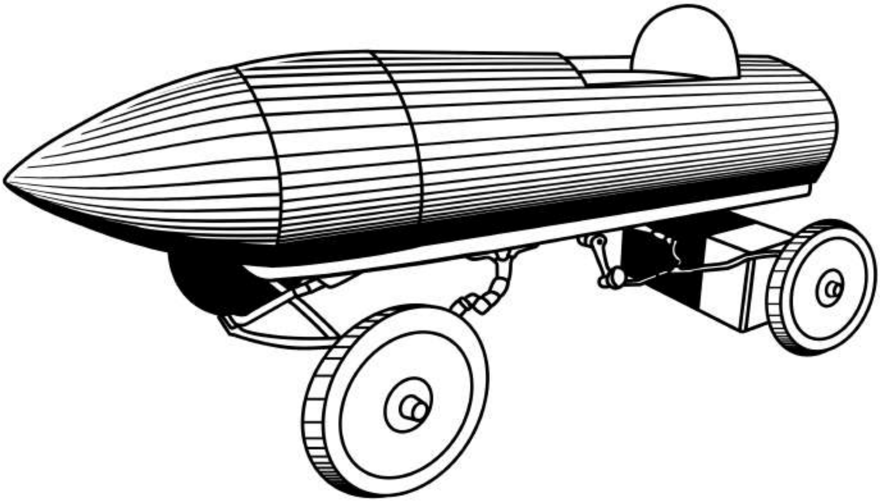


Figure 2.50: La Jamais Contente electric vehicle, 1899

A hybrid car may still seem like a curiosity to some, but in fact this invention is not so new, — its history actually dates back to the 19th century. The first hybrid car was created over 100 years ago. Probably the whole world knows about an electric car driven by an electric motor, but what about its hybrid counterpart, what kind of life has befallen it? The first hybrid car was built in 1899 by Ferdinand Porsche "Lohner-Porsche Mixte" and was accepted by the public at the Paris Motor Show in 1901. While the hybrid car was certainly very practical, demand quickly waned. But why?

We have already mentioned the starter motor that was probably the killer of the electric car, but in fact there was a more obvious killer — the Ford T car. Henry Ford used an assembly line to assemble cars, thereby significantly reducing the cost of the vehicle. And then it was up to ingenious commerce: the great millionaire refused the margin and made money on trade, making the first people's car in history. Ford T turned out to be three times cheaper than an electric car — the winner of the competitive race was determined.

2.7 Were the cars safe?..

If modern automakers would remember the instances of cars a hundred years ago, they would certainly be horrified: shoe brakes and a solid steering column, like a pike, aimed at the driver's chest. But until the 1950s, such cars were mass-produced. Back then, automakers cared little about the safety of drivers and passengers; the emphasis was primarily on speed and to a lesser extent on convenience. And only closer to the middle of the 20th century did it become obvious that the safety of the car is perhaps the most important stage of production and it should be given great attention. Let's look back at the fascinating history of the development of automotive safety systems and how they have affected the modern industry.

Until the early 20th century, the first ICE car bodies looked more like horse-drawn carriages. Some models at first did not even have a floor, but only passenger seats, a roll-up roof and a primitive protective front bulkhead. The engine on such cars was usually located directly under the seats. By the beginning of the 20th century, there were about twenty main types of bodies, which, following the example of carriages, had French names.

Gradually, more durable metals began to be used in body structures. Around the same time, there was a confrontation between open and closed bodies. The first sports car built in the USA, the "Mercer Type 35 Raceabout" in 1911, had an open body, while the English "Lanchester" featured a completely hermetically sealed version. In subsequent years, automakers began to lean towards closed bodies, as they were more practical and reliable, with open options becoming more of a design element.

In 1905, German mechanical engineer Frederick R. Simms installed the **first bumpers** on Simms-Welbeck cars, which incorporated pneumatic elements to absorb impact energy. Thus, bumpers became an integral part of the car, with American companies like Duesenberg and Lincoln giving their bumpers a distinctive look.

Surprisingly, at the beginning of the last century, some cars could reach speeds of 100 km/h, necessitating the installation of an effective braking system. **Disc brakes**, more familiar to us, were patented by Englishman Frederick Lanchester in 1902. However, they were not widely used due to the noise and rattle they produced.

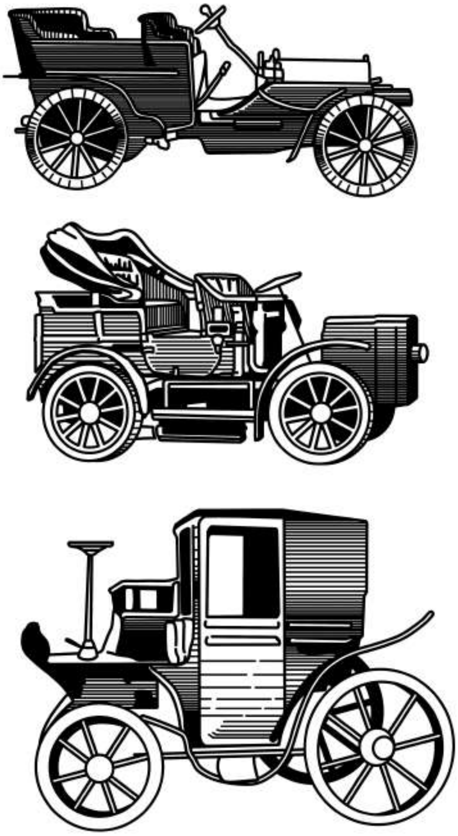


Figure 2.51: Car bodies

Until the 1940s and 50s, **band brakes** and later **drum brakes**, whose pads were securely hidden inside the drums, gained popularity. Initially, brakes were installed only on the rear wheels. The braking system for all four wheels appeared in 1910 on the Arrol-Johnston model. Drum brakes on all wheels first appeared on the Lancia Lambda in 1922. **The first hydraulic brakes** appeared in 1921 on the Model A Duesenberg. The hydraulic brake system required considerable effort when pressing the pedal, so in 1923 Louis Renault invented and patented the first mechanical booster, which was installed on all Renault production cars. The first dual-circuit braking system (mechanical and hydraulic) began to be installed in 1966 on Volvo cars.

With the increase in speed and the number of cars, engineers began to implement systems to facilitate driving and eliminate many dangerous moments on the road. By 1916, most American cars were already equipped with **windshield wipers**, and by the early 1920s, the first **electric headlights** were being installed. In 1938, all production Cadillac cars had rear-view mirrors, wipers, and fog lights.

In 1928, **impact-resistant windshields** were installed on Ford Model A cars. Going back to 1887, Scotsman John Boyd Dunlop invented and patented the bicycle tire, and his invention quickly migrated to cars, making travel much safer and more comfortable. In 1904, Continental developed the first ribbed tires, significantly improving the car's handling. The first radial-ply tires were developed by Michelin in 1946 and are still widely used today. In 1912, American Edward Budd invented the first all-metal body in his workshop, which was widely used only in 1928.

With the popularity of racing, car manufacturers began to pay more attention to the ease of control of the car, leading to cars with load-bearing bodies being made in the 1930s. This was achieved by combining an all-metal body with a rigid frame, increasing the maintainability and durability of the bearing elements and noticeably decreasing the total mass of the vehicle. The first mass-produced monocoque car was the 1934 front-wheel drive Citroën Traction Avant. Later, instead of a load-bearing body, cars were equipped with a load-bearing frame, providing a lower car seat, additional body protection, and noise reduction. The first such vehicles were the serial "Volkswagen KdF" released in 1939.

By the 1950s, cars gradually became not just a luxury item but a common means of transportation, requiring increased safety for the driver and passengers through the introduction of various systems and devices, which would later be called passive safety.

The first patent for a car seat belt was issued to American Edward Claghorn in 1885. At the beginning of the 20th century, racers and aircraft pilots most often used seat belts. The situation changed radically when, in 1959, Volvo began to equip its production cars, starting with the Volvo PV 544, with a reliable three-point seat belt. The development was not innovative: the first three-point seat belts appeared as early as 1902. However, Niels Bohlin, an engineer at the Swedish company, managed to significantly refine the technology, making it as reliable and convenient as possible. Bohlin developed the belt for over a year, during which time he studied about $20,000$ accident reports. In 1962, he received a patent for the three-point seat belt. Following Volvo, by 1964 all new American cars were equipped with three-point seat belts, and two years later, the device became standard in the USA.

In the mid-20th century, many automakers began to conduct their own crash tests of cars.

In the late 1940s, Volvo developed the concept of a high-strength carcass ("habitat"), which reduced the impact force.

By the early 1950s, Béla Barényi, an employee of Daimler-Benz who would later be recognized as the "Father of passive car safety," proposed the revolutionary idea of a safe cabin construction.

Figure 2.52: Volvo PV 544

The concept was a combination of an impact-resistant "habitable pod" with ample space for passengers and collapsible energy-absorbing zones at the front and rear of the vehicle. The technology aimed to convert kinetic energy into the deformation force of body elements. Thus, the contact time with an obstacle increased by about ten times, inversely reducing the loads.

The development gained public recognition and underwent many crash tests, including frontal collisions. The first cars to use energy-absorbing body elements were American Packard premium-class vehicles, produced since 1952. The first car to fully incorporate the principles of passive safety proposed by Barényi was the Mercedes-Benz W111, which hit the market in 1959.

In 1954, Barényi proposed a prototype steering column with a collapsible element, which, in the event of a collision, did not cause severe injuries to the

Figure 2.53: Béla Barényi

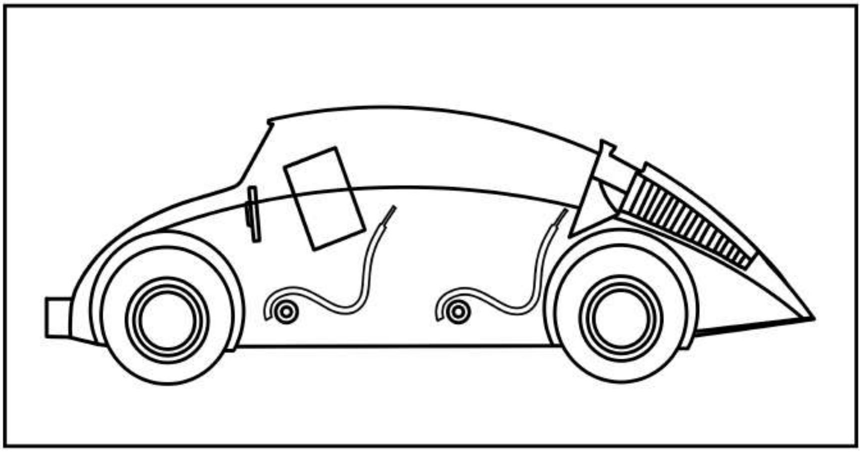


Figure 2.54: Impact-resistant capsule

driver. The development remained overlooked for some time but eventually demonstrated its full advantage over telescopic rigid structures.

In 1956, Ford began to serially equip its cars with five-point seat belts, but the technology did not become widespread. In 1960, Volvo started installing a padded front panel in its production cars, reducing facial and chest injuries. The company was also one of the first to test child safety seats in 1964 and install them in their vehicles.

It's important to note that until the early 1960s, child seats were not considered a safety feature; they were used to raise the child to the eye level of other passengers. In 1978, the United States became the first country to enact a law requiring the transportation of children in special restraint seats. Another significant Volvo invention in 1967 was seat headrests, protecting passengers' necks and heads in rear impacts, which became a US standard two years later. These headrests remained on the market for nearly 30 years, and only in 1995 did Saab introduce active models on the Saab 9-5 car, which, due to inertia, activated a mechanism bringing the headrest closer before the head tilted back, thus reducing the impact force. Saab was also the first company to install energy-absorbing windshields on production cars in 1971 and side protection beams in the doors in 1977.

Gradually, the material of car bodies became more durable and lighter, and interior elements were made from crushable or soft materials. However, perhaps the most important element of passive safety remains airbags. The first experiments with inflatable airbags for the driver and front passenger were conducted by American engineers in 1968, and only five years later did they appear in General Motors and Chevrolet production cars. They were first installed on the Ford Taunus 20M P7B and Oldsmobile Toronado but did not gain widespread popularity until over a decade later.

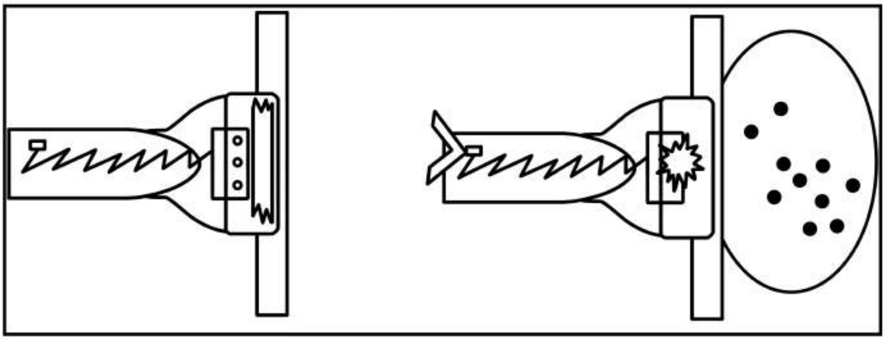


Figure 2.55: Airbag in action

For the first time, an electronically controlled airbag was introduced in Mercedes S-Class cars in 1980. Since then, airbag technology has undergone many changes, such as analyzing the position of passengers before impact and adjust-ing the operation algorithm depending on the situation, but the general principle of operation has remained unchanged to this day. The great success of the front airbags led to the introduction of side airbags in the 1994 Volvo 850 model, in combination with stiffeners installed in the front doors. The company went even further by installing inflatable curtains in the S80 sedan to protect passengers in a side impact. In 1996, Kia equipped its Sportage SUV with knee airbags.

In the mid-2000s, at the initiative of international organizations for pedestrian protection, safety hoods began to be installed on cars, which, with the help of squibs, are able to automatically rise to protect pedestrians during a collision. The Jaguar XK and Citroën C6 were among the first cars to use this technology. In 2012, Volvo began to install airbags for pedestrians built into the hood on the V40 model.

Until the beginning of the 21st century, developments were mainly carried out to improve passive safety by enhancing the car body and various systems; and not in vain — for half a century, the structural safety of cars has grown several times, saving thousands of lives. But the evolution does not stop there: existing technologies are constantly being improved, new materials are being created, and so on. Recently, emphasis has been placed on the introduction of active safety systems — technologies that reduce the proportion of the human factor during an accident. Although this stage is still in its active development, much has already been accomplished.

It all started back in the 1970s when Mercedes-Benz and several other companies began developing an electronic anti-lock braking system (ABS) that prevents the wheels from locking up during emergency braking, thereby helping the driver maintain control. This technology first appeared on the Mercedes-Benz 450 SEL model in 1978 and was subsequently installed on many domestic and foreign cars.

From 1987 to 1992, Mercedes-Benz, in collaboration with Bosch, developed an electronic stability control system for cars during sharp maneuvers called the "Electronic Stability Program" (ESP). In 1995, the technology was refined and renamed "Electronic Stability Control" (ESC). The first cars equipped with such systems were the Mercedes A-Class starting in 1997.

2.8 From cruise control to autonomy

In a big city, moving from point A to point B is always fraught with difficulties: driving a car means getting stuck in a traffic jam, and opting for public transport involves waiting, enduring inconvenience, transferring, and waiting again... and there's still a chance of ending up in a traffic jam. Even the metro has its challenges. It would be nice not to worry about these issues! Just get into the car, enter the destination address — and that's it, then read or listen to music, work on the computer, or gaze out the window—the car will take you to your destination via the optimal route. In science fiction, this idea has long become a cliché — but what about its practical implementation?

It all began in the distant 1930s when General Motors engineers conceived two ideas that were brilliant for their time.

The first idea was that cars would be controlled by radio signals, allowing them to maintain distance on the highway and avoid accidents.

The second was even more intriguing — to enable unmanned trips, special tracks would need to be built in the shape of skateboard ramps. Cars would travel in the middle, and if they began to veer towards the edge, the force of gravity would guide them back into the road's depression.

Despite initial skepticism, these ideas provided a significant push for technology to move in the right direction.

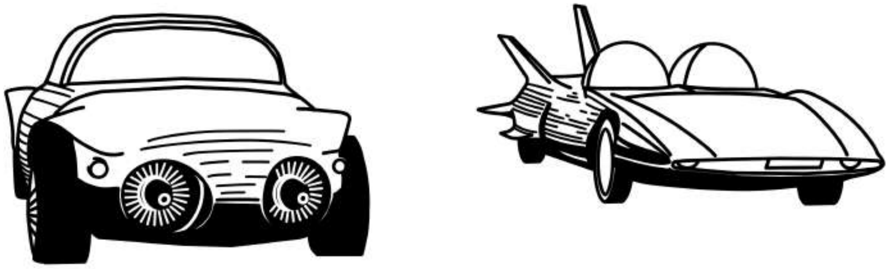


Figure 2.56: Firebird 2

In the 1950s, General Motors tested the "smart" car Firebird II, which, besides a new braking system, had magnetic sensors and interacted with a "smart" road (with an electric cable under the asphalt). The General Motors Firebird II, a 1956 model, was a four-seater with independent suspension. Underneath the titanium body lay a 200 HP Whirlfire GT-304 gas turbine engine, electrical package, and an integrated air conditioning system comparable to those at the start of the 21st century. The Firebird II, in terms of design and ergonomics, followed the 1953 car model, which was dubbed a "jet plane on wheels" (developers and engineers were indeed inspired by the fighter jet concepts of that time). However, the Firebird II was the first to apply a structure for future highway travel — a complex control system intended to interact with an electrical wire embedded in the road surface to send signals and serve as a guide for the latest cars. It was presumed that the electromagnetic field would minimize road hazards by reducing the human factor. At the time, it was an audaciously futuristic model that wowed at exhibitions but never entered mass production. Version #3 of this concept car was equipped with the first autopilot system, known today as cruise control. This system maintained a constant speed and significantly eased the driver's tasks on the road. However, full autopilot was still far off.

In 1995, Mercedes introduced the first practical solution to a technology that simplifies the process of parking called Parktronic. The system consisted of several ultrasonic sensors and a beeper indicator. The principle of operation was quite simple: the sensors measured the distance to the obstacle, and the beeper, by changing the sound signal, warned the driver about when to stop.

The highways of the future were built in Europe and the USA. The first production car that really interacted with them was the Citroen DS, the legendary passenger car that took third place in the ranking of cars of the century. Its low-power 75 HP engine did not stand out in those days, but the car was distinguished by an advanced transmission, combined with steering, brakes, and hydropneumatic suspension. This design was ahead of the automotive industry's development for many years. The Citroen DS was able to interact with the highway using an electrical signal, but there was no question of any independent autopilot — it was more for fun. By the way, its incredible popularity, advanced technologies, and the albeit relatively illusory autopilot made this Citroen Fantomas' flying car.

Experiments with onboard computers in the 60s and 70s were carried out, but they never entered series production. It is worth recalling the experimental Chrysler Plymouth, which was equipped with an onboard computer (well, as much as you could call the device that occupied half the rear seat an "onboard computer") and a generator to power the system, brought to the roof of the car. Laboratory tests were carried out for 10 years, but there was no question of any serial production.

Nevertheless, neither engineering thought nor the imagination of futurists stopped for a minute — mankind was looking for cars not only for luxury or a means of transportation but also for a smart assistant that could make life easier, make roads safer, and work for a person. This desire was reflected in the movies — after several films with “talking” cars, the real hits were a series of films about James Bond with his fancy cars and, of course, the legendary “Knight Rider.” The smart, humorous KITT car based on the Pontiac Firebird Trans AM not only reached speeds of under 500 km/h and was practically invulnerable but also knew how to talk, drive on full autopilot, and control all electronic devices from a distance.

Self-driving car history began in 1961 when Stanford student James Adams created and tested the first self-driving cart. It was controlled by a conventional signal, through a cable. But the second prototype of the Stanford cart was radio-controlled.

Figure 2.57: Trackless trolleys

This experiment did not go unnoticed, and in the 1970s, the well-known practical scientist John McCarthy made his adjustments to the device of the trolley, modernizing it with the help of a technological vision system. Now the trolley could move independently and be guided by the white line. The prototype was also equipped with a rangefinder, video cameras, and four channels for data collection. But this was not enough for the inquisitive mind of McCarthy — back in the 70s, he tried to create a three-dimensional mapping of the area.

After McCarthy's success, engineers focused on creating 100 percent autonomous vehicles without remote control. Scientists in the United States and Japan made significant progress, but the real breakthrough was made by German researchers led by Ernst Dickmanns.

The "Smart" Dickmanns car — this title is awarded to the automated Mercedes-Benz Vario. The impressive size of this van made it possible to place a huge computer system, and the silicon brain began to control the movement of a 5-ton iron monster. Dickmanns' first drone became the prototype of modern robocars — for the first time, computational mechanisms and an eye movement simulation system were used. These innovations made it possible to form a learning model for a car that independently assesses the situation and makes decisions.

Daimler-Benz paid close attention to Dickmanns' developments and launched the Prometheus project, the main goal of which was to improve drones and achieve unprecedented road safety. The project started in 1987, and during its existence (8 years), more than 1 billion dollars were spent. "Prometheus" went down in history as the most expensive project in the development of robocars of the twentieth century. However, the investment was well spent.

The results of the Prometheus project and the developments of Dickmanns were used for the serial production of the 1995 S-Class Mercedes. These cars were equipped with a more advanced cruise control system, which allowed them to adapt to the average speed of traffic flow and maintain a safe distance between vehicles.

By the mid-90s, two robotic drones, VaMP and VITA-2, were presented to the world. They were successfully tested at a test site in the Paris region, during which they:

- moved at speeds up to 130 km/h completely on autopilot;
- independently rebuilt and changed the row;
- monitor the distance and movement of other road users;
- overtaking the cars in front.

The mid-1990s saw the rise of GPS technology with artificially reduced accuracy of up to 100 meters. Permission to use the system for civilian purposes allowed the introduction of new automated systems in cars. In 1997, Mercedes, and two years later the BMW E38 and Toyota Celsior, began using adaptive cruise control technology to automatically maintain a set speed. Cruise control depends on the ABS and ESC systems — if they fail, the function is disabled. In 1997, the European committee for conducting independent crash tests of cars with an assessment of passive and active safety Euro NCAP began its work. The Volvo S40 became the first 4-star car for adult occupant protection, and the first 5-star winner was the Renault Laguna in 2001. In 2007, Volvo introduced blind spot monitoring technology on the S80 sedans, and a year later it equipped the XC60 with another innovation: Autonomous Braking.

To better understand what is meant by "autonomous" vehicle, let us introduce the 6 levels of driving automation based on the Society of Automotive Engineers (SAE) classification.

- Level 0 means that the vehicle is fully manual, even if there are systems that can assist the driver, such as the emergency braking system.
- Level 1 provides driver assistance, such as cruise control or adaptive cruise control where the vehicle can maintain a safe distance behind the next car.
- Level 2 offers partial driving automation. This means the existence of advanced driver assistance systems (ADAS), and the vehicle can control both steering and accelerating/decelerating. Examples include Tesla's Autopilot and Cadillac's (General Motors) Super Cruise systems.
- Level 3 represents a significant jump in the "brain" of the vehicle. In this level, "conditional driving automation" has some kind of "environmental detection" capabilities and can make informed decisions, like accelerating past a slow-moving vehicle. However, these vehicles still require human override in the event of an emergency. The first vehicle in the world supposed to have this level of automation was the 2019 Audi A8; however, due to the lack of government regulations, the carmaker decided to delay the market introduction of their Traffic Jam Pilot autonomous tech.
- Level 4 assumes High Driving Automation. Here, the car does not require human interaction in most circumstances.
- Level 5 vehicles do not require human assistance at all; they may not even have steering wheels or acceleration/braking pedals.

In May 2022, Mercedes-Benz became the world's first manufacturer to be approved by German transport authorities to legally operate its L3 Drive Pilot on the country's public roads, sold as an option on Mercedes-Benz S-Class and Mercedes EQS. This means that those with L3 Drive Pilot are legally allowed to eat, draft emails, or watch videos on the Autobahn. However, since L3 autonomy is conditional, if the vehicle loses the environmental or locational conditions to operate at L3, it will prompt the driver to take control within ten seconds. If the driver fails to respond in that time, the car will automatically turn on the emergency lights and decelerate to a full stop on the side of the road, then unlock the doors in case first responders need access to the cabin.

At CES 2023, Mercedes-Benz further announced that it has become the first manufacturer to receive L3 certification in the United States, from the state of Nevada. However, since L3 approval is granted at a state level in the US, the system is only considered L3 in Nevada for now. Nonetheless, the OEM says its Drive Pilot is fully prepared to deliver L3 autonomous driving in all 50 states.

One of the first prototypes of autonomous vehicles was proposed by Google — the Google Car. This was a mini car with an unprecedented level of autonomy. Designed for two people, it features two engines, custom body materials, all-electric propulsion, speeds of up to 25 miles per hour (just over 40 km/h), control from a start button, and does not require a person other than as a passenger. Naturally, it is integrated with Google services — on the central console, you can watch videos and movies on YouTube, work with mail, and surf in Chrome. Interestingly, the car was also built by Google, as previous partners Lexus and Toyota understandably imposed many restrictions on risky experiments. It is extremely difficult to enter the mass market of personal vehicles, and in December 2016, the project was renamed Waymo and became a new startup company that is part of Alphabet. Waymo underwent further tests on its cars on public roads after its separation from Google. In June 2022, Waymo announced a partnership with Uber, under which the former will integrate its autonomous technology into Uber's freight truck service. On December 13, 2022, Waymo applied for the final permit necessary to operate fully autonomous taxis, without a backup driver present, within the state of California.

In 2004, the first-ever auto-competition with the participation of DARPA robocars took place, where drones persistently declared themselves.

In 2012, AUDI tested its drone. The car on autopilot developed a speed of up to 193 km/h, navigated turns perfectly, and accelerated nicely on the track.

In 2015, the first mass-produced autonomous cars appeared — the Tesla Model S, which traveled on roads independently. Along with Waymo cars, they are considered the standard of unmanned automobile technology.

The period of 2016—2017 is when most major auto companies announced the development of their own prototypes of robocars and plans for their serial production.

In general terms, the automatic control system functions as follows: with the help of sensors, radars, and cameras (and in different projects, machines are equipped with various sets of sensors), the system gathers information from the surrounding environment — it "sees" road signs and markings, turns, obstacles, and monitors the behavior of pedestrians on the sidewalk to anticipate if someone might try to cross the road. GPS helps to determine the current location on the map and the destination for passenger delivery; internal sensors inform the car about its own state — how much fuel is left, if there are any malfunctions, etc. All this data is constantly processed by the central computer, which drives the car along an automatically plotted route, making decisions in emergency situations or when new information arises, such as a sudden traffic jam ahead that necessitates rerouting.

Of course, the possibility of violating the rules of the road is not incorporated into the computer. Another issue, however, is that other traffic participants can violate these rules: human drivers and pedestrians. A person is inventive, often irresponsible, and prone to instantaneous impulses—in the event of a collision with unpredictable behavior that is not provided for by the rules, the computer may freeze.

Apparently, due to the notorious human factor, for twenty years the development of robotic cars has not progressed beyond experimental models (which, by the way, always have a person behind the wheel — to take control in case of an emergency).

Another difficulty that prevents the widespread introduction of driver robots is the condition of the road surface, markings, and signs. Here the situation is approximately the same as with computer handwriting recognition: what is clearly readable for a person can be full of ambiguities, or even completely unintelligible gibberish, for a machine. Some developers suggest placing electromagnetic markings directly on the roadbed alongside the usual visual system of road signs. But all this notification must be constantly maintained in an excellent, visible, and readable state — otherwise, where a person orients themselves according to the situation, the computer may behave unpredictably, misinterpreting some instructions.

2.9 Conclusion

In reality, the history of just one passenger car alone could fill not just a chapter in a book, but an entire multi-volume collection of works. The first mode of transportation capable of carrying humans and goods without relying on muscular or draft power dates back to the late 18th century, but the conventional understanding of an automobile emerged in the late 19th century. Therefore, let us now recall the production passenger cars that played a significant role in the history of global automotive manufacturing.

Undoubtedly, the most significant role belongs to the Benz Patent-Motorwagen. It is the world's first automobile with an internal combustion engine. German engineer Karl Benz built his tricycle in 1885, and he officially received a patent for his invention on January 29, 1886. From that date onwards, the history of automobiles, as we commonly understand it, began.

In the early 20th century, two models emerged — the Cord L-29 and the Citroën Traction Avant — which became the first mass-produced cars with front-wheel drive. The American Cord was introduced in 1929, while the French Citroën made its debut in 1934. However, it was the Traction Avant that became the first widely available car with front-wheel drive. Prior to that, various experiments were conducted, and front-wheel drive racing cars were created, but they did not reach the production line.

A little later, in 1936, the Mercedes-Benz 260 D was introduced. Just like the very first mass-produced car, it was the Germans who pioneered the world's first series-produced passenger car with a diesel engine, which had been previously used primarily in trucks.

It is interesting to note that along with the Mercedes-Benz model, the Hanomag Rekord Diesel Typ D 19 A, a diesel-powered passenger car, was also showcased. However, the Mercedes-Benz 260 D was the first one to go into production. Starting from 1932, the Citroën Rosalie was manufactured with a diesel engine, but it was only installed in the commercial versions, which doesn't grant it the title of being a pioneer in diesel passenger cars.

The Jensen FF, which was produced from 1966, marked the next milestone in the global automotive industry. This British model became the world's first production passenger car with full-time four-wheel drive. Until then, all-wheel drive had been primarily associated with off-road vehicles. However, Jensen Motors decided to challenge stereotypes and created a road-going grand tourer with full-time four-wheel drive.

In 1996, the General Motors EV1 was introduced, becoming the world's first production electric car. While vehicles powered by electric motors had been around since the early days of self-propelled transportation, the American coupe was the first relatively mass-produced model with an electric motor. The Tesla Model S and Nissan Leaf came much later, becoming the first truly popular electric vehicles.

Another significant event came in 1997 with the debut of the Toyota Prius. Despite its unassuming appearance, the model became the world's first mass-produced hybrid car. The sedan, equipped with a gasoline-electric powertrain, became a harbinger of the technologies we see today. Nowadays, almost every major car manufacturer offers a hybrid model in their lineup.

In 2013, the Hyundai conglomerate made a breakthrough by introducing the Hyundai ix35 FCEV crossover. This vehicle became the world's first mass-produced hydrogen-powered car. The model was propelled by an electric motor powered by fuel cells, which generated electricity from stored hydrogen in tanks. Although hydrogen-powered technology has not yet become widely adopted, several brands have started producing hydrogen fuel cell vehicles for commercial sale.

Indeed, in modern times, automakers are primarily focusing on hybrids and electric vehicles, although traditional internal combustion engines powered by gasoline and diesel fuel will remain prominent for a long time. However, it is not unlikely that contemporary technologies will enable another significant leap forward in the coming years, opening up new chapters in the history of global automotive manufacturing.

It is true that in recent times, there has been a constant stream of news related to autonomous vehicles. Major automotive companies are actively engaged in the development of self-driving cars and associated technologies. This might lead many people to falsely believe that the history of autonomous transportation began in the 21st century. However, few are aware that the first attempts to create fully autonomous vehicles were made as early as the 1980s. For example, if one were to search the archives of The New York Times using the query "unmanned vehicles," a significant amount of material from up to 50 years ago would appear.

There are different accounts of when the first fully autonomous vehicles appeared. However, it is a fact that initially, all such developments were created for military purposes. In the early 20th century, the first research in the field of unmanned aerial vehicles began. In 1916, Archibald Low created the first drone — a radio-controlled airplane. During World War I, aerial torpedoes and self-propelled German mines were already actively used.

As we can observe, most of the technologies used in modern autonomous vehicles (radars, lidars, various sensors, satellite navigation systems, onboard computers, cameras, etc.) were developed over 20 years ago. So why haven't autonomous vehicles become widespread in our lives yet? Essentially, this question was addressed by Dickmanns in one of his works. He stated that for a self-driving car to perform all the tasks that a professional driver can do, it needs to learn how to process a vast amount of information and apply it effectively in complex driving situations. Autonomous vehicles, like any computer system, operate within rules that they can never break, while human drivers don't always adhere to traffic regulations. Therefore, the main challenge for the industry in the future will be to "humanize" autonomous vehicles by training them to interact with human drivers in conventional cars.

The emergence of mathematical modeling of traffic flows

> The car was invented as a convenient place to sit out traffic jams.
>
> —Evan Esar

3.1 Introduction

It is hardly necessary to emphasize the importance of transportation in the life of modern humans. When did the first cars appear on the roads? Despite the twentieth century being rightfully considered the era of motorization, the first cars, running on steam, appeared on roads much earlier, and not just as prototypes, but in practical use, for example, as public transport. Three steam wagons created by Gurney in 1826 for intercity communication made 396 regular daily runs in 1831, totaling 3644 miles (5683 km) [249].

However, throughout the nineteenth century, the appearance of a car on the street was an outstanding event, and horse-drawn transport remained dominant. At the beginning of the twentieth century, the automotive industry achieved significant success. The Peugeot company produced three hundred cars in 1899 [121], but they remained more of an elite consumption item than mass transport. Automakers and urbanists were still faced with many questions related to both the choice of automotive strategy in general and the integration of cars into an urban environment unsuitable for mechanical transport.

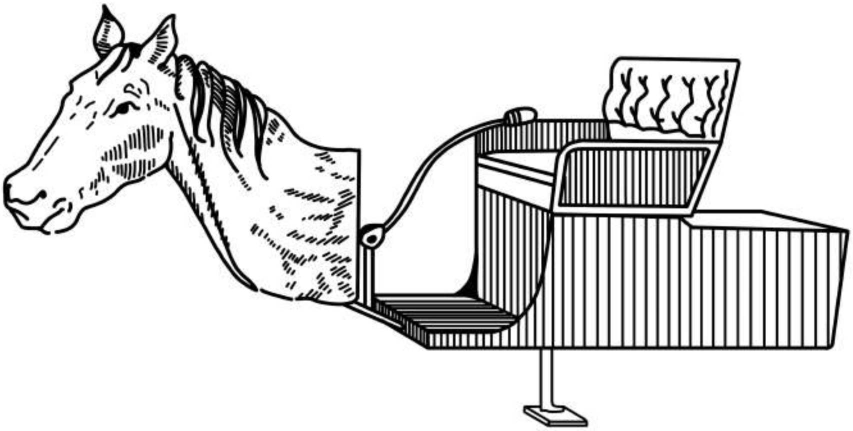


Figure 3.1: The Horsey Horseless was an early automobile created by Uriah Smith in 1899

Cartage transport was not in accord with cars on the same streets. Horses were afraid of this new type of transport, which not only looked unusual but also produced a lot of noise and strong smells. The appearance of a car on the street could provoke an accident. Automakers tried to solve this problem, for example, by mounting an imitation of a horse's face on the front of the car [249]. However, this did not solve the problem, as such vehicle models did not become common.

The issue of engine type was also on the agenda. In the first decade of the twentieth century, engineers developed and improved not only internal combustion engines but also steam and electric ones. Which type of engine would be the most convenient in operation, producing less noise and odor, weighing less, and providing sufficiently long operation without refueling? [249]

Lawyers, as well as engineers, were closely involved in the motorization issues of various countries. The first problem they faced was how to integrate cars into the legal framework of Europe and America. In 1909, the Paris Conference on Motor Vehicles was held in France, the leader in the European automotive industry. Legislators were primarily concerned with the following issues focusing on international automobile traffic: What technical requirements should a car meet? What qualifications should a driver have? What traffic rules should apply to all countries that have signed the convention? Following the conference, the convention with respect to the international circulation of motor vehicles was signed [52], introducing the first unified traffic signs and approving samples of international certificates.

On the European continent, divided into many national states, the issues of reflecting motorization in international law were quite acute. The situation was different in the United States of America. The problems of Europe, generously rugged by roads and borders, were far from the United States, a vast territory where the increasing rate of motorization required the development of a road network. In this context, the Federal Aid Road Act of 1916, also known as the Bankhead–Shackleford Act, was enacted [289]. This act was related primarily to the financing of road construction [3]. It did not concern the creation of a unified road system, which would have been too radical a decision, but rather the financing of rural road construction from federal funds. Only roads that fell under the definition of "rural post road" were financed; the rest of the roads could not receive funding, although they might have been important [51]. According to the Act, during the first year, only 5 million dollars were allocated, and in 1917 the United States entered the First World War.

In order to find solutions to most of these problems, primarily related to road construction, a serious analysis of the current situation on the roads was required. Accordingly, there was a need to develop new methods to model and analyze road traffic and forecast its future needs. Is it useful to build a new automobile road? If so, what characteristics should it have? If the construction of a new road is not possible, is it possible to modify the current road network?

What traffic mode should be chosen on existing roads and those under construction? These needs led to the emergence of a new branch of mathematical modeling — mathematical modeling of traffic flows.

The first work on traffic flow modeling appeared in the nineteenth century and was applied to railway transport by the Viennese engineer Eduard von Lill [169]. He described the structure of railway passenger traffic in the direction of Vienna-Brunn-Prague and derived his famous law on the movement of passenger transport.

In his model, von Lill used Origin-Destination (OD) matrices, referred to in some literature as "matrices of the correspondences". These are square matrices that describe the movement of objects from one transport zone to another. The mathematical dependencies developed by von Lill were subsequently widely used in the description and calculations of urban passenger flows and, due to their similarity to the law of gravitational gravitation, were called the gravitational model. In this model, the correspondence values are directly proportional to the numbers of departures from one transport region to another and inversely proportional to the distance between these regions. However, often in such models, they are inversely proportional to the exponential or quadratic function of the distance.

The gravity model shows the dependence of the number of travelers between two places on travel-related properties of the places and the travel distance:

$R = \frac{Q \cdot Z}{D^2}$,

where R is the number of travellers per unit time, Q is the "travel value" of the source location, Z is the "attraction value" of the destination location, D is the distance between locations. Measures of travel value and attraction value can be related to population numbers, household income, tourism or other services offered by the places. This rather outdated and simple travel law is nowadays replaced by comprehensive studies and statements in supply-related destination management on the one hand and demand-related travel analyses on the other.

The issues involved in urban planning were considered by the Russian scientist Grigory Dubelir [69] in the book "Urban streets and paves" (1912). This book is often considered the first publication on the subject and was mainly devoted to issues of urban planning and transport network, as well as materials for road construction. Urban Planning is the first part of this book. It describes the layout of city streets both in terms of ease of movement, and in terms of ecology and economy. In the chapter “Street network location systems”, devoted to urban planning issues, the author introduced several types of systems: radial, rectangular, diagonal, and modern. After the October Revolution, Dubelir participated in the commission of the GOELRO — the State Commission for Electrification of Russia. In 1920, the GOELRO plan developed by the commission was presented at the VII All-Russian Congress of Soviets, and Dubelir was one of the authors of the transport electrification section, and an active organizer of the entire project as a whole [111, 110].

Figure 3.2: Grigory Dubelir

The First World War, which began in 1914, accelerated the development of the automotive industry, as its development was of strategic importance. However, by the end of the war, many warring countries felt the opposite effect. For example, the military department of the Russian Empire signed an agreement on the construction of car factories designed to produce 1, 500 vehicles for military purposes in Moscow, Yaroslavl, and Rostov-on-Don. However, due to the economic difficulties of wartime and the revolutions of 1917, the project was not implemented, and moreover, fewer cars were producing in the Russian Soviet Federative Socialist Republic (RSFSR) for a long time than before the war. There were many reasons for this, but one of the main ones was the post-war devastation [150].

Traffic flow theories seek to describe in a precise mathematical way the interactions between the vehicles and their operator (the mobile components) and the infrastructure (the immobile component) [66]. Hereinafter, it is clearly shown that the development of mathematical modeling in a particular country is connected with the level of its' motorization.

3.2 The 1920s. Breakthrough in motorization

Compared with the previous period, the 1920s were a time of much more intense automobileization. This process was more intense in the USA than in European countries, devastated by the recently ended First World War. According to the U.S. Bureau of Public Roads, in 1914 in the States, $1,711,339$ cars and trucks were registered (of 2.5 million cars and trucks worldwide), and by 1919, the total number of cars and trucks had risen to $7,558,848$. At the same time, there was a more than noticeable gap with Europe, where even in France — the flagship of the automotive industry — by the end of the war, only $120,000$ vehicles were registered [213]. The explosive rates of automobileization in the USA necessitated road construction. Moreover, automobileization provoked the construction of roads, as opposed to the existence of roads allowing automobileization. This opinion was expressed by the revolutionary, party worker, and expert in the development of automobile transport, Nikolay Osinsky, in the 1920s [27]. Since 1921, engineer and politician Thomas Harris MacDonald, a road construction activist, appeared in print and on the radio, arguing that good roads are a mandatory component of human rights.

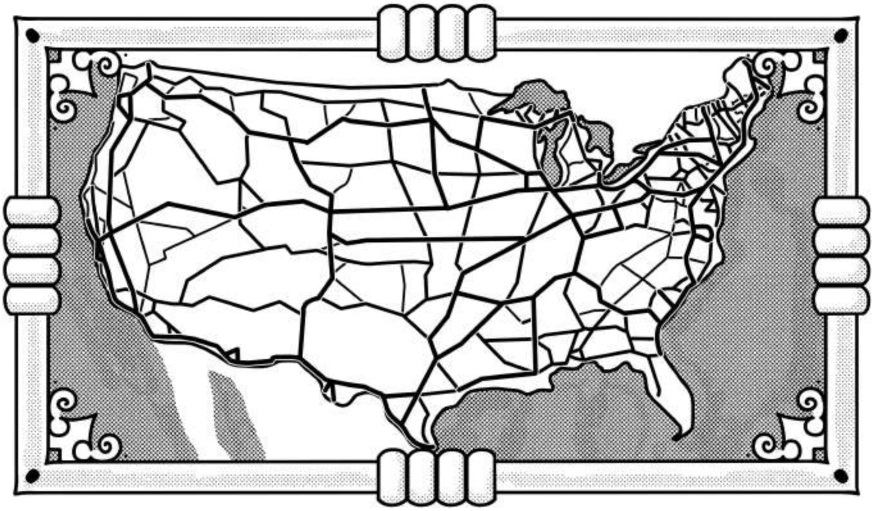


Figure 3.3: Pershing map

The "Federal Highway Act of 1921" [51] can be highlighted as the start of the national road network construction. During the realization of the project, attempts were made to consider not only civil and economic but also military needs. T. H. MacDonald, appointed Chairman of the Bureau of Public Roads in 1921, made a request to the army. At this request, General John Pershing created a colossal map of 200 thousand miles of interconnected main roads, known as the Pershing Map [279].

The absence of centralization in the existing road system defined the issue of collaboration between MacDonald and AASHO, which specified the Federal Highway Act of 1921. The proposed road network was henceforth considered not only as a postal road network but also as a highway system. This new tendency contributed to the development of new terminology such as the term "Interstate Highway System". From this moment, federal intervention in road construction became less limited from a financial perspective. In detail, the issue was to unite the highways into a stable system capable of solving problems of a national scale. The road network construction required specific works on their design, thereby setting problems that required mathematical methods for their solution.

t the end of the 20s, cities were greatly affected by the problem of motorization, which required improvements in road and traffic conditions, and led to global changes in technical standards, mainly designed for horse-drawn transport. Due to the massive increase in traffic on the streets, it became necessary to simulate traffic to study the capacity of roads and intersections, i.e., the maximum number of cars passing through a given period per unit of time, as well as to optimize traffic to improve the situation on the roads. The first attempts to study capacity and use this concept date back to the thirties. Capacity was understood as the intensity of traffic at which traffic difficulties became apparent.

In 1928, Sigvald Johannesson [143] made an attempt to determine the capacity of the road based on the average minimum distance between the centers of cars. In his opinion, the capacity of the road is reached when any further increase in traffic, other factors unchanged, causes a decrease in speed.

The specific feature of this period was also the fact that the capacity of the roads was important not only because of congestion but also because of surface problems. Asphalt-like road surfaces were at the starting point of their development, and dirt, macadam roads, and other more fragile types of pavement were quickly destroyed by motor traffic, which raised the capacity issue again.

While the USA engineers had already faced the problem of road capacity, Europe's automotive industry — growing or recovering — could not yet fill the roads to such an extent, so the road traffic agenda was not much different from the pre-war one. In 1925, the French automobile industry produced only 177, 000 cars. On April 24, 1926, the International Convention on Motor Traffic [92] was issued in Paris, signed by more than fifty states of the New and Old World, including the USA and the USSR, thereby uniting the traffic of the whole world in one legal field. In general, the participants of the 1926 conference dealt with approximately the same issues as in 1909: What characteristics should a car have to be allowed to move on public roads? What should the identification marks of a car, truck, or motorcycle be? How to distinguish one from the other, and the other from the third? What qualifications should the driver have? The 1926 Convention differs from the 1909 Convention in its elaboration, as well as taking into account the technical innovations that appeared over the past 17 years, such as car mirrors.

In 1927, the same dynamics in the growth of the number of vehicles remained: there were 27.5 million cars in the world, of which 23.4 million were in the USA, and 3.1 million in Europe [27].

Three-color traffic lights began to be used also in the twenties. The first ones appeared in 1920 in the USA in Detroit, in continental Europe in 1922, and in England in 1927. The mechanism of the driver's reaction to various light signals, especially amber, became the subject of research by many developers of mathematical models of traffic [171].

Generally, the 1920s were not rich in the development of mathematical modeling of road traffic, but the growth in the level of automobileization in this period — consistent in Western Europe and explosive in the United States — created the need for such work in the coming decades.

3.3 The 1930s. Scientific study of traffic flow

The 1930s are considered the most recognized years of the beginning of the scientific study of traffic flow.

October 24, 1929, was remembered on both sides of the Atlantic as Black Thursday, the day the New York Stock Exchange crashed and the Great Depression began. The widespread economic crisis affected not only the United States but also the countries on the other side of the Atlantic. Bank collapse, a sharp drop in the level of production (industrial goods to a greater extent than short-term products), and rising unemployment were the phenomena that set the tone for any industry of the thirties. However, in connection with our book, not so much the causes of the Great Depression or the mechanisms of its deployment are important, but rather the ways in which the states tried to cope with what was happening.

Franklin Delano Roosevelt, the elected President of the United States in 1933, tried to solve the problems of the Great Depression through integrated macroeconomic regulation of a market economy. In addition to regulatory methods such as setting a minimum wage and a maximum working day [201], the government implemented a number of government order programs. Public works were applied in many areas, such as housing construction, land reclamation, and, in particular, road construction [94].

The Roosevelt government believed that the construction of highways would not only provide jobs but also help revitalize the country's economic life as a whole. Therefore, despite the economic downturn, the total number of highways grew (from $193,049$ miles in 1930 to $214,000$ in 1935), while the share of financing from the state fell sharply (from $\$136,681$ thousand in 1930 to $\$24,307$ thousand in 1935), and the financing from the federal level increased (from $\$99,839$ to $\$218,112$ respectively) [43]. In 1938, a law on federal highways was passed under the name of its author Thomas Harrison MacDonald. The experience of German autobahns was taken into account when he designed this project. In 1939, Fairbank and MacDonald issued a report entitled "Toll Roads and Free Roads," the first formal description of what subsequently became the interstate network. To further develop this promising concept, President Roosevelt appointed a National Interregional Highway Committee led by MacDonald and Fairbank. The resulting report, Interregional Highways, proposed a $39,000$-mile highway system, three-fourths of which was rural [270]. However, construction began only in the forties due to the US entry into World

War II [84].

Due to the general economic downturn, in the 30s, mathematical modeling developed more rapidly than road construction. The first model to be mentioned is the model of Bruce Douglas Greenshields [112]. Greenshields used the latest technical achievements. In 1933, he used photography with complex shutter speed settings to measure the speeds of individual cars and the distances between them. The pictures were subsequently superimposed with vertical division, so the exact vehicle distance could be measured. This supplement made the obtained images discrete. Apart, Greenshields took an effort to provide space that is not distorted during photographing. Especially inventive was the technique to avoid the car merging with the background in the photo. A wide white canvas was stretched along the road against which the dark car was always clearly visible [155].

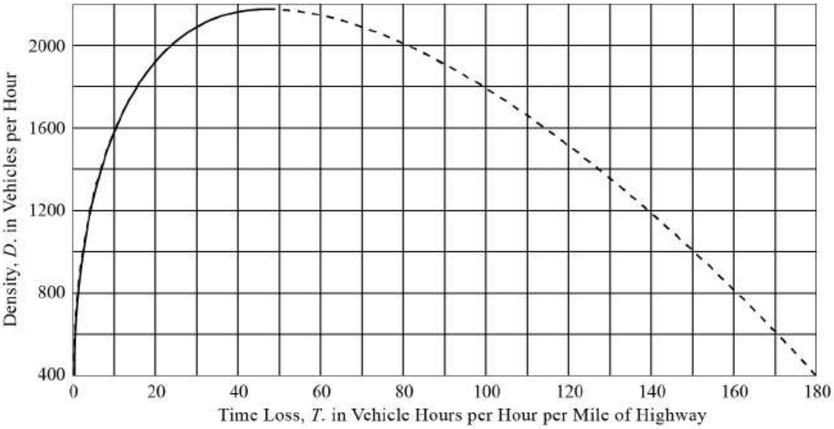


Figure 3.4: Time lost by vehicles due to reduced speed on a congested two-lane highway

Greenshields approximated the dependence of the distance between cars from speed as a linear function $S = 6.9 + 0.226v$, where the distance is expressed in meters, and the speed in kilometers per hour [112]. The number 0.226 in this formula later became interpreted as the reaction time of the driver, and the Tanaka model [137] widely uses it.

However, neither the Johannesson model nor the Greenshields model took into account such an important traffic variable as pedestrians. Regarding traffic safety, the issue of pedestrians was very hot in the thirties. According to National Safety Council data, in 1937 the mortality rate in road accidents reached

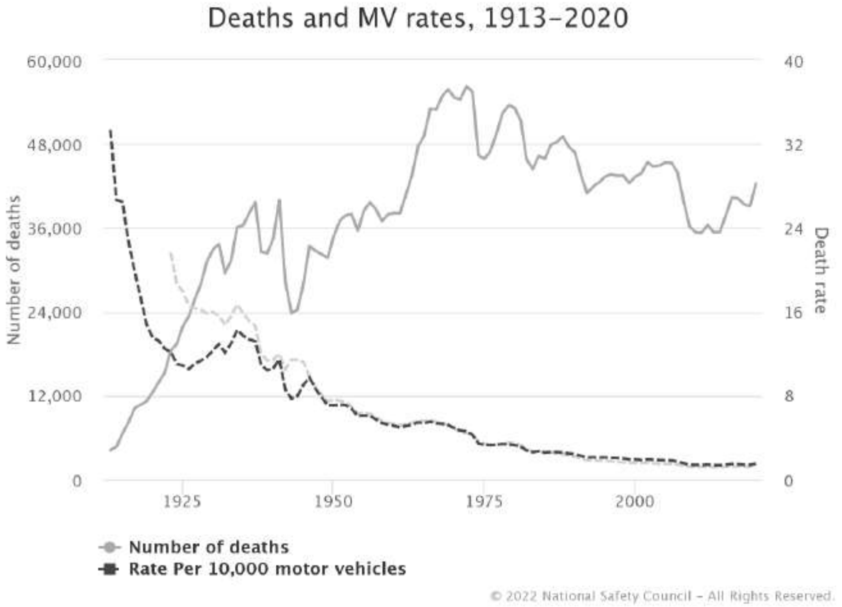


Figure 3.5: Motor-vehicle deaths and rates, United States, 1913—2020 (National Safety Council Data [2])

a peak for the entire twentieth century — 30.8 deaths per 100, 000 population [57]. Perhaps this problem influenced the emergence of mathematical models that take into account pedestrians.

The first to address this question was British scientist William Frederick Adams. In his work "Road traffic considered as a random series" [4], he derived an expression for the mean delay to a pedestrian caused by a stream of traffic on a main highway, in which the successive gaps (in time) are independent random variables, each with a negative exponential distribution. It was assumed that each pedestrian waits for a gap which is larger than a prescribed safe interval before attempting to cross. Adams also suggested that road traffic could be represented by a Poisson distribution.

A similar problem was later considered by Garwood, although he expressed it through the movement of a single vehicle [103]. This distribution was later called the Garwood distribution.

Despite the economic difficulties, the thirties were a time when the number of individual cars began to grow rapidly in Europe. In the countries of the first tier of industrialization (Great Britain, France), the number of cars doubled. At

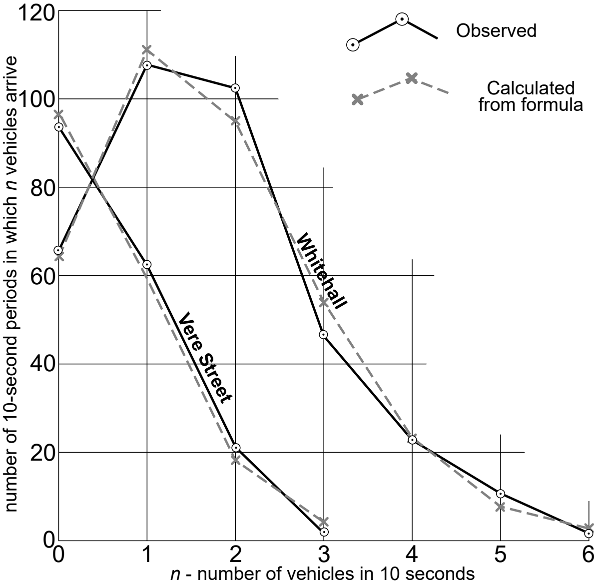


Figure 3.6: Distribution of numbers of vehicles arriving during periods of 10 seconds [4]

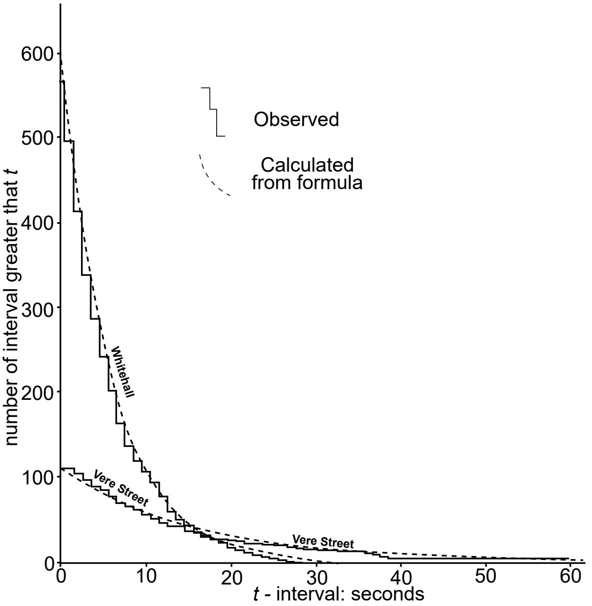


Figure 3.7: Distribution of intervals elapsing between arrivals of successive vehicles [4]

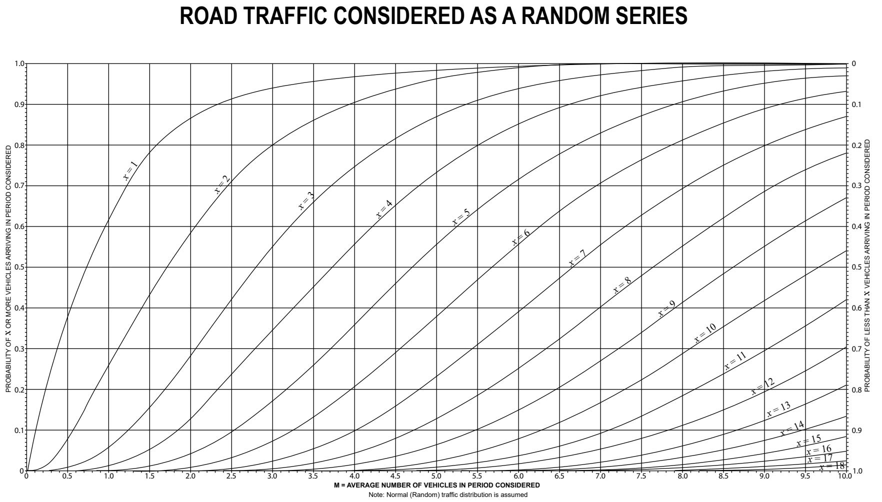


Figure 3.8: Probability that X or more vehicles will arrive in a given time using the figure can be calculated as follows: Through the value of m (average amount of passing vehicles throught this road) on the base line a vertical line should be drawn to intersect the curve for the given value of X. From the point of intersection a horizontal line must be drawn, and the required probability may be read off where this line cuts the left-hand scale [4]

the same time, traffic congestion in US cities started to cause concern for the chief of the Bureau of Public Roads. The inadequacy of urban planning for heavy motor traffic led to a drop in quality of life and a decrease in real estate prices in the city center. In general, these phenomena manifested themselves in the sixties, but Thomas Harris MacDonald had already preached it in the thirties. As mentioned at the beginning of this section, this decade is considered the start of the scientific study of traffic flows. Regression methods (used by Grinshields) and queuing theory (Poisson processes, used by Adams) are still widely used in this area. Moreover, the attempts made by Adams would be proven in subsequent decades.

3.4 The 1940s. The World War II decade

The Second World War, which began in 1939, strongly affected the pace of road construction in the United States and even more so in European countries. Funding for road construction from all sources decreased by about two and a half times. If in 1940, 11, 549 highway miles were built, only about a thousand miles less than in 1935, then in 1945 only 3, 035 miles were built [43].

However, in general, World War II had a positive effect on the American economy. A number of researchers believe that participation in the war really allowed the USA to cope with the Great Depression. This enabled the full implementation of the project developed by Thomas MacDonald [279]. Work began in 1947 and yielded tangible results: from 233, 772 miles of highways in 1945, their total length grew to 640, 753 by 1950 [43].

In 1946, a year before the start of large-scale work on the interstate highway system, the first mass studies were conducted on the issue of driver reaction time. For example, the Ohio State Highway Administration conducted research on over 1, 000 people to determine the effect on stopping distance of the time taken to apply the brake at different speeds after the driver detected danger. The average reaction time for men was 0.57 seconds, for women 0.62 seconds. It was justifiably assumed that under normal conditions the reaction time would be longer since, during the tests, the drivers could foresee the appearance of danger. The author of these studies was Harry Neal [195].

The post-war world also needed a review of international law. The Paris Convention of 1926 was already quite outdated, and it was necessary to create a more complex, comprehensive legal system that would better cope with the changing technological situation and, in general, greatly develop the system of road law. The Geneva Convention on Road Traffic was adopted on September 19, 1949 [275]. A very important change was the fact that the Geneva Convention on Road Traffic paid much more attention than previous documents to driver behavior on the road [303]. In general, it can be said that the Geneva Convention on Road Traffic had seriously expanded the agenda of international road law. However, some of the steps taken within its framework were not effective enough. The Convention had been ratified by 80 countries, but some participating countries might not have ratified individual annexes and protocols. For example, only 34 countries that did not include the USSR, the United States of America, and Great Britain signed the protocol on road signs. The post-war world was in a condition of the greatest variety of road signs in the entire history of the twentieth century. The symbolic system of signs was adopted in the USSR and continental Europe, textual in the USA, Australia, and New Zealand, and mixed (texts and symbols at the same time) in the UK and some countries of Asia and South America. This diversity existed before the Vienna Convention on Road Traffic, adopted in the late seventies. But while Europe was overcoming post-war devastation, and the United States was emerging from the Great Depression through military loans, the international rules of the road issues were not paramount.

Despite the fact that the United States' post-war economy developed much more dynamically than Europe's, the boom in mathematical modeling occurred on both sides of the Atlantic at the same time — in the next decade.

3.5 The 1950s. Macroscopic and microscopic models in traffic

The fifties were really explosive in terms of road construction, road traffic, and its modeling. In the United States, four times more highways were built in one decade than in the twenty pre-war years. In 1950, 6, 666 passenger cars were sold. Compared to 1945, gasoline consumption almost doubled in 1950 and continued to grow .

Regarding the mathematical models of traffic flows, during the fifties, two main families of traffic modeling were identified for various tasks: microscopic and macroscopic. In this decade, the United States and Great Britain remained leaders in mathematical modeling.

In 1951, Goodman considered the Erlang distribution of intervals between vehicles and obtained a formula for the distribution of the number of vehicles in synchronous counting, called the generalized Poisson distribution, further investigated in the works of Whittlesey [292], Haight [116], and other scientists. Usually, the Erlang distribution is used in cases where the process can be represented as the sum of elementary sequential components distributed according to the exponential law. In the same year, Tanner, studying the Garwood distribution, obtained a new method that made it easy to transition to the general case [269], which Mayne [181] did in 1954—1958, providing the appropriate formulas for an arbitrary traffic flow on the main street.

In 1952, the English scientist John Glen Wardrop (1922—1989) presented his two equilibrium principles relating to the concept of Nash equilibrium from game theory, developed independently of each other [286].

However, in transportation networks, there are multiple actors involved, which makes the analysis complex. In a transportation network, Wardrop's principle [286] states that each user behaves in a rational and selfish manner. Rational because they evaluate alternatives and selfish because they choose the best route for themselves. Their goal is to minimize the generalized cost of the journey. Therefore, each user will switch to an alternative route as long as the generalized cost of that route is lower than the initial generalized cost.

1. The first principle aligns closely with the ideas expressed in the 1920s by Frank Knight [152] (1885—1972), who co-authored with Wardrop. Its

statement is as follows: *"For each origin-destination pair, the used paths have the same generalized cost, and this cost is lower than the costs of unused paths. If multiple routes are used for an origin-destination pair, their generalized costs are equal."* This principle is also known as "user equilibrium (UE)," "selfish Wardrop equilibrium," or simply "Wardrop equilibrium."

2. The second principle is also well-known as "system optimal (SO)" or "social Wardrop equilibrium" and can be formulated as *"optimum is reached only with the joint efforts of all participants in the flow"*. For example, this principle supports centrally controlled vehicles.

In both cases, each driver assumes that their individual influence is so small that it will not have any impact on the overall traffic situation. We will focus on the equilibrium of individual cars. A simpler formulation of this principle would be: *"the travel time between two points will be the same regardless of the chosen route."* Indeed, if there is a route with a shorter travel time, some cars will use it, and eventually, a new equilibrium will be established.

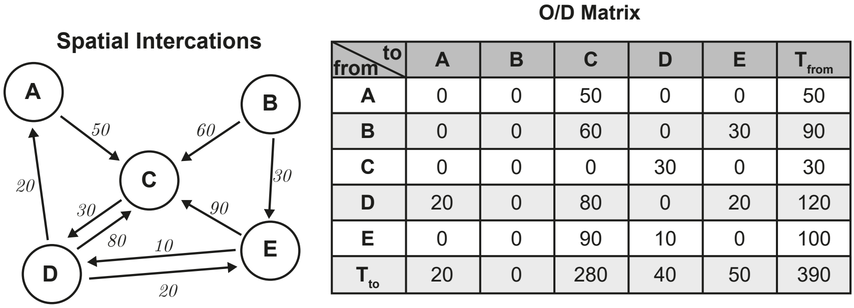
Up to now, these principles remain one of the simplest and most understandable when describing the propagation of moving objects over a transport network. Time for a trip on all the routes currently used will always be no more than the time for a trip along unused routes; each of the participants in the stream independently of the others at any time tries to choose the most optimal trajectory. When using physical analogies, then the transport flow model can be described by the viscous fluid propagation model. However, even with clear hydrodynamic analogies in this conception, the first hydrodynamic model appeared later in this decade.

During the same year of 1952, American scientist from Brown University, Martin Joseph Beckmann, formulated and developed a model that implemented Wardrop's principles [24]. The model was based on the assumption of the existence of a functional monotonously increasing dependence between the intensity of the traffic flow on a certain section of the road and the travel time along it. The model did not limit the throughput of roads in any way, it allowed arbitrarily large intensities of movement along it, introducing only an additional delay in such cases. The model did not describe traffic congestion in any way, that is, situations in which the intensity and flow rate are small. However, Beckmann's work greatly influenced transportation science. He considered the problem of mathematical modeling of traffic from the point of view of economics, being engaged in the problems of transportation and location of production, which was extremely important for the dynamically developing American economy

Figure 3.9: Martin Joseph Beckmann

[29].

As mentioned before, the construction of the Origin-Destination (OD) matrices is one of the central and most complex tasks in all studies related to significant structural or parametric changes in urban transport traffic. With all the diversity of approaches to the formation of such matrices, one can clearly distinguish between them into two large classes: extrapolation and probabilistic.



O/D Matrix

from \ to	A	B	C	D	E	T_{from}
A	0	0	50	0	0	50
B	0	0	60	0	30	90
C	0	0	0	30	0	30
D	20	0	80	0	20	120
E	0	0	90	10	0	100
T_{to}	20	0	280	40	50	390

Figure 3.10: A conceptual diagram of an origin-destination matrix

Extrapolation methods are based on the use of survey data of the existing state of distribution of passenger and transport flows between the corresponding areas, using proportional growth factors for predictive calculations. In the early 50s, this methods had become widespread. The Detroit method was first applied in the design of the highway system in 1953 and, along with the use of growth factors for individual districts, used the city's growth rate as a whole [74, 60]. Unfortunately, the prediction by using this method was not always reliable, and the calculation error would increase if city growth rates differ from individual areas of growth.

In 1954, Thomas J. Fratar developed an extrapolation iterative method for OD matrices generation. Here is a description of Fratar in an article describing his method: *"He led numerous research in the field of planning for projects of motorways, airports, sea and railways. During World War II, he served the U.S. Army as an engineer for evaluating the operation of the Iranian State Railway. Prior to that, he was associated with the Yale Automobile Transport Bureau of Scientific Research related to automobile transport. The traffic distribution method described in this article was used to analyze traffic in the Cleveland metropolitan area for Kuyahoga County, Ohio"* [97]. A single step in the Fratar balancing provides input data to the next one, until equality between the counted value in advance turnover transport area and the calculated amount of correspondence. This method is the most prevalent among the numerical extrapolation methods of calculating the OD matrices [261, 228].

In 1954, Bailey considered an ordinary system of group queuing theory [19], which had a great influence on Bisi, who proposed to consider the passing of road intersections in terms of this theory [250]. It is a system with a Poisson input stream and variables bounded above by the size of a group of applications. Using the method of generating functions, the expressions could be obtained for the average value and variance of the number of customers in the queue for the Markov chain, nested by the moments immediately preceding the moments when the groups of applications were finished [133]. This model was more realistic than the model with two discrete parameters proposed by Beckmann, McGuire, and Winsten later in 1956.

In 1955, the British scientist from the University of Manchester, Sir Michael James Lighthill and his student Gerald Beresford Whitham developed the first macroscopic mathematical model of transport flows, that is, the model considering the flow of vehicles as a whole [167]. The same conclusions had been independently reached by Paul I. Richards [246], therefore the model is often called the Lighthill-Whitham-Richards (LWR) model [104]. The LWR model has taken a step from the static functional dependencies of traffic flow parameters to a description of their dynamic relationship in time and coordinate. In

Figure 3.11: Michael James Lighthill

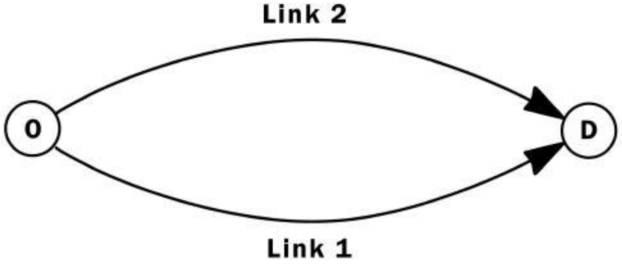
these papers, the motion on a one-way infinite road from the point of view of hydrodynamics was considered. The authors showed that transport processes in continuous media are a suitable tool for modeling congestion. Models with two-lane roads began to be developed only in the 60s.

In 1955, the work of Daniel L. Gerlough and André Schuhl [108, 254] confirmed the hypothesis of Adams and it became the catalyst which increased the use of the Poisson distribution in the theory of traffic flows.

In the same year, France adopted a law on the establishment of a system of state toll roads, which was supposed to provide the country with a system of expressways. After the implementation of the project, almost every resident of France should live within less than an hour's trip from the expressway. The construction of the system began in 1961 [119].

Beckmann, McGuire, and Winsten in 1956 [23] formulated and analysed the first model of origin-destination flows (demand) and user-equilibrium route flows for a congested road network. The concept of the equilibrium distribution of traffic flows is as follows: in a state of equilibrium no one can change his way so that the "price of the trip" is reduced and, thus, no one has the motivation to change his path. The Beckmann model main idea is that the solution of the minimization problem is the Nash-Wardrop equilibrium. The solution of this mathematical programming problem can be facilitated due to the Beckmann transformation [259].

Beckmann introduced mathematical transformations, that satisfy the condition of maintaining equilibrium in a network connecting the nodes.



In the example shown in the figure, there are two paths – Link1 and Link2. Delay times associated with transport distribution are represented by the formulas $t_1 = 2 + x_1$ and $t_2 = 1 + 2x_2$. Let the $O \to D$ link have 5 flow units, i.e. $x_1 + x_2 = 5$. Then the equilibrium conditions can be expressed as $t_1 \leqslant t_2$ if $x_1 > 0$ and $t_1 \geqslant t_2$ if $x_2 > 0$. Then, under the condition of non-zero flows, it is easy to obtain that $t_1 = t_2$, from which the flow and time variables $x_1 = 3$, $x_2 = 2$, $t_1 = t_2 = 5$ are calculated.

If we set the minimization problem as follows:

$$\min z(x) = \int_0^{x_1} (2 + \omega) d\omega + \int_0^{x_2} (1 + 2\omega) d\omega$$

under the conditions $x_1 + x_2 = 5$, $x_1, x_2 \geqslant 0$, then its solution will give the same result. What does these integrals mean? Just the total time spent by drivers.

The year 1956 was just as important in the context of road construction as in this year President Dwight D. Eisenhower signed the Federal-Aid Highway Act of 1956 on June 29 [91]. The project involved the creation of a system of interstate highways, which was described in the publication “General Location of National System of Interstate Highways,” also known as the Yellow Book [241]. Charles Erwin Wilson, known as “Engine Charlie,” who led General Motors for a long time, assisted Eisenhower in the planning [99]. Most of the highways had

Figure 3.12: Charles Erwin Wilson

to meet the standards necessary for the highways to fulfill their purpose, thus, to provide high-speed traffic. In this regard, the movement of slow-moving or oversized vehicles was prohibited or restricted on highways. Also, according to the introduced standards, traffic lights were not used on highways, and those areas where traffic lights were still used were considered non-conforming to standards. On the highways, the upper speed limit was implemented, which, however, increased. In addition to the economic and social significance of the project, one should not overlook the fact that the Dwight D. Eisenhower National System of Interstate and Defense Highways was designed with military needs in mind. Eisenhower, being the commander of the expeditionary forces of the Allies in Europe, drew attention to the German autobahn system as a strategic object and, when developing the interstate highway project, paid great attention to its possible use for military traffic [183].

The fifties were the time when big business contributed to the history of mathematical modeling.

The American giant, General Motors Corporation, in its research laboratories located in Detroit, began research on modeling traffic. The group of scientists developing mathematical models of traffic for General Motors consisted of Denos C. Gazis, Robert Herman, Renfrey B. Potts, and Richard Rotheri. In

1959, the first three scientists mentioned above proposed one of the first microscopic models of a single-lane transport stream, which allows obtaining a fundamental diagram — the relationship between the intensity of the flow of vehicles and the density [107]. Microscopic mathematical models, in which vehicles are treated as separate elements, are used mostly for single vehicle movement predictions and for modeling on subsystems, whereas macroscopic models are used in global optimization problem solutions.

In addition to the traffic research in the laboratories, General Motors Corporation in December 1959 held the First International Symposium on the Theory of Traffic Flow. In the same year, a group of scientists formed on the basis of Research Laboratories of General Motors Corporation won the Lanchester Prize for Operations Research Award for three articles on traffic flow research, including the article mentioned above [35].

During the fifties, the processes that Thomas Harris MacDonald investigated in the 30s were accelerated in their development. The obstructed traffic in the city center contributed to the new reality, which further the urban researcher Jane Jacobs would call the death of American cities. The declining real estate prices in the city center and the relocation of the middle class to the suburban regions led to the central areas of many cities beginning to turn into slums. The need to solve this issue boosted the need for mathematical modeling of traffic, including at the micro level.

In conclusion, the fifties are characterized by the allocation of two main vectors of road traffic modeling development — hydrodynamic and microscopic. The development of new methods for Origin-Destination matrices calculation complemented these methods and allowed predicting the change of customer preferences as well as their number. The confirmation of Adams' hypothesis of the Poisson nature of traffic flows clearly separated a new approach to modeling transport processes, based on the queuing theory, which is popular nowadays, especially when modeling traffic at controlled intersections.

3.6 The 1960s. Computer simulation

Sixties continued trends inherent in the previous decade. The number of motor vehicle users continuously grew. In 1960, there were 87 million driver licenses in the USA and this number continued to grow, however the production of cars in 1960 remained at the same level of 1950. This could be explained both by significant imports, and long-life operation of cars. Construction of roads become more intense since 1956, when the project of the construction of interstate highways system was launched. The length of highways built during the year grew, and the amount of money spent on construction by the federal authorities grew even more: from \$ 756, 926 thousand in 1955 to \$ 2, 272, 518 thousand in 1960. Including state investments, in 1960, \$ 3, 264, 033 thousand was spent on the construction of the highway system [43]. Mathematical modeling of traffic had also been actively developed. There had been more proposed models and studies related to traffic in this decade than in the previous one.

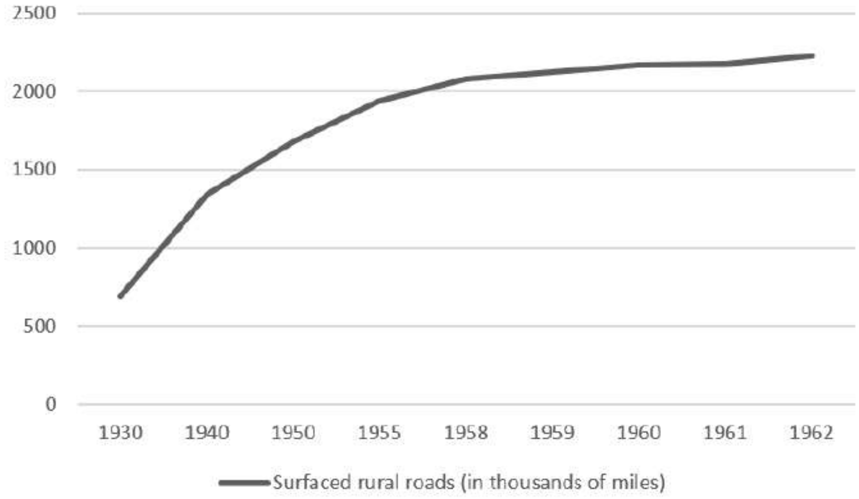


Figure 3.13: Road length statistics

In general, the number of motor vehicles, both private and state-owned, was growing worldwide. From 1955 to 1967, the number of cars owned by the UK increased from 4, 956 thousand to 12, 487 thousand, which is two and a half times, and the number of cars owned by France increased from 3, 700 thousand to 12, 430 thousand, which is a little more than three times. In total, in 1968 in the UK, there were 2, 225.2 thousand units of vehicles, and in France, in the same year, there were 2, 075.6 thousand vehicles [122].

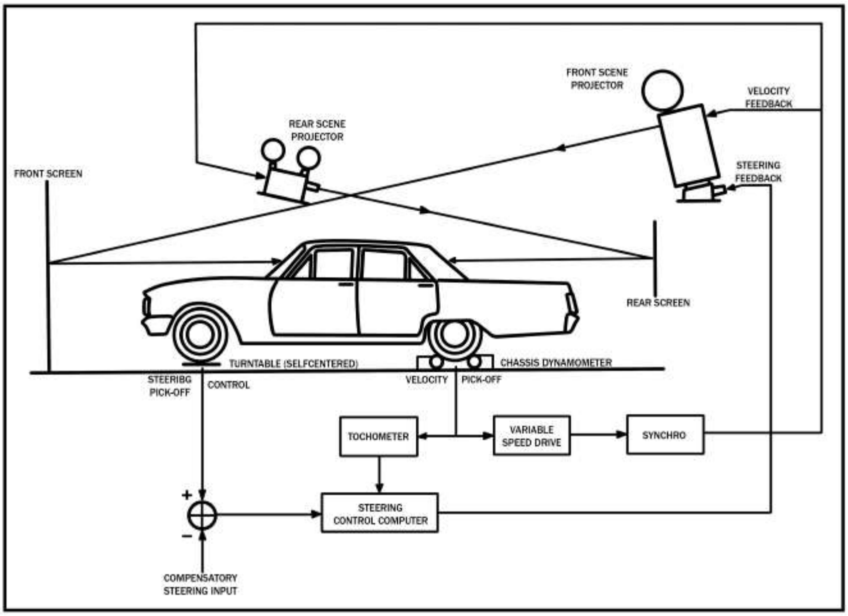


Figure 3.14: Block diagram of driving simulation laboratory

In 1960, three important studies from different areas were released. The first was the work of Slade Hulbert and Charles Wojcik, who worked on the first proposal to use driving simulators [131]. The car simulator should include a real car, the rear wheels of which were supported by steel dynamometer rollers. This simulator was provided by two films, one of which depicted what happens in front of the driver, and the second showed what was going on behind him. Depending on these images, the driving was performed. When turning the steering wheel, the frames of the front film should move, the engine noise should be real, and with the help of flywheels placed on the rollers, the inertia of the car should be reproduced, which was supposed to give the impression of an engine and a change in speed. Based on the technical base of those times, these studies yielded very modest results. Most of such works at the time that were devoted to the simulation of automobile traffic concentrated on questions of methodology, the order of research, and potential applications.

The second study was the work of Gordon Newell, relating to one of the most important areas of application of mathematical modeling of traffic, namely location modeling and traffic light programming. He proposed a model for the sequential arrangement of traffic lights [200], which provided an overall average speed of vehicle traffic.

The third work was also devoted to the problem of traffic light optimization, which in this decade became actively developed. Perhaps this was due to an increase in the technical capabilities of traffic control. For example, in Moscow until 1956, there were traffic lights that required the direct participation of a policeman to switch them [171]. Denos Gazis, Robert Herman, and Alexei Maradudin continued to work at the research laboratories of General Motors. Their research was devoted, among other problems, to the duration of the amber traffic light [106]. The goal of the amber phase of the traffic light is to make cars behave as if the traffic light in front of them is red, but if the stop cannot be safely made, the car can carefully drive through the road intersection. Confusion often arises due to incorrect use of this signal. If this phase is too long, drivers continue to use it as a green traffic light, and if it's too short, they treat it as red. When the duration of the amber phase is correctly matched, this interval performs both necessary functions. To solve the problem of perceiving the amber signal as a call to move faster, Drew suggested using a countdown traffic light.

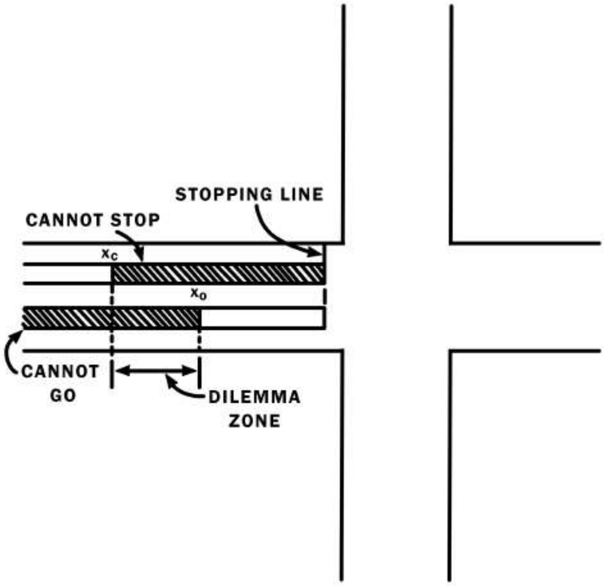


Figure 3.15: Schematic diagram showing the "dilemma zone" near an intersection

The studies proposed in the article showed that of approximately 70 intersections studied, only one had a correctly selected duration of the amber phase. This study shows that at the research laboratories of General Motors, not only beneficial and useful research for the development of engines was conducted. The research goals of some studies went beyond the direct interests of the financing corporation.

In 1961, Newell continued his research. He proposed a microscopic mathematical model (in which vehicles are treated as separate elements) [199], named after him, which is one of the first non-linear models of optimal speed. For each driver in this model, a certain “safe” speed depending on the distance to the vehicle ahead (the leader) can be obtained.

In the same year, Ilya Prigogine, the Belgian scientist of Russian origin, for the first time described the kinetic equation for traffic flow [234]. Unlike hydrodynamic models, kinetic models are based on the description of the dynamics of the phase flow density and arise with the aggregation of molecular dynamics. This means that the density of the distribution of vehicles is studied both in coordinate and speed. This model arose with the aggregation of molecular dynamics, that is, a model that describes each particle separately. The transport flow in this model is described by a kinetic equation of the Boltzmann type, in which, instead of the “collision integral of gas particles”, the “vehicle interaction integral” is used. As the Euler equations are derived from the Boltzmann equations, so the macroscopic model in this approach is derived from the kinetic model.

1963 was marked by two very important events — the publication of the first generalizing work: “Mathematical theories of traffic flow” by Frank Haight and the first major project on computer-based traffic management in the city of Toronto.

The book “Mathematical theories of traffic flow” by Frank A. Haight [115] from Royal Institute of Technology of Stockholm, was one of the first books systematizing knowledge of transport flows. Haight examined the modeling of road traffic flows and placed upon the general theoretical framework for this purpose, based on the theory of probability and queuing theory. Through this book, the mathematical modeling of traffic flows emerged as an independent branch of applied mathematics. Nevertheless, this book was a bit archaic, as it contained already outdated methods of modeling traffic flows, not including detailed descriptions and mathematical apparatus of existing and actively developed hydrodynamic and kinetic models.

At the same time as systematization, large-scale works on the practical application of the accumulated knowledge were conducted. In Toronto, Canada, the first large traffic control system was created, which included a computer [139].

Customers were made by Corporation of Metropolitan Toronto, the work was performed by the American company Traffic Research Corporation, New York, USA, and the author of the report on the work done was an engineer and researcher Neal A. Irwin. He wrote that a series of quantitative experiments were carried out in 1960—1961 when developing the control system to find out what type of traffic control minimizes vehicle delay. In the pilot project, the area under management was 1.7×0.5 square miles, with 16 intersections with traffic lights. The system used 39 detectors, indicating the intensity of movement and other important management indicators. The project was recognized as successful, and the Corporation of Metropolitan Toronto decided to extend the system to the entire capital region, leaving Traffic Research Corporation as a technical consultant.

In 1964, more than a hundred intersections with traffic lights came under the control of the system, and this number continued to grow. Such computer systems designed to control traffic in particularly busy areas appeared in 1965—1969 in many cities, including Tokyo, Munich, London, Madrid, and New York [25].

Since 1970, Japan has become one of the leading countries in traffic control automation, where these systems have been installed in all major cities of the country, and a similar system has been created for the Japanese expressway network.

Perchonok and Herst in their articles from 1964 to 1965 continued the vector begun in the forties studies of Neal [223, 222]. They were among the first to introduce the driver error factor in modeling movement. Hurst argued that a normally thinking driver should and will take risks. "Advice to the driver not to risk is equivalent to an offer not to travel on the road. Since the driver does not intend to adopt such a recommendation, he must learn to distinguish acceptable risk from unacceptable. Many road accidents are the result of an incorrect risk assessment by normal emotionally stable drivers with satisfactory motivation" [66].

By the sixties, it was clear that the Detroit method of forecasting was not always reliable, and errors in calculation would increase if the growth rate of the city differed from the growth rate of individual districts. In 1966, the work of the British Alan G. Wilson appeared, who worked at the same time at Rutherford Laboratory, University of Cambridge, and Oxford University [295]. Wilson's papers were based on entropy methods in mathematical modeling, including, among other things, calculations of the OD matrices. Entropy methods are based on the theory also used in statistical physics — entropy is a measure of uncertainty of the system. The main hypothesis of his model was an assumption that in a macro system, the equilibrium state is reached at the maximum of its entropy under the conditions of the finiteness of the resources contained in the system. Different variants of these conditions lead to problems of mathematical programming that differ in complexity [294, 293].

The advent of computers has allowed the production of complex numerical experiments using simulation processes, and, most importantly, it became possible to take into account the random nature of traffic flows. Simulation is usually necessary in cases when the studied systems cannot be analyzed using direct or formal analytical methods.

In 1967, apparently, the first computational mathematical model of traffic flow was created by Drew, Meserolo, and Buhr [67] from Texas Transportation Institute, Texas A&M University College Station. The authors studied the process of leaving the highway from an adjacent entrance.

This particular stage of road traffic was chosen, apparently, not by chance — in Drew's further work, he discussed the growth of road accidents at the entrance to the highway. The study was fairly general in nature. Authors provided the possibility of changing the input parameters of the structure of the highway, which made it possible to simulate the operation of various highways [66]. Variables such as the number of lanes, the length of the highway segment, the number of entrances and exits, their location and length, the length of each acceleration section, the beginning of the slope, and the slope of the road were studied. Initially, each vehicle was assigned parameters such as its starting point, actual speed, and required speed. The movement of vehicles was carried out according to the authors' logic, essentially setting the micromodel of traffic.

Martin Joseph Beckmann, mentioned earlier, continued his activities. From 1967 to 1987, he was a consultant at the General Motors research laboratory and participated in research with Denos Gazis, Robert Herman, and Renfrey Potts. Around the same time, he worked as an editor for the Transportation Science journal. He published numerous articles related to transportation problems, their calculations, and their impact on the economy. Many researchers note Beckmann's activities, particularly his decisive influence on the fact that mathematical modeling of road traffic emerged as a separate area of applied mathematics [29].

In the late sixties, the first Japanese studies on mathematical modeling of road traffic appeared. In 1967, Japanese specialists from Tokyo University, Inose and Hamada, together with Professor Fujisaki, published an article, "Theory of road-traffic control based on macroscopic traffic model" [136]. Apparently, this article was the first to explicitly express the concept of macromodelling in road-traffic control. The model ignored random factors in human behavior and vehicles and assumed a saturated traffic flow, that is, the traffic flow was always equal to the capacity of each intersection.

Since the end of the fifties, Japan had experienced a period called the "Japanese economic miracle" in historiography. The priority in the development of Japanese industry shifted from labor-intensive light industry, which required a lot of raw materials and resources, to heavy industry, in the direction of high-tech areas of production such as electronics and the development of new materials. Coupled with the active adoption of the American methodological framework, this development vector of the Japanese economy created ideal conditions for the development of mathematical modeling of traffic and the practical application of these developments. Since 1970, Japan had become one of the leading countries in automating traffic management, where systems similar to the system in Toronto were installed in all major cities of the country, and a similar system was created for the network of Japanese high-speed roads [104]. The need for the introduction of such systems was evident in statistics — in 1968, there were $2,055.8$ thousand passenger cars in Japan, almost as many as in Germany ($2,862.2$ thousand), and the number of buses and trucks in Japan even surpassed the values in the United States. The total number of vehicles in Japan ($4,085.8$ thousand) was twice the number of vehicles in the UK ($2,225.2$ thousand) or the number of vehicles in France ($2,075.6$ thousand). At the same time, the market share of Japanese cars, not imported from America, grew. Japan developed its own automotive industry, and the introduction of advanced technologies allowed a reduction in labor and time spent on the production of one car by about $5-6$ times [122].

At the end of the sixties, one of the largest conferences devoted to the issue of road traffic was held — the Vienna conference of 1968 [194]. Its decisions, together with later additions, continue to be valid. It operates in the vast majority of countries in Europe and in many countries in Asia, Africa, and the Americas. The Vienna Convention was an indicator that international road legislation had undergone many changes. International road law had been further expanded even in comparison with the Geneva Convention of 1949 and dealt with many issues affected by mathematical modeling of road traffic, such as the relationship between driver behavior and the occurrence of congestion. However, the

Convention almost did not address issues related to traffic management (standards of road planning, traffic lights, etc.). In fact, the Convention did not place these issues on the agenda. The Convention did not come into force until the late seventies, as it was amended in 1971 and, in connection with Article 47, Paragraph 1, the Convention becomes effective twelve months after ratification [194].

The publication of a major generalizing work by Donald Drew "Traffic Flow Theory and Control" [66], which is one of the largest monographs on the subject, completed the development of mathematical modeling of traffic in the sixties. This book considered both physical and psychological features of vehicle movement, as well as various mathematical models of traffic flows, including an energy approach to the service level. The problem of visibility and illumination on roads with different curvature was also studied in detail, apparently posed by the author himself.

The author assessed the effectiveness of the road-street system according to three main criteria: traffic intensity, average speed, and the number of road accidents. All these data needed to be collected and processed in terms of mathematical statistics. Control calculations on the main streets should be made at least once every two years, cars should be divided according to the direction of movement, and the minimum duration of the calculation was 24 hours. For counting, different special tools such as manual, pneumatic, photo, sound, and infrared counters were used. When designing the road, the author proposed using the estimated hourly traffic flow for a given year, which forced consideration of the methods of predictive modeling, which at that time were elementary. A significant part of the monograph was devoted to all possible distributions, analysis of their interactions, as well as their applications in mathematical modeling of traffic flows. One of the chapters was devoted directly to computer modeling of traffic flows.

The result of the sixties for mathematical modeling of road traffic was primarily that it emerged as a separate area of applied mathematics. Some associate this with the appearance of the work of Haight [115], others with the activity of Beckmann [29]. However, the general research vector was new technical means that allowed the transition from purely mathematical methods to simulation modeling.

3.7 The 1970s. Systematization and expansion

In the seventies, high-speed highway systems projects that were adopted in the fifties and were actively implemented in the sixties started to yield results. The American Dwight D. Eisenhower National System of Interstate and Defense Highways had been partially completed by this time. In France, by 1970, 1010 kilometers of toll highways built under the 1955 project were commissioned. In the same year, the French government allowed the participation of private companies in the projects with concessionaires' rights [73]. Mathematical modeling in the seventies was also characterized by the continuation of the generalization of research experience, the appearance of textbooks, and the refinement and improvement of previously developed models.

The beginning of the decade was marked by the appearance in 1971 of another work, written not by a representative of an educational institution (like most of the articles and studies examined), but by a representative of a corporation, Harold J. Payne. However, this time the corporation was not General Motors, but Electronic Security Systems Corporation. Perhaps this is due to the fact that the mathematical modeling of traffic was from this moment of interest not only to manufacturers of transport but also to manufacturers of electronic control systems. The production of automatic traffic control systems was supposed to open a new niche for this business. Payne's model describes vehicle movements using a convection-type differential equation. The movement equation was deduced from the leader-following model. Payne's model should be understood as the conservation law, but due to the fact that speed does not depend on density, the right-hand side of the equation was introduced — conservation of momentum. In addition, the model had the following important advantage over previous models: Payne refused the assumption that the desired speed was achieved instantaneously [219]. The desire to "smear the discontinuities of solutions" that later in this decade moved from the LWR model to the Whitham model was motivated partially by the introduction of diffusion corrections into the Payne model.

A number of scientists realized that knowledge in the field of mathematical modeling of traffic needed to be conveyed to the general reader. First of all, traffic practices were considered — not only drivers but also people designing roads, urban spaces, as well as public transport routes. A tutorial was needed, and it was written by Renfrey B. Potts, who was one of the researchers of microscopic models at the base of the research laboratories of General Motors, and by Robert M. Oliver [233]. They managed to create a mathematical base for practitioners in the field of road construction and road maintenance. For example, the problem of choosing the fastest route was considered. The book was written for a wide range of readers. At the end of each chapter, to consolidate the materials, problems were presented that the reader was asked to solve, and solutions to problems were provided in the appendix. The book sold well for several years but, unfortunately, did not become a widespread textbook [35].

In 1972, Japanese researchers Inose and Hamada continued their work and wrote the book "Road Traffic Control" [135]. In this book, the macroscopic approach was used as an effective and almost universal method for studying and optimizing traffic management. The same year, together with Professor Fujisaki, they obtained a patent on a road traffic control system [135]. This system provided an apparatus for effecting optimal road traffic control of a road network in a traffic area. Such a traffic area was divided into a plurality of sub-areas, and a district controller was provided in each sub-area.

In 1974, Nebraska completed the construction of all highways provided by the Federal-Aid Highway Act of 1956 in the state. It was the first state to complete the work [79]. However, the completion of work throughout the United States was still far off. The originally planned system was built only in the early nineties. Moreover, considering all the additions to the original plan made during the construction process, the work cannot be considered completed even now.

Following the appearance of the LWR model, the step mentioned back in 1955, but finally proposed in 1974, was to take into account the "farsightedness" of drivers by adding diffusion terms corresponding to the fact that as the flow density increases, drivers reduce speed, and with a decrease in density — increase [290, 291]. Gerald Beresford Whitham, one of the authors of the LWR model, released a book "Linear and nonlinear waves". In this book, among other things, were described observations that the behavior of nonlinear waves during explosions according to the differential equations arising is similar to the process of congestion propagation on a single-lane road. During the Cold War, when active research was being conducted on nuclear weapons, such parallels were not difficult to arise. The described model was called the Whitham model.

The systematization of already accumulated knowledge and the development of the conceptual apparatus and terminology was continued. In 1975, an American researcher, Robert E. Shannon from the University of Alabama in Huntsville, published a generalizing work "Systems simulation: The Art and Science", in which he introduced the fixed-step method and "step method up to the next event" [258]. The terms "δt method" and "special state method" were also used. The methods proposed by Shannon allowed for more precise traffic flow computer simulation since computer models must have discreteness.

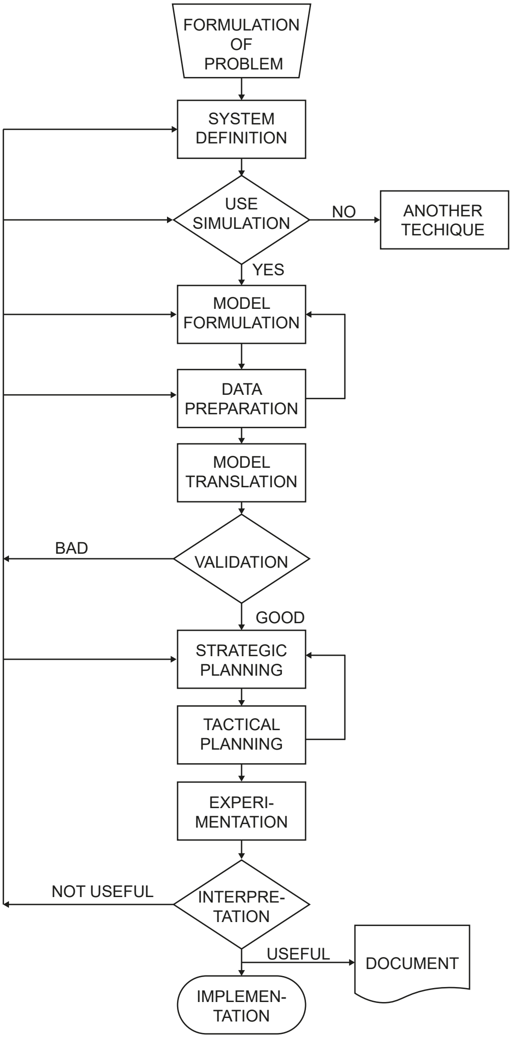


Figure 3.16: Simulation process by Shannon

In the same year, Paveri-Fontana continued the development of kinetic theory [217]. He proposed an improvement of the Prigogine kinetic equation. As in the original equation, and in the equation proposed by Paveri-Fontana, vehicles were considered as point objects. Modification of these equations taking into account the non-point sizes of vehicles would be made only in 1996 in the works of Helbing [120]. Prigogine's model included several weaknesses. In particular, the distribution of the desired velocities in it was fixed and did not depend on the evolution of the system. This motivated Paveri-Fontana to develop his own improved version of the model. In his version, he expanded the space of states by including the function of the generalized distribution of the desired speeds.

Japanese scientists continued to bring proposed models to life. In the 1970s, Inose and Hamada continued active collaboration with Edward C. Posner, who also published some articles financed by the Office of Ordnance Research, United States Army. This collaboration bore results: in 1975, the new book "Road traffic control" was published. In this book, the macroscopic method of control was expanded to mixed models based on cellular automata.

The book focused on the issues of traffic light control, which, according to the authors, form the core of all traffic control. This issue was raised due to the fact that in some areas, the density of traffic lights was so high that a car passed through 2 adjacent traffic lights in about 10 seconds, which forced the consideration of collective traffic light control systems. One of the chapters of the book, as expected, was devoted to the issues of simulation of traffic flows using computers. The authors divided simulation models into three categories: microscopic, which considered individual vehicles, macroscopic, representing groups of several vehicles, and liquid, considering the traffic flow as a fluid. In modern literature, the models that H. Inose and T. Hamada call liquid models are considered macroscopic. The methods described by the authors in simulation modeling are very close to the more modern theory of cellular automata. In particular, the authors introduced probabilistic factors of lane change.

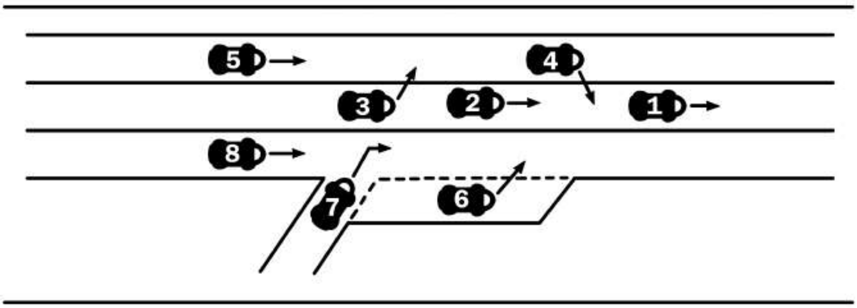


Figure 3.17: Vehicle layout: microscopic model

Since micro-level simulations could not provide the characteristics of the entire road network, large-scale simulation work had to be carried out, which led to the use of macroscopic models. Macro modeling usually operated with approximate models, which, nevertheless, remained quite satisfactory in traffic conditions, due to the fact that the main time delays were caused not by the interaction of cars with each other, but by the peculiarities of the traffic light phases.

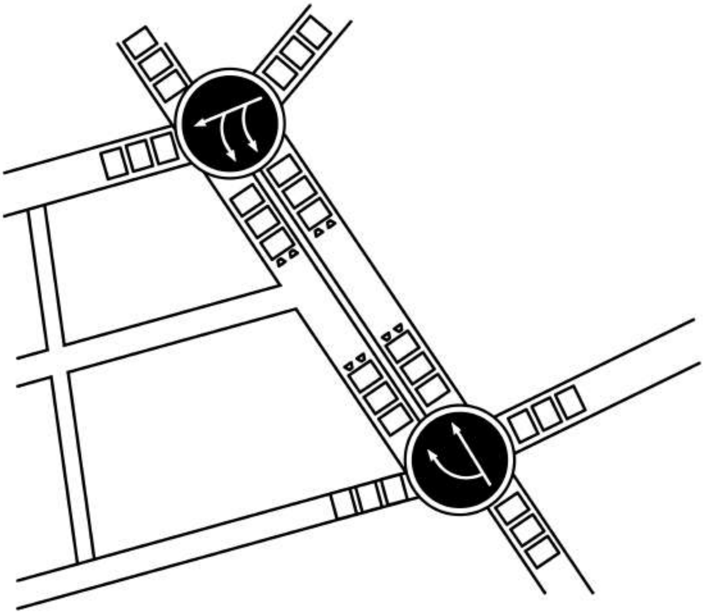


Figure 3.18: This is a part of a mnemonic diagram used in the traffic management system in Yokohama and its surroundings. Recommended directions are indicated by arrows, corresponding to the functions of similar indicators located at intersections. The congested state is indicated on the mnemonic diagram by three signals, while the average speed is displayed by two signals of different colors. One signal represents speeds below 24 km/h, and the other represents higher speeds

The authors used the following macroscopic modeling method: the roads were divided into sections several tens of meters long, and the traffic flow was represented by the number of cars on the section and their average speed. Thus, a discrete distance along the road was set with a step equal to the section length, and the cars within the section did not differ. The time step in this model was chosen of the same order as the minimum value of the passage time of the section.

This book also highlights the role of graph theory in traffic modeling, including the addition of techniques such as the introduction of "fictitious" vertices corresponding to independent traffic light phases, and the study of the topology of traffic graphs.

In 1979, in the Soviet Union, a large work "Mathematical Modeling and Assessment of Traffic Conditions of Cars and Pedestrians" was published by V. M. Kislyakov, V. V. Filippov, and I. A. Shkolyarenko [151]. The book carried out a generalization based on research both in the USSR and abroad on the subject of assessing the conditions of safe movement of cars and pedestrians. It presented new indicators of such an assessment, suitable for a scientifically based forecast, taking into account the probabilistic nature of automobile and pedestrian traffic. The book discusses traffic conditions both in the presence of traffic light control and without it. The assessment was based on an analysis of the conditions for the performance of typical maneuvers by drivers in a convenient and safe manner. In this book, the traffic situation was defined as the probabilistic state of the road traffic process, but the probabilistic factor of the drivers themselves was not considered in this context. This book was one of many works on traffic modeling in the Soviet Union. The growing interest in this area of applied mathematics in the USSR was connected, first of all, with the preparations for the 1980 Olympic Games [171].

The interest in the study of transport systems in the world in the 1960s and 1970s led to the financing of research contracts. A number of authoritative scientists in the field of mathematics (known specialist in mathematical statistics Breiman [31]), physics (Prigogine [234]), and automatic control (Athans [16]) made a great influence on further research [137]. However, in the seventies, there was a decline in the active development of new methods for mathematical modeling of traffic flows, and most of the research was related to the improvement of methods that already existed at the moment and their transfer to new technologies. Perhaps this is due to the fact that the transfer of existing models to new methods of computer modeling was more interesting for most researchers. Or, perhaps, this is due to the fact that previously developed models considered less modern calculation tools, and therefore had many assumptions. Moreover, often these assumptions had characteristics incompatible with the real traffic flow: such as the possibility of an infinite flow in the Beckmann model or neglect of the vehicle dimensions in kinetic models.

3.8 Conclusion

From the twenties to the seventies of the twentieth century, mathematical modeling of traffic underwent significant development. From the initial attempts to simulate traffic, researchers developed a specific methodology, conceptual framework, and ultimately established mathematical modeling of traffic as a distinct area of applied mathematics. Notably, starting from the end of the fifties, businesses began to contribute to the development of traffic modeling, leading to the emergence of various models, such as microscopic ones, developed not in academic settings but in corporate laboratories.

This period is significant as it marked the time when this field was not only born but also developed with clearly defined problems, conceptual, and methodological frameworks.

This was extremely timely, as in the subsequent decades, the workload for scientists increased. In the late ’80s and early ’90s, many countries, especially the US, experienced a significant increase in the number of vehicles, leading to deteriorating transportation conditions and making transportation issues a matter of national security [12]. These challenges led to new developments. Prominent physicists and laboratories, like those in Los Alamos, began to study these problems using computer technology. For example, the TRANSIMS (Transportation Analysis and Simulation System) project, an integrated set of tools for regional transportation system analysis based on cellular automata, marked a new paradigm in modeling individual vehicles and their multimodal transfers based on synthetic populations and their activities [262]. The widespread use of computers necessitated new, better, and more convenient methods for the computational modeling of traffic flows.

Von Neumann introduced the concept of cellular automata back in the ’50s [197]. From the Nagel-Schreckenberg model [193] in 1992, the active use of cellular automata theory in transport modeling began. Most models assume that a single cell can contain no more than one vehicle [104]. In cellular automata models, a crucial question is: *what is the cell?* Assuming that the cell is a fixed part of the roadway that a vehicle passes through in a certain time, the resulting dynamic system is a discrete analog of continuous models.

Many achievements of this period remained relevant in the future. Currently, the primary focus is on studying dynamic flow behavior, noting that the main complexities in dynamic systems' behavior result from the properties of the interaction between vehicles in models. Thus, more modern models and heuristics based on a mix of macroscopic (for optimization) and microscopic (for real driver behavior and congestion modeling) approaches have begun to develop. Kinetic models in modern transportation science are primarily used for investigating crowd behavior, for instance, to simulate entrances to and exits from venues like stadiums [173]. Models intermediate between hydrodynamic and kinetic, like the so-called mesoscopic model used by RAS academician Chetvertushkin's team at the Keldysh Institute of Applied Mathematics, Russia, are also prevalent in contemporary literature [153].

Since then, numerous articles and books have explored discrete micro-level models of traffic flows. Initially, these projects were very costly due to significant expenses on computer equipment and software development, necessitating additional actions to divide the task into subtasks for separate calculation. However, with the advent of cheaper and more advanced computing capacities, more general and complex models have emerged, including models incorporating various driver behavior strategies.

In conclusion, the development of mathematical modeling of road traffic from the 1920s to the 1970s was commensurate with transport development, and active studies in traffic modeling coincided with the emergence of large road construction projects. Mathematical modeling established itself as a distinct area of applied mathematics, ready to meet the challenges of the coming decades with a diverse array of methodologies and theories.

History of vehicle infrastructure

> In closing, we cannot refrain from mentioning one other striking advantage that the motocycles possess, namely, their cleanliness. With motocycles in use, the problem of street cleaning will be greatly simplified, and the health of our cities promoted [21]
>
> —Samuel Barton, 1897

4.1 Introduction

The origin and development of cities in literature are often linked to the stages of economic development and transportation systems. The central hypothesis is that a city represents the spatial manifestation of a particular economic structure and production relations. Initially, the city center emerges as the industrial production hub, leading to increased social stratification and demographic growth. It evolves into a public hub, managed by private entities, serving as a financial and advertising center. The spatial effects of industrial capitalism include:

- City expansion;
- Multifunctional use of land;
- Mass transportation of goods (including for production purposes).

The development of a city is inextricably linked to the development of transportation. This raises questions: How should a city's transportation network be planned? How will vehicles interact with each other and with other road users? Expanding our view, what implications does this have within a country?

In this chapter, we will explore the ways in which transportation infrastructure has developed and the corresponding rules of road traffic on both city and country scales.

4.2 Urban traffic network

Cities have traditionally served as transport hubs, especially during the colonization era, facilitating wealth transfer from the New World to Old World markets. For instance, major U.S. cities like New York, Boston, San Francisco, and New Orleans, located at river and ocean confluences, enabled efficient goods transportation by water. Conversely, small towns were scattered globally, mainly serving agricultural areas. Due to high transportation costs, it was economically sensible to locate commercial and manufacturing centers near the majority of agricultural workers.

The introduction of railroads gradually reduced transportation costs and increased city flexibility. By the late 19th century, cities such as Chicago and São Paulo became significant railroad hubs. This trend persisted into the 20th century, as railroads substantially lowered transportation costs and diminished the advantages of water transport, particularly for domestic transportation. This transformation's impact is evident worldwide. In the 19th century, urban areas prospered if they offered favorable conditions for producers, making national transportation infrastructure pivotal in urban landscape shaping. Transportation sector advancements will continue to facilitate city growth.

The post-World War I rise of the automobile industry initiated a new era of urban employment focused on road development driven by car production growth. The expansion of bus production and public transportation development led to urban periphery expansion, road network growth, and increased parking space demand. Since the mid-1950s, the surge in privately owned vehicles in cities has made private transportation a significant component of the urban economic system, akin to a "drug."

These developments resulted in:

- Increasing congestion and traffic safety issues;
- Reduced transport accessibility to the city center;
- Increased temporary movement losses;
- Land use changes due to increased demand for the urban transport network, aligning with the service sector;
- Central squares' deterioration.

Hence, the development of the urban transport network in Western countries was driven by the widespread use of personal vehicles as a means to develop the urban economic system, providing:

- Increased consumer sector accessibility;
- No interference with goods flow;
- Enhanced influence of city centers;
- Emergence of new city areas (large city with suburbs, urban agglomeration).

Simultaneously, public transport systems (especially rail, metro, and city trams) received a new development impetus in large cities, typically overwhelmed by private vehicles.

4.2.1 Urban planning

But how did urban planning develop in the context of cities, and where did it begin?

One of the earliest known utopian treatises dedicated to describing an ideal city was created by Italian architect Antonio Averlino, known as Filarete, between 1461 and 1463. Titled "Treatise on Architecture," it details the city of Sforzinda, named after Filarete's patron, Duke Francesco Sforza of Milan. The city, surrounded by a fortress wall, takes an eight-pointed star shape. Watchtowers are erected at each octagon vertex, with avenues extending from the center to these points. The city connects to the outside world through a canal system. The main square, featuring the cathedral, is centrally located in Sforzinda. Notably, Filarete's project incorporates various astrological symbols throughout its design.

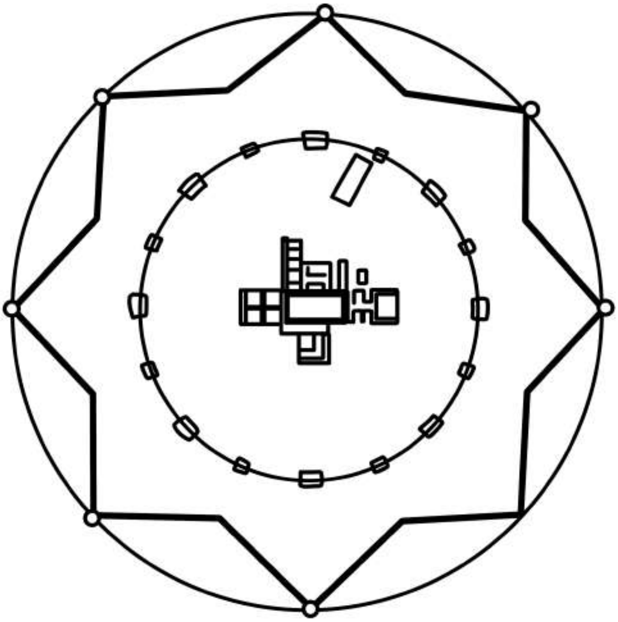


Figure 4.1: Filarete, «Plan de Storzinda»

Symmetrical plans for ideal cities were proposed by several influential figures in architecture. Leon Battista Alberti, known for his work on modern architecture, presented such plans in his ten books of treatises titled "De re aedificatoria" around 1450. Leonardo da Vinci, renowned for his artistic and scientific prowess, contributed to the concept of ideal cities with symmetrical layouts.

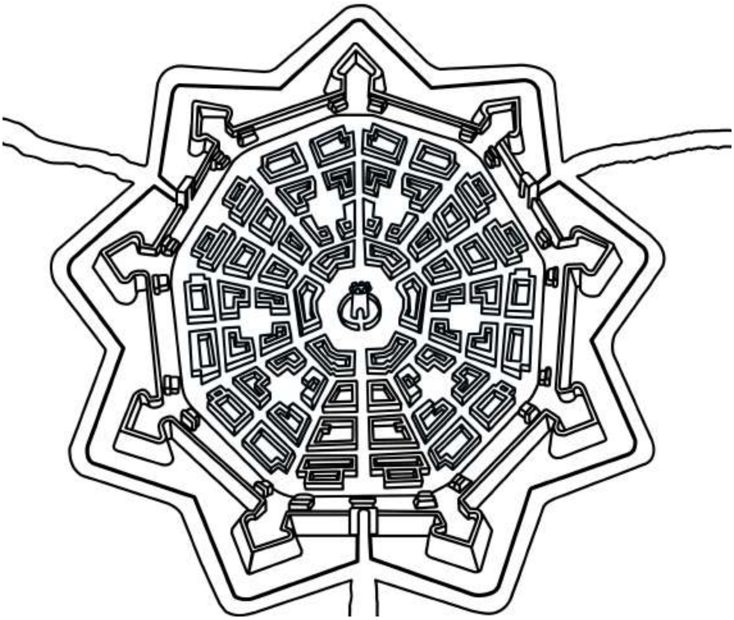


Figure 4.2: Fortified town of Palmanova

Giorgio Vasari, Antonio da Sangallo the Younger, Baldassare Tommaso Peruzzi, and Vincenzo Scamozzi are among other notable architects who explored and advocated for symmetrical plans in their works. These individuals significantly contributed to the development of architectural theories and ideals during their time.

The socio-theological utopia influencing further urbanism ideology development is detailed in Tommaso Campanella's famous work "La città del sole" (The City of the Sun, c. 1602).

The "ideal city" of Chaux, an 18th-century project of an ideal city and utopian socialism by architect Claude-Nicolas Ledoux (1736—1806), dubbed "the architect of the Enlightenment", is a major precursor of utopianism.

One notable example of massive city rebuilding in recent history is Paris. The Paris visitors see today, with its wide boulevards and grand views, results from a concentrated and centrally planned rebuilding campaign during the Second Empire (1852—1871), with some projects completed in the early Third Republic (1871—1940).

Napoleon III and Haussmann's undertaking was enormous. Both were familiar with monumental planning concepts, having studied examples like the Campidoglio in Rome during the late Renaissance and witnessed developments like the Place des Vosges in Paris, constructed in the 17th century by King Henri IV. Napoleon III was intrigued by the French tradition of "utopian" urban planning, where architects like Ledoux and Boulle conceptualized ideal cities on paper. However, much of the actual urban renovation focused on individual monuments or monument groups. Clearing space in a congested city to erect a single magnificent monument differed from reorganizing the entire city fabric to position individual monuments in relation to one another along communication and visual access lines.

In addition to ambitions, changing hygiene norms, etc., new transport forms, such as trains and omnibuses, meant more people traveled to Paris from the countryside and moved around the city. David Harvey argued that Parisian property owners had resisted government urban renewal initiatives during the July Monarchy (1830—1848) [5]. The emperor, by contrast, worked effectively to convince the capitalists that it was important to "improve the circulation of merchandise and people throughout the city" [186].

Aerial photographs of Paris show key buildings like the Stock Exchange building set apart with ample space around them, as if to free them not only from barricades but from all possible encumbrances.

The development of transport, leading to increased citizen mobility, sparked discussions about more scientific city planning.

In 1913, George Ford stated in his "The city scientific" speech:

«Except on the aesthetic side, city planning is rapidly becoming as definite a science as pure engineering. In city planning there is, above all, the necessity for a careful analysis of the conditions. It is becoming more and more obvious that the best way to secure a city plan which will be lastingly satisfactory from all points of view and really comprehensive is to put the work in charge of several experts, one an engineer, one an architect, and one, perhaps, a social expert. This group of experts must work together from the start and consult continually with regard to each feature and phase of the city plan. By such co-operation and by standardized procedure it is possible to determine within a comparatively short time a plan which is not only the best for to-day but which is so elastic that any changes during the next fifty or one hundred years can be fitted into it with virtually no loss or

alteration.» [96]

As already mentioned in the third chapter, the Russian scientist Dubelir is considered the founder of mathematical modeling of traffic. In the 1910s, he conducted pioneering studies on adapting urban planning to the inevitable sharp increase in the number of cars. Dubelir wrote that "an increase in the intensity of the movement of carriages, and especially cars, should be taken as the most important factor in the development of cities, to be taken into account in design practice" [68].

The first part of Dubelir's book is titled "City Planning," where he describes the layout of urban streets from the perspective of convenience in movement. He suggests that a certain gap should be allowed between two carriages, and this gap can be considered as the width of the lane necessary for carriage passage, or what he further refers to as the "track," measuring 1.25 sazhens (an old Russian unit of length). This width also allows carriages to make turns since the length of a carriage with horses typically does not exceed 2.30 sazhens.

Dubelir also discusses the number of these tracks for streets with heavy traffic, stating that it can be increased to six, considering the need for parking, passage, and overtaking carriages in each direction. His emphasis on lane width and the number of tracks reflects his considerations for efficient movement within the city.

Additionally, Dubelir addresses the ecological and economic aspects of city planning. He mentions the hygienic and partially fire-related aspects, emphasizing the need to cover wide lanes with improved pavement. However, he acknowledges that such wide coverage would involve significant expenses and, in most cases, would be practically unfeasible. In the chapter "Street Network Layout Systems" dedicated to city planning, the author introduces the following types of systems:

1. "Radial system" (with Moscow being a typical example of such a system, with its radial streets and rings).

2. "Rectangular system" (while the radial system is characteristic of old cities that developed gradually along existing roads, the origin of the rectangular system is associated with the simultaneous planning of entire parts or even whole cities in a short period of time). The most famous city with this street layout is New York.

3. "Diagonal system", considered by the author as the best system for intense traffic until the 1890s. Paris was given as an example of this system. It features diagonal streets connecting different parts of the city but not passing through the center.

4. "Modern planning system" includes "main arteries," several "centers" with radial streets connected by diagonals. Residential streets are located between the main roads, and space is allocated for urban squares.

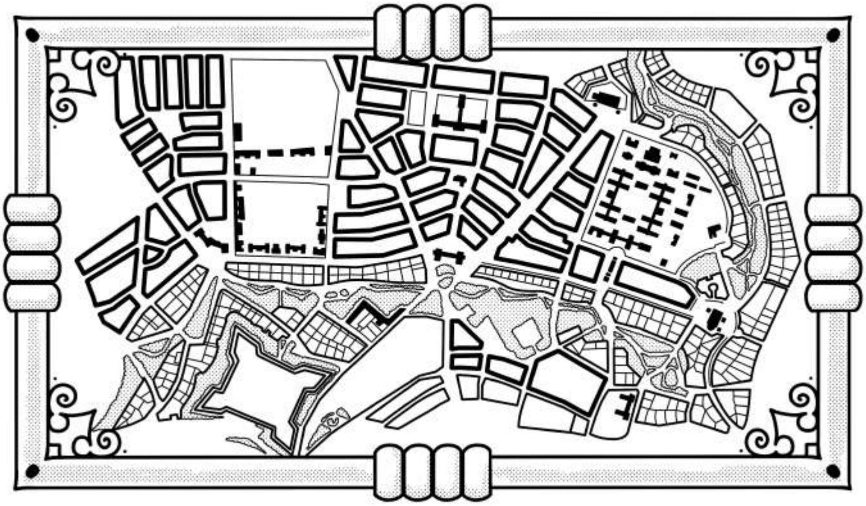


Figure 4.3: Modern planning system

George Sheleikhovsky (1892—1946), an eminent Russian engineer and urban planning scientist, wrote about various aspects of urbanism. He emphasized the importance of designing streets and urban infrastructure with a long-term perspective, projecting them for a period of around 50 years. He argued against placing any restrictions on automobiles as a means of transportation, stating that there is no basis for such limitations when dealing with such long-term planning. Instead, he advocated for designing cities by considering the maximum possible flow of automotive traffic on the main thoroughfares.

Sheleikhovsky also highlighted the need for a synthesis of four key planning elements: residential areas, transportation, street networks, and the overall form of the urban plan. He believed that these four aspects should be integrated and addressed collectively in the planning process [260].

Abram Zilbertal, in his statements, argued that addressing transportation issues

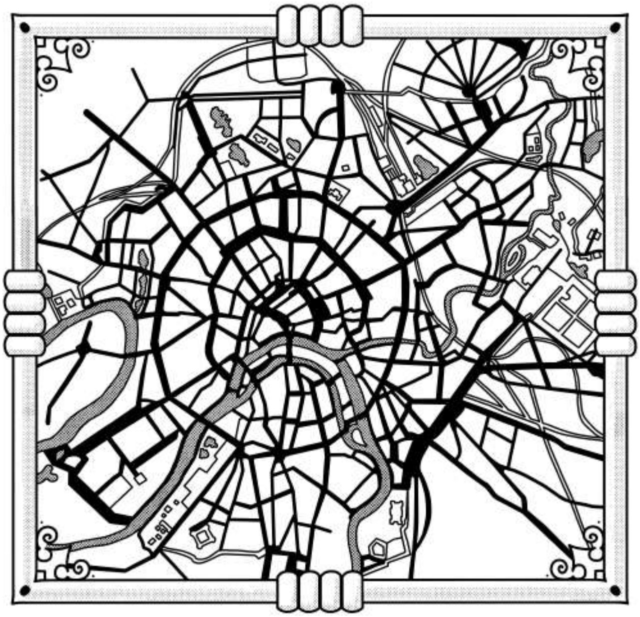


Figure 4.4: Radial road system

is not solely a mathematical problem but is indeed influenced by how highly society values its time and convenience [305].

Sir Raymond Unwin played a significant role in the town planning movement in England. During a time when the first Town Planning Act of 1909 was not yet enacted, urban development was regulated by local bye-laws and regulations, which often lacked creativity and flexibility. However, Unwin navigated these challenges adeptly by securing Private Bills that provided his team with the required powers to implement their plans.

Sir Raymond Unwin's book "Town Planning in Practice: An Introduction to the Art of Designing Cities and Suburbs," published in 1913, focused on what is now known as urban design. The book provided a comprehensive exploration of various aspects of city planning and design.

Unwin began by examining and comparing the plans of different cities, ranging from Edinburgh to Palmyra to Karlsruhe. This comparative analysis allowed him to draw insights and lessons from various urban environments. He emphasized the importance of understanding the site through thorough surveying, taking into account its topography, natural features, and existing infrastructure.

The book also delved into the concept of beauty in urban design, discussing both formal and informal aspects. Unwin highlighted the significance of aesthetics in shaping the built environment and creating pleasant and harmonious spaces for residents.

Regarding the arrangement of roads and buildings, Unwin explored different possibilities and considerations. It's worth noting that at the time of writing, the impact and development of motorized traffic were not as pronounced or anticipated as they would later become. Therefore, the book likely focused more on pedestrian-oriented design principles and traditional modes of transportation.

He writes: *"In residential districts one of the greatest difficulties to be contended with is the constant multiplication of buildings too small in scale to produce individually an effect in the road, and every opportunity should be taken to group buildings so that units may be produced of large scale. Even where it is not possible to avoid much repetition of semi-detached or detached houses, they should be so arranged as to give some sense of grouping. The setback of three or four pairs of houses and the arrangement of a continuous green in front of them, with the proper treatment of the house at each end, which are set forward again to the building line, will of itself produce some grouping ... and in many other ways, especially where it is possible for the site planner to be in touch with the designer of the buildings, much may be done to produce interest and variety in the street pictures, while at the same time maintaining the general sense of unity which is usually so wanting in modern suburban roads ... The tendency of the modern individual has been to build his house in such a way as to emphasise its detachment and difference from all its neighbours, but no beauty can arise from the mere creation of detached units: the result is bound to be monotonous and devoid of beauty."* [277]

4.2.2 Lane markings: let's divide the flows!

In addition to the direct planning of city streets, road markings contribute to the orderliness of traffic. Nowadays, there are numerous road signs available that help organize proper traffic flow on any road segment. However, road markings are essential and indispensable. Just imagine how challenging it would be to navigate a road with three lanes and no markings. During nighttime, road markings also assist drivers. The solid white lines delineate the edges of the roadway, greatly enhancing visibility in low-light conditions.

From the early years of motorization, the most common types of road incidents were pedestrian accidents and collisions. Even when vehicle speeds did not exceed 70 km/h, the need for road markings was evident. Therefore, as soon as it became possible to draw on road surfaces, life-saving lines appeared on the roads. Among the first were lines to separate lanes of oncoming traffic, stop lines to mark intersections and pedestrian crossings. The life-saving road markings owe their existence and development to advocates for traffic safety and have an interesting history.

In the early 20th century, an American individual sought to patent his idea of applying a center line on the road to separate oncoming traffic streams. However, his claim for authorship was rejected on the grounds that the concept of road markings was as old as time and had already been implemented, for example, in the construction of ancient roads in Europe and America. Indeed, light-colored stone lines placed in the center of roads delineated lanes of oncoming traffic in the streets of ancient Greek and Roman cities. The central line made of limestone also survived on roads built by the Aztecs in the early 17th century (Figure **??**). Examples of such markings can still be found in the area of ancient Mexico today.

And yet, the true history of the centerline is best traced back to the emergence of asphalt concrete roads. One of the first proposals for such markings was put forward by Edward N. Hines, a member of the Wayne County Road Commission in Michigan, USA. According to legend, in 1911, he was driving along a narrow two-way road and was greatly concerned about oncoming vehicles. When he came across a milk truck with leaking contents, an idea struck him. According to popular rumors, the spilled stream of white milk on the dark asphalt inspired Hines to use road markings to separate oncoming streams of cars. The first painted center lines were used in 1911 on River Road in Trenton, Wayne County [156]. His belief was that if a line designated a lane, people

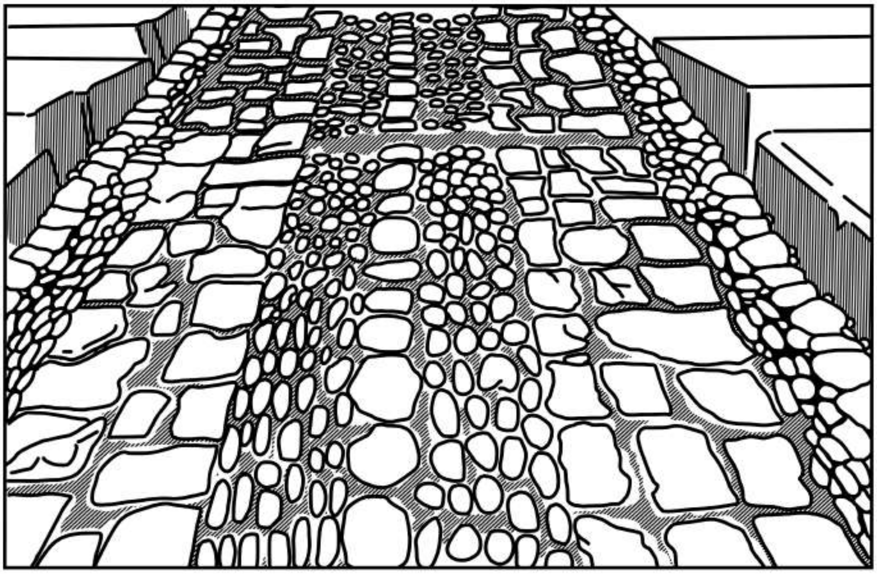


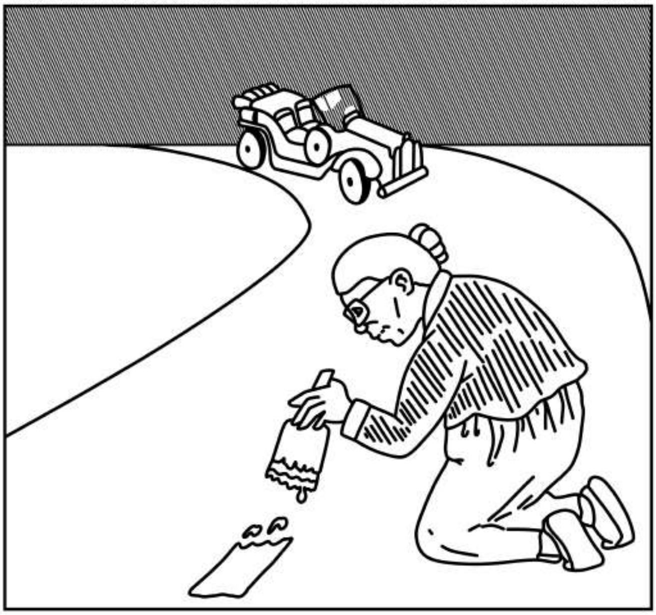
Figure 4.5: Central line on the Inca stone road in Peru

would stay in their respective lanes, leading to a reduction in accidents. And he was right. In recognition of his contributions and ideas, Hines was inducted into the Michigan Transportation Hall of Honor in 1972.

In 1917, in another state, Oregon, Deputy Sheriff Peter Rexford also proposed the idea of using a separating line to make the roads safer. He suggested using a yellow line. However, the local authorities refused to fund the idea. Undeterred, Rexford took matters into his own hands with the support of the sheriff and paid for it himself.

Among the early initiatives for road markings, there is another remarkable story. In 1917, in the city of Indio, California, nurse June McCaroll was driving to attend to a patient when she narrowly avoided a truck coming from the opposite direction. She miraculously swerved and ran off the road. This incident prompted her to come up with the idea of marking the road to separate opposing traffic. June approached the local authorities, who praised her for the good idea but took no action. Undeterred, June took matters into her own hands. She grabbed paint and brushes, got down on all fours, and personally painted the first known four-inch (10 cm) stripe measuring two miles long on Indio Boulevard, which was part of U.S. Route 99 at that time. But the determined nurse didn't stop there. June wrote letters to various agencies and spoke at different public clubs. Finally, her voice was heard, and in 1924, the local legislature passed a law instructing the Road Commission to apply road markings.

In 1918, the United Kingdom introduced the use of white road marking lines for automobile safety. The initial experimentation with painting a center white line took place in Sutton Coldfield, Birmingham. Due to numerous complaints from residents regarding reckless driving and frequent collisions, the Sutton Coldfield Corporation decided to conduct an experiment by painting lines on Maney Corner. The results were highly successful in reducing accidents, leading to the adoption of road marking lines as a standard road safety measure throughout the country in 1926. This innovative idea was soon replicated by other countries, and road lines began to appear worldwide.



During the 1930s, painted lines on roads served multiple purposes beyond indicating road ownership. Solid white lines were used as road dividers, stop signs, and cautionary signals, often with the assistance of police officers directing traffic. This period marked a shift in the perception of painted lines and symbols as more than just visual markers. With the increasing popularity of automobiles and faster interstate travel, a standardized language using various types of lines, colors, symbols, stop bars, painted arrows, and legends was developed to accommodate these developments.

To ensure consistency nationwide, the first Manual on Uniform Traffic Control Devices was published in the USA during the 1930s. This guide established a common language for traffic control and implemented its regulations across the country. Today, this guide continues to regulate all aspects of highway safety, including lines, legends, logos, arrows, reflective sheeting, signs, and more, ensuring uniformity and enhancing driver safety. As regulations increased, so did the implementation of road striping to meet the evolving needs of road users.

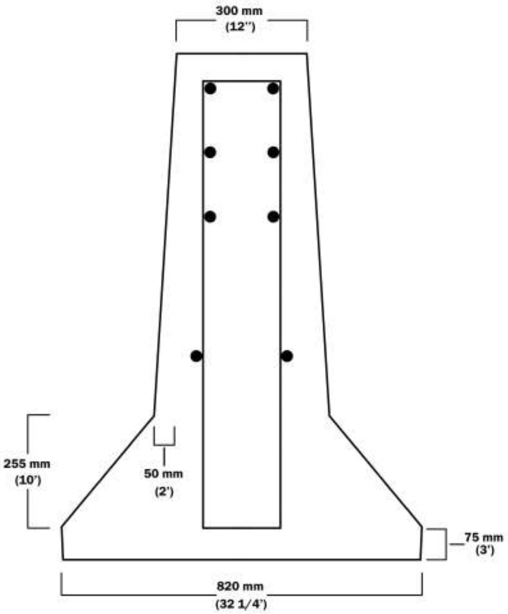
Of course, there were discussions about the fact that a simple line was not a sufficiently visible divider. For example, suggestions were made to use strips made of a different material to enhance their visibility. For instance, in his 1939 speech, M.R. Keefe, Chief Engineer of the State Highway Commission of Indiana, wrote:

> *« This pavement was opened only last fall and up to the present I am disappointed in this type of median dividing strip. The divided lane section is about four miles long, and the median strip is constructed in three different colors over this distance. A recent trip over this road after dark convinced the speaker that there is much to be desired in visibility. The different colors could not be distinguished one from the other. In fact, at times, it was difficult even to see the four- foot median strip. My last inspection was just after a snow fall, and the median strip had been splashed with mud and slush until its visibility was practically zero. I like the suggested design of the speaker— a four-foot bituminous strip dividing his two-lane road. If the lanes are of concrete and the four-foot median strip is bituminous, the contrasting color will do a great deal to eliminate medial friction. We shall watch this construction with much interest. May I venture the prophecy that if this two-lane construction is increased to four lanes later, the four-foot median strip will have proved so satisfactory that it will be left in place rather than replaced with a curb section. »* [147]

Later, especially on highways, median strips became popular for use. The physical separation of traffic lanes through the use of median safety barriers is significantly more secure than just a line on the road.

Median safety barriers can be constructed using various materials such as concrete, steel, and wire rope. The choice of the type of median barrier depends on several factors including traffic volume, traffic speed, types of vehicles, median width, number of lanes, road alignment, crash history, and installation and maintenance costs.

One of the most widely used types of modular concrete or plastic barriers is the Jersey barrier, also known as the New Jersey wall. It was developed in the 1950s and introduced in its current form in 1959 at the Stevens Institute of



Technology in New Jersey, United States. The barrier was developed under the direction of the New Jersey State Highway Department with the purpose of dividing multiple lanes on highways.

Currently, median strips can effectively divide a road into two separate lanes, creating a significant separation between opposing traffic. They serve as a "neutral ground" or barrier between the lanes, enhancing safety and traffic flow.

Interestingly, there is no international English standard for the term. "Median," "median strip," and "median divider island" are common in North American and Antipodean English. Variants in North American English include regional terms such as "neutral ground" in New Orleans usage. In Connecticut, the median of a highway would be the midway. In British English, "central reservation" or "central median" is the preferred usage; it also occurs widely in formal documents in some non-British regions such as South Africa, where there are other informal regional words. "Neutral section" and "central nature strip" are coinages in Australian English.

4.2.3 Crosswalk

The concept of pedestrian crossings also dates back centuries. Ancient Roman roads, for instance, had well-preserved pedestrian crossings. Three rectangular stones remarkably resemble the modern-day “zebra crossing.” However, their functionality went beyond just ensuring safe crossings and forcing riders to reduce their speed. The main purpose was to allow pedestrians to cross the road without soiling their feet, as in ancient Greece and Rome, roads also served as channels for urban waste disposal.

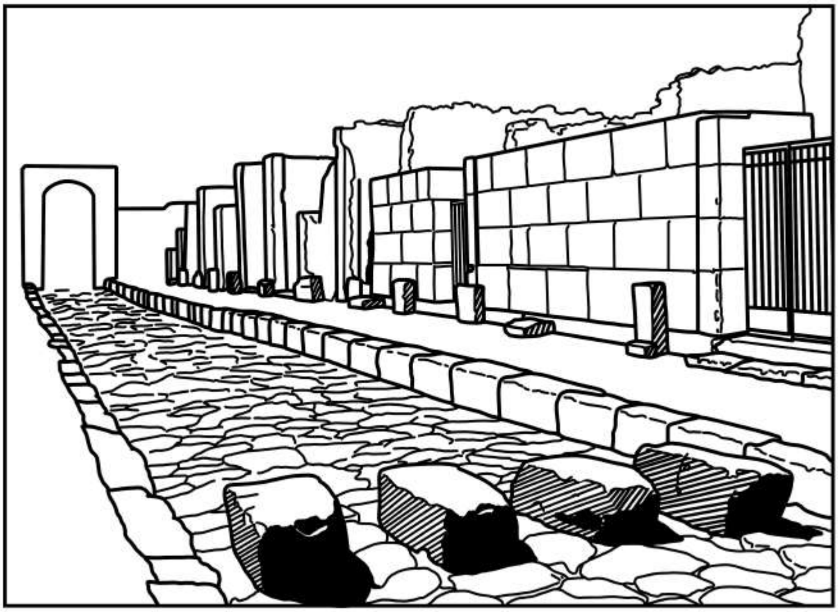


Figure 4.6: Zebra crossing in Ancient Rome

With the advent of trains and automobiles, designated areas for pedestrians to cross the road were initially marked by signs and later by road signs. The road markings on pedestrian crossings first appeared in the form of metal discs. They were highly visible to pedestrians, instilling confidence in their safety. However, they often went unnoticed by drivers, which frequently led to tragic incidents.

Even in the report [224] of 1972, which studied pedestrian accident experience at unsignalized intersections and whether it is less in marked or unmarked crosswalks, the marking was generally done only by two white lines. Authors propose the accident experience covering a 5-year period which was studied at 400 intersections, each having one marked and one unmarked crosswalk crossing the main thoroughfare. They provide the following advantages and disadvantages of the marked crosswalks:

«In general, marked crosswalks have the following advantages.
1. *They may help pedestrians orient themselves and find their way across complex intersections.*
2. *They may help show pedestrians the shortest route across traffic.*
3. *They may help show pedestrians the route with the least exposure to vehicular traffic and traffic conflicts.*
4. *They may help position pedestrians where they can be seen best by oncoming traffic.*
5. *They may help utilize the presence of luminaires to improve pedestrian nighttime safety.*
6. *They may help channelize and limit pedestrian traffic to specific locations,*
7. *They may aid in enforcing pedestrian crossing regulations.*
8. *They may act, in a limited manner, as a warning device and reminder to motorists that this is a location where pedestrian conflicts can be expected.*

Marked crosswalks also exhibit some disadvantages.
1. *They may cause pedestrians to have a false sense of security and to place themselves in a hazardous position with respect to vehicular traffic.*
2. *They may cause the pedestrian to think that the motorist can and will stop in all cases, even when it is impossible to do so.*
3. *They may cause a greater number of rear-end and associated collisions due to pedestrians not waiting for gaps in traffic.*
4. *They may cause an increase in fatal and serious-injury accidents.*
5. *They may cause an increase in community-wide accident insurance rates.*
6. *They may cause a disrespect for all pedestrian regulations and traffic controls.*

Unjustified and poorly located marked crosswalks may cause an increased expense to the taxpayers for installation and maintenance costs that may not be justified in terms of improved public safety. Indeed, such crosswalks may tend to increase the hazard to pedestrians and motorists alike.» [224]

The question of visibility of the crosswalks was widely studied. At 1940, Burton Marsh wrote:

«Our studies lead us to believe that there is need for careful re-examination of the design of pedestrian crosswalks. Particularly where crosswalks are located in the midblock, or at irregular-shaped intersections, it is of very great importance that

the motorist always be able to know in advance where the crosswalk is at which he is expected to yield the right-of-way to pedestrians. Yet, a large number of crosswalks in this country do not satisfy this reasonable requirement. For example, the usual 4- *to* 6*-in. paint lines are virtually invisible from a reasonable distance back of the crosswalk location and from the height at which the motorist is expected to see it. At the near-right side of the crosswalk, if the customary type of paint line is used, it should be in the neighborhood of* 24 *in. wide, except where vehicle speeds are very slow. Furthermore, if snow or ice covers the markings, there is no way at all for a motorist to know of the existence of an irregular-marked crosswalk, although the pedestrian may be expecting the motorist to yield right-of-way there. At night many of the crosswalks are virtually invisible. In England, this has been faced and the Belisha Beacon is one of the important results. Named after the then Head of the Ministry of Transport Mr. Hore-Belisha, it consists of a post with alternate dark and light bands on it and is surmounted by a large orange-colored globe which is lighted at night. By November* 1, 1937, 28, 000 *of these special crossings had been established,* 12, 000 *of which were in London, and evidences were that they were helping to reduce the pedestrian toll. Fatalities had decreased* 25 *per cent and injuries* 12 *per cent, in a partial before and after study. These beacons serve both day and night to help the motorists identify places at which they are to yield the right-of-way. There is a real need for much more effective crosswalk design in this country and it is especially important at irregularly located crosswalks.»* [179]

What did they talk about? A Belisha beacon is an amber-coloured globe lamp atop a tall black and white striped pole, marking pedestrian crossings of roads. Yes, this specific color can definitely help to know where is the potential crosswalk.

In 1948, an experiment was conducted in the United Kingdom to determine the most effective pedestrian crossing markings. Around a thousand variations were analyzed, and it was concluded that a series of wide black and white stripes had the greatest visual impact on human eyes. The laboratory was visited by a Member of Parliament and later by Prime Minister Leonard James Callaghan. It is said that he observed the resemblance of this marking to a zebra, attributing to it the name "the zebra crossing" for the new pedestrian crossing. However, it should be noted that Callaghan himself did not claim authorship of the term.

On October 31, 1951, the black and white markings were applied to pedestrian crossings in the town of Slough, England, which soon led to a significant reduction in pedestrian accidents. On July 8, 1952, such markings were also introduced in Munich. On August 24, 1953, the legislature implemented pedestrian crossings nationwide in Germany. The new markings were highly regarded in

Germany, and the birth of the pedestrian crossing is celebrated as a significant event. Every five years, at least one major German newspaper writes about it to this day.

Indeed, the zebra was not the only animal considered in British road safety experiments. In the 1950s, there were serious debates about replacing the zebra crossing with the "panda" crossing. The panda crossing featured black and white triangles and was accompanied by a traffic light system. Flashing and alternating amber, red, and green lights regulated the sequence of movement for pedestrians and vehicles. These crossings were installed between 1962 and 1967. However, despite its seemingly rational design, the panda crossing was not successful. The distinction between the flashing and pulsating amber phases was subtle yet crucial. By 1967, the panda crossing became a concern for the Ministry of Transport, and a new type of crossing called the "X-way" was introduced. The transition to the new system was not gradual but rather urgent, with the pandas being removed. In 1969, the "X-way" was replaced by the pelican crossing, or the archaic "pelicon crossing" (Pedestrian Light Controlled). In this type of crossing, pedestrians had to activate the traffic lights themselves. There was also a variation of the pelican crossing called the "Pegasus crossing," designed for equestrians. The button to activate the traffic lights was positioned at a height reachable by a rider's hand.

Figure 4.7: The "panda" pedestrian crossing was considered as an alternative to the "zebra" crossing in the United Kingdom during the 1960s

The British experiments attracted widespread attention. In 1968, this road marking was included in the Vienna Convention on Road Signs and Signals: *"For marking pedestrian crossings, it is preferable to use fairly wide parallel stripes along the axis of the roadway."*

The worldwide fame of the “zebra crossing” was also contributed to by The Beatles. On August 8, 1969, the four members of The Beatles posed on a brand-new pedestrian crossing on Abbey Road in London, which led to the EMI recording studio of the same name, for their final collaborative work — the Abbey Road album. This pedestrian crossing attained the status of a historical landmark, and the “zebra crossing” spread to cities all around the world.

Experiments to improve pedestrian crossing markings continue. In the 1990s, the “Toucan crossing” (derived from “two can cross”) was introduced, designed for pedestrians and cyclists. This crossing featured road markings equipped with pedestrian detection sensors that controlled the traffic signals.

The most modern type of pedestrian crossing that has been implemented in England since 2003 is called a “puffin crossing” (derived from the phrase “pedestrian user-friendly intelligent”). The design of a puffin crossing differs from the older pelican crossing in that the pedestrian lights are located on the same side of the road as the pedestrian, rather than across the road. Electronic sensors detect pedestrians approaching the crossing, activate the green light for them, and display a red light for vehicles. The sensors also monitor the presence of pedestrians on the crossing and change the signal accordingly. This system allows for efficient use of the crossing: when no pedestrians are present, the signal remains green for vehicles. Ongoing experiments continue to explore different ways to enhance pedestrian crossings, including the development of innovative markings and the search for optimal lighting solutions.

4.2.4 Does the color matter?

The decisive criterion for choosing the color of road markings is its contrast with the color of the road surface. In the 1930s, even black paint was used in Germany for marking because the concrete surface of the autobahns was light. Yellow and white colors were primarily used as they provided the highest contrast to the gray and black road surface. For a long time in the United States, both yellow and white colors were used for dividing lines.

As we have already mentioned, in 1917, a yellow centerline was painted across the Columbia River Highway. Deputy Sheriff Peter Rexford made the decision to use yellow paint after realizing that white paint was not as visible during dark and stormy nights.

So from 1917 until 1954, both yellow and white paint were used for striping roads in the USA, and cities, counties, and states engaged in ongoing debates about which color was more appropriate.

In 1948, both yellow and white paint were officially recommended and used interchangeably. Yellow paint was commonly used for warning signs and was a familiar sight on American roads. On the other hand, white paint was considered more visible at night and less harmful to road workers, as yellow paint contained a significant amount of lead chromate. However, in 1954, the debate on which color to use for highway center lines was finally settled. Forty-seven states agreed to adopt white as the standard color, while Oregon was the first and last state to use yellow painted lines.

Indeed, it is surprising that the ruling regarding the color of center lines changed again just seven years later. In 1961, the Manual on Uniform Traffic Control Devices (MUTCD) was revised, and it made yellow center lines mandatory for the two exceptions where they had previously only been recommended.

The 1971 edition of the MUTCD introduced significant standards regarding the use of yellow and white lines. It declared that yellow lines would become the standard for centerlines indicating opposing traffic, while white lines would be used for dividing traffic going in the same direction. A single yellow line could also mark the left edge of the roadway on divided highways or one-way roads. The white line would continue to designate the edge of the shoulder.

This transition from white to yellow lines took place between 1971 and 1975. The resurgence of yellow as the preferred color was largely due to its established association with warning signs, and crossing over into opposing traffic was considered a situation that warranted a warning.

The Vienna Convention also initially recommended the use of both colors equally:

«If road markings are painted, they shall be yellow or white; however, blue may be used for markings showing places where parking is permitted or restricted. When both yellow and white are used in the territory of a Contracting Party, markings of the same class shall be of the same colour.» (Article 29.2 *of [276])*

But in 1973, the Geneva Protocol gave preference to the white color:

“The road markings shall be white. The term ’white’ includes shades of silver or light grey. However:

- Markings showing places where parking is permitted or restricted may be blue;
- Zigzag lines showing places where parking is prohibited shall be yellow;
- The continuous or broken line on the kerb or on the edge of the carriageway to show that standing or parking is prohibited or restricted shall be yellow.” [236]

Amendment which entered into force on 28 March 2006 does not change this general rule.

From that moment on, white paint has been used for permanent road markings in all European countries. An exception is made for permanent yellow markings, which indicate bus and taxi stops, as well as areas where parking or stopping is prohibited.

Temporary road markings can be painted in yellow (Germany, Estonia), orange, or red (Austria, Switzerland) and are used during road repairs and traffic reorganization.

When both permanent and temporary road markings are present, it is necessary to follow the temporary markings. They are typically applied using short-lived paint that will fade away or be removed by road services upon completion of the repairs.

4.2.5 Other road markings

Road markings play a crucial role in traffic control plans as they define the layout of the road surface and offer visual guidance to road users. From their initial implementation to the present day, road markings have become an integral part of transportation infrastructure.

One of the first studies about the impact of road markings width and configuration on driver behavior was conducted in 1986 [70], in which the authors examined the effectiveness of 10 temporary marking treatments on various measures of driver performance (speed and distance, erratic maneuvers, and subjective comments and ratings of the treatments by the drivers).

Studies related to the general visibility of road markings were mainly focused on determining the maximum detection distance for road markings and the minimum levels of retroreflectivity required by drivers in dry and wet conditions, as well as other factors affecting road marking visibility.

One of the first such studies was conducted in 1999 [306], in which authors set up a field study to determine how various types of road marking arrows affect recognition distance. The results showed that the elongated full-scale arrows provide significantly longer recognition distances than their standard full-scale counterparts.

A great deal of work has been done over the past century to prove the worth of the humble painted pavement marking. A summary of some of the findings was presented in a paper [36] of 2004 by Bob Carnaby.

In 1930, a British inventor named Percy Shaw patented a light-reflecting device for improving visibility on roads, known as “Cat’s Eye.” It gained widespread use in the United Kingdom during World War II, when the use of cat’s eyes allowed roads to remain visible to drivers even during blackout conditions.

In the 1950s, Elbert Dysart Botts, a paint chemistry specialist at Caltrans (California Department of Transportation), came up with the idea of using glass beads in road paint to improve visibility. To prevent the layer of water covering the markings from hindering visibility, the retroreflective elements were raised a quarter of an inch (approximately 6 mm) above the road surface. This technique resulted in another effect: when driving over such markings, the driver would hear a dull thud. However, this was considered advantageous as it served as a warning to drivers who may not have noticed the markings, alerting them to the fact that they were crossing over the lines.

Ceramic or plastic markers were originally attached to the road using special nails until Herb Roney, a former student of Botts, invented a special epoxy resin that securely bonded the markers to the road surface. Starting from 1966, these markers, known as “glass beads by Potters Beads,” began to be used on roads in the United States, and later in other countries around the world.

Initially, dashed lines were introduced as a cost-saving measure for paint, but their utility quickly expanded beyond just financial savings. It was discovered that dashed lines could effectively communicate messages to drivers. In 1956, the use of dashed lines began, accompanied by a new set of rules for overtaking other vehicles on the road. These dashed lines added complexity to the existing traffic guidelines in the United States, but they were a necessary addition. As double lines were implemented on two-way roads, dashed lines were combined with solid lines to serve as a means of keeping vehicles on their respective sides of the road and indicating when it was permissible or not permissible to pass.

Lines, legends, arrows, and various markings, along with raised pavement markers, play a crucial role in modern transportation systems. They are extensively utilized on major roads across Europe, America, and many other countries worldwide. These markings, typically in white and yellow colors, serve a variety of purposes, including providing directional guidance, indicating interstate routes, marking school zones, defining stop bars at intersections, and much more. Parking lots also make use of markings, with red often indicating fire lanes where parking is prohibited, and blue designating spaces for handicapped parking. These lines and markings are pervasive in our daily lives and will continue to be essential elements of transportation infrastructure for many years to come.

Indeed, in some countries, green and red paint is used to designate lanes for public transport, bicycles, or electric vehicles. In the United States, designated lanes for public transport, often known as HOV (High Occupancy Vehicle) lanes or carpool lanes, are reserved for vehicles carrying more than one passenger.

The Vienna Convention on Road Signs and Signals in 1968 stated: *«The length and width of markings vary according to purpose, although no exact figures for* *size are stated; roads in built-up areas should use a broken line for lane division, while continuous lines must only be used in special cases, such as reduced visibility or narrowed carriageways. All words painted on the road surface should be either of place names or words recognizable in most languages, such as "Stop" or "Taxi".»*

Road markings continue to evolve, introducing new types, colors, and ongoing experiments with materials and application technologies. Currently, the concept of "waffles" is gaining popularity. The marking consists of a grid of intersecting lines in white, bright yellow, or orange colors. Its purpose is to inform drivers that the intersection is congested and prone to traffic jams. If a driver is unsure whether they can cross the "waffle" section without stopping, they should refrain from crossing the stop line and wait until the preceding vehicle completely clears the intersection. Thus, the "waffle" marking aims to address the issue of self-blocking caused by inconsistent flow coordination between streets. This type of marking is used in Japan, China, and other countries.

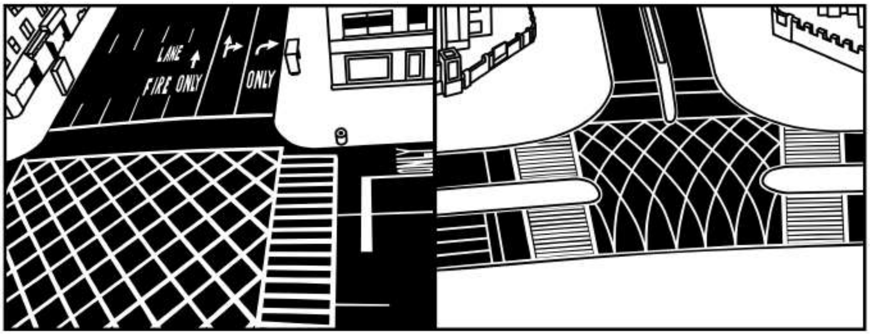


Figure 4.8: Waffle marking

In the United States, a new road marking has been introduced, previously tested in the United Kingdom and Australia. Within 150 meters of a dangerous intersection or pedestrian crossing, a white zigzag line is painted on the road. When drivers encounter this "crazy marking" in the middle of the road, they instinctively start braking. This marking is employed in high-accident areas to draw attention and does not impose any specific requirements on drivers.

4.2.6 What is the markup made from?

Initially, road striping relied on simple single-part paints. However, with the increase in automobile traffic worldwide, the issue of durability arose. Painted lines would quickly wear off in high-traffic areas, posing a significant hazard as drivers heavily relied on road markings for guidance. To enhance durability and reflectivity, glass beads by Potters Beads were sometimes added to the paint as it dried. However, this method did not provide sufficient durability for heavy traffic conditions, leading to the development of alternative products.

The introduction of two-part paints, also known as epoxy or plural components, marked a significant advancement in road striping materials. These paints, formulated to be harder and more durable than standard paints, are comparable in toughness to the materials used in constructing fiberglass boats, where epoxy is combined with a hardener or catalyst.

By incorporating glass beads, these hybrid paints offered longer-lasting stripes and pavement markings while improving visibility. The combination of durable coatings and reflective glass beads became a popular choice and continues to be widely used in road striping applications.

Thermoplastics represent a third, highly durable class of road markings. Developed during World War II in response to solvent shortages, thermoplastics are durable plastic materials with unique properties. They can be heated to a liquid state and then cooled to a solid state, a process that can be repeated multiple times.

Thermoplastics are applied in liquid form and rapidly solidify upon cooling to form durable road markings. Glass beads can be embedded during the hot application to enhance nighttime visibility and increase longevity.

A key advantage of thermoplastics is their ability to be applied in a thicker layer compared to standard paint, typically 125 mils, contributing to their long-lasting nature and durability on road surfaces.

Thermoplastic resins used in road markings are typically derived from modified esters sourced from gum or tall-oil rosins, as well as aliphatic C5 synthetic hydrocarbons. They need to be heated to temperatures exceeding $200°C$ for application and solidify almost instantaneously upon cooling, minimizing traffic disruptions.

The longevity of thermoplastic markings depends on their thickness, with a 1 mm thickness typically lasting around three years and a 3 mm thickness lasting approximately five years.

Preformed thermoplastic pavement markings, sometimes called “tape” (but not to be confused with preformed polymer tape), are thermoplastic cut into final shapes by manufacturers and ready to be positioned on an asphalt or concrete pavement surface. This eliminates the need for metal stencils and on-site melters. The ability to be melted or extruded, allowed to cool, and melted again, makes preformed thermoplastic both possible and practical.

Typically, preformed thermoplastic markings can last 3 to 6 years, and much longer in low traffic or parking lot environments. The most common applications of preformed thermoplastic pavement markings are found at intersections as transverse markings such as stop lines, legends, crosswalks, arrows, bike lane symbols, and accessibility symbols.

4.2.7 Electric roads

Two main approaches are generally considered to provide a certain level of autonomy to a vehicle: increasing the energy storage capacity or improving the density and coverage of the energy charging infrastructure. The specific energy density of fossil fuels is approximately 30 times higher than that of current lithium-ion batteries used in electric vehicles. Fossil fuels offer about 2.5 kWh/kg of useful mass-specific energy (considering a 20% efficiency for the thermal powertrain) compared to around 80 Wh/kg for batteries (accounting for efficiency and the additional mass due to conditioning). For example, the 22 kWh battery of a Renault Zoe weighs 300 kg. To achieve an equivalent energy, a full tank of 50 liters of fuel (approximately 36 kg or 75 kWh) would require a battery weighing over a ton! However, some manufacturers, such as Tesla, have pursued this path, incorporating batteries with capacities of up to 85 kWh in their vehicles. This approach involves significant additional costs.

While the electric accumulator was invented by Gaston Planté in 1859, its evolution over the past century has been relatively limited. For reference, early electric vehicles, such as the 'Jamais contente,' used lead-acid batteries with an energy density of around 20 Wh/kg, and the first industrialized electric vehicles achieved around 50 Wh/kg, compared to about a hundred today. The progress made in 100 years has only been a factor of 2 to 3! However, research in the field of energy storage is advancing, and new technologies are being developed. The prospects are therefore encouraging, but they still fall far short of the performance offered by fossil fuels.

Indeed, larger capacity batteries imply longer charging times or higher charging power, which can pose challenges for managing the long-term charging of vehicles, especially during peak demand periods. More powerful chargers also raise significant technical challenges, such as the need for large charging cable sections, complex connector safety systems, and, depending on the power requirements, the necessity of integrating a cooling system into the cable to limit conductor mass, improve ergonomics, and facilitate handling.

Hybridization, on the other hand, is a viable short-term transitional solution. It requires no infrastructure investment and does not significantly alter user habits. Plug-in hybrids represent the second stage of the transition towards fully electric transportation, allowing for a gradual investment in charging infrastructure to maximize the benefits of operating in all-electric mode.

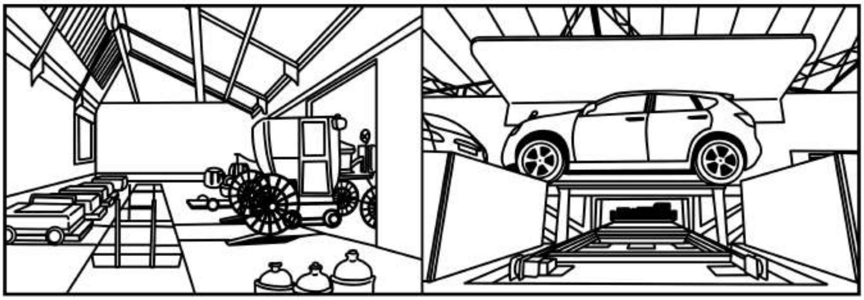


Figure 4.9: Battery exchange station, a) Krieger Parisian taxi, 1898, b) Betterplace, 2010

Another approach would be to provide fast and regular energy replenishment, commonly referred to as "topping up." This operating mode appears to be particularly well-suited for urban public transportation, where mission profiles are known and controlled, enabling sufficient charging points along a route to guarantee the vehicle's full autonomy. However, this is not the case for private vehicles with variable and unpredictable mission profiles.

Battery swapping could be another possibility, along with the widespread development of fast charging stations. Battery swapping, conceived in 1896 and implemented in 1898 for Parisian taxis, between 1910 and 1927 in the United States, and more recently by companies like Betterplace and Renault, appears to face challenges related to standardization among vehicle models and manufacturers, in addition to the relatively complex mechanics of swapping stations.

The transfer of energy without contact is generally referred to as Wireless Power Transfer (WPT), Contactless Energy Transfer (CET), or Contactless Power Transfer (CPT) by English speakers. These terms refer to the general principle rather than the specific energy transfer technology. However, there is often misuse of terminology, with the most common term, WPT, sometimes specifically referring to induction-based technology. The techniques employed for long-distance energy transfer are diverse, ranging from Nikola Tesla's early work in the 1900s with Tesla coils, which allowed energy transmission over long distances using the Earth's atmosphere as an electrical conductor, to highly directional energy transmission using laser technology.

The underlying physical principle is generally based on electromagnetic radiation, although there is also a solution based on mechanical waves. There are at least four different technologies, at various stages of development, that pri-

Vectors	Frequency	Transmissible power	Distance	Performance (%)	Application examples
Acoustic	20 kHz-100 kHz	0.01-1000 W	0.2-1 mm 1-300 mm	$1-80$	Biomedical Sensor for space nuclear
Laser	10^{12}Hz 700-1400 nm	1-100 W	du m au km	$20-30$	
Inductive (magnetic)	20 kHz - 10 Mhz	1 W-200 kW	0.2 mm-2 m	$\geq 80-90$	Biomedical Battery chargers for EVs and mobile devices Actuators
Capacitive (electric)	100 kHz - 10 Mhz	1-50 W	0.1-0.5 mm	$50-80$	Mobile telephony Sensors

Table 4.1: Contactless energy transfer vectors

marily depend on the wavelength used to transmit energy. These technologies include transmission vectors such as lasers, electromagnetic fields, and acoustic waves. Transfer through electromagnetic fields can be achieved by exploiting either the magnetic field vector or the electric field vector.

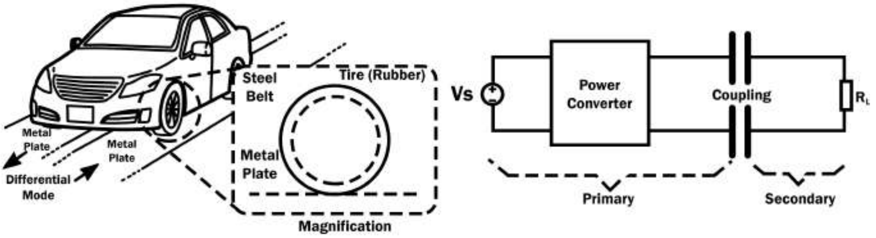


Figure 4.10: Principle of capacitive energy transfer for moving electric vehicles

The contactless energy transfer through capacitive coupling, as illustrated in Figure **??**, can be considered for powering a moving vehicle. Recent studies in this area, particularly by Toyota and the University of Auckland, have shown promise. This method has the advantage of being less sensitive to misalignment compared to induction-based transmission. However, its implementation is complex and requires high voltages across the coupling capacitors. Additionally, the distance between the plates results in low capacitance, necessitating operation at very high frequencies, beyond the megahertz range.

Energy transfer by induction facilitates the exchange of energy between two systems without any electrical connection. Based on the principle of electromagnetic induction discovered by Michael Faraday in 1831, its primary application is in the classic transformer, which features a completely closed magnetic circuit to efficiently channel the generated flux. In wireless energy transmission, the magnetic circuit can be partially or completely removed, increasing the distance between the primary and secondary coils. The opening of the magnetic circuit, which served to perfectly channel the field lines in traditional transformers, leads to significant flux leakage.

The adoption of inductive coupling for energy transmission has grown significantly, driven by the development of high-frequency converters. Numerous applications of this transmission method are being explored or have already been introduced to the market. An example is low-power chargers (such as those for mobile phones or tablets) that adhere to a specific standard known as Qi. This standard defines chargers of less than 5 W with a coil distance ranging from 5 to 40 mm and is supported by a consortium of over 200 companies worldwide.

The transportation sector also utilizes inductive coupling technologies, especially for auxiliary power supply in maglev trains. Additionally, wireless power transfer is being considered for electric or hybrid vehicles, including buses, trucks, and cars. A distinction is made between stationary charging (contactless charging, also known as “static” charging) and charging while the vehicle is in motion (charging or power supply, also known as “dynamic” charging). Static charging systems have seen significant development since the 1990s, with projects like the French Praxitèle and research at the University of Auckland. There is extensive literature on this topic, with many systems developed, leading to a standardization phase in 2014 that involved major automobile manufacturers and some equipment suppliers. One challenge is the large number of developed systems that lack compatibility. Moreover, none of the proposed solutions can be directly applied to a moving vehicle.

Contactless charging while stationary offers the advantage of avoiding cable handling and being less sensitive than wired chargers. However, this charging method still faces disadvantages associated with stationary charging, such as concentrated energy at charging points, complex queue management, and recharge times (and therefore immobilization) equivalent to those of wired chargers. The primary benefits lie in the convenience of use and the aesthetic appeal of the road infrastructure with less conspicuous charging stations.

Indeed, the technology enabling charging while a vehicle is in motion completely changes its characteristics. In a scenario where the road infrastructure is equipped with transmitting coils activated upon a vehicle’s power request, the vehicle’s range becomes primarily limited by the size of the equipped infrastructure.

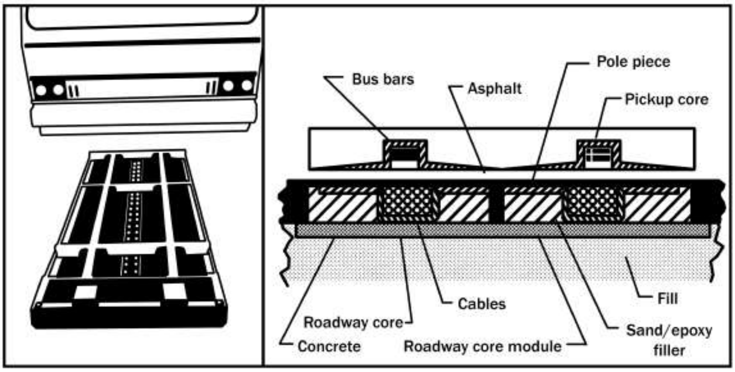


Figure 4.11: Program PATH: (a) receiving coil, (b) cross-sectional view of the road and the receiving coil

The concept is not new; prototypes were developed in the 1980s—1990s, notably as a result of the Californian RPEV (Roadway Powered Electric Vehicle) Project in Santa Barbara, which involved the University of Berkeley. This project led to the creation of an electrified road section that powered a 35-seat mini-bus (Figure ??). The system was capable of transmitting a power of 200 kW with a maximum current of 2000 A, operating at a frequency of approximately 180 to 400 Hz, and with an air gap of about 7.5 cm.

The main constraints of the system were the noise generated by the converters, the efficiency (approximately 60%, considered viable), and the weight. The onboard equipment weighed 850 kg, and each ground coil weighed 400 kg. The onboard coil was 4.3 m long and 1 m wide, while the primary coils were 2.8 m long and 51 cm wide. The equipped road section was 210 m long and included sensors to detect the vehicle's arrival and supply power to the coil underneath it.

The Serpentine project at EPFL (École polytechnique fédérale de Lausanne), carried out in Lausanne, Switzerland, led to the creation of a demonstrator. An autonomous capsule was wirelessly powered during its movements on the pilot site in Ouchy in 2001. This project demonstrated that energy transfer through induction can be achieved with minimal losses, with an announced efficiency of 96%. However, the demonstrator was dismantled in August 2004 due to regulatory issues. The vehicles involved were fully automated and driverless, raising regulatory concerns at the time.

In the 2010s, the Korea Advanced Institute of Science and Technology (KAIST)

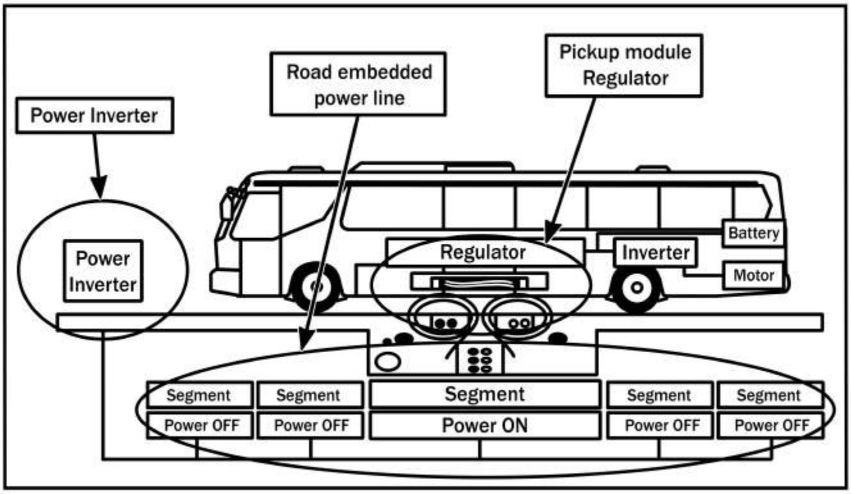


Figure 4.12: Overview of the OLEV system

introduced a contactless power transfer system for a bus in motion. This system was also tested on off-road vehicles and a tourist train. The system claimed an efficiency exceeding 80% and a power capacity of up to 75 kW. Similar to the California RPEV project, only the segment beneath the vehicle was powered (Figure **??**), with a current of up to 200 A at a frequency of 20 kHz.

Bombardier has been developing a contactless power transfer solution for electric vehicles since 2010, resulting in the creation of two demonstrators in different locations (Germany and Belgium). One demonstrator is based on a tramway, and the other on a bus as part of the Flander's Drive project, which includes partners such as Volvo, VanHool, NXP, and others. The project's results have been made public, and the main findings regarding on-the-go energy transfer are as follows: the maximum speed achieved with functional energy transfer is 70 km/h, primarily limited by the power setpoint reaching time of 50 ms for each coil. The transmitted power is 80 kW with an efficiency ranging from 88% to 90%, and an air gap of 10 cm. The vehicle's coil, measuring 2×1 m, is mobile and retracts when not in use to preserve the vehicle's ground clearance. The lateral positioning tolerance is a maximum of 40 cm. The coils are sequentially powered as the vehicle passes over them, detected by an inductive sensor.

Highways England began a dynamic wireless power transfer project in 2015 but canceled it in early 2016 for budgetary reasons. In June 2013, the Swedish Transport Administration initiated an electric road project aimed at developing electrified roads. This project, involving pre-commercial procurement, aimed to gather decision data on electric road platforms in Sweden and lay the foundation for a fossil-fuel-free transportation infrastructure by 2030. A recent pilot project was launched in June 2020 in Lund, Sweden, utilizing patented technology developed by Evolution Road and Lund University. This technology enables the transmission of up to 300 kW of power to vehicles through a retractable pick-up that interacts with a metal rail embedded in the road.

Trafikverket, the agency responsible for Sweden's highway network, has decided to install the country's first "electric road" for vehicle charging in the county of Örebro, located approximately 100 km west of Stockholm. The primary objective of this installation is to promote the use of electric trucks. The selected route for the electric road is a 21 km stretch of the two-lane E20 highway between Hallsberg and Örebro, a key transportation corridor for freight between northern and southern Sweden. While the specific technology to be used has not yet been determined by Trafikverket and the government of Örebro County, they have set a deadline of 2025 to have the electric road operational.

Electric roads have made significant progress within the research community but are now facing a critical phase known as the "valley of death." The foundational technologies for transferring power dynamically from the road to moving vehicles have been developed through numerous research projects worldwide, often with substantial public funding support. While electric road systems are nearing the stage of real-world testing on public roads, they still have a considerable distance to cover before becoming a fully operational, large-scale commercial system.

4.2.8 Smart roads?

Smart highways and roads incorporate electronic technologies to enhance traffic control through features such as intelligent traffic lights and adaptive street lighting. These advanced road systems also enable real-time monitoring of road conditions, traffic levels, and vehicle speeds.

The primary concept behind solar road panels is to utilize road space for electricity generation using photovoltaic panels instead of traditional concrete or asphalt surfaces. Solar road panels can power LED lights to create dynamic road markings like lane indicators or display warning messages such as "Reduce Speed" signs. They can also generate energy for heating elements to clear ice and snow from the road surface and employ wireless charging technology, enabling electric vehicles to recharge their batteries while driving over the panels.

Critics have identified several challenges and limitations associated with solar roadways, including higher costs compared to conventional solar power infrastructure and lower productivity due to the panels' horizontal positioning on the road surface. The need for thicker glass to withstand traffic weight and the lack of proper cooling mechanisms can affect the performance of solar roadways by reducing transparency and leading to overheating.

Revamping the entire US interstate system with solar-powered roads and inductive charging capabilities for electric vehicles would be an immense infrastructure project, with its feasibility and completion time being subjects of debate.

The solar road concept combines photovoltaic (PV) roads, incorporating solar panels and batteries to capture and store sunlight, and inductive roads that can wirelessly charge electric vehicles. The Netherlands initiated a project called Solaroad in 2014, testing a solar-powered bike path, while the Korea Advanced Institute of Science and Technology (KAIST) in Korea has worked on inductive charging technology, successfully charging a bus using this method.

Both PV roads and inductive charging technologies face challenges related to efficiency and potential obsolescence. The efficiency of PV cells can be affected by various loss mechanisms, and the rapid evolution of PV technology means that major solar projects may become outdated by the time they are completed.

An experimental 1-kilometer road in France called Wattway, inaugurated in December 2016, fell apart by August 2018. The road generated only half of the anticipated electricity and experienced substantial deterioration, leading to its labeling as a fiasco by Le Monde.

In the Netherlands, an experimental project known as Smart Highway was initiated in 2012, focusing on the implementation of several innovative technologies to develop "smart roads." On the N329 highway, approximately 100 km southeast of Amsterdam, road markings were applied using fluorescent paint. These markings would absorb sunlight during the day and glow for up to 10 hours at night, providing illumination after dark. Additionally, dynamic road markings were created using thermochromic paint, which responded to changes in external conditions. For instance, when the temperature fell below 3℃, snowflake symbols that were previously invisible would appear, alerting drivers to the potential for icy conditions. In 2012, this project was honored with the Dutch Design Awards and the INDEX Award, both of which recognize projects that have the potential to significantly improve quality of life.

In April 2014, a pilot section of the highway in Brabant, Netherlands, was officially inaugurated to showcase this technology. However, the glow-in-the-dark paint ceased to function after two weeks due to moisture-related issues.

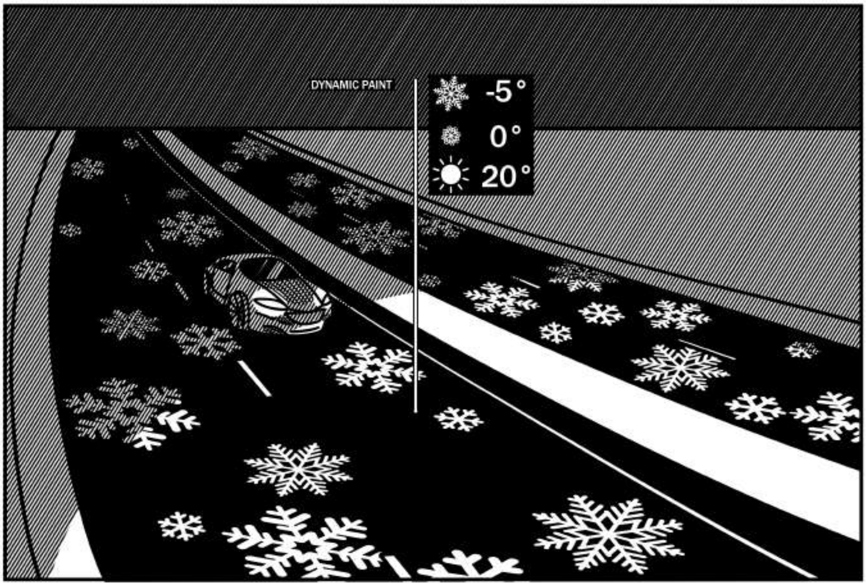


Figure 4.13: Dynamic road markings in the Smart highway concept

Smart roads play a crucial role in the European Union's plan known as "Cooperative Intelligent Transport Systems" (C-ITS), which aims to enable the sharing of information among road users and traffic managers to coordinate their actions. Through innovation, cooperation, connectivity, and automation, Europe aims to make its roads smarter. To facilitate the harmonization of investments and regulatory frameworks across European countries, the European Commission adopted a European Strategy for C-ITS on November 30, 2016.

As part of this initiative, European countries, along with market operators, have established the C-ROADS platform. This platform serves as a common framework for sharing projects and experiences related to the innovation of smart roads, allowing for collaboration and knowledge exchange among participating countries. The aim is to drive advancements in smart road technologies and create a more connected and efficient transport system throughout Europe.

According to the ITS Industry Report 2019 by INDRA [231], smart road technologies are part of a rapidly expanding sector with significant market prospects and opportunities. The report estimates a projected annual growth rate of 7% for this market in the coming years, with a forecasted market size in Europe exceeding 2 billion euros by 2022. According to the latest research report of February 1, 2023, the smart road market size is 110.5 billion by 2030 and is expected to grow at a CAGR of 26.6% during 2020 to 2030.

At a national level, the Italian government has taken significant steps towards the implementation of smart roads. In April 2018, the Italian Ministry of Infrastructure and Transport (MIT) introduced the Smart Road Decree, marking the official beginning of smart road initiatives in Italy. Additionally, the Italian Parliament included references to smart roads in the 2018 Budget Law, emphasizing their importance at a national level.

The Smart Road Decree establishes functional standards to create more connected and safer roads in Italy. These roads will be equipped with technologies that enable real-time communication with vehicles, providing information on traffic, accidents, weather conditions, and other relevant data to enhance travel comfort and infrastructure management.

The MIT has identified initial interventions that will focus on sections of the motorway network, including new construction and areas requiring significant maintenance. The plan aims to prioritize the Italian infrastructures within the Trans-European Transport Network (TEN-T) and the entire national motorway network. Over time, the services will be extended to cover the entire integrated national transport system.

Aligning with the directives set by European and national institutions, Anas S.p.A., a company within the FS Italiane group responsible for managing state-owned roads and motorways, has developed its Smart Road Plan. This plan envisions investing approximately one billion euros over the next decade to establish an efficient road network that is prepared for future challenges.

These efforts in Italy demonstrate the country's commitment to embracing smart road technologies and creating a modern and advanced infrastructure network.

4.3 Traffic code

4.3.1 Emergence of order on the roads

Transportation problems in large cities existed long before the advent of automobiles.

From the ancient manuscripts discovered by archaeologists, it can be learned that even in ancient times, there were restrictions on road traffic. For example, the king of Assyria, Sennacherib (705—681 BC), prohibited leaving carts on the main streets of his capital, warning that anyone who dared to violate the royal decree would be impaled on a stake near their own house. This punishment was in line with the customs of that time.

For millennia, one of the most important regulators of human transportation behavior has been the establishment of specific rules that govern these behaviors. There existed a system of requirements that reflected the development of the state, its roads, vehicles, and the legal sphere. The emergence and development of norms and requirements were driven by the societal needs for safe, fast, and uninterrupted transportation. As mechanical transportation developed, more and more individuals became active participants in the movement of vehicles and the utilization of roadways. This expansion of transportation involvement resulted in a greater need for regulations and safety measures to ensure the well-being and orderly conduct of all individuals on the roads.

As mentioned in the first chapter, one of the earliest instances of road signs appeared on Roman roads. Special signs were installed along the roads indicating towns, distances to various destinations, and directions for turns. There was even a dedicated agency responsible for these matters. As for traffic rules, even during the time of Julius Caesar in 47 BC, the movement of vehicles was prohibited during the first 10 hours after sunrise to organize traffic flow on certain narrow streets of Rome, which already had a population of around one million at that time. At the same time, one-way traffic for horse-drawn carriages was introduced. We can get an idea of the traffic conditions in Ancient Rome from the complaints of Juvenal, who wrote about urban traffic issues:

«If duty calls him, the rich man will be borne through the yielding crowd, and pass rapidly over their heads on the shoulders of his tall Liburnian, and, as he goes, will read or write, or even sleep inside his litter, for his sedan with windows closed

entices sleep. And still he will arrive before us. In front of us, as we hurry on, a tide of human beings stops the way; the mass that follows behind presses on our loins in dense concourse; one man pokes me with his elbow, another with a hard pole; one knocks a beam against my head, another a ten-gallon cask. My legs are coated thick with mud; then, anon, I am trampled upon by great heels all round me, and the hob-nail of the soldier's caliga remains imprinted on my toe.» [145]

In the 12th century, specific rules of road traffic had already emerged on the roads of the Middle Ages. These rules were documented, among other sources, in the Sachsenspiegel or Saxon Mirror. Compiled in 1235 by Eike von Repgow, the Sachsenspiegel is considered one of the earliest examples of German legal literature written in the vernacular. It covered various aspects of medieval law, including regulations related to road traffic.

«The king's highway shall be wide enough that one cart can leave room for another: an empty cart shall make way for a full one, and the one with a smaller load for the more heavily laden one. Riders yield to carts, and pedestrians to riders. If, however, they are on a narrow road, or on bridge, and if someone is being pursued on horseback or on a foot, then the cart should stop so the pursuer can pass. The cart that reaches a bridge first, crosses it first, whether it is empty or full.» [164]

It is interesting to note that according to this document, traffic was prescribed to be left-hand driving. According to popular belief, this was likely due to the fact that many travelers carried weapons at that time. It was believed to be more convenient for them to quickly access their weapons and defend themselves by passing on the right side.

By 1271, according to the accounts of the explorer and traveler Marco Polo, a developed road network already existed in Asia. He wrote:

«There is another regulation adopted by the grand khan, equally ornamental and useful. At both sides of the public road: he causes trees to be planted, of a kind that become large and tall, and being only two paces asunder, they serve (besides the advantage of their shade in summer) to point out the road (when the ground is covered with snow); which is of great assistance and affords much comfort to travellers ...where it is impossible to have trees, he orders stones to be placed and columns to be erected, as marks for guidance. He also appoints officers of rank, whose duty it is to see that all these are properly arranged and the roads constantly kept in good order.» [229]

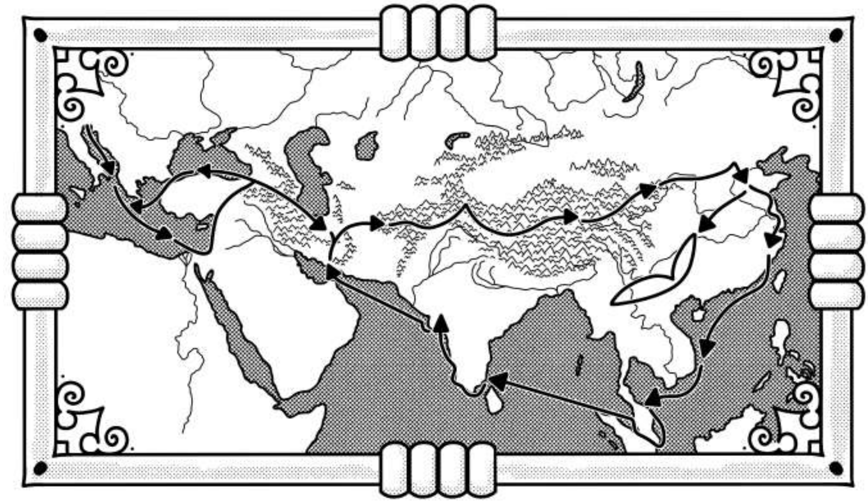


Figure 4.14: Marco Polo. Travels Between 1271—1295

In contrast to creating shaded roads in desert regions, cities faced the issue of illumination. One of the earliest official ordinances regarding nighttime lighting was a decree by the parliament in 1558. In Paris, it was decided to light resin pots on every street corner from 10 PM to 4 AM. This measure aimed to provide some form of illumination during the night hours to improve visibility and enhance safety in the city streets.

In 1822, regulations closely resembling modern traffic rules were established in Great Britain. These rules determined which side of the road vehicles should keep to, specified the appropriate speed for traveling, and set maximum load limits for carriages based on the width of the wheel rims. These regulations aimed to bring order and standardization to road traffic, ensuring consistent practices and promoting safety on the roads.

Concerned about the safety implications of motor vehicles, the British Parliament enacted the Locomotive Act in 1865. Under this Act, a person was required to walk in front of any motorized vehicle for a distance of at least sixty yards while carrying a red flag as a warning of imminent danger [9]. When pedestrians or other road users demanded it, the motorized vehicle had to come to an immediate stop. It was strictly forbidden to open steam valves or emit sharp whistles if horses or people were nearby. Furthermore, in urban areas, it was prohibited to pollute the air with smoke and gas. The Act specified that "every locomotive propelled by steam or any other than animal power to be used on any turnpike road or public highway shall be constructed on the principle of consuming and so as to consume its own smoke" [109]. According to the same Act, the maximum speed for self-propelled vehicles was set at 6 km/h on state and major roads, and 3 km/h on country roads and within populated areas. However, it's worth noting that this law was repealed in 1896, indicating that there were subsequent changes and updates in regulations governing vehicle speed limits.

Of course, public safety was not the only motivation for this law, since the railroads benefited from this restriction on motor vehicle traffic. In fact, laws restrict motor vehicle use far beyond safety requirements. That's a bad thing. Still, the sentiment is good: the public has a right to be warned of potential dangers before society gets used to the introduction of new technologies.

The understanding that speed limits enhance road safety predates the advent of the automobile. In the 16th century, for example, it was prohibited to drive carriages faster than 6 km/h in France. This measure was implemented due to concerns about the danger posed to pedestrians by excessive speeds, considering the prevalence of carriages on the roads.

Similarly, in 1687, in Rhode Island, North America, a penalty was imposed for riding too fast after a child was fatally struck by a galloping horse rider.

Steam carriages, as early forms of automobiles, started encountering road accidents from their initial journeys. In April 1834, John Scott-Russel established the Steam Company of Scotland, which operated steam-powered road transport between Glasgow and Paisley. These carriages covered the distance in 34 minutes, reaching speeds of up to 27 kilometers per hour. Although the stagecoaches could accommodate 26 passengers, they became extremely popular and often became overcrowded. Six carriages of the same type were built for this purpose in Edinburgh.

However, on July 29, 1834, one of the wheels of a heavily loaded carriage collapsed, causing the boiler to rupture and explode, resulting in the deaths of five people. This incident is considered to be the first fatal automobile accident in history. As a result of this tragedy, further steam stagecoach journeys were subsequently banned due to safety concerns.

Steam-powered vehicles, despite having safety features such as safety valves and control systems, posed inherent risks in the design of their boilers.

On August 31st, 1869, a tragic incident occurred involving Mary Ward, a scientist, who fell from an experimental steam car built by her cousin, William, the 3rd Earl of Rosse. The accident took place as the vehicle was turning a corner near the church in Birr, County Offaly, Ireland, the hometown of the earl. Unfortunately, the rear wheel of the car ran over Mary, resulting in a broken neck and instant fatality. Mary Ward is sometimes recognized as the first person in recorded history to have died as a result of a car accident.

In the second half of the 19th century, the organization of traffic flow and pedestrian movement became an increasingly prominent issue in cities. In a Paris guidebook from 1862, it was noted that the city streets experienced an "incredible" amount of traffic, with over $10,000$ vehicles passing through them daily.

Among the many brilliant ideas of the genius Leonardo da Vinci was the concept of separating transportation and pedestrian flows in space. The very first pedestrian and transportation interchange at different levels was constructed in 1851 in Central Park, New York City, marking a pioneering achievement in urban planning and design.

4.3.2 Left or right?

The question of which side of the road to adhere to during travel was particularly relevant for the development of road traffic rules. Let's try to analyze the historical roots of different approaches to the order of movement on the road.

In the existing regulations of old times, there was no norm specifying how one should position themselves on the roadway. The behavior on the road was determined by the social status of the individuals moving on the road. For example, a nobleman would never yield the road to a peasant.

In 1300, Pope Boniface VIII issued a bull in which he instructed pilgrims traveling to Rome to adhere to the left side of the road. However, this instruction applied only to the traffic on the San Angelo Bridge. Another example can be found in the rules of knightly tournaments, which required knights to approach each other, leaving their opponent on the right side.

Indeed, Great Britain is considered the main "culprit" behind the practice of driving on the left side of the road, which later influenced many countries around the world. According to one theory, Britain adopted this order from maritime rules, where passing ships would give way to each other by keeping to the right. In 1756, an ordinance was issued requiring drivers to keep to the left on the London Bridge. In 1772, the rule was extended to Scotland. In 1773, due to increased horse traffic, the UK government recommended left-hand traffic, formalizing the practice in the General Highways Act, which later became part of the Highway Bill of 1835, although some regions of the Kingdom maintained their own rules until that date. Indeed, the left-hand traffic system was also adopted on railways. In 1830, on the first railway line between Manchester and Liverpool, the traffic operated on the left-hand side. The influence of Great Britain had an impact on the traffic regulations in its colonies, leading to the adoption of left-hand traffic in countries such as India, Pakistan, and Australia. In 1859, Sir John Rutherford Alcock, Queen Victoria's ambassador, convinced the authorities in Tokyo, Japan, to adopt left-hand traffic as well.

Another theory explains the introduction of left-hand traffic in England as a concern for pedestrians, who typically walk on the right side of the sidewalk or road to protect themselves from the whip or lash that carriage drivers usually hold in their right hand. However, this theory is doubtful.

According to this theory, after the practice of riding with weapons and suspecting every passerby as an enemy ceased on the roads, a spontaneous shift towards right-hand traffic began. This was mainly related to human physiology and the significant difference in strength and dexterity between the two hands when controlling heavy horse-drawn carriages with multiple horses. When passing each other on a narrow road, it was easier to guide the carriage to the right side or edge of the road by using the stronger right hand on the reins to control the horses. It is likely that this simple reason led to the tradition and eventually the norm of road traffic. This norm eventually solidified as the norm of right-hand traffic.

In the case of France, the organization of traffic has a long history. It was Henri IV, on December 16, 1607, who established the principle of unobstructed and organized traffic: *"We want and it pleases us that when the streets and roads are crowded or obstructed, our grand Voyer or his deputies shall order individuals to remove the said obstructions."* [225] The ordinance of February 4, 1786, provides more specific guidelines and states that certain vehicles will have priority at intersections: *"His Majesty has ordered and commands that all road users, carters, carriers, and others shall yield the road and make way for all couriers and travelers going on post."*

More or less implicitly, the direction of traffic was on the right side since the Proclamation of the Provisional Executive Committee on May 30, 1793. The first article stated: *"Carters, wagoners, and other carriage drivers who frequent the highways shall consistently keep to their right without turning or deviating unless forced to do so by an obstacle, so that their vehicles and those coming from the opposite side, who will also keep to their right, pass each other on the left."* [205]

Users and certain regulations often favored the median use of the roadway as well. In 1804, Napoleon Bonaparte made it mandatory for military personnel to travel on the right side of the road, or even in the center. At that time, only right-side traffic was accepted in Paris. The armies of the Republic, soon transformed into the armies of the Empire, contributed to the spread of the right-side norm. One of the most significant factors was Napoleon's concern for having a swiftly maneuvering army and efficient logistics. However, there is actually no direct connection between the direction of traffic on roads and tactical dispositions in combat. In fact, the decree of August 28, 1808, defined the allocation of a portion of the roadway to each vehicle, although still in a somewhat imprecise manner: "carters, wagoners, and carriers shall be required to yield half of the pavement to the vehicles of travelers."

Further spread of the traffic order, strangely enough, correlated with major politics in the early 19th century. Countries that supported Napoleon, such as Holland, Switzerland, Germany, Italy, Poland, and Spain, switched to right-side traffic. On the other hand, countries that resisted the Napoleonic army, like Britain, Austria-Hungary, and Portugal, remained "left-sided." But is this correlation causality? Certainly not necessarily. Therefore, contrary to a persistent belief, Napoleon Bonaparte did not arbitrarily impose right-hand traffic on most of continental Europe. Simply put, if French customs spread during his conquests, it was because there was a French specificity. It is difficult to accept, as one can find on certain websites, that France, as the eldest daughter of the Church, chose right-hand traffic to respect a very small bull of Boniface VIII. Paradoxically, such an argument is never accompanied by the fact that out of superstition, a Catholic would naturally tend to move to the right (towards the Lord). By the way, it is worth noting that Denmark was the first independent country in Europe to truly adopt the right-hand rule in 1793.

If we consider not only countries in Europe as a whole, it can be stated that in Russia, even in medieval times, the rule of right-hand traffic developed spontaneously and was observed as natural human behavior. Just Juel, the Danish envoy to Peter the Great, wrote in 1709 that in Russia, it is customary everywhere for carriages and sledges to pass each other by keeping to the right side.

Approximately 50 years later, the decree of August 16, 1852, transformed the obligation to yield half of the pavement, which had been confirmed by various intermediate texts, into an obligation to keep to the right to free up the roadway during vehicle crossings. It wasn't until the ordinance of August 14, 1893, that the choice of the right side of the roadway for circulation, even in the absence of crossings, was definitively established in France.

However, in England, Portugal, Sweden, and some other countries, the traffic remained left-hand oriented. Austria had an interesting situation, with left-hand traffic in some provinces and right-hand traffic in others. Only after the Anschluss with Germany in the 1930s did the entire country switch to right-hand traffic.

In the beginning, left-hand traffic was also present in the United States. Perhaps it was the "freedom-loving" spirit of Americans that led them to do things differently from what was customary in England. According to numerous internet rumors, the French general Gilbert du Motier, Marquis de Lafayette, who made a significant contribution to the struggle for independence from the British crown, “convinced” Americans to switch to right-hand traffic. Meanwhile, Canada maintained left-hand traffic until the 1920s.

At that time, transportation on the roads was relatively limited, and it moved at a low speed. In reality, the issue of which side of the road to travel on was not as significant. The roads were narrow, and vehicles could pass each other by giving way to those who were stronger, wealthier, or of a higher social status.

4.3.3 The first road and traffic signs

Orientation on the road has been an important issue throughout history. When humans first attempted to venture beyond their place of residence, they had to mark their path. The first signs appeared at road intersections. Often, columns with carved crosses were installed at such crossings. It is believed that the cross encapsulates one of its fundamental purposes as a symbol: the idea of choice.

The example of a marked fork in the road appears in the Book of Ezekiel (Ezekiel 21:19–23 NRSV): *«Mortal, mark out two roads for the sword of the king of Babylon to come; both of them shall issue from the same land. And make a signpost, make it for a fork in the road leading to a city; mark out the road for the sword to come to Rabbah of the Ammonites or to Judah and to Jerusalem the fortified.»*

The first kind of law regarding road traffic was passed in England in 1555, making road maintenance the responsibility of the parish. In 1686, King Peter II of Portugal established the first known Traffic Regulation Act in Europe. This act included the placement of priority signs in the narrowest streets of Lisbon, indicating which traffic should yield to give way. One of these signs still exists today in Alfama, on Rua do Salvador, right next to number 26. Unfortunately, it is located above an electrical box. It states: *«Year of* 1686. *His Majesty commands all coaches, carriages and litters coming from Salvador's entrance to back up to the same part.»*

It is possible that one of the first road signs with a symbol was used on mountain roads in present-day Austria and Switzerland to warn of steep descents. The sign depicted a wheel or a brake shoe and was painted on the side of the road on a rock. Later, shields with the inscription "Place of braking" began to be installed. According to one popular legend, such warnings were put in place after Frederick Augustus II of Saxony became a victim of an accident, as indicated by a memorial plaque located near the Austrian village of Brennbüchel near Imst.

During the era of Cardinal Richelieu in France (1624—1642), decrees were issued that designated the most important squares and road junctions with crosses, poles, or pyramids to facilitate orientation for travelers. In the late 17th century, in England, under the reign of King William III, a special law was enacted that obligated magistrates to install specific indicators. The further dissemination of road information was connected to the development of the postal service. These indicators consisted of small boards attached to poles, allowing travelers to easily read the inscriptions on them. During the same period, milestone markers and sundials also emerged.

Indeed, there were already some road signs before motorization, but their widespread appearance and evolution into a complex system to manage road traffic went hand in hand with the diffusion of modern means of transportation such as bicycles, motorcycles, and automobiles. The powerful and fast automobile asserted itself over previous road users such as coachmen, cyclists, horseback riders, and pedestrians, who had to yield and the road transformed from a space open to all into a lane reserved for automobiles. This led to strong social opposition encountered by the newcomer in road traffic during the initial phase of its expansion [185].

The first set of traffic rules in the world was adopted in France on August 16, 1893 [55]. It was during this time that the Prefect of Police in Paris decided to regulate the street traffic of the newly emerging automobiles. The country already had 600 cars, mostly concentrated in the capital city of France. In the city, a list of requirements was established regarding the movement of mechanical carriages on streets and roads. It was prohibited to drive or park on sidewalks, alleys, and areas designated exclusively for pedestrians. The speed limit in the city was set at 12 km/h, while outside the city, it was limited to 20 km/h.

Indeed, it was during that time that the traffic rules took their final form, legally establishing the concept of a "car driver." They began to be implemented in many countries around the world. For example, the first such rules in England were adopted in 1896. They prohibited the movement of automobiles at a speed exceeding 12 miles per hour (19.65 km/h) in all cases.

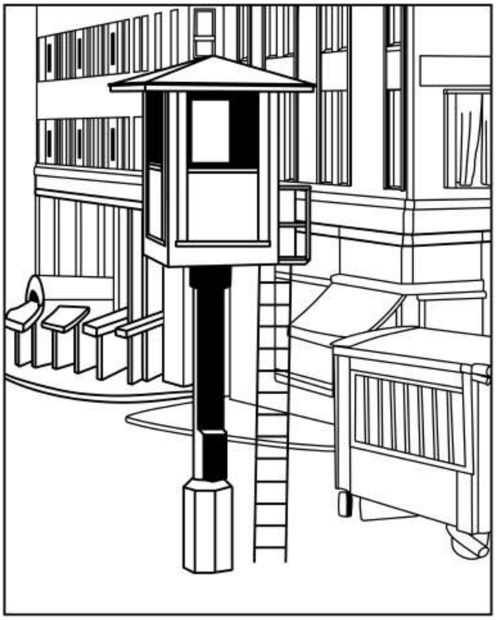
If in the very first traffic rules, priority was given to horse-drawn transport and automobiles were required to yield to horse-drawn vehicles, the Paris Traffic Rules of 1900 already stated: *"It is not allowed to harness a horse and drive it on the streets if it is frightened by the movement of mechanical vehicles; the owner of the horse must refrain from riding it, and if he does so anyway, he violates the order and is responsible for it."*

The Automobile and Touring Clubs deserve credit for the development and advancement of road signage. These clubs were primarily concerned with safety and aimed to prevent dangers and enhance the journeys of their members. They accomplished this by drawing attention to points of interest or scenic viewpoints and by warning of potential hazards such as narrow passages, sharp turns, potholes, or steep gradients. Initially, the Touring Club de France focused on installing signs indicating locations and directions, and by 1908, around 10, 000 such signs had been installed. Obstacle signs were introduced in 1904 to further enhance road safety [71].

The most commonly used shapes for signs in different countries were rectangles and circles. For example, the sign indicating the name of a locality looked quite different from modern signs. It had a blue background with a horizontally positioned arrow and the name written in white. Prior to World War I, circular signs with English and French texts were widely used in England. The rectangle is the simplest and easily manufactured shape. Only later did people start considering that this shape could convey additional information. In Italy, by 1904, there were already 400 signs installed. They were made of sheet metal. During the same period, paint that reflected light from car headlights at night started being applied to the signs. In England and France, mirrors were installed in obstructed and dangerous areas. Later in Italy (in 1910), there were already 11 types of road signs in use. Among the signs were: “Slow down” — a red dot on a white background; “Road in ruins” — a broken arrow in the middle; “Section of road where accidents can occur” — an arrow pointing diagonally downwards. “Dangerous turn” (closed arrow); “Damaged road”; “Uneven road”; “Railway crossing.”

4.3.4 Traffic lights

Alongside the development of requirements for road traffic, it became increasingly clear that road safety could not be ensured through prohibitions and recommendations alone. The first contentious issue emerged at intersections with heavy traffic. For example, in the mid-19th century, the need arose to regulate the movement of horse-drawn carriages, early automobiles, and pedestrians at busy intersections in some major cities around the world. This task was often performed by a traffic regulator or a police officer. While they managed their duties adequately, many inconveniences arose, particularly due to adverse weather conditions. People had to stand at intersections in rain and snow, heat and cold.



The solution was soon found. Experience with railways already existed. There, due to high speeds and significant braking distances, the semaphore was invented. It effectively regulated the priority of trains passing through a single track. The same approach was then adopted for regulating traffic at road intersections.

In 1868, in London, at the square near the building of the English Parliament, a traffic signal was installed. It was designed by John Peake Knight, a mechanical engineer and specialist in railway semaphores. The traffic signal was manually operated and had two positions: the first allowed cautious movement, and the second prohibited movement. Over time, the semaphore arms were replaced with gas pipes. During the day, the signal displayed information using arrows, and at night, it used a gas lamp. However, at a later time (around 1896), the gas lamp of the traffic signal exploded, injuring the police officer who operated it.

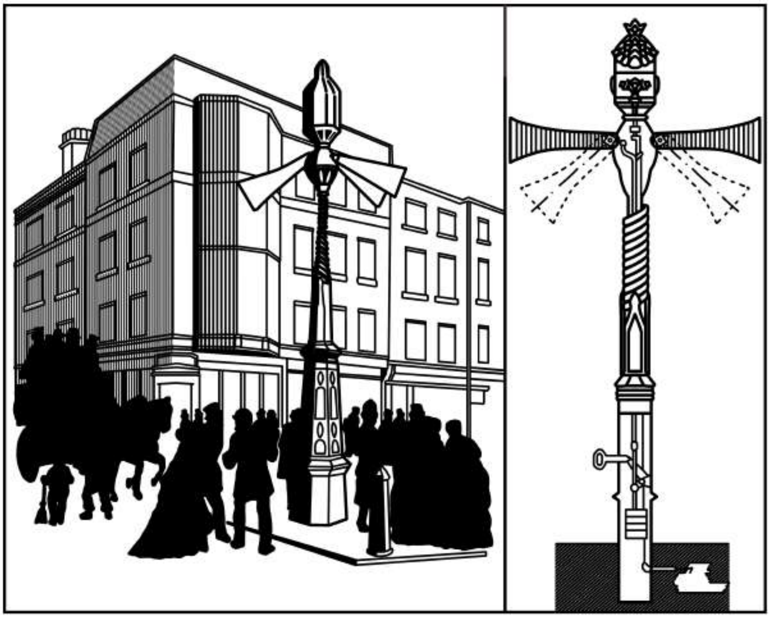


Figure 4.15: Traffic signal

The first automated traffic signal system, capable of switching without direct human intervention, was developed and patented in 1910 by Ernest Sirrine. His traffic signal displayed the inscriptions "Proceed" and "Stop," which were not illuminated. Lester Wire is acknowledged as the inventor of the electric traffic signal. In 1912, he devised a traffic signal with two round electric signals — red and green — but did not patent it. The subsequent year, James Hogue was granted a patent for a manually controlled traffic signal featuring red and green lights, installed in 1914 at 105th Street and Euclid Avenue in Cleveland. This system produced an audible signal when transitioning and was operated by police officers stationed in a glass booth at the intersection.

On May 1, 1917, William Ghiglieri of San Francisco was awarded a patent for the first automatically operated traffic signal using red and green colored lights, with an option for manual operation.

In 1920, Detroit saw the installation of three-color traffic signals incorporating a yellow light, invented by William Potts and John F. Harris. Europe followed suit in 1922, and England in 1927, with the first Soviet traffic signal installed in Leningrad in 1930.

Over time, traffic signals have evolved to not only facilitate pedestrian crossing but also optimize traffic flow, especially at intersections.

4.3.5 Road safety

With the advent of automobiles, it became evident that they posed a significant danger and were involved in road incidents, akin to other means of transportation.

The first recorded road incident involving an automobile and resulting in injury occurred in England on August 17, 1896, during the World Exhibition near London. As reported by the Norwood News on August 22, 1896, three German-manufactured, French-assembled cars were showcased at the Dolphin Terrace, behind the palace.

At this time, Mrs. Bridget Driscoll from Croydon, accompanied by her daughter and a friend, was en route to the premiere of a folk dance ensemble at the Crystal Palace in London. Engrossed in conversation, they overlooked a road sign warning, "Attention! Horseless carriages." Suddenly, an automobile emerged. The inquest revealed Mrs. Driscoll hesitated in front of the car, appearing "bewildered" before being struck.

«"The car then swerved off, and [the] witness looked to see where it was, and it was then going over her mother. (Here witness broke down.) Her mother was knocked down, and the car was at once pulled up," reported the local newspaper, using rather equine terms. Tragically, she was run over and sustained fatal injuries. The driver was acquitted after a trial, as the car's speed was only 4 *miles per hour (*6.4 *km/h). Following a six-hour inquest, the jury concluded it was an "accidental death."»*

Racers, too, faced mortal dangers and injuries. On May 1, 1898, during the Périgueux-Bergerac-Périgueux race, the Marquis de Montaignac became the first driver to die in a car race. Attempting to wave to Mr. Montariol, whom he was overtaking, the Marquis de Montaignac collided with his rival's car, causing both men and their vehicles to crash into a field. Mr. de Montaignac succumbed to his injuries three hours later.

The specifics of the first car-to-car collision remain uncertain. Despite popular tales of a two-car collision in Ohio in 1895, involving the state's only two cars, concrete evidence is lacking, leaving the first two-vehicle crash a mystery in automotive history.

At the dawn of the 20th century, France emerged as a leader in automobile production. In 1904, there were 14 cars and motorcycles per $10,000$ inhabitants. That year, Paris reported 324 road accidents involving motor vehicles and $2,704$ involving horse-drawn carriages.

The hesitancy regarding road safety stemmed from the complexity of the issues at hand. It was unclear whether France should adopt right or left-side driving. As late as 1911, a non-parliamentary commission rejected left-side driving. This indecision also arose from various commission outcomes [98]. The 1904 commission issued a vague regulation, not specifying a maximum speed limit or municipal powers. The 1905 Parisian commission, convened by the Municipal Council of Paris, struggled over the maximum speed limit and police powers. The 1909 commission advised motorists to adhere to local driving customs, while the 1911—1914 commission was evenly divided between those advocating for and against speed limit specifications, provided driver responsibility was simultaneously increased [140].

Insurance companies and mutual insurance associations also influenced the hesitancy. Despite inconsistent profits from automobile risks, they opposed mandatory insurance for motorists, fearing premium regulation. They, along with pro-automobile lobbyists, favored voluntary driver education [299]. France introduced mandatory insurance only in 1958.

Automobile-related lobbying groups noted that the majority of accident victims were pedestrians, a trend that persisted across Europe until the 1980s. They humorously proposed a Pedestrian Code in 1907 and a walking permit in 1909 [192].

On June 6, 1898, *le Journal* published an open letter from Mr. Hugues Le Roux to the police prefect, stating:

«*Paris, June* 6, 1898.

Mr. Prefect of Police,

Yesterday evening at six o'clock, near the Rue de Courcelles, my wife, my children, and I narrowly escaped being crushed by a gentleman driving an automobile at the speed of a locomotive.

Of course, it was impossible to catch up with him. The officer I approached — asking if this gentleman lived in the neighborhood and if we had a chance of finding him — replied, "Alas, Sir, we are powerless against these people. They know they will escape by fleeing..."

Mr. Prefect of Police, it is not in six months, but tomorrow, that you must compel these reckless individuals to prominently display a number that will allow them to be identified after their escape.

In the meantime, I am among those who believe that safety does not exist in the streets of Paris. And since your officers claim to be powerless, I have the honor to inform you that starting today, I will carry a revolver in my pocket and I will shoot the first mad dog who, riding in an automobile or a petrol tricycle, flees after endangering my family or myself. Please accept, Mr. Prefect of Police, the expression of my sincere regards.

Hugues Le Roux.»

Indeed, the impact of the famous writer's pamphlet was not as negative as one might think.

4.3.6 Let's count the cars

Businessmen Daniel and Hermann Beissbarth, close friends of Daimler mentioned in the second chapter, were automotive pioneers in Bavaria at the end of the 19th century. On April 14, 1899, at the Munich Police Department, they received the very first vehicle registration plate, number "1". This plate is now exhibited in the German Museum, alongside Germany's first driving license.

They were read a special directive regarding the rules of city traffic, which stated, among other things, that "within the city, it is prohibited to exceed a speed of 12 kilometers per hour, to drive through the English Park and the reserve, as well as to approach carriages and horses too closely."

The first precursor of modern number plates appeared in France in 1783, when King Louis XVI ordered Parisian coachmen to display a badge with their name and address on their carriages to reduce crime in the streets of Paris. Subsequently, in 1901, a nationwide number plate system was introduced in France. In New York, on April 25, 1901, all car owners were required to register their vehicles and display number plates, which cost one dollar. In 1903, the UK issued its first registration numbers. In Russia, car license plates appeared in 1904, not in the capital at that time, St. Petersburg, but in Riga.

Initially, there were no license plate standards, and every motorist could make or purchase a plate from craftsmen and attach it to the car. It was customary to indicate the number and the name of the city or region where the car was registered on the number plates. Initially, cars were assigned sequential numbers based on their registration in the city. The thing is, traveling between cities was not a common occurrence. For example, V.I. Lenin in Moscow used to drive a Rolls-Royce, which was registered and given the number 236 before the outbreak of World War I.

By the license plate, it is easy to determine the owner or driver of a vehicle, for example, in the case of traffic rule violations or road accidents. This is especially relevant if the driver intends to flee, as mentioned in Le Roux's letter.

There are many fascinating legends associated with license plates. It is claimed that in 1901, Berlin merchant Rudolf Herzog, with the permission of the city administration, attached the number IA1 to his car. IA represented the initials of his young wife Johanna Anker, and the number one signified that she was his first and only wife. However, it was just a regular license plate. When looking at photographs from that time on the streets of Berlin, one can also notice other cars with numbers like "IA36" or "IA109." It seems there were many lovestruck drivers.

4.3.7 Speed

Speed is one of the most valuable qualities of transportation. With the invention of the automobile, speed began to increase rapidly. Racing drivers set speed records.

On July 22, 1894, Le Petit Journal held an "International competition on horseless vehicles." One of the motivations for organizing car races was the popularity of bicycle races that were frequently held in the 1880s and 1890s.

The competition was exclusively reserved for inventors and builders of motorized cars. Those considered as such were the owners of cars who had made significant modifications to their vehicles, rendering them a distinct type according to the jury's discretion. There were 102 registrations, but only 21 competitors participated in the final race from Paris to Rouen, covering a distance of over 126 km. This was not a speed race but rather an endurance race, as the maximum speed allowed was limited to 12.5 km/h.

Jules-Albert de Dion organized the first speed race in 1895, which was more than 1, 000 km round trip between Paris and Bordeaux. The race marked the triumph of petrol engines.

The first recorded land speed record dates back to 1898 when French driver Gaston de Chasseloup-Laubat reached a speed of 63.158 km/h in his Jeantaud Duc equipped with a 63 horsepower engine. Just a year later, on May 1, 1899, near Paris, Belgian driver Camille Jenatzy set a new record by reaching a speed of 105.88 km/h in an electric car called "La Jamais-Contente." This marked the first time a car exceeded the 100 km/h mark. In 1902, the record set by "La Jamais-Contente" was surpassed by a steam-powered car owned by Léon Serpollet, which reached a speed of 120.77 km/h.

The longest-standing steam land speed record was achieved in 1906 when the Stanley Steamer, driven by Fred Marriot, reached a speed of 195.65 km/h. In 1909, the Blitzen-Benz, a 200 horsepower (150 kW) internal combustion engine car, became the first car to surpass the 200 km/h limit. On November 8, Victor Hémery broke the world speed record at the Brooklands Circuit, covering a kilometer at an average speed of 202.700 km/h.

The pursuit of setting records was not solely for advertising purposes. It served

Figure 4.16: Mr. Bouton and Mr. Count de Chasseloup-Laubat in a steam-powered dog-car, 1885

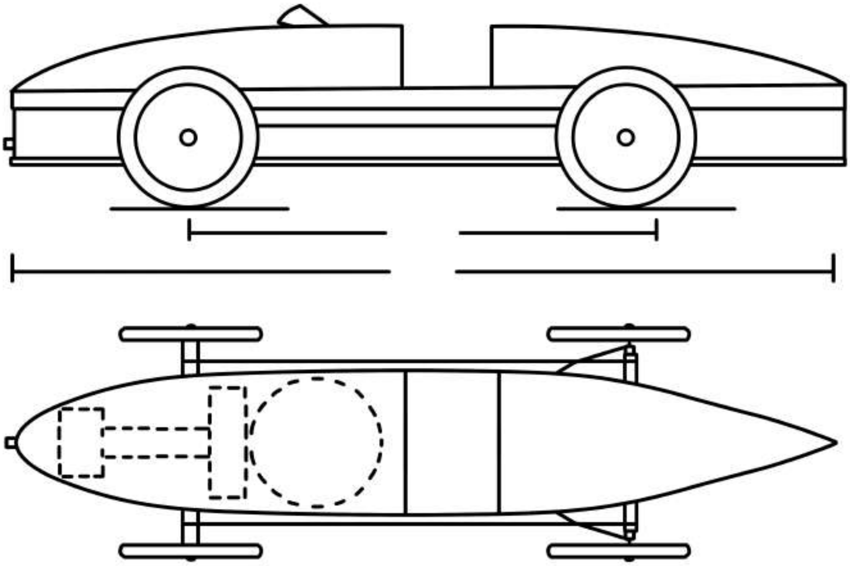


Figure 4.17: Stanley Steamer

as a means to explore new car designs that were safer, more fuel-efficient, and more comfortable. However, alongside the desire to increase vehicle speed,

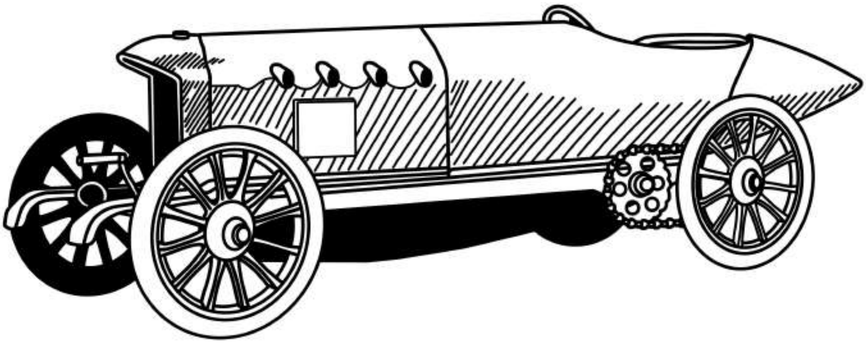


Figure 4.18: Blitzen-Benz of Bob Burmann, 1911 (228.1 km/h)

there was also a need to control and limit it within certain boundaries to prevent accidents and mishaps.

The understanding that speed limits contribute to traffic safety existed long before the invention of automobiles. As early as the 15th century, certain French cities restricted the speed of horse-drawn carriages to a walking pace, not exceeding 6 km/h, recognizing that surpassing this limit posed a danger to pedestrians.

The resistance towards this new mode of transportation was evident in the first traffic regulations published around the turn of the 19th and 20th centuries. Speed limits were set arbitrarily and were often much lower than the actual capabilities of the cars themselves.

Almost every country aimed to restrict the speed of vehicles. Furthermore, different cities within the same country could enforce varying maximum speed limits. For instance, in Germany, the speed limit in Berlin and Potsdam was set at 25 km/h, while in Dresden and Leipzig, it was limited to 20 km/h. Germany, in some cases, exhibited an official and bureaucratic tendency, filling regulations with minute instructions on what drivers should or should not do. Speed limits in cities were sometimes reduced to less than four miles per hour. In certain provincial towns, a complete prohibition on higher speeds was in place to avoid, as the German officials would say, "frightening children and horses."

In 1902, in the United Kingdom, the House of Commons passed regulations that established a maximum speed of 16 km/h for automobiles with metal-rimmed wheels and a load on the axle between 3 and 6 tons. For speeds exceeding this range, the limit was set at 13 km/h. In Austria's traffic regulations, the speed limit in populated areas was restricted to 15 km/h, while outside of urban areas, it was set at 40 km/h. In Italy, according to the regulations approved in 1905, the maximum speed within populated areas was not to exceed 12 km/h, while outside of populated areas, it was limited to 40 km/h. In Madrid, in 1903, the permitted speed was no more than 8 km/h. In Switzerland, the regulations allowed for a speed limit of 30 km/h on roads in flat areas outside of populated areas, described as "like a horse trotting," while for mountain roads, the limit was set at 10 km/h.

Stomach wrote in The Lancet in 1902:

"Our present highways were laid out for horse traffic, and in the palmy days of coaching the best coaches travelled at a rate of ten miles per hour including stoppages. If a motor-car goes faster than 15 *or* 18 *miles per hour the road is being used, very generally, for a purpose which its angles and gradients and breadth do not suit ...Here it is the heavy of a good motor running safely at a very high speed along a clear course, we greatly doubt whether the motorists will gain much by the abolition of the* 12 *miles an hour rule — which, indeed, as it stands, is too often regarded as a licence for overriding the common law in regard to driving to the common danger."* [267]

But, as said John Scott-Montagu in 1904, *«"Speed" and danger are not equivalent terms, any more than "slowness" and "safety"; and the use of discretion by the driver is a better guarantee of safety to the public than the mere limitation of miles per hour.»* [255]

To conclude the discussion about speed limits, it will be necessary to cite Xenophon Huddy, who in 1905 concluded his article about the status of the motor car by the following:

«One who operates an automobile should possess something more than expert knowledge of the machine's construction, or the best mode of operating to obtain the greatest power and speed. A motor vehicle is not a machine of danger when controlled by an intelligent prudent driver ...It is evident therefore, that it is in the manner of driving the machine, and that alone, which threatens the public safety. The ability immediately to stop, its quick response to guidance, its unconfined sphere of action, seem to make the automobile one of the least dangerous of conveyances if properly driven.» [130]

4.3.8 Who can drive a car?

The number of motor cars and motor cycles registered in December 1907, in the United Kingdom was 123, 973 (Lord Montagu gives the number at 125, 320).

Society is starting to realize that driving a car is a complex task and that a car is a source of increased danger. Therefore, the person who operates it must possess certain knowledge and skills in this field. When a person first sits behind the wheel of a car, they face numerous difficulties and challenges. The very first of these is overcoming the psychological barrier. How can one control such a roaring and growling “monster”? Skeptics claimed that a person in an open car would not be able to withstand a speed exceeding 30 km/h, as the air pressure would be so strong that it would cause suffocation. That is why the early “drivers” were perceived as extraordinary individuals, brave and physically strong. Their external accessories reflected this perception: leather jackets and pants, high-laced boots, gauntlet gloves, driving goggles, caps, and helmets with earpieces. Even after the introduction of windshields and enclosed cabins, such driver attributes persisted for a long time.

As Clemens wrote in 1906, *«An ideal driver of motor cars must be an excellent judge of pace and distance, keenly on the alert to seize every opportunity, and ready to act quickly on any emergency. All these qualities call for perfect and binocular vision, and must of necessity cause considerable eye and nerve strain even by day. ...; and I certainly think that all would-be motor drivers ought to produce satisfactory evidence of good vision before being granted licences.»* [50]

When and why did driver’s licenses appear? As we have already discussed, many aspects related to automobiles during its early stages were borrowed from horse-drawn transportation: horsepower, carriage, brakes, and traffic rules. While the automobile was still an exotic means of transportation, driver’s licenses simply did not exist. The skilled craftsmen who built the first models were the ones who drove them. This continued until the automobile shed its image as a technical curiosity and began to assert itself as a widely accessible, advantageous, and popular mode of transportation.

The birth of driver’s licenses was influenced by three factors. The first factor was that operating and driving an automobile turned out to be a complex task. Even today, with significant advancements and increased technical knowledge among people, not everyone can simply sit behind the wheel and drive.

Another factor that led to the introduction of driver's licenses was the occurrence of the first accidents involving automobile drivers. Due to a lack of specialists knowledgeable about automobiles, not every place could establish driving schools. However, the demand for the new mode of transportation was so high that many people purchased automobiles without knowing how to handle them properly and relied on instructions to learn. This situation particularly arose in the United States, a rapidly developing industrial country, where there were many individuals behind the wheel who lacked sufficient training and had poor knowledge of traffic laws, resulting in road accidents. As a preventive measure, special licenses were introduced for drivers of commercial vehicles in New York in 1899. These licenses ensured that the person behind the wheel was familiar with traffic rules and laws.

The third factor, often considered the "culprit" behind the introduction of driver's licenses, was the implementation of vehicle taxes. In 1910, a journalist in Germany proposed that, when paying the annual tax, the vehicle owner should also undergo a practical examination to assess their ability to operate the vehicle.

At the turn of the 19th and 20th centuries, automobile garages and driving clubs were being established everywhere, aimed at training automobile drivers. The Motor Carriage Supply Company of London offered the first driving lessons in Britain in June 1900. The Liver Motor Car Depot and School of Automobilism of Birkenhead was the first organization in Britain to officially call itself a driving school. It was established in May 1901 by William Lea, with Archibald Ford serving as its Chief Instructor.

To obtain a certificate (diploma) for the right to drive an automobile between 1900 and 1909 in European countries, strict requirements were put in place for examination candidates. The oral exam required demonstrating knowledge of the vehicle's mechanisms, traffic signs, and how to behave in specific situations. The practical exam assessed the candidate's ability to start the engine, move off, and maneuver on a road with busy traffic.

The first driving license was issued in 1888 when the district administration department of the Grand Duchy of Baden in Germany issued a document to Friedrich Karl Benz, the owner of the "Rheinische Gasmotorenfabrik" (Rhenish Gas Engine Factory). The document allowed him to take a test drive with his patented motorized carriage in the vicinity under the condition that he would be held responsible for any incidents caused by the motorized vehicle. Thus, Friedrich Karl Benz became the first holder of a driver's license.

In 1896, Henry Ford acquired the first automobile license in America. However, it was not a license in the sense and quality we recognize today but rather a "permit" to operate an automobile under specific conditions. For instance, Henry Ford was mandated not to exceed a speed of 12 km/h in "open country" and 6 km/h in populated areas, and to equip the vehicle with a "moderate sounding signal."

Following the enactment of a law on August 16, 1893, in France, it was resolved that an examination would be conducted for anyone desiring to drive an automobile on the streets of Paris. Administered by the Chief Mining Engineer, responsible for the safety of the steam boilers in steam-powered automobiles, the examination was available to individuals not younger than 20 years old. Candidates were required to demonstrate their ability to *"start and stop the engine, operate the machinery, fix minor breakdowns, and most importantly, know how to avoid an explosion."* Upon successfully passing the examination, the driver received a document (akin to a modern driver's license) certifying that *"the bearer, in the presence of the examiner, demonstrated composure, presence of mind, sharp eyesight, the ability to adjust the speed and direction of the vehicle depending on the circumstances, and a willingness to participate in road traffic in public places."* Marie de Rochechouart de Mortemart, more commonly known by her married name, the Duchess of Uzès, gained renown as the first woman to obtain the certificate of competency for driving vehicles in 1897. The regulations for automobiles continued to evolve, leading to the introduction of technical passports for automobiles in France in 1899.

In 1900, the "Mandatory Decree on the Regulations for Passenger and Cargo Traffic in the City of St. Petersburg by Automobiles" was enacted, necessitating a special permit from the city administration for drivers. Applicants were subjected to rigorous medical examinations. Even then, the city administration provided drivers with a special booklet containing printed rules for driving automobiles and an attached numbered ticket (metal plate) for each vehicle. The minimum age requirement for drivers was set at 17 years old.

How did it work abroad in 1906, based on the article by Edward Manson? [178]

In Austria, driving licenses were necessary and were only issued to individuals over eighteen years of age who had passed the official examination. *«The latter* *involves knowledge of the vehicle's mechanism and a practical test of the ability to drive. The license ...shows which kinds of vehicles the holder is authorized to drive. The military are examined by their own authorities, and foreign drivers are exempt from these regulations during a stay of three months if they have an official driver's license from a country with similar regulations and which exercises reciprocity; otherwise, they must apply within eight days.»*

In Brussels, drivers of motor cars and motor cycles with more than two wheels must be eighteen years of age. Drivers' licenses and examinations were abolished in 1900, though they were said to have been satisfactory, and the police would not be sorry to see them reintroduced. It was suggested that their abolition was because some members of the Legislature were largely interested in the motor-car industry.

In France, an examination of drivers, conducted under the direction of local officials of the Mines Department, is enforced. It is supposed to last about fifteen minutes and includes a drive around streets with a viva voce examination on construction, working, and repairs. Licenses issued by certain foreign countries are usually accepted by the French officials — as, it is stated, are those given by the British Automobile Club in practice. *«A Government Departmental Committee (*1905*) has recommended instituting an age limit for drivers of motor cars (eighteen years for cars up to* 35 *HP and twenty-one for cars above) and enforcing the production of a medical certificate before a driving license can be issued.»*

In Germany, *«Motor vehicles may only be driven by persons who thoroughly understand their mechanism and handling, and can show a certificate to that effect from an officially recognized expert.»* No one under eighteen years of age may drive a motor vehicle, especially a motorcycle, except with the consent of the police authorities and the person's legal representatives. Drivers are responsible for ensuring their vehicles are in proper order and that the registration certificate is carried. Drivers of foreign cars had to have their driving license endorsed by a German authority.

In New Zealand, there were no driving licenses at that time.

In ten states of the USA, driving licenses were necessary. In three of these states, the licenses were valid for a year only, and in four, some examination or evidence of competency was enforced, in some instances for hired drivers only.

In 1915, during the Regular Session of the Indianapolis Common Council, the following conditions were proposed for public hack driving [54]:

«(b) No person shall be so licensed unless they fulfill the following qualifications:

1. *They must be of the age of* 21 *years or over.*

2. *They must be of sound physique, with good eyesight and not subject to epilepsy, vertigo, heart trouble, or any other infirmity of body or mind which might render them unfit for the safe operation of a public hack.*

...

4. *They must not be addicted to the use of intoxicating liquors.*

...»

4.3.9 International trips

At the beginning of the current century, one of the road signs featured a horizontally positioned arrow with the name of the locality. The sign had a rectangular shape, and the arrow and text were drawn on its background. The top part of the sign usually displayed the emblem of the club that installed it (at that time, tourist clubs were primarily responsible for installing signs), as well as the name of the person who donated funds for the production of the sign.

The circumstances surrounding the appearance of one of its main elements, the arrow, are interesting. The arrow had already served humanity as a weapon in hunting and wars. Even earlier, the meaning of the arrow was clear to everyone — it was affixed to posts, trees, drawn on stones, and indicated direction. Starting from the Middle Ages, clock hands were shaped like spears or arrows. The arrow was widely used in cartography and surveying. The tip of the arrow was an important element.

This is one of the earliest symbols that was understood by everyone.

The second half of the 1890s and the first decade of the 20th century marked the beginning of meaningful actions, collaborative efforts, and organizations of motorists.

Motoring magazines started to be published, and motorists' clubs were established. In 1895, an Italian tourism club installed 40 cast iron signposts on the coast of Senegal.

On August 4, 1898, at the Casino Bourgeois in Luxembourg City, the Ligue Internationale des Associations Touristes (LIAT) or International Touring Organizations was founded. It was an international federation of motoring organizations created to represent the interests of national automobile associations and touring clubs. On May 30, 1919, it changed its name to the Alliance Internationale de Tourisme (AIT).

By the early 1900s in Paris, the LIAT Congress began discussing standards for road signage. In 1909, four pictorial symbol signs were chosen as a standard by nine European governments in their respective areas. It was proven that these symbols were perceived faster than words, which was especially important in

Europe, where various nations with their own languages coexisted closely.

In the United States, during the early 1900s, there was a growing demand for signs to accommodate the expanding automobile industry. Drivers were frequently getting lost due to the lack of signage, and the existing signs were often damaged or inadequate. This led to a growing awareness of the need for proper signage.

As early as 1899, a precursor group to the American Automobile Association (AAA) was formed with the objective of placing signs on busy roads to assist travelers in finding their way. In 1905, the Buffalo Automobile Club in New York State established a network of signs, and shortly after, the Automotive Club of California began placing signs on major highways in the San Francisco area. In some cases, colored bands were wrapped around utility poles to serve as makeshift signs.

Despite the establishment of individual countries' laws on automobile traffic, one of the most crucial issues was the recognition of vehicle passports and driver's licenses when crossing national borders. It was the Swiss jurist Meili who became the proponent of such traffic regulations, advocating for the creation of a uniform automobile regulation to be applied throughout the European continent.

To address these concerns, national automobile clubs joined together on June 20, 1904, to form The Association Internationale des Automobile Clubs Reconnus (AIACR), known in English as the International Association of Recognized Automobile Clubs. During its early congresses, the association focused on resolving issues related to road traffic and the standardization of road signs. France played a leading role in these matters, as the association was founded in Paris. However, the first president of the association, serving from 1904 to 1931, was Baron Étienne van Zuylen van Nyevelt from Belgium.

Thus, during the 1st International Congress of Tourism in 1905, a call was made for the harmonization of national regulations regarding road traffic. AIACR took decisive action in June 1907 during the Homburg meeting, coinciding with the Prix de l'Empereur race, which had Prince Heinrich, brother of the emperor, as its honorary president.

It is logical that when the brother of the emperor is involved, governments also participate in solving transportation issues. In 1907, the French Minister of Transport, Bertin, agreed to chair an international conference of motorists. At this conference, it was decided to consider the issue of international agreements for regulating road traffic. The outcome of this international conference of the AIACR was a resolution adopted on June 15, 1907, which stated, among other things: *"Recognized clubs have decided to request their governments to align a series of police and legislative regulations related to motorism with international rules."* Thus, the first step was taken towards creating unified requirements in road traffic.

In 1909, an international multi-day automobile race, named after Prince Henry, was scheduled to take place. It was decided to equip the entire race route with new international signs. In relation to this, the Austrian "Allgemeine Automobil-Zeitung" reported on May 2, 1909, that the Ministry of Labor should provide the race route of Prince Henry with precautionary signs in the spirit of international messages, as it would be conducted on state roads.

The race was planned to start in Berlin, then pass through Germany, Hungary, Austria, and finish in Munich. It was this event that significantly contributed to the development and standardization of road signs at the beginning of the century.

From October 5 to 11, 1909, the first International Conference on the Issues of Automobilism was held in Paris. On October 5, representatives from 16 European countries, along with the United States of America, convened in the halls of the Ministry of Foreign Affairs at Quai d'Orsay. While some delegations had official negotiation mandates from their governments, others attended as observers. The delegations comprised senior officials from specialized ministries and representatives from various Automobile and Touring Clubs. At that time, there were no dedicated transportation ministries, and road traffic matters could fall under several ministries, such as the Ministry of Interior, Police, Public Works, etc. The discussions focused on a proposal by the French government and some suggestions from the German government. Two approaches were explored to advance the internationalization of road traffic: the creation of new international provisions on one hand, and the reciprocal recognition of national provisions on the other hand. However, harmonizing national regulations faced limitations due to significant differences.

The main principle among the conference participants was non-interference and respect for the laws and customs of sovereign states. It was on these principles that the first Convention on Road Traffic, known as the “International Agreement on Road Traffic,” was developed and adopted.

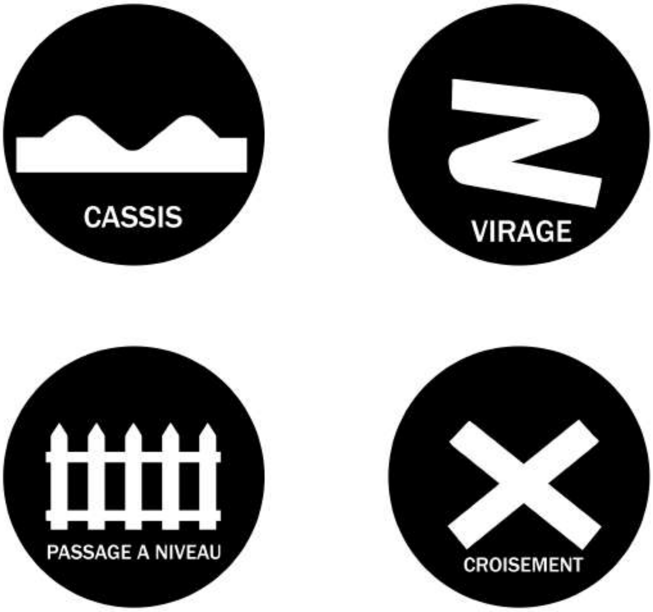


Figure 4.19: First 4 international road signs

4.3.10 Convention of 1909

The majority of European countries adhered to the “Convention with Respect to the International Circulation of Motor Vehicles” on October 11, 1909. Ten states had already ratified it when it entered into force on May 1, 1910, with others joining later.

Article 1 outlined the conditions that a motor vehicle must meet “to be internationally admitted to circulate on public roads.” Every vehicle was required to be equipped with a braking device and, if it weighed more than 350 kilograms, a reverse gear. Furthermore, “an examination before the competent authority or before an association authorized by it” was mandatory. During the examination, special attention had to be paid to ensuring the safe functioning of the vehicle, as well as to the fact that it “does not frighten saddle or draft animals by noise” and “does not seriously inconvenience pedestrians by smoke or steam.” [185].

Article 2 detailed the requirements concerning the driver of the automobile. Several countries had not set a minimum age for driving a motor vehicle, but now the age was fixed at 18 years. Furthermore, the driver had to “possess the qualities that provide sufficient guarantee for public safety,” and these qualities had to be confirmed by an authority or an authorized association. The German delegation wished for this passage to be drafted in a stricter and more concrete manner. In their opinion, *«The candidate must know how to drive perfectly. They must have a good understanding of the vehicle’s components and their functioning. Additionally, they must have the physical and intellectual abilities to calmly, safely, and attentively operate the vehicle, even in the busiest streets and under challenging circumstances. They must have perfect vision and hearing and possess the moral qualities that provide sufficient guarantee to the public.»* However, this proposal faced opposition from countries like Great Britain, Belgium, and the Netherlands, which did not have any driving examinations in place.

Those who met the requirements stated in Articles 1 and 2 could obtain an “International Certificate of the Road.” As stated in the third article, this certificate was valid for one year and recognized in all contracting states, with the holder not being required to undergo another examination. The fourth article introduced the plates indicating the country of origin of the vehicle, which are still in use today. Henceforth, any vehicle leaving its national territory had to display these plates: F for France, D for Germany, GB for Great Britain, and so on. Switzerland was assigned the letters CH (Confoederatio Helvetica) because the letter S (for Sweden) had already been allocated.

The fifth article of the convention addressed warning devices (lanterns, horns), while the sixth article covered specific provisions for motorcycles. The eighth article paved the way for the internationalization of road signs. An agreement was reached on four obstacle signals, incorporating one of the proposals from the previous year's Road Congress. However, the harmonization of regulations outlined in the convention did not proceed any further. On the contrary, Article 7 required vehicle drivers to strictly adhere to local customs. In the final article of substantive content, Article 9, the principle of territoriality was once again reaffirmed in a broad sense, with each automobile driver being obligated to comply with the rules of the country they are in. In this regard, the Conference had yielded to the power of circumstances. Nevertheless, a declaration of intent was adopted, with the goal remaining the establishment of a uniform international law.

In Article 8 of the Convention, in addition to information about the rules for the placement of road signs, it stated:

«Each of the contracting States undertakes to see, within the limits of its authority, that no notices directing attention to dangerous places shall be put up on roads other than the notices of which the representation is given in an annex to this Convention ...Nevertheless, the Governments of the contracting States may agree in common to modify this system of notices. There would be advantage in adding to this system of notices a notice directing attention to a custom-house and ordering a halt, and also another notice directing attention to a toll-house or an office for the collection of town dues.»

The ratification of the agreement was supposed to take place by March 1, 1909. Twenty countries signed the Convention on Motor Vehicle Traffic. The Convention was ratified by the United Kingdom and several of its dominions, as well as Prussia, Austria, Hungary, France, Spain, Italy, and several other European countries. Russia submitted its ratification documents on March 5. Some countries, such as Greece, the Netherlands, Portugal, and Romania, declined to ratify the Convention. Finally, the first actual road signage appeared in 1912, along the route from Paris to the seaside resort of Trouville-sur-Mer.

The need for international road signs primarily arose to warn drivers about potential dangers they might encounter along their travel route.

The four obstacle signals declared binding, first by the Road Congress and then by the International Convention, were initially proposed by the Touring Club de France (TCF).

These signs included: guarded level crossing, intersection of roads, transverse road irregularity, and dangerous turn.

Let's explore in more detail why these specific warning road signs were the first to emerge.

The road sign initially referred to as "winding road" at the beginning of the century was actually the "Dangerous Turn" sign. Like many other signs, it resulted from shortcomings in road construction. The Z-shaped design of the sign originated from a bent arrow. It became essential for indicating high-speed travel, sharp turns, and poor visibility. Failing to adjust speed could lead to veering off the road or colliding with oncoming vehicles or carriages. This sign was particularly necessary on roads passing through forests, mountains, and hilly terrains. The zigzag line on the sign warned drivers to "be cautious, reduce speed, as the road ahead sharply changes its direction." In addition to the risk of not fitting within the turning radius on such sections, there was a high chance of colliding with oncoming traffic. To eliminate the risk of vehicles being forced off the roadway or into the oncoming traffic lane on turns, switchbacks were introduced. Borrowed from railways, switchbacks were first used in Italy in the early 20th century.

The "Uneven Road" sign, or as it was initially called, the "Transverse Road Unevenness" sign, arose from deficiencies in road construction and maintenance. After all, roads in the early days of motorization were much worse than they are now. Roads often crossed small streams and even entire rivers without bridges or culverts, resulting in deep ruts and potholes on the surface. Without warnings about these road irregularities, drivers could experience serious accidents or damage their vehicles.

The "Level Crossing" sign, shaped like a grid, indicated a railway level crossing. Trains, due to their much larger mass compared to their braking capability, have a significantly longer braking distance than road vehicles. With few exceptions, trains do not stop at level crossings and rely on vehicles and pedestrians to clear the tracks in advance. So, why was a sign needed in addition to the barrier or gate? The reason is that inattentive drivers, without reducing their speed, often broke through the barrier and drove onto the railway tracks. Furthermore, the absence of illumination at level crossings at night posed an additional danger.

Indeed, in Europe’s multicultural context, the need for international communication led to the search for road signs in the form of icons and symbols. The four adopted obstacle signals in 1909 were precisely such pictograms, represented as highly simplified icons.

Producing signs and placing them along the road became one of the tasks that associations or companies like the tyre manufacturer Michelin assumed. Typically, clubs put their names on the signs they sponsored, providing them with a way to advertise themselves. Along with signs highlighting dangers, most indicated directions or places of interest.

The Kaiserliche Automobil Club installed hundreds of these signs every year. In 1913, $1,383$ signs were placed mainly in front of level crossings and curves that were impossible to see in advance, slightly less frequently in front of intersections and speed bumps [1]. As a rule, the signs prominently displayed the name of the company that had donated them; in some cases, they featured the club’s colors. In turn, companies in the automotive industry quickly realized that it was a way to mark their presence and demonstrate that they were close to their customers. By 1925, the Allgemeiner Deutscher Automobilclub had installed a total of $36,000$ signs on German roads.

During 1911—1914, the tire manufacturer Michelin placed 30,000 enamel signs on French roads, in addition to the $23,000$ signs already provided by the Touring-Club de France by 1908. These signs indicated the beginning and end of built-up areas and urged motorists to be cautious, among other things, for children [71].

As a rule, a local club was assigned a specific stretch of road or area where they were responsible for installing and maintaining road signs. Automobile clubs promoted the automobile and defended it against attacks from conservative forces and often rival factions.

Unlike the European signaling system, written signage managed to persist for a much longer time in the United States of America, as English was the “language of the road” that everyone understood.

Together with the previously used circular and rectangular signs, triangular signs began to be used as well, although at that time the shape of the sign did not yet carry any specific meaning or convey a particular message. The triangular shape of the sign was officially introduced at the 2nd International Conference in 1914. The triangular shape indicated that it and the information on it were intended for drivers of self-propelled vehicles. Public organizations still handled the production of road signs. By the beginning of World War I, the number of signs had increased to 9, and they could easily be divided into 3 groups: direction signs, speed limit signs, and warning signs. By 1915, over 1,000 signs had been installed in Switzerland.

Indeed, the need for uniformity in road traffic signs became apparent as motorists started traveling across borders. If each country had radically different signs that were unintelligible to drivers from other nations, it would create confusion and pose a risk to road safety. Therefore, there was a growing realization of the importance of standardizing road traffic signs even before World War I. After the war, discussions regarding standardization shifted to a new institutional setting that had emerged in the aftermath of the conflict. This marked a continuation of the efforts to establish international agreements and frameworks for road traffic signs and further contributed to the development of standardized signage systems.

While there was early international harmonization for road traffic signs, leading to the development of a global system of road signage in the second half of the 20th century, the harmonization of national traffic regulations remained rudimentary [185]. The 1909 Convention confirmed the principle of territoriality, stating that states had the option to inform incoming motorists about local regulations. For instance, at the Swiss border, foreign motorists were provided with a booklet containing the main traffic rules of all cantons and a map of roads prohibited for car traffic. However, intercultural misunderstandings were inevitable. In July 1910, along Lake Lucerne, an American tourist was fined for speeding twice within a few kilometers. He refused to pay one of the fines, arguing, among other things, that he had not been able to read the corresponding signs. Even diplomats became involved in the case, with the American embassy requesting a detailed explanation from the Swiss Federal Council regarding the "rude behavior" of the Swiss police officers. Such problems could arise, especially when a foreign motorist was involved in an accident, as the local law on liability would apply to the driver, which still holds true today. The first international attempts to harmonize mandatory insurance and civil liability laws date back to the interwar period [185].

The convention was also silent on one point: the desired reduction of customs formalities had not been achieved because national regulations were too divergent. While in most states, automobiles had to go through customs based on their weight at the border crossing, Britain and Denmark already had very liberal regulations at the beginning of the century (with a declaration stating that after the stay, the driver would leave the country again in the vehicle, which would then be considered a mere piece of luggage). In most other European countries, the motorist would receive a pass after paying the customs duty, and then return it when leaving the country for good, after which the amount paid would be refunded. To prevent abuse, cars were sealed at the border. The pass was valid for several months, and in most cases, during this period, it was possible to cross the border as often as desired.

Unfortunately, the implementation of the first international road traffic agreement was not achieved. The cannons of World War I shattered the peaceful silence of many countries, dividing them into two irreconcilable camps. There was no longer any focus on road signs, requirements for vehicle equipment, or driver's licenses.

In many European countries, the internationally adopted signs were being installed, but they were also accompanied by their own "local" or "national" signs.

The international road traffic sign system experienced significant development during the interwar period, with the League of Nations playing a coordinating role. At the same time, there was a slight decline in the influence of associations, as it was now the public authorities who took greater responsibility, including legislative aspects, for the content and implementation of road signs. During the interwar period, the challenges of urban traffic demanded the development of a comprehensive signaling system, and city associations made numerous proposals in this regard. The first pedestrian crosswalks and traffic signals also emerged during this time. To navigate the urban traffic jungle, one had to understand the semiotics of the road. Motorists acquired the necessary knowledge through driving schools, while children learned about traffic rules through road safety lessons.

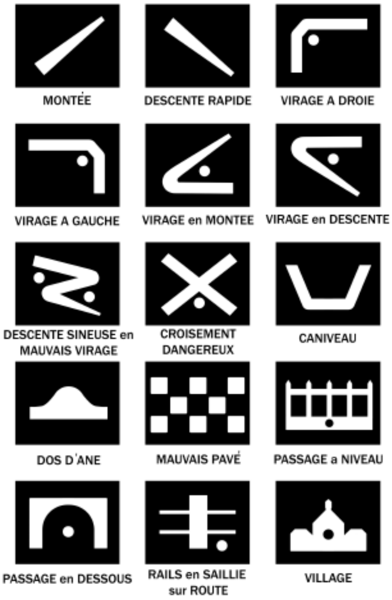


Figure 4.20: Road signs in France

4.3.11 League of Nations and road sign wars

After World War I, there was a significant shift in international cooperation with the establishment of the League of Nations. Proposed by American President Woodrow Wilson, the League was created after lengthy negotiations in Paris. It served as an institutional framework for international discussions on various practical matters. The League's founders believed that advancements in technology were rendering national borders less significant, and they saw the need for a new, comprehensive organization to oversee these developments, including ongoing efforts to achieve harmonization of road signs among nations.

In the terminology of the League of Nations, the term "communications" encompassed not only telecommunications but also transportation and electricity. "Transit" referred to the passage of goods, services, or people through states that were not their origin or destination. The Committee for Communications and Transit was established following the First General Conference on Communications and Transit, one of the initial conferences organized by the League. The conference took place in Barcelona between March 10 and April 20, 1921, and brought together representatives from forty-four states.

Although the Committee on Road Traffic was not initially included among the original committees, the subject soon gained prominence and was added to the agenda. The Committee on Communications and Transit (CCT) started discussions on driving licenses in March 1922, recognizing the close connection between licenses and road safety. Deputy Director Butler of the International Labour Organisation, representing the International Transport Workers Federation, proposed the idea of an international driving license for chauffeurs to the League of Nations. The CCT expanded this request and recommended the organization of a conference to comprehensively revise the 1909 Convention on Motor Traffic. To facilitate this revision, the CCT established a Committee of Enquiry for Road Traffic, responsible for drafting the revised version, including provisions on road traffic signs.

During the discussions within the Committee, members deliberated on the advantages and disadvantages of the existing sign system established in the 1909 Convention, as well as the proposal to adopt a single sign for all dangerous traffic situations, already implemented in Scandinavia. The Swedish government had informed the other signatories of the 1909 Convention about its plan to introduce a red-edged triangle as a universal danger sign, a decision supported by the automobile clubs of Denmark, Finland, and Norway.

The Committee decided to retain the four original signs and introduce a fifth sign depicting a locomotive to indicate unguarded level crossings. This addition was necessary due to the increasing number of level crossings without gates across Europe. The 1909 Convention had already allowed for the possibility of modifying or adding to the existing signs. Article 8 stated, “Governments of the Contracting States may agree in common to modify this system of notices.” In addition to the four signs, the Committee suggested including a sign for indicating a customs house and a sign ordering a halt as beneficial additions to the existing system.

Contrary to the Scandinavian representatives who advocated for the adoption of the unique red triangle, professional chauffeurs expressed satisfaction with the existing signs. A. Förstner, the representative of the International Federation of Professional Motor-car Drivers, stated that the signs adopted in the 1909 Convention were excellent and supported their widespread use. To accommodate these differing opinions, Pflug proposed adding the Scandinavian triangle to the list of international signs, to be used for dangers not covered by the other signs.

Figure 4.21: The signs presented include three variations of the distinctive red triangle used in the Scandinavian countries. These variations indicate a 20 km speed limit, a 4.5 ton weight limit, and the name of a town

It was understood that states were not obligated to introduce all the signs, nor were they prohibited from introducing additional ones if they deemed them necessary. However, the Road Committee emphasized that signs with the same meaning as the standard signs should not use alternative symbols to convey their message. Bilfeldt, representing the Nordic countries, wholeheartedly supported the proposal, and the amendment was adopted [253].

Despite the previous decision to include the Scandinavian red-edged triangle as an additional sign for conveying specific dangers, the Scandinavian members persisted in their efforts to promote its use as a unique danger sign. In October 1925, during a meeting of the Conseil Central de Tourisme International, which served as a cooperative body between automobile and touring clubs, the Scandinavian members once again proposed the adoption of the red-edged triangle.

However, their proposal faced resistance from representatives of other clubs who were not willing to accept it. In a demonstrative act, the Scandinavian members walked out of the meeting, expressing their dissatisfaction with the refusal to embrace the red-edged triangle as a universally recognized danger sign.

Indeed, the topic of traffic signs was finally addressed at a conference held in Paris in 1926, focusing on the revision of the 1909 Convention. As a result of the conference, the adopted convention was divided into two separate documents. One document covered administrative regulations, while the other document dealt specifically with the rules of the road, including road danger signs, which increased from four to six.

Figure 4.22: Danger sign

In 1926, the four danger signs that had been created and used since 1909 were definitively changed from the shape of a disk to a triangle, a form that is still in use today. An additional sign depicting an unguarded level crossing was added. Then, in 1928, at the request of Switzerland, one-way, direction, parking, and no parking signs were adopted.

During the following five-year period, the League of Nations undertook the task of international road sign standardization as part of its broader objective to establish itself as the central platform for discussing all international aspects of road traffic, particularly within a European context.

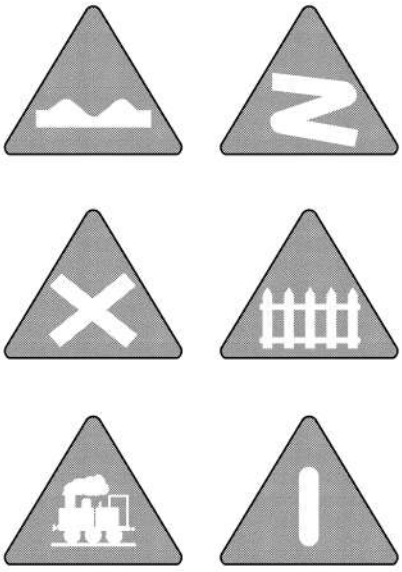


Figure 4.23: Layout of traffic signs defined in the Czech Republic in 1935

In 1931, an international convention on the standardization of road traffic signs was signed. However, since 1935, traffic signage in the United States has been uniform but fundamentally different from European practices. This divergence in traffic signage between the United States and Europe has persisted and continues to this day.

4.3.12 Road sign epidemic

Let’s focus on several important issues regarding the development of traffic regulation.

After World War I, there was a revival of automobile manufacturing in all countries. New roads were constructed, and new brands and models of cars emerged, leading to changes in their appearance and design. For instance, initially, the driver would sit at a considerable height from the ground, similar to a coachman on a carriage, to have a better view of the road. However, over time, the driver’s seat was placed lower. Consequently, road signs and indicators also had to be positioned lower.

Starting from the 1920s, due to the increasing number of accidents on the roads, private organizations began to emerge in different countries. The members of these organizations were responsible for providing medical assistance to those involved in accidents.

In the United States, the American Automobile Association (AAA) introduced the first roadside assistance service in April 1915. It started with a group called the “First Aid Corps,” consisting of five motorcyclists from the Automobile Club of St. Louis. They patrolled the city streets on Sundays, offering free minor engine and tire repairs to stranded motorists, regardless of whether they were members or non-members.

In the United Kingdom, both the RAC (Royal Automobile Club) and The Automobile Association (AA) have been providing repair services to their members. They offer on-the-spot repairs, towing services to local garages or the driver’s home if nearby (often within a limit of 20 miles), and in some cases, they also provide onward journey services such as offering rental vehicles. The AA was formed in 1905, while the RAC has had a longstanding association with the Royal Automobile Club.

Similarly, in Germany, the Allgemeiner Deutscher Automobil-Club (ADAC) started offering a similar roadside assistance service in 1927.

In response to the significant rise in road accidents during the blackouts, the London “Safety First” Council was established on December 1, 1916. In 1917, the council began collecting accident data and advocated for the licensing of all drivers. As part of their efforts, a campaign was launched to change the pedestrian rule so that pedestrians would face oncoming traffic. This campaign proved to be highly successful, resulting in a 70 percent reduction in fatal accidents caused by pedestrians stepping into the path of vehicles within a span of 12 months.

In Europe, in the mid-1920s, a true "epidemic of road signs" began, as it was called at the time. There was a debate in the press about the effectiveness of using a large number of signs on the roads. It was lamented that road signs were overcrowded with advertisements. Signs with only text started to appear, such as "Avoid dust." Complete sentences began to be written on signs. Of course, jokes were not uncommon either. Signs with "STOP" were installed, and when a driver approached the sign and stopped, they could read in smaller letters below: "...during lunch and dinner time."

The excessive amount of signs was, for example, discussed in 1923 by Karl Lehmann [162]:

«Most highway sign posting is expensive and poor advertising. ...Recently in Orange and Lake Counties, Florida, more than $5,000$ *unsightly signs have been removed from the highways by the road crew and other good citizens. The highways belong to us all and each one of us has as good a right to take down unsightly signs on the highway as some one else has to put up such signs. If a few hundred more good citizens would stop and take down these signs along the highways we would soon see great improvement in all parts of our state.»*

Even in 1932, Alvah Lauer wrote: *«It is hardly necessary to mention the hazards introduced by pseudo-warning signs put up by roadside lunch houses, oil stations, etc. The general psychological effect of adaptation increases the danger. In common parlance this means the same as that describing the ultimate condition of the shepherd boy's parents when he kept called "Wolf, wolf." They refused to become alarmed. This principle is known in psychology as adaptation. Not only is it necessary to have the signs real but there is an optimal point in the amount of bona fide markings. Too many signs or markers, like too many laws, are not always taken seriously. The optimal point of marking is a matter to be further investigated.»* [160]

In the 1920s, significant progress was made in the field of road traffic regulation.

This period was even referred to as the "influence of anonymous folk creativity on road signs." However, in the end, common sense prevailed, expressed through the idea of installing road signs that could be understood by everyone through graphical representations.

In France, a system of road signs was established. The small cast iron plates with illegible small fonts that were previously used were replaced with larger signs made of wood or enameled steel. In addition to the text, they started indicating the category of the road, for example, the letter "N" indicated a road of national significance, while "D" indicated a departmental road.

In Italy, a wide campaign for the installation of road signs and indicators began in 1919, immediately after the end of World War I. Well-known automobile companies such as FIAT and Pirelli were involved in financing this initiative. By 1922, Italy had already installed 4, 308 road signs and indicators throughout the country.

In Germany in the 1920s, automatic traffic lights, road signs, and road markings began to appear. In the existing regulations of the country, the signs were categorized as follows:

1. Warning signs: These were round signs with a blue background and white symbol. According to the 1926 Convention, they were later changed to a triangular shape.

2. Prohibitory signs adopted by several states of the Weimar Republic. They were rectangular in shape and featured black symbols on a yellow background.

3. Speed limit signs were also adopted by several states of the Weimar Republic. They had a rectangular shape, with white text on a blue background.

4. City signs. This group of signs was used to regulate traffic in cities and urban areas. The signs had an arrow shape with a circle in the middle, indicating a specific instruction or requirement.

There were debates about how to depict signs and whether it was better to use symbols or words on the signs. At that time, there was no clear consensus on this matter. In Germany, experiments were conducted, and it was believed that signs with a red border and black symbols on a white background were most easily perceived by drivers. These signs were equally visible during the day and at night.

The process of creating various signs varied. In 1925, the French newspaper "Revue Automobile" announced a competition to design a sign for unguarded level crossings.

A design proposal featuring an image of a locomotive was soon submitted. This sign, depicting a locomotive, has remained unchanged and has been preserved to this day.

New signs would sometimes first appear in the regulations of individual countries and then be considered by the Convention. For instance, the "Pedestrian Crossing" sign was adopted in the Traffic Regulations of Paris, approved on February 16, 1925.

The red circle (modern "No Entry" sign) to prohibit the passage of vehicles on a street was first used in France as well. If traffic was only permitted on a certain section of the street, a portion of the circle would be painted white.

The prescribed direction of movement was indicated by a white arrow on a blue background, and parking spaces were marked by a white letter "P" on a circular blue background.

In Italy and France, starting from 1926, danger warning signs had a white symbol on a dark background. This color combination was believed to provide better readability, and the chosen paint colors were more durable.

In 1927, the Swiss City Union published its "Proposals for the Unified Regulation of Traffic in Swiss Cities." While the 1926 Convention mentioned that the triangular shape of a sign only applied to automobile drivers, the proposals recommended that signs regulating street traffic should be exclusively circular in shape. It was also recommended to use black symbols on a white background for these signs.

Road signs in the United States, like their European counterparts, were divided into warning signs, regulatory signs, and guide signs. The group of warning signs had a diamond shape, hanging from a point. Regulatory signs were rectangular in shape and positioned vertically.

Indeed, one of the main differences between the American sign system and the European system from the beginning was the extensive use of text on signs in the United States. Many signs in the American system included written messages, providing specific information or instructions to drivers. Interestingly, the use of text on signs varied somewhat freely between different states, despite the standardized sign shapes and designs.

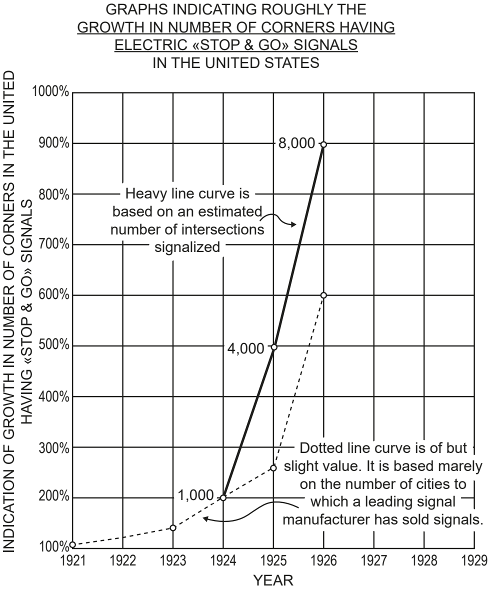


Figure 4.24: Graphs indicating roughly the growth in number of corners having electric «stop & go» signals in the United States

4.3.13 Red Light, Green Light

Alongside road signs and markings, traffic signal regulation began to develop during this period.

The real commercial development and sale of electric traffic signals in the USA began modestly around 1921. The accompanying graph (Figure ??) indicates that the development was very slow until 1923 or 1924. Since that date, as the chart shows, there has been a rapid development in this new field. But, as Marsh stated in 1927, *«Development costs and costs of making the very rapid improvements in signals and in control equipment, have been and continue to be so great that it is certainly safe to say that the industry has not been a highly profitable one.»* [180]

However, even in 1927, there were claims that the design of the traffic signal units themselves had progressed considerably towards standardization. *«The use of green for "go" and red for "stop" has practically become standardized in this country.»* [180]

However, the amber light was used for too many different purposes, leading to frequent misunderstandings.

«The amber has been used for at least the following purposes: Clearing the intersection, warning of change in signal color, providing an exclusive pedestrian period, and as a special signal to make left turns. It has also been used in combination with the red and green lights, with various significances.

It is quite natural, therefore, that the public does not know what to do on the amber light; and, hence, the amber light has not to date "worked out" satisfactorily. Numerous traffic officials have experimented with different meanings of the amber, and hence, could not educate the public in one definite meaning for this light.» [180]

Many have wondered why specific colors were chosen to regulate traffic, particularly red and green as the primary signals in traffic lights.

Modern signaling systems across all transportation modes utilize light signals. In railway transportation, five colors are used: red, yellow, green, blue, and white.

For automotive transportation light signaling, three colors are employed: red, yellow, and green. The color red is universally recognized as a signal for danger or warning. Conversely, green signifies safety and calmness, making it the natural choice for the go-ahead signal in traffic lights. While it may appear that these colors were arbitrarily selected, the choice, especially for red, was not random.

Red symbolizes danger across many cultures. Except for stop signs (not stop lights), red has signified stop well before the invention of cars. In train signals, red's use dates back to when mechanical arms were lifted and lowered to show whether the track ahead was clear.

The red color has the longest wavelength in the visible spectrum, making it visible from greater distances than other colors. This visibility is crucial in poor conditions, such as fog, which absorbs blue and green light, making green appear yellow and yellow seem red. Misinterpreting yellow as red and green as yellow in fog poses no traffic danger.

Initially, green indicated caution in early railway lights, with a clear or white light for the "all-clear" signal. However, tragic accidents occurred when engineers confused stars with the all-clear signal, leading to green being reassigned to mean "go." Railways then adopted a green and red light system for train movement signals, prioritizing safety.

Until the mid-1900s, stop signs were not exclusively red; many, alongside yield signs, were yellow due to visibility issues at night. The trend of yellow stop signs began in Detroit in 1915. Interestingly, Detroit installed its first electric traffic signal five years later, pioneering the use of an amber light at Michigan and Woodward Avenues.

Advancements in materials and technology enabled the production of highly reflective signs, allowing red to reclaim its place as the dominant color for stop signs. Yellow, with the second-highest visible wavelength after red, became associated with caution, as seen in school zones, buses, crosswalks, and other cautionary areas.

In the early days of U.S. traffic lights, the green signal was at the top. However, it was later determined that the red signal, being more critical for safety, should be more prominent. Recently, lenses for red signals have been made larger to emphasize their importance.

4.3.14 Urban or non-urban?

From April 20 to 24, 1926, a conference in Paris focused on coordinating and approving the new 1926 International Convention relating to Road Traffic.

The Convention stated that each state should regulate its internal traffic and establish a common model for marking dangerous locations.

It approved six road signs, standardizing their shape and dimensions, though the background color remained arbitrary due to divided opinions. Countries with consistent winter snow cover argued that a white background would reduce visibility for drivers.

The Committee highlighted the importance of traffic signs and continued their development post-Convention. A significant advancement was the explicit inclusion of an urban dimension, differentiating between urban and non-urban signs. This focus was evident with Dr. De Schulthess, representing the International Union of Towns, attending the Committee's fifth session. A Swiss committee, including local authorities and various interest associations, had created a set of urban traffic signs. These signs were to be presented at an IUT meeting in July.

The involvement of the International Union of Towns and various associations underscored the Road Committee's role as a primary platform for international road traffic discussions. This role extended beyond traffic signs to include commercial road traffic and tax exemptions for vacationing vehicles.

The Road Committee faced the challenge of balancing two perspectives regarding urban traffic signs. On one hand, it acknowledged the significance of adapting traffic regulations to local conditions, a task best managed by local authorities. However, this approach encountered difficulties due to the varying positions of local authorities within the state hierarchy of the League of Nations member states. Achieving complete uniformity in urban traffic signs became challenging because of the differing levels of centralization across countries. For instance, in Belgium, urban traffic regulations fell entirely under the jurisdiction of local authorities, with central authorities having no legal basis to intervene in this local competence.

However, the Committee also recognized the need to mitigate excessive confusion for road users caused by local variations. They understood the importance of ensuring that foreign travelers could easily understand and interpret road signs while traveling abroad. This led to the push for sign unification to facilitate quick and clear interpretation, with a particular focus on tourism in regions where cross-border tourist traffic was significant. This European context was crucial to the debate, as Europe had a large number of car owners and a high density of national borders. European countries could not make decisions about road signs in isolation due to the interconnected nature of traffic flows and the challenges posed by divergent traffic regulations across borders. Countries like New Zealand, an island state, could argue that they saw no need to align their existing signs with those agreed upon within the League of Nations framework. Similarly, India noted that its self-contained road system, at its current stage of development, would be less affected by the disadvantages of non-unification compared to Europe.

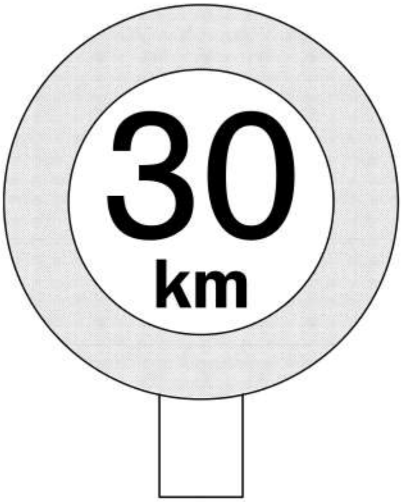
Finding a balance between universal standards and local flexibility was a challenging task. The discussion on speed limits exemplified this delicate balance. The issue of speed limits highlighted various conflicts stemming from the increasing dominance of automobiles on the roads. All member states of the League of Nations had established lower speed limits within urban areas compared to rural areas. The impact of speed on accidents became a growing public concern, leading to awareness campaigns.

In the United States, public outcry against the dangers of high automobile speeds evolved into an automobile industry-led initiative to condemn reckless driving instead. This shift in perspective led to a new understanding of the relationship between automobile speed, accidents, and individual responsibility, which gradually took shape during the 1930s.

Dr. S. Stimmer, representing the Association for the Promotion of Traffic Safety, expressed his strong belief in the necessity of implementing lower maximum speed limits as a critical step towards unifying European traffic rules. In an emotionally charged letter, he argued that the absence of a general limit on excessive speeds was the root cause of the alarmingly frequent and devastating accidents.

However, Committee member Crespi held a different perspective, and he expressed his concerns to Committee secretary Romein. According to Crespi, the automobile clubs disagreed with such measures, as they believed they hindered human progress. Crespi directly contradicted the notion that motorized traffic caused more fatalities compared to non-motorized traffic. He further contended that advancements in braking systems had significantly reduced the need for speed limits.

After careful consideration, the Committee reached a unanimous decision that local authorities should have complete freedom to establish speed limits based on their specific local conditions. While some members believed that implementing a general speed limit at the national level could be beneficial, others argued that speed was just one aspect contributing to accidents and should not be the sole focus. This debate reflected the broader challenge of harmonizing road traffic rules, as witnessed during the 1926 conference in Paris. However, the Committee did not encounter any difficulties in agreeing upon the signs that indicated speed limits. In fact, the shape of the proposed signs underwent a change from rectangular to round, a decision that was confirmed at the 1931 conference in Geneva.



The harmonization of shapes and colors was seen as a solution to the concern that an excessive number of signs could potentially distract drivers and create new hazards on the roads. Instead of imposing a limit on the number of symbols used in road signs, standardizing shapes and colors offered an alternative approach. There was widespread support for increasing the number of internationally standardized signs, highlighting the need to develop a typology of signs.

In terms of colors, there were variations between dark signs on a light background and light signs on a dark background, as outlined in Table 1. Regarding shapes, the triangle was designated for indicating dangerous traffic situations in accordance with the provisions of the 1926 Convention on Motor Traffic. Initially, the Committee recommended rectangular plates for displaying speed limits, while all other signs were to be circular with a diameter of sixty centimeters.

Table 4.2: Road sign colours, selected European states, 1927[253]

Country	Colour	Rationable
France, Italy	white on a dark blue background	visibility, durability
Germany	black on a white background with red edges	most conspicuous
Switzerland	black on a white background	–

By the time of the European Conference on Road Transport in 1931, the discussions regarding the standardization of traffic signs, primarily in Europe, had reached their peak. The outcome of the conference was the establishment of a convention that divided signs into three distinct categories, each with its own exclusive shape. Triangular signs were designated for indicating dangerous situations, circular signs denoted obligations, and rectangular signs provided indications or information.

In terms of color usage, the convention emphasized that the color red should predominantly be used in prohibition signs. While traffic signs did not achieve complete uniformity across Europe, a brief overview of the available alternatives demonstrates the extent to which the conference successfully mitigated the potential chaos in traffic signs.

The headquarters of the League of Nations in Geneva served as the venue for the conference, where traffic signs were a central topic of discussion. Following the 1926 Conference in Paris, the Committee had made significant efforts to establish a basis for consensus on traffic signs. Now, the moment had arrived to confront the reality. The conference chair acknowledged that while most state representatives naturally believed their own system to be optimal, it was necessary to question such positions and be open to concessions.

The final report explicitly stated that the outcomes of the conference applied to both main roads and urban traffic, challenging the previous notion put forth in the 1909 Convention that traffic signs were generally unnecessary in towns or villages. This conviction was reiterated during the conference itself, as it was believed that motorists would exercise caution in crowded areas regardless of the presence of traffic signs.

During the conference, the drafts that had been prepared by the Committee remained largely unchanged, signifying the success of their years-long work. However, a notable contrast emerged in the discussions surrounding commercial road transport, which was considered a significant theme on the conference agenda but was ultimately abandoned altogether. Ultimately, the individual states retained a considerable degree of maneuvering space with regard to traffic signs. The situation ran counter to the position of the IUT as articulated by De Schulthess, who proclaimed his union ready to make the Road Committee the supreme arbiter in road signaling and related regulations.

The Committee's viewpoint, expressed prior to the landmark conference, was indeed aligned with its own opinion that international agreement was the most effective method to prevent the proliferation of localized traffic solutions. Recognizing the importance of swift action, the Committee emphasized that further delay would only increase the challenges of resolving existing divergences. However, the Committee members were cautious, as past negotiations had made them well aware of the unpopularity of centralizing authority.

During the conference, the recommendation was made that no new signs should be introduced without prior consultation with the League's Road Committee. While De Schulthess and the IUT were concerned about replacing local variations in traffic signs with national variations, the Road Committee favored soft harmonization methods rather than imposing a top-down approach. Nonetheless, through this approach, the Committee managed to establish clear limits on the uncontrolled proliferation of traffic signs.

During the Interwar years, the efforts to standardize road signs internationally culminated in the 1931 conference. The system that was established during this conference remained largely unchanged thereafter. While some modifications were implemented, they did not involve a complete overhaul of the system. The Committee responsible for road signs aimed to preserve its legacy by monitoring compliance with the 1931 conventions and encouraging more countries to join. In essence, the conference solidified the existing framework for discussing and adopting internationally agreed-upon road signs, ensuring its continued relevance in the field.

After the 1931 conference, urban themes continued to be a prominent focus dur-

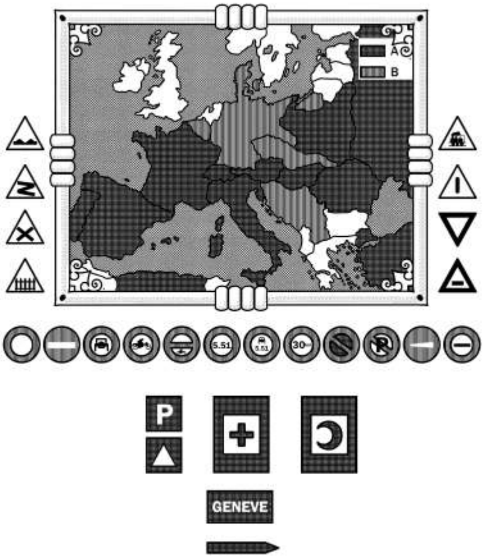


Figure 4.25: CONVENTION ON THE UNIFICATION OF ROAD SIGNALS. (Geneva, 30. III. 31.)
A.—Countries which have ratified or acceded to the Convention, **B**.—Other countries utilising the system of road signals laid down in the Convention, **C**.—Danger signals. **D**.—Signals giving definite instructions, **E**.—Signals merely providing information

ing the Committee's sessions. The distinction between urban and non-urban traffic became more pronounced. A growing concern emerged regarding the safety of pedestrians in urban areas. The majority of the discussions revolved around finding ways to protect pedestrians in traffic and promoting their proper behavior by staying on sidewalks or at the edges of the road, minimizing disruptions to motorized vehicles. The importance of regulating traffic discipline in urban settings became increasingly recognized.

One specific issue that gained attention was the use of horns. A new sign prohibiting the use of horns was introduced, reflecting a broader societal sentiment against unnecessary noise. The debate on noise reduction primarily centered on traffic noise, with the automobile horn being seen as a symbol of uncivilized behavior and a target of criticism in the anti-noise movement. Efforts were made to discourage the use of horns for anything other than accident prevention, aligning with the broader goal of creating a more peaceful and orderly urban environment.

The Road Committee reached a significant milestone in urban traffic sign standardization by agreeing upon a sign that symbolized the urban/non-urban divide. This sign, proposed by the Belgian government, featured prominent white letters on a blue rectangular background. It not only displayed the name of the town but also marked the boundary between urban and non-urban areas, indicating the different rules and regulations that applied within and outside of urban limits.

Initially, the standardization efforts of the Road Committee focused primarily on signs for main roads outside urban areas. However, by the mid-1930s, the Committee expanded its scope to include urban traffic signs as well. This development was accepted and supported by representatives of municipal authorities, as well as automobile and touring clubs from various countries. The League of Nations had established itself as a platform for international coordination and harmonization of road traffic signs, demonstrating its role in shaping European governance in this domain.

If we consider the contradictions that arise within a single country, the United States serves as a prime example. After World War I, it was faced with significant disparities in legislative road acts among its states. The first conference on roads and signs in the country was held in 1924. Building on the previously adopted 1921 "Federal Highway Act," a series of laws and regulations concerning road traffic was developed. It affirmed the sovereign right of individual states to regulate and control traffic. Recommendations included the issuance of driver's licenses, setting the minimum speed limit at 15 miles per hour and the maximum speed limit at 35 miles per hour, and standardizing road signs and signals. In 1930, a national conference published a list of street signs in the United States. Due to the unity of language, creating a system of road signs did not require the complex negotiation procedures that the League of Nations faced.

Unfortunately, many agreements reached in Europe were not immediately implemented. The looming difficulties leading up to the Second World War, and then the war itself, slowed the standardization of road signs, customs and border agreements, and, of course, the condition of the roads.

The subsequent development of traffic rules generally followed more international principles. In Europe, this was facilitated by the formation of the European Union. The increase in speed and improvement of international communication, later augmented by the Internet, along with globalization, also played a significant role.

4.4 Road network within the country

In this chapter, we will examine the development of the road network within a country. Since it is impossible to cover all countries within a single book, we will focus on just one case study: the United States.

We will explore the most prominent and active period from 1916 to 1941.

The year 1916 is chosen because it marks the beginning of comprehensive federal funding for road construction with the passage of the Federal Aid Road Act of 1916. This act signifies the full-scale federal financing of road infrastructure and reflects the priorities of government policy. The year 1941 serves as the upper boundary of the study, as it marks the entry of the United States into World War II, during which road construction was largely halted and later reevaluated after the war. Therefore, 1941 is the year when highway construction in the United States, to a large extent, ceases to exist in the form it took in the early decades of the twentieth century.

It is crucial to consider that the political culture in the United States was dynamic and experienced significant fluctuations during the first half of the 20th century. These political changes justify dividing this section into subsections, each showing how road construction strategy was not only part of the political course but also a vivid demonstration of the political patterns characteristic of each period.

4.4.1 In the Era of the Highway: preconditions for the transition to federal highway regulation in the United States

Features of the management of road construction in the United States before 1916

The history of federal road system construction is closely connected to the constitutional principle of power division between the federal center and state authorities. Traditionally, road construction was managed by local authorities and landowners adjacent to the roads, falling under the purview of local and state governments. However, at the beginning of the 20th century, the situation changed as financing and planning began to be controlled by federal ministries.

Rural roads were poorly maintained, as the road maintenance tax was minimal and could be substituted by a few days of labor, where local residents, typically lacking the necessary qualifications, engaged in road repairs. By the mid-19th century, roads between states were also under local control, with insufficient funds for repair and development. Despite several federally funded or executed road projects prior to the early twentieth century, roads were often of poor quality, poorly maintained, and new roads were rarely built. The existing routes were not sufficiently interconnected, with only $150,000$ miles of the more than 2.2 million miles of public highway in 1904 having adequate coverage for the time [256]. Railroads were the primary means of transportation, but their inadequacies became apparent by the early 20th century.

Firstly, agricultural producers needed better roads for transporting goods to railway stations. Poor-quality highways led to inflated costs of agricultural goods due to the time and expenses incurred in transporting products over long distances to reach stations, hampering farming development and raising prices of agricultural commodities [8].

Secondly, the industry required good roads for transporting coal from stations to factories. While industrial development and motorization positively impacted urban roads, they also restricted industrial expansion within city limits and tied the industry to the railway network. Additionally, various industries directly involved in road construction or the automotive business, including tire manufacturers, glass producers, road construction equipment manufacturers, concrete producers, quarry owners, and representatives of the oil industry, were also stakeholders. Some of these groups had representatives in the government willing to lobby for their interests [218]. This interest was evident in railroad trips organized by The Good Roads Movement in collaboration with the Illinois Central Railroad in 1901, promoting the importance of quality road construction [8].

The advent of motorization in the late 19th century highlighted the inadequacies of the rural road system. By 1900, there were $8,000$ automobiles in the United States [8]. The difficulty of traveling on poorly maintained roads during rainy weather, which caused muddy conditions, or the excessive dust on dry roads, along with the damage heavy vehicles inflicted on fragile gravel road surfaces, were significant issues.

The Good Roads Movement, initiated by motorists and bicyclists in the 1880s, advocated for quality road construction and quickly gained nationwide prominence. However, the unity between farmers and motorists in the fight for road construction eventually dissipated due to the issue of road damage caused by automobiles [232].

The growing public concern over road construction issues led to discussions at the state government level. The first Road Convention of the State was convened in Iowa City, Iowa, in 1883, and the first national road conference took place in 1894 with representatives from eleven states in attendance. States Highway Departments began to be established, with the first one created in Massachusetts, playing a crucial role in developing a national highway system [8].

At the federal level, the Office of Road Inquiry (ORI) was established in 1893, evolving into the Bureau of Public Roads (BPR) by 1915. The BPR was the first agency dedicated to addressing road construction issues at the federal level but lacked the authority to directly manage road construction or influence state-level road projects. Its role focused on studying road construction methods and disseminating research findings, respecting the constitutional authority of the states [176]. The ORI, initially consisting of only two individuals, received minimal funding but launched campaigns to collect statistics and publish bulletins in collaboration with The Good Roads Movement. One of the most effective programs was "The object lesson road," constructed using ORI funds as a demonstration of construction methods [8].

The idea was to construct small road sections using ORI funds as a demonstration of construction methods. Only minor road segments were built to avoid accusations of federal influence on state road policies, although many governors requested more of such projects to be built in their states [8]. OPR also provided advisory services directly to the States Highway Departments and other local agencies. From 1906 to 1916, federal highway engineers visited 144 counties in 28 states [256].

In 1905, the Bureau of Chemistry's testing division merged with OPRI, and Logan Waller Page became the head of the newly formed OPR. Page's extensive experience in laboratory work was crucial as OPR intensified its study of road materials' properties during the first decade of the century. The research base of OPR was among the most significant in the world, comparable to that of the French École nationale des ponts et chaussées. Engineers from OPR and later BPR developed specialized equipment for testing road surfaces, which was also used outside the bureau. Logan Waller Page contributed to the creation of such equipment, including the Page Impact Machine [176]. The services of the Bureau were utilized by government officials, contractors, and material manufacturers, with the material testing provided by the Bureau being free of charge [256]. After researching various options, asphalt was chosen as the preferred pavement material [212]. These issues were crucial for the economy of the United States, as the lack of quality roads hindered the development of the country's internal market, impeded industrial growth, and negatively impacted the standard of living for farmers.

States Highway Departments faced challenges in improving roads due to either conservative management or a lack of engineers, despite the continuous advisory services provided by OPR and the establishment of engineering training courses. The existing road network was poorly designed in terms of both location and construction methods. As a result, a significant amount of money was wasted on road maintenance because the roads were not properly drained, poorly located, and constructed with inadequate slopes. Addressing these issues required the implementation of national construction standards, which OPR did not have the authority to enforce. Despite their efforts to promote scientific methods of road construction, mistakes continued to be reproduced in many districts [256].

OPR closely collaborated with public movements to publish a sort of manifesto on road construction issues. In 1910, The American Highway Association was established, bringing together numerous local communities. The following year, a congress of the association was held in Virginia, which issued a resolution declaring the key technical and administrative principles and requirements for road construction. The demands were addressed to state and federal authorities and included the standardization of traffic regulations, control of road construction by state authorities, and mandatory technical road maintenance. According to the resolution, experienced engineers should oversee construction, and states were expected to support construction efforts, including enacting laws allowing for the use of convict labor on road projects. While not all of these requirements were implemented, they reflected the issues of road construction and public expectations at the time [8].

The United States Post Office Department played a significant role in lobbying for the interests of rural road construction. The challenges faced by the postal service in rural areas greatly influenced the format of future federal support for road construction. However, the postal service not only encountered problems but also had the potential to be part of the solution. On July 28, 1916, the Postmaster General was authorized by Congress to conduct research to determine how the postal service could contribute to the realization of agricultural products and direct transactions between producers and consumers. The investigation was intended to explore possibilities for expanding the Parcel Post System's operations [264].

The system that was created was called "Rural Express" and involved motorized delivery of agricultural products from producers to stores. Herbert Clark Hoover, who would later become the President of the United States, was involved in the development of this service while serving as the Head of the United States Food Administration. The goals of the service, as described in an article in Engineering News-Record, included stimulating increased agricultural production, improving the standard of living for farmers, and preserving already produced goods. Hoover claimed that up to 50% of the produced goods were spoiled before reaching the market. He also asserted that "Rural Express" would reduce the need for livestock, allowing farmers to utilize the land previously used for grazing to expand their cultivation areas in the future. Hoover emphasized that this system would lead to the development of internal trade and lower prices for products by reducing transportation costs [243].

The poor condition of roads in rural areas not only affected the economic wellbeing of farmers but also left them socially isolated. Bad roads hindered communication even with their neighbors. It was only from 1893 that mail delivery to rural areas began, but only if the roads met the requirements. The condition of the roads determined whether farmers could receive newspapers and engage in correspondence [8]. Indeed, the "Rural Free Delivery of US Mail" and "Rural Express" programs had a significant impact on the shift towards federal funding for road construction.

As a significant number of farmers expressed interest in participating in the free mail delivery program, they demanded improved road conditions in their districts. In response, Congress addressed this issue by passing the Post Office Appropriation Bill in 1912, marking the first attempt at federal financing of road construction [264]. However, the majority of states declined to participate in the program, citing conflicts with state legislation or a lack of authority to implement such measures. Consequently, only a limited number of roads were constructed under this program. Despite these challenges, the program provided valuable insights into the difficulties the federal government would encounter when interacting with local authorities. A key takeaway was relatively straightforward: interactions should be established between federal authorities and states, rather than between federal authorities and counties. In 1912, there were 48 states and approximately $3,000$ counties, making communication significantly more complex with such a large number of entities. This experience would later lead to an increased influence of the State Highway Departments.

The beginning of federal funding

«At The Third American Road Congress in November 1914, *during a speech by Dorsey William Shackleford, the Chairman of the Committee on Automobile Highways of the House of Representatives, it was expressed that "Whatever doubts may have existed heretofore, it is now generally conceded that the federal government has the right to construct and maintain roads used for federal purposes. From this, it follows that it also has the right to assist the states in the construction and maintenance of roads that are partially intended for federal use [257]." The society was already prepared to acknowledge that federal funding for road construction was an acceptable measure, and the government cautiously took its first steps in that direction, as stated by Frederic Logan Paxson, "...following the movement instead of leading it"»[218].*

The comprehensive federal funding began with the Federal Aid Road Act of 1916. It did not aim to create a unified road system, as that would have been too radical. There was still a strong belief in Congress that the road network should primarily complement the railway system, focusing on roads from farms to stations. Transcontinental roads or interstate highways were seen as luxuries for wealthy tourists [257]. Dorsey William Shackleford, a member of the House of Representatives and the Road Committee, was an active supporter of this view on transcontinental roads. In his speech at The Third American Road Congress, he proposed dividing road construction supporters into two categories: supporters of "touring-roads," whose slogan, according to Shackleford, was "See America First." He criticized them by stating, *«The 'tourist road' class is made up largely of wealthy automobile owners who desire to spend a part of their leisure time in touring the country. They are backed up by the road machinery and material manufacturers, who look upon Uncle Sam as a 'good thing,' a liberal purchaser, and one who would be a valuable customer if he would only go into the construction of 'national highways.'», «The 'touring-roads' class demands that the United States shall limit its road activities to the construction and maintenance of a few 'ocean-to-ocean' and 'across-country' highways of great perfection and then leave the rest of the people to build their own roads or do without, as they may choose.»*[257] Shackleford opposed these individuals by contrasting them with the proponents of "business-roads" who recognized that automobile roads were just a part of the complex and interconnected transportation network in the United States. According to this perspective, the automobile roads should be integrated into this network. Shackleford, as it was commonly stated at the time, supported roads "from farm to market" or "from farm to station." An important characteristic of this viewpoint was that automobile roads were seen as complementary to the railroad network rather than duplicating it. Shackleford emerged as one of the key advocates for this perspective in Congress.

Indeed, supporters of an independent highway network argued for its necessity in promoting industrial development. However, as evident from the quote provided above, this led to accusations that proponents of an autonomous road network were seeking to develop their businesses at the federal government's expense.

Thus, by the time of World War I, there were two main opinions on what kind of roads to build. The first opinion was that short roads should be constructed in rural areas to support farming and small businesses, and these roads should integrate into the existing transportation system of the country. The second opinion argued that extensive highways were necessary for the benefit of big business and motorists, including the development of full-fledged autonomous road networks, possibly even transcontinental highways.

In the future, the problems with railroad transportation during the First World War highlighted the need for an integrated highway network, not only for heavy industry but for any sector of production and any category of transportation service consumers that were previously reliant on railroads. The struggle between these two viewpoints continued throughout the second decade of the 20th century.

The Federal Aid Road Act of 1916 primarily focused on funding the construction of rural roads using federal funds, thus primarily satisfying the proponents of "farm-to-market" roads. The act was authored by Senator John Hollis Bankhead, who firmly supported these positions, and the co-author of the act was the aforementioned Dorsey William Shackleford. How was the Act created? The bill was discussed in the Joint Committee on Federal Aid in the Construction of Post Roads, which was established in 1912 following the passage of the Post Office Appropriation Bill in 1912. The committee reviewed data collected by the Office of Public Roads (OPR). This data allowed them to draw conclusions about how the condition of roads affected the level of education for children in rural areas, the overall quality of life in rural areas, property values, vehicle wear and tear, and motorization rates. The findings were mostly discouraging and led the committee to conclude that federal support for road construction was necessary.

The discussion of the project primarily focused on plans for constructing rural roads. This initiative faced opposition from the American Automobile Associa-

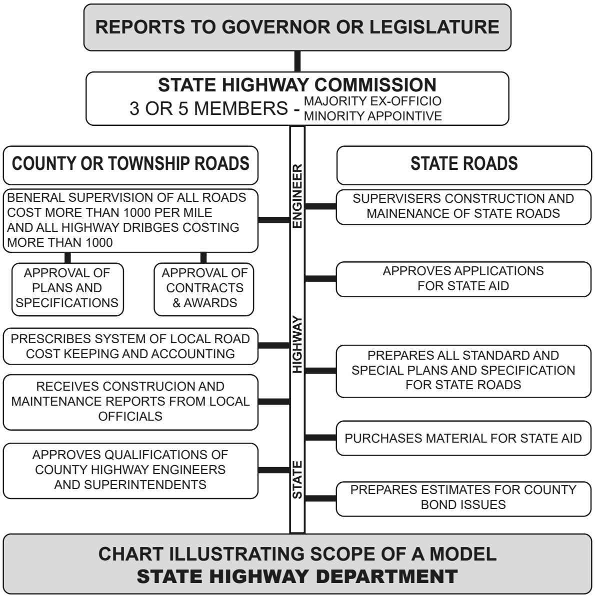


Figure 4.26: Although the date is not known, it can be assumed that Director Page made up this model of a State highway department about the time that he made his recommendation on the 1916 Federal Aid Road Act.

tion (AAA), which favored the creation of a national road network. However, with support from the American Association of State Highway and Transportation Officials, USDA Head David Houston, Postmaster General Albert Sidney Burleson, and the Office of Public Roads (OPR), the bill was enacted on July 11, 1916.

The Federal Aid Road Act of 1916 [264] established principles of state and federal cooperation in financing road projects. The decision-making process for funding involved coordination between federal government representatives, led by the USDA Head, and state representatives from the State Highway Departments. The act encouraged the formation of state-level organizations for road construction, as 17 states lacked such departments or similar agencies at the time of its enactment. To qualify for federal assistance, a state needed to have such an agency [296]. The focus was on funding individual projects for constructing, repairing, or improving specific roads or groups of roads based on their importance within the state, rather than creating a unified road system. The budget for each project was approved by the Secretary of the Treasury, with appropriations distributed among the states based on factors like area, population, and road network extent.s

Local State Highway Departments, responsible for both construction and ongoing maintenance of the roads, managed the efforts. This arrangement was logical as road ownership remained with the states, aligning with the act's goal to promote self-help and enhance state control over road policies [296]. If reports of inadequate maintenance reached the USDA Head, they would alert the local State Highway Departments. If improvements were not made within four months, the USDA Head would cease approving projects from that department until the roads were properly maintained. Overall, construction was under joint federal and state control, but the roads remained state-owned. States initiated projects, selected routes, established standards, and maintained the roads, consistent with their constitutional rights, while the federal government retained supervisory roles.

Two key provisions underscored the constitutional autonomy of states in the Federal Aid Road Act of 1916 and subsequent legislation leading to the Hayden-Cartwright Act of 1934: half of the funding came from state funds, and project proposals and development were managed by State Highway Departments.

Construction of roads within the National Forests, which are federal lands and not part of any state, was funded separately. The approval process for these projects considered the ratio of the estimated road cost to the anticipated revenue from accessing the resources the road would provide, with the cost not exceeding 10% of the projected revenue. In contrast, projects submitted by State Highway Departments were not subject to such financial constraints but faced other criteria later deemed irrational and overly restrictive.

The entire process, from project development to implementation, was overseen by the Office of Public Roads (OPR). State Highway Departments could access informational and personnel support from the agency at any stage.

The financing under the Federal Aid Road Act of 1916 was primarily for "rural post roads," defined as roads used or intended for mail carriage [264]. Other roads, despite their potential significance, were ineligible for funding under this act. A mandatory condition was that all roads financed under the act had to be toll-free. Initially, $5 million was allocated, with appropriations expected to increase to $25 million. However, the United States' entry into the First World War in 1917 temporarily halted road development plans.

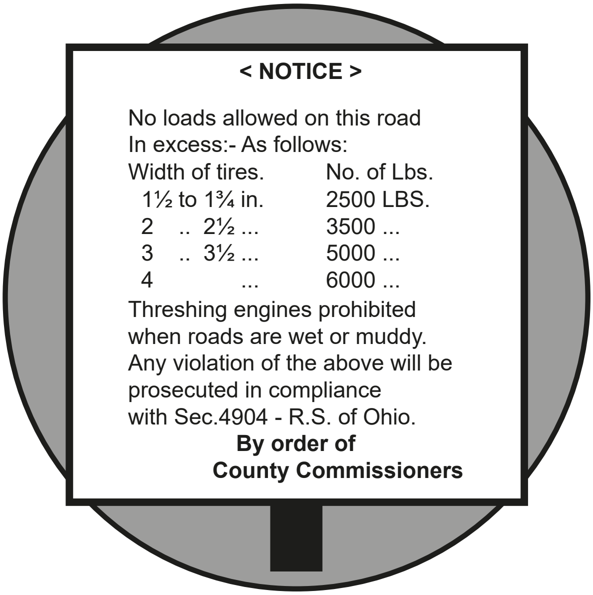
The World War I as a factor of the development of road construction in the USA

The transition of the American economy into wartime production and consumption resulted in a significant increase in the volume of goods and people being transported. The transportation demands of the war overwhelmed the railway system to such an extent that the government nationalized it from December 1917 to March 1920. The war caused shortages of people, money, materials, and railway cars, affecting not only road construction but also military needs. It became evident that the railroads could not handle the military traffic, leading to some of the burden being shifted to highways. Naturally, the gravel surfaces, which could barely withstand regular automobile traffic, were severely damaged by the increased military traffic [175]. Hundreds of miles of roads were destroyed by trucks in just weeks or months, and even asphalt roads suffered damage. Sometimes, a newly constructed road in good condition could become impassable within a few days, even with traffic volumes of fewer than ten trucks per day [129].

Nevertheless, automotive transport provided significant support in alleviating congestion at railroad terminals. Instead of waiting for recipients to pick up their goods themselves — a very outdated procedure — freight deliveries and paid storage at terminals were introduced. Railroad traffic was completely controlled by the Railroad War Board, and this, combined with the army's needs, made it difficult for private businesses, including farmers, to access railroad transportation. This situation pushed farmers to rely on trucking for transportation from rural areas to cities and prompted industrialists to use automobile transport between cities.

Indeed, World War I served as a catalyst for the emergence of commercial freight transportation, further highlighting the need for quality roads in the business and farming sectors. In 1918, automobile transport was quite expensive, and high-quality roads had the potential to reduce these costs.

In the tactical perspective, the process of increasing freight transportation further burdened the road infrastructure. The deterioration of gravel roads led to rapid vehicle wear and high repair costs for states. Additionally, the Railroad War Board rarely allocated railway cars for transporting repair materials, necessitating their transport by trucks. This created a vicious cycle: high road traffic required repairs, but repairing the roads necessitated even more traffic. This problem prompted the Bureau of Public Roads (BPR) to explore construction



< NOTICE >

No loads allowed on this road
In excess:- As follows:

Width of tires.	No. of Lbs.
1½ to 1¾ in.	2500 LBS.
2 .. 2½ ...	3500 ...
3 .. 3½ ...	5000 ...
4 ...	6000 ...

Threshing engines prohibited
when roads are wet or muddy.
Any violation of the above will be
prosecuted in compliance
with Sec.4904 - R.S. of Ohio.

By order of
County Commissioners

Figure 4.27: A road limit was one answer to the problem of heavy trucks breaking up the road surface

methods that would eliminate the need for transportation. Such methods were found, and roads made of sandy-clay materials were constructed using materials readily available on site. Although the roads built using this method were of relatively low quality, they were acceptable for routes that didn't transport heavy industrial goods, such as routes in the southern states [176].

Indeed, there were situations where road quality became a matter of national security. The war in France necessitated the use of automobiles, and a military order in September 1917 declared that the army expected $30,000$ trucks. Consequently, the question arose regarding their transportation from the factory in Toledo, Ohio, to the ports. Since the railways were overloaded, and trucks were one of the few categories of military products capable of independent movement, it was decided to establish special military routes.

The military routes posed a significant challenge for the Highway Departments of the states through which the routes passed. The situation was further complicated by the fact that the transportation of trucks took place during the autumn-winter period in mountainous terrain, requiring constant clearing of the roads after snowfall. However, this experience demonstrated that it was feasible to maintain a functioning road during winter, and that reliable communication between states was necessary for military purposes. The project proved successful, as on December 17, 1917, all trucks except one reached Baltimore [218]. From this point, a discussion arose about certain roads being national security assets, and there were calls from the American Association of State Highway Officers (AASHO) to designate military roads for each state to assess maintenance costs in advance. Additionally, the active collaboration between the State Highway Departments and the federal government, represented by the OPR, strengthened the positions of the departments and prompted more comprehensive work in analyzing the transportation situation and classifying existing roads. Not every road was suitable for military routes, which necessitated careful evaluation by the Departments. The strengthening of the positions of the State Highway Departments is important in the context of highway development because, according to the FARA 1916 and similar projects, it was expected that the initiative for submitting projects would originate from the departments. In the future, it would be the responsibility of the State Highway Departments to initiate and present projects, while the federal authorities would only approve or disapprove the projects [8].

On December 9, 1918, Logan Waller Page, the director of the Bureau of Public Roads, passed away due to a heart attack. In early 1919, Thomas Harris MacDonald became the new leader of the bureau. At that time, the Bureau of Public Roads had a poor reputation among highway officials, whose interests were represented by AASHO. The reputation was tarnished because Page was very insistent on complying with federal requirements and intolerant towards those who disagreed with his scientific views. MacDonald, on the other hand, with his technical education and experience as an engineer, proved to be an ideal successor. He was a "man of the region" and believed that the problem of developing the highway network should be addressed only in close collaboration with state authorities and public organizations.

In a letter from Thomas Harris MacDonald to USDA Head David Houston on March 20, 1919, MacDonald expressed the belief that the relationship between the departments should become more friendly. This could be achieved through decentralization of responsibility and an increase in engineers' salaries. MacDonald believed that the relationship between the federal center and the implementers on the ground needed to be liberalized, and he proposed the creation of an advisory committee consisting of members of AASHO.

In the same year, on February 28, the "Post Office Appropriation Bill," prepared by Logan Waller Page, came into effect, amending the FARA 1916. It expanded the concept of "rural post road," allowing funding to be allocated for the con-

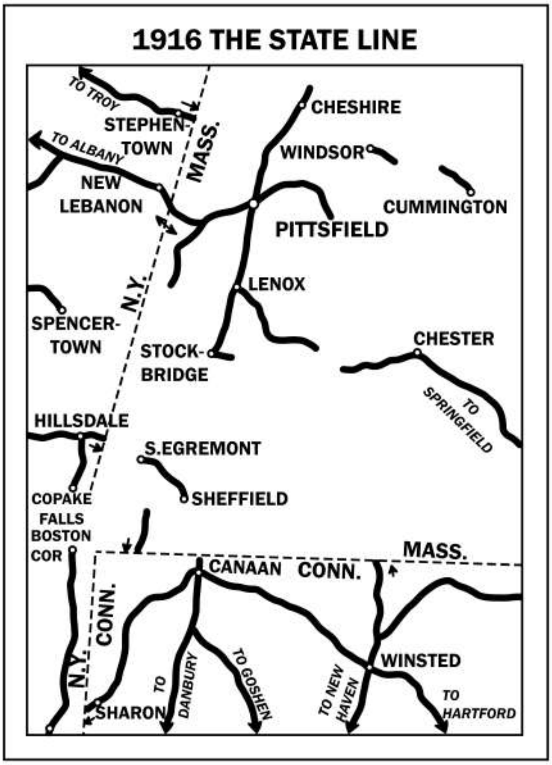


Figure 4.28: A drawback of the 1916 Federal Air Road Act was that federally aided roads need not have been connected.

struction of roads not only used for mail delivery. MacDonald emphasized the significance of this transformation in his article “Four Years of Road Building Under the Federal Act.” [175]. The new chief of the Bureau of Public Roads, Thomas Harris MacDonald, was highly active in the press and on radio. He went beyond merely informing about new construction methods and actively advocated for the creation of a national highway system [157].

Indeed, the Bureau of Public Roads was at the forefront of this effort, undertaking tremendous work to improve the condition of roads in the United States. It conducted research on construction methods, collected information on the state of roads across the country, organized public campaigns to popularize advanced road construction methods, provided guidance to State Highway Departments and the USDA, and trained civil engineers in road construction.

The war significantly altered public perception regarding the purpose of road construction. Before the war, most members of Congress viewed roads from the station to the farm as merely extensions of the rail network. However, after the transport collapse of 1918, this view fell out of favor. As reported by the *Engineering News-Record, «which had formerly favored federal aid, came out for a national highway system declaring that the federal-aid projects were so scattered that there was no hope that they could ever be connected into a workable national system»*[285]. The necessity to develop not only individual high-quality roads but also a cohesive road network became increasingly evident.

Federal Aid Highway Act of 1921 and transition to federal highway planning

The road construction management system in the United States was prepared to begin creating a national highway network. Most states had met the federal government's requirement to establish State Highway Departments. State authorities also took the necessary financial measures to ensure their ability to fund their share of the road construction projects' costs. Typically, this involved securing special loans [218].

MacDonald recognized that an analysis of the current state of affairs was crucial for creating a national highway network. In 1920, a classification of all roads in the United States was conducted based on their importance and service characteristics. Drawing from the experience gained during the war when the road network was ill-prepared for military traffic, MacDonald sought assistance from the military.

In 1921, MacDonald submitted a request to the United States Army Corps of Engineers to determine which roads would have priority during times of war. Based on the report received, a road system map was created, consisting of $200,000$ miles of interconnected major roads, also known as the "Pershing Map." General John Joseph Pershing himself presented the results to Congress in 1922 and supported the creation of a federal highway network. A key insight from the report was that a system capable of meeting the country's industrial and commercial needs would also adequately serve military requirements [8].

The program developed by MacDonald in close collaboration with the AASHO was outlined in the Federal Aid Highway Act of 1921 (FAHA). This act shifted the focus from postal roads to "highways," and it introduced the term "Interstate highways system." The FAHA of 1921 defined the concept of a highway and addressed various aspects such as the right of passage, bridges, drainage structures, signs, fences, and protective structures along the highway. However, it excluded all highways or streets within a municipality with a population of 2, 500 or more, as determined by the latest available census. This exclusion applied unless the section of the mentioned highways or roads had houses spaced more than 200 feet apart along a continuous stretch of one mile [264].

With the Federal Aid Highway Act of 1921, federal intervention in road construction extended beyond financing. The objective was to unify highways into a system capable of addressing nationwide concerns, as outlined in the act's sixth section. This system was to be created through collaborative efforts between the federal government and state administrations. They needed to jointly determine which roads would comprise the highway system, with the condition that the share of highways in each state's road network should not exceed seven percent.

Indeed, the influence of the war is evident in the Federal Aid Highway Act of 1921. The act specifically stipulated that control over the highways, which during the years of the First World War was under the purview of the Council of National Defense, would be transferred back to the head of the United States Department of Agriculture (USDA).

In the FAHA of 1921, compared to the FARA of 1916, there was increased attention given to the durability of road pavement and its ability to withstand heavy loads - the very parameters by which the American road system did not pass scrutiny in 1917—1918 [264].

Throughout 1922 and 1923, the development of the highway network project progressed in Congress with invaluable assistance from the Pershing Map and the Bureau of Public Roads, culminating in its completion in November 1923. The proposed total length of the roads in the project was 272, 000 kilometers, accounting for 5.9 percent of all public roads [264].

Over a span of thirty years, from the late 1890s to the early 1920s, there was a shift in road construction from being primarily undertaken by local authorities (often by local residents themselves) to federal planning of the highway network, considering the social, economic, and strategic interests of the country.

Following the enactment of the FAHA of 1921, significant efforts were made to implement its provisions. The Bureau of Public Roads (BPR) regularly summarized the progress in numerous brochures. By May 30, 1922, construction of 17,039 miles of highways had been completed, with an additional 14,491 miles under construction, excluding roads on federal lands. The BPR focused on developing advanced construction methods while ensuring the majority of the roads were suitable for rural areas.

Thus, the emphasis was not only on the technological aspects of road construction but also on simplicity and cost-effectiveness in operation. It was crucial to adhere to proper construction techniques and ensure proper drainage. Additionally, it was important for the roads to be capable of handling wear and tear without requiring frequent repairs.

The Bureau continued its active promotional efforts by traveling throughout the country with models of roads and bridges, providing commentary on the construction technology involved and offering warnings about the most common mistakes.

Indeed, one of the most significant contributions was that the BPR continued consultation with the State Highway Departments in the development and creation of the highway network. The BPR served as the foremost expert organization in the United States on this subject matter. The story of this stage of creating the US highway network can be concluded with the opinion of MacDonald, that these efforts constitute perhaps the greatest road construction program in world history ever undertaken under unified control [176].

4.4.2 The Roaring Twenties

Problems and challenges of federal road construction planning after World War I

Indeed, the 1920s were a defining period in the development of the American highway network. It was during this decade that the conditions were established, allowing the United States to become a "nation on wheels." The creation of an interconnected highway system was instrumental in achieving this transformation, and it was in the 1920s that the primary directions and methods of its development were established.

Federal funding for road construction became part of a larger effort to stimulate the American economy through government contracts. This process began during World War I and continued for some time after its conclusion [159].

The Bureau of Public Roads played a leading role in advancing road construction in terms of both scientific research and the development of specifications for surfacing, drainage, and other engineering structures.

The economic and demographic situation in the states was also a crucial factor that needed to be considered in the development of the national highway system, as required by the FAHA 1921. A dedicated group was established to develop an evaluation system for projects submitted by the State Highway Departments. This group was led by Edwin James, a professional highway engineer at the Bureau of Public Roads (BPR). Using data from the United States Census Bureau, James' group analyzed each county in every state based on demographic indicators and the level of industrial or agricultural development. Based on this information, the James group identified the most economically important routes for each state and county and compared them with the projects submitted by the State Highway Departments. In most cases, the projects proposed by the Bureau and the states aligned, and any differences were resolved through conferences. The work on developing the national highway system was completed in July 1923 [8]. This system is often referred to as the "Federal Aid System" or the "seven percent system," as according to the FAHA 1921, it was intended to include seven percent of the total mileage of state highways.

The development of federal road construction faced difficulties not only due to the aftermath of the First World War or the tensions between federal and state authorities but also due to the peculiarities of the economic development of the United States.

The consequences of material and equipment shortages during the First World War would continue to be evident in the research strategy of the Bureau of Public Roads (BPR) for some time. In the early 1920s, the BPR published brochures describing construction methods that did not require transporting materials, following the principle of *"everything you need can be found right on the roadside"* [176]. These statements become more understandable when considering the wartime experience, where any trip of a truck loaded with construction materials would lead to the destruction of the road surface.

The early 1920s were marked not only by the passage of the FAHA in 1921 but also by an economic downturn. The recession, which lasted from January 1920 to July 1921, was characterized by an extreme decline in GDP and an unemployment rate of around 8%[118]. Unemployment was caused by the transition of the economy from wartime production to peacetime activities, leading to a surplus of labor in the market due to demobilization. The state of the economy also affected road construction, as federal assistance was reduced by implementing a cost limitation of $20, 000 per mile for construction projects [8].

However, there was no complete elimination of federal assistance programs. The FAHA 1921 act was passed, and the federal funding program was extended until 1924. This was due to the recognition that even proponents of "farm-to-market" roads acknowledged the need for a national road system during the war to achieve national goals. Furthermore, during the pursuit of these goals, the road network suffered significant damage, and it would be unfair to burden the states alone with the responsibility of rebuilding the destroyed roads. This sentiment was expressed by John Hollis Bankhead, one of the authors of FARA 1916, in his report to the Senate on January 27, 1919. Bankhead had advocated for limited federal involvement in road construction prior to the war [20]. Indeed, the government saw the development of road construction as an opportunity to provide employment for veterans and reduce unemployment. Some researchers highlight this as a reason for the adoption of the FAHA 1921 [128].

In 1919, road construction activity intensified; however, it coincided with the seasonal transportation of coal. This overlap resulted in a shortage of open railcars available for transporting construction materials, causing the construction process to proceed slowly. Many contractors were unable to complete their work before winter, and by the end of the season, only about half of the projects approved by the USDA Head for 1919 were completed [8].

Despite the challenges of 1919, in 1920 the USDA sponsored even more road construction projects. The road contract market became saturated, and prices for contractor services increased, leading to delays in the completion of ongoing projects. Additionally, the construction process was prolonged due to a mismatch between the volume of work and the available equipment, exacerbated by a shortage of equipment necessary to handle the high number of projects.

Chief MacDonald called for a halt in approving new projects and advocated using the economic downturn as an opportunity to meticulously conduct engineering work [114]. Moreover, this "wait-and-see" approach was justified because the proposed FAHA 1921, being prepared during that period, allowed for the possibility of spending more money per mile of road. This meant that more complex and costly engineering solutions could be implemented, or higher-quality road surfaces could be used, providing a compelling reason to await the passage of the act.

Thus, the recession of 1920—1921 had a significant impact on the road construction sector, with only a fraction of the planned projects being completed in 1920. However, industry leaders at both the federal and state levels were able to analyze the situation and draw several important conclusions.

The first conclusion was the need for mechanization in the industry to mitigate the rising costs of construction. The second conclusion recognized the necessity of expanding businesses associated with road construction, such as the production of construction materials and road machinery, as well as raising industry standards. Finally, it became evident that well-developed administrative mechanisms for road construction were necessary to handle the influx of projects.

Despite the setbacks in the 1920 construction season and the loss of appropriations for 1922 (which were influenced not only by the recession but also administrative difficulties), federal financing programs were not terminated. In fact, Congress continued to increase the amounts of federal allocations, signaling a commitment to the future of road construction [8].

Institutional interaction in the field of road construction

The central agency that played a pivotal role in national road construction and was involved in almost every decision was the Bureau of Public Roads (BPR).

The interaction between the Bureau of Public Roads (BPR) and State Highway Departments was constant, active, and fruitful. These State Highway Departments were responsible for road construction and maintenance. By 1916, almost every state had established a Highway Department or Council with similar functions. By the 1920s, most of these departments included professional highway engineers [76], often trained under the auspices of the BPR. However, these engineers did not always have decision-making authority in all aspects of their work [113]. State Highway Departments frequently sent requests to the BPR for information on road construction methods or direct instructions. In turn, the BPR requested up-to-date data on road conditions and statistical information from the Departments. The BPR influenced the policies of the Departments through its control over the approval of projects submitted by the Departments to receive federal assistance. Federal resident engineers, who not only assisted the Departments in their work but also carried out supervisory functions, played a role in this process. For example, they monitored the proper maintenance of roads built under federal funding programs. The lack of maintenance could lead to the suspension of funding for projects from states that failed to fulfill their obligations.

Herbert Clark Hoover's contribution also included convening the First National Conference on Street and Highway Safety. This conference, which began its work in December 1924 in Washington, D.C., was a notable example of "consensus politics." The conference delegates represented a wide range of state and private agencies, public movements, and associations. State commissions, insurance companies, railroad companies, safety councils, chambers of commerce, trade unions, women's clubs, automobile associations, and car manufacturers — all were represented at the conference, aiming to facilitate a comprehensive discussion of the issue and maintain a balance of interests [8].

In the 1920s, the design standards for American road systems in the United States were de facto established by two individuals: BPR Chief MacDonald and BPR Chief of Design Herbert S. Fairbank. Ultimately, state governments had to raise their standards sooner or later under the pressure exerted by the BPR [256]. Professor of History Frederic Logan Paxson, who witnessed the development of the American road system, noted that the expansion of the highway network itself weakened the autonomy of the states [218].

It is also necessary to mention that the process of road construction was closely associated with constant interaction with military agencies, the Council of National Defense, and the Department of Defense.

Starting from 1921, as mentioned earlier, the Army transferred equipment and materials (mostly trucks and dynamite) to the State Highway Departments that were no longer needed after the end of military operations [264]. According to the amendments to the Federal Aid Highway Act of 1921, which came into effect in 1925, the USDA continued its stable cooperation with the United States Department of Defense. The transfer of military equipment is no longer mentioned in the 1925 Act, but the transfer of explosives remains relevant [264].

Development of the national road system

The 1920s were highly favorable for road construction. Under the Federal Aid Highway Act (FAHA) of 1921, the funding limit was raised to $20, 000 per mile, facilitating the construction of high-quality roads capable of handling heavy traffic. After 1923, Congress began planning the budget for highway construction several years in advance, up to 1925, enabling states to make long-term plans without the fear of sudden funding shortages. This budget approval format was particularly convenient for states where legislatures met only once every two years. State Highway Departments refined their organizational procedures after the setbacks of 1920 and 1921 and fostered relationships with local construction industry representatives, including contractors for future projects. Following the recession of 1921, material prices and wages decreased again. All these factors combined to create a prosperous period for road construction after 1922 [8].

By 1925, over a quarter of the planned mileage outlined in the 1923 project had been constructed, totaling 178, 000 miles [274]. However, the development of the automotive industry demanded continued progress. The system planned in 1923 was clearly insufficient by 1925, as the United States had nearly 20 million automobiles on the roads, placing increased strain on the existing infrastructure. Motorists desired more interconnected routes, and this demand extended across the entire country. In the 1920s, the automobile ceased to be a luxury of wealthy city dwellers and became accessible to people without substantial capital, such as farmers. The widespread adoption of automobiles is evident from the fact that in 60% of the states, the number of registered vehicles exceeded 100, 000 [274].

Yes, the creation of a national highway network involved not only the construction of roads but also the development of uniform construction standards. The Bureau of Public Roads (BPR) collaborated with the American Society for Testing and Materials (ASTM) to establish these standards. The BPR played a crucial role in codifying the highway system by creating a unified registry that included the numbering and classification of highways.

The process of establishing a unified numbering and marking system for the interstate highway system began in 1924 when the American Association of State Highway Officials (AASHO) directed a resolution to the head of the U.S. Department of Agriculture (USDA), emphasizing the need for such actions. The task of developing such a system was undertaken by a council within the BPR, led by Chief MacDonald, with Edwin James serving as secretary. Prior to this, Edwin James had been in charge of the group responsible for creating the national highway system under the Federal Aid Highway Act of 1921.

The group decided to limit the numbered routes to those included in the federal aid system. Six meetings were conducted with regional commissions to identify the most important routes to be designated and ensure their integration into a unified system. Finally, on November 11, 1926, the USDA Head approved nearly ten thousand miles of the designated system. The approved routes were marked with black and white signs displaying the route number [8].

The creation of the numbered system facilitated the identification of routes and therefore the use of the highway network [218]. The emergence of specialized touring guides, known as Automobile Blue Books, also contributed to making the use of the highway system more accessible to travelers. The publication of the Automobile Blue Book was handled by the American Automobile Association (AAA).

An important task for the BPR was to unify disparate routes into a single highway network, as prescribed by the FAHA of 1921. In the BPR's annual report for 1922, Chief MacDonald stated that despite the construction of 10, 000 miles of highways during the reporting year, which was equivalent to three transcontinental routes, one could not yet speak of a transcontinental highway system in the United States, nor even of the existence of a unified and functioning road system [237]. Creating such a system was the primary goal facing the authorities in the field of road construction in the 1920s.

How was this task supposed to be addressed? The primary method of road construction was known as "roadbuilding by stages." This approach, proposed by BPR Chief MacDonald in 1923, involved dividing the road improvement process into stages and intentionally delaying certain construction phases [237]. Initially, it was feasible to effectively drain and level the road, allowing for the installation of a cost-effective sand-clay surface in the first stage. In later stages, this surface could be upgraded to higher-quality, more durable materials such as blocks, concrete, or asphalt [237]. Although this staged construction did not extend the overall length of the highways, it enhanced the quality of the sections already built. This method facilitated the rapid improvement of the system's poorest parts, particularly in the southern and prairie states. By 1926, 11% of the total length of US highways had been improved using this method. In contrast, the northeastern states utilized this method less frequently, as they were constructing roads with surfaces better suited for heavy industrial traffic and continuous vehicle movement, such as concrete and brick (26% of the total improved mileage in 1926).

The strategy of staged construction was largely dictated by limited funding, as explicitly stated in the annual reports of the BPR [237]. Simultaneously, it responded to the country's increasing motorization, marked not only by a growing number of vehicles but also by rapid technological advancements in automobiles. Staged construction conserved federal funding by avoiding the installation of pavement that would soon become outdated [8].

In 1925, the need for a new federal funding act arose, as no appropriations had been planned beyond that year. Amendments to the acts of 1916 and 1921 proposed allocating $75 million annually for federal highway construction programs until 1927, a measure extended through 1929 in 1926 [264]. This stable funding allowed State Highway Departments to focus on both quick, cost-effective projects and more strategic, long-term investments. The amendments of 1925 and 1926 ensured that the Departments would have continued access to these funds, enabling them to complete their projects.

By 1925, the planning and construction of the road network had been effectively streamlined. This was evident from the stable financing and the strong connections between public institutions, federal agencies, state departments, and contractors. The construction of each new road segment was met with enthusiasm from local residents and marked by celebratory events, including ribbon-cutting ceremonies and speeches by politicians [218]. The favorable economic conditions, the need to accommodate an increasing number of automobiles, the development of intercity freight and passenger transportation, established interagency relationships (with checks and balances within the sector), and the political acumen of road construction leaders all contributed to the highly productive nature of the 1920s. By 1929, over 180,000 miles of the federal aid system (the national road system) had been constructed [274], with 90% of that mileage improved: properly aligned, drained, and surfaced with quality materials [8]. In 1923, the BPR stated that 90% of the people in the United States lived within 10 miles of the developing system. The completion of this system was bound to significantly impact the country's economy and improve the quality of life for Americans [238]. By 1929, the federal aid system, as planned after the enactment of the Federal Aid Highway Act (FAHA) in 1921, had been completed just over a quarter of its intended scope.

4.4.3 The Great Depression

Reflection of the economic crisis on road construction

The Great Depression was a severe crisis that began with the collapse of the New York Stock Exchange on October 29, 1929, and spread to almost all sectors of the American economy and politics, including road construction, which was affected by both the economic downturn and governmental measures to address it.

First, we must examine how the economic crisis directly impacted road construction. According to federal assistance programs, road construction funding was split equally between federal funds and state-allocated funds. States were particularly hard-hit, especially those that derived their road construction funds from property taxes or income taxes [8]. Additionally, state funds generated through bond issuance suffered as well. Consequently, substantial expenses arose as interest payments on these bonds had to be made, totaling $90 million by 1932 [8]. However, revenue from vehicle registration and transportation fees did not decline as significantly as other sectors [274]. Furthermore, revenues from fuel taxes actually increased during the early years of the Great Depression. By the end of 1929, fuel taxes were collected in all states [32], and by 1932, these revenues had exceeded $807 million. Nevertheless, due to the sharp decline in overall revenues, these funds were often diverted to non-highway purposes such as unemployment benefits.

Representatives from the automotive and related industries were dissatisfied with this situation, leading to changes in state legislation that explicitly prohibited the non-targeted use of fuel tax revenues [8]. In response, the Hayden-Cartwright Act was enacted in 1934, stipulating that states not spending transportation taxes on highways by June 30, 1935, would be ineligible for federal assistance [264].

From this, it is clear that a new goal for federal authorities emerged during the Great Depression: to ensure that transportation taxes, such as fuel tax and vehicle registration fees, were used solely for highway purposes. This goal was achieved with the implementation of the Hayden-Cartwright Act.

The decline in state revenues also compromised their ability to contribute equally to the federal funding system for highway construction. Since 1916, half of the funds in federal assistance programs were expected to come from state budgets, maintaining their constitutional autonomy. However, reduced tax revenues made it challenging for states to meet these contributions [240]. The Emergency Relief and Construction Act (FERCA) of 1932 addressed this by providing funds to states in advance, with the obligation to repay them after 1938 through the retention of a portion of funds allocated within the same federal assistance program [264]. With the passage of the Hayden-Cartwright Act in 1934, all previous advances were converted into grants, thus allowing a clearer distinction between federal funds and state autonomy and removing the perception that federal funding of highway construction violated state autonomy.

The crisis also affected road construction through the contraction of industries linked to construction materials. The reduction in production within the steel, lumber, and cement industries impacted road construction, which had already begun to decline by early 1929.

The reduction manifested in a decrease in the mileage of new roads constructed and roads improved within the framework of staged roadbuilding [239]. However, the impact on road construction was not as severe as in many other sectors. The reduction in funds only significantly affected road construction by 1932, but by 1934, both government expenditures and toll revenues had exceeded the 1929 levels and continued to grow until the end of the decade. This economic resilience in the construction industry was unique during the Depression period [47], attributable to two factors: transportation revenues, which formed the basis for state highway department funds, did not decrease as sharply as other tax types, and federal funds continued to support road construction throughout the crisis.

In response to the early challenges faced by road construction, increased federal funding was introduced. In 1929, Congress raised appropriations for 1932 and 1933 to $125 million, which was two-thirds more than the previous decade's standard of $75 million [264]. However, states struggled to match this funding due to financial difficulties related to bond obligations and declining tax revenues. The equivalence of financial contribution was a crucial condition for federal road construction funding since 1916 [8]. This issue was addressed by the Emergency Relief and Construction Act (FERCA), passed on July 21, 1932, which allocated $300 million from The Reconstruction Finance Corporation funds to provide assistance to those in need and alleviate the hardships caused by unemployment.

Under the FERCA, the federal government planned to advance \$120 million to the states specifically for road construction purposes.

The development of the federal financing system progressed with the passage of the National Industrial Recovery Act (NIRA) during the Roosevelt administration. This act was a key component of the New Deal, which aimed to address the economic downturn and the humanitarian crisis resulting from widespread unemployment [264].

An important enhancement to the federal financing system was the expansion of the term "highway." It was no longer limited to highways located outside of certain-sized populated areas but now also included highways within those areas. Additionally, funds were allocated for the construction of secondary and access roads, previously excluded from federal assistance programs.

The next significant legislation was the Hayden-Cartwright Act, which addressed emergency highway construction and related projects. This act allocated \$200 million, distributed by the USDA Head to the States Highway Departments. The USDA Head's approval of specific projects submitted by these departments served as a commitment to finance those projects.

The Hayden-Cartwright Act also allowed for the use of funds in projects aimed at enhancing transportation safety and convenience, indicating that road safety was becoming a vital factor in road construction strategies, alongside the NIRA.

It was also mandated that taxation related to transportation should be used specifically for transportation needs. States failing to allocate transportation taxes and fees accordingly would be excluded from federal road construction funding programs.

Public works funding, including for road construction, continued through the Emergency Relief and Appropriation Act of April 8th, 1935. This act aimed to allocate $\$4,000,000,000$ to increase employment through construction and other socially significant projects. Specifically, $\$800,000,000$ was intended for highways, roads (not qualifying as highways), streets, and crossings. The requirement for states to contribute an equivalent share of funding was removed, affirming the provision introduced by the Hayden-Cartwright Act.

The President's powers were expanded, allowing the establishment of rules and directives for the expenditure of funds. For instance, the President could determine the maximum working hours and wage rates for each state. Under this act, the President exerted significant control over wages in public works, setting wage rates directly for any project.

The law also emphasized making transportation safer, particularly by addressing the elimination of railroad level crossings [264].

In 1938, Congress passed the Federal Aid Highway Act of 1938, which allocated $100 million for the fiscal year ending on June 30, 1940, and $115 million for the following fiscal year. The law specified separate amounts for approaches, farm-to-market roads, and roads within the free mail delivery system, and allocated specific amounts for traffic safety [264].

The FAHA of 1938 tasked the BPR with studying the feasibility of constructing an interstate system. This system would consist of superhighways crossing the country from east to west and north to south. Unlike the highways included in the federal aid system, which were primarily planned at the state level, the superhighways were intended to span the entire country from border to border and coast to coast [264].

This led to the presentation of the report "Toll Roads and Free Roads" by Chief MacDonald in 1939. It contained analyses and forecasts regarding the use of superhighways and their potential financial profitability. Based on current traffic data, MacDonald concluded that no highway would generate enough demand to fully cover its costs, but the construction of these highways could still be a profitable investment for the government [7].

Road construction was an arena where many political processes unfolded during the Great Depression. It saw the strengthening of presidential power, exemplified by the Emergency Relief and Appropriation Act of April 8, 1935. There was a pressing need to address unemployment, and even a shift in attitudes towards private property, as many Works Progress Administration (WPA) projects were implemented on privately owned lands [128].

Researchers and observers of the Great Depression concur that this period marked a significant move towards centralization, a trend long resisted by the political structure of the United States. This shift was evident in the increased control federal authorities had over working conditions in both private and state projects, as well as in financial terms. Between 1930 and 1940, federal assistance to states (excluding loans and direct federal expenditures such as public works) rose from approximately \$135 million to over \$580 million. As part of this process, the financing of road construction became a means through which states lost financial autonomy. In 1930, over half of the federal funds allocated to states were directed towards highway construction [128].

The funding for road construction increased throughout the decade, reflecting the priorities of the federal government in this sector. Alongside enduring tasks such as supporting mail delivery and agricultural activities, the decade solidified additional goals like road safety and addressing the legacy of horse-drawn traffic.

Despite the financial implications, the model for providing assistance preserved state autonomy in initiating commitments, a condition firmly upheld by Chief MacDonald [128].

However, the primary goal set by the US government and Congress for road construction was employment.

Following Roosevelt's re-election for a second term, the Federal Emergency Relief Administration (FERA) of 1932 was succeeded by the Works Progress Administration (WPA) under Harry Hopkins in 1935. Road construction was prominently included in the WPA's projects. The Division of Engineering and Construction achieved significant milestones, completing the construction and improvement of over $600,000$ miles of highways, roads, and streets, and the construction or reconstruction of over $116,000$ bridges and viaducts spanning more than 700 miles by June 1941. Over 4 billion dollars were spent on these infrastructure projects [128].

Road construction was notably effective in providing employment opportunities, with construction of highways, roads, and streets accounting for 44% of all workers employed concurrently on Division projects [128].

In conclusion, the government leveraged road construction as a dual-response to the crisis: as a sector for employment and as an area where new labor protection norms were applied. Across three public works programs, road construction demonstrated significant results in terms of completed projects and job creation. Harold LeClair Ickes, in his memoir "Back to Work: The Story of the PWA," credited this success to the effective leadership of the Bureau of Public Roads (BPR), which had been organizing road construction for over a decade.

The Bureau regularly conducted audits to ensure the efficient utilization of federal aid funds [47].

Highways and military strategic planning

Indeed, road construction has always been seen as vital for national security. It has been considered important for the defense of the country, as well as for facilitating its settlement, thereby preventing occupation and claims by foreign governments [128]. After the experiences of World War I, periodic revisions of strategically important highways were conducted by the Department of the Army and the Bureau of Public Roads. The 1939 revision, prompted by the looming threat of World War II, revealed that out of the 74, 600 miles of strategically important routes, many thousands did not meet the standards in terms of width, surfacing, or other parameters. The most critical issue for troop transportation was found to be the condition of bridges: 2, 400 bridges were unable to withstand the load set by the AASHO standards.

Despite Congress's awareness of these issues, highway programs were not coordinated with port development and military base programs. Consequently, many military bases were left in a state of logistical isolation, unable to upgrade the roads leading to the bases because they were not part of the federal aid system [8].

Thus, despite the collaboration between the Bureau of Public Roads and the Department of the Army during and after World War I, the creation of an effective national defense highway system was not achieved, largely due to inadequate communication with state authorities, who could have included the necessary routes in the federal aid system.

Modern transportation systems and their modelling

> The waiting time for the bus is directly proportional to the adverse environmental conditions
>
> —Popular wisdom

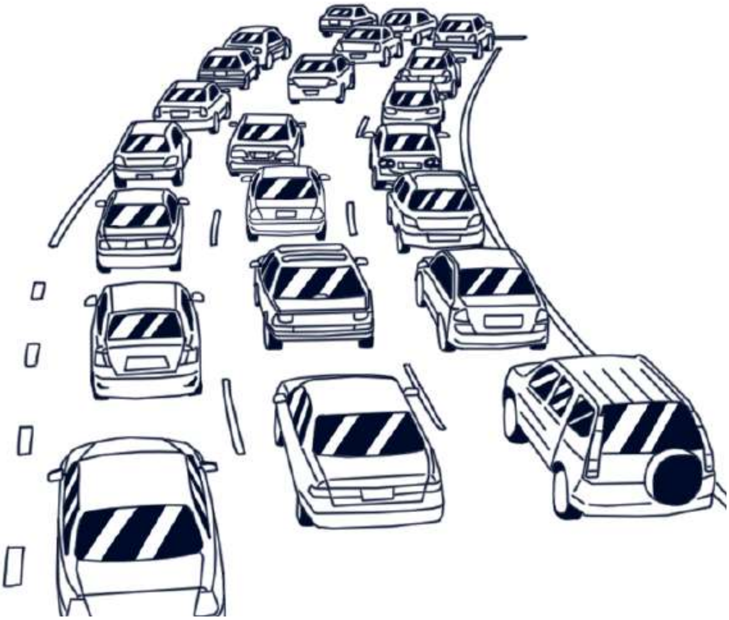


Figure 5.1: Traffic jam

5.1 Transport topology paradoxes

The Network Design Problem is one of the most discussed problems that was set before the appearance of the transport simulation as a discipline.

Where do traffic jams come from? We will demonstrate "on the fingers" one of the main causes of traffic jams, which is important to know. Suppose we have a flow of 2000 cars per hour from point A to point B, and there are two roads (routes) from A to B. The first road is "fast" and takes 20 minutes (if there are no traffic jams), but no more than 1200 cars per hour can pass along it. The second road is a "bypass," which can also pass 1200 cars per hour, but the route takes 40 minutes. Initially, all drivers will try to pass on the "fast" road, which will not be able to serve the entire flow. A queue will start to form. As a result of the increase in cars in the traffic jam, the travel time on this road will also increase for newly arrived motorists. At some point, it will seem to some of the users that the traffic jam has grown so much that it is faster to use the "detour" route. So we get an equilibrium in which both alternatives promise the same gain (or loss). It is clear that the selfishness of the drivers prevents the traffic from being immediately distributed so that all the extra cars go via the "detour" route. A traffic jam is a punishment for selfishness — its role is to equalize the possible gain between all options.

5.1.1 Braess's Paradox

In mathematics, and more specifically in game theory, the Braess paradox states that adding a new road to a road network can reduce overall performance when moving entities choose their routes individually. This comes from the fact that the Nash equilibrium of such a system is not necessarily optimal. This paradox was highlighted in 1968 by the German mathematician Dietrich Braess (born in 1938) [30]. He gave an example of a network where the addition of a new route resulted in an increase in travel time on the network structure for all road users.

Figure 5.2: Dietrich Braess

We define the transport network in the form of a weighted directed graph. Let the delays on the arcs of this graph be expressed by the following formulas (Figures ?? and ??):

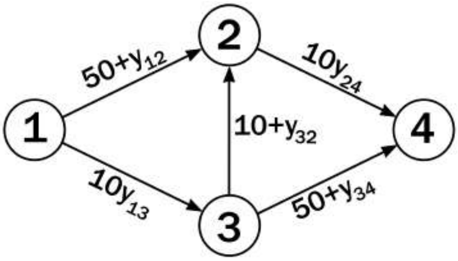


Figure 5.3: Braess's Paradox: with central road

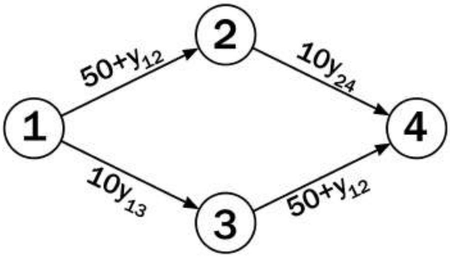
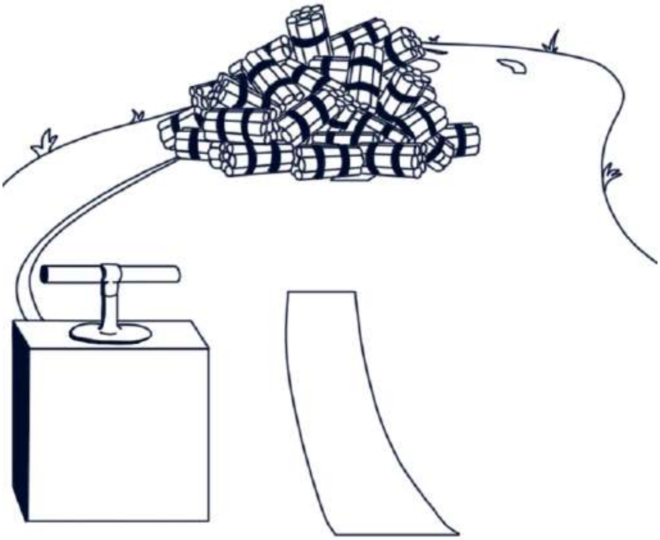


Figure 5.4: Braess's Paradox: without central road

Let the intensity from point 1 to point 4 be 6 vehicles per unit time. Denote by y_{ij} the intensity of traffic along a given arc $i \to j$, and the weights (travel times) at arcs depend on corresponding flows. In the Beckmann model, equilibrium is achieved under the following condition: all paths (for a given Origin-Destination (OD) pair $1 \to 4$) used must have the same "length," otherwise, drivers will choose routes with less travel time.

In total, there are 3 paths from 1 to 4: $1 \to 3 \to 4, 1 \to 3 \to 2 \to 4, 1 \to 2 \to 4$. If all 6 vehicles per time unit are distributed as 2 vehicles per path in unit time, the travel times will equal 92 minutes ($10 \cdot (2+2) + (50+2) = 10 \cdot (2+2) + (10+2) + 10 \cdot (2+2) = (50+2) + 10 \cdot (2+2)$). Such a distribution is the only Nash-Wardrop equilibrium in this network.



Let us close the central arc 3 → 2. The configuration with the flow of 3 vehicles per time unit at 1 → 3 → 4 and 1 → 2 → 4 paths will be the one equilibrium. In this case, the travel time is $10 \cdot 3 + (50+3) = (50+3) + 10 \cdot 3$ = 83 minutes.

That means everyone is better off! It is noteworthy that if we invite people to play such a repetitive game (such games, for example, were played with students of various universities), in which the choice of the route is determined by each player based on previous draws, then the system does indeed converge to the Nash-Wardrop equilibrium in both cases (with road 3 → 2 and without it). This is a general fact of the evolutionary theory of potential games.

We have just described an example of such a paradox. The mechanism of the formation of the Braess paradox largely echoes the prisoner's dilemma of game theory: everyone chooses the optimal strategy for themselves, which only worsens the optimal equilibrium.

This paradox has been encountered in actual road structures in the United States, Great Britain, etc. Here is a list of the most well-known networks with Braess's paradox in real life:

- In Seoul (South Korea), an improvement in traffic around the city was observed when an expressway was removed during the Cheonggyecheon restoration project.

- In Stuttgart (Germany), after investments in the road network in 1969, the situation did not improve until a newly constructed section of road was again closed to traffic.
- In 1990, the closure of 42nd Street in New York reduced congestion in that area.
- In 2008, Youn, Gastner and Jeong pointed to specific routes in Boston, New York and London where this might indeed be the case, and designated roads that could be closed to reduce travel times.
- In 2011, following the closure of Interstate 405, the absence of heavy traffic in a wide area is potentially seen as the most recent example of Braess's paradox at work.

As Roughgarden and Valiant showed in their 2006 paper [278, 248], this paradox is very likely to occur in random graphs. The solution to this paradox is not obvious, as it might seem at first glance. Roughgarden showed in his paper that the problem of finding inefficient edges in a lattice is NP-hard, i.e., practically unsolvable. On the other hand, Milchtaich showed in his paper [187] that the only lattice in which Braess's paradox cannot be realized is a lattice of parallel edges. A consequence of the above is that poorly designed addition of new roads to a road network can increase travel overload in that network.

5.1.2 Other paradoxes and ways to resolve them

The Downs-Thomson paradox, first demonstrated by Downs in 1962, states that the equilibrium speed of vehicles in a transportation network is determined by the average "door-to-door" speed of equivalent trips in public transport [245].

The Pigou-Knight-Downs paradox suggests that increasing road capacity does not always reduce overall vehicle delays. This phenomenon occurs because traffic may simply shift from other roads to the improved road, thereby increasing congestion [288]. This paradox is, in fact, a consequence of the Downs-Thomson paradox.

The Downs-Thomson paradox emerges due to passengers transitioning from public to personal transport in response to delayed demand. This shift decreases public transport ridership, reducing operator profits and compelling them to increase service intervals. This change forces more passengers to switch to private vehicles. Concurrently, the road conditions worsen as drivers, believing in improved road capacity during peak hours, begin to use the roads at these times. These factors disrupt the transport balance, causing a surge in vehicle flow on the extended road, leading to greater congestion and deteriorating public transport services.

The Downs-Thomson paradox is not universal; it applies primarily in areas with well-developed public transport systems and where the existing road network is no longer sufficient to handle the traffic flow. Both experimental and mathematical evidence supports this paradox.

The Lewis-Mogridge position, formulated in 1990 [188], observes that building more roads invariably leads to increased traffic, filling the new capacity. This observation directly references the "Iron Law of Congestion", introduced into transport literature by Anthony Downs [65]. The position describes a process where traffic increases until it occupies all available road space, with the speed advantages of new routes diminishing within months or even weeks. While new roads may alleviate congestion in specific areas, in most cases, the congestion merely shifts to other transport hubs.

According to the Lewis-Mogridge position, it is incorrect to conclude that road construction is generally unjustified. Instead, developers and builders should consider the entire transport system holistically, ensuring that the principles of traffic flow are well understood before proceeding with construction. This position has been used to explain issues caused by private vehicles, particularly urban road and highway congestion, and to highlight the success of central London's toll system.

Mogridge, a British transport analyst, also noted that investment in expanding a city's road network inevitably leads to its saturation and a reduction in the system-wide average speed, including both roads and public transport. These dynamics relate to the Downs-Thomson paradox. However, according to Anthony Downs, these relationships between the average speeds of personal and public transport apply specifically when public transport has dedicated lanes during peak hours. For instance, since 2001 in central London, approximately 85% of all morning rush hour commuters use public transport, which occupies 77% of the trip in a separate lane, while only 11% use private cars. Once the ease of travel by land-based public transport matched that of the metro, the travel times for both modes of transport became roughly equal.

One method to regulate traffic flows is the introduction of tolls. The impact of tolls varies, as seen in different cities. For example, Rome's city center experienced a 15—20% reduction in road traffic and only a 6% increase in public transport usage after implementing an entrance fee. Conversely, the introduction of a toll on Melbourne's roads effectively addressed traffic congestion. However, the M1 toll motorway in Hungary serves as a negative example, where high tolls led to significantly underused roads.

In 1974, Peter Steenbrink published "Optimization of transport networks" [265], focusing on the theoretical and applied aspects of optimizing large transport systems.

Steenbrink describes a cost-minimization heuristic method, initially proposed by Barbier in 1966 and later extended by Haubrich, to optimize the Dutch railway network. This method, similar to the continuous optimal adjustment method, begins by calculating the initial load of the network, considering all possible arcs. Each arc is then evaluated for potential redistribution of flow to minimize the objective function (arc exclusion). After completing the elimination process, a new network is established, and the process repeats until the objective function converges. This method aims to identify and resolve transport topology paradoxes.

5.2 Modern models of traffic flows

5.2.1 The development of traffic models

The paradoxes of transport topology are just one of the reasons why it is necessary to develop an adequate mathematical apparatus for modeling traffic flows.

In the 1950s, in connection with studies of the processes occurring during a bomb explosion, there was a rapid development of gas dynamics (generalized solutions of conservation laws, stable difference schemes for calculating solutions). Simultaneously, the first macroscopic (hydrodynamic) models appeared, in which traffic flow is likened to the flow of a "motivated" compressible fluid (Michael Lighthill and Gerald Whitham, Paul Richards), alongside the first microscopic models (following the leader), where the equation of motion for each car is explicitly written out (Denos Gazis [107], Louis Pipes [227], etc.). The Lighthill–Whitham–Richards (LWR) model (1955) [168] likens traffic flow to that of a compressible liquid and is described by the conservation law of the number (per unit density) of cars. This model postulates the existence of a functional dependence (equation of state) between the flow of cars (= speed × density) and density. This dependence is often referred to as the fundamental diagram (usually a concave function). Indeed, motivation in the simplest models is "wired" into this dependence.

Remember the term "fundamental diagram", this concept remains one of the key concepts in traffic flows to this day, which is not surprising. After all, it shows how efficiently, for example, a road section is used.

In subsequent years, the class of micro- and macro-models was significantly expanded. The next step after the introduction of the LWR model, mentioned as early as 1955 but finally proposed in 1974, was to account for the "farsightedness" of drivers by adding diffusion terms corresponding to the fact that drivers slow down when the traffic density ahead increases and accelerate when it decreases. This model was called the Whitham model.

At about the same time, a new model was created — the Payne model [220], which no longer assumed that the desired speed was achieved instantly. In the modern macroscopic approach [18], traffic flow is often described by a nonlinear system of hyperbolic equations (for density and flow velocity) with diffusion. In this case, the equation of state is "sewn up" into the second equation of this system, reflecting the drivers' desire to move at the desired speed.

In 1975 [217], Stefano Paveri-Fontana proposed an improvement to Ilya Prigogine's kinetic equation. Both the original equation and the Paveri-Fontana equation considered cars as point objects, neglecting the size of the cars. The modification of these equations, taking into account the non-point dimensions of cars, was made later (in 1996) in the works of Dirk Helbing. Today, kinetic models are primarily used not to simulate the movement of vehicles on roads but to study crowd behavior, such as simulating entrances and exits from venues like stadiums. For instance, during the construction of facilities for the Winter Olympic Games in Sochi, the Helbing model was applied. Currently, there are also models that are intermediate between hydrodynamic and kinetic, such as a so-called mesoscopic model used by the team of Boris Chetverushkin at the Institute of Applied Mathematics of the Russian Academy of Sciences.

Back in the 1950s, John Von Neumann introduced the concept of cellular automata [282]. In 1992, the Nagel-Schreckenberg [193] model was formulated, marking the beginning of the active use of the theory of cellular automata in transport modeling. These models, sometimes of little scientific value, are very convenient for computer calculations.

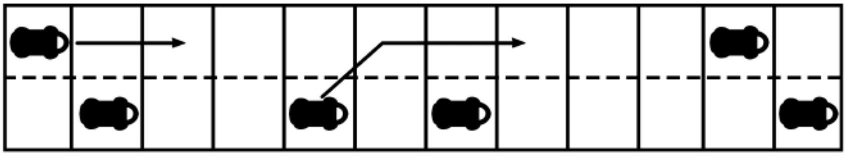


Figure 5.5: Cellular automata

In models of this type, both the road and time are divided into discrete intervals. Typically, it is assumed that no more than one vehicle can occupy a single cell at any given time. A key question in cellular automaton models for traffic is the definition of a cell. If the cell represents a fixed segment of the lane that a car covers within a specified time, this results in a dynamic system that serves as a discrete analogue to continuous models.

Since the advent of computers, numerous articles and books have been published on discrete micro-modeling of traffic flows. Initially, these projects were costly due to significant expenses associated with computing equipment and software development. This forced developers to break the task into subtasks and compute them separately. However, with the reduction in costs and improvements in computing power, more comprehensive and complex problems and models have been developed, including those incorporating various driver behavior strategies.

Subsequently, researchers shifted their focus to studying the dynamic behavior of traffic flows. It has been noted that many complexities in system dynamics arise from the interaction properties between vehicles in the models.

In the late 1980s and early 1990s, due to a significant increase in vehicle numbers and the resulting deterioration of the transport situation, the study of transportation systems in the United States was deemed a national security issue. The best "physical minds" and the computing resources of the Los Alamos National Research Laboratory were dedicated to this task. For instance, the nonprofit project TRANSIMS (Transport Simulation Analysis System) is an integrated set of tools designed to conduct regional analyses of transportation systems using cellular automata. This system introduces a new paradigm for modeling individual vehicles and their multimodal transportation based on synthetic populations and their activities. Unlike other traffic constellation models, TRANSIMS maintains a consistent and continuous representation of time. Its microsimulation-based, time-dependent routings also differ from other aggregate models. The core of this approach involves loading the road network with traffic and iterating until the Nash equilibrium is reached. The system includes submodules for population synthesis, activity generation, route planning, and microsimulation. The input data for the next iteration is the output data from the previous one. We will revisit this project in Section ??.

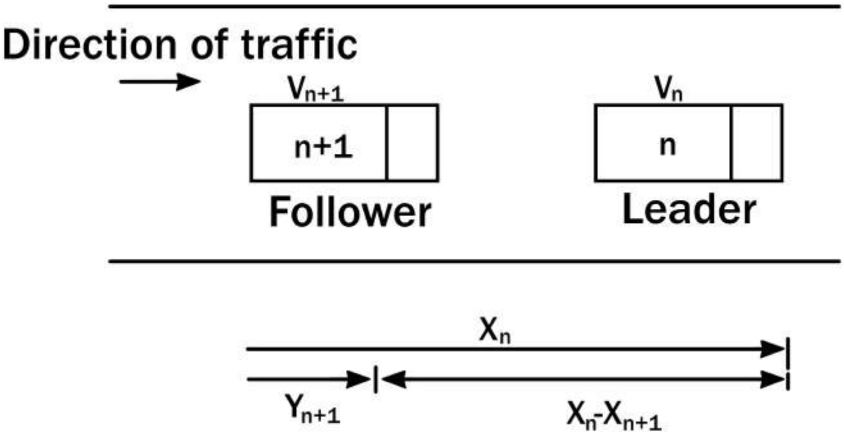


Figure 5.6: Notation for car following model

The modern microscopic approach is dominated by "intelligent driver" type models, in which a car's acceleration is described by a function of its speed, the distance to the car in front (the leader), and the relative speed to the leader (Martin Treiber, 1999). Additionally, in such models, time may flow discretely, and the dynamics of car movement can be stochastic (Markov). Typically, these models are referred to as cellular automata models.

5.2.2 Urban traffic flows

We have discussed macroscopic and microscopic models, but what exactly do these models address? What are the microscopic and macroscopic parameters of traffic flow?

Macroscopic parameters include:

- Average speed, measured in kilometers per hour;
- Density, indicating the number of vehicles in a 1 km long segment of road;
- Flow, defined as speed × density, representing the number of vehicles crossing a section of the road per hour.

Microscopic parameters include:

- Time headway, the time gap (in seconds) between successive vehicle arrivals;
- Space headway, the distance between two consecutive vehicles;
- Individual vehicle parameters: position, speed, acceleration, etc.

In light of the above, we note that the problem of mathematically rigorous justification of the kinetic model, based on microscopic data, as well as the problem of substantiating the macroscopic model based on the kinetic one, remains open. Furthermore, in modes corresponding to the "phase transition" in traffic flow, such justification seems fundamentally impossible: it is not feasible to perform the necessary scaling, transition to the dynamics of averages, or use the ergodicity of the system (as the invariant measure is not unique). Additionally, it is unclear how to terminate (close) the moment chain of linked equations. In such modes, one can only loosely speak about the similarity of models.

Transport phenomena are complex, nonlinear, and depend on the interactions among a large number of vehicles. Unlike rigid bodies that interact mechanically, vehicles react to individual drivers' actions, forming groups and exhibiting shock wave propagation effects.

Urban traffic flows (TF) exhibit several characteristics. Firstly, they are stochastic; the behavior of TF can only be predicted probabilistically. TF moves along a network, each with specific characteristics that allow for a relatively precise description.

Secondly, TFs are non-stationary, with fluctuations in characteristics occurring in daily, weekly, and seasonal cycles.

Thirdly, TF is characterized by incomplete controllability. Even with complete information about the flows and the ability to inform drivers about necessary actions, these instructions are merely advisory. Thus, achieving the global optimum of any control criterion is highly problematic.

Fourthly, TF is defined by multiple quality criteria, such as travel delay, average speed, predicted number of accidents, and the amount of harmful emissions released into the atmosphere. Most of these characteristics are interrelated, making it impossible to isolate any single one.

The fifth characteristic of road traffic as a control object is the complexity, and often impossibility, of measuring even the main characteristics that determine the quality of control. Thus, assessing traffic intensity requires either traffic flow sensors in all directions, aerial photography data, or laborious manual surveys.

Finally, it is fundamentally impossible to conduct full-scale experiments in traffic control. This limitation is due to the need to ensure traffic safety, the material and labor costs of the experiment (such as changing road markings and signs), and the fact that significant changes in the traffic organization scheme affect the interests of many participants.

These difficulties in formalizing traffic flow processes have significantly hindered the progress of scientific research to meet the practical demands.

To devise effective strategies for managing traffic flows in a metropolis, and to design optimal road networks and traffic organizations, it is essential to consider a broad range of traffic flow characteristics and the patterns of influence from external and internal factors on the dynamic characteristics of a mixed traffic flow.

5.2.3 What is the fundamental diagram?

The classical paradigm in traffic flow theory, mathematically defined as a functional dependence of the traffic flow's speed on its density (the dependence of the behavioral dynamics of a traffic flow on the nature of its properties), remains a key heuristic principle for developing theoretical models. An example is the modern three-phase theory of Kerner.

Developed by Boris Kerner between 1996 and 2002, this theory primarily explains the physics of the transition from free to dense traffic (traffic breakdown) and the resulting spatio-temporal structures in dense traffic on highways. Kerner describes three phases of the traffic flow, while classical theories based on the fundamental diagram consider two phases: free flow and the so-called dense flow (congested traffic). According to Kerner, two phases are distinguished within a dense flow: a synchronized flow and a wide moving cluster of cars (wide moving jam). Thus, there are three phases of the traffic stream, each defined as a specific state of the traffic stream, considered in both space and time.

- Free Flow F: In a low-density traffic flow, drivers can almost freely set their desired speed. There is a positive correlation between the volume of traffic, measured as the number of cars passing through a specific point per unit of time, and the density of traffic, measured as the number of cars per unit length of the road.
- Wide Moving Jam J: This denotes the rearward front of a broad moving cluster. Here, cars exiting the cluster accelerate to match the speed of free or synchronized traffic, which maintains a consistent average speed while moving upstream, passing through all bottlenecks on the freeway.
- Synchronized Flow S: This refers to the rearward front of the synchronized flow, where cars accelerate to match the speed of the free flow. Unlike a wide moving jam, synchronized flow doesn't share its characteristic properties. Specifically, the trailing edge of a synchronized flow

often remains stationary near a freeway bottleneck.

Spontaneous formation of a dense flow, or a spontaneous phase transition, can occur over a broad spectrum of flow values in free traffic flow. Based on empirical measurement data, Kerner concluded that there exists an infinite number of free-flow highway capacities. These infinite throughput values range between the minimum and maximum throughput values.

Kerner's theory has received substantial criticism and is often characterized as "empirical" rather than "scientific", as it is derived from observations without evidence supporting its generality. Nevertheless, the most significant aspect of Kerner's theory is its explanation of the empirical nucleation nature of traffic breakdown at road bottlenecks via the $F \to S$ transition. This empirical nucleation nature of traffic breakdown cannot be explained by earlier traffic flow theories, including the two-phase traffic flow models studied in [149].

The fundamental diagram is also utilized in the work of the Californian school led by Pravin Varaiya and Alexander Kurzhanski [280]. Using a cellular automata model, they propose a method for optimal control of traffic lights and highway entrances in California. One of the Californian team's objectives was to optimize traffic on high-speed highways, achieved as follows: access to the highways was restricted, reducing the "bottleneck" effect that occurs when lanes are occupied at the entrance. Surprisingly, this led to the highways operating more optimally, and drivers spent time only at the exits, significantly reducing their time costs.

It follows from the fundamental diagram in Figure **??** that different (usually two) densities, and consequently different speeds, correspond to the same value of car flow. Clearly, the more advantageous mode is the one with the higher speed. The task of controlling flow through traffic lights or highway entrances is to ensure that the average driver spends most of their time in modes corresponding to higher speeds.

5.2.4 What are the principles for traffic modelling?

Analysis of the foundations of traffic flow theory and its comparison with practical requirements indicate the need to apply new principles in developing transport modeling theories. Carlos Daganzo's first attempt to deviate from the clas-

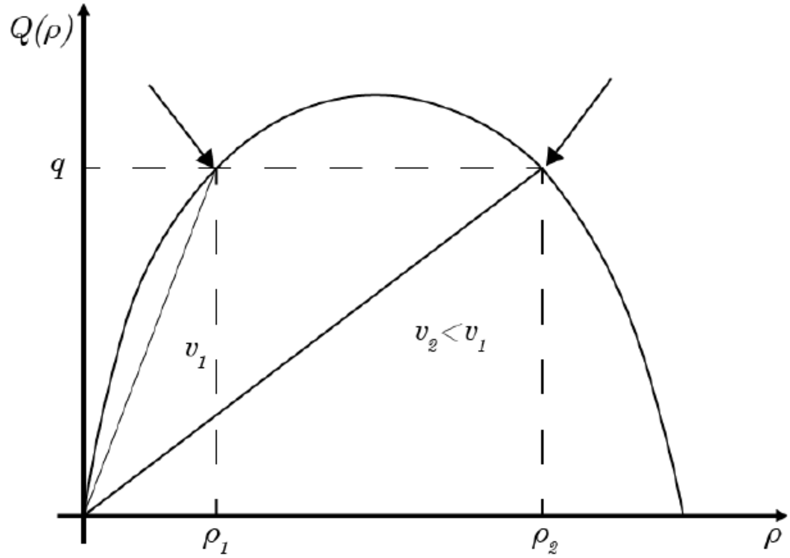


Figure 5.7: Fundamental diagram: the left state is faster then the right under the same flow

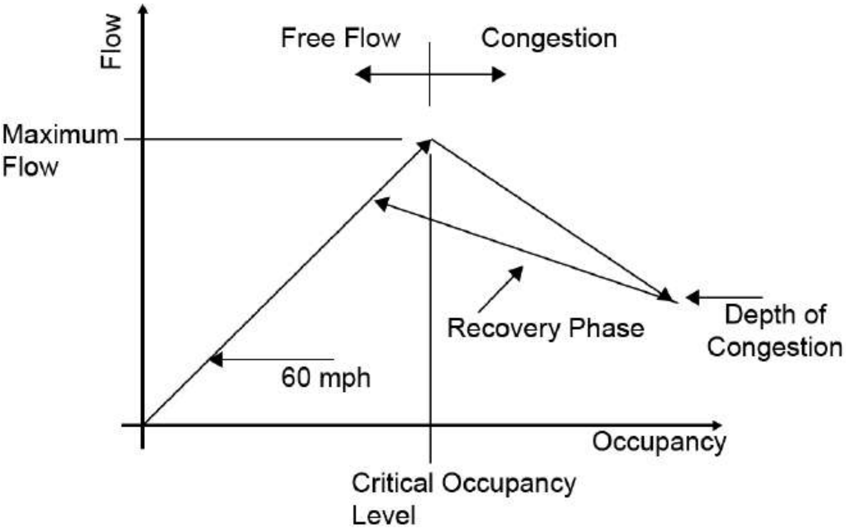


Figure 5.8: Model of congestion. If occupancy is maintained below critical level, section operates at 100% efficiency and the speed is at 60 mph

sical approach involved considering the impact of road narrowings as a basis for studying traffic dynamics and selecting a vehicle queue as the elementary unit for modeling dense traffic flows. However, this approach has not yet yielded significant results. What do we need to move forward?

- A more in-depth analysis of transport modeling fundamentals, which will enable us to identify key features of the classical paradigm and compare them with those of the new paradigm. For example, this might involve shifting the emphasis in road congestion norms from empty roads to overcrowded roads.
- The development of adequate mathematical models to describe and predict modern empirical observations of local traffic flow compactions at highway entrances and exits that disrupt normal traffic patterns.
- Research and modeling of phase transitions, including both direct and reverse transitions associated with the hysteresis effect, exploring the physical mechanisms of this phenomenon, and further developing Daganzo's ideas on the role of bottlenecks.

The features and limits of applicability of the following applied transport models are revealed:

- Equilibrium models assume constant density and flow speed over time. These models distribute vehicles across various available routes to a destination. Participants in the flow are categorized according to the cost functions corresponding to each edge in the network, with the goal of minimizing the total price functional. The challenge here lies in selecting adequate cost functions.
- Hydrodynamic models characterize vehicles by their type of reaction to surrounding traffic: instantaneous, elastic, or with relaxation; or by a model of variable demand for movement. The main focus of these models is the origin-destination matrix, which varies depending on the time of day, day of the week, and seasonality. The minimizing functional in these models includes an unknown demand function.
- Optimization models determine the optimal speed for a single-lane flow based on flow density. These include optimal speed models and leader-following models, which are dynamic equilibrium models that integrate traffic demand functions with a microscopic view of vehicle movement along individual trajectories.
- Kinetic models employ various kinetic equations for the distribution function of vehicles in terms of speed and position coordinates on the track. An essential feature of these models is their stochastic nature: the demand function is derived from the distribution of cost functions over their arguments, which are considered random.

5.3 Multistage traffic flow models

The first comprehensive study of the urban transport system was conducted by the United States in Detroit in 1953, followed by Chicago in 1956. Since then, the general research approach has remained fundamentally unchanged, focusing on an aggregate four-stage computerized transport model.

5.3.1 Main history turns

There are 4 historical periods in the development of transport models:

1. Development in the 1950s—1960s occurred in response to the acceleration of highway construction and advancements in computer technology.
2. Development in the 1970s—1980s was spurred by criticism of complex methods.
3. The 1980s—1990s saw development in response to criticism of static route analysis methods (trip-based analysis).
4. In the 1990s, development was driven by concerns about environmental pollution and the implementation of transport demand management policies (travel demand management).

Prior to the 1950s, travel meter data was employed to analyze trips. However, this approach was only sufficient when considering daily periods, and any forecasts were coarse, relying on trend analysis. With the acceleration of road construction in the 1950s, there emerged a need for more sophisticated predictive tools and economic forecasts. Concurrently, the development of computers provided the means to process the vast amounts of data required to model the entire urban system. The developers of the initial models were primarily engineers with "positivist" perspectives, who utilized physical laws to plan urban systems. For instance, gravitational interaction was used to model the distribution of traffic flows across a network. The key technique developed during this period was the "complex four-stage model".

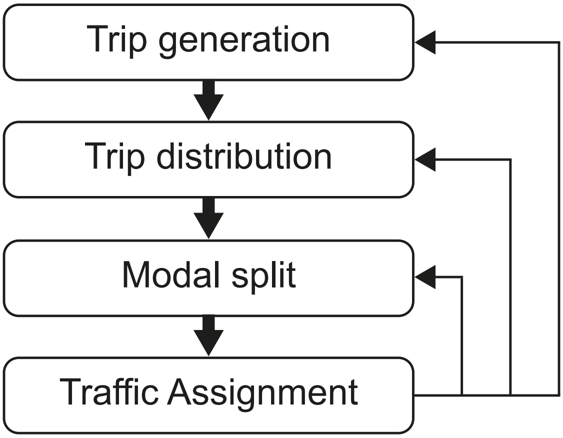


Figure 5.9: Four-stage transport model

5.3.2 The core of the four-stage transport model

The foundational assumptions of the four-stage model developers were as follows:

- Future land use patterns can be forecasted irrespective of changes in the transport system.
- Driver behavior can be predicted using household data collected from a specific area, such as an urban district.
- The correlation between household characteristics and driving behavior will persist over time.
- Drivers select routes with the goal of minimizing time and cost of travel.
- The interconnections between urban areas, average characteristics of a typical weekday, and peak hours, provide a sufficient representation for the purposes of enhancing the transport system.

In the 1970s, amidst political crises across the developed world, the conclusion of the post-war boom, fuel shortages, and high inflation in the UK led to civil unrest, encompassing transportation issues. During this era, a new research approach emerged, less reliant on computers and more inclusive of public participation, encompassing a broader range of evaluation criteria such as environmental and social justice issues. This was a shift from the previous development period when the focus was primarily on economic evaluations. However, despite some changes in the transport planning process, the computer models used essentially remained the same, retaining the complex four-stage model developed in earlier decades. In 1973, Lee published a highly critical paper [161], pointing out seven fundamental flaws: excessive complexity, computational coarseness, large amounts of input data, a mismatch between the theory used in models and the actual behavior of road users, difficulties in interpreting simulation results, the requirement to adjust the model to obtain realistic results, and high cost. In response to this criticism, three new analytical methods were developed:

1) Models depicting the mutual influence of transport and land use.
2) Downscaling methods, such as developing route selection models based on individual travel objectives instead of area zoning.
3) Micro-modeling methods that enhance the complex procedure of demand distribution, considering the behavior of the driver at the level of an individual vehicle or a small group of vehicles.

Despite improvements in modeling technology during the 1970s and early 1980s, complex four-stage models continued to be used in practice. Therefore, in 1986, Atkins issued new criticisms of the redundancy, inefficiency, and wastefulness of the main methods in transportation planning [17]. In response to these criticisms, the 1980s and 1990s saw the development of new modeling tools: **dynamic methods** and **activity-based methods**.

The activity-based approach replaces the observation of the trip with a detailed consideration of the activities leading up to the trip. The underlying hypothesis is that to understand (and influence) driving behavior, it is essential to comprehend the activities of a person using the transport system. Therefore, these methods focus on the activities of household members rather than on travel as an isolated event.

Dynamic methods were developed to counter the “static” approach of previously used methods. Trip analysis was based on single-section data at a single time point, but critics argued that traffic behavior and traffic flows change over time.

The development period in the 1990s coincided with fundamental changes in transport policy in the developed world, including the Clean Air Act Amendment (CAAA, 1990) and the Intermodal Surface Transportation Efficiency Act (ISTEA, 1991) in the United States. These Acts facilitated the expansion of the public transport system, the provision of dedicated lanes, the elimination of congestion, and the introduction of toll rates to counter pollution problems. They also allowed for flexibility in transferring funds between trunk investment funds and public transport investment funds. The CAAA and ISTEA challenged transport professionals by requiring forecasts of the amount of exhaust pollution in the atmosphere. To advance transport modeling, the US Government launched the Travel Model Improvement Program (TMIP). Although work is still ongoing, the TRANSIMS micro-simulation model has been developed in some detail, enabling the calculation of predictions regarding the distribution of exhaust gases in urban spaces.

5.3.3 Main bricks of modern models

Modern models typically comprise the following interrelated structural components:

- Models of transport demand volumes;
- Models of the structure of transport demand;
- Models of traffic flows on the road network.

Despite the variety in the mathematical tools employed, most modern transport models involve the execution of standard operations and calculations:

- Zoning of the study area;
- Creation of graphs for the road network and route networks of public transport;
- Determination of total volumes of transport demand by transport areas for various travel purposes;
- Modeling the structure of transport demand by travel destinations, time of day, and modes of transport;
- Modeling the distribution of traffic flows along the road network of the study area and calculating the parameters of these flows;
- Calculation of road transport performance indicators and evaluation of the effectiveness of potential control measures.

The zoning of the study area involves the following steps: The simulated territory is divided into transport areas, each considered both a generator and an absorber of transport and passenger flows. Zoning is based on ensuring that, if possible, each area has a uniform layout and development, and possesses a single functional specialization (e.g., residential, industrial, commercial, non-industrial areas, etc.). When defining the boundaries of these regions, it is advisable to consider natural and artificial barriers (such as rivers and railways), the existing administrative zoning of a city or region (since statistics are usually collected by administrative units), and to avoid elongated areas to simplify modeling and ensure more uniform transport accessibility within each area. An example of purely administrative division is the Paris district division shown in Figure **??**, while a purely graph-based division is depicted in Figure **??**.

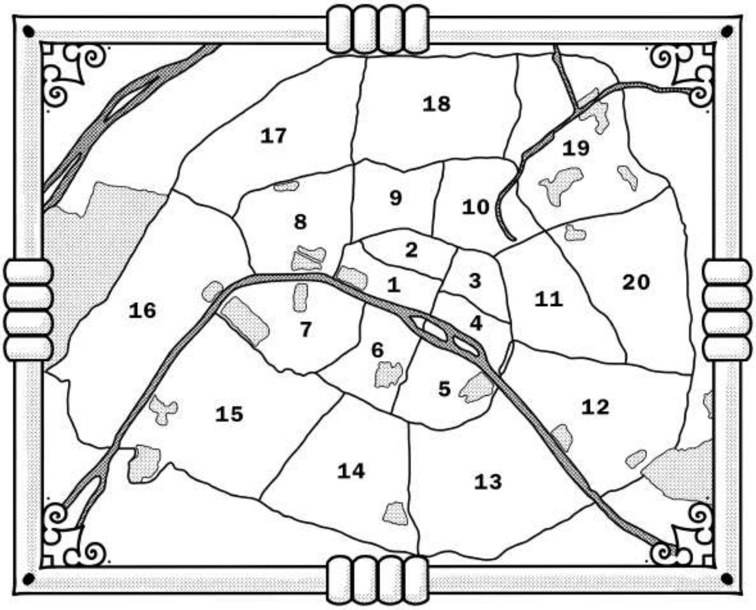


Figure 5.10: Arrondissements of Paris

Determining the total volumes of transport demand involves modeling the total number of trips between transport areas over a specified period. To construct the origin-destination matrix, various forms of the gravity model are typically employed. These models suggest that the number of trips between districts depends on factors such as population, the number of workplaces, and the presence of trade and service enterprises, which vary according to the purpose of the trip. Additionally, the measure of remoteness considered is usually not the physical distance but rather the average travel time or the value of the generalized costs of the trip, which includes both the monetary costs of a ticket or fuel and the estimated cost of travel time. At this stage, the time cost values

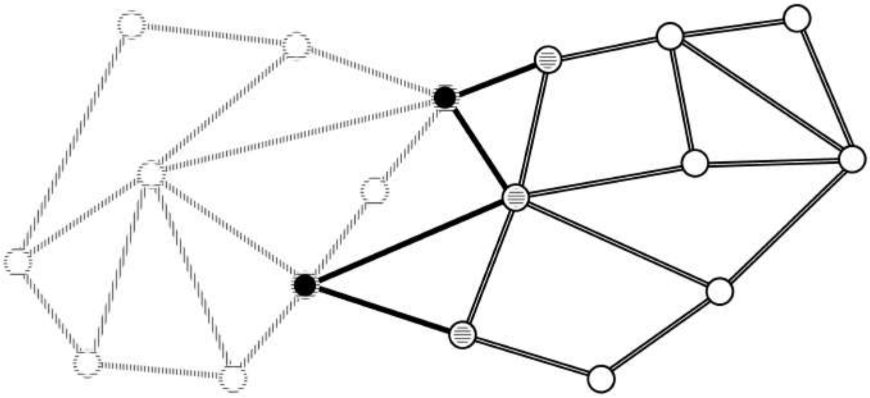


Figure 5.11: Graph zoning

derived from transport surveys or expert assessments are utilized. The simulation of travel time along routes occurs during the stage of calculating road transport performance and assessing the effectiveness of potential control measures. Based on this data, an origin-destination matrix is calculated for different travel purposes.

The **structure of the transport demand** among the population is shaped by the aggregate of individual decisions regarding the trip. These decisions include the choice of destination, departure time, mode of transport, and route. Besides general patterns of individual preferences and subjective factors, qualitative characteristics significantly influence the choice of travel method. For instance, public transport is commonly chosen for commuting to work, while private cars are preferred for social trips. During the journey, the choice of travel method may change due to situational factors or circumstances unrelated to the trip. Factors such as changes in parking costs, fuel prices, personal well-being, and weather conditions also influence this choice. The possibility of multimodal trips, where one journey involves multiple modes of transport including the use of intercept parking systems ("Park and Ride"), should also be considered.

Currently, empirical models are being supplanted by models based on probabilistic discrete choice — stochastic models that use the value of generalized costs as a criterion for selecting travel time and mode of transport. Based on these models, OD matrices are compiled for the studied periods and types of vehicles.

An important aspect of transport demand is its derivative nature. The primary purpose of transport is to facilitate access for the population to various activities and for goods to markets and raw materials to production sites. Thus, a trip does not satisfy needs by itself but merely creates the conditions for satisfying them. Consequently, it is more accurate to view a trip not as an independent benefit but as a resource expenditure necessary to acquire other benefits, thereby rendering it a negative utility.

This perspective allows us to establish two criteria for consumer choice in the transport services market:

- the benefit expected by the consumer from the activity accessed through the travel. This criterion typically characterizes the destination and time of departure.
- the total amount of all types of costs incurred by the consumer in relation to the trip. This criterion also characterizes the time of departure, the type of transport, and the specific route.

It is assumed that consumers always opt for alternatives that offer the highest ratio of expected benefits to total transport costs. Since the benefit expected from a trip depends on a broad range of factors unrelated to the transport system's operation, it is not explicitly present in the model but is included in the calculations as an aggregate indicator of a transport area's attractiveness to users. The consumer choice criterion is thus the reduced sum of all costs associated with the trip (the so-called generalized cost of the trip), composed of direct cash expenses, time costs, vehicle operation costs, and other expenses converted into a single dimension — typically time or monetary units.

The road network is modeled by a directed graph. Each edge of the graph is assigned attributes such as length, maximum speed, maximum throughput, width, and number of lanes in a given direction. The nodes describe the maximum throughput, traffic direction priorities, allowed turns, and traffic light control modes. Transport areas are defined on the road network graph as nodes connected by "dummy" edges, which are assigned a minimal length and substantial bandwidth.

The modeling of traffic and passenger flow parameters on public transport routes relies on data concerning the structure of transport demand, travel purposes, types of travel, and time periods. Initially, the volumes of transport demand are distributed over the shortest routes between areas on the road network graph. Subsequently, considering the limited capacity of nodes and edges, they are redistributed throughout the graph. The traffic flow parameters (speed, density, intensity) calculated for each road network section are used to determine the time spent on trips between transport areas.

When developing a mathematical model, finding a balance between the level of detail and computational resources is crucial. The level of detail is generally determined by numbers of:

- transport areas;
- groups of users within the transport system;
- purposes of travel;
- modes of transport;
- time periods;
- objects of road infrastructure.

The level of detail in modeling each of these aspects depends on the characteristics of the transport program being developed. For instance, to evaluate the effectiveness of large-scale and strategic measures (such as urban planning decisions), it may be beneficial to model population mobility as realistically as possible. In this scenario, the model would describe a wide range of socio-economic consumer groups and various travel purposes. However, to remain within the constraints of available computing power, the road network would be modeled with less detail. The model would consider only a few large transport areas, and the road network graph would depict only the main traffic flows.

It is now clear that balancing transport capacity and demand cannot be achieved solely by constructing new roads. Consequently, the focus has shifted towards managing transport demand and enhancing traffic organization. Transport demand management employs mechanisms beyond traditional modeling approaches, such as encouraging the redistribution of trips across different transport modes, including multimodal journeys and combinations of private and public transport, as well as car sharing for joint or alternating trips.

Incentives for time-sharing are implemented to distribute traffic more evenly throughout the day and minimize congestion during peak hours. Since transport demand is a derivative of main activities, redistributing trips over time necessitates modeling the primary activities of the population.

As the demand for more sophisticated transport system models increases, so too do the methods of mathematical modeling and the availability of computing resources. This progress enables the adoption of the most advanced mathematical modeling techniques in transport system software.

5.3.4 TRANSIMS structure

For example, studies utilizing the TRANSIMS package are based on a virtual city model that includes comprehensive information about residents, their daily transport activities, and detailed transport infrastructure. Residents plan trips to meet their mobility needs within the city. The movements of each resident in the simulated urban area across all types of transport systems are modeled on a per-second basis. The aim is to replicate realistic urban driving dynamics so that users can accurately estimate vehicle emissions and fuel consumption. Like the four-stage model, the TRANSIMS model operates sequentially (not in parallel), with explicit feedback providing insights into how individual traffic participants react to the fulfillment of their transport preferences.

TRANSIMS comprises five modules: population synthesizer, activity generator, route planner, microsimulator, and feedback block.

The first module, the population synthesizer, is used to create an artificial population for households in the study area. It aggregates census data (summarized by census area or block/group) and disaggregated data, utilizing public population data from the public use microdata samples (PUMS) census records. Synthetic households are generated from this aggregated census data. Household attributes identified a priori for analysis may include gender, age, education, employment, income, and vehicle type. The modeler can select from any of the available options.

The second module, the household activity generator, determines the "potential" daily activities of each synthetic individual within each synthetic household. Inputs for this module include the synthetic population, regional activity surveys, and transport system data (encompassing both land use and transport network). Transport activity studies should span at least 24 hours, be representative of a population sample, and encompass all activities and trips of each family member. The total number of trips for each traffic participant is determined by counting the location changes during their daily traffic activity. Essentially, the list of daily routes outlines the daily journey chains of each road user, akin

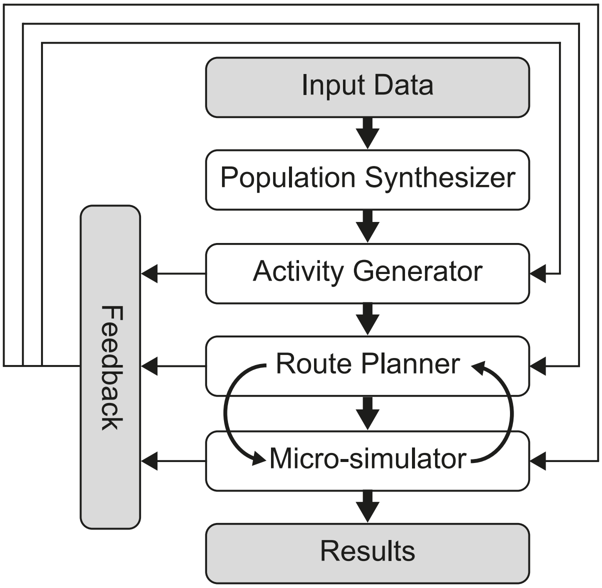


Figure 5.12: Basic TRANSIMS methodology

to the information in a traditional travel diary.

The third module, the route planner, assigns transport routes to each trip generated in the second module to meet the final goals of the traffic participants. It is important to note that the route attributes in TRANSIMS include not only road sections but also information about the mode, mode changes, parking locations, and data about fellow travelers.

The fourth module, the micro-simulator, uses route plans from the route planner as input and models transport networks at a microscopic level of detail, despite the low fidelity of most traffic models. This module models the interaction between demand (desired travel destinations of the synthetic population and travel destination locations) and resources (the transport system's capacity to meet this demand) for each mode throughout the entire simulation period. The dynamics of the transport system over the entire 24-hour period are modeled, requiring transport network inputs for each mode.

The feedback module is used to achieve an equilibrium distribution of traffic flows in the network and facilitates data exchange between the activity generator, route planner, and micro-simulator. The process of achieving equilibrium is iterative: if certain routes become unfeasible at the next step, the data is sent back to the activity generator, which then determines alternative options and redistributes them across the network using the route planner. Additionally, if the intermediate routes are not feasible within the specified time intervals, this information is also sent back, and the route generator seeks alternative travel routes. The output of the microsimulation module can be aggregated at any level, allowing information about each traffic member on a second-by-second basis or aggregating data simultaneously over space (from individual roads to the entire transport network) and time (from 1 second to a 24-hour period), depending on the modeler's desires.

After the simulation is completed, the output from the micro-simulation is used to calculate various performance parameters. Since the module is based on micro-simulation and each participant in the movement is modeled explicitly, the user has a lot of choice in which metrics to use.

5.4 Customer preferences

We have already discussed terms such as customer preferences and utility functions. The concept of a utility function is particularly important in modeling transport systems and in solving transport optimization problems.

A utility function is a mathematical tool used to represent consumer preferences across a set of valid alternatives. The numerical values of the function help order these alternatives according to the consumer's degree of preference. A higher value indicates a stronger preference.

Not every preference relation can be represented by a utility function. However, in economic models, such a function typically exists. For example, in a basic path selection problem, the utility function might be represented by travel time (more specifically, the negative value of this time). However, some researchers prefer to express everything in monetary terms.

5.4.1 Time is money

People in developed countries spend many hours of their lives traveling to and from work, and waiting for taxis and trains. Thus, it would be beneficial to reduce this time. The time of a consumer becomes a commodity to purchase. Consumers follow microeconomic rules to choose their transportation method, incorporating travel and waiting times into the total cost of consuming a good. This fact is often overlooked, but consumers would prefer to spend as little time in transportation as possible.

In most economic models, goods are categorized into leisure, which adds utility when consumed, and work, which is the source of income to purchase goods to increase utility. There is a trade-off between leisure and work: the more time spent working, the less time is available for leisure. It is crucial for each individual to determine how to allocate their available time between leisure and work. In transport investments, the concept of "time value" (*value of time, VOT*) is widely used. Although each action has its own value for an individual, modeling a group's behavior to determine estimates presents challenges: how to obtain sources of individual distributions? These can be acquired through population surveys (method of *declared preferences*), as well as econometric models (method of *revealed preferences, RP*). Neither method is perfect, and the results vary from one study to another.

Becker's [22] 1965 model is classical, and time optimization is its main goal. He equates the VOT to the individual wage rate. Lisco (1968) [22] considers the value of transit time to be half of the individual wage rate. For Watters (1992) [287], this value ranges from 30 to 60 percent of the wage rate.

Johnson's (1966) [144] utility function, which uses work time, results in leisure time differing from that of other authors. Oort (1969) [204] proposes including transportation time in his model, where working time increases as travel time decreases. He considers travel time as a type of unpleasant activity that negatively affects the utility function. Mohring et al. (1987) [189] introduce passenger disutility as a linear function of time spent walking to and from bus stops, waiting for service, and time spent actually traveling. For de Serpa (1971) [58], the utility function depends on activities and goods. Evans (1972) [80] considers a utility function as a function with arguments that are the times spent in activities.

In modern articles, Diaz [142] observes that time is now considered the main component of utility, and all activities lead to changes in utility value. Thaithatkul et al. [271] describe a ride-sharing system considering a linear utility function depending on waiting time and travel time. Lenoir and Laplace [163] consider the price of traveling time and the activities and conditions that can affect it. Schakenbos et al. [252] study transfers in public transport, showing that the disutility during an interchange depends on travel time, waiting time, and headway.

Results obtained using RP data do not allow for accurate calculation of general trends in people's preferences due to the low accuracy of the data collected and strong variation. Using *Stated Choice (SC)* data to determine the cost of time has been the standard since the mid-1980s. This type of survey processes more data and also obtains preference patterns not only in waiting or travel times but also in the type of transport used. Since the mid-1990s, algorithms have appeared that allow for correction and detection of correlations in responses from each person, which reduced some of the advantages of using the SC data method, but it still remains the preferred method.

One limitation of SC analysis is the complexity and sometimes impossibility of estimating the value of small time savings. Critics argue that the treatment of small time changes estimated in SC studies calls for reconsideration of the use of RP data.

Studies showing the VOT as primarily based on an economic analysis of utility functions, *the value of travel time changes (VTTC)* (the monetary measure of the value that people attach to changes in travel time), also consider the psychological aspect of choice preferences. VTTC is a key factor in evaluating and comparing different transportation projects. Saving time on the road often constitutes a large part of the economic benefits of a project. Thus, overcoming a given constraint is crucial for cost-benefit analysis. Several countries are conducting national studies to evaluate official VTTC, which can be used to assess and prioritize transport development projects. In the design of most SC experiments, citizens choose between possible scenarios evaluating a time-cost trade-off. Such methods have been used, for example, in the Netherlands, Norway, the UK, Denmark, and Sweden.

To ensure studies are comparable, relevant normalization must be provided. The corresponding values are provided by such authorities as the "Commissariat général du plan" in France or the Department for Transport in the UK. Algers et al. [10] describe the first Swedish national VOT study, which followed methods used in the UK and Netherlands studies.

5.4.2 How time can be unlinear?

Many of you are likely familiar with the experience of waiting for transportation. How much more frustrating would it be if Uber changed the predicted pick-up time from one minute to two minutes? From 9 to 10 minutes? Now consider this scenario: you request an Uber, see that the estimated time of arrival is 10 minutes, and, expecting a long wait, decide to smoke your pipe. But unexpectedly, the car arrives much sooner than anticipated. Like many aspects of our lives, things are rarely straightforward. Often, it is our subjective perceptions that influence our decisions.

Research conducted in the mid-1990s confirmed and quantified non-linearities in the value of time (VOT) and revealed the following:

- The loss of time is perceived as more significant than its acquisition;
- The impact of time or cost savings is relative to the total time or money

spent on a trip;

- Smaller increments of time have a reduced value of travel time cost (VTTC);
- There may be preferences for current time departures.

Further studies with a more complex analysis of responses about stated choise (SC) in general corroborated these findings. However, the real question is whether the data obtained correlate with reality. If they do, how should we account for them? If not, what experiments should be conducted to ensure they align with reality? How should we incorporate these effects when evaluating a study?

Hultkranz and Mortazavi [132] used Swedish data to reveal non-linearities of VOT. Börjesson and Eliasson [28], based on data collected through two surveys in 2007 and 2008, clearly demonstrated the non-linearity of the VTTS, noting that commuting/school trips have about 30% higher VTTS than other private trips. Yang et al. [302] considered the Cobb–Douglas type production function for taxi-customer interactions, taking into account walking time and taxi waiting time.

In the past decade, models based on error defined in the log time and log cost difference scale, with the value of time distributed randomly in the population, have emerged. These models are more challenging to understand and implement in simulations, but currently, they most comprehensively explain all observed effects. Multinomial logit models and more recent mixed logit models are also widely used to evaluate VTTC. There is still no clarity in the literature regarding the definition and classification of the main modeling approaches used for the data sets described above. Generally, methods of parametric and non-parametric estimation are distinguished. More informative parametric models allow the VTTC to vary depending on the covariates, which is significant given the detected non-linearities. Based on national VTTC studies, two primary approaches can be identified. These parametric approaches, named Random Utility (RU) and Random Valuation (RV), differ in where the random component is applied. In RU, it pertains to the difference between the utilities of the trip options. In RV, it relates to the difference between the proposed assessment threshold and the actual value of the cost of travel time, implied in changes in time and proposed cost. From a deterministic perspective, these approaches can be considered equivalent and both can be derived from standard microeconomic theory. The significance of the differences between them has yet to be clarified in contemporary literature, and the approach is usually chosen based on empirical data or the preferences of the researchers.

5.4.3 And nothing else matters?

However, even if we use only time as a measure of utility, why do people use different modes of transport and how do they choose among them?

Transport plays a key role in territorial development. Urban passenger transport is an integral part of the city's unified transport system. Within city limits, urban passenger transport fulfills several main functions: delivering passengers with minimal time costs, providing information and service services, and facilitating movement for socially vulnerable categories of citizens.

The ratio of forms of movement is a clear indicator of the development of the urban transport system, as it reflects the total load of transport on the environment. *"Competitive and attractive public transport prevents the increase in transport energy consumption, emissions, and noise. It also reduces traffic jams and facilitates the movement of different categories of the population"* [75].

Most studies on travel mode choice are conducted in developed or high-income countries, with fewer studies focusing on developing countries. Additionally, limited research addresses the factors that affect a commuter's travel mode choice. Factors such as connectivity or reach, accessibility to a specific mode of transportation, information, time satisfaction, user attendance, comfort, security and safety, and environmental impact have been identified as influential in the choice of a commuter.

Fu and Juan [100] conducted a study on psychosocial factors influencing public transportation usage in several Chinese cities; it also revealed that differences in satisfaction were observed to be prominent between genders. Neely [196] also found that factors such as sociodemographic, built environment, latent attitudes, and trip characters influenced the choice of transport mode for long-distance, intercity travel from New England to New York in the United States, with age being a less significant factor for business trips. Reliability, privacy, comfort, availability, safety, and attitudes towards public transportation were key factors in transportation preference of commuters in the Chicago area in the United States. Market segmentation, such as age, household income, and vehicles per household, affected ferryboat commutes; it was also found that sensitivity to travel costs was the same across all market segments. Those who preferred automobiles emphasized stress-free travel that is non-work related. Income group was also a factor — lower income groups were inclined to choose public transportation compared to those in a higher income-earning group.

Gender also plays a role in transport mode choice. Thai, Malaysian, and Indian females were more inclined to prefer public transportation compared to men. Conversely, Indonesian, Japanese, and Taiwanese men were more likely to prefer driving themselves.

Regarding age, Irish households with more adults preferred to use public transportation, such as taxis or buses. Elderly Malaysians also favored public transport more than young people, provided that the service is improved. Younger Japanese people, on the other hand, are more mobile and mostly preferred to use their own vehicles. In Libya, households with more adults have a positive correlation with the household's income and car ownership.

Availability and access to public transport were also key factors for Thai people, and the distance to an available mass transit was considered important. Availability of parking space also influenced car usage. A study conducted by Diaz [61] showed that for inter-island travel between Manila and Busuanga Island in the Philippines, high-income earners were not particular about the fare difference in transportation. The perception of Indonesian commuters also impacts preference. Ride comfort and convenience were the major factors affecting the preference of Malaysians and Taiwanese. A study by Okamura [203] et al. in Manila revealed that Filipinos consider the Jeepney acceptable for short trips since it is cheap and they are familiar with it, but avoid it for longer trips because it is noisy and dangerous. These perceptions were obtained based on passengers' lifestyles.

A person's behavioral intention may also affect transport mode choice. The socio-demographic profile of commuters also plays a role in a commuter's preference, such as age, gender, personal income, or occupation. Although some factors are common among various demographic profiles (such as in Asian or developing countries), the intentional behavior towards preference in transportation still varies. Level of service, accessibility, and supply of public transportation can influence preference as well.

There are different sources of available statistics. The Filosofi (Fichier Localisé Social et Fiscal) data set collects income data of tax-registered people in France.

It is published as open data three years after the acquisition of the data. The most recent data set has been published in 2022 and therefore contains income information of the population in 2019. Specifically, the data set provides the centiles of the distribution of declared income and disposable income. While the former mainly considers gross salaries, the latter takes into account deductions due to taxes, social security, state insurances, as well as social benefits. The distributions are given on the level of regions and municipalities, but one year later (i.e., four years after acquiring the data), a more fine-grained data set is published that provides distributions on the level of IRIS. The latter data set, however, does not provide further sociodemographic information, while the municipality-based version also offers income distributions by household size and a few other household-level sociodemographic attributes. The Enquête globale de transport (EGT) is a *household travel survey* (HTS) conducted in the Île-de-France region, mainly during the year 2010. The survey has the classical structure of a household travel survey: Each member of a household is asked about their activities and travels during one particular reference day. Such surveys make it possible to estimate models of daily travel patterns, including the type of activities, the mode of transport for their connecting trips, and more.

The choice of traffic also is influenced by and influences the time of the trip. Figure **??** provides the input data to the secondary location assignment algorithm by Horl and Axhausen (2020) [125]. For each mode, all trips from the regional HTS are divided into travel time bins such that each bin contains at least 200 observations. Inside of the brackets, the resulting number of bins is shown. The mean for the travel-time-binned distance distributions is shown. For "car driver" and "public transport", the shaded area shows the 90% confidence bounds of the respective travel-time-binned distributions.

5.4.4 Will I take bus or train?..

The urban passenger transport system in large cities involves the interaction of various types of transport. This necessity arises from the inability of a single mode of transport to meet all the mobility needs of the population, given the distinct technical and economic characteristics of each mode. Consequently, it is essential to develop an urban passenger transport system where all modes — metro, buses, trams, high-speed trams, trolleybuses, and suburban railways — interact as efficiently as possible to create a unified movement space.

The primary challenge in creating such an integrated system is the presence

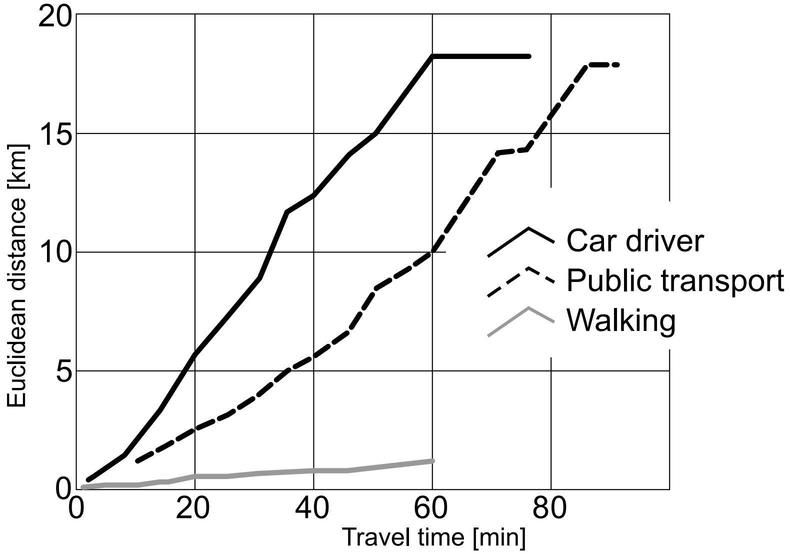


Figure 5.13: Correlation between chosen traffic mode, travel time and distance

of numerous independent carriers and transport companies, each owned by different entities. Vukan Vuchik [283] notes that this situation "*... led to the fact that each trip by several modes of transport (or several routes) turned into a loss of time for the passenger due to inconsistent schedules, as well as the need to re-pay for travel*".

In such a disintegrated public transport system, each component operates independently, aiming solely to transport the maximum number of passengers at minimal costs. This approach fails to consider the capabilities and advantages of each transport type from a unified perspective, resulting in a lack of integration and emergence.

Multi-criteria analysis is employed for research in transport and logistics. Decision-making commonly utilizes methods such as AHP PROMETHEE. The application of multi-criteria analysis in transport research primarily focuses on the following areas:

- Selection of a transport project;
- Selection of a route in a transport network;
- Choice of transport mode;

- Study of carriage operations;
- Assessment of environmental impacts.

According to Avishai Ceder's 2016 book [42], three types of demand variations related to attributes can be identified as follows.

Service dependency

Transit demand is the output of the transportation-demand analysis process. The model-split procedure determines the share of transit demand, out of the total demand in a city, by using the utility of transportation modes. The utility of a transit mode is derived from current transit system attributes, such as travel time, number of transfers, and convenience. It is evident that transit network design would result in a new set of utilities and consequently a new figure for transit demand. This variation of transit demand is termed service dependency. Figure **??** illustrates the concept of service dependency in relation to the bus network design problem (BNDP).

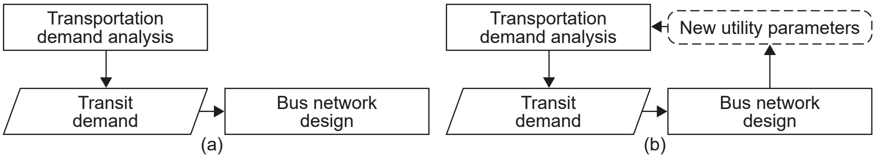


Figure 5.14: Service dependency (fixed vs. elastic demand) related to the BNDP. (a) Fixed demand consideration for bus network design and (b) elastic demand consideration for bus network design.

Temporal variation

Passenger demand fluctuates throughout the days of the week and the weeks of a month, influenced by people's activities and seasonal needs. These temporal variations can shape a continuous demand function characterized by peak and off-peak periods. An important aspect of these variations is seasonal demand, which follows an annual pattern. For example, summer tourism often leads to a significant increase in transit demand. Generally, in many cities, recreational trips increase during summer, whereas educational and work trips decrease, resulting in a completely different demand pattern compared to other seasons. This multi-period demand consideration was studied by Chang and Schonfeld [46], who developed an analytic model for optimizing bus service, taking into account the time dependency and elasticity of demand characteristics. Their model, however, was designed specifically for feeder bus systems and is not applicable to entire networks. Although their approach marked a progression in this area, the simplifications in network description and decision variables limit its use in real-time transit networks. For instance, dividing a city into triangular zones and optimizing line spacing for a feeder network is not a realistic assumption, particularly in cities with irregular networks.

Randomness

Random demand variation is another aspect of changes in transit demand. This was explored by Yan et al. [300] in their study of intercity bus routing and scheduling. However, random fluctuations in demand are minor compared to the substantial changes based on time of day. For instance, the demand fluctuations between two days during peak (or off-peak) hours are significantly smaller than the differences between the observed peak and off-peak demands within a single day. Therefore, due to the complexity involved in considering demand distribution functions, this type of variation is typically deemed insignificant in the analysis of network (routes) design problems.

A realistic approach to transit-demand characteristics can make the problem more closely align with reality. Numerous cities worldwide experience demand variations due to tourist arrivals and other related activities. These variations are currently the primary reason for changes in the schedules of existing transit services. There is a clear need to investigate how adjustments in transit routing can help achieve a better network-design solution when taking into account demand variations.

Another aspect of customer behavior is their readiness for ride-sharing.

5.5 Let’s share our ride

The introduction of ride-sharing is expected to not only diminish traffic jams and decrease passengers’ travel times but also reduce total vehicle run times and minimize extra stops along the way.

Furuhata et al. [101] propose a classification of existing ride-sharing systems into six classes based on various criteria:

- **Dynamic real-time ride-sharing** matches drivers and passengers in real-time (en-route) and includes strategies for routing, scheduling, and pricing.
- **Carpooling** involves more predictable matchings, such as regular trips with similar origin-destination (OD) pairs.
- **Long-distance ride-match** offers more flexible departure times and OD pairs but is comparable in trip length.
- **One-shot ride-match** is a hybrid of carpooling and long-distance ride-match.
- **Bulletin-board** is a board-based ride-sharing system where conditions are negotiated among participants.
- **Flexible carpooling** iis a semi-organized service that typically does not use a matching agency and is based on a first-come, first-served basis at designated meeting points.

Ride-sharing services face three major challenges: designing attractive mechanisms (pricing, incentives), arranging rides properly (using user profile information, multi-hop rides), and building trust among travelers in online systems.

Various challenges have emerged with the advent of dynamic ride-sharing, including economic, social/behavioral, institutional, and technological challenges.

The psychological and environmental challenges are widely discussed in the literature. Merat, Madigan, and Nordhoff [184] focus on the social and psychological factors of the ride-sharing for automated vehicles and consider users' motivations to use the AV and the ride-sharing depending on various factors. Epperson [78] examines methods, including pricing policies, to attract people to ride-sharing systems to reach a critical mass. Levofsky and Greenberg [165] study the reasons for the success or failure of ride-sharing systems that existed 15 years ago. Deakin, Frick, and Shively [56] report that most people prefer anonymous locations, such as parking lots and major intersections, over disclosing their precise start and destination points. Chan and Shaheen [45] focus on the historical and psychological aspects of ride-sharing primarily in North America. Murray and Steele [191] propose a system and method for facilitating and encouraging ride-sharing based on awards. Bistaffa et al. [26] concentrate on social ride-sharing, where groups are formed based on desired starting points, destinations, and time windows, considering social network constraints to reduce transportation costs. Kamar and Horvitz [146] address coordination challenges among self-interested individuals aimed at minimizing transportation costs and environmental impact. They propose two ride-sharing optimizations: (1) generating ride-share plans for groups of agents, and (2) clustering agents into ride-share groups, where a ride-share plan is defined as a single trip or a chain of trips near the origins and destinations of other trips. Jacobson and King [141] focus on the fuel consumption model in ride-sharing systems, noting that the monetary value of fuel savings to passengers is at least as significant as the value of time. Caulfield [38] examines ride-sharing patterns in Dublin and estimates the environmental benefits in terms of emission reductions and total vehicle kilometers traveled. Andréasson [14] provides an overview of operational issues in ride-sharing, including implications for passenger safety and security. Adler, Miculescu, and Karaman [6] consider a class of problems applicable to vehicle platooning and passenger ride-sharing in autonomous systems, noting that sending more vehicles at once allows for more energy-efficient transportation but requires some vehicles to wait for others to arrive.

The technological challenges, especially those related to ride-sharing matching methods, are closely linked with mathematical and computational methods. Modern literature includes a variety of articles, patents, and methodologies that highlight these challenges.

Tom LeWinson and Sharon LeWinson [166] describe ride-sharing applications in their patent, detailing specific logic for various scenarios. The first method involves searching for users whose starting points are within a specified distance from the seeker's route. The second method involves searching for users whose destinations are within a specified distance from the seeker's route. The third method combines the first two methods.

Maciejewski and Nagel [177] incorporate multiple pick-up and drop-off locations in their approach. However, the comprehensive nature of their proposed simulation limits its scalability. Kornhauser et al. [154] employ pixelization of the city to facilitate ride-sharing, basing matches on waiting times at departure points to consolidate more passengers into a single vehicle.

Fagnant and Kockelman [82] focus on determining the optimal fleet size for shared autonomous vehicle (SAV) services, under the following conditions:

1. Current passengers' trip duration may increase by no more than 20%.
2. Current passengers' remaining trip time may increase by no more than

40%.

(3) New traveler's total trip time may increase by no more than the greater of 20% of the total trip time without ride-sharing or 3 minutes.

(4) New travelers will be picked up within the next 5 minutes.

(5) Total planned trip time to serve all passengers must not exceed the sum of the remaining time for current trips, the time for the new trip, and an additional 1 minute for drop-off if not pooled.

The search process prioritizes trips with the earliest final drop-off time.

The methods of ride-sharing are widely used, for example, of such modern systems, as Uber Pool etc. Such systems widely use the modern informatics, mathematics and economic techniques.

Ride-sharing methods are extensively utilized in modern systems such as Uber Pool, employing advanced techniques in informatics, mathematics, and economics. Recent articles on ride-sharing discuss various methods of combinatorial optimization, artificial intelligence, Tabu search heuristics, machine learning techniques, clustering, and queuing theory.

Based on the state of the art [101], we categorize ride-sharing based on trajectory shapes:

1 Identical ride-sharing. Both the origin and destination of current trip *a* and new passenger *b* are identical.

Figure 5.15: Identical ride-sharing

2 Inclusive ride-sharing. Both the origin *B* and destination *B'* of new passenger *b* are on the way of the original route $A \rightarrow A'$ of a vehicle *a*.

Figure 5.16: Inclusive ride-sharing

3 Partial ride-sharing. The origin *B* of new passenger *b* is on the original route $A \to A'$, but the destination *B'* is not. Ride-sharing covers only part of passenger *b*'s trip.

Figure 5.17: Partial ride-sharing

4 Detour ride-sharing. Neither the pick-up nor drop-off locations of passenger *b* are on the original route of vehicle *a*.

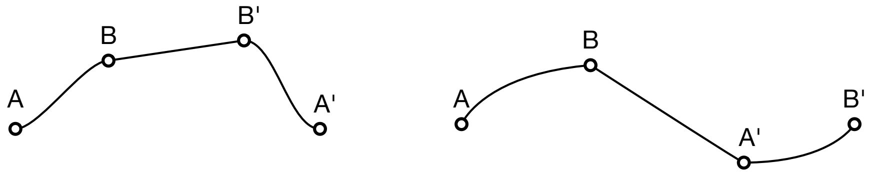


Figure 5.18: Detour ride-sharing

Detour ride-sharing, by taking a detour, covers both the origin and destination locations, accommodating the largest number of potential shared trips.

Why optimize if we can simply select the best passenger for our trip?

Let us provide an easy example of a gap between greedy and optimal solutions, as shown in Figure **??**. Let the vehicle be at point *O* with a planned trip from *O* to *D*. Suppose there are clients at stations 1, 2, and 3 who wish to travel to station *D*. Using a greedy method, the most suitable client will add the least detour to the vehicle's trajectory. After choosing such a client from station 1, no more clients can be added to the trajectory.

If an acceptable detour for ride-sharing is 10%, then the optimal ride-sharing strategy can be described as follows: Clients from stations 2 and 3 will be taken together on the same trip from *O* to *D*. This arrangement will bring about a reasonable delay within the stated constraints.

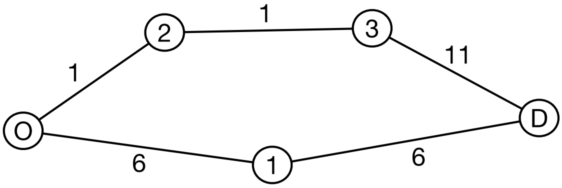


Figure 5.19: Ride-sharing. Case 1

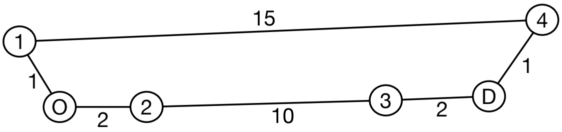


Figure 5.20: Ride-sharing. Case 2

This family of methods can also miss many potential sharing scenarios. For example, consider the current scheduled trajectory from *O* to *D*, as shown in Figure **??**. The client at Station 1 wants to go to Station 2, and their entire route coincides with the already scheduled one. However, this client will not be picked up because both their origin and destination are not within a 2-minute ride from *O* and *D*.

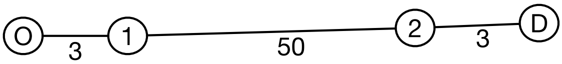


Figure 5.21: Ride-sharing. Case 3

5.6 Optimisation methods in transport

5.6.1 Disruption management systems

For many cities around the world, public transport serves as the backbone for commuters, making its reliability crucial for society. Especially in large cities, these systems transport millions of people daily, and any disruptions during peak hours on high-capacity lines such as rail or metro can significantly impact numerous travelers. In Stockholm, Sweden, metro and commuter train disruptions are managed in several ways: Bridging buses are deployed to replace the rail-based service along the disrupted lines. Additionally, travelers who can demonstrate that they would have been at least 20 minutes late using the best public transport alternative can receive reimbursement for their travel expenses in alternative modes such as taxis or private cars. The travelers must arrange these alternative modes of transport themselves. These reimbursements are costly for public transport operators, who are fined by the public transport authority (PTA) for any substantial delays. In turn, the PTA reimburses delayed travelers; in 2014, the administration paid out a total of SEK 9.7 million in reimbursements, which more than tripled to SEK 29.1 million by 2018.

Depending on the type of disruption, various techniques are employed to restore service and minimize negative impacts for travelers. Pender et al. [221] conducted a survey among 71 international public transport agencies regarding their strategies for managing unplanned service disruptions. The responses indicate that approaches vary depending on whether the disruption is a train failure or a line blockage. For line blockages, bus bridging is the prevailing strategy (85% of respondents). However, different approaches are often used depending on the location, time, duration, and type of event. These include single tracking/bypassing (51%), rerouting trains (23%), diverting to other lines (35%), or using parallel public transport (47%), and improving the frequency of parallel public transport (17%). Hiring taxis is employed by only a small fraction of respondents (5%).

One drawback of the bus bridging approach is that the extra buses and drivers must either be kept on stand-by or moved from their original lines. Additionally, it often takes a significant amount of time for the buses to arrive, and their normal schedules need to be revised. Consequently, passengers often experience substantial waiting times, which fosters a negative attitude towards bus bridging.

Various forms of collaboration between public transport and taxi services are being explored worldwide. According to Ibrahim [134], about half of the taxi industry's revenue in Sweden comes from contract work and about 60% in the Netherlands. This contract work involves services for passengers who have difficulty using regular public transport. Some cities, such as Singapore, have included demand-responsive service in the public transport system for a special fee. Stiglic et al. [266] investigated the potential of an integrated ride-sharing and public transport system to achieve reliable and efficient transfers between the two forms of services. Through mathematical and computational analysis, the authors found that drivers' willingness to ride-share is critical to the overall system performance.

The sharing of taxi trips to alleviate disruptions is also gaining attention in the literature. The limited literature studying the integration of demand-responsive transport (DRT) into PTA disruption management shows the potential of such solutions, but does not systematically evaluate a wide range of possible solutions, often evaluating only certain aspects (such as recovery time or cost effects) while ignoring others (such as passenger waiting time distributions). In general, collaboration on disruption management between the PTA and DRT services could potentially reduce the negative impacts for travelers as well as decrease costs for all stakeholders. According to Zeng et al. [87], tram operators in Munich and Berlin handle short-term disruptions through collaboration with a local taxi operator to provide faster rail replacement service. Since taxis operate throughout the entire city, they can be assembled considerably faster. In addition, using taxis allows the tram system control centers to concentrate on recovering the system.

Let's consider an example of research conducted by Tatiana Babicheva and her colleagues in the field of transport optimization: Matej Cebecauer, Wilco Burghout, Erik Jenelius, and David Leffler.

In the paper "Integrating Demand Responsive Services into Public Transport Disruption Management" [41], various disruption management strategies using demand-responsive transport and bridging buses were evaluated. Simulation experiments were performed on a case study based on public transport smart card data and taxi probe data in the Stockholm region, Sweden. Starting from a baseline scenario where only bridging buses were used, we compared taxi-based DRT scenarios starting from typical unrestricted taxi-DRT without ride-sharing, then adding ride-sharing, limiting the service area, and finally proposing a cooperative scheme where we tried to take advantage of high-capacity bridging buses as well as the flexibility of DRT.

The demand was considered during an afternoon peak and was considerably significant.

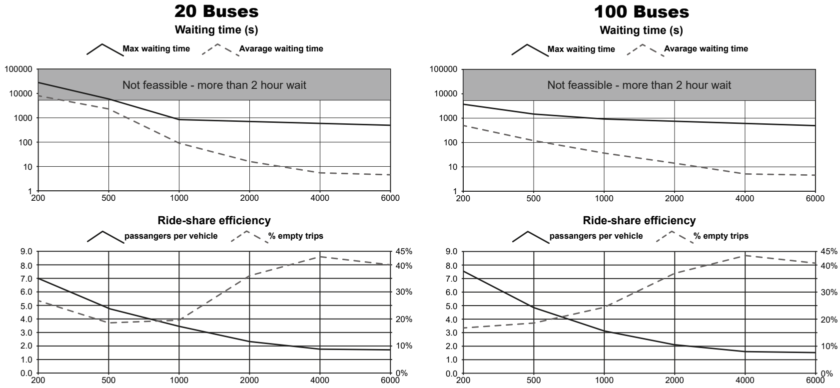


Figure 5.22: Results for scenario 6: Mixed buses and DRT for various fleet sizes. Left: 20 buses, right: 100 buses. DRT fleet size varies between 200 – 6000 vehicles (x-axis).

Figure ?? illustrates the results for a scenario in which buses and DRT collaborate. The results indicate that waiting times are significantly reduced by combining the high capacity of buses with the flexibility of DRT (featuring ride-sharing and a limited service zone). Even a fleet of 200 DRT vehicles and 20 buses (left column in Figure ??) achieves marginally acceptable waiting times at an operating cost of 140 kSEK. Increasing the fleet size further reduces waiting times but at a steeply rising cost to the operator and decreasing efficiency in terms of passengers per vehicle and the percentage of empty vehicle trips.

We will not delve into technical details, but what conclusions can we draw?

The results of the computational experiments lead to the following main conclusions:

- Operating buses alone leads to waiting times exceeding two hours, even with relatively large reserve bus fleets.
- Taxi DRT without ride-sharing and without limiting the service to the

disruption zone also results in unacceptable waiting times and exorbitant costs due to the large fleet size required to meet the demand.

- Implementing ride-sharing for demand-responsive vehicles has a significant impact, as it enables rapid decreases in waiting times, delays, and costs for all involved parties.
- Restricting the service area to only the disruption zone significantly reduces waiting times and costs, and enhances the efficiency of ride-sharing.
- Coordinating bridging buses and demand-responsive transport offers the best trade-off between costs and waiting times. In this scenario, a small number of DRT vehicles, along with a limited number of buses, can provide a sufficient level of service for disruption management at relatively low costs compared to independent strategies.

5.6.2 Park and ride

How do mathematical optimization methods assist when public transport is already functional? Despite the efficacy of public transport systems, the annual population growth and the increasing number of vehicles continue to strain the transport infrastructure of large cities. Issues such as disruptions in vehicle operations, problems with vehicle storage, the absence of organized parking in city centers, and traffic congestion during peak hours are prevalent. City authorities employ various strategies to address these challenges, including allocating special lanes for public transport, revising route patterns, updating the vehicle fleet, and implementing punitive measures such as fines for improper parking and vehicle towing to penalty lots.

Many city authorities focus on establishing new parking zones, particularly park-and-ride facilities. While the economic efficiency and social significance of creating park-and-ride parks are clear, assessing their environmental impact requires further research. An environmental audit is increasingly relevant for evaluating the operations of these facilities.

Park-and-ride facilities are car parks located near stops, stations, or transport hubs where drivers can leave their vehicles and continue their journey via public transport. The primary goal of these parking lots is to alleviate urban traffic congestion. According to some sources, park-and-ride parking originated in China for cyclists, while other sources claim that these facilities first appeared in the US in the mid-1970s alongside carpooling initiatives. Further studies suggest that the planning of the park-and-ride (P&R) system began in the 1920s in the United States as parking systems independent of the transportation infrastructure. Regardless of their origin, P&R systems have been recognized as essential in various countries and cities worldwide, such as Oxford and Nottingham in the UK, and in Hungary including its capital Budapest [216], as well as in West-European cities like Munich, the Netherlands, and Belgium [304].

Some early mentions in scientific articles provide insights [77]: *«As part of a strategy to reduce vehicle emissions, local authorities could increase the number and capacity of Park and Ride schemes. Oxford's Park and Ride scheme began in* 1973 *and has expanded to serve around* 4, 000 *cars each day. Its growth has contributed to preventing a rise in the number of vehicles entering the city each day, attracting about* 15—20% *of radial traffic flows during the morning commuting peak. However, the impact on air quality is not totally positive since a significant proportion of Park and Ride trips are ones that would not otherwise be made, or which would otherwise be made entirely by public transport. Some people will drive further to use Park-and Ride facilities. Unless restraint measures are introduced in the city centre (e.g, reduced parking spaces and/or increased charges), the road space freed by people switching to Park and Ride may simply be filled up by new local trips.»*

As part of the Dallas Urban Corridor Demonstration Program, a park-and-ride facility was developed to encourage automobile users to switch to transit [59]. The program aimed to enhance the service level on the North Central Expressway and divert some of the 160, 000 trips projected for 1990, achieving improvements in air quality and energy consumption. The success of the initial park-and-ride facility led to the development of additional facilities.

Statistics from 1974 reveal: *«An important consideration in determining the size of the park-and-ride lot and its distance from downtown is the distance the typical park-and-ride user is willing to travel from home to the lot, and from the lot to the downtown destination. A survey in the northeastern United States conducted among interchange parkers revealed that more than* 50 *percent of those who used public transit drove less than* 5 *miles from their point of origin to the change-of-mode facility. More than* 80 *percent drove less than* 10 *miles.»*

A 2002 rail survey undertaken by the Greater Wellington Regional Council (GWRC) was the primary data source used to define catchments. Distance as recorded in the survey is straight-line distance rather than road-based. As such, all distances recorded here are straight-line. Based on the 2002 rail survey, ap-

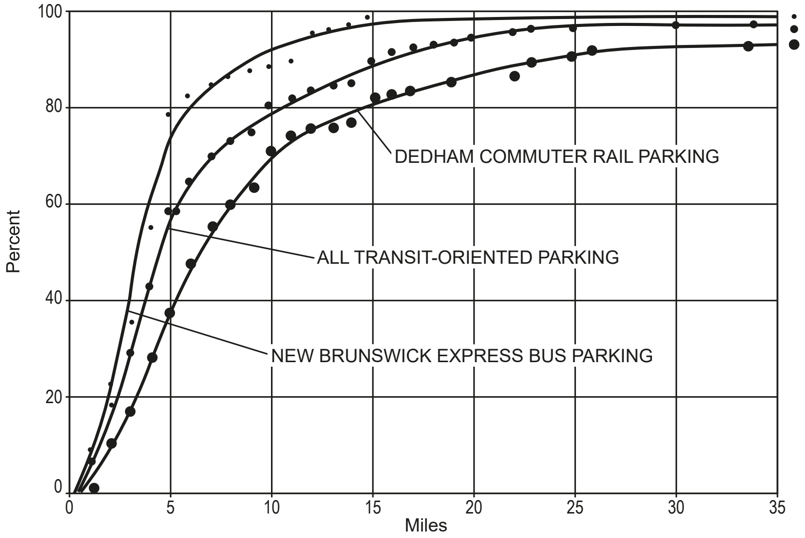


Figure 5.23: Distance from home to interchange parking area in New Jersey

proximately 50% of car access to all rail stations in the Wellington Region is within 1.85 km, 75% within 3.5 km and 90% within 6.5 km. The maximum distance passengers are prepared to travel by car to access rail is 32 km and the minimum is around 250 m (this also includes car drop-off). Car passenger distance ranges are 1.6 km for 50%, 3.1 km for 75% and 5.9 km for 90% of demand. [211]

P&R has transitioned from an independent system to an integral part of the transportation infrastructure, as well as the parking policies included in sustainable mobility plans (SUMP) [190], [124], [273].

Intercepting parking should possess the following characteristics:

1. be located within walking distance of public transport hubs;
2. be accessible to all car owners;
3. offer free parking for individuals continuing their journey by public transport, or provide significantly lower rates than those of typical city parking lots.

Planning the P&R system varies by city. In European cities, for instance, finding space for P&R systems is challenging due to limited available space, often already occupied. Thus, planning begins with enhancing the coverage of the

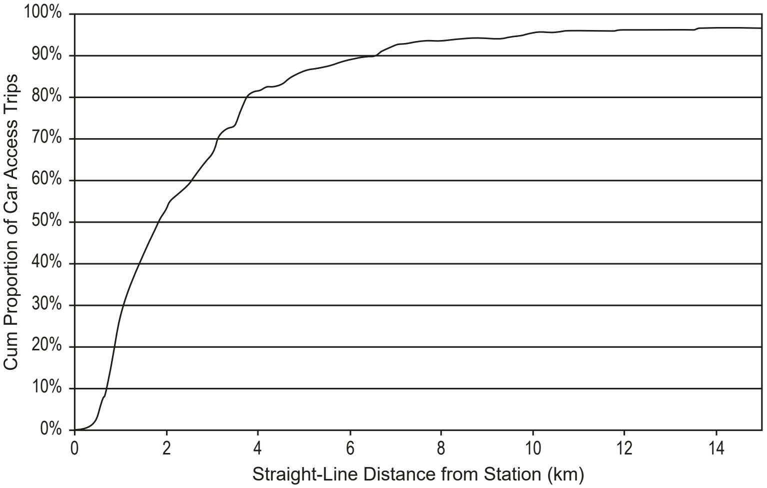


Figure 5.24: Car access cumulative trips by straight-line access distance (all stations).

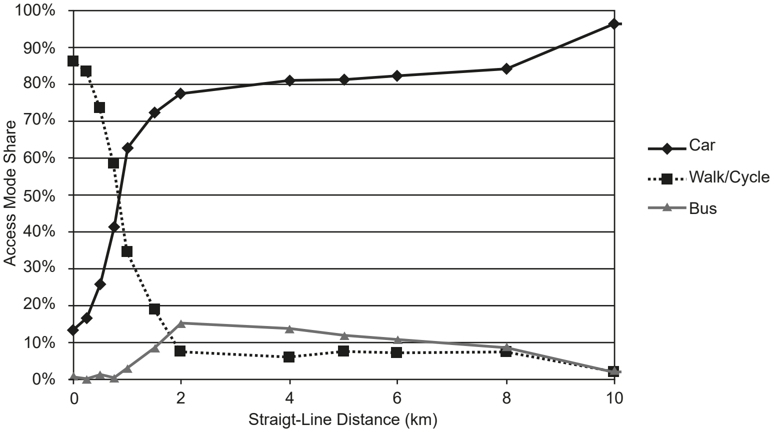


Figure 5.25: Access mode share by distance (all stations).

second leg of the trip via public transport. Conversely, in Latin American cities, where P&R is a relatively new concept, planning starts with identifying the location for future facilities. Many cities implement this system to decrease private vehicle trips to the city center, thereby reducing pollution. Planning for the P&R system may include the integration of electric and autonomous vehicles in smart cities. There is no one-size-fits-all approach to planning; it depends on the type of city and the transport policies related to the P&R system [207].

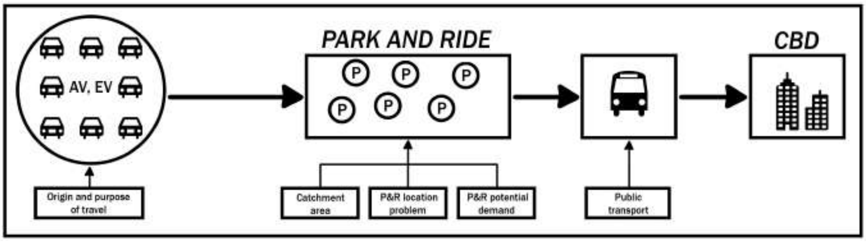


Figure 5.26: Planning a Park and Ride system.

The facilities within the P&R system are a critical aspect of planning and improving the system. The attractiveness of a facility to potential users depends on its location. If the distance or travel time from the P&R to the destination is too long, it may deter use; conversely, if it is too close to the destination, the system may be unnecessary. For planners, choosing a location involves more technical considerations, including travel times, costs, and land availability for construction. Therefore, determining the optimal location requires studying various methods, methodologies, and criteria [127].

In a highly cited article by Parkhurst in 1995 [215], the following benefits and disbenefits of park and ride systems were discussed:

Summary of benefits

1. Park and ride not only contribute to economic and environmental improvements but also garner political support for broader town center transformations like pedestrianization.
2. Implementing park and ride systems can alleviate local congestion, reduce energy consumption, and decrease air pollution by diverting some urban traffic.
3. Parking facilities on the city outskirts increase the overall supply, freeing up central land for more economically beneficial uses. Although reducing central parking supply is a potential benefit, practical examples in the UK are scarce due to the dominance of privately controlled parking.
4. Restraining traffic and parking in city centers does not necessarily reduce overall access. Park and ride systems enable increased traveler numbers without the congestion and environmental impact associated with additional car arrivals.
5. Motorists appreciate park and ride for stress-free travel, often utilizing bus lanes. Upon arrival, they may enjoy preferential access to shopping areas and avoid the challenges of multi-storey parking. Positive experi-

ences with public buses could encourage broader use of public transport.
6. By enhancing central core accessibility, park and ride systems may help mitigate pressures for suburban development and out-of-town relocations.

Summary of dis-benefits

1. A significant drawback of park and ride is the environmental impact of constructing large car parks on sensitive land, often encroaching on green belts at the urban fringe.
2. A more critical shortcoming lies in the limited effectiveness of park and ride in reducing downstream traffic. Park and ride schemes in Oxford and York, for instance, have not led to long-term traffic reductions but rather maintained a congested equilibrium [215].
3. While a cost-benefit assessment may justify subsidies, there has been a historical burden on urban residents for park and ride costs. Some schemes create equity issues by discouraging non-car arrivals, denying locals the chance to walk or cycle to the site, despite contributing financially.
4. Even though the total market for bus travel within the site may increase, subsidies are typically needed for the park and ride site itself. This can lead to competition for marginal users, impacting public transport revenue, bus energy efficiency, and possibly requiring more public subsidies.
5. Increased accessibility to the urban center, while beneficial, can also result in more trips from greater distances. Park and ride's association with increased parking stock creates additional trip-end opportunities, potentially leading to increased rural development as the penalties of accessing the city are reduced.

On the one hand, the operation of park and ride (P&R) facilities negatively impacts the environment. Transport is a constant source of noise, so parking areas are strategically located on the periphery of residential zones, away from parks, playgrounds, and similar areas. However, the environmental efficiency of using P&R systems is undeniable, as emissions from driving are significantly higher compared to those from parked vehicles. But where should these parking lots be located? Here, mathematics proves invaluable.

The optimal location of P&R facilities can be analyzed through a geographic information system. Planners in Delaware City evaluated several potential sites based on criteria such as proximity to the central business district (CBD) or primary activity centers, minimal lot competition, travel characteristics to the

CBD, maximization of the service area population, location relative to transit services, and frequency of transit service, thus determining the optimal locations for P&R facilities [81]. Multicriteria methods were also employed, using a set of primary and secondary criteria to gather expert opinions. The results indicate that the main criterion for P&R facility location should be accessibility to public transport, suggesting that these facilities should be situated near public transport infrastructure [209], [208]. For instance, positioning them close to railway stations maximizes the potential to divert private vehicle users from the transport network [126]. Additionally, transportation costs per unit distance were considered to determine suitable locations [62], [170]. In smart cities, a transport network incorporating various modes and P&R provides data on the number and location of P&R facilities, including mode choice under different conditions and the impact of travel time due to P&R implementation in a real environment [226]. Furthermore, multi-objective spatial optimization modeling methods can be applied, incorporating three fundamental criteria in the P&R system: covering as much potential demand as possible, situating P&R facilities close to major roads, and integrating these facilities within an existing system [89].

Demand is a critical parameter for locating facilities, described as a function of distance and coverage. A discrete linear model for locating P&R facilities demonstrates the flexibility and utility of the modeling approach developed to address a wide range of planning issues [88]. Additionally, mode choice according to P&R usage rates can maximize benefits and minimize social costs [284]. Linear models have been utilized to designate a series of P&R facilities in an average city, aiming to establish criteria that best approximate reality [206]. A mixed linear programming formulation determines the location of a fixed number of P&R facilities to maximize their usage [15]. An evaluation for P&R reliability analysis is used to locate facilities in a stochastic P&R system, where travelers have the option of car mode or P&R mode. The results show that the reliability of P&R facilities is significantly influenced by parking capacity, frequency, and metro fare [85]. A case study recommended locating facilities 5—6 km from the city center, except when geographic barriers exist [34], and aimed to reduce the possibility of P&R facilities contributing to congestion [263].

The complexity of locating P&R system facilities in a city's urban environment increases with the number of criteria included, such as demand, connectivity, transportation design, and economic viability [123]. Various methods and methodologies, including mathematical models and computer programs, are used to study the location of the P&R system [72] [301]. For example, a two-level programming model for P&R localization captures the interactions between decision-makers and travelers to maximize total social welfare [86].

On the other hand, the location of the P&R system also depends on the transit service levels serving these facilities [251]. A study on the location of all facilities in a city's urban area is already considered a planning tool combined with transportation policies, including minimizing operating deficits and adding decision variables such as transit and parking fees [39] [40]. Additionally, combined modal split and traffic assignment models can establish a two-level mathematical programming model to determine the optimal location and capacity of P&R [48].

5.7 Smart city, smart transport?

The concept of a **Smart City** is widely discussed in the literature. Breux and Diaz describe three aspects of smart city concepts: "An efficient city, rational and leader", "A more stimulating city where life is good", "A sustainable city".

«A data-driven city is a city that intelligently uses data to better deliver critical services. Transparency, open data, and innovation are all important parts of the modern civic identity [...] which is focused on strengthening its position as a tech leader. However, being a data-driven city is really about more efficiently and effectively delivering the core services of the city: smarter, risk-based resource allocation, better sharing of information agency-to-agency to facilitate smart decision-making, and using the data in a way that integrates in the established day-to-day patterns of city agency front line workers. Being data-driven is not primarily a challenge of technology; it is a challenge of direction and organizational leadership». [95]

An essential component of the smart city is smart transport.

An important factor in the sustainable socio-economic development of a city, region, or country as a whole is the uninterrupted functioning of the transport system. Carrying passengers is the main function of urban public passenger transport, which determines the population's quality of life and creates a barrier-free environment for societal development.

5.7.1 Personal Rapid Transit or even autonomous taxis?

One direction of smart transport focuses on the development of Personal Rapid Transit (PRT) systems.

PRT systems employ several driverless vehicles for on-demand passenger transportation from one station to another. Unlike traditional mass transportation systems, which use fixed routes and stops, PRT is comparable to a taxi service, offering non-stop on-demand transport. Most research on PRT systems assumes dedicated infrastructure (guideways) for PRT vehicles. Due to the high construction cost and visual intrusion of such guideways, as well as the rigidity of routes provided, only a few implementations of PRT services have ever been completed, e.g., Morgantown PRT, Heathrow.

Autonomous or driverless vehicles have also seen rapid development recently. Autonomous taxis (aTaxi) can be considered a generalization of PRT systems, as they provide door-to-door on-demand transportation using existing roads. Additionally, there is a current trend toward “Transport as a Service” (TaaS) or “Mobility as a Service” (MaaS) approaches: cars (and bicycles, etc.) are shared using increasingly convenient online platforms. This revolutionizes the market for passenger transportation and provides new ways to match planned and unplanned travel demand with transport service availability.

Let’s discuss the history of PRT systems.

In 1953, Edward Haltom observed limitations in conventional large monorail trains, such as the Wuppertal Schwebebahn. He noted that the time required for these trains to start and stop meant that a single line could only accommodate a limited number of vehicles per hour, typically between 20 and 40. To achieve efficient passenger movements on such a system, the trains had to be large enough to carry hundreds of passengers, necessitating the construction of large and sturdy guideways capable of supporting the weight of these massive vehicles. However, the need for extensive infrastructure resulted in high capital costs, making these monorail systems economically unattractive. This realization prompted Haltom to explore alternative transportation concepts that could offer greater flexibility and cost-effectiveness.

Haltom shifted his focus towards developing a transportation system that could operate with shorter timings, allowing for smaller individual cars while maintaining overall route capacity. By reducing the size of the cars, the weight at any given point was minimized, resulting in smaller and less expensive guideways. To address congestion at stations, Haltom's system employed "offline" stations that allowed mainline traffic to bypass stopped vehicles. His design, called the Monocab system, featured six-passenger cars suspended on wheels from an overhead guideway. However, like most suspended systems, it faced challenges with switching arrangements, as switching from one path to another required moving the rail, a slow process that limited the potential headways.

Donn Fichter, a city transportation planner, initiated research on Personal Rapid Transit (PRT) and alternative transportation methods around 1953. In 1964, Fichter published a book [93] proposing an automated public transit system specifically designed for areas with medium to low population density. A key argument in the book emphasized the need for flexibility and significantly improved end-to-end transit times compared to existing systems to convince people to switch from cars to public transit. Fichter believed that only a PRT system could provide the desired flexibility and performance. While a few other urban and transit planners also discussed the topic and some initial experiments took place, PRT remained relatively unknown at the time.

During the late 1950s, the issues associated with urban sprawl became increasingly evident in the United States. As cities improved their road infrastructure and reduced transit times, suburbs began to emerge at greater distances from city centers, resulting in population migration away from downtown areas. The lack of pollution control systems, coupled with the rapid increase in car ownership and longer commutes, contributed to significant air quality problems. Additionally, the movement to the suburbs resulted in a decline in capital investment in downtown areas, contributing to the rapid decay of urban areas in the US.

Mass transit systems were seen as a solution to address these challenges. However, during this period, the federal government exacerbated the problem by focusing on funding the development of the Interstate Highway System while scaling back support for mass transit. As a result, public transit ridership in most cities experienced a sharp decline.

In 1962, President John F. Kennedy urged Congress to address these issues, leading to the formulation of plans. In 1964, President Lyndon B. Johnson signed the Urban Mass Transportation Act of 1964 into law, establishing the Urban Mass Transportation Administration (UMTA). UMTA was created to fund the development of mass transit systems, similar to how the Federal Aid Highway

Act of 1956 had supported the construction of the Interstate Highways. UMTA aimed to provide financial assistance for the capital costs associated with building new transportation infrastructure.

However, urban planners who were aware of the Personal Rapid Transit (PRT) concept expressed concerns that building more transportation systems based on existing technologies would not adequately address the problem, as highlighted by Fichter earlier. They believed that new systems would need to offer the flexibility and convenience comparable to that of a personal car.

In 1966, the United States Department of Housing and Urban Development was tasked with studying innovative urban transportation systems that could transport people and goods quickly, safely, without polluting the air, and in a manner that aligned with sound city planning principles. The resulting report, published in 1968, recommended the development of PRT alongside other systems like dial-a-bus and high-speed interurban links.

During the late 1960s, the Aerospace Corporation, an independent non-profit corporation established by the US Congress, invested considerable time and resources into PRT research. They conducted extensive theoretical and systems analysis work, although they were restricted from selling their findings to non-federal government customers. In 1969, members of the Aerospace Corporation's study team published the first widely publicized description of PRT in Scientific American [117]. In 1978, Irving et al. published a widely cited technical book [138] on the subject. These publications triggered a "transit race" reminiscent of the space race, with countries worldwide eagerly joining what seemed to be a promising market with immense potential.

The oil crisis of 1973 had a significant impact on the cost of vehicle fuels, which sparked increased interest in alternative transportation solutions.

In 1967, the Aramis project was initiated by aerospace giant Matra in Paris. With an expenditure of approximately 500 million francs, the project aimed to create a "virtual train" system. However, Aramis faced control software issues that caused cars to collide, leading to its failure during qualification trials in November 1987. Despite considerable efforts, the project ultimately did not succeed.

From 1970 to 1978, Japan implemented the "Computer-controlled Vehicle System" (CVS). In a full-scale test facility, 84 vehicles operated at speeds of up to 60 kilometers per hour on a 4.8-kilometer guideway, achieving one-second headways during tests. Another version of CVS operated publicly for six months from 1975 to 1976, featuring 12 single-mode vehicles and four dual-mode vehicles on a 1.6-kilometer track with five stations. This version of CVS successfully transported over 800, 000 passengers. However, the Ministry of Land, Infrastructure, and Transport of Japan ultimately deemed CVS unsafe under existing rail safety regulations, particularly regarding braking and headway distances, leading to its cancellation.

On March 23, 1973, Frank Herringer, the administrator of the U.S. Urban Mass Transportation Administration (UMTA), testified before Congress about initiating a program for a high-capacity Personal Rapid Transit (HCPRT) system with half to one-second headways. However, the HCPRT program was redirected into a more modest technology program, allegedly due to lobbying efforts from interests concerned about potential obsolescence if a genuine PRT program gained visibility, as suggested by J. Edward Anderson [11]. Consequently, those interested in HCPRT were unable to secure research funding from UMTA.

In 1975, the Morgantown Personal Rapid Transit (PRT) project was completed. It featured five off-line stations that allowed for non-stop, individually programmed trips along a 14.0-kilometer track. The system utilized a fleet of 71 cars. While the Morgantown PRT demonstrated crucial characteristics of PRT, it is not officially considered a PRT system due to its vehicles being too heavy and carrying too many passengers. When the system carries a large number of passengers, it operates in a point-to-point manner rather than running continuously from one end of the line to the other. During periods of low usage, all cars make a full circuit, stopping at every station in both directions. The Morgantown PRT system remains in operation at West Virginia University in Morgantown, West Virginia, serving approximately 15, 000 riders per day as of 2003. It successfully showcases automated control, but its high operational and maintenance costs, particularly related to the steam-heated track, have hindered its expansion to other locations.

Currently, the oldest operating automated public transportation system is located at West Virginia University. Constructed in the 1970s, it was designed to connect three university campuses situated in different parts of the city. This transportation system is relatively small and straightforward, consisting of a single, non-branching, and non-intersecting line that spans approximately 6 kilometers with five stops. The small vehicles travel along the line at an average speed of 22 km/h, similar to trams.

From 1969 to 1980, Mannesmann Demag and MBB collaborated to develop the Cabinentaxi urban transportation system in Germany. The two companies formed the Cabintaxi Joint Venture and successfully created an extensive PRT technology that received full development approval from the German government and its safety authorities. The system was intended to be implemented in Hamburg, but budget constraints led to the cancellation of the proposed project before construction could begin. With no other potential projects in sight, the joint venture dissolved, and the fully developed PRT technology remained uninstalled. In 1985, Cabintaxi Corporation, a US-based company, acquired the technology and continues to be active in the private-sector market for transportation systems.

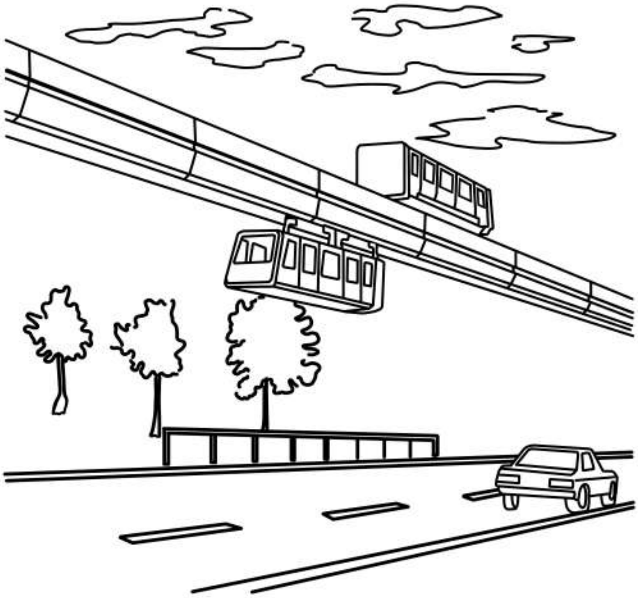


Figure 5.27: CaninenTaxi. Cabins were located immediately on both sides of the flyover: above and below it

In 1979, the three-station Duke University Medical Center Patient Rapid Transit system was inaugurated. This system ceased operations in 2009 to accommodate the hospital's expansion.

During the 1990s, Raytheon heavily invested in a system known as PRT 2000, which was based on technology developed by J. Edward Anderson at the University of Minnesota. However, Raytheon was unable to fulfill a contract to install the system in Rosemont, Illinois, near Chicago, as the estimated costs soared to US$50 million per mile. This was reportedly due to design modifications that increased the system's weight and cost compared to Anderson's initial design. In 2000, the University of Minnesota regained the rights to the technology, which were later acquired by Taxi2000.

In 1999, the ParkShuttle system, designed by 2getthere, commenced operations in Rotterdam's Kralingen neighborhood with 12-seater autonomous buses. The system expanded in 2005, and second-generation vehicles were introduced to cover five stations along a 1.8-kilometer route, incorporating five grade crossings with conventional roads. The service operates on a scheduled basis during peak times and on-demand at other times. Additionally, in 2002, 2getthere deployed twenty-five 4-passenger "CyberCabs" at the Floriade horticultural exhibition in the Netherlands. These vehicles transported visitors along a 600-meter one-way track spiraling to the summit of Big Spotters Hill, featuring only two stations. This six-month endeavor aimed to evaluate public reception of PRT-like systems.

Furthermore, there are several automated metro systems worldwide, such as the complete metros in Copenhagen, Dubai, and Haifa (with the metro-funicular "Carmelit"), as well as specific lines in cities like Paris, Lausanne, Nuremberg, Taipei, and London (Docklands Light Railway). Numerous older subway networks in Europe and Asia plan to transition to full automation gradually. The Moscow monorail system is also capable of automatic operation, though it currently operates with drivers.

Many large airports and industrial complexes utilize people movers to transport groups over short distances. Typically, these consist of small trains or individual carriages that automatically traverse a simple route.

While these developments are impressive and intriguing, they do not encapsulate our initial vision of personal rapid transit (PRT). Despite some systems featuring compact dual cabins instead of carriages, like the American SkyTran project, they do not fulfill certain criteria for PRT. What are these criteria? Let's discuss that now.

A genuine personal automated transport system must first travel on its own dedicated highway, free from level intersections with either vehicular or pedestrian traffic. Such a highway will likely be elevated, though it may descend to ground level or include underground sections in areas devoid of roads and buildings. The highway for automated transport, particularly at ground level, must be safeguarded against human intrusion, foreign object falls, and possibly even precipitation (notably in winter) to a greater extent than railway tracks.

Moreover, PRT must be completely personal. This implies that the system should be available on demand, transporting an individual or a small group (such as a family or friends) directly to their destination without the presence of other passengers or the need for intermediate transfers.

Stops should be located within walking distance from anywhere in the city. To surpass public transportation, the PRT network must be significantly denser, with stops no farther apart than those for buses or trams. Ideally, the PRT line would run along every street in the city, save for minor lanes, dead ends, and inter-house driveways, with stops every 100—200 meters. In some instances, stops could even be situated inside buildings — such as office and shopping centers, metro stations, railway stations, and airports.

To also outperform private vehicles, PRTs (cabins, wagons) must be swift, and the network must possess high bandwidth. "Fast" means an average city speed of at least 60 km/h, and on non-stop highways, up to 150—200 km/h.

The network's capacity refers not only to speed but also to its ability to handle large passenger flows, such as transporting football fans home after a match — a task not always feasible for the subway. There must be no traffic jams; the computer controlling the PRT system must optimize routes around congested network sections and adjust them if necessary. Additionally, waiting and boarding areas should not be overcrowded; a high number of stops can help disperse crowds. The system should also enable overtaking and swiftly transport a required number of cars to any city area.

Practically speaking, such "transport of the future" is not feasible or justified in every city or state. It is needed in large metropolitan areas with acute transportation issues that cannot be resolved by existing modes of transport. Moreover, constructing a PRT system demands advanced technology and substantial capital investments. Although a kilometer of PRT highway is cheaper than that of a metro, the network's density is much greater.

An important, though not mandatory, aspect of PRT organization is the "open architecture" principle of the network. This allows various organizations to build their own PRT sections and connect them to the citywide network, facilitating different models of cars and equipment from various manufacturers. All network components must be compatible, akin to Lego pieces.

There will inevitably be challenges in introducing a new mode of transportation to the urban environment.

Initially, it may not be possible to build an effective control system and extensive network with numerous alternative routes, leading to traffic jams and crowds. Early adopters of PRT might jostle at stops and navigate the city in frustration as the system's automation struggles to find an exit to the desired highway.

It's also crucial to consider the PRT network's access system, which relies on maintaining a safe distance between cars. Some developing systems allow for close convergence and even temporary coupling on similar routes. However, in Japan, for example, the standards for spacing PRT units are as stringent as for rail transport. This was a contributing factor to the closure of one of the world's first PRT systems, CVS, after several years of test operation; it was deemed potentially unsafe due to insufficient spacing between the trailers to decelerate in emergencies.

Another hurdle is the public's initial skepticism and fear of entrusting their safety to automation. The thought of speeding above the ground in a robot-controlled cabin can be daunting, especially considering that even simple systems like elevators can malfunction. In a more complex system, malfunctions could be more frequent and varied, potentially leading to people becoming hostages or victims of faulty automation. For instance, Morgantown's small automated transit system experiences breakdowns several times a week.

Furthermore, people often prefer their own cars over public transport for status and comfort, so they might not readily switch from their personal vehicles to PRT. The perceived lack of prestige in public transport can be partially addressed by offering luxury, driverless limousines. However, amenities remain a challenge for public transport, as vandalism is inevitable, and PRT vehicles could deteriorate as quickly as subway cars or minibuses, if not faster.

Despite these challenges, personal automatic transport is not fundamentally unattainable.

It is important to mention that currently, there is no transport system that fully qualifies as PRT, though several projects are underway that aim to achieve this in the future.

The most renowned of these initiatives is the ULTra ("Urban Light Transport") experimental transport system at London Heathrow Airport. In January 2003, the prototype ULTra system in Cardiff, Wales, was certified to carry passengers by the UK Railway Inspectorate on a 1 km test track. ULTra consists of four-seater cars that move on rubber-tired wheels on a dedicated overpass. Currently, they operate solely within the airport, but plans are underway to extend service to nearby hotels. The developers anticipate that, eventually, the cars will navigate city streets, although they will require human control outside the overpass.

A project closely resembling ULTra is the CyberCab, developed by the Dutch company 2getthere for Abu Dhabi Airport and the adjacent eco-city of Masdar. It was announced that Masdar would exclusively use CyberCab electric cars for transportation, replacing traditional streets with footpaths. The system operates in an undercroft beneath the city and was intended as a pilot for a more extensive network that would include freight transport. However, the expansion was halted after the pilot's inauguration due to the high costs of constructing the undercroft, leading to proposals for alternative electric vehicles.

The third operational PRT project is Vectus, a Korean-Swedish collaboration. In June 2006, Vectus Ltd commenced the construction of a 400 m test track in Uppsala, Sweden. The system was showcased at the 2007 PodCar City conference in Uppsala. A 40-vehicle, 2-station, 4.46 km system named "SkyCube" was launched in Suncheon, South Korea, in April 2014.

In the 2010s, the Mexican Western Institute of Technology and Higher Education initiated the LINT ("Lean Intelligent Network Transportation") project and constructed a 1/12 operational scale model. This evolved into the Modutram system, with a full-scale test track in Guadalajara that became operational by 2014.

Beyond the aforementioned, numerous projects exist solely in conceptual form or as isolated prototypes. These ventures are highly varied, encompassing monorail systems, magnetic levitation, and even the concept of enclosing passenger cabins within sealed tubes. Some projects consider dual-mode vehicles, capable of both automated and manual control, or integrating conventional cars into the PRT network by mounting them on special "pallets" that travel along PRT highways. The German MAIT project envisions a "container principle," where both goods and people, contained in modular units, are transported within the system.

5.7.2 Empty Vehicle Redistribution

The primary goal of a transportation operator offering Personal Rapid Transit (PRT) services is to use the available vehicles efficiently. This entails maximizing the quality of service while minimizing the total operational cost, which includes reducing energy consumption, vehicle servicing, and other related expenses.

A key challenge in such systems is twofold: assigning vehicles to passengers effectively and managing the redistribution of empty vehicles.

Passenger demand fluctuates significantly throughout the day and week, often showing a strong directional bias during peak hours (e.g., in the morning, most travel originates outside and is directed toward the city center). This variability leads to a mismatch in the number of vehicles available at starting points and destinations. Consequently, redistributing empty vehicles is crucial to align vehicle availability with passenger demand. Analyzing trip data can enhance the statistical prediction of future trip requests.

The changing pace of life in large cities, increased population mobility, the rapid growth of urban agglomerations, and longer travel distances have underscored the importance of communication speed as a key factor in travel method selection. This necessitates a reevaluation of urban transport system development concepts. Often, constraints prevent the implementation of entirely new systems, necessitating the optimization of existing systems under current conditions. Urban areas are experiencing road network congestion and rising motorization levels, adversely affecting the transport system and resulting in increased economic losses for cities and regions. In this context, the creation of an integrated and adaptive traffic organization model becomes increasingly relevant, considering the optimal use of all available resources. This includes leveraging public railways for intra-city passenger transport, establishing express routes during peak hours to facilitate passenger flow and alleviate road network congestion, and promoting off-street transportation options to encourage the use of public urban transport by enhancing its appeal.

An important aspect of PRT systems is the empty vehicle redistribution (EVR) problem, which seeks to determine the optimal routes for a fleet of vehicles to satisfy customer demand. In the Vehicle Routing Problem (VRP) with time windows, passengers must be picked up and/or delivered within specified time frames.

EVR methods can be categorized into two main groups, each addressing different challenges. The first group consists of reactive redistribution strategies, which respond to immediate demand when a passenger arrives. These algorithms dispatch an empty vehicle based on specific passenger needs (e.g., the one with the longest waiting time or the one closest to the vehicle). The second group includes proactive redistribution strategies, which involve sending empty vehicles in anticipation of forecasted near-future demand, addressing the needs of the entire system.

In some PRT systems, operations are distinguished by different modes. The demand mode operates without heuristics or knowledge of future demand, relying solely on current demand, which aligns the problem with dynamic scenarios. In the scheduled mode, future passenger demand is either known or predicted with strong heuristics, aligning the problem with static deterministic scenarios.

Kek, Cheu, and Chor [148] developed two staff-based relocation techniques termed "shortest time" and "inventory balancing." The shortest time relocation calls an empty vehicle from the nearest station to minimize travel time. Inventory balancing involves redistributing vehicles from stations with a surplus to those with a shortage. This model employs a "virtual station status," which accounts for the actual number of vehicles at a station, the weighted total scheduled for return in a given time interval, and those reserved and thus unavailable for immediate use. When the "virtual station status" falls outside the station's threshold values, a relocation request is triggered.

The following is a discussion of various methods employed in both academic literature and practical applications.

- **Reactive (Sending) Algorithms**

  Basic Allocation (BA): When there are waiting passengers and empty vehicles at the same station, the empty vehicles are assigned to the passengers who have been waiting the longest, without redistributing empty vehicles to other stations. This strategy primarily serves as a baseline for evaluating more sophisticated redistribution methods.

  Simple Nearest Neighbours (SNN): This involves reallocating the closest empty vehicles to serve the passengers who have been waiting the longest. This algorithm has been utilized by Andreasson [13], Fatnassi et al. [90], Fagnant and Kockelman [83], and Lees-Miller.

Heuristic Nearest Neighbours (HNN): Similar to SNN, this method reallocates the nearest empty vehicles to the longest-waiting passengers, but it also considers the travel time for vehicles to reach the passengers. This approach aims to improve upon SNN by accounting for vehicle travel time. The distinction between SNN and HNN can be considerable. For instance, in a network shaped like a long edge with one empty vehicle at one end and two waiting passengers at opposite ends, SNN would send the vehicle to the currently longest waiting passenger regardless of travel time. In contrast, HNN would direct the vehicle to the passenger who, upon the vehicle's arrival, would have waited the longest. This method is referenced by Kek et al. [148], and Bell and Wong [298].

Send The Nearest (STN): This method reallocates the nearest empty vehicles based on the closest pairs of available vehicles and waiting passengers. Stations with waiting passengers are ranked according to the travel time from the nearest available vehicle. Babicheva et al. describe this method.

- **Proactive (Redistribution) algorithms:**

Index Based Redistribution (IBR): This method redistributes empty vehicles based on the maximum station index. The concept of the station index is akin to the "virtual station status" as described by Kek et al. [148], but it employs a different index calculation method that focuses on the number of vehicles without considering passenger waiting times. Babicheva et al. also describe this method.

Surplus/Deficit vehicle Redistribution (SDR): This strategy involves redistributing empty vehicles from stations with the greatest surplus to those with the greatest deficit. Andreasson [13] employed this method. The surplus or deficit of vehicles at a station is defined as the number of available vehicles (currently or within a specified time horizon) minus the number of waiting passengers (currently or within the same time horizon).

Why are different algorithms used? Can't we just find the best one and use it? To understand this, consider an example.

Imagine a single-line network with two stations and two vehicles (Figure ??). Each station has one waiting client. The travel time between the two stations is 5 minutes.

In this scenario, the HNN method yields the same results as the IBR method, because predicted clients are not considered in this case study. The Surplus/Deficit redistribution method is only applicable when vehicles are located at the stations, making it unsuitable for this scenario.

The SNN, IBR, and STN algorithms were evaluated and compared to determine which algorithm provides the best maximum waiting time, the best average waiting time, and the highest number of satisfied clients.

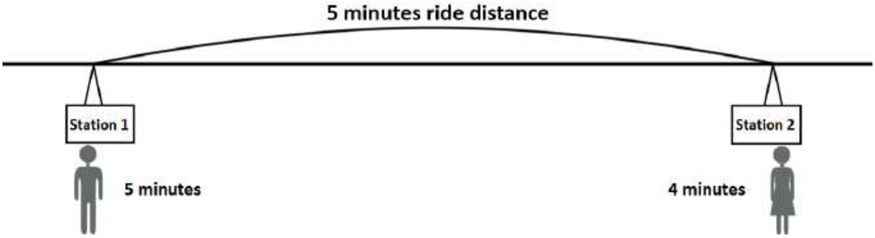


Figure 5.28: The simple line network.

We will now consider four cases with different locations of two vehicles:

Case 1. The first vehicle is located a 3-minute ride from Station 1, and the second vehicle is located between Station 1 and Station 2, with a ride time of 1 minute to Station 1 and 4 minutes to Station 2.

Case 2. The second vehicle is located a 7-minute ride from Station 1, and the first vehicle is located a 1-minute ride from Station 1.

Case 3. The first vehicle is located between Station 1 and Station 2, with a ride time of 1 minute to Station 1 and 4 minutes to Station 2. The second vehicle is located a 5-minute ride from Station 2.

Case 4. The first vehicle is located between Station 1 and Station 2, with a ride time of 4 minutes to Station 1 and 1 minute to Station 2. The second vehicle is located a 2-minute ride from Station 2.

The location of the first vehicle in Case 1 is illustrated in Figure **??**.

The **SNN** algorithm will, at each optimization step, serve the longest waiting client first with the nearest empty vehicle. The longest waiting client is at Station 1, and the nearest vehicle is Vehicle 2, which will thus be assigned to the client at Station 1. Then, there is one waiting client at Station 2 who has not yet been assigned. To this client, the remaining unassigned empty Vehicle 1 will

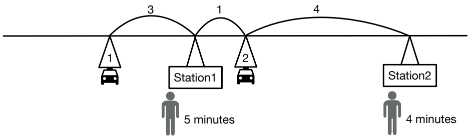


Figure 5.29: The sample line network. Case 1.

be assigned.

Therefore, the waiting times for the clients, at time of departure, will be:

- For the client at Station 1: 6 minutes (5 min + 1 min drive time).

- For the client at Station 2: 12 minutes (4 min + 8 min drive time).

Thus, for **SNN**, the maximum waiting time would be 12 minutes, and the average waiting time would be 9 minutes.

The **STN** algorithm will, at each optimization step, evaluate all the vehicle-station travel times (in drive time) and assign the nearest vehicle-station pair. In this case:

- The V1-St1 travel time is 3 minutes.

- The V1-St2 travel time is 8 minutes.

- The V2-St1 travel time is 1 minute.

- The V2-St2 travel time is 4 minutes.

Thus, the algorithm will assign Vehicle 2 to Station 1.

In the next optimization step, the algorithm will assign the remaining Vehicle 1 to the remaining client at Station 2.

The dispatching scheme in this case is therefore the same as for SNN. Thus, the waiting times for the clients, at the time of departure, will be:

- For the client at Station 1: 6 minutes (5 min waiting + 1 min drive time).

- For the client at Station 2: 12 minutes (4 min waiting + 8 min drive time).

Thus, for **STN**, as in **SNN**, the maximum waiting time would be 12 minutes, and the average waiting time would be 9 minutes.

The **IBR** algorithm will sort all the stations by their index [**IBR**], and assign the nearest vehicle to the station with the highest index.

The index for Station 1 is the (dis-)utility of the longest waiting client at this station, including the time it takes for the nearest vehicle to arrive at the station. It will take 1 minute for the nearest vehicle to reach Station 1 (Vehicle 2), thus the index will be equal to $u(5+1) = u(6)$. The index for Station 2 is $u(4+4) = u(8)$.

The utility function is monotonically increasing, so $u(8) > u(6)$. Based on the algorithm, the station with the maximal index (Station 2) will be served by the nearest vehicle (Vehicle 2).

In the next optimization step, Station 1 will be served by Vehicle 1.

Thus, the waiting times for the clients, at the time of departure, will be:

- For the client at Station 1: 8 minutes (5 min waiting + 3 min drive time).

- For the client at Station 2: 8 minutes (4 min waiting + 4 min drive time).

Thus, for **IBR**, the maximum waiting time would be 8 minutes, and the average waiting time would be 8 minutes.

Therefore, in Case 1, the **IBR** method shows the best results for both the maximal and average client waiting times.

As shown in Table **??**, there is no single algorithm that excels in all evaluated cases. For instance, the IBR algorithm performs best in cases 1 and 4, while the SNN algorithm excels in cases 3 and 5, and the STN algorithm in cases 2 and

3. This illustrates that even in straightforward evaluation scenarios, different situations necessitate distinct algorithms. We therefore conclude that a combination of algorithms may offer a more resilient strategy and reduce sensitivity to specific demand characteristics and vehicle distributions within the network. Additionally, we observe that the sequential assignment of vehicle-client pairs, characteristic of greedy algorithms, frequently results in suboptimal solutions.

Table 5.1: Average and maximal waiting times, in minutes.

	Case 1		Case 2		Case 3		Case 4	
	Avg	Max	Avg	Max	Avg	Max	Avg	Max
IBR	**8**	**8**	13.5	16	11.5	15	**7.5**	**9**
SNN	9	12	13.5	16	**7.5**	**9**	**7.5**	**9**
STN	9	12	**8.5**	**12**	**7.5**	**9**	8.5	12

As shown in Table **??**, when considering the problem with time windows, there is similarly no single algorithm that is universally superior across all cases. In each scenario, the number of clients not served within their designated time windows — assumed to be 10 minutes — is minimized using either the IBR, SNN, or STN algorithms.

Table 5.2: Number of clients not served within the time windows.

	Case 1	Case 2	Case 3	Case 4
IBR	**0**	2	1	**0**
SNN	1	2	**0**	**0**
STN	1	**1**	**0**	1

In all proposed methods, dispatching depends on a core parameter, which may be the station index, the number of clients, the maximum waiting time, or the minimum distance to an available vehicle, given that clients are present at the station. The mixed methods aim to dispatch vehicles on a greedy basis. After determining each station's index, which is contingent upon the chosen redistribution strategy, we can identify the set of vehicles available for redistribution. Greedy methods, which proceed by matching the first-sorted elements step by step, can exhibit good convergence. However, they are unlikely to consistently yield the optimal solution and may result in unnecessary trips by empty vehicles.

To improve the solution, we represent the data from the two sets in the form of a bipartite graph, with edges corresponding to travel times from the available vehicles to the stations requiring service.

We aim to match vehicles to the top N stations in the network. To minimize the average waiting time for clients, we must find matchings that minimize the total weight sum of the used edges. Although this problem is NP-hard in general — particularly when weights depend on the number of passing vehicles — polynomial solutions exist for certain cases where weights are constant. For instance, if the number of serving stations equals the number of available vehicles, we can employ the Hungarian algorithm. Next, we explore heuristic methods for solving the matching problem.

Each call of the matching function in the network will serve the top N stations based on the selected parameter. The number of matchings will not exceed the number of available vehicles.

Consider the following example to demonstrate that a greedy algorithm may not yield the optimal solution. We adopt the SNN strategy, which is based on the maximal client waiting time.

Figure **??** depicts a simple network with two stations and two available vehicles, represented as the bipartite graph in Figure **??**.

Figure 5.30: The sample network.

The greedy approach can be applied in two ways: choosing the client with the maximum waiting time at each step or selecting the vehicle with the shortest path to any waiting client. In the first scenario, the client at Station 1 with the longest wait is served by the nearest vehicle (2). The next client, at Station 2, is then served by Vehicle 1. The total waiting time is $(5 + 2) + (4 + 8) = 19$ minutes, and the total travel time for empty vehicles is $8 + 2 = 10$ minutes. In the second scenario, Vehicle 2 is chosen to serve Station 1 because it has the shortest time to reach the nearest client. Subsequently, Vehicle 1 serves Station 2. Both scenarios result in the same outcome.

The optimal solution would pair Vehicle 2 with Station 2 and Vehicle 1 with Station 1, resulting in waiting times of 8 minutes for Station 1 and 7 minutes for Station 2. Consequently, both the average and maximum waiting times are reduced.

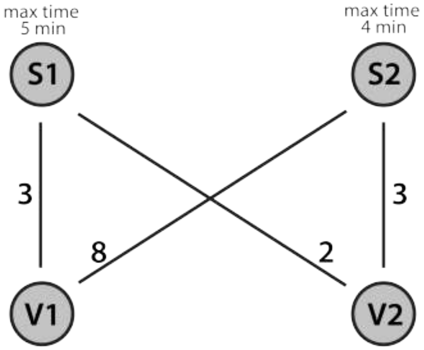


Figure 5.31: The bipartite graph.

This demonstrates that no single algorithm can always provide the best solution at every moment. Hence, even small problems in traffic optimization are extensively discussed in literature and are subjects of various PhD theses.

5.7.3 Smart vehicles in real life

Currently, the economic analysis of extensive autonomous vehicle systems is not feasible under existing laws, but the experimentation phase is advancing rapidly.

Autonomous mobility is no longer a futuristic fantasy but a present reality and a significant challenge for the RATP Group, the world's fourth-largest public transportation operator. The RATP Group aspires to be the partner of choice for smart, human, and sustainable cities that are receptive to transformation and innovation, while being mindful of their urban model's required evolution.

How are these experiments conducted? Typically, they involve autonomous shuttles that operate on a predetermined route. A notable project is the collaboration between the City of Paris, the City of Vincennes, and the RATP Group. In November 2017, the autonomous shuttle experiment in Vincennes commenced, running a 6 km route with eight stops from Château de Vincennes station (Line 1) to Parc Floral (Paris 12th).

These experiments continued even during the recent health crisis, with appropriate precautions. On weekends from $2:00$ PM to $5:30$ PM (during the curfew), travelers enjoyed this innovative and eco-friendly mobility service, operated by three shuttles: two EZ10 shuttles from EasyMile and one Autonom® Shuttle from Navya. These French-made shuttles are fully electric and autonomous, initially accommodating up to four passengers in line with Covid-19 health protocols. Under normal circumstances, the shuttles now carry up to 11 passengers.

The third phase of this project includes on-road testing in the densely populated urban areas of Marigny and Maréchaux in Vincennes.

Since its initiation, the trial of autonomous shuttles in Bois de Vincennes has been a significant success, transporting over $40,000$ passengers, covering $11,000$ km at an average speed of 13 km/h, and recording no safety incidents.

Moreover, the trial has facilitated the creation of the new role of "safety driver" or security operator. The RATP group has trained approximately twenty safety drivers during the Bois de Vincennes trial, who have acquired comprehensive knowledge of autonomous vehicles.

Why are these safety drivers essential? These RATP agents are tasked with manually taking control of the vehicle in the event of any issues. They remain vigilant and carry a sizable control console with them.

What other projects are currently in progress or envisioned for the near future?

On February 2, 2021, the RATP Group, Arval France, and the Municipality of Rueil-Malmaison launched an autonomous shuttle service on public roads, linking the Rueil-Malmaison train station to Arval France's headquarters. The autonomous shuttle EVAA (EZ10 Gen 3) was introduced in January 2020, permitting Arval employees to test the shuttle on a predefined route. Following this, the RATP Group carried out several months of trial runs without passengers. In 2022, two shuttles were operational on public roads from Monday to Friday, 8 am to 8 pm (and during curfew, from 7 am to 6 pm). Managed by RATP, these shuttles provided exclusive transport for Arval employees from the Rueil-Malmaison train station to the company's headquarters.

Since September 2020, RATP Dev, a subsidiary of the RATP Group, has been piloting an on-demand autonomous shuttle service in partnership with ArchParc, a business park in Haute-Savoie, known as the "French Geneva." The trial, set to run until December 2021, is operated by Alpbus, a subsidiary of RATP Dev, in collaboration with the Swiss company Bestmile, and is part of the echosmile project funded by the European Union and led by ArchParc. The service, which includes 7 stops (Vivacy, Meggitt, ABC, Esplanade, Botanic, Communauté de Commune, Athéna) within the technopole, employs the EZ10 Gen 3 shuttle from EasyMile, a company based in Toulouse. The first phase of the trial, from September to December 2020, was highly successful: the shuttles traveled 1000 km without any road safety incidents and used only half of their battery capacity each day.

A potential critique of such shuttles is their relatively low speeds, which range from 13 to 18 km/h. This is intentional to maintain the highest safety standards during the trials. Nevertheless, speeds are being incrementally increased without compromising the safety of passengers and the public.

5.7.4 Science fiction becomes reality

Autonomous vehicles serve to complement existing modes of transportation, promote tranquility in city centers, reduce greenhouse gas emissions, repurpose parking areas, and decrease noise pollution.

Researchers, practitioners, and policymakers worldwide are captivated by this emerging reality. To make it even more palpable, continued collaborative efforts are essential.

The stakes are considerable: it's about understanding how these new vehicles will integrate into traffic regulations, ensuring that this innovative mode of transportation is profitable, practical, secure, and indispensable. More than that, it's about how this mode of transportation will transform the lives of our current and future generations.

We, as theorists and practitioners of intelligent mobility, pledge to do our utmost.

Bibliography

[1] In: *Jahrbuch des Kaiserlichen Automobil-Clubs und der mit ihm im Kartellvertrag stehenden deutschen Automobil-Clubs.* (1914).

[2] National Safety Council (NSC). *Car Crash Deaths and Rates.* 2021. URL: https://injuryfacts.nsc.org/motor-vehicle/historical-fatality-trends/deaths-and-rates/.

[3] US DoT. America's Highways: 1776-1976. "Federal Highway Administration". In: (1976).

[4] W. F. Adams. "Road traffic considered as a random series". In: *Journal of the Institution of Civil Engineers* 4.1 (1936), pp. 121–130.

[5] Michael Adcock. "Remaking Urban Space". In: *Baron Haussmann and the Rebuilding of Paris* 1870 (1996), pp. 25–36.

[6] Aviv Adler, David Miculescu, and Sertac Karaman. *Optimal policies for platooning and ride sharing in autonomy-enabled transportation.* 2016.

[7] United States Public Roads Administration. "Toll Roads and Free Roads". In: *Message from the President of the United States* 1939.76th Cong., 1st Sess., H. Doc. 272 (1939).

[8] United States. Federal Highway Administration. *America's Highways, 1776-1976: A History of the Federal-aid Program.* U.S. Department of Transportation, Federal Highway Administration, 1977.

[9] John Agnew. "Steam engines on UK roads, 1862–1865: Banning orders, agricultural locomotives and the 'red flag'Act". In: *The International Journal for the History of Engineering & Technology* 90.1 (2020), pp. 53–74.

[10] S. Algers et al. "The national Swedish value of time study". In: *PTRC European Transportation Forum, Warwick.* 1995.

[11] John Edward Anderson. "The Historical Emergence and State of the Art of PRT Systems". In: *INFRASTRUCTURE (NEW YORK, NY)* (1996).

[12] P. Andreas and R. Price. "From war fighting to crime fighting: transforming the American national security state". In: *International Studies Review* 3.3 (2001), pp. 31–52.

[13] Ingmar Andréasson. "Reallocation of empty personal rapid transit vehicles en route". In: *Transportation research record* 1838.1 (2003), pp. 36–41.

[14] Ingmar Andréasson. "Ride-sharing on PRT". In: *Automated People Movers 2005: Moving to Mainstream.* 2005, pp. 1–7.

[15] Felipe Aros-Vera, Vladimir Marianov, and John E Mitchell. "p-Hub approach for the optimal park-and-ride facility location problem". In: *European Journal of Operational Research* 226.2 (2013), pp. 277–285.

[16] M. Athans et al. "Stochastic control of freeway corridor systems". In: *Decision and Control including the 14th Symposium on Adaptive Processes, 1975 IEEE Conference on.* Vol. 14. IEEE. 1975, pp. 676–685.

[17] Stephen T Atkins. "Transportation planning models-what the papers say". In: *Traffic engineering & control* 27.9 (1986).

[18] AATM Aw and Michel Rascle. "Resurrection of" second order" models of traffic flow". In: *SIAM journal on applied mathematics* 60.3 (2000), pp. 916–938.

[19] N. T. J. Bailey. "On queueing processes with bulk service". In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1954), pp. 80–87.

[20] John Hollis Bankhead. "Post Office appropriation bill [fiscal year]: Report № 666 to Accompany H. R. 13308: [Submitted by Mr. Bankhead. Jan 27, 1919] // 65th Congress, 3d Sess. Washington, 1919". In: *Senate Reports (Public)* 1 (1919).

[21] Samuel M Barton. "The evolution of the wheel: velocipede to motocycle". In: *The Sewanee Review* 5.1 (1897), pp. 48–62.

[22] G. S. Becker. "A theory of the allocation of time". In: *The economic journal* (1965), pp. 493–517.

[23] M. Beckmann, Ch. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation.* 1956.

[24] Martin Beckmann. “A continuous model of transportation”. In: *Econometrica: Journal of the Econometric Society* (1952), pp. 643–660.

[25] W. D. Berg and P. J. Tarnoff. “Research in urban traffic management”. In: *Transportation research record* 603 (1976), pp. 55–56.

[26] Filippo Bistaffa et al. “A cooperative game-theoretic approach to the social ridesharing problem”. In: *Artificial Intelligence* 246 (2017), pp. 86–117.

[27] Vitalijj Bondavev and Vasilij Shishka. “On one discussion about the development of Soviet automotive industry in the late 1920s (in Russian)”. In: *Bulletin of Perm University. Series: History* 1(18) (2012).

[28] M. Börjesson and J. Eliasson. “Experiences from the Swedish Value of Time study”. In: *Transportation Research Part A: Policy and Practice* 59 (2014), pp. 144–158.

[29] David Boyce. “Martin J. Beckmann—Robert Herman Lifetime Achievement Award in Transportation Science Transportation Science Section Operations Research Society of America April 25, 1994 Boston, Massachusetts”. In: *Transportation Science* 28.3 (1994), pp. 177–179.

[30] Dietrich Braess, Anna Nagurney, and Tina Wakolbinger. “On a paradox of traffic planning”. In: *Transportation science* 39.4 (2005), pp. 446–450.

[31] L. Breiman. “The Poisson tendency in traffic distribution”. In: *The Annals of Mathematical Statistics* 34.1 (1963), pp. 308–311.

[32] Roy F Britton. “Highway Taxation: Present Status and Probable Future Trends”. In: *The ANNALS of the American Academy of Political and Social Science* 187.1 (1936), pp. 79–87.

[33] JSP Buckland. “Thomas Savery His Steam Engine Workshop of 1702”. In: *Transactions of the Newcomen Society* 56.1 (1984), pp. 1–20.

[34] EN Burns. “Priority rating of potential park-and-ride sites”. In: *ITE journal* 49.2 (1979).

[35] LH Campbell and PG Taylor. “Renfrey Burnard Potts 1925–2005”. In: *Historical Records of Australian Science* 25.2 (2014), pp. 291–305.

[36] Bob Carnaby. “Road marking: high priority road safety, or just road maintenance?” In: *Proceedings of the Australasian road safety research, policing and education conference.* Vol. 8. 2. Monash University. 2004.

[37] Sadi Carnot. *Reflexions on the motive power of fire: a critical edition with the surviving scientific manuscripts.* Manchester University Press, 1986.

[38] Brian Caulfield. “Estimating the environmental benefits of ride-sharing: A case study of Dublin”. In: *Transportation Research Part D: Transport and Environment* 14.7 (2009), pp. 527–531.

[39] Joana Cavadas and António Pais Antunes. “An optimization model for integrated transit-parking policy planning”. In: *Transportation* 46.5 (2019), pp. 1867–1891.

[40] Joana Cavadas and António Pais Antunes. “Optimization-based study of the location of park-and-ride facilities”. In: *Transportation Planning and Technology* 42.3 (2019), pp. 201–226.

[41] Matej Cebecauer et al. “Integrating Demand Responsive Services Into Public Transport Disruption Management”. In: *IEEE Open Journal of Intelligent Transportation Systems* 2 (2021), pp. 24–36.

[42] Avishai Ceder. *Public transit planning and operation: Modeling, practice and behavior.* CRC press, 2016.

[43] United States. Bureau of the Census, Edwin D Goldfield, et al. *Statistical Abstract of the United States, 1964.* US Government Printing Office, 1964.

[44] Louis Alexandre de Cessart. *Description des travaux hydrauliques: ouvrage imprimé sur les manuscrits de láuteur. Texte.* Vol. 2. Baudouin, 1808.

[45] Nelson D Chan and Susan A Shaheen. “Ridesharing in north america: Past, present, and future”. In: *Transport Reviews* 32.1 (2012), pp. 93–112.

[46] S. K. Chang and P. M. Schonfeld. “Multiple period optimization of bus transit systems”. In: *Transportation Research Part B: Methodological* 25.6 (Dec. 1991), pp. 453–478.

[47] Charles W Cheape. *Building the American Highway System: Engineers as Policy Makers. By Bruce E. Seely.(Philadelphia: Temple University Press, 1987. xv+ 315 pp. $32.95.)* 1988.

[48] Xinyuan Chen, Zhiyuan Liu, and Graham Currie. "Optimizing location and capacity of rail-based Park-and-Ride sites to increase public transport usage". In: *Transportation Planning and Technology* 39.5 (2016), pp. 507–526.

[49] Rudolf Clausius. *The mechanical theory of heat.* Macmillan, 1879.

[50] Cecil Clements. "Errors of vision as a factor in motor car accidents". In: *British Medical Journal* 2.2397 (1906), p. 1636.

[51] Sixty seventh congress. *An Act to provide that the United States shall aid the States in the construction of rural post roads, and for other purposes.* Technical report. 1921.

[52] "Convention with Respect to the International Circulation Of Motor Vehicles". In: *American Journal of International Law* 4.S4 (1910), pp. 316–328. doi: 10.2307/2212082.

[53] Brian Corfield. "Thomas Newcomen the Man". In: *The International Journal for the History of Engineering & Technology* 83.2 (2013), pp. 209–221.

[54] Indianapolis Common Council. "December 6, 1915: Regular Session". In: *Proceedings of the Indianapolis Common Council and City-County Council* (1915).

[55] Line Courteaud and Christian Roux. "Thème, histoire et problématique en éducation à la sécurité routière". In: *Questions Vives. Recherches en éducation* 9.19 (2013), pp. 141–157.

[56] Elizabeth Deakin, Karen Frick, and Kevin Shively. "Markets for dynamic ridesharing? Case of Berkeley, California". In: *Transportation Research Record: Journal of the Transportation Research Board* 2187 (2010), pp. 131–137.

[57] Car Crash Deaths and Rates.

[58] A. C. DeSerpa. "A theory of the economics of time". In: *The Economic Journal* 81.324 (1971), pp. 828–846.

[59] JJ DeShazo. "Park and ride: the Dallas program". In: *Transport Research Board Special Report* 153 (1975), pp. 32–38.

[60] Donald R Deskins Jr. "Race, residence, and workplace in Detroit, 1880 to 1965". In: *Economic Geography* 48.1 (1972), pp. 79–94.

[61] C. E. D. Diaz. "Mode choice of inter-island travellers: analyzing the willingness of ferry passengers to shift to air transportation". In: *J. East. Asia Soc. Transp. Stud.* 9 (2011), pp. 2058–2073.

[62] Ian SJ Dickins. "Park and ride facilities on light rail transit systems". In: *Transportation* 18 (1991), pp. 23–36.

[63] Henry Winram Dickinson. *James Watt: craftsman and engineer.* Cambridge University Press, 2010.

[64] Henry Winram Dickinson. *Matthew Boulton.* Cambridge University Press, 2010.

[65] Anthony Downs. "The law of peak-hour expressway congestion". In: *Traffic Quarterly* 16.3 (1962).

[66] D. R. Drew. *Traffic flow theory and control.* Tech. rep. 1968.

[67] D. R. Drew, T. C. Meserolo, and J. H. Buhr. *Digital Simulation of Freeway Merging Operation.* Tech. rep. 1967.

[68] GD Dubelir. "City planning". In: *Saint Petersburg: Slovo (in Russian)* (1910).

[69] Grigory Dubelir. *Urban streets and paves (in Russian).* Kiev, 1912.

[70] Conrad L Dudek, R Dale Huchingson, and Donald L Woods. "Evaluation of temporary pavement marking patterns in work zones: proving-ground studies". In: *Transportation research record* 1086 (1986), pp. 12–20.

[71] M. Duhamel. *Un demi-siècle de signalisation routière: naissance et évolution du panneau de signalisation routière en France, 1894-1946.* Presses de l'école nationale des ponts et chausées, 1994. isbn: 9782859782207. url: https://books.google.fr/books?id=FRnGMgEACAAJ.

[72] Michael Duncan and Robert K Christensen. "An analysis of park-and-ride provision at light rail stations across the US". In: *Transport Policy* 25 (2013), pp. 148–157.

[73] Jacques Editeur Scientifique DURAND et al. *Les grandes liaisons routières: histoire d'un schéma.* 1972.

[74] S. M. Easa. "Urban trip distribution in practice. I: Conventional analysis". In: *Journal of Transportation Engineering* 119.6 (1993), pp. 793–815.

[75] Y Eero. “General indicators of sustainable development in Helsinki”. In: *Helsinki Information Center Statistics 24/2006.* http://www3.turku.fi/ctw/helsinki_sustainable_development_indicators.pdf. 2006.

[76] Luedke A.L. Eldridge M.O. Clark G.G. “State Highway Management, Control and Procedure”. In: *Public Roads* 1.10 (1919), pp. 29–103.

[77] DM Elsom and H Crabbe. “Implementing the national air quality strategy in the United Kingdom”. In: *WIT Transactions on Ecology and the Environment* 14 (1970).

[78] K. Epperson. *Santa Barbara Dynamic Ridesharing Pilot Program. Technical Report. SBCAG Traffic Solutions. Santa Barbara, CA, USA.* 2015.

[79] Kevin EuDaly et al. *The complete book of north American railroading*. Voyageur Press, 2009.

[80] A. W. Evans. “On the theory of the valuation and allocation of time”. In: *Scottish Journal of Political Economy* 19.1 (1972), pp. 1–17.

[81] Ardeshir Faghri et al. “Integrated knowledge-based geographic information system for determining optimal location of park-and-ride facilities”. In: *Journal of urban planning and development* 128.1 (2002), pp. 18–41.

[82] Daniel J Fagnant and Kara M Kockelman. “Dynamic ride-sharing and fleet sizing for a system of shared autonomous vehicles in Austin, Texas”. In: *Transportation* (2016), pp. 1–16.

[83] Daniel J Fagnant and Kara M Kockelman. “The travel and environmental implications of shared autonomous vehicles, using agent-based model scenarios”. In: *Transportation Research Part C: Emerging Technologies* 40 (2014), pp. 1–13.

[84] Herbert S Fairbank. “Interregional Highways: Indicated by State-Wide Highway Planning Surveys”. In: 123.456 (1940), pp. 789–777.

[85] Wenbo Fan. “Reliability analysis of stochastic park-and-ride network”. In: *Journal of Modern Transportation* 20 (2012), pp. 57–64.

[86] Wenbo Fan et al. “Bilevel programming model for locating park-and-ride facilities”. In: *Journal of Urban Planning and Development* 140.3 (2014), p. 04014007.

[87] Yan Fang and Amy Z Zeng. “Long-term collaboration mechanism for disruption recovery service in public tram systems”. In: *Procedia Computer Science* 60 (2015), pp. 1337–1346.

[88] Bilal Farhan and Alan T Murray. “Distance decay and coverage in facility location planning”. In: *The Annals of Regional Science* 40 (2006), pp. 279–295.

[89] Bilal Farhan and Alan T Murray. “Siting park-and-ride facilities using a multi-objective spatial optimization model”. In: *Computers & Operations Research* 35.2 (2008), pp. 445–456.

[90] Ezzeddine Fatnassi, Olfa Chebbi, and Jouhaina Chaouachi Siala. “Two strategies for real time empty vehicle redistribution for the personal rapid transit system”. In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013).* IEEE. 2013, pp. 1888–1893.

[91] 1956 Federal-aid highway act of 1956. Public Law 627 pages 374–402. “Public Law 627”. In: (1956), pages 374–402.

[92] L’Assemblée fédérale. *Convention internationalerelative à la circulation automobile.* Technical report. 1926.

[93] D Fichter. *Individualized automatic transit and the city, BH Sikes, Chicago, IL.* 1964.

[94] Alexander J Field. *A great leap forward: 1930s depression and US economic growth.* Yale University Press, 2011.

[95] Michael Flowers. “Beyond open data: The data-driven city”. In: *Beyond transparency: Open data and the future of civic innovation* (2013), pp. 185–198.

[96] George B Ford. “The city scientific”. In: *Engineering Record* 67.May (1913), pp. 551–552.

[97] T. J. Fratar. “Vehicular trip distribution by successive approximations”. In: *Traffic Quarterly* 8.1 (1954), pp. 53–65.

[98] Patrick Fridenson. “La société française et les accidents de la route (1890-1914)”. In: *Ethnologie française* 21.3 (1991), pp. 306–313. ISSN: 00462616, 21010064. URL: http://www.jstor.org/stable/40989400 (visited on 01/24/2023).

[99] Ann Fetter Friedlaender. *The interstate highway system: a study in public investment.* Vol. 38. North-Holland, 1965.

[100] X. Fu and Z. Juan. "Exploring the psychosocial factors associated with public transportation usage and examining the 'gendered' difference". In: *Transp. Res. Part A Policy Pract.* 103 (Sept. 2017), pp. 70–82.

[101] Masabumi Furuhata et al. "Ridesharing: The state-of-the-art and future directions". In: *Transportation Research Part B: Methodological* 57 (2013), pp. 28–46.

[102] Larry Galehouse, James S Moulthrop, and R Gary Hicks. "Principles of pavement preservation: Definitions, benefits, issues, and barriers". In: *TR News* 228 (2003).

[103] F. Garwood. "An application of the theory of probability to the operation of vehicular-controlled traffic signals". In: *Supplement to the Journal of the Royal Statistical Society* 7.1 (1940), pp. 65–77.

[104] A. V. Gasnikov et al. *Introduction to Mathematical Modeling of Traffic Flows (In Russian).* Ed. by A. V. Gasnikov. Moscow: MCCME, 2013.

[105] Hubert Gautier. *Traite de la construction des chemins... de ceux des Romains et de ceux des Modernes.* André Cailleau, 1721.

[106] Denos Gazis, Robert Herman, and Alexei Maradudin. "The problem of the amber signal light in traffic flow". In: *Operations research* 8.1 (1960), pp. 112–132.

[107] Denos C Gazis, Robert Herman, and Renfrey B Potts. "Car-following theory of steady-state traffic flow". In: *Operations research* 7.4 (1959), pp. 499–505.

[108] D. L. Gerlough. "The Use of the Poisson Distribution in Hightway Traffic". In: Connecticut, 1955, pp. 1–58.

[109] Alexander Glen. *The Highway Acts, 1862-1878, the Locomotive Acts, 1861-1878, and the General Provisions of the Turnpike Continuance Acts, 1863-1878.* Knight and Company, 1879.

[110] "GOELRO plan. RSFSR Electrification Plan. 2nd edition. (In Russian)." In: (1955).

[111] "GOELRO plan. RSFSR Electrification Plan. Report to the VIII Congress of Soviets of the State Commission on Electrification of Russia." In: *Moscow. State Technical Publishing House* (1920), p. 96.

[112] B. D. Greenshields. "The Photographic Method of Studying Traffic Behavior". In: vol. 13. Highway Research Board Proceedings. 1933.

[113] MacDonald Th. H. *Engineer as a Leader in Public Service: to Society for Promotion of Engineering Education.* 1923, p. 12.

[114] MacDonald Th. H. "Federal-Aid Accomplishments". In: *Public Roads* 3.32 (1920), p. 14.

[115] F. A. Haight. *Mathematical theories of traffic flow.* Tech. rep. Royal Institute of Technology, Division of Mathematics, Stockholm, Sweden, 1963.

[116] F. A. Haight. "The generalized Poisson distribution". In: *Annals of the Institute of Statistical Mathematics* 11.2 (1959), pp. 101–105.

[117] William F Hamilton and Dana K Nance. "Systems analysis of urban transportation". In: *Scientific American* 221.1 (1969), pp. 19–27.

[118] Timothy J Hatton and Mark Thomas. "Labour Markets in the Interwar Period and Economic Recovery in the UK and the USA". In: *Oxford Review of Economic Policy* 26.3 (2010), pp. 463–485.

[119] Carola Hein. "Maurice Rotival: French planning on a world-scale (Part I)". In: *Planning Perspectives* 17.3 (2002), pp. 247–265.

[120] D. Helbing. "Gas-kinetic derivation of Navier-Stokes-like traffic equations". In: *Physical Review E* 53.3 (1996), pp. 23–66.

[121] Robert Herman. "Theory of traffic flow; proceedings of the Symposium on the Theory of Traffic Flow, held at the General Motors Research Laboratories, Warren, Michigan, 1959". In: *International Symposium on the Theory of Traffic Flow, 1st, 1959, Warren, Michigan, USA.* 1961.

[122] Ueno Hiroya. "Models and Predictions of Automotive Industry." In: (1970).

[123] José Holguín-Veras et al. "Park-and-ride facilities in New York City: Economic analyses of alternative locations". In: *Transportation research record* 2276.1 (2012), pp. 123–130.

[124] José Holguı et al. “User rationality and optimal park-and-ride location under potential demand maximization”. In: *Transportation Research Part B: Methodological* 46.8 (2012), pp. 949–970.

[125] Sebastian Hörl and Kay W Axhausen. “Relaxation-discretization algorithm for spatially constrained secondary location assignment”. In: *99th Annual Meeting of the Transportation Research Board.* 2020.

[126] Mark W Horner and Sara Groves. “Network flow-based strategies for identifying rail park-and-ride facility locations”. In: *Socio-Economic Planning Sciences* 41.3 (2007), pp. 255–268.

[127] Mark W Horner and Tony H Grubesic. “A GIS-based planning approach to locating urban rail terminals”. In: *Transportation* 28 (2001), pp. 55–77.

[128] Donald S Howard et al. “WPA and federal relief policy”. In: (1943).

[129] Prevost Hubbard. “Efficiency of Bituminous Surfaces and Pavements under Motor-Truck Traffic”. In: *Public Roads* 1 (1919), pp. 2007–17.

[130] Xenophon P Huddy. “The Motor Car’s Status”. In: *The Yale Law Journal* 15.2 (1905), pp. 83–86.

[131] Slade Hulbert and Charles Wojcik. “Driving simulator research”. In: *Highway Research Board Bulletin* 261 (1960), pp. 1–13.

[132] L. Hultkrantz and R. Mortazavi. “Anomalies in the value of travel-time changes”. In: *Journal of Transport Economics and Policy (JTEP)* 35.2 (2001), pp. 285–299.

[133] N. V. Pershakov I. A. Gudkova K. E. Samouylov. “Batch Service Queue Model Design and Analysis (in Russian)”. In: *Proceedings of Peoples’ Friendship University of Russia. Series: Mathematics, informatics, physics* (3–4) (2007), pp. 46–52.

[134] Muhammad Faishal Ibrahim. “Improvements and integration of a public transport system: the case of Singapore”. In: *Cities* 20.3 (2003), pp. 205–216.

[135] H Inose and T Hamada. “Road traffic management”. In: *Sangyo-Tosho, Tokyo* 23 (1972).

[136] H. Inose, H. Fujisaki, and T. Hamada. “Theory of road-traffic control based on macroscopic traffic model”. In: *Electronics Letters* 3.8 (1967), pp. 385–386.

[137] H. Inose and T. Hamada. *Road traffic control (In Russian).* Moscow, 1983, p. 248.

[138] Jack H Irving et al. “Fundamentals of personal rapid transit”. In: (1978).

[139] N. A. Irwin. “The Toronto computer-controlled traffic signal system”. In: *Traffic Control.* Springer, 1965, pp. 209–218.

[140] Aubin J. “L’automobile devant les Chambres et devant l’opinion”. In: (1908).

[141] Sheldon H Jacobson and Douglas M King. “Fuel saving and ridesharing in the US: Motivations, limitations, and opportunities”. In: *Transportation Research Part D: Transport and Environment* 14.1 (2009), pp. 14–21.

[142] S. R. Jara-Díaz. “Allocation and valuation of travel time savings”. In: *Handbook of Transport Modelling* 2 (2008), pp. 363–379.

[143] Sigvald Johannesson. “Highway economics”. In: (1931).

[144] M. B. Johnson. “Travel time and the price of leisure”. In: *Economic Inquiry* 4.2 (1966), pp. 135–145.

[145] Decimus Junius Juvenal. *The satires of Juvenal.* BoD–Books on Demand, 2020.

[146] Ece Kamar and Eric Horwitz. “Collaboration and Shared Plans in the Open World: Studies of Ridesharing”. In: *the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09)* (2009).

[147] MR Keefe. “Beyond the Two-Lane Road”. In: (1939).

[148] Alvina GH Kek, Ruey Long Cheu, and Miaw Ling Chor. “Relocation simulation model for multiple-station shared-use vehicle systems”. In: *Transportation research record* 1986.1 (2006), pp. 81–88.

[149] Boris S Kerner. “Criticism of generally accepted fundamentals and methodologies of traffic and transportation theory: A brief review”. In: *Physica A: Statistical Mechanics and its Applications* 392.21 (2013), pp. 5261–5282.

[150] Tatiana Kharlamova. “Motorization of the country in the 1920-1930s: discussion experience (in Russian).” In: *News of Moscow State Technical University of Mechanical Engineering (MAMI)* 3.2 (2012), pp. 274–281.

[151] Filippov V.V. Kislyakov V.M. and Shkolyarenko I.A. “Mathematical modeling and assessment of traffic conditions of cars and pedestrians (In Russian)”. In: *M.: Transport* (1979).

[152] F. H. Knight. “Some fallacies in the interpretation of social cost”. In: *The Quarterly Journal of Economics* 38.4 (1924), pp. 582–606.

[153] A. V. Kolesnichenko and B. N. Chetverushkin. “Derivation of hydrodynamic and quasi-hydrodynamic equations for transport systems based on statistics of Tsallis (In Russian)”. In: *Preprints of Keldysh Institute of Applied Mathematics* 0 (2014), pp. 8–32.

[154] A Kornhauser et al. *Uncongested Mobility for All: New Jersey’s Area-wide aTaxi System. Princeton University. Princeton, New Jersey.* 2013.

[155] Reinhart D Kuhne. “Foundations of Traffic Flow Theory I: Greenshields’ Legacy–Highway Traffic”. In: *Symposium on the Fundamental Diagram: 75 Years (Greenshields 75 Symposium) Transportation Research Board.* 2008.

[156] Tom Kulsea Bill & Shawver. “Making Michigan Move: A History of Michigan Highways and the Michigan Department of Transportation. Lansing: Michigan Department of Transportation. p. 10. OCLC 8169232. Retrieved January 28, 2023 – via Wikisource.” In: (1980).

[157] Anna Petrovna Kuznetsova. “Automobile highways in the whirlwinds of history”. In: *CAD and GIS of highways* 2 (9) (2017), pp. 66–75.

[158] John Gray Landels. *Engineering in the Ancient World, Revised Edition: With a Revised Preface, a New Appendix, and a New Bibliography.* Univ of California Press, 2000.

[159] Andrey Viktorovich Laponov. “The Great Depression and its Overcoming”. In: *World of the New Economy* 2 (2013), pp. 34–39.

[160] Alvah R Lauer. “Improvement in highway safety”. In: *National Academy of Sciences, National Research Council, Highway Research Board Proceedings* 12 (1932), pp. 389–99.

[161] Douglass B Lee Jr. “Requiem for large-scale models”. In: *Journal of the American Institute of planners* 39.3 (1973), pp. 163–178.

[162] Karl Lehmann. “Let Us Beautify Our Good Roads”. In: *Proceedings of the Florida State Horticultural Society.* Vol. 36. 1923, pp. 11–14.

[163] N. Lenoir and I. Laplace. “Beyond traditional value-of-time: passenger behavior for multimodal door-to-door travels in the age of information technologies”. In: *European Transport Conference 2017.* 2017.

[164] Malcolm Letts. “The Sachsenspiegel and its Illustrators”. In: *LQ Rev.* 49 (1933), p. 555.

[165] Amber Levofsky and Allen Greenberg. “Organized dynamic ride sharing: The potential environmental benefits and the opportunity for advancing the concept”. In: *Transportation Research Board 2001 Annual Meeting.* 2001, pp. 7–11.

[166] Tom Lewinson and Sharon Lewinson. *Method of selective ride-sharing among multiple users along an optimized travel route.* US Patent App. 11/735,552. Apr. 2007.

[167] M. J. Lighthill and G. B. Whitham. “On kinematic waves: II. Theory of traffic flow on long crowded roads”. In: vol. 229. Proc. R. Soc. London, Ser. A. 1955, pp. 281–345.

[168] Michael James Lighthill and Gerald Beresford Whitham. “On kinematic waves II. A theory of traffic flow on long crowded roads”. In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 229.1178 (1955), pp. 317–345.

[169] Eduard Lill. *Das Reisegesetz und seine Anwendung auf den Eisenbahnverkehr mit verschiedenen auf die Betriebsergebnisse des Jahres 1889 bezugnehmenden statistischen Beilagen in Tabellen und bildlicher Form.* Spielhagen u. Schurich, 1891.

[170] Tian-Liang Liu et al. “Continuum modeling of park-and-ride services in a linear monocentric city with deterministic mode choice”. In: *Transportation Research Part B: Methodological* 43.6 (2009), pp. 692–707.

[171] Oleg Loginovsky and Alexander Shinkarev. “Development of approaches to the management and organization of traffic in large cities (in Russian).” In: *Bulletin of the South Ural State University. Series: Computer Technologies, Management, Electronics* 14(4) (2014).

[172] MB Lomonosov. *On the layers of the earth and other works on geology (In Russian).* Ripol Classic, 2013.

[173] J. Lorenz et al. "How social influence can undermine the wisdom of crowd effect". In: *Proceedings of the National Academy of Sciences* 108.22 (2011), pp. 9020–9025.

[174] C.F. Lüder. *Christian Friedrich von Lüder Oberamtmanns in Castellaune, vollständiger Innbegriff aller bey dem Strassenbau vorkommenden Fälle, samt einer vorausgesetzten Weeg-Geschichte und einem Verzeichnis der unentbehrlichen Weeg-Gesetze.* Garbe, 1779. URL: https://books.google.fr/books?id=ywN0nOlrsm8C.

[175] T MacDonald. "Four Years Of Road Building Under The Federal-Aid Act". In: *Public Roads* 3.26 (1920), p. 3.

[176] Thomas Harris MacDonald. *The Bureau of Public Roads and Its Exhibit: Supplementing Exhibit of the Bureau of Public Roads at the Brazil Centennial Exposition, Rio de Janeiro, Brazil, 1922-1923.* US Government Printing Office, 1922.

[177] Michał Maciejewski and Kai Nagel. "Towards multi-agent simulation of the dynamic vehicle routing problem in MATSim". In: *International Conference on Parallel Processing and Applied Mathematics.* Springer. 2011, pp. 551–560.

[178] Edward Manson. "The Regulation of Motor-Cars at Home and Abroad". In: *Journal of the Society of Comparative Legislation* 7.2 (1906), pp. 333–355.

[179] BURTON W Marsh. "Pedestrian research". In: *Proceedings of the 19th Annual Meeting of the Highway Research Board.* 1940, pp. 340–346.

[180] Burton W Marsh. "Traffic control". In: *The Annals of the American Academy of Political and Social Science* 133.1 (1927), pp. 90–113.

[181] A. J. Mayne. "Some further results in the theory of pedestrians and road traffic". In: *Biometrika* 41.3/4 (1954), pp. 375–389.

[182] John Loudon McAdam. *Remarks on the present system of road making: with observations, deduced from practice and experience, with a view to a revision of the existing laws, and the introduction of improvement in the method of making, repairing, and preserving roads, and defending the road funds from misapplication.* Longman, Hurst, Rees, Orme, and Brown, 1823.

[183] Dan McNichol. "Dwight D. Eisenhower: The Catalyst Who Changed America (Second of two articles)". In: *Transportation Builder* 17.7 (2005).

[184] Natasha Merat, Ruth Madigan, and Sina Nordhoff. "Human Factors, User Requirements, and User Acceptance of Ride-Sharing in Automated Vehicles". In: *International Transport Forum Roundtable on Cooperative Mobility Systems and Automated Driving.* 2016, pp. 6–7.

[185] Chistoph Maria Merki. "L'internationalisation du trafic routier avant 1914". In: *Relations internationales* 95 (1998), pp. 329–348. ISSN: 03352013, 21052654. URL: http://www.jstor.org/stable/45345118 (visited on 01/26/2023).

[186] John M Merriman and Jean-Pierre Bardos. *Aux marges de la ville: faubourgs et banlieues en France, 1815-1870.* Ed. du Seuil, 1994.

[187] Igal Milchtaich. "Network topology and the efficiency of equilibrium". In: *Games and Economic Behavior* 57.2 (2006), pp. 321–346.

[188] Martin JH Mogridge. *Travel in towns: jam yesterday, jam today and jam tomorrow?* Springer, 1990.

[189] H. Mohring, J. Schroeter, and P. Wiboonchutikula. "The values of waiting time, travel time, and a seat on a bus". In: *The RAND Journal of Economics* (1987), pp. 40–56.

[190] Vladimir Molan and Jelena Simićević. "PARK-AND-RIDE SYSTEM: URBAN PARKING MANAGEMENT POLICY." In: *International Journal for Traffic & Transport Engineering* 8.4 (2018).

[191] Thomas Murray and Frederick Steele. *System and method for facilitating ridesharing.* US Patent App. 10/464,737. June 2003.

[192] Spinga N. "L'introduction de l'automobile dans la société française entre 1900 et 1914, mémoire de maîtrise, Université Paris X-Nanterre." In: (1973).

[193] Kai Nagel and Michael Schreckenberg. "A cellular automaton model for freeway traffic". In: *Journal de physique I* 2.12 (1992), pp. 2221–2229.

[194] United Nations. "Convention on road traffic". In: *Vienna Convention on Road Traffic,* (1968).

[195] Harry E Neal. "Developments in Highway Signs and Markings". In: (1946).

[196] S. Neely. *Factors Influencing Mode Choice for Intercity Travel from Northern New England to Major Northeastern Cities*. https://escholarship.org/uc/item/6zw5v1jz. 2016.

[197] L. Von Neumann. "First draft of a report on the EDVAC". In: *The Origins of Digital Computers*. Springer, 1982, pp. 383–392.

[198] "New innovations in pavement materials and engineering: A review on pavement engineering research 2021". In: *Journal of Traffic and Transportation Engineering (English Edition)* 8.6 (2021), pp. 815–999. ISSN: 2095-7564. DOI: https://doi.org/10.1016/j.jtte.2021.10.001. URL: https://www.sciencedirect.com/science/article/pii/S209575642100101X.

[199] G. F. Newell. "Nonlinear effects in the dynamics of car following". In: *Operations research* 9.2 (1961), pp. 209–229.

[200] G. F. Newell and et al. "Queues for a fixed-cycle traffic light". In: *The Annals of Mathematical Statistics* 31.3 (1960), pp. 589–597.

[201] Anthony Patrick O'Brien. "A behavioral explanation for nominal wage rigidity during the Great Depression". In: *The Quarterly Journal of Economics* 104.4 (1989), pp. 719–735.

[202] JTTE Editorial Office et al. "New innovations in pavement materials and engineering: A review on pavement engineering research 2021". In: *Journal of Traffic and Transportation Engineering (English Edition)* 8.6 (2021), pp. 815–999.

[203] T. Okamura et al. "Passengers' attitudes to the service items of jeepneys in Metro Manila by different lifestyles". In: *J. East. Asia Soc. Transp. Stud* 10 (2013), pp. 1384–1395.

[204] C. J. Oort. "The evaluation of travelling time". In: *Journal of transport economics and policy* (1969), pp. 279–286.

[205] Jean Orselli. *Usages et usagers de la route, mobilités et accidents 1860-2008*. La Documentation Française, 2009.

[206] J Ortega, J Tóth, and T Péter. "Estimation of parking needs at Light Rail Transit System stations". In: *Conference on Transport Sciences*. Vol. 11. 2019.

[207] Jairo Ortega, János Tóth, and Tamás Péter. "Planning a park and ride system: A literature review". In: *Future Transportation* 1.1 (2021), pp. 82–98.

[208] Jairo Ortega et al. "An integrated approach of analytic hierarchy process and triangular fuzzy sets for analyzing the park-and-ride facility location problem". In: *Symmetry* 12.8 (2020), p. 1225.

[209] Jairo Ortega et al. "Using best worst method for sustainable park and ride facility location". In: *Sustainability* 12.23 (2020), p. 10083.

[210] N August Otto. "Nikolaus August Otto". In: *Encyclopædia Britannica* (2012).

[211] Vaillancourt P. et al. *PARK & RIDE SURVEY*. 2014.

[212] L.W. Page. *Roads, Paths and Bridges*. Farmer's practical library, ed. by E. Ingersoll. Sturgis & Walton, 1912.

[213] Jean Panhard. "Petite histoire de l'automobile en France". In: *Culture technique* 19 (1989), pp. 29–32.

[214] Evangelos Papadopoulos. "Heron of Alexandria (c. 10–85 AD)". In: *Distinguished figures in mechanism and machine science*. Springer, 2007, pp. 217–245.

[215] Graham Parkhurst. "Park and ride: Could it lead to an increase in car traffic?" In: *Transport policy* 2.1 (1995), pp. 15–23.

[216] Graham Parkhurst and Jeremy Richardson. "Modal integration of bus and car in UK local transport policy: the case for strategic environmental assessment". In: *Journal of Transport Geography* 10.3 (2002), pp. 195–206.

[217] SL Paveri-Fontana. "On Boltzmann-like treatments for traffic flow: a critical review of the basic model and an alternative proposal for dilute traffic analysis". In: *Transportation research* 9.4 (1975), pp. 225–235.

[218] Frederic L Paxson. "The highway movement, 1916-1935". In: *The American Historical Review* 51.2 (1946), pp. 236–253.

[219] H. J. Payne. "Model of freeway traffic and control". In: *Mathematical Model of Public System* (1971), pp. 51–61.

[220] Harold J Payne. "FREFLO: A macroscopic simulation model of freeway traffic". In: *Transportation Research Record* 722 (1979).

[221] Brendan Pender et al. "Disruption recovery in passenger railways: International survey". In: *Transportation research record* 2353.1 (2013), pp. 22–32.

[222] K. Perchonok. "The Measurement of Driver Errors". In: (1964), pp. 1–15.

[223] Kenneth Perchonok and Paul M Hurst. "Effect of lane-closure signals upon driver decision making and traffic flow." In: *Journal of applied psychology* 52.5 (1968), p. 410.

[224] I PHASE. "Pedestrian crosswalk study: accidents in painted and unpainted crosswalks". In: *Highway Research Record* 406 (1972), p. 1.

[225] Cyrille Piatecki et al. *Pourquoi conduit-on à droite?: une investigation historique dans la génèse d'un conflit des normes*. Laboratoire d'économie d'Orléans, LEO, 2006.

[226] Cristobal Pineda et al. "Integrated traffic-transit stochastic equilibrium model with park-and-ride facilities". In: *Transportation Research Part C: Emerging Technologies* 71 (2016), pp. 86–107.

[227] Louis Albert Pipes. "Car following models and the fundamental diagram of road traffic". In: *Transportation Research/UK/* (1966).

[228] E. Pogrebnyak and N. Samoilenko. "Analysis of formation methods of the correspondence matrix of the city transport network (in Russian)". In: *Municipal economy of cities, technical and scientific digest* 69 (2006), pp. 121–126.

[229] Marco Polo. *The Travels. This is a translation of a fourteenth-century manuscript in the collection of the Bibliothéque Nationale, Paris.* New York: Orion Press.

[230] Antoine-Rémi Polonceau. *Notice sur quelques parties des travaux hydrauliques*. chez Carillan-Goeury libraire des ponts et chausses, 1829.

[231] Andrea Pompigna and Raffaele Mauro. "Smart roads: A state of the art of highways innovations in the Smart Age". In: *Engineering Science and Technology, an International Journal* (2021).

[232] Isaac B Potter. *The gospel of good roads: A letter to the American farmer*. 2. League of American Wheelmen, 1891.

[233] Renfrey Burnard Potts and OLIVER RM. "Flows in transportation networks". In: (1972).

[234] I. Prigogine. "A Boltzmann-like approach to the statistical theory of traffic flow". In: *Theory of traffic flow, proceedings* (1961), pp. 158–164.

[235] Thomas Procter. *A vvorthy vvorke profitable to this whole kingdome. Concerning the mending of all high-waies, as also for waters and iron workes. By Tho: Procter Esquire.* London: Printed by E. Allde & are to be sold at his house on Lambard-hill, neere Olde-fish-streete, 1607. URL: http://name.umdl.umich.edu/A10145.0001.001.

[236] "Protocol on Road Markings, Additional to the European Agreement supplementing the Convention on Road Signs and Signals, of 1 March 1973". In: (1973).

[237] Bureau of Public Roads. *Bureau of Public Roads Annual Report*. Washington: U.S. Department of Commerce, Bureau of Public Roads; U.S. Government Printing Office, 1922.

[238] Bureau of Public Roads. *Bureau of Public Roads Annual Report*. Washington: U.S. Department of Commerce, Bureau of Public Roads; U.S. Government Printing Office, 1923, p. 3.

[239] Bureau of Public Roads. *Bureau of Public Roads Annual Report*. Washington: U.S. Department of Commerce, Bureau of Public Roads; U.S. Government Printing Office, 1929.

[240] Bureau of Public Roads. *Bureau of Public Roads Annual Report*. Washington: U.S. Department of Commerce, Bureau of Public Roads; U.S. Government Printing Office, 1930, p. 3.

[241] United States. Bureau of Public Roads. *General Location of National System of Interstate Highways: Including All Additional Routes at Urban Areas Designated in September, 1955*. Vol. 37. US Government Printing Office, 1955.

[242] Sorin RAŢIU. "The history of the internal combustion engine". In: *Annals of the faculty of engineering Hunedoara* (2003), pp. 145–148.

[243] “Regional Highway Transport Directors Meet”. In: *Engineering News-Record* 81.13 (1918), p. 599.

[244] Carlton Reid. “Hardtop History”. In: *Roads Were Not Built for Cars: How Cyclists were the First to Push for Good Roads & Became the Pioneers of Motoring*. Washington, DC: Island Press/Center for Resource Economics, 2015, pp. 110–122. ISBN: 978-1-61091-688-2. DOI: 10.5822/978-1-61091-688-2_7. URL: https://doi.org/10.5822/978-1-61091-688-2_7.

[245] Rodrigo Rezende Amaral and El-Houssaine Aghezzaf. “The Downs-Thompson paradox in multimodal networks”. In: *2nd International Conference on Traffic and Transport Engineering (ICTTE)*. City Net Scientific Research Center Ltd. Belgrade. 2014, pp. 76–85.

[246] P. I. Richards. “Shock Waves on the Highway”. In: *Operations Research* 4 (1956), pp. 42–51.

[247] Henry William Robinson. “Denis Papin (1647-1712)”. In: *Notes and Records of the Royal Society of London* 5.1 (1947), pp. 47–50.

[248] Tim Roughgarden. “On the severity of Braess’s paradox: designing networks for selfish users is hard”. In: *Journal of Computer and System Sciences* 72.5 (2006), pp. 922–953.

[249] Alexey Rubets. *The history of automobile transport in Russia (in Russian)*. Academia, 2003.

[250] T. L. Saaty. “Seven more years of queues. A lament and a bibliography”. In: *Naval research logistics quarterly* 13.4 (1966), pp. 447–476.

[251] MA Sargious and N Janarthanan. “Forecasting demand for the park-and-ride mode and determining the optimal location of stations”. In: *Canadian Journal of Civil Engineering* 10.4 (1983), pp. 695–702.

[252] R. Schakenbos et al. “Valuation of a transfer in a multimodal public transport trip”. In: *Transport policy* 46 (2016), pp. 72–81.

[253] Frank Schipper. “Unravelling hieroglyphs: Urban traffic signs and the League of Nations”. In: *Métropoles* 6 (2009).

[254] A. Schuhl. “The probability theory applied to distribution of vehicles on two-lane highways”. In: *Poisson and Traffic. Eno Foundation for Highway Traffic Control. Sangatuck, CT* (1955), pp. 59–75.

[255] John Scott-Montagu. “Automobile Legislation: A Criticism and Review”. In: *The North American Review* 179.573 (1904), pp. 168–177.

[256] Bruce E Seely. “Engineers and Government-Business Cooperation: Highway Standards and the Bureau of Public Roads, 1900–1940”. In: *Business History Review* 58.1 (1984), pp. 51–77.

[257] Dorsey Shackleford. “Federal Road Legislation”. In: *Proceedings of the third American Road Congress under auspices of American Highway Association, American Automobile Association, Detroit, Mich.* 1913, pp. 55–64.

[258] R. E. Shannon. *Systems simulation; the art and science*. Tech. rep. 1975.

[259] Y. Sheffi. *Urban transportation networks*. New Jersey: Prentice-Hall, 1985, p. 399.

[260] GV Sheleikhovsky. “Composition of city plan as a transport problem”. In: *Manuscript, Leningrad* (1946).

[261] V. I. Shvetsov. “Mathematical Modeling of Traffic Flows (In Russian)”. In: *Automatics and telemechanics* 11 (2003), pp. 3–46.

[262] L. Smith et al. *TRANSIMS: Transportation analysis and simulation system*. Tech. rep. Los Alamos National Lab., NM (United States), 1995.

[263] Robert J Spillar. “Park-and-ride planning and design guidelines, 1st ed.” In: (1997).

[264] United States. *The Statutes at Large of the United States*. C,. 37–50. U.S. Government Printing Office, 1913–1939.

[265] Peter A Steenbrink. *Optimization of transport networks*. London; New York: Wiley, 1974.

[266] Mitja Stiglic et al. “Enhancing urban mobility: Integrating ride-sharing and public transit”. In: *Computers & Operations Research* 90 (2018), pp. 12–21.

[267] HOUR-GLASS STOMACH. “justification.” In: *The Lancet* (1902), p. 684.

[268] Orest Subtelny. *Ukraine: A history*. University of Toronto Press, 2009.

[269] J. C. Tanner. “The delay to pedestrians crossing a road”. In: *Biometrika* 38.3/4 (1951), pp. 383–392.

[270] Brian D Taylor. “When finance leads planning: Urban planning, highway planning, and metropolitan freeways in California”. In: *Journal of Planning education and research* 20.2 (2000), pp. 196–214.

[271] Phathinan Thaithatkul et al. “Simulation approach for investigating dynamics of passenger matching problem in smart ridesharing system”. In: *Transportation Research Procedia* 21 (2017), pp. 29–41.

[272] William Thomson. “XV.—On the Dynamical Theory of Heat, with numerical results deduced from Mr Joule’s Equivalent of a Thermal Unit, and M. Regnault’s Observations on Steam”. In: *Earth and Environmental Science Transactions of The Royal Society of Edinburgh* 20.2 (1853), pp. 261–288.

[273] Maine Department of Transportation. *PARK & RIDE SURVEY FINAL REPORT*. Technical report. 2014.

[274] United States. Department of the Treasury. Bureau of Statistics et al. *Statistical Abstract of the United States*. Washington: U.S. Government Printing Office, 1962, p. 553.

[275] Nations Unies. *Convention sur la circulation routière*. The White House. Report to Congress. 1949.

[276] ol. 1091. Vienna. 8 November 1968 United Nations Treaty Series. “CONVENTION ON ROAD SIGNS AND SIGNALS”. In: (1968).

[277] Raymond Unwin. *Town planning in practice: An introduction to the art of designing cities and suburbs*. T. Fisher Unwin, 1913.

[278] Gregory Valiant and Tim Roughgarden. “Braess’s paradox in large random graphs”. In: *Random Structures & Algorithms* 37.4 (2010), pp. 495–515.

[279] Frank E Vandiver. *Black Jack: The Life and Times of John J. Pershing*. Vol. 1. Texas A&M University Press, 1977.

[280] Pravin Varaiya. “Freeway Performance Measurement System (PeMS), Version 3 Phase II”. In: (2004).

[281] Rudi Volti. *Cars and culture: The life story of a technology*. JHU Press, 2006.

[282] John Von Neumann. “Probabilistic logics and the synthesis of reliable organisms from unreliable components”. In: *Automata studies* 34 (1956), pp. 43–98.

[283] V. R. Vuchik. *Transport in cities convenient for living*. 2011, p. 576.

[284] Judith YT Wang, Hai Yang, and Robin Lindsey. “Locating and pricing park-and-ride facilities in a linear monocentric city with deterministic mode choice”. In: *Transportation Research Part B: Methodological* 38.8 (2004), pp. 709–731.

[285] “Wanted—A Connected National Highway System”. In: *Engineering News-Record* 81.14 (1918), p. 608.

[286] J. G. Wardrop. “Some theoretical aspects of road traffic research”. In: vol. 1. 3. Thomas Telford-ICE Virtual Library, 1952, pp. 325–362.

[287] B. Watters. “Values of travel time savings used in road project evaluation: a cross-country/jurisdiction comparison”. In: *Papers of the Australasian Transport Research Forum October 1992, Canberra, Volume 17, Part 1*. Vol. 17. 1992.

[288] Bert van Wee, Jan Anne Annema, and David Banister. *The transport system and transport policy: an introduction*. Edward Elgar Publishing, 2013.

[289] Richard F Weingroff. “Federal Aid Road Act of 1916: Building the foundation.” In: *Public Roads* 60.1 (1996), pp. 2–6.

[290] G. B. Whitham. *Linear and nonlinear waves*. Modern Book Incorporated, 1975.

[291] G. B. Whitham and G. Beresford. “Linear and Nonlinear Waves”. In: *Wiley-Interscience of waves on a running stream. Q. J. Mech. Appl. Maths* 22 (1974), pp. 221–233.

[292] J. R. B. Whittlesey and F. A. Haight. “Counting distributions for an Erlang process”. In: *Annals of the Institute of Statistical Mathematics* 13.1 (1961), pp. 91–103.

[293] A. G. Wilson. “Further developments of entropy maximising transport models”. In: *Transportation Planning and Technology* 1 (1973), pp. 183–193.

[294] A. G. Wilson. “The use of the concept of entropy in system modelling”. In: *Journal of the Operational Research Society* 21 (1970), pp. 247–265.

[295] AG Wilson. *Gravity Models in the Physical and Social Sciences. I. Some Comparisons*. Tech. rep. MAU-N-13, unpublished note, Ministry of Transport, 1966.

[296] P Wilson. "Operations of Bureau of Public Roads Under the Federal Aid Road Act". In: *Public Roads* 1.10 (1919), pp. 22–24.

[297] Jaime Wisniak. "Benoit Paul Emile Clapeyron: A Short Bibliographical Sketch". In: *The Chemical Educator* 5.2 (2000), pp. 83–87.

[298] Ka-Io Wong and Michael GH Bell. "The optimal dispatching of taxis under congestion: A rolling horizon approach". In: *Journal of advanced transportation* 40.2 (2006), pp. 203–220.

[299] Bartfeld Y. "L'assurance obligatoire dans les transports publics automobiles de voyageurs et marchandises, Paris, Domat-Montchrestien." In: (1936).

[300] S. Yan, C.-J. Chi, and C.H. Tang. "Inter-city bus routing and timetable setting under stochastic demands". In: *Transportation Research Part A: Policy and Practice* 40.7 (Aug. 2006), pp. 572–586.

[301] Hai Yang and Judith YT Wang. "Park-and-ride location and price optimisation in a linear monocentric city with logit-based mode choice". In: *Transportation in the Information Age: Proceedings of the 7th Conference of Hong Kong Society for Transportation Studies.* 2002, p. 345.

[302] Hai Yang et al. "Equilibria of bilateral taxi–customer searching and meeting on networks". In: *Transportation Research Part B: Methodological* 44 (8-9 Sept. 2010), pp. 1067–1083.

[303] Raymund T Yingling and Robert W Ginnane. "The Geneva Conventions of 1949". In: *American Journal of International Law* 46.3 (1952), pp. 393–427.

[304] Toon Zijlstra, Thomas Vanoutrive, and Ann Verhetsel. "A meta-analysis of the effectiveness of park-and-ride facilities". In: *European Journal of Transport and Infrastructure Research* 15.4 (2015).

[305] Abram Khaimovich Zilbertal. "Problems of urban passenger transport [Problemy gorodskogo passazhirskogo transporta]". In: *Moscow, Gostransizdat* (1937).

[306] Helmut T Zwahlen, Thomas Schnell, and Sarah Miescher. "Recognition distances of different pavement arrow designs during daytime and nighttime". In: *Transportation research record* 1692.1 (1999), pp. 119–128.

Thank You for Finishing the Book!

We hope this book has been a valuable resource for you.

Loved It?
Take a moment to leave a quick, honest review—just scan the QR code below with your phone, and it will take you directly to the review page on Amazon. It's fast, easy, and helps others discover this book!

� Scan to Review:

SCAN ME

Not Completely Satisfied?
We'd love to hear how we can improve! Share your feedback directly with us:
⊠ Write to us at: tatiana@babicheva.org

Thank you for being part of our journey to help create the next generation of math champions!