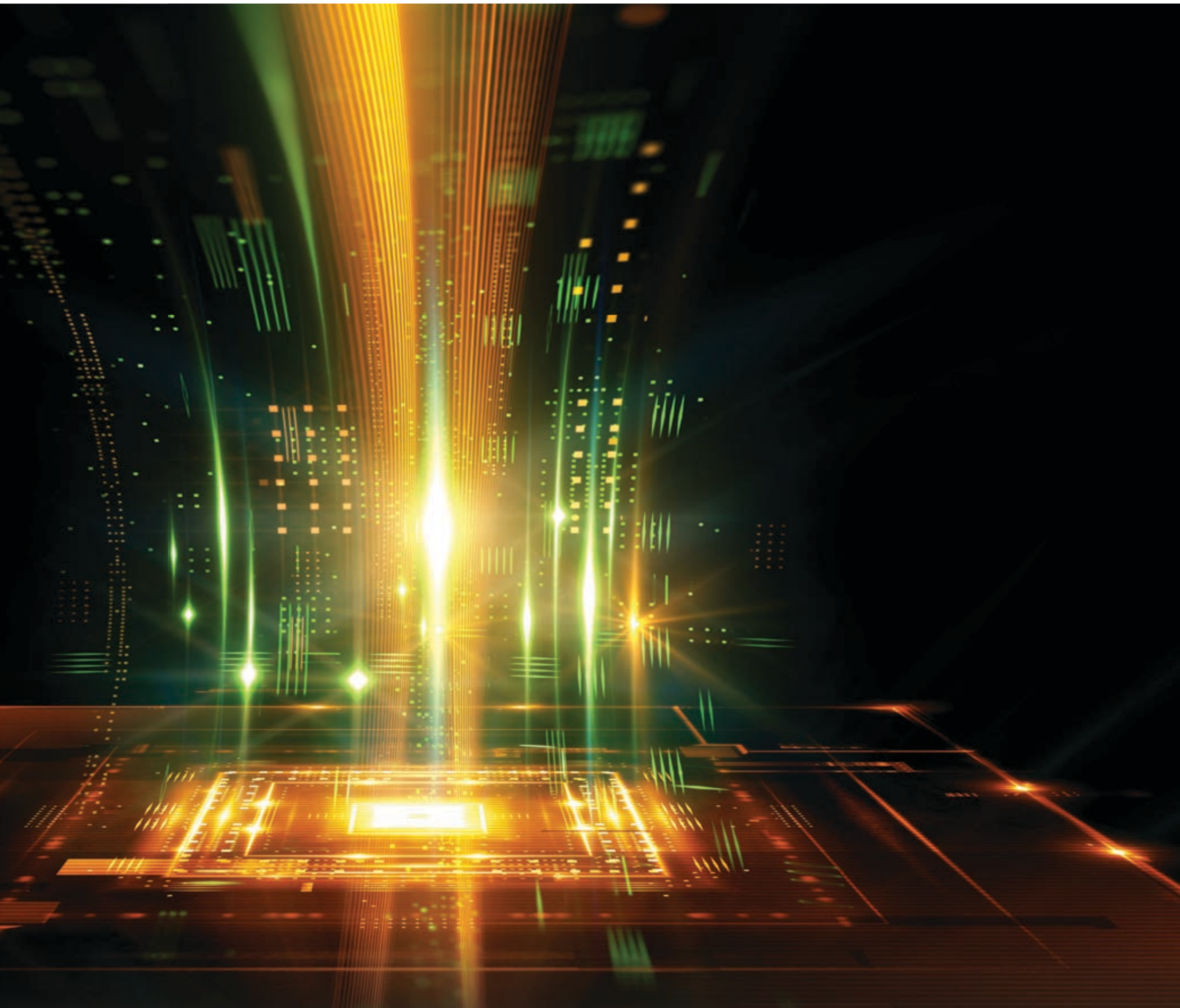


CYBER-PHYSICAL SYSTEMS FOR CLEAN TRANSPORTATION

Funded by the Erasmus+ Programme of the European Union

Cyber-Physical Systems for Clean Transportation

RTU Press

Riga 2022

Cyber-Physical Systems for Clean Transportation. Riga, RTU Press, 2022. – 391 p.

The book provides an overview of hybrid and electric cars, explains the principles of their operation, and discusses energy-saving technologies in transport, mathematical aspects of multimodal transportation, specifics of applying the project approach to the development of cyber-physical systems for clean transportation, intelligent information technologies and systems in transport, EMC related aspects of cyber-physical systems in cars, and road traffic cyber-physical systems microsimulation.

Editors:
Andrii Hnatov — Kharkiv National Automobile and Highway University
Shchasiana Arhun — Kharkiv National Automobile and Highway University
Kateryna Danylenko — Kharkiv National Automobile and Highway University

Language Editor — Daina Ostrovska
Layout Design — Ģirts Semevics
Published by — RTU Press
Phone: +371 67089123
e-mail: izdevnieciba@rtu.lv

ISBN 978-9934-22-676-2 (pdf)

Acknowledgements

This publication has been prepared with the financial support of European Union. The authors from Riga Technical University are responsible for the content of this document. The publication reflects the views only of the authors that cannot be regarded as the European Union's official position.

The book has been developed within the frame of the ERASMUS+ CybPhys 609557 project. It is a textbook that can be used by students and specialists dealing with electric and hybrid transport, electrical engineering and applied informatics. The book provides an overview of hybrid and electric cars, explains the principles of their operation, and discusses energy-saving technologies in transport, mathematical aspects of multimodal transportation, specifics of applying the project approach to the development of cyber-physical systems for clean transportation, intelligent information technologies and systems in transport, EMC related aspects of cyber-physical systems in cars, and road traffic cyber-physical systems microsimulation.

Key Action: KA2– Cooperation for innovation and the exchange of good practices

Action: Capacity building in higher education

Action Type: Joint projects

Deliverable: Thirteen electronic books for the courses in the declared directions

Funded by the
Erasmus+ Programme
of the European Union

Project Scientific Manager: Nadezhda Kunicina

Project Manager: Anatolijs Zabasta

Editors: Andrii Hnatov, Shchasiana Arhun, Kateryna Danylenko

Institution: Riga Technical University

Summary: Cyber-physical systems for clean transportation

The content of the material is aimed at specialists in electromechanics. For its proper comprehension certain knowledge is required in special courses (course unit): motor vehicle construction and layout, electrical car, electric motor drive, discrete automation units, automatic control systems. With regard to mechatronic systems for particular purpose, from the control point of view, the ways of implementing control actions are paid special attention to.

The methodology of presenting the material expects a sequence of information introduction: the purpose and classification of systems and their component parts; element configuration; the system structure and their functioning; the examples of current industrial prototype systems; conceptual technical solutions involving modern technologies.

For better mastering of the material, the book contains schematic illustrations and structural diagrams of real life machines. In the text, technical definitions of functional elements and processes are given in brackets with the names that are understandable for both electrical specialists and mechanical engineers. The text of the manual provides links to original and additional sources of information. At the end of each part, a list of questions for self-control is given. Also, at the end of the book, there is a list of acronyms and abbreviations.

Contributors

Nadezhda Kunicina, Professor, Lead Researcher of the Division of Industrial Electronic Equipment, Institute of Industrial Electronics and Electrical Engineering, Riga Technical University, Faculty of Power and Electrical Engineering, 12/1 Azenes Str., Room 503, Riga, LV 1048, Latvia, tel. +371 67089052, Nadezda.Kunicina@rtu.lv

Anatolijs Zabasta, Lead Researcher of Institute of Industrial Electronics and Electrical Engineering, Faculty of Power and Electrical Engineering, Riga Technical University, 12/1 Azenes Str., Room 503, Riga, LV 1048, Latvia, tel. +371 67089051, Anatolijs.Zabasta@rtu.lv

Leonids Ribickis, Professor, Rector of Riga Technical University, Head of Institute of Industrial Electronics and Electrical Engineering, Faculty of Power and Electrical Engineering, Riga Technical University, Scientific Head of Electromechatronics Scientific Laboratory, 1 Kalku Str., Room217, Riga, LV 1658, Latvia, tel. +371 67089300, Leonids.Ribickis@rtu.lv

Jeļena Pečerska, Lead Researcher of the Institute of Information Technology, Faculty of Computer Science and Information Technology, Riga Technical University, 2 Daugavgrivas Str., Room 431, Riga, LV 1048, Latvia, Jelena.Pecerska@rtu.lv

Andrejs Romanovs, Lead Researcher of the Institute of Information Technology, Faculty of Computer Science and Information Technology, Riga Technical University, 2 Daugavgrivas Str., Room 429, Riga, LV 1048, Latvia, Andrejs.Romanovs@rtu.lv

Andrii Hnatov, Dr. of Science, Professor of Head of Vehicle Electronics Department, Kharkiv National Automobile and Highway University, 25 Yaroslav Mudry Str., Kharkiv, Ukraine, 61002, tel. +380 667430887, kalifus76@gmail.com

Arhun Shchasiana, Dr. of Science, Professor of Vehicle Electronics Department, Kharkiv National Automobile and Highway University, 25 Yaroslav Mudry Str., Kharkiv, Ukraine, 61002, tel. +380 993780451, shasyana@gmail.com

Oleksandr Dziubenko, Ph. D., Assistant Professor of Vehicle Electronics Department, Kharkiv National Automobile and Highway University, 25 Yaroslav Mudry Str., Kharkiv, Ukraine, 61002, tel. +380 667684116, dzyubenko.alan@gmail.com

Nataliia Rudenko, Ph. D., Assistant Professor of Foreign Languages Department, Kharkiv National Automobile and Highway University, 25 Yaroslav Mudry Str., Kharkiv, Ukraine, 61002, tel. +380 683490745, rud_nat@ukr.net

Yuriy Borodenko, Ph. D., Assistant Professor of Vehicle Electronics Department, Kharkiv National Automobile and Highway University, 25 Yaroslav Mudry Str., Kharkiv, Ukraine, 61002, tel. +380 983629112, docentmaster@gmail.com

Kateryna Danylenko, Head of International Relations Office, Kharkiv National Automobile and Highway University, 25 Yaroslav Mudry St., 61002, tel: +380 57 707-36-63, international@khadi.kharkov.ua

Natalia Morkun, Dr. of Sciences, Professor, Head of Automation, Computer Sciences and Technologies Department, Kryvyi Rih National University, 11 Vitalii Matusevytch Str., Krivyi Rih, Ukraine, 50027, tel. +380 975554633, nmorkun@knu.edu.ua

Iryna Zavsiehdashnia, PhD, Associate Professor of Automation, Computer Sciences and Technologies Department, Kryvyi Rih National University, 11 Vitalii Matusevytch Str., Krivyi Rih, Ukraine, 50027, tel. +380 979010061, zavsiehdashnia@knu.edu.ua

Vladimir Sistuk, Ph.D., Assistant Professor of Automobile Transport Department, Kryvyi Rih National University, 44 Pushkin Str., 50000, tel: +38098 76114099, sistuk07@gmail.com

Yurii Monastyrskyi, Dr. of Sciences, Professor, Head of Automobile Transport Department, Kryvyi Rih National University, 44 Pushkin Str., 50000, tel: + 380 9876114099, monastirskiy08@ukr.net

Sergey Ruban, PhD, Associate Professor of Automation, Computer Sciences and Technologies Department, Kryvyi Rih National University, 11 Vitalii Matusevytch Str., Krivyi Rih, Ukraine, 50027, tel. +380 974807674, serhii.ruban@knu.edu.ua

Vitaliy Tron, PhD, Associate Professor of Automation, Computer Sciences and Technologies Department, Kryvyi Rih National University, 11 Vitalii Matusevytch Str., Krivyi Rih, Ukraine, 50027, tel. +380 961149797, vtron@knu.edu.ua

Joan Peuteman, Professor, KU Leuven, Department of Electrical Engineering, Faculty of Engineering Technology, Campus Bruges, Spoorwegstraat 12, 8200 Bruges, joan.peuteman@kuleuven.be

Contents

Chapter 1:

Shift towards an increase of (partially) electrically driven transportation

Joan Peuteman
KU Leuven

1.1. General transportation issues

When considering transportation, a distinction can be made between the transportation of humans, animals and goods. It is also possible to make a distinction between private means of transportation and public transport systems.

Additionally, a distinction can be made between transportation over land (using roads, railways, canals and rivers), transport over seas and, finally, air transport. In all these cases humans, animals and goods can be transported.

All these means of transport already exist for centuries and even millenia. Air transport (when excluding space travel) is actually the newest form of transportation. An important milestone was the use of a hot air balloon by the Montgolfier brothers in 1783, in France. Another very important milestone was the construction of the first practical aircraft in 1903 by the Wright brothers. Especially the success of the Wright Flyer was the start of an ever increasing aircraft industry.

Crucial for all means of transport is the fact that energy is needed. In ancient times, a lot of muscle power, by humans or animals, has been used. When considering ships used to transport humans or goods over sea, traditionally a combination of muscle power (by rowing) and wind power (by using sail ships) has been used.

When considering the energy sources, a lot changed during the First Industrial Revolution. The use of steam engines also had an important impact on the transportation sector. Steam ships and steam trains became very popular. Even steam driven cars have been developed. Approximately one century later the development of internal combustion engines (e.g. using gasoline or diesel) also had a huge impact on the transportation sector.

Since the First Industrial Revolution, the use of fossil fuels has been very important. Coal is burned to produce the heat which is needed to produce steam. Gasoline, diesel and other fuels are produced based on crude oil. Also natural gas is a fossil fuel which can be used to drive, e.g., a car. Unfortunately, the use of fossil fuels has a number of limitations and disadvantages. Mankind consumes fossil fuels much faster than nature can produce them, which means that the depletion of fossil fuels is a threat. By burning fossil fuels, exhaust gases have a negative impact on the environment and climate (e.g., global warming due to CO_2).

Several attempts arise to reduce the use of fossil fuels also when considering the transportation sector. Quite often these attempts are linked with the introduction of (partially) electrically driven systems. Although it is impossible to give an exhaustive overview of the use of electrical applications in the transportation systems, a number of topics are considered here which also draw the attention of the academic world. More electric aircrafts and more electric ships will be discussed. The use of electric trains is already a number of decades a common practice in a lot of countries and regions worldwide. While the use of electric trains is a mature technology, the use of electric cars and hybrid cars is actually an emerging technology.

1.2. More electric aircrafts

The propulsion mechanisms of traditional aircrafts are realised by propellers or jet engines using kerosene as energy source. But an aircraft has a large number of other loads and actuators which need to be powered. In general, they are powered by hydraulic, pneumatic and electric sources. In recent times, there is an evolution towards an increased use of electric sources and reducing the hydraulic and pneumatic sources.

By using electrical systems for the aircraft actuation systems, for environmental control systems, for fuel pumping and for wing ice protection, the airplane can become more energy efficient and quieter. By using electrically driven systems, reductions of the weight, volume and even cost of the installations can be obtained. In general, electrical systems are also reliable.

When considering an airplane, a large lifetime is needed for all components (especially in comparison with domestic, industrial and automotive applications). Airplane mechanisms must operate in a harsh environment which also imposes strict requirements on the components.

In the future, perhaps commercial all electric aircrafts will appear. In this situation, also the propulsion will be electrically driven. When considering the Solar Impulse, actually a solar powered aircraft (with batteries) already exists. Notice, however, that the Solar Implulse is an experimental airplane and still a long way is needed to realise commercial all electric aircrafts.

In case of an all electric aircraft, it is a challenge to store enough energy in a compact way (also the weight must be limited). Significant improvements in rechargeable batteries and fuel-cell technologies are needed. Since full electrically driven airplanes with battery or fuel-cell based energy storage is probably something for the far future, hybrid technologies are expected to be an important intermediate solution.

1.2.1. Electrical, pneumatic, hydraulic and mechanical systems

Consider an airplane driven by a gas turbine engine. The main goal of the engine is the propulsion of the airplane. By using a gearbox, the engine also drives an electrical generator. By converting mechanical into electrical energy a power source is available to supply a broad range of electrical loads (e.g., lighting, in-flight entertainment, etc.). By extracting high pressure air from the gas turbine, the pneumatic system is fed. The pneumatic system is used to obtain cabin pressurization, air-conditioning, etc. Using a gearbox, the engine drives a hydraulic pump allowing to use hydraulic actuators. Finally, also mechanical loads like fuel pumps and oil pumps are driven by the gas turbine engine. This approach is visualised in Fig. 1.1.

In the so-called More Electric Aircraft, the gas turbine engine will drive an electrical generator which converts mechanical power into electrical power like all electrical generators. But the electrical generator not only feeds the traditional electrical loads. For instance, also the fuel pumps and oil pumps are electrically driven. This approach is visualised in Fig. 1.2.

By using one single power source, i.e., electrical power, an increased efficiency and reduced fuel consumption can be obtained. By reducing the number of mechanical and hydraulic systems, a reduction of the weight of the airplane can be obtained. In general, electrical systems are not only reliable, they also require less maintenance.

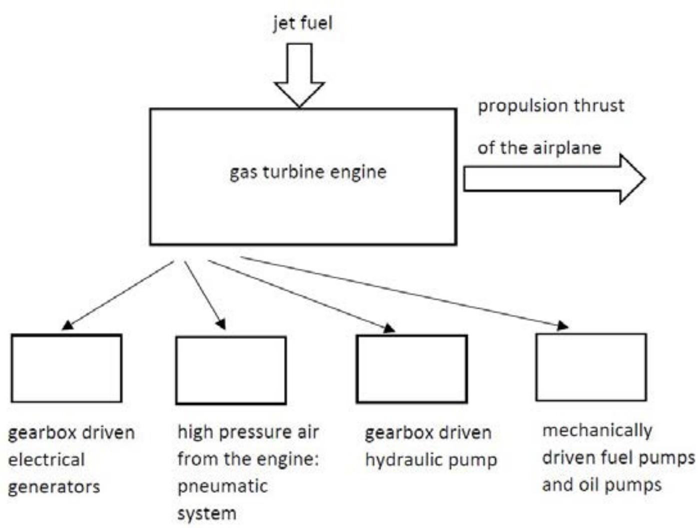


Fig. 1.1. Airplane with electrical, pneumatic, hydraulic and mechanical applications.

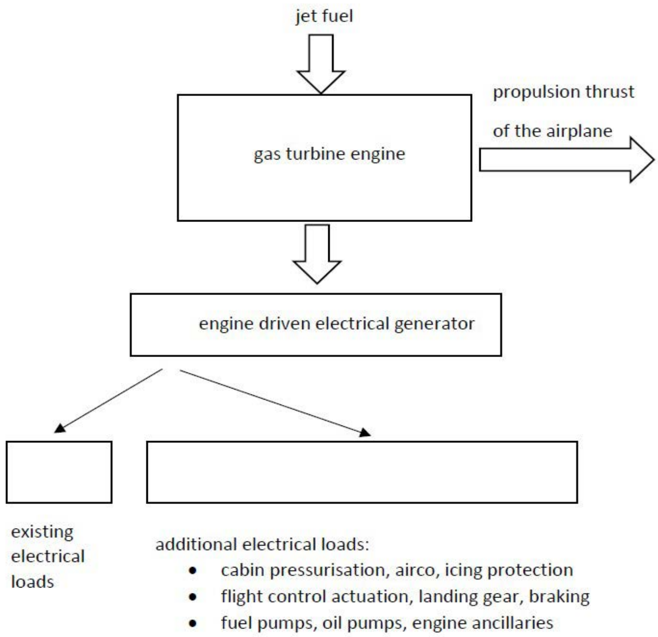


Fig. 1.2. More electric aircraft.

1.2.2. Electrical power systems

Traditionally, in an airplane a number of different voltage levels exist. For low power loads, a DC bus of 28 V DC is used. On larger airplanes, also voltage levels of 270V DC (+/– 135V DC) are used. In case of AC loads, a 115V AC system with a frequency of 400 Hz is used. By using 400 Hz instead of 50 Hz or 60 Hz, electrical machines are less heavy and smaller.

By using more electrical loads, higher voltage levels are needed to reduce the currents. By reducing the currents, heat losses in the cables are limited without increasing the cable section and the cable weight. Voltage levels of 540 V DC (+/– 270 V DC) are used. When considering AC based power distribution, a 230 V AC system with a frequency of 400 Hz is used. In order to obtain a fixed frequency of 400 Hz, a synchronous generator needs a fixed speed of rotation.

Due to the use of power electronic converters, it is also possible to use a grid with a fixed voltage level of 230 V AC in combination with a variable frequency (e.g. a frequency which ranges from 350 Hz to 800 Hz). The advantage of this approach is that a fixed speed of rotation of the synchronous generator is not required. The variable frequency is not a problem, since a power electronic converter converts the frequency to the frequency needed by the load.

1.2.3. The use of AC and DC buses

Figure 1.3 visualises a possible configuration for a More Electric Aircraft power system containing AC buses and DC buses. In Fig. 1.3, two synchronous generators (SG1 and SG2) are used which are driven by the jet engines. The first generator feeds AC bus 1, and the second one feeds AC bus 2. Notice the presence of AC loads where the Wing Ice Protection System (WIPS) is mentioned separately.

In addition to AC bus 1 and AC bus 2, also AC bus 3 is available. This AC bus 3 is used for flight critical actuation systems: for instance, electro-mechanical actuators which are driven by permanent magnet motors. These permanent magnet motors need a controlled frequency (obtained by frequency converters) to control its speed.

By using a transformer and a rectifier, the AC voltages of AC bus 1 or AC bus 2 are converted to a DC voltage. DC bus 1 and DC bus 2 are obtained allowing to feed DC loads. When using an inverter, it is possible to feed e.g. permanent magnet motors to drive actuators.

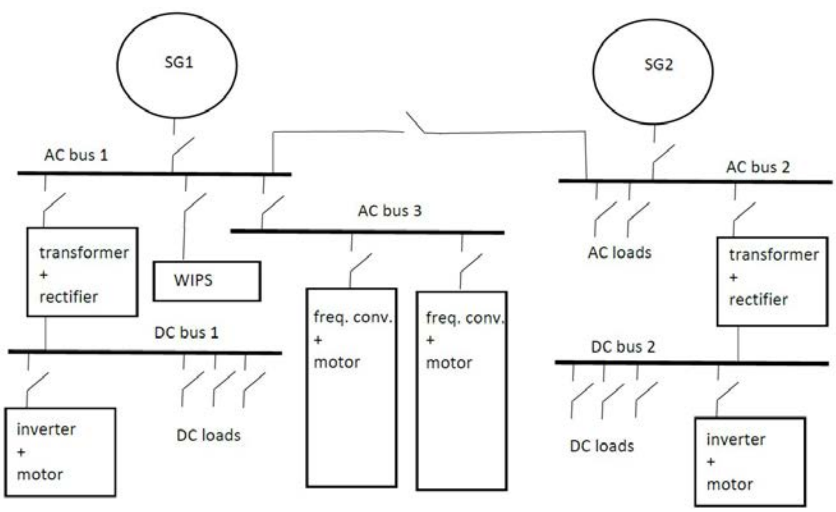


Fig. 1.3. More electric aircraft power system.

1.3. More electric ships

Traditionally, the propulsion of a ship is based on a propeller (or propellers) driven by a diesel combustion engine. Electrification not only appears on board of airplanes, but also when considering ships, there is a rise in electrical applications including electrically driven propulsion systems.

A ship propulsion motor is typically a high-torque but low-speed application. When considering electrical ship propulsion motors, AC asynchronous and AC synchronous motors (with an excitation winding) are dominant. Research is going on to develop permanent magnet motors and superconducting motors to drive the ship. It is a challenge to provide the high torque at low speeds using a compact motor having a high efficiency. Not only the development of appropriate electrical motors is important, decent power electronic converters are needed to control the speed of the motors.

Since large electrical powers (several tens of MW) are needed to drive a large ship, the voltage level of the supplying electrical grid must be sufficiently high in order to limit the heat losses in the cables. Medium voltage levels of e.g. 3.3 kV, 6.6 kV, 11 kV, 13.8 kV, 15 kV and even 20 kV are realistic.

1.3.1. Diesel-electric propulsion

Figure 1.4 visualises the basic prinicple of diesel-electric propulsion of a ship (electric propulsion based on energy storage using batteries or fuel cells is not applicable for larger applications and distances, since the energy storage needs volumes and weights which are too high). A diesel engine is driving a synchronous generator generating a three phase voltage. Using a transformer, the voltage level can be changed, and using a frequency converter, a controlled frequency can be obtained which controls the speed of the electrical motor and the propeller.

The use of diesel-electric propulsion, as visualised in Fig. 4, has a number of advantages. It is possible to have a lower fossil fuel consumption due to the possibility to optimize the loading of the diesel engines. When reducing the fuel consumption, also the exhaust of CO_2 and other harmful emissions reduces. In general, a high reliability can be obtained and maintenance costs are reduced.

It is possible to mount the diesel engine far away from the propellers, since power distribution is possible using the electrical grid. In the diesel-mechanical approach, the diesel engine must be mounted very close to the shaft of the propeller implying a limited spacial flexibility. Using an electrical motor, lower propulsion noise and less vibrations are obtained. Electrical motors in combination with

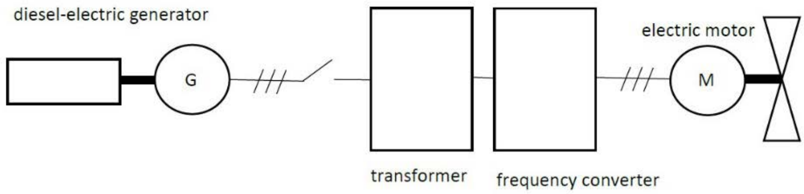


Fig. 1.4. Diesel-electric propulsion of a ship.

decent power electronic converters allow to obtain high torques at a low speed.

Notice that the use of a diesel-electric propulsion also has a number of disadvantages. The investment costs are high, since a lot of components are needed (electrical generator, electrical equipment, transformers, power electronic converters, etc.).

A large number of different motor types can be used, but AC asynchronous motors and AC synchronous motors are the most common ones. The development of permanent magnet motors aims to obtain smaller and lighter motors having a high efficiency. The development of High Temperature Superconducting Motors is also going on. Not only the development of appropriate electrical motors is important, decent frequency converters are needed to control the speed of the AC motors.

1.3.2. Frequency converters

The three most important types of frequency converters are voltage source inverters (VSI), current source inverters (CSI) and cycloconverters. Figure 1.5 visualises the main structure of a voltage source inverter. The AC grid voltage will be rectified, and using a sufficiently large capacitor, a constant DC voltage U is obtained. By using a PWM inverter, a controllable AC voltage is obtained to feed the motor (e.g. an induction motor). The PWM voltage approximates a sine shaped (three phase) voltage. The frequency and the amplitude of the approximated sine shaped voltage can be chosen in order to have the appropriate motor speed.

Two level PWM inverters are frequently used in the industry. When considering larger powers, three level and more general multi level PWM inverters are also used. By using multi level PWM inverters, the switching losses are reduced and still a decent approximation of a sine voltage is obtained. A detailed description of multi level PWM inverters is beyond the scope of the present text.

Voltage source inverters are the most frequently used type of frequency

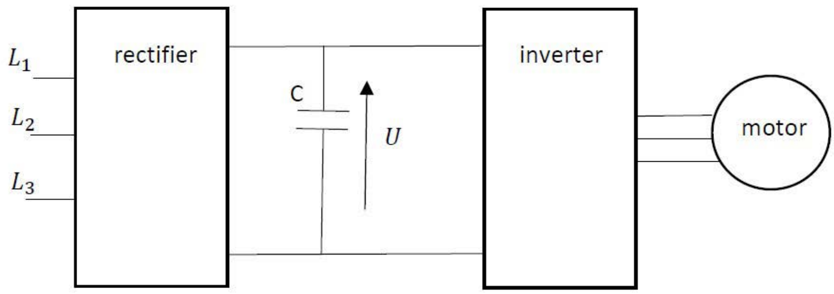


Fig. 1.5. Voltage source inverter.

converters. But especially when considering high powers, and the propulsion of a ship requires high powers, also current source inverters are used. Figure 1.6 visualises a current source inverter.

The current source inverter is also fed by a (three phase) AC grid. No constant DC voltage but a constant DC current *I* is obtained. Using an appropriate inverter, it is possible to feed the AC motor and control its speed and torque. Similar with the VSI, a detailed discussion of the CSI is beyond the scope of the present text. Notice, however, the CSI is able to recuperate energy (kinetic energy of the rotating machine is converted into electrical energy and sent back to the AC grid) with a minimum of power electronic components (also a VSI is able to recuperate energy, but more power electronic components are needed). This energy recuperation property is especially important when dealing with high power applications.

A third type of frequency converter is the cycloconverter which is also mainly used for high power applications. Figure 1.7 visualises the internal structure of such a three phase cycloconverter. Notice on the left the feeding three phase grid and on the right the electrical motor (e.g. a synchronous motor) which will be fed by a controllable frequency. The converter itself contains for each output

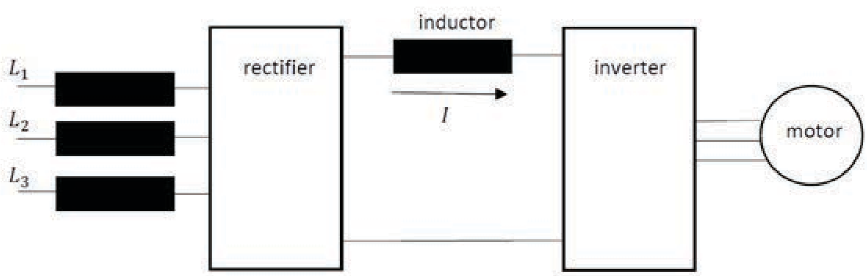


Fig. 1.6. Current source inverter.

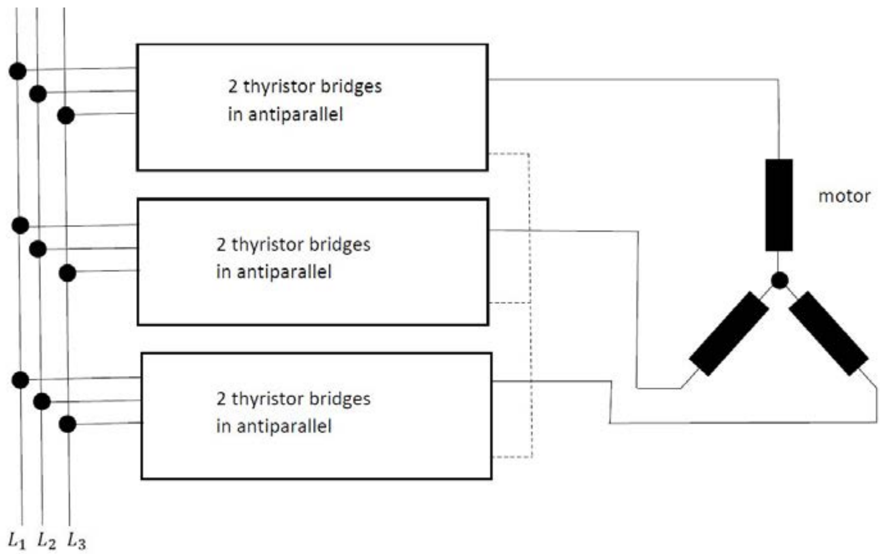


Fig. 1.7. Cycloconverter.

phase two thyristor bridges in antiparallel, which allows to exchange power in two directions. A discussion of the internal working principles of these thyristor bridges falls outside the scope of the present text.

Cycloconverters are used to obtain output frequencies which are lower than the original grid frequency (e.g. output frequencies between and 20 Hz in case the original grid frequency equals 50 Hz). This implies that the motor will have a low rotational speed which is indeed needed when considering the propulsion of a ship. With such an approach, a gearbox between the motor and the propeller can be avoided.

Cycloconverters are used for large powers to drive motors with a low rotational speed. This implies that the driving torques will be high as required for ship propulsion. In general, cycloconverters have a high efficiency, since only one energy conversion is needed (in comparison with VSI and CSI converters where two conversions are needed: AC to DC and DC to AC). Cycloconverters have the disadvantage of being complex systems. Moreover, the current extracted from the feeding grid is far from sinusoidal.

1.4. Electric trains

In the nineteenth century, trains were driven by steam locomotives. By burning coal (or another fuel), heat is obtained to produce steam. The steam is used to drive reciprocating pistons which drive the wheels of the locomotive. Since steam locomotives need a lot of maintenance and due to their low efficiency, diesel engine driven locomotives appeared in the twentieth century. Also locomotives driven by electrical motors appeared and became very popular.

Similar with the diesel-electric driven ships, also diesel-electric locomotives exist. A diesel engine drives an electrical generator, and the generated electrical power is used to feed an electrical traction motor. In the present text, only pure electric driven locomotives will be considered. The electric energy can be stored using batteries and supercapacitors, but especially electric locomotives fed by a third rail or overhead lines are dominant.

1.4.1. Power supply systems

When considering trains, trams or metros, the power supply mainly originates from a third rail or overhead lines. An exhaustive overview of all existing systems worldwide is simply impossible. In the present text, we restrict ourselves to an overview of a number of quite frequently used approaches.

First, a distinction can be made between DC and AC sources. Due to modern power electronic converters, it is realistic to feed a DC motor by a DC source, to feed an AC motor by a DC source, to feed a DC motor by an AC cource and to feed an AC motor by an AC source.

DC sources of 600 V to 1000 V are often used for metros, suburban railways and in general light rail traffic. In case larger powers are needed, 1500 VDC or 3000 VDC overhead lines are commonly used. The use of DC power sources is often a heritage from the past when mainly series excited DC traction motors were common. Traditionally DC motors were used, since their speed could be controlled using switches and resistances. The series excited DC traction motor had the additional advantage to have a large torque at low speeds which allows to accelerate the locomotive and the wagons.

There is also a variety when considering the commonly used AC sources. A typical heritage of the past is a single phase 15 kV system with a frequency of . This voltage was appropriate to feed a universal motor which is actually a series excited DC motor. Such a universal motor also operates when fed by an AC voltage. Due to problems with the commutator, a frequency of was needed instead of 50 Hz or 60 Hz.

A world standard is the use of a 25 kV single phase voltage with a frequency of 50 Hz or 60 Hz. In some really heavy duty situations a voltage level of 50 kV is used. By increasing the voltage level, the transmission efficiency increases, since lower currents are needed to provide the same power. This allows to reduce the copper losses or to reduce the cross section of the cables. On the other hand, increasing the voltage level also increases the dangers.

When using DC systems, it is more difficult to break (fault) currents, since there are no zero crossings of the current as it is the case for sinusoidal currents. Especially when voltage levels are higher, interrupting DC fault currents is a challenge. This explains the limitation to 3 kV voltage levels (although attempts existed to use higher voltage levels). When using AC systems instead of DC systems, it is not only easier to break currents. It is also possible to adapt the voltage level to the needs by using transformers. Unfortunately, the series impedance of the conductor system is higher when considering AC systems. Indeed, the impedance contains an inductive part which increases the impedance implying larger voltage drops.

The power suppy systems depend on the country (or region) and are also determined by choices made in the past. For instance, a 3 kV DC supply is used in Belgium and Latvia. In Belarus, a 25 kV AC system with 50 Hz is used. In Ukraine, both 3 kV DC and 25 kV AC systems are used depending on the region.

1.4.2. DC and AC power supplies

When using a third rail for power supply, normally a DC power supply is used. Overhead lines can be used in case of AC and DC power supplies. The power supply is used to feed the power electronic converter and the motor (the drive system of the train). The overhead line is able to provide power, since it is fed by a substation. The current provided by e.g. the overhead line flows through the pantograph of the train, provides the drive system (power electronic converter and motor) with power and flows back to the substation by making electrical contact with the axle (using axle brushes), wheels and finally the running rails. The running rails, which have the earth potential, are actually electrically connected with the substation. In this way, a closed electrical circuit is obtained.

Figure 1.8 visualises an overhead line providing a voltage of e.g. 3000 VDC. Notice that with the pantograph and using a circuit breaker the locomotive can be connected and disconnected from the overhead line. After a filter, a DC-DC converter is used to obtain a controllable DC voltage which allows to control the speed of the DC motor. The DC motor can be a DC series motor, but also a separately excited DC motor can be used (in that situation an additional power electronic converter is needed to feed the excitation winding with a controllable voltage). Finally, notice that axle brushes are used to make an electrical connection with the axle, wheels and, finally, the running rail. Figure 1.8 is very basic and more details are

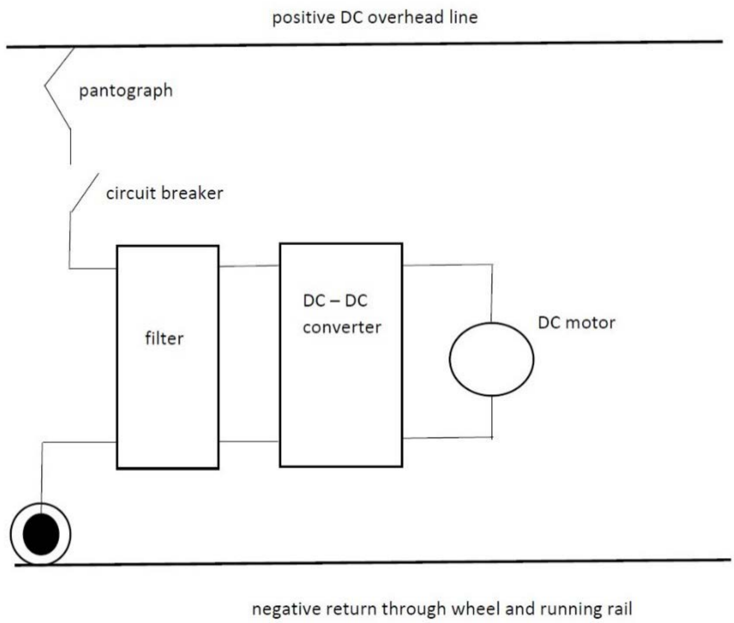


Fig. 1.8. Basic structure of a locomotive fed by a DC overhead line.

needed in real life. Notice also that an AC motor can be used by replacing the DC-DC converter by a DC-AC converter.

Figure 1.9 visualises an overhead line providing a single phase AC voltage of e.g. 25 kV. Notice again that with the pantograph and using a circuit breaker the locomotive can be connected and disconnected from the overhead line. Notice that axle brushes are used to make an electrical connection with the axle, wheels and, finally, the running rail. A transformer is used to reduce the AC voltage level, and, in general, the winding ratio of the transformer can be adapted to the needs. An AC-DC converter is used to allow a controllable DC voltage feeding the DC traction motor.

Figure 1.10 is very similar with Fig. 1.9. By using a frequency converter (e.g. a voltage source inverter can be used), a three phase voltage having an adjustable frequency and an adjustable RMS value can be obtained to control the speed of the AC traction motor.

1.4.3. Physical characteristics of the use of railway systems

The use of railway systems is common practice in a vast number of countries. Railway systems provide a number of important advantages. Due to the low

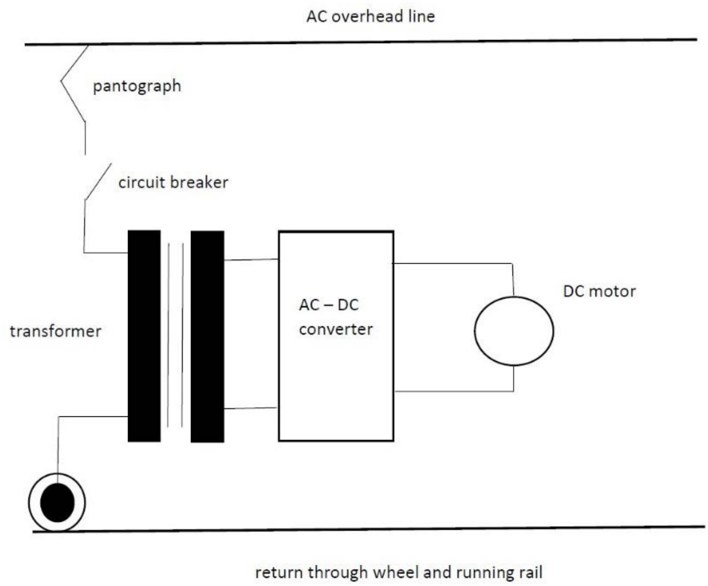


Fig. 1.9. Basic structure of a locomotive fed by an AC overhead line with a DC motor.

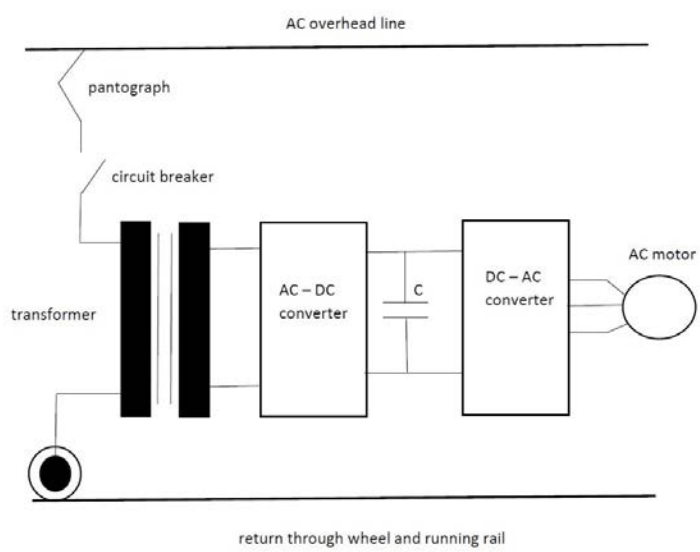


Fig. 1.10. Basic structure of a locomotive fed by an AC overhead line with an AC motor.

coefficient of friction between the wheels and the running rails, a low rolling resistance is obtained. This property allows an efficient propulsion and a high energy efficiency. By using railway systems, high speeds of operation are possible.

The low coefficient of friction between the wheels and the rails also implies some disadvantages. There is a risk of slip, the acceleration rate is limited and also the braking rate is limited. The weather has a large impact on the contact surface which influences the coefficient of friction.

When electrical locomotives are used, regenerative braking is possible. While braking the train, the traction motor operates as a generator and kinetic energy is converted into electrical energy. The generated power is injected into the overhead line and other trains can consume that power. By using overhead lines or a third rail system, no large electrical storage using e.g. batteries is needed.

A railway system uses rails, which implies that no steering is needed and the train has a predictable motion. Moreover, high standards of safety are possible. On the other hand, the rails are fixed and the train can only be used where rails are available. Trains actually have a limited network flexibility. Especially buses and cars using the road provide the advantage to have a large network flexibility. Car drivers can take the route they like and need.

1.5. Electric cars

Mainly due to the large flexibility, the use of cars is very popular. Vast majority of households in an industrialized country have one or more cars. Today, vast majority of cars use a fossil fuel like gasoline or diesel as an energy source. But around 1900, the situation was very different. For instance in the USA in 1900, 4192 cars were registrated: 1681 steam driven cars, 936 gasoline cars, and 1575 electric cars. These numbers show that electric cars were much more popular than gasoline cars. The first car to exceed a speed of 100 km/h was the electric car of Camille Janetzy (1899 in France) (https://en.wikipedia.org/wiki/La_Jamais_Contente).

After 1900, the situation changed. The discovery of new oil wells, dropping prices of gasoline (and fossil fuels in general) and the compact energy storage obtained by these fossil fuels implied a vast dominance of fossil fueled cars. Also today, cars using fossil fuels are still dominant.

In order to reduce the dependence on fossil fuels, since a number of decades there is again a growing interest in the use of electric cars. Electric cars are also an opportunity to reduce the emission of harmful exhaust gases (like CO_2). Although

Tesla (https://www.tesla.com/) is probably the best known company to develop and manufacture electric cars, a large number of (traditional) car manufacturers are also developing and selling electric cars.

Since energy storage using batteries is less compact than energy storage using fossil fuels like gasoline, also hybrid cars are available on the market. By combining a combustion engine and an electric motor (which is also able to function as a generator), the advantages of electric cars and fossil fuel powered cars can be combined. It is important to realise a sufficiently large range, to increase the energy efficiency, to reduce the exhaust of harmful exhaust gases and keep the car affordable.

When considering the transportation of humans, animals and goods, a large variety of means of transportation exist. All means of transportation have their advantages and disadvantages, and in all cases technical challenges remain.

Bibliography

A. Boglietti, A. Cavagnino, A. Tenconi, S. Vaschetto, The Safety Critical Electric Machines and Drives in the More Electric Aircraft: a Survey, 35th Annual Conference of IEEE Industrial Electronics, 3-5 November 2009, Porto, Portugal, pp. 2587–2594.

M. Chymera, C. J. Goodman, Overview of Electric Railway Systems and the Calculation of Train Performance, IET Professional Development Course on Electric Traction Systems, 5–8 November 2012, London, UK.

L. Dimitrov, S. Kanturska, Features in the selection and operation of AC Motors for electric propulsion system in ship, 15th International Conference on Electrical Machines, Drives and Power Systems (ELMA), 1-3 June 2017, Sofia, Bulgaria, pp. 228–232.

B. Gemmell, J. Dorn, D. Retzmann, D. Soerangr, Prospects of Multilevel VSC Technologies for Power Transmission, IEEE/PES Transmission and Distribution Conference and Exposition, 21–24 April 2008, Chicago, IL, USA.

T. J. McCoy, J. V. Amy jr., The State-of-the-Art of Integrated Electric Power and Propulsion Systems and Technologies on Ships, IEEE Electric Ship Technologies Symposium, 20–22 April 2009, Baltimore, MD, USA, pp. 340–344.

J. Pollefliet, Elektronische vermogencontrole, deel 1: vermogenelektronica, Academia Press, Gent, Belgium, ISBN 90 382 0657 7, 2004 (in Dutch).

J. Pollefliet, Elektronische vermogencontrole, deel 2: elektronische motorcontrole, Academia Press, Gent, Belgium, ISBN 90 382 0661 5, 2004 (in Dutch).

B. Sarlioglu, C. T. Morris, More Electric Aircraft: Review, Challenges, and Opportunities for Commercial Transport Aircraft, IEEE Transactions on Transportation Electrification, vol. 1, no. 1, june 2015, pp. 54–64.

F. Schmid, C. J. Goodman, Overview of Electric Railway Systems, 13th IET Professional Development Course on Electric Traction Systems, 3–6 November 2014, Londen, UK.

S. Soylu (Ed.), Electric Vehicles – The Benefits and Barriers, IntechOpen, Rijeka, Croatia, 2011, ISBN 978-953-307-287-6.

P. Wheeler, Technology for the More and All Electric Aircraft of the Future, IEEE International Conference on Automatica (ICA-ACCA), 19–21 October 2016, Curico, Chile.

Bibliography: websites

The Solar Impulse: https://en.wikipedia.org/wiki/Solar_Impulse

Electric locomotives for trains: https://en.wikipedia.org/wiki/Electric_locomotive

The railway Technical Website: http://www.railway-technical.com/infrastructure/electric-traction-power.html

Electric Traction Systems and their advantages: https://www.electronicshub.org/electric-traction-system/

Chapter 2:

Hybrid and electric cars

Joan Peuteman
KU Leuven

2.1. Prospects for the introduction of electric vehicles

In a lot of countries, an electrification of the fleet of cars is going on. When designing the drivetrain, several alternatives arise which include the use of different energy storage solutions. Often, a distinction is made between six types of drivetrains.

- ICE: only an internal combustion engine is used to drive the car. At present, this is the most common type of drivetrain. It is also still the most common type of drivetrain used in newly sold cars. Athough the use of an ICE alone is still dominant, the rise of other types of drivetrains imply that the market share of entirely ICE driven cars will decrease in the future.

- HEV: hybrid electrical vehicle. The drivetrain combines the use of a combustion engine and an electrical motor in order to combine the advantages of both technologies. Several approaches exist, but an important distinction is the difference between a series HEV drivetrain and a parallel HEV drivetrain. Energy is stored using fossil fuels and a battery.

- PHEV: plug-in hybrid electrical vehicle. Similar with a HEV, the drivetrain combines the use of a combustion engine and an electrical motor. It is important that the car can be connected with the electrical power grid which allows additional charging of the battery.

- REEV: range extended electrical vehicle. The car is only driven using an electrical traction motor fed by a battery. The battery can be charged using a plug-in connection with the electrical power grid. There is also an internal combustion engine which is used to drive an electrical generator which also allows to charge the battery (which feeds the electrical traction motor). Instead of a battery, also a fuel cell can be used.

- BEV: battery electrical vehicle. The car is only driven using an electrical motor fed by a battery. There is no internal combustion engine implying that the battery is charged using a plug-in connection with the electrical power grid.

- FCEV: fuel cell electrical vehicle. The car is only driven by an electrical motor, i.e., there is no internal combustion engine. Instead of a battery, a fuel cell is used to feed the electrical motor. The energy is stored using hydrogen. By making a plug-in connection with the electrical power grid, hydrogen can be produced. Alternatively, hydrogen can be bought at a hydrogen refuelling station.

At present, ICE based drivetrains are dominant but conventional HEV vehicles (e.g. the Toyota Prius) gain importance. When considering BEVs, the cars developed, manufactured and sold by Tesla are well known. But (almost) all car manufacturers perform research in the development of an increasing electrification of their drivetrains. The future of the market is unclear implying that different manufacturers have different opinions on that future. There are important differences in the enthusiasm with which the manufacturers perform reseach in the different types of drivetrains. A typical distinction is the choice between battery based or fuel cell based (using hydrogen) developments.

2.1.1. Introducing electrical vehicles to society

The electricifation of the fleet of cars in society is a long term evolution. From the point of view of consumers, buying and using these newer types of cars, today mainly the so-called ‘early EV adopters’ is important. In general, these ‘early EV adopters’ are high-income and well-educated consumers. Basically a distinction can be made between the ‘trendy greens’ and the ‘TCO sensitives’. The ‘trendy greens’ use some type of electrical vehicle because they are environmentally conscious, they are willing to try new technologies and they like to be trendy. The ‘TCO sensitives’ mainly care about the total cost of ownership (TCO). They are willing to change their travel habits in order to save money. Of course, often a combination of the ‘trendy green’ and the ‘TCO sensitives’ motivation is noticed.

In order to reduce emissions of CO_2 in the atmosphere or in an attempt to be less dependent on the use of fossil fuels, a lot of governments stimulate the electrification of the fleet of cars. Several approaches exist. For instance in urban areas, electrical vehicles have preferential parking permits or they have the permission to drive in taxi and bus lanes (especially during rush hours this accounts for a considerable time saving). The financial aspect is also important. From a technical point of view, electrical vehicles are still more expensive than cars having only an internal combustion engine. By giving subsidies when buying an electrical vehicle, the price gap can be reduced. Additionally, exemptions from purchase tax, VAT, registration tax, etc. can be an important incentive. During the life time of the car, exemptions from circulation tax, toll road charges, etc. are also important incentives. Governments can also impose fleet emission regulations, which discourages the use of traditional cars having only an internal combustion engine (or additional taxes can be imposed).

Governments can also stimulate the use of electric vehicles by providing the required infrastructure (electrical charging stations, appropriate parking spots, etc.). By stimulating the use of electrical cars, also all car manufacturers are stimulated to develop, manufacture and sell electric cars. In case a country has its own automotive industry, research and product development are important to pioneer the technology and keep the value chain in the country.

The drivers of the cars are not always the owners. An important part of the passenger cars in Europe belong to a corporate fleet. Also these companies can be stimulated to buy (more) electrical vehicles. Especialy when the cars have predictable driving patterns, the use of electrical vehicles can be a possibility. Due to the predictable driving patterns, the range requirements are known. This (partially) remedies the main advantage of a BEV which is battery fed, i.e., the limited range before recharing of the battery is needed.

As already mentioned, the total cost of ownership (TCO) of a car is very important. Governments have an impact by giving subsidies, reducing taxes, etc. But also the pure financial aspects of the technical solutions are very important. The evolution of (fossil) fuel prices and of battery pack prices (batteries are expensive, which implies that a reduction of the battery pack prices is very important) have a major impact. Charging infrastructure (slow and fast chargers) for the batteries is mandatory. Not only a standardization of the battery charging infrastructure is important, but also the cost of installing, using and maintaining the charging infrastructure has a huge impact. The price of a kWh is also important.

How fast the electric power trains will dominate the fleet of cars is unknown. But it is (almost) definite that electric vehicles will become an important part of the everyday life.

2.1.2. Overview of the main drivetrains

2.1.2.1. Drivetrain based on an internal combustion engine

At present, the majority of the cars are entirely driven by an internal combustion engine (ICE) as visualised in Fig. 2.1. An internal combustion engine is driving the wheel axle, and a transmission (gearbox) is needed, since the speed of rotation of the combustion engine is much higer than the speed needed by the wheel axle.

In order to reduce the dependence on fossil fuels and to obtain a reduction of exhaust gases (especially in urban areas a better air quality is desired), the use of this type of drivetrain will be discouraged in the future. Notice, however, that the use of the drivetrain in Fig. 2.1 also has a number of advantages. Fossil fuels are stored in a tank which allows a very compact energy storage (implying a large

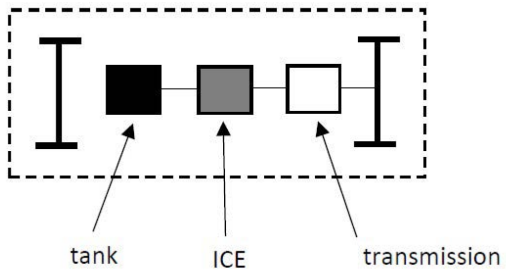


Fig. 2.1. Drivetrain based on an internal combustion engine.

range). Additionally, refueling gasoline or diesel goes very fast and there is an existing infrastructure of petrol stations (i.e. there is no dependence on electric infrastructure including battery charging installations). Research is going on to reduce the fuel consumption of internal combustion engines (including automatic start/stop), to reduce the emission of exhaust gases.

2.1.2.2. Hybrid electric vehicle and plug-in hybrid electric vehicle

Figure 2.2 visualises a plug-in hybrid electric vehicle. Notice the transmission which drives the wheel axle. The transmission is driven by an internal combustion engine and an electrical motor (a parallel configuration is visualised in Fig. 2.2). The combustion engine is the primary mover which is supported by a somewhat smaller electrical motor.

The combustion engine also drives an electrical generator. By converting mechanical energy into electrical energy a battery will be charged. The battery allows the use of an electrical motor, and a power electronic converter is used to control the speed of the electrical motor.

In the case of a plug-in hybrid electric vehicle (PHEV), the electrical power grid infrastructure allows to charge the battery in, e.g., a garage. In case the plug-in facility is not available, a hybrid electric vehicle (HEV) is obtained.

By combining an electrical motor and an internal combustion engine, a higher efficiency for the combustion engine can be obtained. A reduction of the emission of exhaust gases is also possible. Although the combustion engine is still the primary source of propulsion, fully electric driving is possible (e.g., at a lower speed or in case of a lower distance in an urban area).

Due to the fuel tank and the combustion engine, it is still possible to use the existing infrastructure of petrol stations. Due to the compact energy storage,

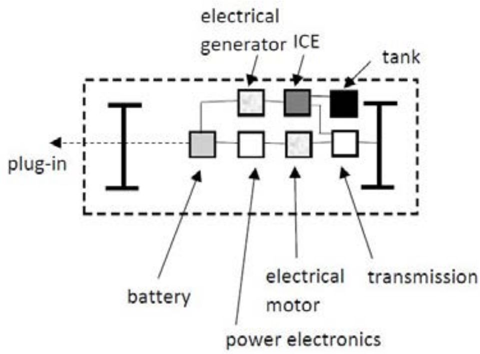


Fig. 2.2. (Plug-in) hybrid electric vehicle.

a large range is still available. This large range is not obtained in case of fully electric driving. The combustion engine is still the primary source of energy, i.e., the car still mainly relies on fossil fuels and exhaust gases are produced.

2.1.2.3. Range extended electric vehicle

Figure 2.3 visualises a range extended electric vehicle (REEV). Actually, a series hybrid configuration is obtained. Notice the presence of a fuel tank which allows the use of an internal combustion engine to drive an electrical generator. The speed of the combustion engine is not related with the speed of the electrical motor (and the speed of the car). This allows to choose an optimal working point for the combustion engine which increases the efficiency, reduces fuel consumption and reduces the emission of exhaust gases.

The electrical generator converts mechanical energy into electrical energy to allow battery charging. Battery charging is also possible by using a plug-

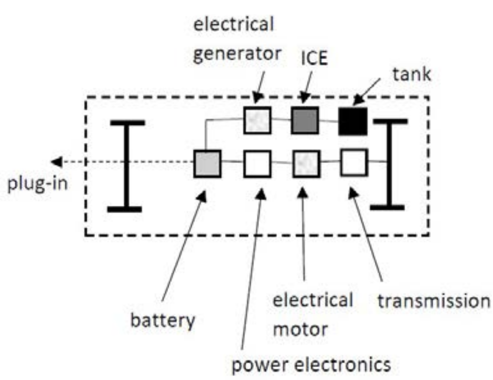


Fig. 2.3. Range extended electric vehicle.

in connection with the electrical power grid. The battery allows the use of an electrical motor, and a power electronic converter is used to control the speed of the electrical motor.

2.1.2.4. Battery electric vehicle

Figure 2.4 visualises a battery electric vehicle (BEV). A configuration without a combustion engine is obtained, i.e., pure electric traction is obtained. A large battery is needed to store a sufficient amount of energy which allows to feed the electrical motor (very often lithium-ion batteries are used). A power electronic converter is used to control the speed of the electrical motor.

It remains a challenge to store a sufficiently large amount of energy to obtain an acceptable range. Increasing the range of a BEV is still a technical challenge. Batteries are charged using slow and fast chargers fed by the electrical power grid. Even with fast chargers, charging the battery sufficiently fast remains a technical challenge. Even with fast chargers, charging times of 20 to 30 minutes must be taken into account (in case of slow charging up to eight hours of charging can be needed). At present, still a lot of countries lack charging infrastructure, and investments are needed to increase the number of chargers.

When neglecting the environmental impact of the electrical energy production, a BEV is a zero emission car. The efficiency of an electrical motor is much higher than the efficiency of an internal combustion engine. In case the electrical energy is generated using thermal power plants, the rather low efficiencies of these power plants must be taken into consideration. Actually, the final goal is generating the electrical energy using renewables like solar energy, wind energy or hydroelectric energy.

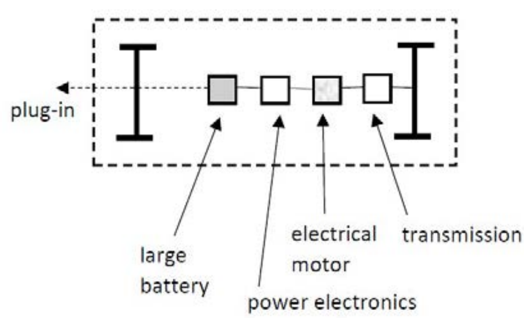


Fig. 2.4. Battery electric vehicle.

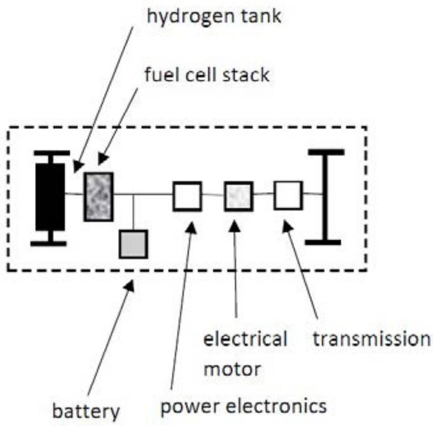


Fig.e 2.5. Fuel-cell electric vehicle.

2.1.2.5. Fuel-cell electric vehicle

Figure 2.5 visualises a fuel-cell electric vehicle (FCEV). The wheel axle is driven by an electrical motor using a transmission to obtain the required speed. The use of a Proton Exchange Membrane Fuel Cell where the fuel cell stack is fed by hydrogen is the most popular. A hydrogen tank is needed to store the hydrogen.

Also with the fuel-cell electric vehicle it is a challenge to obtain a sufficiently large range. Since hydrogen is a very thin gas, even with high storage pressures ranging from typically 350 bar to 700 bar, the energy storage capacity is limited.

In case no harmful emissions occur to produce the hydrogen, a FCEV is a zero emission car which is useful in e.g. urban areas. In comparison with the charging process of a battery, refueling with hydrogen is fast (it only takes a few minutes). At present, the infrastructure to generate, transport and distribute hydrogen is still missing in the vast majority of the countries. When considering hydrogen based infrastructure, safety issues need to be taken into account.

2.1.3. Battery charging

2.1.3.1. Lithium-ion batteries

When considering a HEV, PHEV, REEV and especially BEV, the use of batteries is very important (even in the FCEV in Fig. 2.5 some battery energy storage is considered). At present, a lot of research is going on to improve the battery technology. Batteries must be reliable and able to store a sufficient amount of energy in a compact volume (with a more or less limited weight). Batteries must have a sufficiently large round-trip efficiency, i.e., the losses while charging and discharging the battery must be limited. Batteries are quite expensive (price drops are welcome) implying also that a long life expectancy is desired.

At present, the lithium-ion battery technology is dominant, and a distinction exists between the use of small-format cells and large-format cells. The small-format cells (which are used in the BEVs of Tesla) are actually lithuim-ion batteries which are already used for consumer electronics. This implies that these small-format cells are produced at a large scale which reduces the price. An appropriate battery management system (BMS) with decent cooling is very important. When (over)-heating of the cell occurs, oxygen dissociates which can create a self-sustaining thermal reaction.

The vast majority of the BEV manufacturers use large-format cells. The energy density is lower implying that they are less vulnerable to overheating. Unfortunately, they are more expensive, since they do not benefit from the large scale production like the small-format cells do.

The high purchase price of electric car configurations still hampers their breakthrough. Infrastructure costs (e.g., installing a charging point in the garage), maintenance costs, fuel costs, insurance costs, etc. all have an impact on the total cost of ownerschip (TCO), but the price of the battery packs is really an important part of the total cost of ownership. A reduction of the battery prices is mandatory in the future.

2.1.3.2. Charging infrastructure

Petrol stations are widespread around the world allowing to fill the tank with gasoline, diesel, etc. This infrastructure is useful for drivetrains having only an internal combusion engine, for (plug-in) hybrid electric vehicles and for range extended electrical vehicles. A fuel-cell electric vehicle needs hydrogen refueling stations which are not commonly available.

Also, when considering the infrastructure needed to charge batteries, still a lot of infrastructure is missing. PHEV, BEV (and possibly also a REEV) need battery chargers. When using wired charging, plugging in of the car is often performed at an appropriate station. The appropriate charging station can be provided by the public sector, provided by private companies (e.g. intended to be used by their employers during the working hours or intented to be used by customers who pay for the service) or be owned by the car driver himself (charging at home which is commonly slow charging).

In order to obtain the situation that electric vehicles are not only used for local urban travel, i.e., to allow intercity travel, fast charging stations are needed along the highways. Due to the lack of fast charging stations along the highways, BEVs are often the second car of a household. The BEV is used for daily use when crossing shorter distances, and the 'first' car is used for longer intercity trips.

The time needed to charge a battery mainly depends on the power level extracted from the grid and the battery size (i.e. the amount of energy to be stored in the battery). As already mentioned, a distinction exists between slow and fast chargers.

Multiple types of plugs and sockets are used implying that standardisation is very important and needed. In case of slow charging, a European standard plug (Type 2 "Mennekes") is most common. In case of fast charging, three standards are important: the Tesla Supercharger, the European/US CCS "Combo" and the Japanese CHAdeMO.

The existence of charging stations with standardised plugs and sockets is very important. But also identification and payment systems must be accessible for as many drivers as possible. The use of open source protocols is needed to allow the use of the same identification, communication and payment systems for all charging stations (also if these charging systems are owned by different organisations). A cooperation between governments and private organisations can boost the development of the required infrastructure.

2.1.3.3. Alternatives to wired battery charging

Instead of using charging stations which allow to charge the batteries, batteries can also be swapped. When reaching a battery swapping station, the partially discharged battery of the car is removed and another fully charged battery is inserted. Such a fast battery swapping approach avoids the long battery charging times.

Another interesting approach is the use of wireless charging. By mounting coils in the road and feeding them with an alternating current, electromagnetic fields are generated. The energy in these electromagnetic fields are used to charge the battery of the car while driving (again avoiding battery charging times in e.g. a garage). Research is going on, and it has only been realised in a few pilot locations.

The use of plug-in hybrid electric vehicles, hybrid electric vehicles, range extended electric vehicles, battery extended vehicles and fuel-cell electric vehicles has a future. An overview of these drivetrain types has been considered. The need to have battery charging instructure is also discussed. Other chapters will discuss technical details.

2.2. Charging the batteries of electric vehicles

The rise of electric vehicles which need battery based energy storage implies the need for a decent charging infrastructure. Several charging modes exist and an overview of these modes and the technical aspects are considered.

At present, the majority of the cars (traditional cars) use an internal combustion engine to drive the car. By filling the tank of the car with gasoline or diesel (which only requires a few minutes), actually a MJ energy transfer occurs. Indeed, one liter of gasoline accounts for 34.2 MJ and one liter of diesel accounts for 35.8 MJ.

Suppose a car has a tank filled with 50 liters of gasoline. This accounts for 1710 MJ which equals 475 kWh (1 kWh = 3.6 MJ). Suppose a battery is used instead of gasoline, and the battery is charged using a single phase grid of 230 V and a current of 16 A is extracted from the grid (with a unity power factor). The battery is charged with a power of 3.68 kW implying that 129 hours are needed to store the desired 475 kWh. This numerical example shows that it is a challenge to build a charging infrastructure which charges the batteries with an acceptable speed.

2.2.1. Time needed to charge a battery

It is quite realistic to assume that an electrical vehicle consumes approximately 0.2 kWh to drive 1 km. In case of a 1 kW charger, 1 hour charging is needed to provide a range of 5 km, i.e., a charging speed of 5 km/h is obtained.

Suppose the charging infrastructure is fed by a single phase AC grid with an rms voltage of 230 V. Suppose a 16 A current (with a unity power factor) is extracted from the grid. Charging occurs with a power of 3.68 kW implying a charging speed of 18.4 km/h. Actually this charging procedure can be considered to be slow.

The so-called ‘semi-fast’ charging can occur in case of a single phase AC grid of 230 V with a current of 32 A. A charging power of 7.36 kW implies a charging speed of 36.8 km/h. A three phase grid with a 400 V line voltage and a current of 16 A implies a charging power of 11.1 kW which gives a charging speed of 55.4 km/h. When using the same three phase grid with a line voltage of 400 V, a current of 32 A allows a charging power of 22.2 kW. In that situation, a charging speed of 110.9 km/h is obtained.

Fast charging assumes charging powers which exceed 50 kW. In such a situation, charging speeds which exceed 250 km/h are obtained. Especially this 'fast charging' approach needs an appropriate charging unit in combination with a power grid which is able to supply these powers.

2.2.2. Charging batteries of an electric vehicle

A distinction is made between four charging modes.

- Mode 1: Slow charging from a regular 16 A socket (single phase or three phase). Charging mode 1 is simple and cheap and applicable everywhere (from a technical point of view). Since no protection device is mounted into the cable, the electrical installation must take care of the safety regulations (e.g. overload protection and earth leakage protection). In case the battery of a car must be charged, the rectifier and the battery charger are mounted in the car. Notice, however, that in a lot of countries 'mode 1' charging is not allowed for electric vehicles.

- Mode 2: Slow charging from a regular socket (single phase or three phase). AC currents up to 32 A are possible, but quite often the maximally consumed current is lower (e.g. 10 A). Charging mode 2 is simple and cheap and applicable everywhere. A protection device is built into the cable. The ICCB (In Cable Control Box) contains an overcurrent protection and a residual-current circuit breaker. The rectifier and the battery charger are mounted in the car (there is an AC connection with the car). Notice, however, that slow charging of the battery occurs, i.e., loading the battery can take several hours.

- Mode 3: Slow or fast charging using a dedicated EV multi-pin socket. The socket is a fixed part of the electrical installation. The multi-pin socket contains control and protection functions (overcurrent protection and protection against residual currents). The control function realises communication between the socket and the car in order to allow a safe charging process. The rectifier and the battery charger are mounted in the car (there is an AC connection with the car). When applicable, there is an option to charge the battery during the night when electrical energy is cheaper.

- Mode 4: Fast charging using special charger technology (e.g. CHAdeMO). The rectifier and the battery charger are integrated in the charging point. This implies that no rectifier and charger are needed inside the car, i.e., there is a DC connection with the car. A mode 4 charger is expensive implying that it is normally not available at home. Mode 4 chargers are typically available in public parking lots. The large powers also require an electrical grid which is able to provide these powers (e.g. a three phase grid with 400 V line voltage able to provide currents of 80 A). When using lithium-ion batteries, fast charging does not have a negative impact on the life expectancy of the battery (e.g. by loading a 24 kWh battery with a power between 36 kW and 48 kW no problems arise, in case the power is too high, still a decrease of the life expectancy will occur).

2.2.2.1. Charging procedure with proximity pilot and control pilot

Figure 2.6 visualises charging infrastucture used to charge the batty of an electric vehicle. Notice the charging point with a socket outlet which allows to plug-in a cable. The other side of the cable contains a connector which allows to connect with the vehicle inlet.

When considering mode 3 charging, Fig. 2.7 visualises the conductors of a dedicated EV multi-pin socket/plug. Notice the neutral conductor N and the phase conductors L1, L2, L3 (in case of three phase charging) which allow to transport the AC power (energy) from the socket/plug to the connector/inlet of the electrical vehicle used to charge the battery. Notice also the control pilot CP

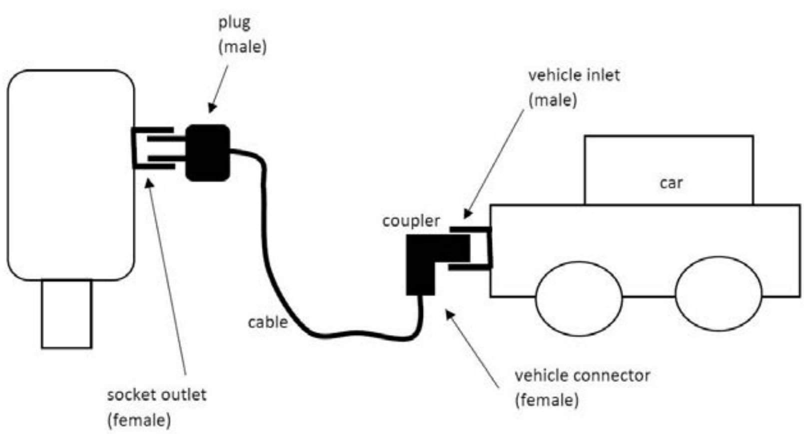


Fig. 2.6. Charging infrastructure.

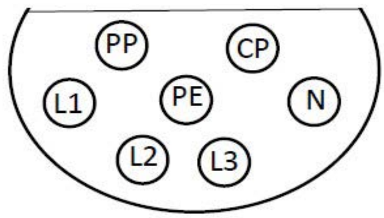


Fig. 2.7. Conductors of a dedicated EV multi-pin socket/plug.

and the proximity pilot PP as well as the protective earth conductor PE.

Figure 2.8 visualises the connection between the socket/plug and the connector/inlet of the car with the cable. In the present text, we will restrict ourselves to the basic CP and PP applications. Between PP and PE, a resistor R_{pp} has been mounted in the plug. The resistance of this resistor R_{pp} gives information concerning the maximum current which is allowed in the cable.

The proximity pilot PP is not a wire run in the cable (contrary to the control pilot CP which is a dedicated wire in the cable), only a resistor R_{pp} in the plug is connected between PP and PE. The larger R_{pp}, the smaller the allowed current in the cable, as listed in Table 2.1.

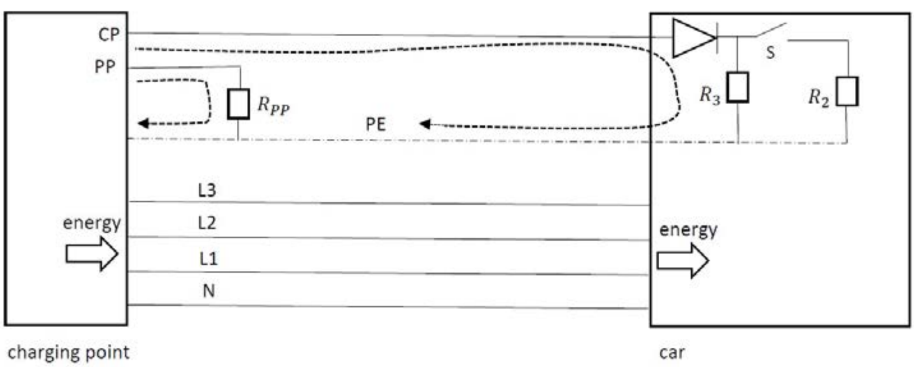


Fig. 2.8. Connection between charging point and car.

The CP is used to communicate with the electric car as visualised in Figs. 2.8 and 2.9. The charging point contains an oscillator generating a PWM voltage with an amplitude of 12 V. Resistance $R_1 = 1$ kΩ. In case the vechicle is not connected with the cable, resistances R_2 and R_3 are not present implying that the voltage measurement provides $V_a = 12$ V. In case the vehicle is connected with the cable and switch S is open, a series connection of $R_1 = 1$ kΩ and $R_3 = 2.74$ kΩ is obtained. When neglecting the voltage drop across diode D, $V_a \cong 9$ V is obtained. An open switch S is an indication

Table 2.1. PP Resistor Indicating the Allowed Cable Current

Allowed cable current, A	PP resistance R_{pp}, Ω
13	1500
20	680
32	220
63	100

from the vehicle that the vehicle is not ready to charge the battery. When the vehicle is ready to charge the battery, switch S will be closed, which implies that $R_2 = 1.3$ kΩ is connected in parallel with R_3. In that situation, $V_a \cong 6$ V implying that the battery can be charged (in case $R_2 = 270$ Ω, $V_a \cong 3$ V, which means the battery is ready to be charged and ventilation is needed during the charging process).

2.2.2.2. Mode 4 fast charging approaches

When considering fast charging, mainly four standards are important:

- ChAdeMo (CHArge de MOve) used by Nissan, Peugeot, Citroën, Kia, etc.
- CCS (Combined Charging System) used by BMW, GM, Daimler, Volkswagen, Ford, etc.
- Tesla Superchargers used by Tesla
- Guobiao recommended standard 20234 (China)

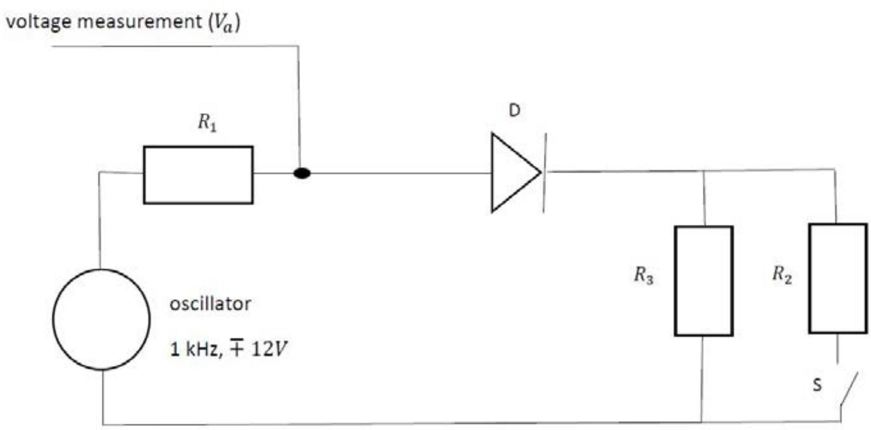


Fig. 2.9. Working principle of the control pilot CP.

CHAdeMO is a Japanese standard using a DC connection to charge the battery of the vehicle. In case of a 500 VDC voltage with a current of 125 A, a power transfer of 62.5 kW can be realised. The newer CHAdeMO 2.0 version allows a power transfer of 400 kW by using 1000 VDC and 400 A.

The CCS standard also allows high-power DC fast charging. Powers between 80 kW and 350 kW are realistic. There also exists a CCS 2.0 version, and a CCS 3.0 version (which will be backwards compatible) is expected. The CCS 3.0 version is intended to allow reverse power transfer, inductive charging and wireless charging communication.

The Tesla Supercharger allows powers between 72 kW and 250 kW (Supercharger V3) using a voltage level of 480 VDC. The Tesla Supercharger allows to charge Tesla Model S, Tesla Model X and Tesla Model 3.

2.2.3. Charging a lithium-ion battery

With some simplification, charging a battery contains two important phases, as visualised in Fig. 2.10. During the first phase, the battery will be charged using a constant charging current. This constant charging current is high implying a battery charging process which is as fast as possible. Having reached a specified voltage level, the battery is further charged by applying a constant voltage level. As the state of charge of the battery increases, the charging current will decrease.

The transition from constant current charging to constant voltage charging usually occurs when the battery reaches a state of charge which equals 80 % of the total battery capacity. Below a state of charge of 80 %, the battery will be charged as fast as possible. Above the state of charge of 80 %, the charging speed gradually decreases.

In case of fast charging, only the first phase of the charging process is considered. In case the charging current is not too high, no damage occurs to the battery. Despite this observation, it is really bad for the battery if only fast charging is used. It is important that from time to time also (slower) charging with a constant voltage is performed.

2.2.4. Alternatives to battery charging

Since charging a car battery requires quite a lot of time (even in the case of ‘fast’ charging), alternatives which allow to save time are welcome. Two typical alternatives are battery swapping and wireless battery charging.

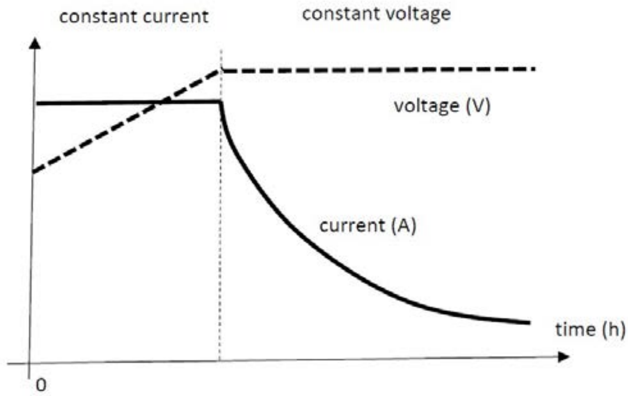


Fig. 2.10. Charging a battery.

2.2.4.1. Battery swapping

When the state of charge of a car battery is low, the battery can be replaced by a fully charged battery (with a high state of charge) which allows the driver to continue its itinerary. This approach accounts for a number of advantages.

- Battery swapping requires only a few minutes.
- After swapping the battery, a fully charged battery is available. When sufficient battery swapping stations are available, actually an unlimited driving range is obtained.
- The driver does not own the battery. The battery swapping company owns the batteries, and the company is responsible for charging the batteries. Taking care of the battery life, providing the charging infrastructure and providing a sufficient number of batteries are the responsabilities of the battery swapping company.
- The battery swapping company can also discharge a number of batteries and inject power into the public power grid. By combining grid to vehicle (G2V) and vehicle to grid (V2G) power exchanges, the Smart Grid philosophy can be supported. When there is an excess of electrical power in the grid (e.g., during the night when the power consumption is low in combination with a large power production of, e.g., wind turbines), the battery swapping company consumes (cheap) power and charges the batteries. Where there is a lack of electrical power in the grid (e.g., during the day when the power consumption is high in combination with only limited power production by, e.g., wind turbines), the battery swapping company injects (expensive)

power in the grid by discharging batteries.

When considering battery swapping, there are also a number of disadvantages.

- A sufficient number of battery swapping stations must be available to provide the desired 'unlimited driving range'.

- The investment costs to provide the required battery swapping stations are high (charging infrastructure, a sufficient number of batteries, swapping infrastructure, etc.). Finally, all costs made by the battery swapping company must be billed to drivers.

- Standardization of the swapped batteries is mandatory but not obvious. For each car, more than one identical battery is needed.

2.2.4.2. Wireless charging

Wireless charging is based on electromagnetic induction. A primary coil is placed on the surface or beneath it. By sending an AC current in the primary coil, an electromagnetic field is obtained which allows to transfer energy to a secondary coil. Voltages are induced in the secondary coil which is mounted under the vehicle. This means the primary coil is the transmitter which is fed by a power supply. The secondary coil is the receiver and using a system controller, it is possible to charge the battery.

The use of wireless charging, also called inductive charging, has a number of advantages:

- There is no physical contact with the verhicle avoiding wear. Since there is no physical contact, high safety properties are obtained.

- Flexible power ratings are possible.

- The approach is weather resistant. Even the presence of snow, dust, sand, etc. has no negative impact.

- Ideas exist concerning the realisation of wireless charging while the car is in motion (dynamic charging) by integrating several primary coils inside the paving.

The use of wireless charging also accounts for a number of disadvantages:

- Although there is no wired connection with the car, a fixed grid connection is needed with the primary coil.

- The transfer efficiencies are quite low. Research is going on to increase the efficiency (using thin coils, using sufficiently high frequencies, optimizing drive electronics).

- The infrastructure needed to realise wireless charging is expensive in comparison with traditional wired charging.

- A breakthrough of the system requires a sufficient number of wireless charging stations and a sufficient number of induction-recharge ready electrical vehicles. At present, both are uncommon.

2.2.5. Battery charging and the integration of renewables

An increased use of electrical vehicles implies an increased need for electrical energy. In case this electrical energy has been generated by e.g. thermal power plants using fossil fuels, they will emit CO_2. In case the additional emissions of CO_2 are the same or larger than the CO_2 emissions of the combustion engine, the transition from combustion engine driven to electrically driven vehicles is not useful or it is even harmful (from the point of view of the total CO_2 impact on the climate). This implies that it is useful to generate the electrical energy in an environmentally friendly way.

Integration of photovoltaic panels in the electrical vehicle charging system is an option. The price reductions of the photovoltaic panels imply that this option also becomes possible from an economical point of view. Moreover, photovoltaic systems do not require a lot of maintenance (e.g. no moving parts).

A distinction can be made between grid connected systems and standalone systems (no connection with an existing power grid). Standalone systems are used e.g. in remote areas where no grid is available. In more densely populated areas, grid connected systems are commonly used.

An increased use of electric vehicles, with batteries charged with energy originating from the public power grid, is an additional load for that electrical grid. This is the case especially during the day when the electrical power consumption is already high. In case the power needed by the charging points originate from photovoltaic panels, no additional load of the power grid occurs.

Instead of mainly charging the batteries during the night, it is an option to charge the batteries mainly during the day. During the day, the electric vehicle is parked idly in the parking area implying that time is available for charging. Especially in hot climate countries having a lot of sunlight, the photovoltaic panels provide the needed power. When the panels provide more power than needed by the batteries, the excess of power can be injected into the grid when the electricity tariff is high.

In case the solar irradiance is small, the photovoltaic panels are not able to supply the power needed by the battery charging infrastructure. The power grid must provide part of the power which is an additional load for that grid.

2.2.5.1. Grid connected battery charging infrastructure integrating photovoltaic panels

A typical battery charging system which integrates photovoltaic panels is visualised in Fig. 2.11. The installation contains three main components:

- A DC-DC converter which contains a maximum power point tracker (MPPT) ensures that the photovoltaic array provides the maximum power with the available solar irradiance.

- A bidirectional charger which allows to charge or to discharge the battery of the electric vehicle.

- A bidirectional DC to AC inverter allows to inject power into the public AC power grid. In case charging the battery needs more power than generated by the photovoltaic array, power can be extracted from the grid.

Notice the presence of a DC bus which is connected with the DC-DC converter, the bidirectional charger and the DC to AC inverter. A computer system or microcontroller functions as a central controller. The central controller determines the power flows of the DC-DC converter, the bidirectional charger and the DC to AC inverter.

2.2.5.2. Modes of operation

Considering the charging infrastructure in Fig. 2.11, mainly five modes of operation exist. In Fig. 2.12, the first mode of operation is visualised where all power generated by the photovoltaic array is injected into the battery. No power exchange with the public power grid occurs. In case the solar irradiance changes,

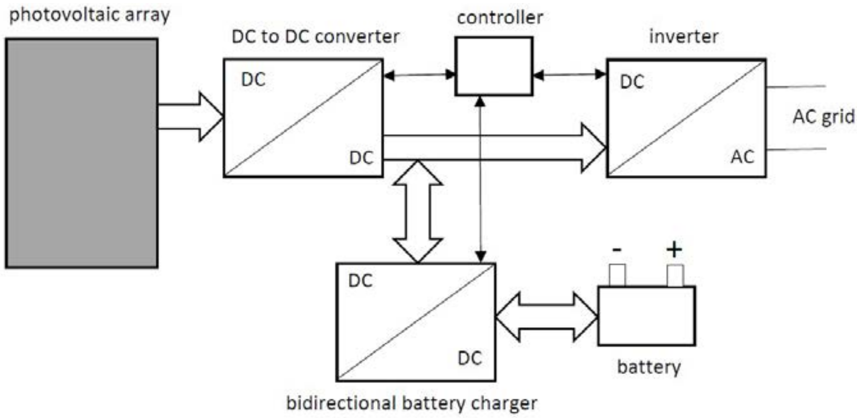


Fig. 2.11. Grid connected battery charging infrastructure integrating photovoltaics.

the generated power changes and the power injected in the battery changes. The DC to DC converter and the battery charger are used.

Figure 2.13 visualises a mode of operation where all power generated by the photovoltaic array is injected into the battery. Additionally, also the grid contributes to the battery charging power. The power derived from the grid depends on the power generated by the photovoltaic array. Since the solar irradiance changes, also the power generated by the photovoltaic array changes. By monitoring the power originating from the photovoltaic array, the power originating from the grid can be adjusted to ensure the required power/energy for the battery of the electrical vehicle.

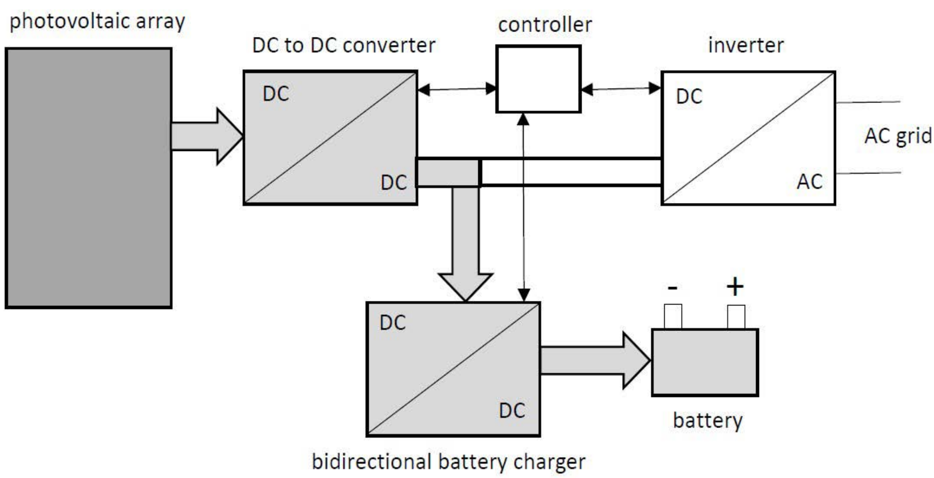


Fig. 2.12. Battery charging using energy from photovoltaic panels.

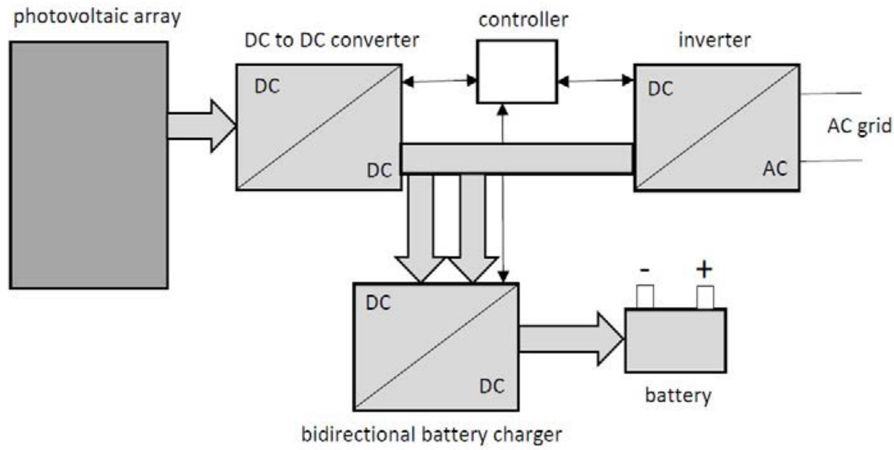


Fig. 2.13. Battery charging using energy from photovoltaic panels and the grid.

In case the photovoltaic array is not able to generate power (e.g. during the night), it is still possible to charge the battery using the grid connection. The AC power is converted into DC power. The DC voltage is converted by the battery charger in order to obtain the voltage suitable to charge the battery (see Fig. 2.14).

Figure 2.15 visualises a situation where no electric vehicle is available. The power generated by the photovoltaic array will be entirely injected into the AC power grid. Even if a vehicle is available, it can be useful to use this mode of operation. This situation occurs when the battery is already fully charged or when the feed-in-tariff for the power grid is sufficiently high. This high feed-in-tariff implies the decision to sell the electrical energy instead of charging the battery.

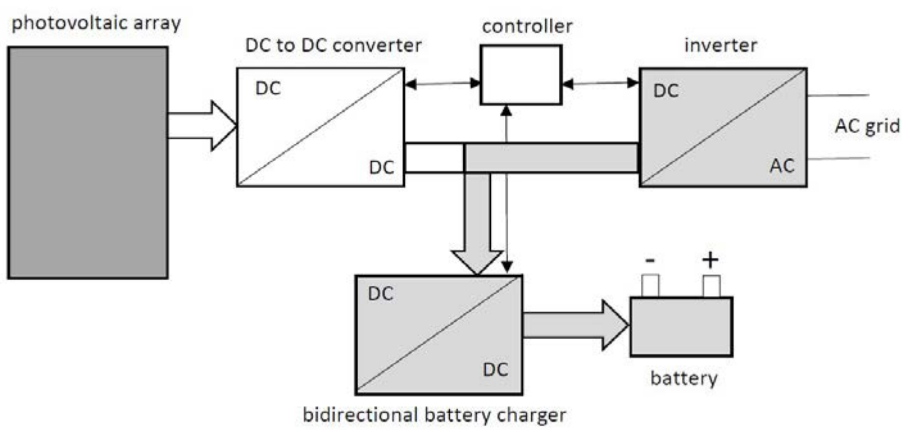


Fig. 2.14. Battery charging using energy from the grid.

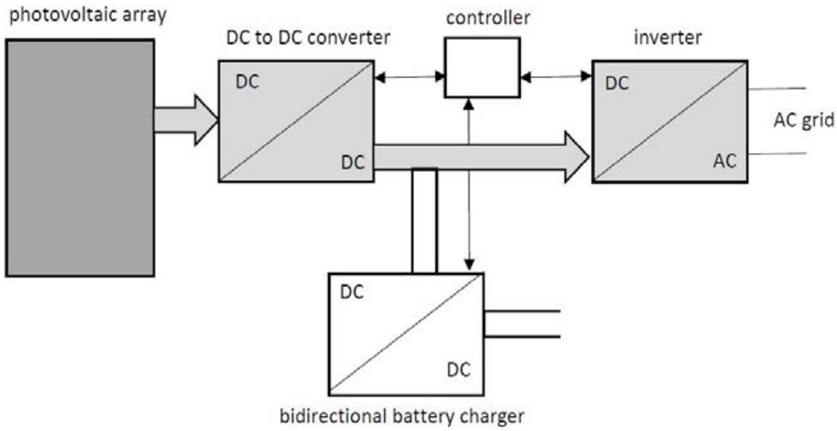


Fig. 2.15. Injection of all power into the AC power grid.

Finally, Fig. 2.16 visualises a situation where no power is generated by the photovoltaic array. Instead of charging the battery of the car, the battery is discharged and the power is injected into the AC power grid. This Vehicle to Grid mode (V2G) allows to support the power balance of the grid. This situation will mainly occur during the day when the feed-in-tariff of electrical power is high.

2.2.5.3. Photovoltaic standalone charging infrastructure

In case there is no connection with a power grid, a standalone system is obtained. Using the power of a photovoltaic array, it is still possible to charge the battery of a vehicle. It is possible to realise a connection (without additional storage) between the photovoltaic array and the electric vehicle (containing a

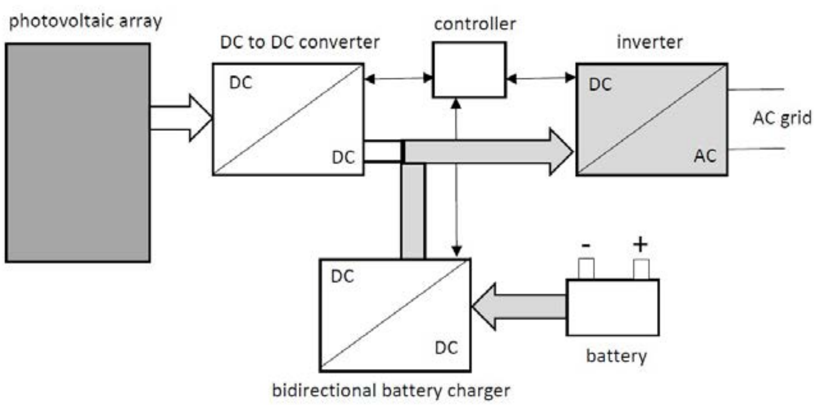


Fig. 2.16. Discharging the battery to inject power into the AC power grid.

battery). Especially when the solar irradiance is small, the generated power will be small and the charging of the battery will be too slow.

As visualised in Fig. 2.17, standalone battery charging systems generally have a sufficiently large energy storage device (which is, e.g., a stationary lead acid battery). When the photovoltaic array produces power, the DC to DC converters (the DC to DC converter connected with the photovoltaic array also has a maximum power tracker) allow to charge the battery of the electric vehicle. When the photovoltaic array produces more power than needed to charge the lithium ion battery of the vehicle, the excess of power will be stored in the large energy storage device (e.g. a lead acid battery). If sufficient energy has been stored in the lead acid battery, it is even possible to charge the battery of the vehicle without power originating from the photovoltaic array. This implies that it is also possible to charge the battery of the vehicle during the night.

When considering standalone photovoltaic based battery charging systems, not only the configuration visualised in Fig. 2.17 is possible. Figure 2.18 visualises a hybrid system. Using a lead acid battery with a large storage capacity (upper part of Fig. 2.18), it is possible to charge the lithium ion battery of the electrical vehicle (also during the night). This approach is not new and corresponds with the approach shown in Fig. 2.17.

The configuration reflected in Fig. 2.18 also uses the power generated by the photovoltaic array to produce hydrogen (lower part of Fig. 2.18). The hydrogen allows to store energy. Using a fuel cell fed by hydrogen, a DC voltage is obtained. Using a DC to DC converter, it is possible to charge the battery of the vehicle.

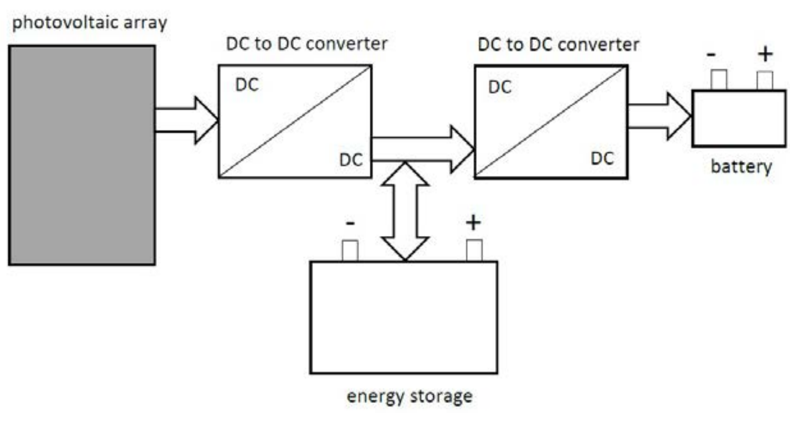


Fig. 2.17. Standalone photovoltaic based battery charging system.

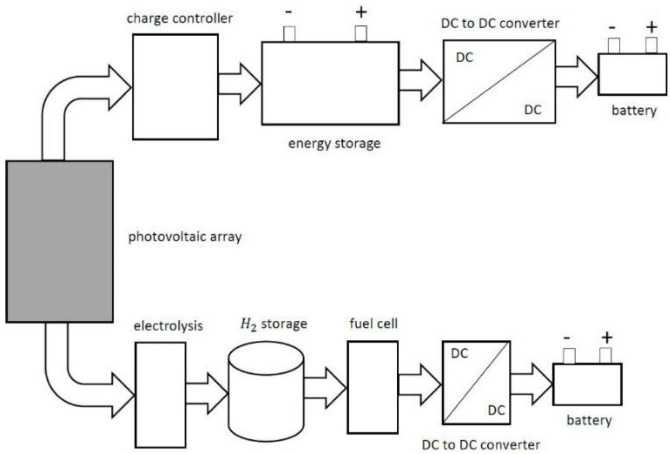


Fig. 2.18. Hybrid photovoltaic based battery charging system.

2.3. Challenges of introducing electric vehicles

A lot of people share the opinion that electrical propulsion (also when talking about cars used for passenger transport) has an important future. Sometimes overenthusiastic people neglect the challenges of introducing an electric fleet of cars.

The development of decent batteries is mandatory. Batteries must be able to store a sufficient amount of energy in a compact volume, and the weight must be acceptable. The efficiency when charging and discharging the battery needs to be sufficiently high. The use of the battery must be reliable and safe. A large life expectancy is also important. Finally, the choice of the raw materials determines the price of the battery.

Batteries need to be charged implying the need for a decent electrical grid infrastructure including a proliferation of charging stations. A high energy efficiency is needed while charging the battery. Also the propulsion systems of the car, i.e., the electric motor and the power electronic converter, need a high efficiency. The efficiency of the gearbox is an important parameter.

The introduction of full electric vehicles where the motor and the power electronic converter are fed by a battery (or a fuel cell) will gain importance, but a lot of engineers and scientists share the opinion that the use of an internal combustion engine will not disappear that fast. When considering internal combustion engines, improvements concerning efficiency and a reduction of the emission of exhaust gases are expected. Not only conventional internal combustion engine cars but especially hybrid electric vehicles have a future.

Whatever technology has been used, the car needs a good dynamic (i.e. acceleration and top speed) performance. The car must be able to operate in harsh environments (changing ambient temperatures, rain, snow, etc.), and the noise production must be limited (e.g. reducing vibrations).

2.3.1. All-climate battery technology

A large number of different battery technologies exist. A description of all these battery technologies, including their advantages and disadvantages, is beyond the scope of the present text. Whatever battery technology has been used, the car drivers need:

- a sufficiently large range (especially 'range anxiety', i.e., the fear that the battery will be discharged before reaching the destination, discourages the use of electric vehicles);

- fast charging stations (e.g., with charging powers up to 350 kW, a 60 kWh battery can be charged in approximately 10 minutes).

Reaching this large range and realising a fast charging behaviour is especially a challenge in the case of cold weather. Due to the low temperatures, the range decreases and the required charging time increases. Efforts are made to develop the so-called all-climate battery technology to have a sufficiently large range and a sufficiently small charging time.

The all-climate battery we consider here, is a self-heating lithium-ion battery. This battery can warm up rapidly (e.g. needing only tens of seconds) from the low ambient temperature to room temperature. Once warmed up, the desired large range and the fast charging behaviour are obtained.

2.3.1.1 Lithium-ion batteries

Rechargeable lithium-ion batteries are used in a large range of applications. They can be found in laptops, MP3 players, cellphones, etc., but they are also used in electric vehicles. By using lithium-ion batteries, a quite compact energy storage is obtained.

A lithium-ion battery contains a number of cells. Each cell contains a cathode (positive electrode), an anode (negative electrode) and an electrolyte. The positive

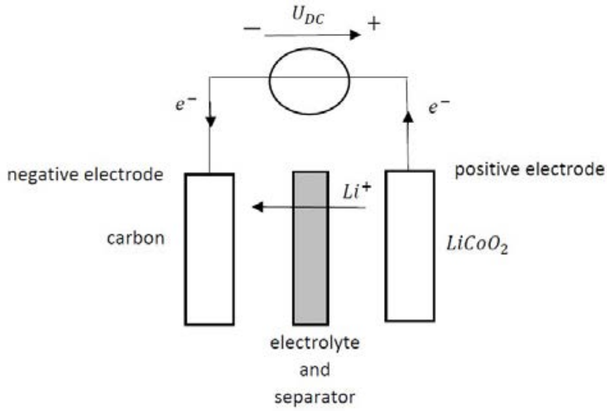


Fig. 2.19. Charging a lithium-ion battery.

electrode is typically made of lithium-cobalt oxide $LiCo0_2$ (the positive electrode can also be made of lithium iron phosphate $LiFePO_4$), and the negative electrode is made of carbon (graphite). Different electrolytes are used, but the electrolyte is not that important to understand the working principle of the battery.

To load the battery, a DC voltage source U_{DC} is needed as visualised in Fig. 2.19. Due to this voltage source, electrons are extracted from the $LiCo0_2$ electrode and sent to the carbon electrode. Lithium ions Li^+ flow from the $LiCo0_2$ electrode to the carbon electrode through the electrolyte. More precisely, during the charging process of the battery the chemical reaction equals

$$\mathrm{LiCoO_2} \rightarrow \mathrm{CoO_2} + \mathrm{Li^+} + \mathrm{e^-}. \tag{2.1}$$

Carbon electrode (C_6) is receiving Li^+ ions and electrons e^- implying the chemical reaction

$$\mathrm{C_6} + \mathrm{Li^+} + e^- \rightarrow \mathrm{LiC_6}, \tag{2.2}$$

By combining these two reactions, the charging reaction equals

$$\mathrm{C_6} + \mathrm{LiCoO_2} \rightarrow \mathrm{LiC_6} + \mathrm{CoO_2}. \tag{2.3}$$

The discharging process of the lithium-ion battery is visualised in Fig. 2.20, and the chemical reactions equal

$$\mathrm{LiC_6} \rightarrow \mathrm{C_6} + \mathrm{Li^+} + e^- \tag{2.4}$$

and

$$CoO_2 + Li^+ + e^- \rightarrow LiCoO_2. \quad (2.5)$$

By combining these two reactions, the discharging reaction equals

$$LiC_6 + CoO_2 \rightarrow C_6 + LiCoO_2\ , \quad (2.6)$$

which restores the original carbon electrode and $LiCo0_2$ electrode.

2.3.1.2. Properties of lithium-ion batteries

Lithium-ion batteries need an electronic controller to regulate the charging and the discharging processes. The controller avoids overcharging and overheating in order to prevent an explosion of the battery. Due to the remaining danger of a 'thermal runaway' of the lithium-ion battery (especially in case of overcharging or an internal malfunction causing a short circuit), they are not allowed on passenger airplanes. Lithium-ion batteries are used in electrical vehicles and the risks are acceptable (also when using gasoline, safety isssues arise).

In comparison with the traditional lead-acid batteries, lithium-ion batteries are much more compact and lightweighted. Lithium-ion batteries are reliable and they do not suffer from the so-called memory effect as it is the case for e.g. a nickel-cadmium battery (nickel-cadmium batteries need to be fully discharged from time to time). Lithium-ion batteries need no cadmium which is known to be a toxic heavy metal.

Lithiun-ion batteries are often used in electric vehicles. In order to obtain a large range, it is important to be able to harvest the braking energy. Harvesting the braking energy using a lithium-ion battery is only possible when the battery is

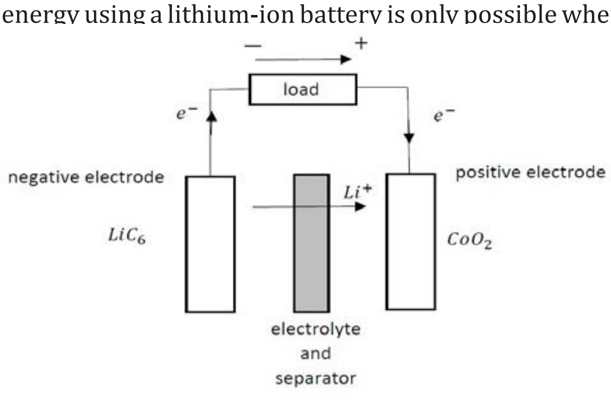


Fig. 2.20. Discharging a lithium-ion battery.

not too cold. This appropriate temperature property can be obtained by using a so called all-climate battery.

2.3.1.3. Structure of an all-climate battery

As already mentioned, the cell of a lithium-ion battery contains a cathode, an anode and an electrolyte. When considering an all-climate battery, a fourth component has been added, i.e., a thin nickel foil inside the cell. Figure 2.21 visualises the working principle of an all-climate battery. The nickel foil is coated with thin materials on both sides accounting for electrical insulation. The nickel foil is sandwiched between two anode layers.

The all-climate battery cell has three terminals. There are positive and a negative terminals like in all batteries (indicated by + and – in Fig. 2.21). But additionally, a third terminal is available which is called the activition terminal (ACT).

The working principle of an all-climate battery is straightforward. In case no heating is needed, the switch S will be open and a traditional battery operation is obtained. In Fig. 2.21, the Thevenin equivalent circuit of the battery is shown. The battery is feeding an electrical load or a charger is charging the battery.

In case heating is needed (e.g. the cell temperature equals –20 °C), swith S will be closed. All current generated by the cell flows in the nickel foil (represented by in Fig. 2.21) which produces heat to warm up the cell. When the cell temperature reaches a sufficiently high temperature (e.g. –5 °C) no additional heating is needed and switch S can be opened again. Heating up the battery cell from e.g. –20 °C to –5 °C only needs a number of seconds (e.g. 15 seconds).

The all-climate battery, visualised in Fig. 2.21, realises an internal heating of the battery. This approach is more appropriate than external heating of the battery. When using external heating, the heating speed is much lower, since it is important to prevent local overheating near the surface of the cell. Heating is speeding up by using more heating power, but this accounts for hot spots near the surface.

2.3.2. Extreme fast charging of car batteries

To realise a breakthrough of electric vehicles, it is not sufficient to have decent batteries which allow a compact energy storage. It is also important that these batteries can be charged sufficiently fast. In order to a obtain fast charging (or extreme fast charging) of the batteries, also adequate charging infrastructure is

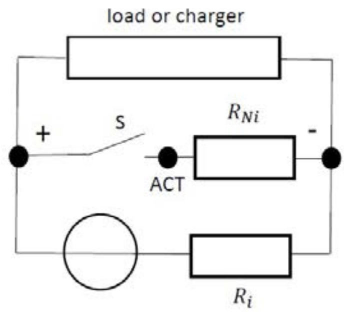


Fig. 2.21. All-climate battery.

needed. The charging infrastructure will extract large powers from the power grid implying that also a decent electrical power grid is needed. Possibly, even a medium voltage grid connection is used.

When speaking about fast (or extreme fast) charging, charging powers of 50 kW or even 350 kW are considered (in case of 350 kW, only a limited number of minutes are needed to charge the car battery). In such a situation, the power to the vehicle battery is provided by a DC connection. A power converter is used which is situated outside the vehicle.

2.3.2.1. Fast DC battery chargers

A DC battery charger can be built as visualised in Fig. 2.22. The charger contains two power conversion stages. The first stage is a three phase rectifier, and the second stage is a DC to DC converter which also realises a galvanic isolation. In order to avoid (or reduce) power quality problems in the three phase AC grid, it is important that the rectifier consumes a current which is sine shaped, i.e., a Pulse With Modulated rectifier can be used as discussed in the present text. Notice also the presence of an input filter in order to further improve the shape of the grid current.

The rectified voltage is converted to the required voltage level by using a DC to DC converter. The DC to DC converter realises a galvanic isolation. This galvanic isolation is important, since the vehicle battery is not grounded. Finally, also an output filter has been used to obtain a more constant DC voltage.

The main structure of such a DC to DC converter is visualised in Fig. 2.23. First, the DC voltage is converted to an AC voltage (e.g. in the kHz range which allows the use of a compact transformer) (in case of a 50 Hz voltage, a heavier and more spacious transformer is needed). A transformer changes the voltage level and provides the galvanic isolation. Finally, an AC to DC converter provides

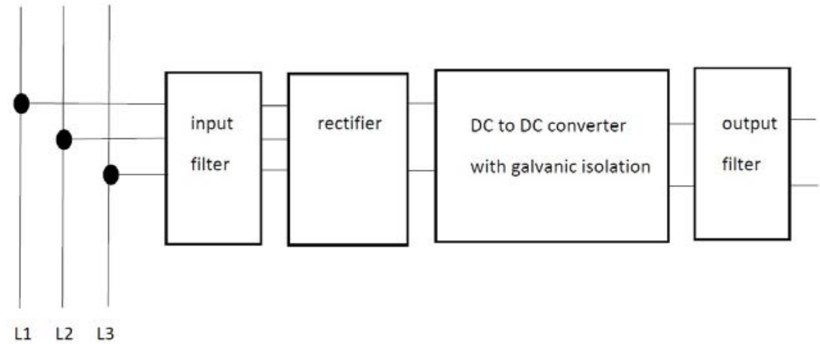


Fig. 2.22. Block diagram of a fast DC battery charger.

the desired DC voltage level.

When considering the fast DC battery chargers, typically the charging procedure starts with 'handshaking' between the battery charger and the car. After insulation testing and communication concerning the maximum charging parameters, the car is able to close the contactor and the real battery charging can start. The charging process typically contains two important phases. During the first phase, the battery will be charged using a constant current (this charging current is high implying a fast battery charging process). During the second phase, the battery will be charged using a constant voltage (as the state of charge of the battery increases, the charging current decreases and the charging process slows down).

2.3.2.2. Pulse width modulated rectifier

Although several configurations exist, Fig. 2.24 visualises a typical Pulse Width

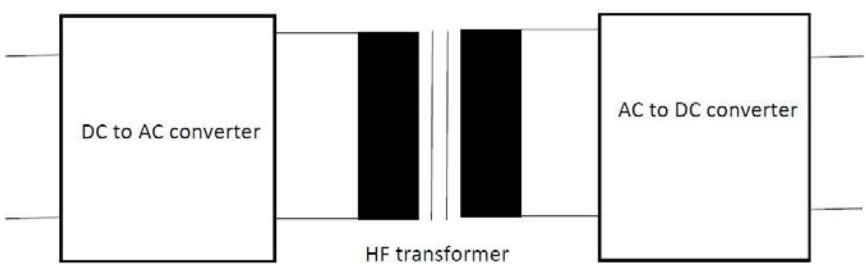


Fig. 2.23. DC to DC converter with galvanic isolation.

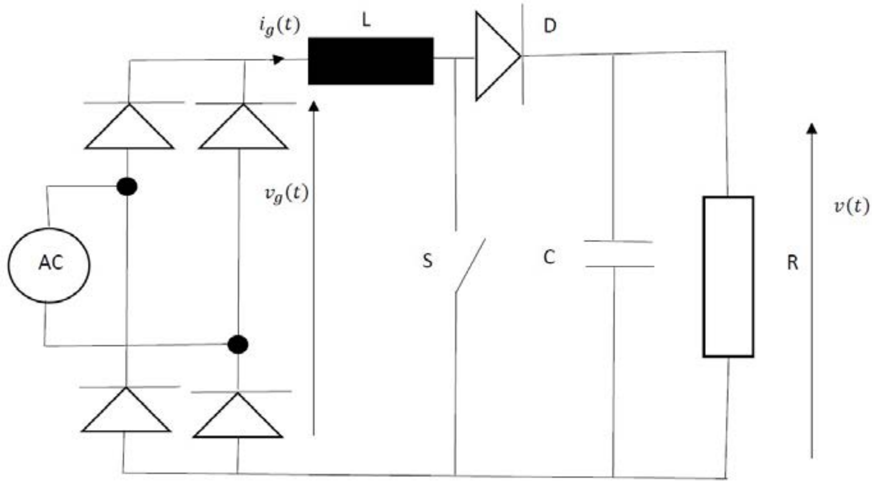


Fig. 2.24. Pulse width modulated rectifier.

Modulated rectifier (PWM). Notice first, the rectifier providing rectified voltage $v_g(t)$ which is visualised in Fig. 2.25. One of the goals is extracting current $i_g(t)$ which has the same shape as $v_g(t)$. This implies that the sinusoidal AC voltage provides a sinusoidal current (and a unity power factor is obtained, since the AC current and the AC voltage have the same phase).

In order to obtain the desired $i_g(t)$-shape, a controller is needed. Current $i_g(t)$ is measured and compared with the required reference current $i_{g,ref}(t)$.

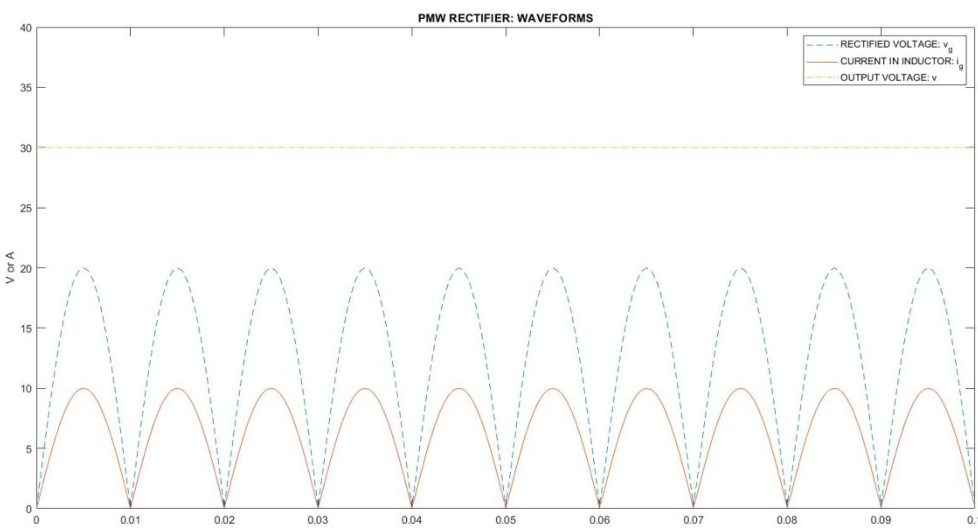


Fig. 2.25. Waveforms of the PWM rectifier.

By opening and closing switch S visualised in Fig. 2.24, current $i_g(t)$ (which is visualised in Fig. 2.25) will be obtained, which is a good approximation of the required reference current $i_{g,ref}(t)$. By closing switch S, current $i_g(t)$ increases. By opening switch S, current $i_g(t)$ decreases. The reference current $i_{g,ref}(t)$ equals

$$i_{g,ref}(t) = \frac{v_g(t)}{R_{AC}} \, , \tag{2.7}$$

where R_{AC} is the appropriate equivalent AC resistor. When assuming no losses occur in the converter,

$$R_{AC} = \frac{V_g^2}{P_{AC}} = \frac{V_g^2}{P_{DC}} = \frac{V_g^2 \, R}{v^2} \, . \tag{2.8}$$

Here, V_G is the RMS value of the grid voltage and P_{AC} is the active power provided by the AC grid. P_{DC} is the DC power dissipated in the load resistor R (with the DC voltage v). In case of a lossless converter, $P_{AC} = P_{DC}$.

Actually, inductor L, diode D and the semiconductor switch S behave as a boost converter. The DC output voltage $v(t)$ across capacitor C and load R is intended to be constant (value v as visualised in Fig. 2.25) and larger than (or equal to) the peak value of $v_g(t)$.

As already mentioned, a controller is needed to vary the duty cycle when opening and closing the semiconductor switch S. The working principle of the Pulse Width Modulated rectifier in Fig. 2.24 is quite straightforward. Suppose that switch S is closed. Voltage $v_g(t)$ stands over inductor L, which implies that

$$L \, \frac{d \, i_g(t)}{d \, t} = v_g(t) > 0. \tag{2.9}$$

When opening switch S, the inductor current $i_g(t)$ will decrease, but it will not drop to zero immediately. Due to diode D, current $i_g(t)$ can still flow. When neglecting the voltage drop across the conducting diode,

$$v(t) = v_g(t) - L \, \frac{d \, i_g(t)}{d \, t} > v_g(t). \tag{2.10}$$

Suppose that switch S is closed during a time interval $\delta \, T_S$ and the same switch is open during a time interval $(1 - \delta) \, T_S$. In a steady state situation, the increase of current $i_g(t)$ equals the decrease, i.e.,

$$\delta \, T_S \, \frac{v_g}{L} = (1 - \delta) \, T_S \, \frac{v - v_g}{L} \, . \tag{2.11}$$

This implies that

$$v = \frac{v_g}{1 - \delta} \, . \tag{2.12}$$

In case the constant output voltage v is larger than the maximum of $v_g(t)$, the duty cycle will depend on the time (actually on the instantaneous value of $v_g(t)$) with

$$\delta(t) = 1 - \frac{v_g(t)}{v}. \tag{2.13}$$

2.3.2.3. Grid infrastructure

In order to charge the batteries of an electric vehicle, the power electronic converters (including battery chargers) need a decent electrical grid infrastructure. A possible grid configuration of a charging station is visualised in Fig. 2.26. Power is supplied by a medium voltage grid, and using a classical 50 Hz transformer a local AC low voltage grid is obtained. Local power generation is possible using photovoltaic panels. Also stationary battery storage can be useful in order to limit and to control the power exchanges between the medium voltage grid and the low voltage grid. Several DC battery chargers (as visualised in Fig. 2.22) are available to charge the batteries of a number of cars simultaneously.

Using an AC grid has a number of advantages. Using AC grids is mainstream technology, i.e., converters, switchgears, protection devices, etc. are available on the market. Well-established standards and practices are available for AC distribution systems.

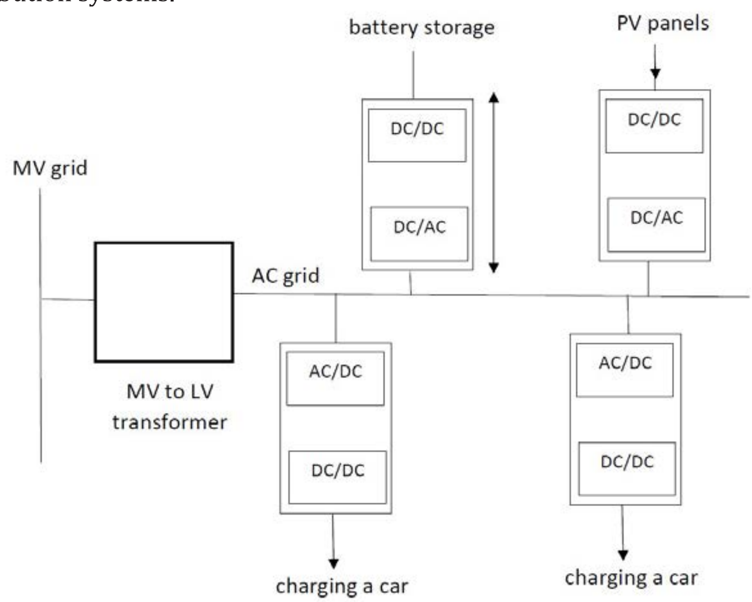


Fig. 2.26. AC grid infrastructure.

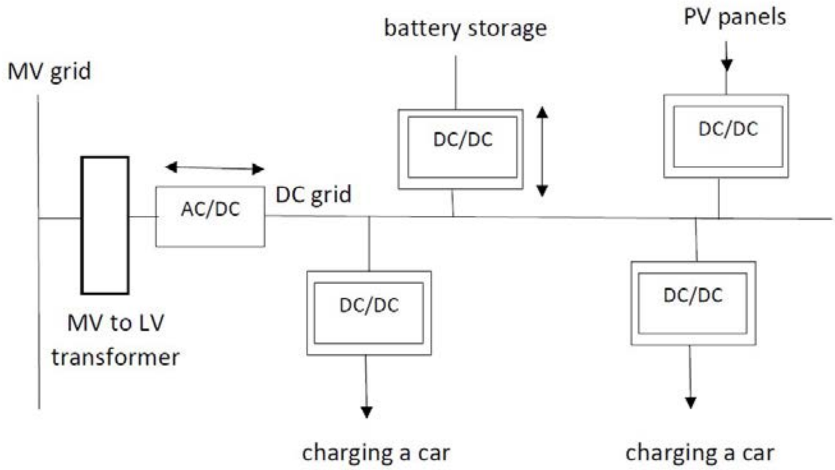


Fig. 2.27. DC grid infrastructure.

Instead of using the local AC low voltage grid infrastructure visualised in Fig. 2.26, also the use of a local DC low voltage grid infrastructure (as visualised in Fig. 2.27) can be useful. By using a local DC grid, the number of power electronic converters will be lower. By reducing the number of power electronic converters, losses decrease, which provides a higher efficiency. In general, the use of a DC grid makes the use of renewables and battery storage easier.

Notice in Fig. 2.27 the connection with the medium voltage grid using a transfomer and a rectifier. Since the installation contains photovoltaic panels and stationary battery energy storage, it can be a good practice to inject power into the medium voltage grid in case there is an excess of energy generation in comparison with the energy needs of the car batteries.

In Fig. 2.27, a separate MV to LV transformer (step-down transformer) and a separate AC to DC converter have been used. A solid-state transformer SST replaces both devices. Possibly, the solid-state transformer allows a bidirectional power exchange (as also visualised in Fig. 2.27). The internal structure of a solid-state transformer is visualised in Fig. 2.28. Figure 2.29 shows the grid configuration containing such a solid-state transformer SST.

The use of a solid-state transformer allows to reduce the losses (as a rule of thumb, the losses are halved). The solid state transformer, visualised in Fig. 2.28, is much smaller than the MV to LV transformer in Fig. 2.27 (the SST contains a HF transformer instead of a 50 Hz transformer) (by using a higher frequency the dimensions of the transformer can be reduced).

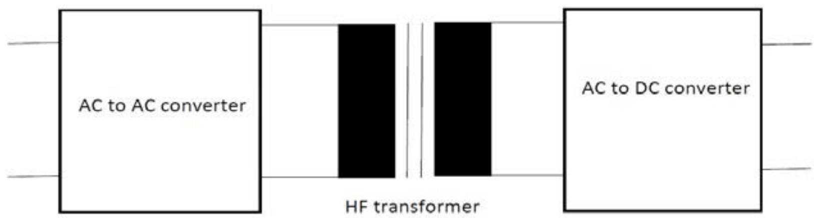


Fig. 2.28. Solid-state transformer.

2.3.2.4. Integrating car battery chargers in a residential environment

Instead of using separate battery charging stations, it is also possible to integrate the car battery chargers in a domestic environment, as visualised in Fig. 2.30. Figure 2.30 shows a grid configuration where both an AC grid and a DC grid are available.

The photovoltaic panels on the roofs generate DC voltages, and using DC to DC converters, the generated powers are injected into the local DC grid. The local DC grid feeds a number of DC to DC converters which can be used to charge car batteries. The local DC grid also feeds other battery chargers conncted with stationary batteries. These stationary batteries can be charged or discharged according to the power needs of the DC grid.

Not only a local DC grid, but also a local AC grid is available. Indeed, the electrical loads in the households are generally fed by an AC grid, no changes

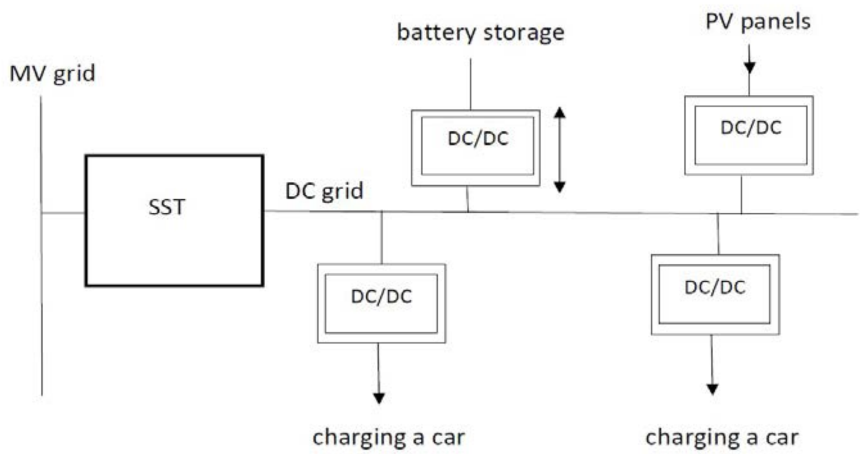


Fig. 2.29. DC grid infrastructure with solid-state transformer.

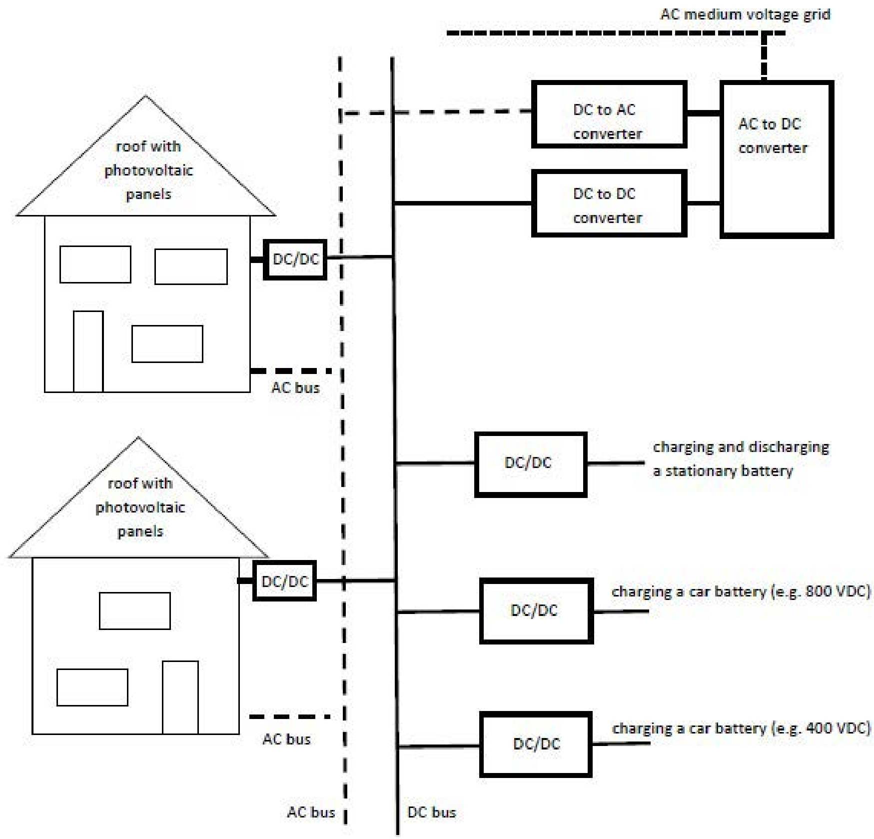


Fig. 2.30. Integrating car battery chargers in a domestic environment.

occur here. Notice also the connection with the medium voltage AC grid. Using a solid-state transformer SST, the local AC grid and the local DC grid are connected with the medium voltage AC grid and power can be exchanged in both directions.

2.4. Hybrid vehicles

Hybrid vehicles are intended to combine the advantages of vehicles having a conventional internal combustion engine and the advantages of electric vehicles. Vehicles having a conventional internal combustion engine have a large range, since petroleum fuels have a high energy density. Electric vehicles have the advantage to account for no/less pollution while driving. Electric vehicles also have a higher energy efficiency than vehicles having an internal combustion engine.

Hybrid vehicles have more than one energy source. They have an internal combustion engine allowing the use of gasoline or diesel. Diesel or gasoline (or another fossil based fuel) is considered to be the primary energy source guaranteeing the large range of the vehicle. In the future, perhaps it will be possible to replace the diesel or gasoline by hydrogen. The primary energy source is mainly the steady state power source.

Hybrid vehicles also have an electrical motor which is generally battery fed. The battery is considered to be the secondary energy source. The secondary energy source is the dynamic power source (e.g. used while accelerating the vehicle). Batteries also allow to recuperate the brake energy. A battery is indeed a bidirectional energy source, since a battery can be discharged and charged (while charging the battery, the electrial traction motor/machine operates as a generator).

A distinction can be made between series hybrid configurations, parallel hybrid configurations, series-parallel configurations and complex hybrid configurations. In the present text, we will mainly focus on the series and the parallel hybrid configurations.

2.4.1. Series hybrid electric vehicle

Figure 2.31 visualises the configuration of a series hybrid electric vehicle. Notice first of all the internal combustion engine connected with a fuel tank containing gasoline or diesel (primary energy source). The combustion engine operates at an optimal working point providing a maximum efficiency. Combustion engines are known to have a low efficiency, but the efficiency depends on the speed and the provided power. By choosing the optimal speed and the optimal power, a maximum efficiency is obtained. Due to the series configuration, the speed and the power of the combustion engine are not proportional with the speed and the needed power of the vehicle.

The combustion engine is driving an electrical generator which generates a three phase AC voltage which will be rectified in order to obtain a DC voltage. The DC voltage feeds a motor controller which controls the behaviour of the electrical traction motor (by adjusting the RMS value and the frequency of the three phase voltage feeding the motor). The traction motor is driving the vehicle using the mechanical transmission system.

In case the combustion engine and the generator provide a power which is larger than the power needed by the traction motor, the excess of power is sent to the battery using the DC to DC converter (dotted arrow in Fig. 2.31). When

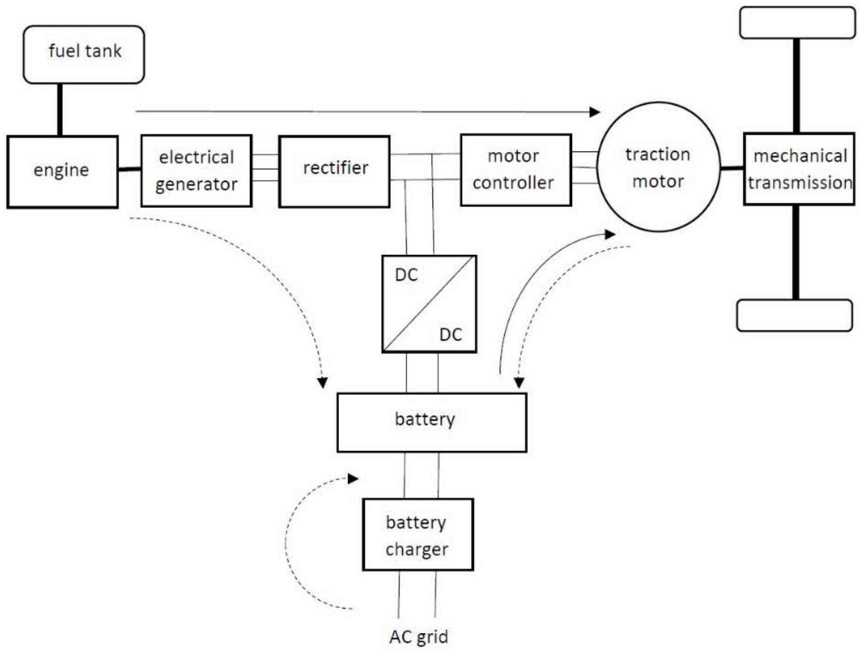


Fig. 2.31. Configuration of a series hybrid electric vehicle.

the vehicle needs more power than provided by the combustion engine, the additional power is supplied by the battery (secondary power source).

The traction motor/machine can also function as a generator while braking the vehicle. The brake energy will be converted into electrical energy by the traction generator. The brake energy will be stored in the battery, i.e., regenerative braking is obtained. This regenerative braking avoids the kinetic energy needs to be converted into heat by using the friction of traditional brakes.

The configuration in Fig. 2.31 is actually a plug-in hybrid electric vehicle, i.e., there is an additional battery charger which allows to charge the battery using power from the public AC grid. Not all hybrid vehicles are equipped with such an additional battery charger.

2.4.1.1. Operation modes of the drivetrain

When considering the drivetrain visualised in Fig. 2.31, a number of operation modes arise.

- Pure electric mode: The combustion engine is turned off and all the driving energy comes from the battery. This mode can e.g. be useful in a densely populated area where exhaust gases need to be avoided and where also noise pollution is undesired.

- Pure engine mode: All the driving energy is coming from the combustion engine, i.e., the batteries are neither charged nor discharged. For instance, in case the battery is empty and still a long trip is needed, based on fossil fuels and the combustion engine the desired range is possible.

- Hybrid mode: The driving energy is coming from the combustion engine and from the batteries. This allows to provide a large power to the traction motor which can be useful during acceleration of the vehicle.

- Engine traction and battery charging: The engine-generator drives the vehicle and the excess of power is used to charge the batteries. This approach allows the combustion engine to operate at its optimal working point (providing a maximum efficiency) when the mechanical power required by the vehicle is lower than the power of the combustion engine.

- Regenerative braking mode: The traction motor/machine is operating as a generator and the brake energy is stored in the batteries (the combustion engine is not used).

- Battery charging mode: The car is not driven (no power is sent to the traction motor), but the engine-generator charges the batteries.

- Hybrid battery charging mode: The traction motor/machine operates as a generator, together with the engine-generator they charge the batteries.

2.4.1.2. Advantages and disadvantages of the series hybrid configuration

The series configuration has a number of advantages. As already mentioned, the combustion engine is able to operate at its optimal working point providing a maximum efficiency. The speed of the combustion engine and the speed of the traction motor (and the vehicle) are not linked with each other. Due to that property, the combustion engine can be designed to behave optimal in a narrow speed and torque region (implying a higher efficiency and lower emissions). This is different from the situation in a traditional car where the combustion engine must operate in a broad range of speeds and torques. The combustion engine cannot behave optimal in such a large region.

The mechanical transmission is entirely driven by the electrical traction motor. Due to the motor controller, the electrical traction motor can operate in a very broad speed range. This implies that no multigear transmission is needed between the electrical motor and the driven wheels.

In case two electrical traction motors are used, each driving a wheel, no differential is needed to deal with the different speeds of the wheels. In case the vehicle contains a four-wheel-drive system, it can be an option to use four electrical traction motors.

The series configuration also has a number of disadvantages. The energy originating from the combustion engine is converted twice (first by the generator and then by the traction motor), which accounts twice for losses. Actually, also losses occur in the rectifier and the motor controller. Moreover, the traction motor must be sized for the full/maximum power. This implies that the generator and the traction motor can be heavy and spacious.

2.4.2. Parallel hybrid electric vehicle: Part 1

2.4.2.1. Basic principle

When considering the parallel hybrid electric vehicle configuration visualised in Fig. 2.32, a number of operation modes arise. Possibly the combustion engine alone is driving the mechanical transmission. Possibly the electrical motor alone is driving the mechanical transmission. Possibly the combustion engine and the electrical motor drive together the mechanical transmission. For shorter distances, mainly full electric mode can be used. Due to the combustion engine, a larger range is obtained, i.e., larger distances are possible.

Additional options (can) exist. Via the mechanical coupling, the combustion engine drives the electrical machine which operates as a generator. The batteries can be charged. When braking, kinetic energy is used to drive the electrical machine as a generator. Kinetic energy is stored into the battery, i.e., regenerative braking is obtained. Indeed, the kinetic energy is not converted into heat by using the friction in traditional brakes.

When considering the mechanical coupling visualised in Fig. 2.32, a distinction can be made between torque coupling and speed coupling.

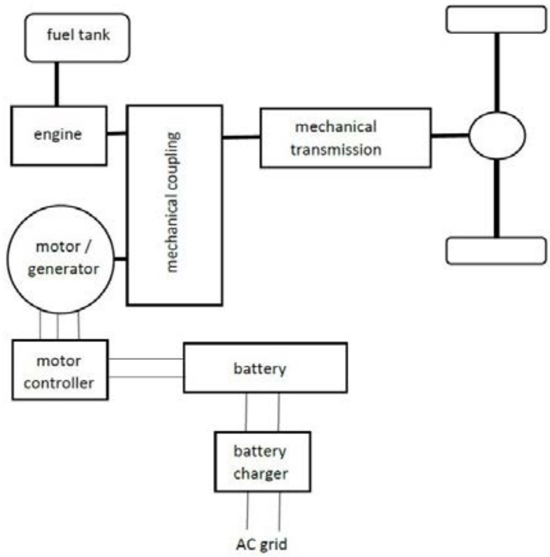


Fig. 2.32. Configuration of a parallel hybrid electric vehicle.

2.4.2.2. Torque coupling and speed coupling

The torque coupler in Fig. 2.33 has two inputs and one single output. The first input has a driving torque $T_{in,1}$ and a pulsation (speed) $\omega_{in,1}$. The second input has a driving torque $T_{in,2}$ and a pulsation (speed) $\omega_{in,2}$. The output provides torque T_{out} at pulsation ω_{out}. Depending on the torque coupling device, there exist constants k_1 and k_2 such that

$$T_{out} = k_1 T_{in,1} + k_2 T_{in,2} \tag{2.14}$$

and

$$\omega_{out} = \frac{\omega_{in,1}}{k_1} = \frac{\omega_{in,2}}{k_2}. \tag{2.15}$$

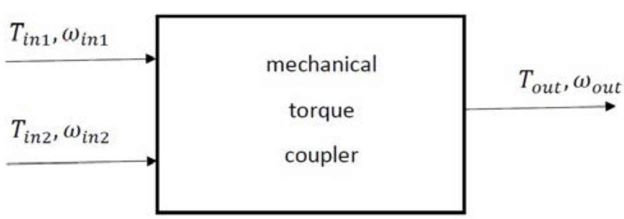


Fig. 2.33. Mechanical torque coupler.

Notice that the two input torques $T_{in,1}$ and $T_{in,2}$ are not related with each other. In case both torques are postive (e.g. $T_{in,1}$ originating from the combustion engine and $T_{in,2}$ originating from the electrical motor), a large output torque can be obtained. Such a large output torque can be useful to accelerate the vehicle or to climb a mountain. Notice that speeds $\omega_{in,1}$, $\omega_{in,2}$ and ω_{out} are proportional with each other. In case one of these speeds has been chosen, all other speeds are fixed.

The speed coupler in Fig. 2.34 has two inputs and one single output. The first input has driving torque $T_{in,1}$ and pulsation (speed) $\omega_{in,1}$. The second input has driving torque $T_{in,2}$ and pulsation (speed) $\omega_{in,2}$. The output provides torque T_{out} at pulsation ω_{out}. Depending on the speed coupling device, there exist constants k_1 and k_2 such that

$$\omega_{out} = k_1\omega_{in,1} + k_2\omega_{in,2} \tag{2.16}$$

and

$$T_{out} = \frac{T_{in,1}}{k_1} = \frac{T_{in,2}}{k_2} . \tag{2.17}$$

Notice that the two input speeds $\omega_{in,1}$ and $\omega_{in,2}$ are not related with each other. Suppose $\omega_{in,1}$ is the speed of the combustion engine and $\omega_{in,2}$ is the speed of the electrical motor. The speed $\omega_{in,1}$ of the combustion engine cannot be too small, but by combining it with a negative speed $\omega_{in,2}$, a really low output speed ω_{out} can be obtained. In order to limit the fuel consumption of the combustion engine, $\omega_{in,1}$ is also not allowed to be too high. By combining a positive $\omega_{in,1}$ with a positive $\omega_{in,2}$, a really high ω_{out} can be obtained. Notice that torques $T_{in,1}$, $T_{in,2}$ and T_{out} are proportional with each other. In case one of these torques has been chosen, all other torques are fixed.

2.4.2.3. Torque coupling: exercises

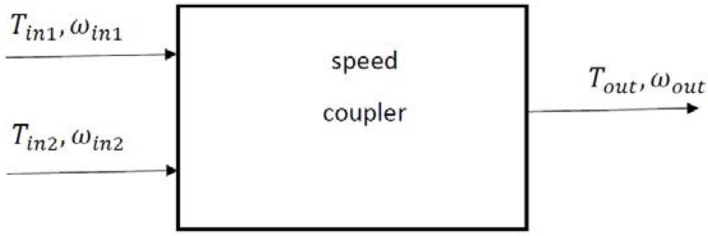


Fig. 2.34. Mechanical speed coupler.

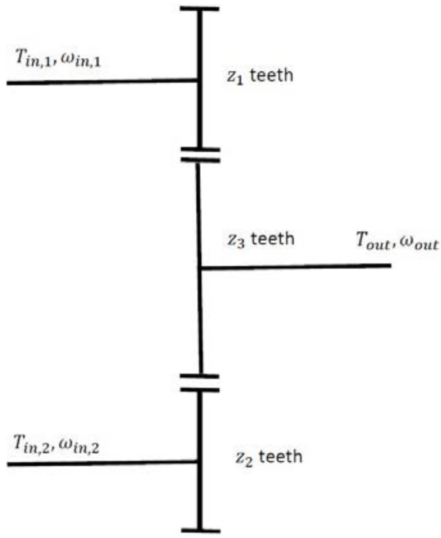


Fig. 2.35. Torque coupler: example 1.

There exists a broad range of torque couplers, and discussing them all is beyond the scope of the present text. A limited number of torque couplers will be shown with the corresponding k_1 and k_2 values. It is left as an exercise to the students to show that the given k_1 and k_2 values are correct.

Exercise 1.

The torque coupler in Fig. 2.35 contains three gear wheels. The first one has z_1 teeth (first input), the second one has z_2 teeth (second input), and the third one (the output) has z_3 teeth. The sizes of the teeth of all gear wheels are the same, which implies that the number of teeth of a gear wheel is proportional with its radius. The peripheral speeds of all gear wheels are the same.

When neglecting all losses, the total input power equals the output power, i.e.,

$$\omega_{in,1}T_{in,1} + \omega_{in,2}T_{in,2} = \omega_{out}T_{out}\,. \tag{2.18}$$

Prove that

$$k_1 = \frac{z_3}{z_1}, \quad k_2 = \frac{z_3}{z_2} \tag{2.19}$$

implying that

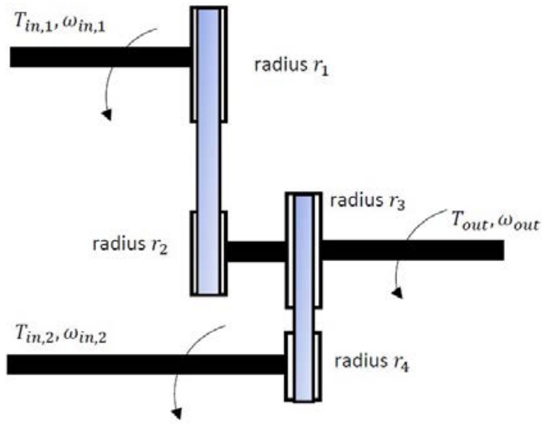


Fig. 2.36. Torque coupler: example 2.

$$T_{out} = \frac{z_3}{z_1} T_{in,1} + \frac{z_3}{z_2} T_{in,2} \tag{2.20}$$

and

$$\omega_{out} = \frac{z_1}{z_3} \omega_{in,1} = \frac{z_2}{z_3} \omega_{in,2}\,. \tag{2.21}$$

Exercise 2.

The torque coupler in Fig. 2.36 contains pulleys having radii r_1, r_2, r_3 and r_4. Due to the belts, pulleys 1 and 2 have the same peripheral speeds and pulleys 3 and 4 have the same peripheral speeds. Pulleys 2 and 3 share the same axis and have the same number of revolutions per minute.

The arrows in Fig. 2.36 indicate the rotational directions. When neclecting all losses, the total input power equals the output power. Prove that

$$k_1 = \frac{r_2}{r_1}, \quad k_2 = \frac{r_3}{r_4} \tag{2.22}$$

implying that

$$T_{out} = \frac{r_2}{r_1} T_{in,1} + \frac{r_3}{r_4} T_{in,2} \tag{2.23}$$

and

$$\omega_{out} = \frac{r_1}{r_2} \omega_{in,1} = \frac{r_4}{r_3} \omega_{in,2} \tag{2.24}$$

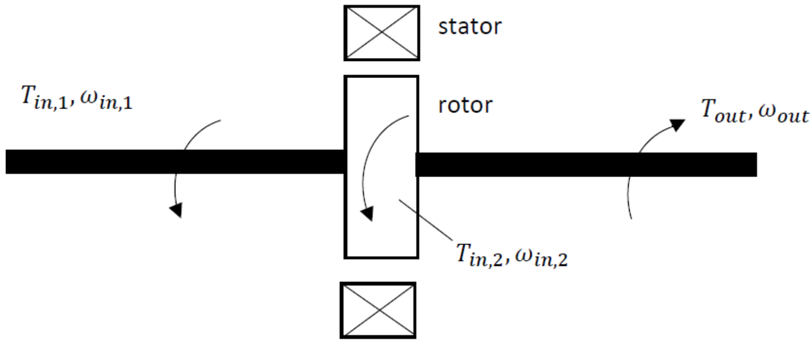


Fig. 2.37. Torque coupler: Example 3.

Exercise 3.

The torque coupler in Fig. 2.37 has one single axis of rotation. This implies that $\omega_{in,1} = \omega_{in,2} = \omega_{out}$, i.e., $k_1 = k_2 = 1$. The output torque $T_{out} = T_{in,1} + T_{in,2}$. The arrows in Fig. 2.37 do not indicate the rotational directions but the direction of the torques. Notice that in the visualisation in Fig. 2.37, the input torques $T_{in,1}$ and $T_{in,2}$ are driving torques, whereas the output torque T_{out} is a counteracting torque.

2.4.2.4. Torque speed characteristics in case of two transmissions

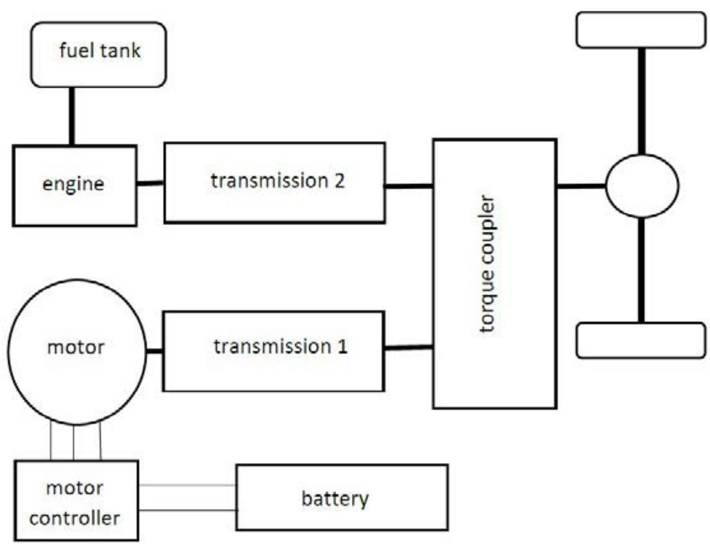


Fig. 2.38. Parallel hybrid configuration with two transmissions.

The configuration visualised in Fig. 2.32 is not the only parallel hybrid configuration. Figure 2.38 visualises an alternative parallel hybrid configuration. Notice a very similar way of the use of the internal combustion engine and the electrical motor. Notice the presence of a torque coupler but notice also the presence of two transmissions (one for the combustion engine and one for the electrical motor). Single-gear transmissions (with only one single speed ratio) or multi-gear transmissions (with several speed ratios) can be used.

In case both transmissions are multi-gear transmissions, a maximum flexibility is obtained but a complicated drivetrain is obtained. In case both transmissions are multi-gear transmissions, the engine and the motor are both able to operate near their optimal working points (e.g. with optimal efficiency).

The maximum flexibility is demonstrated by the torque speed characteristics. The resulting torque speed characteristic is a combination of the torque speed characteristic of the electrical motor and of the torque speed characteristic of the combustion engine. On the left in Fig. 2.39 one can see a typical torque speed characteristic of an electrical motor in case it is fed by a motor controller. Up to the nominal speed, the nominal torque can be obtained. Above the nominal speed, only lower torques are obtained. Notice, however, that the speed range is really large. On the right in Fig. 2.39 one can see a typical torque speed characteristic of a combustion engine. The speed range of a combustion engine is more limited. The shape of the characteristic can be peak shaped or more flat. In the present text, we consider a rather flat torque speed characteristic.

The resulting torque speed characteristic is a combination of the torque speed characteristic of the electrical motor and of the torque speed characteristic of the combustion engine. The speed ratios of the transmissions determine the shape of the resulting torque speed characteristic. A typical result is visualised in Fig. 2.40.

A transmission typically reduces the speed, since the speed of the motor/

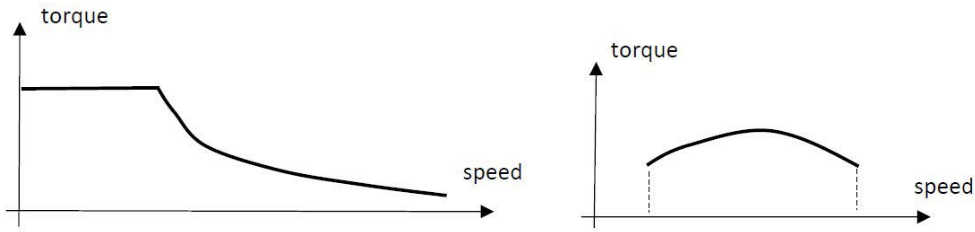


Fig. 2.39. Typical torque speed characteristics.

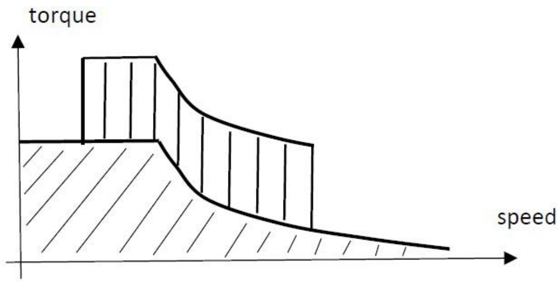


Fig. 2.40. Resulting torque speed characteristic.

engine is higher than the finally needed rotational speed. By reducing the speed, the torque is increased with the same factor (when neglecting the losses, the input power and the output power are the same).

When choosing the first gear of transmission 1, the electrical motor gives a high starting torque to the torque coupler and, finally, to the wheels of vehicle. But that high torque is only available for really low speeds. When choosing a higher gear, the starting torque becomes lower but the torque is available up to higher speeds (compare it with a traditional car, starting occurs in the first gear and as a higher car speed is desired, a higher gear is chosen).

When choosing the first gear of transmission 2, the combustion engine gives a higher additional torque in a more narrow speed range. When choosing a higher gear, the additional torque is lower but available in a broader speed range (the additional torque starts at a somewhat higher speed but it is available up to a really higher speed).

2.4.2.5. Single-shaft configurations

Alternatively to the configuration in Fig. 2.38, also a single-shaft configuration can be used. Figure 3.41 visualises such a single-shaft configuration. Notice the combustion engine providing a torque speed characteristic, as visualised in Fig. 2.39 (on the right). Notice the electrical motor providing a torque speed characteristic, as visualised in Fig. 2.39 (on the left). The rotor of the electrical motor functions as a torque coupler as, visualised in Fig. 2.37. Between the motor and the transmission a total torque $T_{in,1} + T_{in,2}$ is obtained (where $T_{in,1}$ is the torque provided by the combustion engine and $T_{in,2}$ is the torque provided by the electrical motor). The speeds of the combustion engine and the electrical motor are the same, i.e., $\omega = \omega_{in,2} = \omega_{in,1}$. In case the transmission accounts for a speed reduction with factor k,

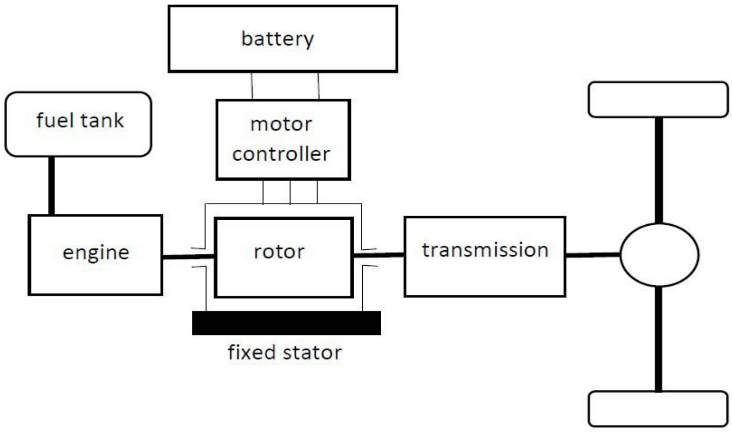


Fig. 2.41. Single-axis configuration with pre-transmission.

$$T_{out} = k(T_{in,1} + T_{in,2}) \tag{2.25}$$

and

$$\omega_{out} = \frac{\omega_{in,1}}{k} = \frac{\omega_{in,2}}{k}. \tag{2.26}$$

In the configuration in Fig. 2.41, the torques of the combustion engine and the electrical motor are modified in the same way by the transmission. The speeds of the combustion engine and the electrical motor are modified in the same way by the transmission. This implies the engine and the motor must have the same

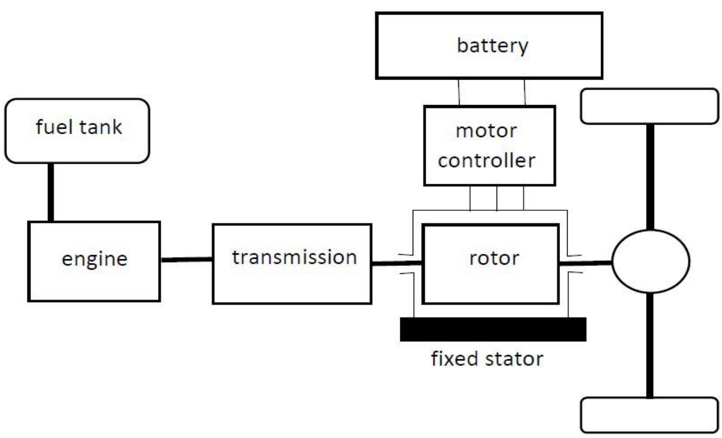


Fig. 2.42. Single-axis configuration with post-transmission.

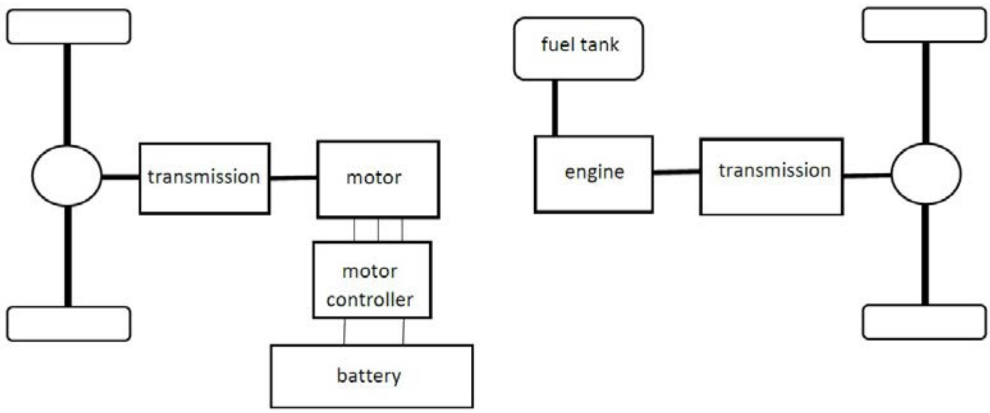


Fig. 2.43. Hybrid configuation with separate driving of the shafts.

speed range. By placing a transmission between the combustion engine and the motor (see Fig. 2.42), only the torque and the speed of the combustion engine are modified. The torque of the electrical motor is directly driven to the wheels (drive shaft). The speed range of a combustion engine is more limited than the speed range of an electrical motor. A multi-gear transmission allows to use the engine over a broad speed range of the vehicle.

When the engine is driving the electrical machine as a generator, it is possible to charge the batteries. In case of a standstill of the vehicle, the electrical machine does not rotate and it is not possible to charge the batteries.

2.4.2.6. The use of separate drivetrains

Figure 2.43 visualises a hybrid configuration where the combustion engine drives the first drive shaft and the electrical motor drives the second drive shaft.

2.4.3. Parallel hybrid electric vehicle: Part 2

2.4.3.1. The use of a speed coupler

The previous paragraphs mainly considered the use of a torque coupler in a hybrid vehicle configuration. Figure 2.34 already mentioned the use of a speed coupler. The speed coupler in Fig. 2.34 has two inputs and one single output. The first input has driving torque $T_{in,1}$ and pulsation (speed) $\omega_{in,1}$. The second input has driving torque $T_{in,2}$ and pulsation (speed) $\omega_{in,2}$. The output provides torque T_{out} at pulsation ω_{out}. Depending on the speed coupling device, there exist constants k_1 and k_2 such that

$$\omega_{out} = k_1\omega_{in,1} + k_2\omega_{in,2} \tag{2.27}$$

and

$$T_{out} = \frac{T_{in,1}}{k_1} = \frac{T_{in,2}}{k_2}. \tag{2.28}$$

Notice that the two input speeds $\omega_{in,1}$ and $\omega_{in,2}$ are not related with each other. Suppose $\omega_{in,1}$ is the speed of the combustion engine and $\omega_{in,2}$ is the speed of the electrical motor. The speed $\omega_{in,1}$ of the combustion engine cannot be too small, but by combining it with a negative speed $\omega_{in,2}$, a really low output speed ω_{out} can be obtained. In order to limit the fuel consumption of the combustion engine, $\omega_{in,1}$ is also not allowed to be too high. By combining a positive $\omega_{in,1}$ with a positive $\omega_{in,2}$, a really high ω_{out} can be obtained. Notice that torques $T_{in,1}$, $T_{in,2}$ and T_{out} are proportional to each other. In case one of these torques has been chosen, all other torques are fixed.

2.4.3.2. The use of a planetary gear

Figure 2.44 visualises a planetary gear which functions as a speed coupler. Notice the sun gear having a radius R_1. The sun gear can be connected with the combustion engine. There is a driving torque $T_{in,1}$, and consider a counterclockwise rotation with pulsation $\omega_{in,1}$. Notice also the ring gear having radius R_2. The ring gear can be connected with the electrical motor. There is driving torque $T_{in,2}$ and consider a counterclockwise rotation with pulsation $\omega_{in,2}$.

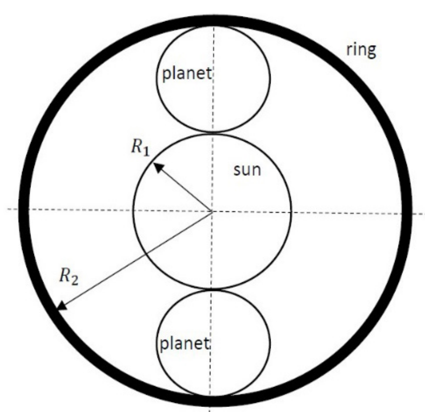


Fig. 2.44. Planetary gear.

A planetary gear unit not only has a sun and a ring. Notice also the planets which are circulating around the sun. The planets are also circulating around their own centre. In Fig. 2.44, there are two planets, but the number of planets is a design parameter. The radius of a planet equals

$$\frac{R_2 - R_1}{2} \tag{2.29}$$

and the centre of the planet has a distance

$$R_3 = \frac{R_1 + R_2}{2} \tag{2.30}$$

to the centre of the sun. The centres of the planets are connected with each other using a carrier (not shown in Fig. 2.44). The carrier rotates with the same centre as the sun and the ring. The carrier is actually the output of the planetary gear which will drive the actual load (the vehicle). Suppose the carrier is rotating in a counterclockwise direction with pulsation $\omega_{out,3}$. The load will imply a counteracting torque $T_{out,3}$ which is a clockwise oriented torque.

2.4.3.3. Torque and speed behaviour of a planetary gear

Based on Fig. 2.44, first, the relationships between the input speeds $\omega_{in,1}$ and $\omega_{in,2}$ (assume both are in a counterclockwise direction) and the output speed $\omega_{out,3}$ (will also be in a counterclockwise direction) will be determined.

Consider the upper planet in Fig. 2.44 and assume it is rotating in a clockwise direction around its own centre with pulsation ω_P. The peripheral speeds of the sun and the planet must be the same. More precisely

$$\omega_{in,1} R_1 = \omega_P \left(\frac{R_2 - R_1}{2}\right) + \omega_{out,3} R_3. \tag{2.31}$$

The peripheral speeds of the ring and the planet must also be the same. More precisely,

$$\omega_{in,2} R_2 = -\omega_P \left(\frac{R_2 - R_1}{2}\right) + \omega_{out,3} R_3. \tag{2.32}$$

By combining these two expressions, ω_P can be eliminated, which leads to expressions

$$\omega_{out,3} = \frac{R_1}{2\,R_3} \omega_{in,1} + \frac{R_2}{2\,R_3} \omega_{in,2} = k_1 \omega_{in,1} + k_2 \omega_{in,2} \tag{2.33}$$

implying the appropriate k_1 and k_2. When neglecting the friction, it is also possible to determine the relationship between the input torques $T_{in,1}$ and $T_{in,2}$ and the output torque $T_{out,3}$. Two relationships are needed, i.e.,

- the resulting torque applied to the carrier must be zero (with respect to the centre of the sun, ring and carrier);
- the resulting torque applied to a planet must be zero (with respect to the centre of the planet).

The first relationship implies that

$$T_{out,3} = T_{in,1} + T_{in,2}\,. \tag{2.34}$$

The second relationship considers torques applied by the sun and the ring on a planet. In case there are n planets, the torque of the sun implies a peripheral force

$$\frac{T_{in,1}}{n\,R_1} \tag{2.35}$$

and the torque of the ring implies a peripheral force

$$\frac{T_{in,2}}{n\,R_2}\,. \tag{2.36}$$

The first force implies a clockwise torque to a planet and the second force implies a counterclockwise torque to a planet (both with respect to the centre of the considered planet). In order to maintain a constant speed of rotation of the considered planet, both torques must be the same, i.e.,

$$\frac{R_2 - R_1}{2}\,\frac{T_{in,1}}{n\,R_1} = \frac{R_2 - R_1}{2}\,\frac{T_{in,2}}{n\,R_2} \tag{2.37}$$

implying that

$$\frac{T_{in,1}}{R_1} = \frac{T_{in,2}}{R_2}\,. \tag{2.38}$$

In combination with relationship $T_{out,3} = T_{in,1} + T_{in,2}$, one obtains that

$$T_{out,3} = \frac{2\,R_3}{R_1}\,T_{in,1} = \frac{2\,R_3}{R_2}\,T_{in,2} = \frac{T_{in,1}}{k_1} = \frac{T_{in,2}}{k_2} \tag{2.39}$$

with the already known values for k_1 and k_2. This shows indeed that a speed coupler is obtained.

2.4.3.4. Applying a planetary gear in a hybrid vehicle

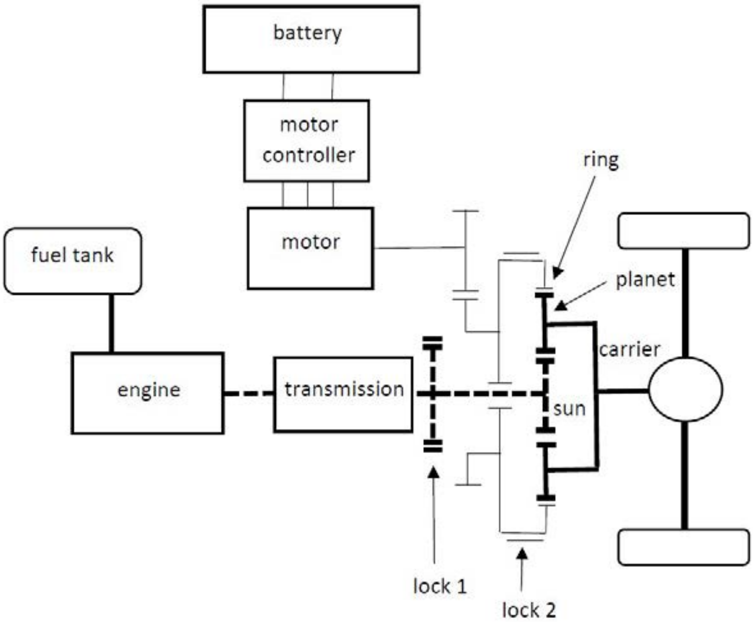


Fig. 2.45. Vehicle equipped with a planetary gear.

Notice in Fig. 2.45 a hybrid vehicle equipped with a combustion engine and an electrical motor. Notice the planetary gear containing a sun gear, a ring gear, planets and a carrier. The combustion engine provides its power to the sun gear. The electrical motor provides its power to the ring gear. Notice also two locks. The first lock allows to lock the sun gear. The second lock allows to lock the ring gear.

When lock 1 and lock 2 are both released (sun gear and ring gear are both able to rotate), the combustion engine and the electrical motor both provide power to the wheels. Hybrid traction is obtained.

When only lock 2 is locked (ring gear is locked and sun gear is able to rotate), only the combustion engine provides power to the wheels. Engine alone traction is obtained. When only lock 1 is locked (sun gear is locked and ring gear is able to rotate), only the electrical motor provides power to the wheels. Motor alone traction is obtained.

The configuration in Fig. 2.45 also allows regenerative breaking. When lock 1 is locked (sun gear is locked and ring gear is able to rotate), the combustion engine must be shut off or the clutch (not shown in Fig. 2.45) must be disengaged.

The electrical machine functions as a generator. The wheels are driving the carrier, the ring gear and the generator. Kinetic and potential energy is converted into electrical energy which is stored in the batteries. Regenerative breaking is obtained.

The vehicle configuration also allows the combustion engine to drive the electrical machine which functions as a generator. Mechanical energy is converted into electrical energy which is stored in the batteries. Battery charging is obtained with energy originating from the combustion engine.

2.4.4. Parallel hybrid electric vehicle: Part 3

Not only a planetary gear can be used to realise a speed coupler, it is also possible to use a transmotor, i.e., an electrical motor with a floating stator. Figure 2.46 visualises such a transmotor which is actually a permanent magnet synchronous motor (PMSM).

2.4.4.1. Permanent magnet synchronous motor

A 'classical' permanent magnet synchronous motor has a fixed stator which is fed by a three phase sinusoidal voltage with frequency f_e, as visualised in Fig. 2.46. Suppose the three phase stator windings, with a three phase current, generate a rotating magnetic field having p pole pairs. The pulsation of that rotating magnetic field equals

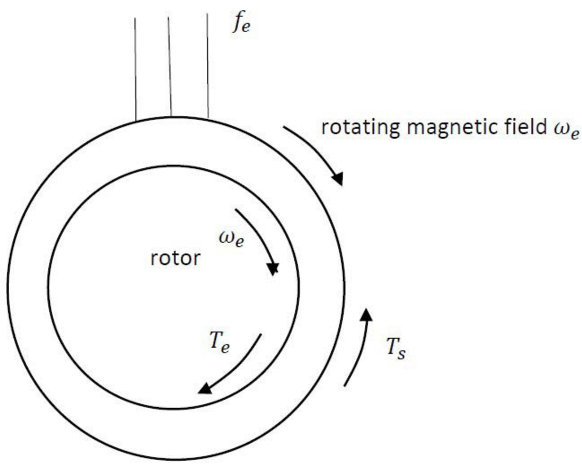


Fig. 2.46. Permanet magnet synchronous motor.

$$\omega_e = \frac{2\pi f_e}{p}. \qquad (2.40)$$

The rotor contains permanent magnets implying that a rotating rotor generates a rotating magnetic field with the same speed as the mechanical speed of rotation. Since the PMSM is a synchronous motor, the rotor will rotate with the same pulsation $\omega_r = \omega_e$. The motor converts electrical energy into mechanical energy, i.e., torque T_e is applied to the rotor providing mechanical power $\omega_e T_e$. Based on the law of action and reaction, torque $T_s = T_e$ is applied to the stator, as shown in Fig. 2.46.

2.4.4.2. Working principle of a transmotor

Figure 2.47 visualises a transmotor which has one single stator and two rotors. The stator (dotted lines in Fig. 2.47) does not really have an 'electrical' goal, but a fixed mechanical enclosure is needed. The transmotor contains an inner rotor and an outer rotor. The inner rotor is driving the mechanical load (port 3). The inner rotor contains permanent magnets (similar with the rotor in a classical PMSM). The outer rotor (also called the 'floating stator') functions like the stator in a classical PMSM. The outer rotor contains a three phase winding and the outer rotor is fed by an inverter providing an AC voltage with a controllable frequency f_e (port 2). Slip rings and carbon brushes are needed to connect the inverter and allow this inverter to provide electrical power to the transmotor (to the outer rotor). The outer rotor is also driven by the combustion engine (port 1) which provides mechanical power to the transmotor.

Consider the outer rotor in Fig. 2.48. The outer rotor is driven by the combustion engine providing driving torque T_{ms} and mechanical speed ω_s (the combustion engine provides mechanical power $\omega_s T_{ms}$ to this outer rotor). The outer rotor is fed by the inverter injecting a three phase current with frequency f_e implying a rotating magnetic field with speed ω_e with respect to

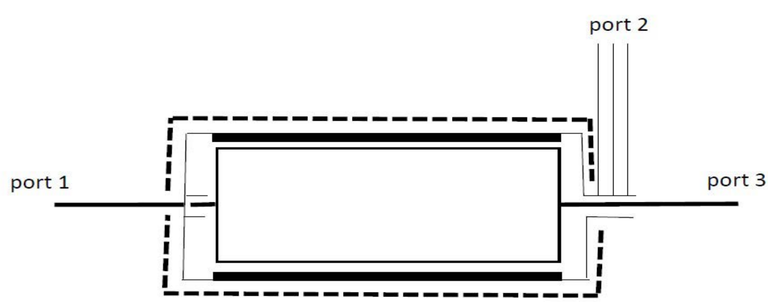


Fig. 2.47. Transmotor.

this outer rotor. This means that a rotating magnetic field with an absolute speed $\omega_s + \omega_e$ is obtained. The inverter provides electrical power $\omega_e T_e$ to the rotor.

A steady state torque is applied to the inner rotor when the inner rotor rotates at mechanical speed $\omega_r = \omega_s + \omega_e$. Driving torque $T_r = T_e$ implies that mechanical power $\omega_r T_r = (\omega_s + \omega_e) T_e$ is provided to the mechanical load (port 3 in Fig. 2.47). The mechanical power at the output is the sum of the mechanical power of the combustion engine $\omega_s T_{ms}$ and the electrical power provided by inverter $\omega_e T_e$ (losses in the transmotor are neglected).

Notice that the speed of the internal rotor is constant, since the driving torque $T_r = T_e$ equals the mechanical counteracting torque T_{mr} of the mechanical load. Since the outer rotor applies a driving torque $T_r = T_e = T_{mr}$ to the inner rotor, the law of action and reaction implies that the inner rotor applies torque $T_s = T_e = T_{mr}$ to the outer rotor. To keep the mechanical speed of the outer rotor constant, the combustion engine needs to apply driving torque $T_{ms} = T_s = T_e = T_{mr}$.

2.4.4.3. Transmotor in a hybrid electric vehicle

Figure 2.49 visualises a hybrid vehicle equipped with a transmotor. In comparison with Fig 2.45, the transmotor replaces the generator/motor and the planetary gear (a planetary gear decreases the vehicle efficiency and increases the overall cost of the verhicle).

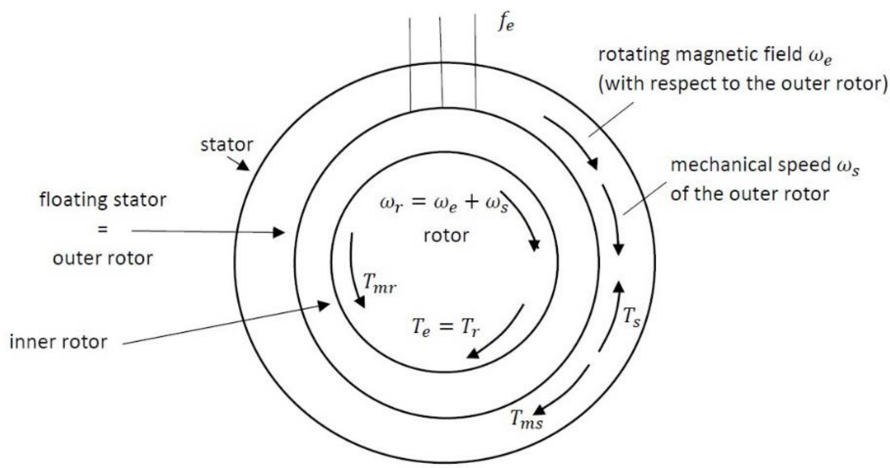


Fig. 2.48. Working principle of a transmotor.

When lock 1 and lock 2 are both released (outer rotor and inner rotor are able to rotate), the combustion engine drives the outer rotor and provides power. The inverter provides electrical power to the outer rotor. The inner rotor receives both powers and drives the vehicle. Hybrid traction is obtained.

When only lock 2 is locked, the inner rotor is locked with the outer rotor. Inner and outer rotors have the same rotational speed and are driven by the combustion engine. There is no electrical power input. Engine traction alone is obtained.

When only lock 1 is locked, the outer rotor is locked, i.e., the outer rotor is not able to rotate. The combustion engine must be shut off or the clutch between engine and transmission (not shown in Fig. 2.49) must be disengaged. There is no power input from the combustion engine. Electrical motor alone traction is obtained.

Regenerative braking is also possible using the configuration in Fig. 2.49. When only lock 1 is locked, the outer rotor is locked. The combustion engine must be shut off or the clutch between engine and transmission (not shown in Fig. 2.49) must be disengaged. There is no power input from the combustion engine. The transmotor behaves as a classical generator. The load is driving the inner rotor of the generator. Mechanical power is converted into electrical power. The generated electrical energy will be stored in the battery.

2.4.5. Combining torque coupling and speed coupling

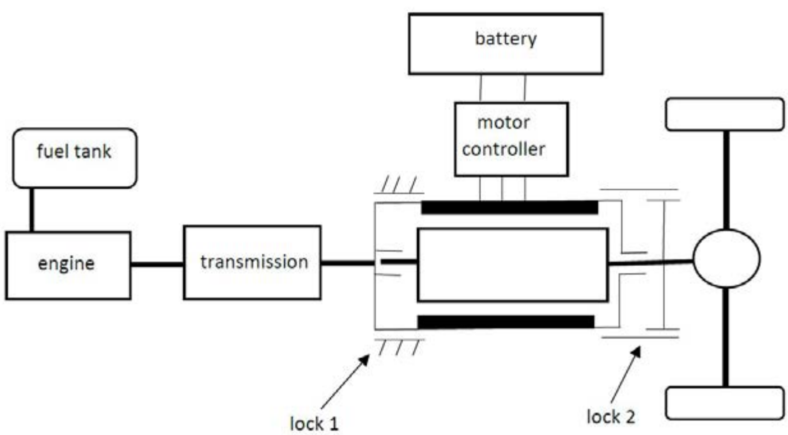


Fig. 2.49. Vehicle with transmotor.

2.4.5.1. Torque coupling and speed coupling using a planetary gear

2.4.5.1.1. Torque coupling

Figure 2.50 visualises a hybrid vehicle with a drivetrain where both torque coupling and speed coupling can be used. Torque coupling is obtained in case

- clutch 1 is engaged (there is a connection) and lock 1 is not locked (the combustion engine, the transmission and the sun gear are able to rotate);
- lock 2 locks the ring gear of the planetary unit, by the sun gear and the carrier gear the power of the combustion engine is sent to the drive shaft;
- clutch 2 is disengaged (there is no connection with the fixed ring gear) implying that the electrical motor can rotate;
- clutch 3 is engaged (there is a connection) implying that the electrical motor is able to drive the sun gear, the carrier gear and the drive shaft.

When considering torque coupling,

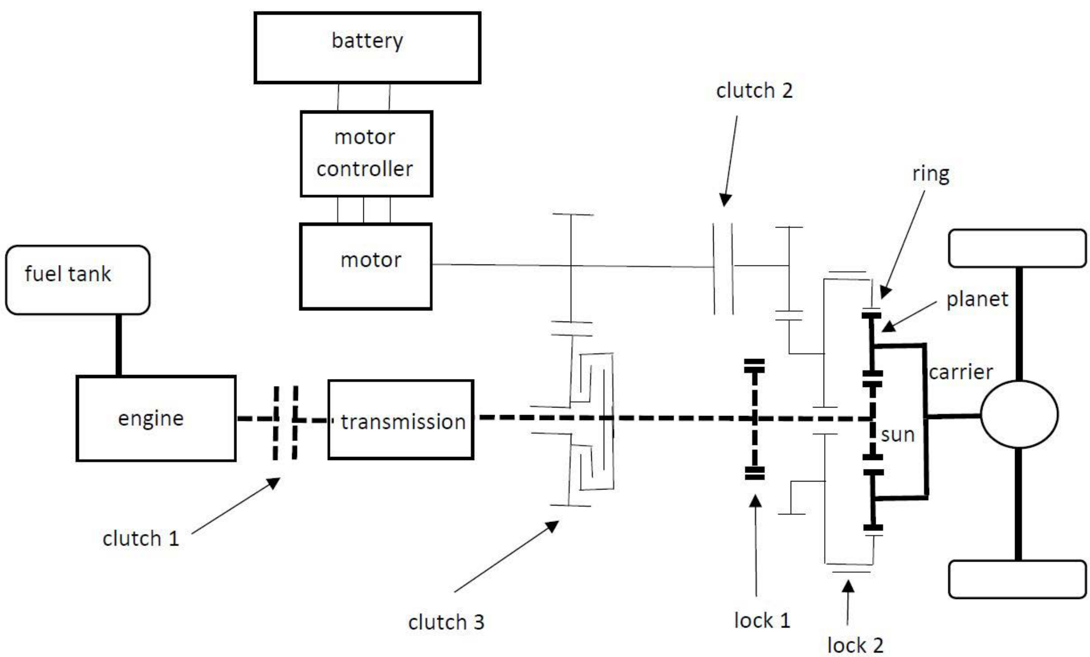


Fig. 2.50. Combining torque and speed coupling in a hybrid vehicle.

- hybrid traction is possible where combustion engine and electrical motor both drive the shaft (combustion engine and electrical motor both apply a torque to the sun);

- the ratio of the speeds of the combustion engine and the electrical motor are fixed, the torques of the engine and the motor are added.

2.4.5.1.2. Speed coupling

The same drivetrain can be used to obtain speed coupling. Speed coupling is obtained in case

- clutch 1 is engaged (there is a connection) and lock 1 is not locked (the combustion engine, the transmission, and the sun gear are able to rotate);

- lock 2 does not lock the ring gear of the planetary unit, by the sun gear and the carrier gear the power of the combustion engine is sent to the drive shaft;

- clutch 3 is disengaged (there is no connection) implying that the electrical motor is not able to drive the sun gear;

- clutch 2 is engaged (there is a connection with the ring gear) implying that the electrical motor drives the ring gear.

When considering speed coupling:

- hybrid traction is possible where combustion engine and electrical motor both drive the shaft;

- the speed coupling is obtained due to the planetary unit, the ratio of the torques of the combustion engine and the electrical motor are fixed, the speeds of the engine and the motor are independent of each other.

2.4.5.1.3. Large flexibility

Not only hybrid traction is possible. By appropriately engaging and disengaging the clutches, also engine alone traction and electrical motor alone traction is possible (and other modes). The choice between torque coupling and speed coupling may depend on the speed of the vehicle.

Suppose the vehicle has a low speed. When using torque coupling, the speed of the combustion engine can be too low. Speed coupling can be used to give the combustion engine a higher speed, since

$$\omega_{out} = k_1\omega_{engine} + k_2\omega_{motor}. \tag{2.41}$$

Indeed, although ω_{out} is low, by giving the electrical machine a negative speed ω_{motor}, a larger combustion engine speed ω_{speed} is possible. Although the vehicle needs a low mechanical power, a larger engine power is possible by sending power to the electrical machine which operates as a generator (power stored in the batteries).

Speed coupling can be useful at high vehicle speed ω_{out}. By avoiding a really high engine speed ω_{engine}, high fuel consumption can be avoided. Indeed, with

$$\omega_{out} = k_1\omega_{engine} + k_2\omega_{motor} \tag{2.42}$$

a high ω_{out} is obtained by having ω_{engine} and ω_{motor} both positive.

Torque coupling can be used when the speed of the vehicle is not too low and not too high. Torque coupling operation mode can be suitable when a high acceleration is needed or during hill climbing. In these situations, a high total ouput torque T_{out} is needed. This high T_{out} is reached when T_{engine} and T_{motor} are both positive with

$$T_{out} = k_1T_{engine} + k_2T_{motor}. \tag{2.43}$$

2.4.5.2.1. Torque coupling

By engaging clutch 1, by disengaging clutch 2 and locking the lock, torque coupling is obtained. The torques of the combustion engine and the electrical motor are added. Since the floating stator (outer rotor) of the transmotor is locked, the synchronous speed of the permanent magnet rotor equals the speed of the combustion engine.

2.4.5.2.2. Speed coupling

By disengaging clutch 1, by engaging clutch 2 and opening the lock, speed coupling mode is obtained. The combustion engine determines the speed of the floating stator (outer rotor), i.e., the combustion engine is driving the floating

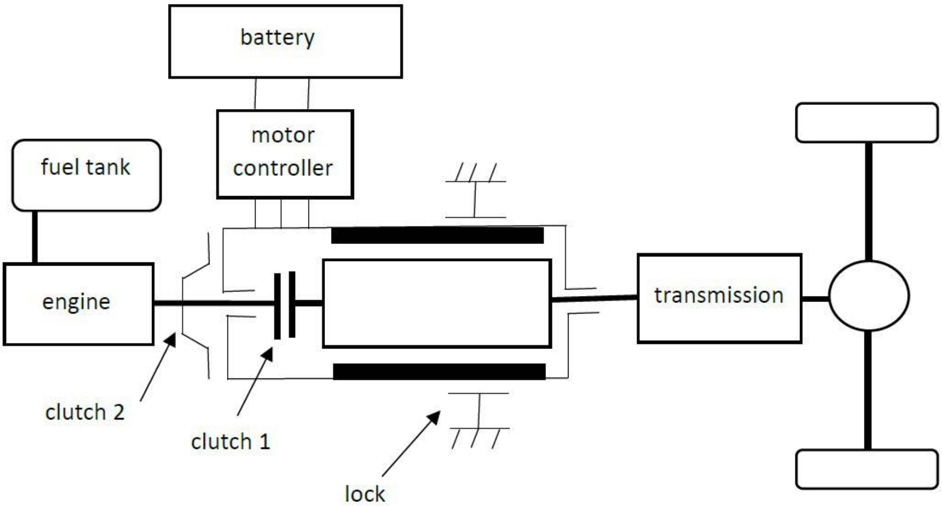


Fig. 2.51. Combining torque and speed coupling in a hybrid vehicle.

stator. The frequency imposed by the inverter on the floating stator determines the speed of the inner rotor (and the drive shaft) with respect to the floating stator. The working principle of the transmotor in Fig. 2.48 is obtained.

2.5. Fuel cell cars

In order to increase the number of electrically driven cars, a wide variety of technical approaches exist. Battery electric vehicles use a battery to store the required energy but, alternatively, energy can also be stored using hydrogen. Hydrogen can be used to feed a hydrogen internal combustion engine, but the use of hydrogen in combination with a fuel cell and an electrical motor is also a well known technology.

In order to reduce the use of fossil fuels, to reduce the emission of harmful exhaust gases and to make cars less noisy, the use of electrically driven cars is an important option. A battery electric vehicle contains a large battery to store a sufficient amount of energy. Using power electronics, the speed of the electrical motor and the speed of the car can be adjusted to the needs of the driver.

Alternatively, the car can be equipped with a hydrogen storage tank allowing to feed the fuel cell stack. Just like a battery, the fuel cell stack is a DC voltage source. Using power electronics, the speed of the electrical motor and the speed of the car can be adjusted to the needs of the driver.

The use of fuel cells dates from the sixties of the previous century. Fuel cells have been used to power a spacecraft, and also the first fuel cell car dates from the sixties. Although fuel cells were considered to be a promising technology by a lot of people, during a long time the use of fuel cells was not really practical. Fuel cells were too expensive and the power density of a fuel cell stack was too low. Due to more recent technological improvements, probably a real future arises for the use of fuel cells (also when conisdering road vehicles).

Prototypes of fuel cell cars are not new at all, but today there also exist commercial vehicles based on fuel cells (e.g. Hyundai, Toyota, Honda). Although the sales figures are still very limited, an important growth is expected.

When dealing with battery electric vehicles, mainly light-duty vehicles are considered. Indeed, mainly private cars are considered. Realising the required energy storage by batteries in case of buses or trucks is far from trivial. When dealing with fuel cell equipped vehicles with hydrogen storage, quite a lot of engineers share the opinion that research is mainly needed towards heavy-duty vehicles. When using hydrogen, more energy can be stored in a smaller volume than it is the case with battery energy storage. The automotive industry used 70 MPa on-board hydrogen energy storage to reach storage densities of approximately 1.4 kWh/kg and 0.8 kWh/L. When considering lithium-ion batteries, only storage densities of approximately 0.24 kWh/kg and 0.5 kWh/L are obtained (source: J. Kurtz et al.). Of course, these energy densities are still lower than, e.g. the energy densities of gasoline with 9.5 kWh/L.

2.5.1. Production of hydrogen

Hydrogen is very common in molecules like water or CH_4, but hydrogen as a gas (i.e. H_2) is almost not available on Earth. Hydrogen as H_2 is a promising energy carrier but it is a secondary source of energy, since it is almost not available on Earth. Based on, e.g. CH_4 or H_2O, it is possible to produce H_2, but this production requires energy. Based on methane (CH_4), a steam reforming process allows to produce hydrogen (H_2). By using electrolysis, water (H_2O) can be split into hydrogen (H_2) and oxygen (O_2).

2.5.1.1. Steam reforming

Today, the majority of the hydrogen production is based on the steam reforming process. At high temperatures (700 °C to 1100 °C), steam (water, H_2O vapor) reacts with methane to produce hydrogen.

$$CH_4 + H_2O \rightarrow CO + 3\ H_2. \tag{2.44}$$

Notice the formation of carbon monoxide CO which is a poisonous gas. By adding a sufficient amount of steam, the carbon monoxide is converted into carbon dioxide.

$$CO + H_2O \rightarrow CO_2 + H_2. \tag{2.45}$$

The total chemical reaction equals

$$CH_4 + 2\ H_2O \rightarrow CO_2 + 4\ H_2, \tag{2.46}$$

which implies that not only hydrogen but also carbon dioxide is produced. In case the hydrogen is used to drive a car using a fuel cell stack and an electrical motor, no harmful CO_2 gas is produced by the car. But the CO_2 gas is produced during the steam reforming process. Possibly the CO_2 is captured and stored, implying that it is not released in the atmosphere.

2.5.1.2. Electrolysis

Today, approximately 5 % of the hydrogen production is based on electrolysis. In its most basic form, electrolysis is used to split up water (H_2O), into hydrogen (H_2) and oxygen (O_2) as visualised in Fig. 2.52. Notice in Fig. 2.52 a bucket filled with pure water. Notice the presence of two identical electrodes (made of an inert material) and, finally, a DC voltage source that is needed to provide the energy needed to split up the water.

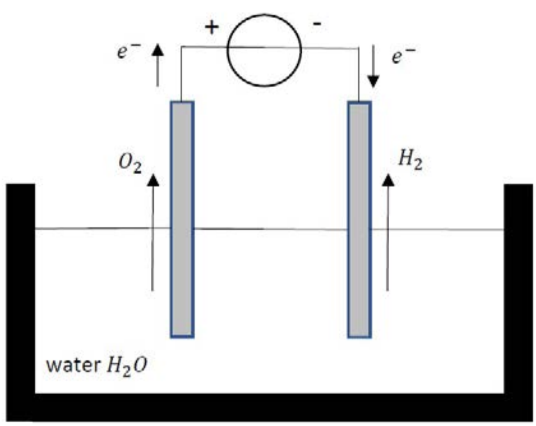


Fig. 2.52. Electrolysis based on pure water.

At the anode, electrons are extracted by the voltage source giving reaction

$$2\,H_2O \rightarrow O_2 + 4\,H^+ + 4\,e^- \tag{2.47}$$

which accounts for the production of oxygen (O_2). At the cathode, electrons are added by the voltage soure giving reaction

$$2\,H^+ + 2\,e^- \rightarrow H_2 \tag{2.48}$$

which accounts for the production of hydrogen, H_2. The overall reaction equals

$$2\,H_2O \rightarrow 2\,H_2 + O_2. \tag{2.49}$$

Notice that both hydrogen and oxygen are produced but the volumetric amount of hydrogen is twice the volumetric amount of oxygen. Actually, the electrolysis process using pure water as visualised in Fig. 2.52 is a slow process (pure water is a bad electrical conductor). The process can become faster (and with a higher efficiency) by adding, e.g. salt (NaCl) or an acid to the water. Indeed, adding an electrolyte increases the conductivity.

As already mentioned, when considering hydrogen production using electrolyses, different approaches exist. The majority of the hydrogen produced by electrolysis is based on the chemical reaction (the chlorakali process)

$$2\,NaCl + 2\,H_2O \rightarrow Cl_2 + H_2 + 2NaOH \tag{2.50}$$

which can be realised as visualised in Fig. 2.53. Often, hydrogen is considered

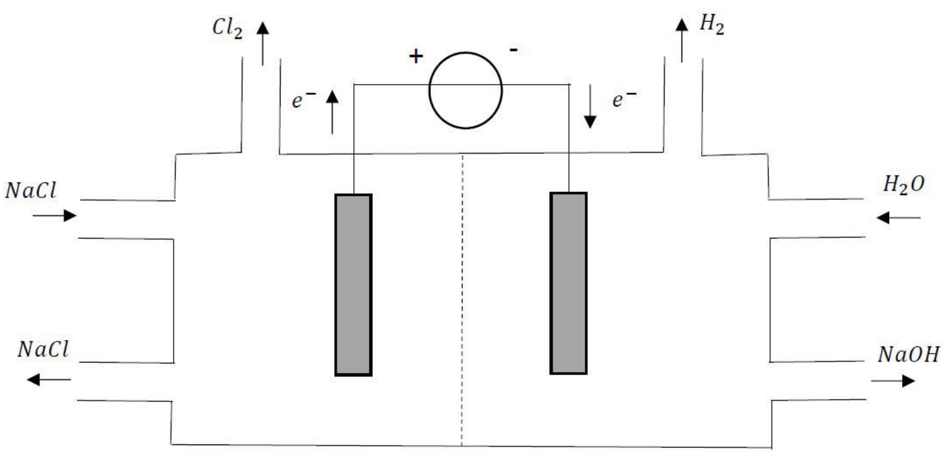


Fig. 2.53. The chlorakali process.

to be a side product of the production of chlorine and caustic soda.

The configuration visualised in Fig. 2.53 shows a bucket containing a left and a right part with a membrane (dotted line) in the middle. Na^+ ions can flow across the membrane in both directions, but other ions like Cl^- and OH^- are blocked by the membrane. On the left side NaCl (actually brine, i.e., salt and water) enters the bucket, part of NaCl reacts, but the majority of the NaCl molecules leave the bucket again. On the right side, water (H_2O) enters the bucket and NaOH is obtained which leaves the bucket. Notice also the production of Cl_2 and H_2.

On the left side, NaCl is split up giving $NaCl \rightarrow Na^+ + Cl^-$. Due to the DC voltage source, electrons are extracted from the left electrode (anode), i.e.,

$$2\,Cl^- \rightarrow Cl_2 + 2\,e^-. \tag{2.51}$$

implying that chlorine gas is produced. The Na^+ ions flow across the membrane to the right part. Additionally, the DC voltage source adds electrons e^- by the cathode. In combination with the water, the reaction

$$2\,H_2O + 2\,Na^+ + 2\,e^- \rightarrow 2\,NaOH + H_2 \tag{2.52}$$

implies that caustic soda and hydrogen gas (H_2) is obtained.

2.5.1.3. Electrolysis: primary energy source

When considering the chemical processes in Figs. 2.52 and 2.53, the primary energy source is the DC voltage source. The electrical energy can originate from thermal power plants (using fossil or nuclear fuels), but a lot of technicians hope to rely on renewable energy. Especially when using photovoltaic energy and energy originating from wind turbines, the power production is very time-dependent, since it depends on the intensity of the sunlight and the wind speed.

Today, the majority of the power generated by photovoltaic panels and wind turbines is injected into the public electrical grid. Especially when the number of photovoltaic panels and wind turbines is increasing, the power grid faces problems to maintain the power balance in the grid. It is indeed important that the generated power always equals the consumed power:

$$P_{gen} = P_{cons}.$$

The power generated by photovoltaic panels and wind turbines is varying and that variation is not correlated with the evolution of the consumed power.

Maintaining the power balance in the grid needs a sufficient number of thermal power plants operating at partial load (or even in standby) to be able adapt the total generated power P_{gen} to the consumed power P_{cons}. Unfortunately, thermal power plants operating at partial load have a lower efficiency.

From that point of view, it is useful to install photovoltaic panels and wind turbines without connecting them to the power grid. The generated energy can be used to produce hydrogen. By storing hydrogen, energy storage is obtained. In case the time-varying renewable energy sources are still connected with the public electrical grid and in case of an excess of power (e.g. when there is a lot of sunlight and wind while P_{cons} is low), the excess of energy can be used to produce hydrogen.

Of course, also a series of financial aspects are relevant. In order to be able to produce hydrogen based on renewable energy sources, a number of financial aspects are important:

- the costs related with the electrical renewable energy sources (like photovoltaic panels, wind turbines) must be reasonable;
- the conversion of the electrical energy into hydrogen based energy storage must be financially affordable;
- the produced hydrogen must be stored, transported and distributed, which also accounts for costs;
- in case the hydrogen is used to supply fuel cells, also the costs related with the production and the use of these fuel cells must be reasonable.

2.5.1.4. Other ways to produce hydrogen

When using electrolysis to produce hydrogen (especially in the case in Fig. 2.52), very pure (but expensive) hydrogen is obtained. Steam reforming and electrolysis are not the only ways to produce hydrogen (although they are at present the most common ones). Alternatively, hydrogen can be produced using photoelectrolysis where the sunlight is used to split the water molecules into oxygen and hydrogen. Biomass gassification is another possibility, and scientists have discovered that some algae and bacteria are able to produce hydrogen.

2.5.2. Storage, transport and distribution of hydrogen

2.5.2.1. Hydrogen storage

Once the hydrogen is produced, it is a challenge to store the hydrogen in a compact and safe way. When burning hydrogen, only water is produced due to the chemical reaction

$$2\,H_2 + O_2 \rightarrow 2\,H_2O\,.$$

By burning 1 kg of hydrogen, 120 MJ of heat will be produced in case the resulting water is released as vapour. If the vapour can be condensed to liquid water, an additional 20 MJ is obtained. This 120 MJ or 140 MJ is approximately three times the energy per unit of mass which is produced when burning gasoline or diesel. Notice, however, that hydrogen is a thin gas at atmospheric pressure and 'real life' temperatures. This implies that hydrogen has a low energy density per unit volume.

In order to store hydrogen in a compact way, a number of approaches exist:

- The hydrogen gas can be stored in pressurized containers (e.g. at a pressure of 300 atmospheres).
- Hydrogen can be stored by absorbing it in a metal, implying that a metal hydride is obtained. The hydrogen can be released by heating.
- Hydrogen can be stored as a liquid by reducing the temperature to –253 °C.

2.5.2.2. Hydrogen transport over long distances

It is important to be able to transport hydrogen over long distances. A possible scenario for the future is installing large amounts of photovoltaic panels in a sunny region like the Sahara. The generated energy is used to produce hydrogen which must be stored and transported to, e.g. Europe.

Liquid hydrogen can be transported using liquid gas tankers (similar to the current transport of liquid natural gas). Of course, special attention is needed for a decent thermal insulation of the tanker in order to be able to transport the cryogenic liquid. Having re-gasified the hydrogen, hydrogen can be pumped through pipelines or it can be transported using road tankers.

2.5.2.3. Distribution of hydrogen

Finally, hydrogen needs to be distributed to the consumers. The applications of hydrogen are very broad.

- Hydrogen can be used in private households for cooking or heating.
- Hydrogen can supply fuel cells to provide electrical energy and heat. The fuel cells can be stationary and can provide a large power if necessary.
- Hydrogen can supply the fuel cell stack in a road vehicle allowing the use of an electrical motor to drive the vehicle. Alternatively, a hydrogen fuelled internal combustion engine can be used.

2.5.2.4. Hydrogen economy

The global approach where hydrogen is produced, transported over large distances and distributed to the consumers is the so-called 'hydrogen economy'. Figure 2.54 visualises the first part of this hydrogen economy where hydrogen is produced and stored (liquid hydrogen storage after the liquefaction process).

Figure 2.55 visualises the second part of the hydrogen economy where liquified hydrogen is transported over long distances. After the second liquified hydrogen storage, the liquified hydrogen is distributed and finally converted back into hydrogen gas. The hydrogen gas is distributed to a broad range of consumers by using road tankers or pipelines. Although there is a very broad range of consumers (Fig. 2.56), it is expected that fuel cell vehicles can become an important application of this hydrogen economy.

2.5.3. Fuel cell vehicles

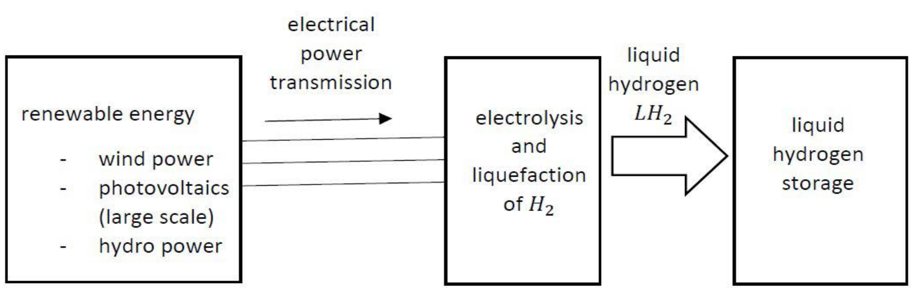


Fig. 2.54. The hydrogen economy: Part 1.

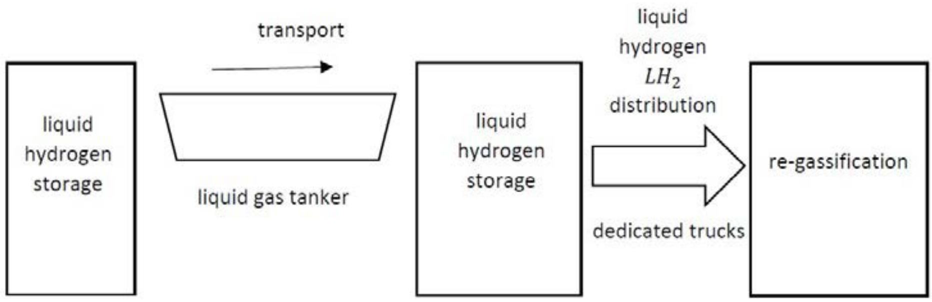


Fig. 2.55. The hydrogen economy: Part 2.

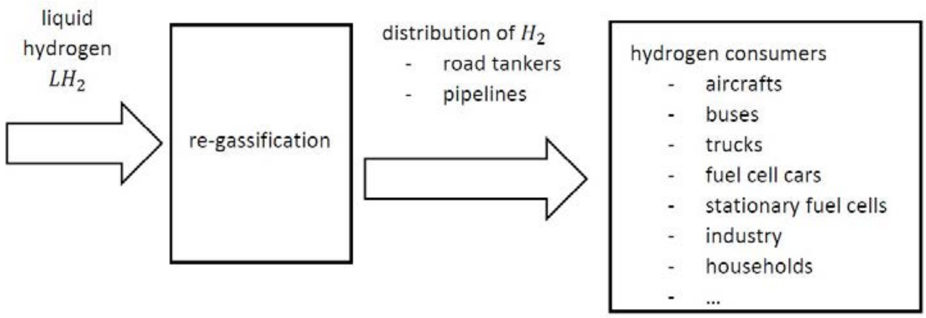


Fig. 2.56. The hydrogen economy: Part 3.

Figure 2.57 visualises a fuel-cell electric vehicle (FCEV). The wheel axle is driven by an electrical motor using a transmission to obtain the required speed. The use of a proton exchange membrane fuel cell where the fuel cell stack is fed by hydrogen is the most popular one. A hydrogen tank is needed to store hydrogen.

Because of no harmful emissions occur to produce the hydrogen, a FCEV is a zero emission car which is useful in, e.g., urban areas. In comparison with the charging process of a battery, refueling with hydrogen is fast (it only takes a few minutes).

2.5.3.1. Fuel-cell technologies

There exist a broad range of fuel-cell types each having their properties, advantages and disadvantages. Not all types of fuel-cells will be discussed here, we restrict ourselves to a selection.

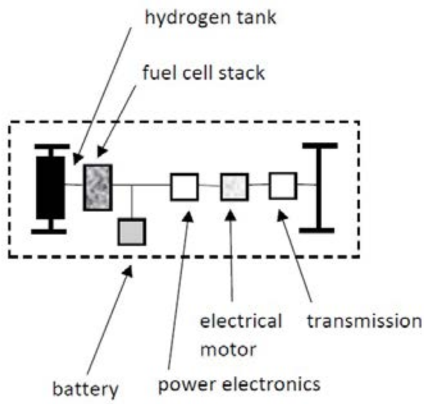


Fig. 2.57. Fuel-cell electric vehicle.

2.5.3.1.1. Proton exchange membrane fuel cell

The working principle of a proton exchange membrane fuel cell (PEMFC) is visualised in Fig. 2.58.

PEMFC contains an oxygen electrode and a hydrogen electrode. Between these electrodes, a proton exchange membrane allows a flow of protons, i.e., H^+ ions. Hydrogen gas is supplied to the hydrogen electrode, and this hydrogen is split into H^+ ions and electrons e^- due to a catalyst (platinum). By crossing the membrane, the H^+ ions move to the oxygen electrode. Notice also that the electrical load (modelled by a resistor R in Fig. 2.58) and electrons move from the hydrogen electrode to the oxygen electrode, i.e., a DC current is flowing in the electrical load. H^+ ions and electrons e^- move to the oxygen electrode and

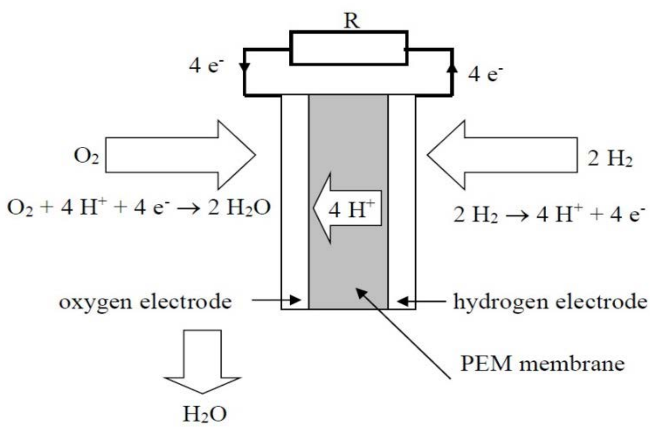


Fig. 2.58. Working principle of a proton exchange membrane fuel cell.

by adding oxygen, water (H_2O) is obtained (also here platinum is needed as a catalyst).

PEMFC is a low temperature fuel cell which operates at temperatures between 40 °C and 60 °C. Due to the low temperature, the PEMFC starts up quickly, which is important in case of a road vehicle. A PEMFC converts the chemical energy available in hydrogen into electrical energy (with an efficiency of approximately 50 % to 60 %) and heat. In case also the heat is used, the CHP (Combined Heat and Power) principle is obtained.

Notice that the electrochemical reaction between hydrogen and oxygen generates only limited voltage (typically even less than 1 V). This implies that a series connection of individual cells is needed (a layered structure is used) to obtain the required voltage level.

2.5.3.1.2. Alkaline fuel cell

Alkaline fuel cell (AFC) is also a low temperature fuel cell like the PEMFC. The AFC has a somewhat higher efficiency than the PEMFC (when considering electrical energy as the useful output), i.e., efficiencies of 70 % are potentially possible. AFC has been used in the sixties in the Apollo-series missions and also in the Space Shuttle. AFC is also useful in the transportation sector.

The working principle of an AFC is visualised in Fig. 2.59.

AFC contains an anode and a cathode. Between these electrodes, an electrolyte allows a flow of hydroxyl OH^- ions. Hydrogen gas is supplied at the anode, and in combination with hydroxyl ions water and electrons e^- are obtained. Notice also that the electrical load (modelled by resistor R in Fig. 2.59) and electrons move from the anode to the cathode, i.e., a DC current is flowing in the electrical load. At the cathode not only oxygen is supplied, but also water is needed. The electrons from the electrical load are needed to obtain the desired hydroxyl OH^- ions. Also in the case of the AFC, platinum is used as a catalyst.

2.5.3.1.3. Direct methanol fuel cell

Direct methanol fuel cell (DMFC), visualised in Fig. 2.60, does not use hydrogen as energy source as it is the case for the PEMFC or the AFC. Methanol CH_3OH is used as energy source and has the advantage to be more energy-dense than hydrogen. Unfortunately, DMFC has a much lower efficiency (approximately

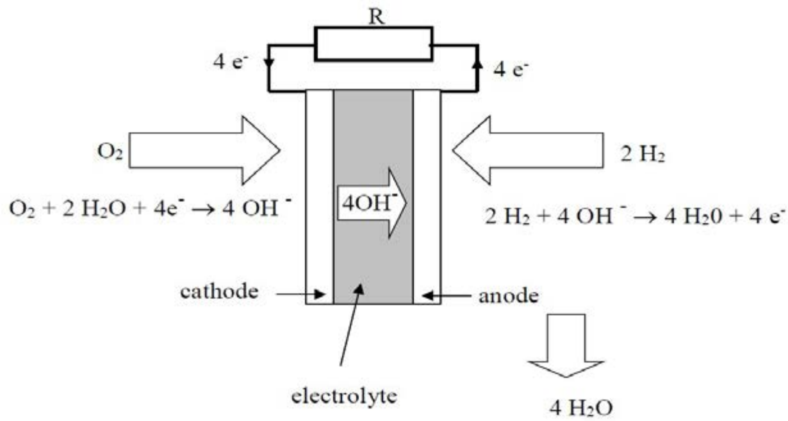


Fig. 2.59. Working principle of an alkaline fuel cell

10 %) than a PEMFC or an AFC. When engineers were able to increase the efficiency of the DMFC, much more applications would arise and, possibly, a methanol economy instead of a hydrogen economy could be developed.

A DMFC contains an anode and a cathode. Between these electrodes, a proton exchange membrane allows a flow of protons, i.e., H^+ ions. Methanol (CH_3OH) is supplied to the anode and in combination with water, H^+ ions and electrons e^- are obtained in combination with CO_2. Notice indeed, that CO_2 exhaust is obtained (contrary to the PEMFC or the AFC).

By crossing the membrane, the H^+ ions move to the cathode. Notice also that the electrical load (modelled by resistor R in Fig. 2.60) and electrons move from the anode to the cathode, i.e., a DC current is flowing in the electrical load. H^+ ions and electrons e^- move to the cathode and by adding oxygen, water (H_2O) is

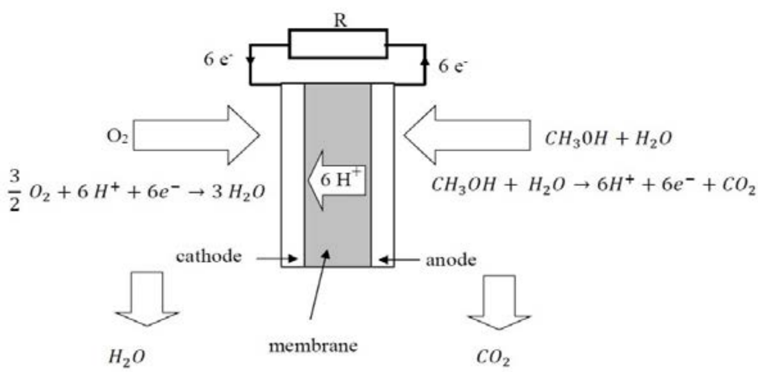


Fig. 2.60. Working principle of a direct methanol fuel cell

obtained (also here platinum is needed as a catalyst).

Due to the compact energy storage using methanol, the DMFC is used in mobile phones, laptops, etc. where smaller powers and smaller amounts of energy are needed (the low efficiency must be taken into consideration). Although some references mention the use of DMFC in transport systems, the PEMFC and the AFC are more important fuel cell types here.

2.5.3.1.4. Other fuel cell types

PEMFC, AFC and DMFC are low temperature fuel cell types. Although there exist other types of low temperature fuel cells, they are the most important types. There also exist so-called high temperature fuel cells like phosforic acid fuel cell (PAFC), protonic ceramic fuel cell (PCFC), molten carbonate fuel cell (MCFC) or solid oxide fuel cell (SOFC). These high temperature fuel cells are used in large scale energy applications including CHP applications (combined heat and power). Due to the long start-up times of these high temperature fuel cells, they generally have no applications when driving road vehicles.

2.5.3.2. The drivetrain of a fuel-cell electric vehicle

A global overview of a drivetrain of a fuel-cell electric vehicle is shown in Fig. 2.61. Notice first of all the fuel cell stack which is a DC voltage source which provides the power/energy needed to drive the vehicle. The fuel cell stack is almost always a PEMFC, and notice also that hydrogen storage is needed (e.g. using a dedicated high pressure tank). A DC to DC converter is used to adjust the DC voltage level of the fuel cell to the desired value, and using an inverter the electrical motor (it is indeed an AC motor) is fed. The output voltage of the inverter has an adjustable RMS value and an adjustable frequency which allows to control the motor behaviour (e.g. the speed).

The configuration in Fig. 2.61 also contains battery energy storage. When braking the car, the kinetic energy is converted into electrical energy by the electrical motor/machine which behaves as a generator. The generated electrical energy is stored in the battery. When needed, it is also possible to charge the battery using energy from the fuel cell.

The battery is also able to assist the fuel cell when the vehicle needs large power (e.g. during acceleration). Possibly in parallel with the battery, additional supercapacitors are available. Also these supercapacitors can store and provide energy (supercapacitors are well suited to provide large power peaks).

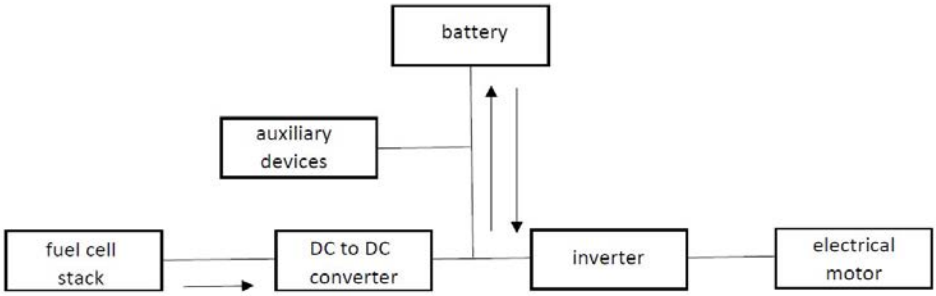


Fig. 2.61. Drivetrain of a fuel-cell electric vehicle

2.5.3.2.1. Plug-in fuel cell vehicles

In the configuration in Fig. 2.61, all energy originates from hydrogen. Alternatively, there also exist plug-in fuel cell vehicles (PFCV). Such a PFCV has a larger battery and a smaller fuel cell. If hydrogen is made of renewable sources and the electrical energy needed to charge the battery comes from renewable sources as well, the vehicle configuration supports the introduction of renewable energy in the society.

2.5.4. Challenges and obstacles when introducing fuel-cell electric vehicles

A breakthrough of fuel cell based vehicles needs a decent hydrogen distribution system as it is available for diesel or gasoline today. At present, the number of refueling points is still too limited. On the other hand, refueling with hydrogen is a fast process requiring only a few minutes (contrary to charging a battery which takes a lot of time).

Quite a lot of people share the opinion that fuel cell vehicles mainly have a future when considering heavy-duty vehicles like buses. The supply of hydrogen is very important, and buses have the opportunity to be able to refuel in one central place, i.e., only one single refilling point is needed. Moreover, fuel cells are expensive (although the prices are decreasing, they are still expensive). The large investment implies that the fuel cell vehicles need to be used for many hours each day in order to be cost-effective. Additionally, a breakthrough of fuel cells will imply larger production numbers and mass production. This mass production is expected to decrease the investment costs when constructing a fuel cell car.

Moreover, the durability of a fuel cell stack must be sufficiently high (especially since a fuel cell stack is expensive). All components of the fuel cell degrade over time from usage. The degradation accelerates when the fuel cell is exposed to extreme conditions or when the performance limits are exceeded.

By overdimensioning the system, it is possible to prevent an exceedance of these limits, but this overdimensioning accounts for a further increase of the investment costs.

To avoid extreme conditions in the fuel cell and increase the lifespan, various sensors are used to monitor the main physical quantities. It is important to maintain the optimal operational conditions. Real-time controllers allow to diagnose the health status of the fuel cell.

In the early stage of fuel cell technolgy, fuel cells barely survived a few hundred hours of operation. Today, attempts are made to reach lifespans of 10 000 hours (the use of durable materials is also needed).

Another aspect is the uncertainty about hydrogen safety. For instance, hydrogen tanks can explode, since the hydrogen is possibly stored at a pressure of 70 MPa (700 times the atmospheric pressure). Moreover, hydrogen is flammable when it is mixed with oxygen. It is a challenge to construct decent hydrogen tanks. In case of a car crash, it is important the hydrogen tank deforms and does not crack, which prevents an escape of the hydrogen. In case the hydrogen tank is exposed to a fire (and a temperature rise occurs), a thermally activated pressure-relief device is needed. This thermally activated pressure-relief device will open an emergency vent when the temperature is too high.

Bibliography

M. H. Rashid (ed), Power Electronics Handbook: devices, circuits and applications, Elsevier, London, 2011, ISBN 978-0-12-382036-5.

L. Ribickis, Power Electronics, RTU Press, Riga, 2015, ISBN 978-9934-10-602-6.

S. Srdic, S. Lukic, Toward Extreme Fast Charging, IEEE Electrification Magazine, vol. 7, no. 1: Overcoming Hurdles for Electric Vehicles, March 2019, pp. 22–31.

G. Tamai, What are the Hurdles to Full Vehicle Electrification?, IEEE Electrification Magazine, vol. 7, no. 1: Overcoming Hurdles for Electric Vehicles, March 2019, pp. 5–11.

X.-G. Yang, S. Ge, N. Wu, Y. Mao, F. Sun, C.-Y. Wang, All-Climate Battery Technology for Electric Vehicles, IEEE Electrification Magazine, vol. 7, no. 1: Overcoming Hurdles for Electric Vehicles, March 2019, pp. 12–21.

A. R. Bhatti, Z. Salam, M. J. Bin Abdul Aziz, K. P. Yee, R.H. Ashique, Electric vehicles charging using photovoltaic: Status and technological review, Renewable and Sustainable Energy Reviews, vol. 54, 2016, pp. 34–47.

D. Dobrzanski, Overview and characteristics of the EV fast charging connector systems, Maszyny Elektryczne, Zeszyty Problemowe, no. 3, 2017 (115), pp. 91-96, retrieved from http://www.komel.katowice.pl/ZRODLA/FULL/115/ref_15.pdf.

D'Ieteren, Info: De elektrische wagen, May 2013, retrieved from http://dbtechnology.be/wp-content/uploads/2014/02/De-elektrische-wagen.pdf (in Dutch).

J. S. Johansen, Fast-Charging Electric Vehicles using AC, master's thesis, Technical University of Denmark, September 2013.

J. Martinez-Lao, F. G. Montoya, M. G. Montoya, F. Manzano-Agugliaro, Electric vehicles in Spain: An overview of charging systems, Renewable and Sustainable Energy Reviews, Elsevier, vol. 77, 2017, pp. 970–983.

B. Rotthier, P. Blockx, P. Van den Bossche, J. Cappelle, Thuis elektrisch opladen: hoe uw wagen verschilt van uw smartphone, retrieved from https://evladen.be/theobrochure.pdf (in Dutch).

Amsterdam Roundtables Foundation, E Volution: Electric vehicles in Europe: gearing up for a new phase?, Amsterdam Roundtable Foundation and McKinsey & Company, Amsterdam, The Netherlands, April 2014.

G. Boyle, Renewable Energy: Power for a Sustainable Future, Oxford University Press, 2012, ISBN 978 0 19 954533 9.

EG & G Technical Services, Fuel Cell Handbook, U.S. Department of Energy, 2004.

C. Gearhart, Hydrogen: What's Different Now, IEEE Electrification Magazine, vol. 6, no. 1: Hydrogen Fuel Cells for Transportation, March 2018, pp. 4–7.

G. Hoogers (ed.), Fuel Cell Technology Handbook, CRC Press, London, 2003, ISBN 0-8493-0877-1.

Jungik Kim, Saehoon Kim, Obstacles to the Success of Fuel-Cell Electric Vehicles, IEEE Electrification Magazine, vol. 6, no. 1: Hydrogen Fuel Cells for Transportation, March 2018, pp. 48–54.

J. Kurtz, M. Peters, M. Muratori, C. Gearhart, Renewable Hydrogen – Economically Viable: Integration in the U.S. transportation sector, IEEE Electrification Magazine, vol. 6, no. 1: Hydrogen Fuel Cells for Transportation, March 2018, pp. 8–18.

T. E. Lipman, A. Z. Weber (eds.), Fuel Cells and Hydrogen Production, A Volume in the Encyclopedia of Sustainable Science and Technology, 2nd edition, Springer, New York, 2019, ISBN 978-1-4939-7788-8.

Y. Manoharan, S. E. Hosseini, B. Butler, H. Alzharani, B. Thi Fou Senior, T. Ashuri, J. Krohn, Hydrogen Fuel Cell Vehicles: Current Status and Future Propect, Applied Sciences, vol. 9, 2296, June 2019, doi:10.3390/app9112296, www.mdpi.com/journal/applsci.

E. Muljadi, S. Hossain, Powering Transportation with Hydrogen Fuel Cells, IEEE Electrification Magazine, vol. 6, no. 1: Hydrogen Fuel Cells for Transportation, March 2018, pp. 2–3.

Bandi Mallikarjuna Reddy, Paulson Samuel, Narapureddy Siva Mohan Reddy, Gorvernment Policies Help Promote Clean Transportation in India, IEEE Electrification Magazine, vol. 6, no. 1: Hydrogen Fuel Cells for Transportation, March 2018, pp. 27–36.

F. Un-Noor, S. Padmanaban,L. Mihet-Popa, M. Nurunnabi Mollah, E. Hossain, A Comprehensive Study of Key Electric Vehicles (EV) Components, Technologies, Challenges, Impacts and Future Direction of Development, Energies, Energies, vol. 10, 1217, August 2017, doi:10.3390/en10081217, www.mdpi.com/journal/energies.

Ali Emadi (ed.), Handbook of Automotive Power Electronics and Motor Drives, CRC Press: Taylor & Francis, London, 2005, ISBN 978-0-8247-2361-3.

A. Ghayebloo, A. Radan, Superiority of Dual-Mechanical-Port-Machine-Based Structure for Series-Parallel Hybrid Electric Vehicle Applications, IEEE Transactions on Vehicular Technology, vol. 65, no. 2, February 2016, pp. 589–602.

H. Nasiri, A. Radan, A. Ghayebloo, K. Ahi, Dynamic Modeling and Simulation of Transmotor Based Series-Parallel HEV Applied to Toyota Prius 2004, 10th International Conference on Environment and Electrical Engineering, Rome, Italy, May 8–11, 2011.

F. Un-Noor, S. Padmanaban, L. Mihet-Popa, M. Nurunnabi Mollah, E. Hossain, A

Comprehensive Study of Key Electric Vehicles (EV) Components, Technologies, Challenges, Impacts and Future Direction of Development, Energies, Energies, vol. 10, 1217, August 2017, doi:10.3390/en10081217, www.mdpi.com/journal/energies.

Bibliography: websites

Lithium-ion batteries: https://www.explainthatstuff.com/how-lithium-ion-batteries-work.html

Lithium-ion batteries: https://en.wikipedia.org/wiki/Lithium-ion_battery

CCS: Combined Charging System: https://en.wikipedia.org/wiki/Combined_Charging_System

CHAdeMO: https://en.wikipedia.org/wiki/CHAdeMO

Tesla Supercharger: https://en.wikipedia.org/wiki/Tesla_Supercharger

Alkaline Fuel Cell: https://en.wikipedia.org/wiki/Alkaline_fuel_cell

Direct Methanol Fuel Cell: https://en.wikipedia.org/wiki/Direct_methanol_fuel_cell

Fuel Cells: https://en.wikipedia.org/wiki/Fuel_cell

Hydrogen: pdf document retrieved from https://www.need.org/Files/curriculum/infobook/HydrogenS.pdf

Hydrogen: https://en.wikipedia.org/wiki/Hydrogen

Hydrogen production: https://en.wikipedia.org/wiki/Electrolysis_of_water and https://en.wikipedia.org/wiki/Chloralkali_process

Chapter 3:

The impact of electric vehicles on the power grid

Joan Peuteman
KU Leuven

3.1. The problem of vehicle emissions

At present, the use of cars equipped with an internal combustion engine dominates the modern society. But concerns about greenhouse gas emissions (like CO_2), air pollution in mainly urban areas and also the dependency on the use of fossil fuels push the society to look for alternatives.

One of these alternatives is the use of electric cars where the power electronic devices and the traction motors are fed by a battery (or alternatively a fuel cell). The limited energy density when battery energy storage is used, discourages the breakthrough of full electric cars. In order to increase the range of the car, hybrid electric cars/vehicles have been developed.

A hybrid electric car combines the use of an electrical traction motor and an internal combustion engine. Due to the compact energy storage when using, e.g. diesel or gasoline, a large range is obtained. The battery fed electrical motor can, e.g. provide additional torque when accelerating the car. Possibly, the electrical motor drives the car when making shorter trips or when using the car in an urban area.

In a hybrid vehicle (hybrid electric vehicle: HEV), the battery can be charged by using the traction motor as a generator which is driven by the combustion engine. In such a situation, the energy stored in the battery originates from the diesel or gasoline tank. A plug-in hybrid electric vehicle (PHEV) is a hybrid vehicle having a larger battery. By making a connection with the electrical power grid, energy from that power grid is used to load the battery. Full electric vehicles also need the power grid to charge the battery (battery electric vehicle: BEV).

In the present text, we use the abbreviation PEV to describe a plug-in electrical vehicle. Two types of PEV can be considered: a PHEV and a BEV.

With a limited amount of PEVs, PHEVs and BEVs the impact on the electrical power grid is limited. But when the numbers of PEVs, PHEVs and BEVs are increasing, their impact on the electrical power grid is also increasing. This impact will be studied here because the power grid must be prepared to face these challenges. The impact on the electrical power grid contains a number of different aspects:

- When considering the power balance of the grid, it is important that the generated and the comsumed power equal each other (here neglecting the heat losses in the grid): more precisely $P_{gen} = P_{cons}$.

- The PEVs are connected with the low voltage distribution grid. In case single phase connections are used to charge the battery of the car, it is important that the charging cars are spread among the three phases to obtain a more or less symmetric load of the three phase grid.

- PEVs account for an additional load of the low voltage grid. This implies an impact on the load of the distribution transformer and the feeders of the grid. Due to the higher current levels, more heat losses occur. Moreover, larger current accounts for voltage drops which affect the resulting voltage levels in the grid.

- In general, the battery charger accounts for non sinusoidal currents which are extracted from the power grid. Indeed, a battery charger is a nonlinear load, as it contains a power electronic converter. Due to the grid impedance, these non sinusoidal currents account for non sinusoidal voltage drops and non sinusoidal voltages at the nodes of the grid. Fourier analysis shows that these non sinusoidal currents and voltages can be considered as the sum of a first order harmonic and higher order harmonics, i.e., harmonic pollution of the grid occurs.

In literature, a large amount of information is available. Different researchers also show different, often complementary, approaches to face the technical challenges of the power grid. In the present text, we will mainly consider two papers which are available in literature. When more detailed information is needed, consult the original papers and the references mentioned in these papers.

3.2. Impact of PEVs on the distribution grid: Paper 1

3.2.1. Initial situation

"Assessment of Plug-in Electric Vehicles Charging Impacts on Residential Low Voltage Distribution Grid in Hungary" by H. Ramadan et al. considers a case study with a low voltage grid which is available in Budapest (Hungary). The structure of the grid is visualised in Fig. 3.1.

The power grid is a radial three phase distribution grid. Line voltages of 400 V are available in combination with phase voltages of 231 V between a phase conductor and the neutral conductor. Without the charging points of PEVs, a total of 139 residential loads are available. These residential loads are single phase loads and they are evenly distributed to the three phases. The average daily load profile of these loads is visualised in Fig. 3.2.

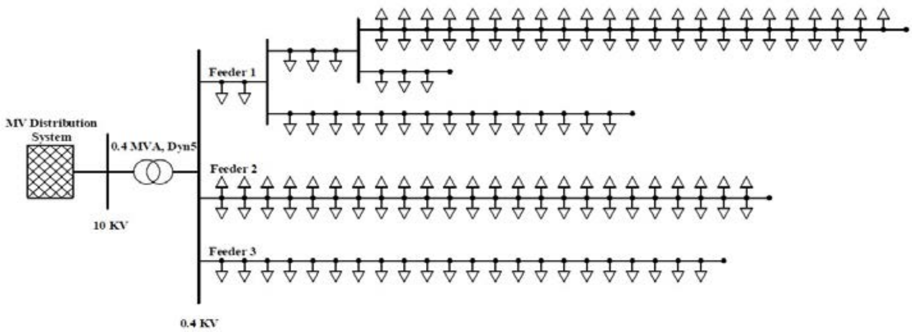


Fig. 3.1. Power grid localized in Budapest (source: Ramadan et al.).

Figure 3.2 visualises the average daily load profile per unit of the grid visualised in Fig. 3.1. Notice a real peak consumption around 8 pm and a smaller peak consumption around 11 am. During the night, the power consumption is small in comparison with the consumption during the day.

A load profile, as visualised in Fig. 3.2, depends on a number of parameters, including the following ones:

- A distinction can be made between load profiles visualising the power distributed by low voltage grids (feeding a large number of small loads) and the total power consumption of e.g. a country. The latter also includes the power consumption of e.g. large industrial factories.

- Load profiles are country dependent.

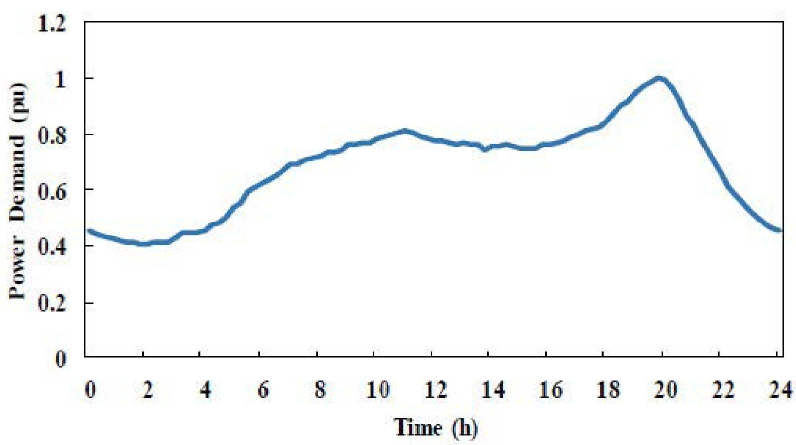


Fig. 3.2. Average daily load profile (source: Ramadan et al.).

- Load profiles depend on the climate and the season. When the weather is warm, people use airconditioning to keep the living rooms cool. When the weather is cold, people use electrical heating.

- Load profiles are influenced by the dominant culture and religion (e.g. different religions have different rest days, and during a rest day the electrical energy consumption is lower, since industrial factories consume less energy).

3.2.2. Additional power consumption by PEVs

In the paper by H. Ramadan et al., the battery of the PEV is loaded by a 3.3 kW single phase charger. The charging is performed with a constant 3.3 kW power at home. Energy E_C needed to fully charge the battery depends on the initial SOC (state of charge) of the battery, the total energy storage capacity C and the efficiency $\eta_{charger}$ of the battery charger. More precisely

$$E_C = (1 - \frac{SOC}{100}) \cdot \frac{C}{\eta_{charger}} . \qquad (3.1)$$

The SOC represents the remaining energy in the battery when plugging in the car. This happens when the car driver returns home after a drive. Due to different driving distances, different SOC values are encountered, and this is probabilistic data. A typical probabilistic density function of the battery SOC when the charging starts, is given in Fig. 3. Notice in Fig. 3.3 that approximately one third of the cars have an initial SOC of about 50 %. Almost no cars have empty batteries or full batteries at the end of the day.

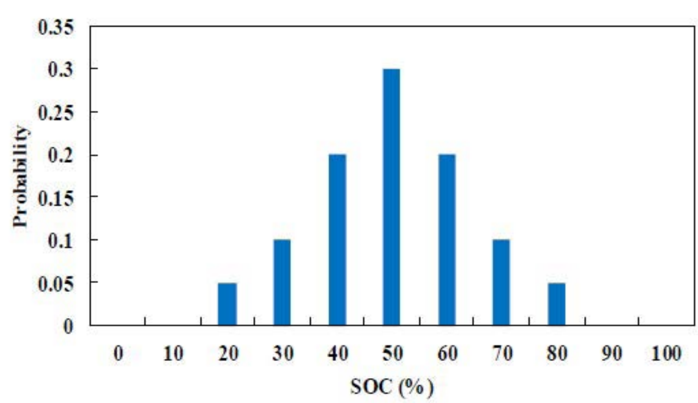


Fig. 3.3. Probability density function of battery SOC (source: Ramadan et al.).

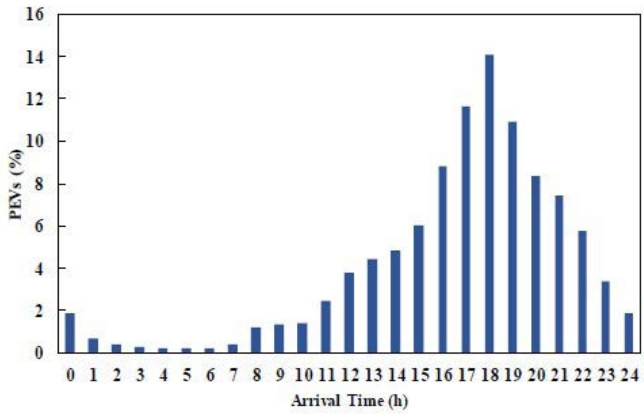


Fig. 3.4. PEV's arrival time (source: Ramadan et al.).

A distinction is made between uncoordinated charging of the batteries and delayed (or coordinated) charging of the batteries. In case of uncoordinated charging, the PEV starts charging when the driver arrives home having finished the last journey of that day. Journey arrival data can be obtained from transportation reports (when considering the USA, the National Household Travel Survey (https://nhts.ornl.gov/) provides a lot of information). Figure 3.4 visualises PEV's arrival time according to NHTS-2009. People mainly arrive home in the early evening.

3.2.3. Uncoordinated charging

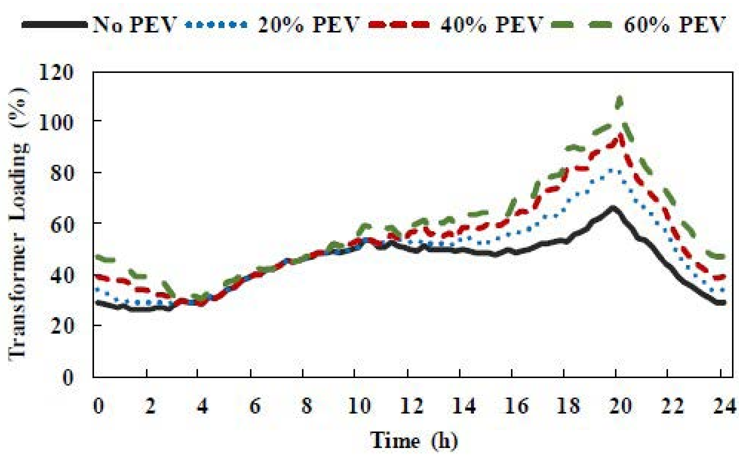


Fig. 3.5 Transformer loading with uncoordinated charging (source: Ramadan et al.).

In case of uncoordinated charging, the arrival times in Fig. 3.4 determine when the battery charging starts. Especially between 5 pm and 7 pm, a lot of cars start charging their batteries, which has an impact on the total daily load profile. Such a total daily load profile is visualised in Fig. 3.5.

The total daily load profile gives the evolution of the power provided by the Dyn distribution transformer in Fig. 3.1. The penetration level of PEVs is important. When no battery charging occurs, the load profile in Fig. 3.2 is obtained. Especially when the penetration level is increasing, the impact on the power grid is also increasing. This consideration leads to the study of three penetration levels: 20 %, 40 % and 60 %. In the afternoon and especially during the evening, additional power consumption occurs. In the late night and the early morning, almost no vehicle charging occurs.

When considering the electrical power generation and the use of all grid components, it is preferable to have a more or less constant load profile, i.e., the ratio between the maximum power consumption and the minimum power consumption approaches as much as possible. When considering the original load profile (without car battery charging), the ratio between the maximum and the minimum power consumption approximately equals 2.5. Due to the uncoordinated charging of batteries, the ratio between the maximum and the minimum power consumption increases and approximately a value of 4.5 is obtained in case of a 60 % PEV penetration level. Indeed, during the night, when the minimum power consumption occurs, only a limited amount of batteries are charged. Around 8 pm, when the original load profile already reaches a maximum, a lot of car batteries are charged. Using a coordinated (delayed) battery charging, a more constant load profile can be obtained.

Notice that the transformer maximum loading in Fig. 3.5 equals approximately 110 % at 60 % PEV penetration level. Due to an overloading of the transformer, the transformer losses increase implying higher temperatures, which reduces the life expectancy.

3.2.4. Coordinated charging

The coordinated battery charging scenario, described in the paper by H. Ramadan et al., mainly loads the batteries during the off-peak period. The battery charging starts at four different instants of time in the off-peak period: 10 pm, 11 pm, 3 am and 4 am. The vehicles having a battery with an initially low SOC get priority and can start earlier with their loading procedure. Approximately around 7 am, all batteries have been charged. The vehicles are ready to be used during the day time.

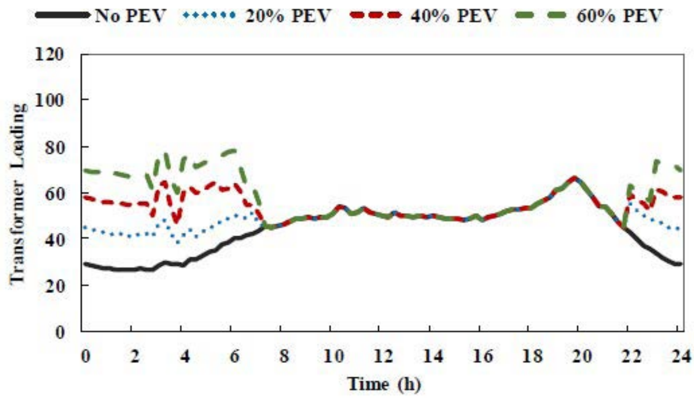
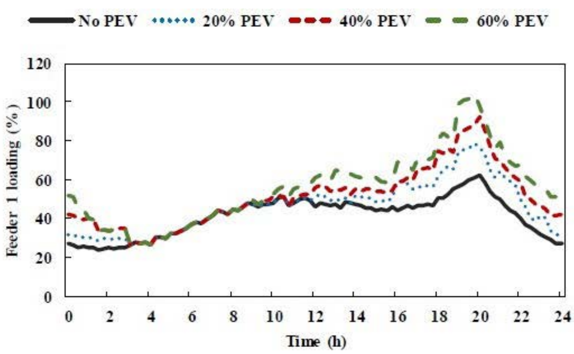


Fig. 3.6. Transformer load profile in case of coordinated charging (source: Ramadan et al.).

Figure 3.6 visualises the load profile of the distribution transformer in case of different penetration levels of PEV battery charging. When considering the case of 60 % of PEV penetration level, the ratio between the maximum and the minimum power consumption approximately equals (which is even lower/better than the situation without battery charging). The maximum power consumption appears around 7 am, the minimum power consumption appears during the day.

Figure 3.7 visualises the load profile of feeder 1 in Fig. 3.1 (which is the most loaded feeder in the grid) in case of uncoordinated and coordinated battery charging. When comparing with the load profile of the distribution transformer, similar conclusions arise.



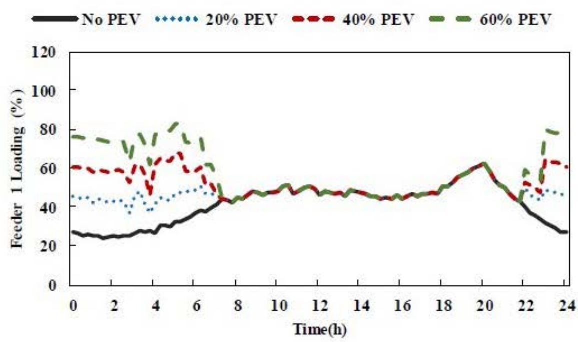


Fig. 3.7. Feeder 1 load profiles in case of uncoordinated (left) and coordinated (right) charging (source: Ramadan et al.).

3.2.5. Voltage deviations

In case the power consumption increases, the current level increases accounting for a larger voltage drop across the grid impedance. Figure 3.8 visualises the voltage deviations in feeder 1. Notice that the coordinated battery charging accounts for lower voltage deviations, which is an important advantage.

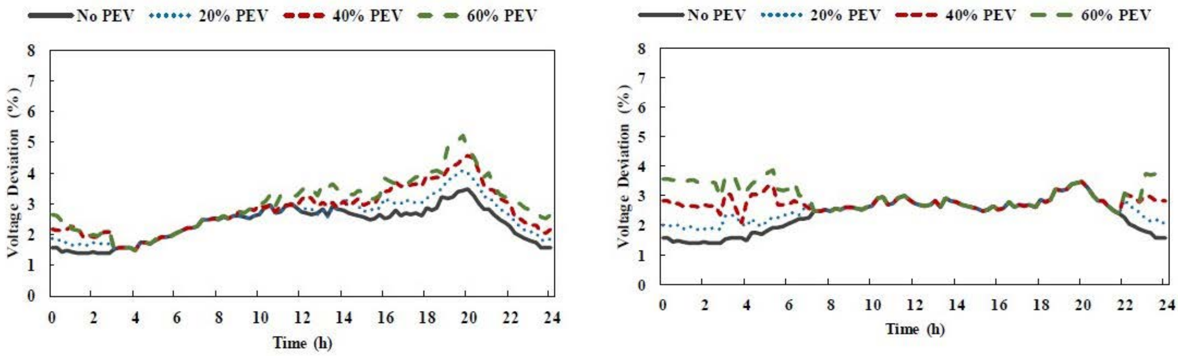


Fig. 3.8. Feeder 1 voltage deviations in case of uncoordinated (left) and coordinated (right) charging (source: Ramadan et al.).

3.2.6. Heat losses

In case the power consumption increases, the current level increases accounting for larger heat losses (Joule effect). By charging the batteries additional heat losses occur, especially when the PEV penetration level increases. By applying coordinated instead of uncoordinated charging of the batteries, a reduction of heat losses is obtained. By reducing the high current values, heat losses reduce, since heat losses are proportional to square of the rms value of the current.

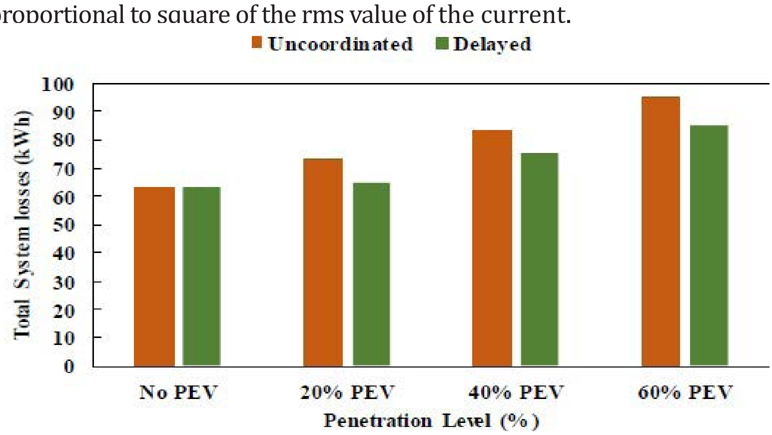


Fig. 3.9. Heat losses in case of uncoordinated and coordinated battery charging (source: Ramadan et al.).

3.3. Impact of PEVs on the distribution grid: Paper 2

3.3.1. Main philosophy

"Probabilistic Analysis of Plug-In Electric Vehicles Impact on Electrical Grid Through Homes and Parking Lots" by S. Rezaee et al. considers the impact of plug-in electric vehicles on the grid. When considering these PEVs, the paper makes a distinction between PHEVs (plug-in hybrid electric vehicles) and BEVs (battery electric vehicles).

These plug-in electric vehicles account for a stochastic behaviour, since the instant of time when they are plugged in has a stochastic nature. The batteries of the vehicles are also charged at different locations which are stochastic in nature. This charging occurs not only in the garages of private households but also at parking lots. At these parking lots, a bidirectional power interface allows to charge the batteries and store energy. This stored energy can be used to drive the vehicle (which is the primary goal), but by partially discharging the battery and injecting power into the grid, dispersed generation is obtained. This means not only parking-to-vehicle P2V but also vehicle-to-parking V2P power exchanges occur. At the private households, only charging of the batteries is considered.

In case of a parking lot, parking-to-vehicle and vehicle-to-parking power exchanges are not equal to grid-to-vehicle and vehicle-to-grid power exchanges (the public grid outside the parking lot is considered). Indeed, in a single parking lot a large number of vehicles are connected with the local grid inside the parking lot. In case the battery of the first car is discharging and this power is immediately sent to the second car, then vehicle-to-parking and parking-to-vehicle power exchanges occur. But neither vehicle-to-grid nor grid-to-vehicle power exchanges occur.

In the nearby future, merely PHEVs will be used to obtain ranges which are comparable with the ranges of traditional cars having internal combustion engines (fueled by e.g. diesel or gasoline). A real breakthrough of BEVs (without combustion engine) will appear when the technology is sufficiently mature. Efforts to develop and sell BEVs already exist (e.g. with Tesla or Nissan as manufacturers) but their impact is still quite limited.

The future rise of PHEVs and BEVs will have a considerable impact on the electrical power grid. Especially the use of fast charging modes (small recharging times from 0.50.5 to 33 hours) requires quite large powers. Although using uncoordinated/uncontrolled/dumb charging of the batteries is the most likely scenario in the near future, it is non sustainable when the number of PHEVs and

BEVs increases considerably. A development of scenarios realising coordinated/optimal/smart charging is necessary. Realising scenarios where a number of vehicles also (partially) discharge their batteries and inject their energy into the grid is an important option.

3.3.2. Parking lot infrastructure

Figure 3.10 provides an overview of the parking lot infrastructure. The parking lot is connected with the public three phase AC grid. A DC-AC electronic power interface (EPI) is an interface between the three phase AC grid and a DC bus with a sufficiently high voltage level. The voltage level in the DC bus is higher than the voltage level available at a private household, since larger powers are needed and the current level must be limited. Larger powers are needed, since a large number of cars will be parked and fast charging must also be possible.

Notice the presence of DC-DC electronic power interfaces providing the appropriate DC voltage level to the vehicles. A number of vehicles are discharging their batteries and deliver power to the parking lot (V2P = vehicle-to-parking-lot). Depending on the SOC of the batteries, the upcoming travel length and also the drivers' willingness a vehicle will or will not provide power to the parking lot. Other vehicles are charging their batteries (P2V = parking- lot-to-vehicle).

Possibly, the infrastructure is equipped with stationary energy storage devices (batteries) and local renewable electrical generating units (e.g. PV panels).

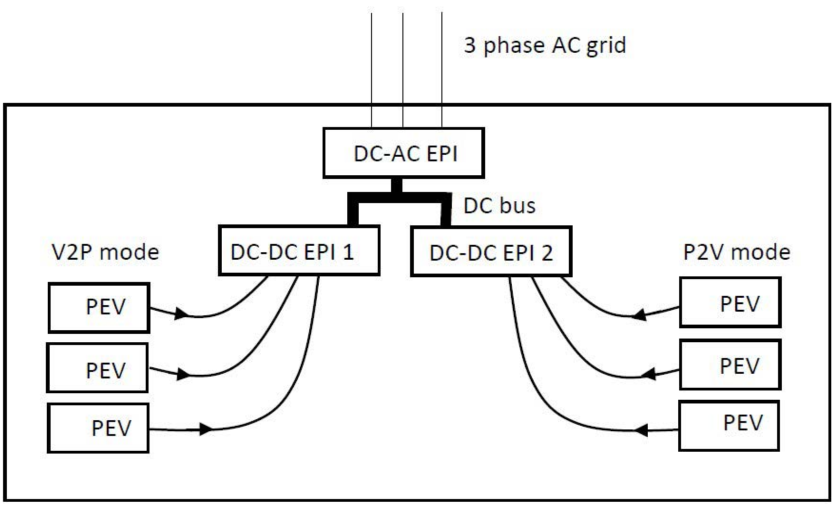


Fig. 3.10. Parking lot infrastructure.

3.3.3. Practical use of a PEV

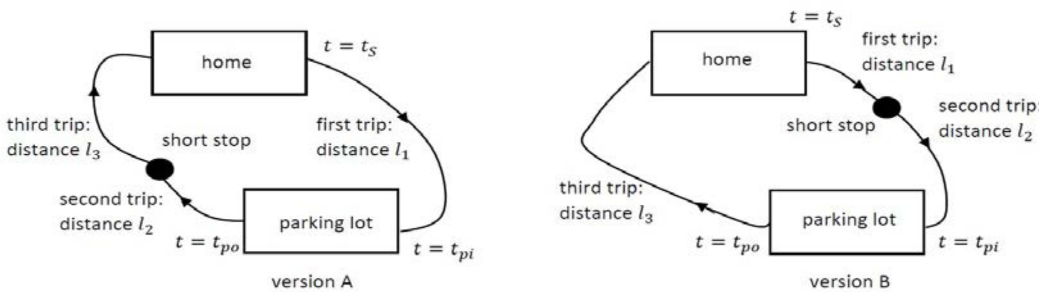


Fig. 3.11. Two versions of a typical PEV daily travel.

The behaviour of each individual PEV, used in an urban environment, is different, which complicates the situation. It is important to have insight in the global behaviour of the vehicles in order to estimate the impact on the electrical grid. For simplicity, the assumption has been made that PEVs have the same behaviour in each 24-hour period. Based on statistics available on http://nhts.ornl.gov, a lot of information is available. The paper by Rezaee et al. assumes that a daily PEV travel contains three trips, as visualised in Fig. 3.11. During the first trip distance l_1 is bridged, during the second trip distance l_2 is bridged and during the third trip distance l_3 is bridged. At the end of the third trip, the vehicle returns home.

As visualised in Fig. 3.11, two versions of the travel map exist. In version A, the vehicle starts at home at $t = t_S$ and has the first stop at a parking lot (at $t = t_{pi}$ the PEV plugs into the electronic power interface). At the parking lot, there is time to recharge the batteries (or partially discharge them). At $t = t_{po}$, the PEV unplugs from the electronic power interface. The second trip starts when leaving the parking lot. Having finished the second trip, it has a short stop (without charging or discharging the battery). Finally, the driver drives home bridging distance l_3. Some homes have a battery charger, other homes have no battery charger.

Version B is very similar, but the short stop takes place after the first trip. Only after the second trip, the parking lot is reached where batteries can be recharged or partially discharged.

Based on statistics available on http://nhts.ornl.gov, a number of averaged parameters are obtained.

- On average, a driver makes three trips a day (which corresponds with Fig. 3.11).

- On average, a driver realises a total daily travel distance of 29 miles (approximately 47 km) (with an average speed of 32 miles per hour, i.e., approximately 51.5 km per hour).

- On average, the short stop equals 10 minutes.

3.3.4. Electrical characteristics of PEVs, PHEVs and BEVs

The all-electric range (AER) is an important parameter of a PHEV. The AER gives the distance (expressed in $mile$ or km) which can be travelled by only using the energy of an entirely charged battery (i.e. without using the combustion engine). The AER mainly depends on two parameters: the nominal energy storage capacity C_N of the battery (expressed in kWh) and the average energy consumption r when driving the vehicle (expressed in $kWh/mile$ or kWh/km). More precisely

$$AER = \frac{C_N}{r}.$$

When considering a PHEV, several approaches are possible when combining the electrical traction motor and the internal combustion engine. A possible (in principle easy approach) is using purely electrical traction from the beginning until the battery capacity reaches its lower limit. Then the combustion engine is used guaranteeing a total range which is larger than the AER.

When considering a BEV, $AER = \frac{C_N}{r}$.

3.3.4.1. Battery charger at home

When considering the driving behaviour sketched in Fig. 3.11 and there is a battery charger installed at home, the battery of the PEV can be charged during the night. Normally, there is enough time for charging the battery, which implies the daily travel starts at $t = t_S$ with a fully charged battery ($SOC = 100$ %). With C_S the battery charge at the beginning of the daily travel, $C_S = C_N$.

When reaching the parking lot, from $t = t_{pi}$ till $t = t_{po}$ the battery can be charged (P2V mode) or partially discharged (V2P mode). The choice between P2V or V2P mode depends on the DOD (depth-of-discharge) of the battery when arriving at the parking lot and the expected distance which will be travelled between leaving the parking lot and arriving back home.

In case Δt_P is the vehicle parking duration, it is clear that $\Delta t_P \geq t_{po} - t_{pi}$. In case of P2V mode, sufficient time is needed to charge the battery until it has a sufficiently high SOC, i.e., time interval Δt_R is needed.

3.3.4.2. No battery charger at home

When considering the driving behaviour sketched in Fig. 3.11 and there is no battery charger installed at home, the battery of the PEV (PHEV or BEV) cannot be charged during the night. This implies the daily travel starts at $t = t_S$ without a fully charged battery ($SOC < 100$ %). The state of charge at $t = t_S$ depends on the driving behaviour the day before. With C_S the battery charge at the beginning of the daily travel, $C_S < C_N$.

In such a situation, especially a BEV absolutely needs a parking lot where the battery can be charged. In this case, P2V mode is needed to charge the battery up to a SOC of 100 %.

3.3.4.3. Four types of PEV travel profiles

S. Rezaee et al. consider four types of PEV travel profiles. Table 1 summarizes the main parameters for these four types of PEV travel profiles.

PEV profile Type 1: Daily travel version A in combination with a battery charger at home.

- Due to the battery charger at home, the battery initial charge $C_S = C_N$, i.e., one starts with a fully charged battery.
- l_b is the daily length of travel before entering the parking lot and with travel version A: $l_b = l_1$.
- l_a is the daily length of travel after entering the parking lot and with travel version A: $l_a = l_2 + l_3$.
- Due to the presence of a battery charger at home V2P and P2V modes are both possible at the parking lot in case of a BEV (having a large battery storage capacity in comparison with a PHEV).

PEV profile Type 2: Daily travel version B in combination with a battery charger at home.

- All remarks of PEV profile Type 1 are still valid except for l_b and l_a. More precisely, $l_b = l_1 + l_2$ and $l_a = l_3$.

PEV profile Type 3: Daily travel version A without a battery charger at home.

- Due to the absence of a battery charger at home, the battery initial charge $C_S < C_N$, i.e., one does not start with a fully charged battery. In case of a BEV (having a large battery capacity), a stochastic analysis based on real life data reveals that the average C_S-value equals $C_S = 0.88\ C_N$. In case of a PHEV (having a smaller battery capacity), a stochastic analysis based on real life data reveals that the average C_S-value equals $C_S = 0.36\ C_N$.

- l_b is the daily length of travel before entering the parking lot and with travel version A: $l_b = l_1$. l_a is the daily length of travel after entering the parking lot and with travel version A: $l_a = l_2 + l_3$.

- Since there is no battery charger at home, no discharging of the battery is allowed at the parking lot (no V2P mode is allowed). Only P2V mode is used, while parked in the parking lot, the battery of the vehicle needs to be charged during Δt_R.

PEV profile Type 4: Daily travel version B without a battery charger at home.

- All remarks of PEV profile Type 3 are still valid except for l_b and l_a. More precisely, $l_b = l_1 + l_2$ and $l_a = l_3$.

Associated with these travel profiles, there is an evolution in the SOC of the

Table 3.1. Overview of PEV Travel Profile Types.

PEV profile type	Battery initial charge $\frac{C_S}{C_N}$ (%)		l_b	l_a	Δt_{eff}	
	BEV	PHEV			BEV	PHEV
1	100	100	l_1	$l_2 + l_3$	Δt_p for V2P Δt_R for P2V	Δt_p
2	100	100	$l_1 + l_2$	l_3	Δt_p for V2P Δt_R for P2V	Δt_p
3	88	36	l_1	$l_2 + l_3$	Δt_R	Δt_R
4	88	36	$l_1 + l_2$	l_3	Δt_R	Δt_R

batteries. As already mentioned, the daily travel starts with an initial battery charge C_S. When connecting with the parking lot (having already travelled over distance l_b) at $t = t_{pi}$, the battery charge equals C_P. During Δt_{eff} power is exchanged between the vehicle and the parking lot using the electronic power interface EPI (charging or discharging the battery). After disconnecting from the EPI, the driver finishes the daily travel profile and when entering home, the battery has battery charge C_L.

When a battery charger is available at home, at $t = t_{hi}$ the PEV is plugged in and the charging procedure starts. In general, at home, slow charging is used. At $t = t_{ho}$, the PEV is unplugged and the home charging of the battery ends, in general, a $SOC = 100$ % is obtained.

3.3.4.4. Mathematical model of power transactions

Power transactions can be modelled by mathematical expressions.

The daily travel starts with an initial battery charge C_S. To reach the parking lot, the vehicle crosses distance l_b. With r being the average energy consumption of the car (expressed in $kWh/mile$ or kWh/km), the remaining battery capacity when entering the parking lot is given by

$$C_P = C_S - l_b.r. \qquad (3.2)$$

When considering the required energy C_R available in the battery when leaving the parking lot, a distinction is needed between the situation where a home charger is available and the situation where no home charger is available. First, consider the situation where a home charger is available. The required battery charge C_R equals

$$C_R = l_a.r + 0.2.C_N. \qquad (3.3)$$

After leaving the parking lot, the vehicle will cross distance l_a requiring an amount of energy which equals $l_a.r$. Notice that an additional amount of energy $0.2.C_N$ will be stored to compensate unforeseen situations (e.g. a larger travelling distance). In case $C_R > C_P$, charging of the battery is needed, i.e., P2V mode is used (during Δt_R). In case $C_R > C_P$, discharging of the battery is possible, i.e., V2P mode can be used (during Δt_p).

Consider also the case when no home charger is available. While parked in the parking lot, the battery of the car will always be charged (always P2V mode) and

$$C_R = C_N \tag{3.4}$$

value is preferred. The battery is fully loaded when leaving the parking lot.

By defining ΔE_P as the energy exchanged with the parking lot (expressed in kWh), a distinction can be made between the V2P mode and P2V mode. In case of V2P mode, $C_R > C_P$ and

$$\Delta E_P = (C_P - C_R).\eta_P > 0. \tag{3.5}$$

With η_P being the power conversion efficiency at the parking lot, ΔE_P is the net energy injected into the DC bus of the parking lot. Due to losses during the conversion, $\Delta E_P < C_P - C_R$. When assuming that the battery discharging occurs with power P_P (measured at the parking lot side), the required discharging time

$$\frac{|\Delta E_P|}{P_P}. \tag{3.6}$$

In case of P2V mode, $C_R > C_P$ and

$$\Delta E_P = \frac{C_P - C_R}{\eta_P} < 0. \tag{3.7}$$

Due to losses during the conversion, $|\Delta E_P| > |C_P - C_R| = C_R - C_P$. When assuming that the battery charging occurs with power P_P (measured at the parking lot side), the required loading time

$$\Delta t_R = \frac{|\Delta E_P|}{P_P}. \tag{3.8}$$

Since Δt_{eff} equals the actual parking power transaction duration, it is clear that $t_{po} = t_{pi} + \Delta t_{eff}$ (which refers to the parking lot). In case charging at home is possible, mathematical expressions allow to model the energy transactions at home. More precisely,

$$\Delta E_H = -\frac{0.8C_N}{\eta_H} < 0 \tag{3.9}$$

represents the energy exchanged at home (measured at the grid side of the home) while charging the battery. Indeed, after leaving the parking lot the battery charge $C_R = l_a r + 0.2C_N$. By traveling home, an amount of energy $l_a r$ is consumed leaving a battery charge which equals $0.2C_N$. To obtain a fully charged battery the next day, an additional battery charge which equals $0.8C_N$ is needed.

Losses of the home charger are taken into consideration by using the power conversion efficiency η_H. Similar with the convention used at the parking lot, ΔE_H is negative, since power/energy is extracted from the grid to load the battery.

Since P_H is the power charging rate for the battery charger at home, charging time $|\Delta E_H|/P_H$ is needed. This implies that

$$t_{ho} = t_{hi} + \frac{|\Delta E_H|}{P_H}. \tag{3.10}$$

Notice that t_{hi} is the time when the PEV plugs into the battery charger and t_{ho} is the time when the PEV plugs out of the battery charger. In order to have a fully charged battery, t_{ho} must be an earlier time stamp than $t = t_S$.

Notice that ΔE_P and ΔE_H are both related to the energy exchange (and power exchange) with the electrical grid. Indeed, the study of the behaviour of PEV, BEV and PHEV is performed in order to estimate the impact on the electrical power grid.

3.3.5. Stochastic modelling

When considering the two versions of the travel maps visualised in Fig. 3.11, stochastic modelling is required.

- The daily travel starting time is different for different vehicles. In an urban context, a Rayleigh probability density function (pdf) describes the probability of a particular t_S (we use the notation $P_S(t_S)$). To give an idea, especially between 6 am and 7 am, a lot of cars are leaving their garage.

- When considering the total daily travel path $l = l_1 + l_2 + l_3$, Fig. 3.12 visualises (using a bar graph) the percentage of vehicles driving such a distance leading to a probability density function $P_T(l)$. It is also possible

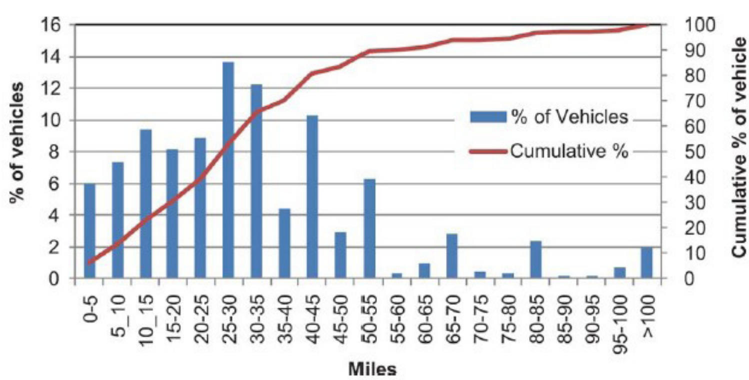


Fig. 3.12. Percentage of vehicles driving a number of miles per day (source: Darabi et al.).

to consider distances l_1, l_2 and l_3 separately. Figure 3.13 visualises the probability density functions of l_1 and l_2.

The horizonal axis of Fig. 3.13 on the left actually visualises l_1/l. The length of the first trip l_1 ranges from 0 % to 100 % of l giving a normal distribution $P_{T1}(l_1)$. The horizontal axis of Fig. 3.13 on the right visualises $l_2/(l - l_1)$ which also ranges from 0 % to 100 %. The probability density function $P_{T2}(l_2)$ is obtained. Once l, l_1, l_2 are determined, also $l_3 = l - l_1 - l_2$ is determined.

When considering the two versions of the travel maps visualised in Fig. 3.11, stochastic modelling is required. We already considered t_S, l, l_1, l_2 and l_3. Also the parking stop duration Δt_P is important, since sufficient time is needed to load the battery. In "Profile of Charging Load on the Grid Due to Plug-in Vehicles" by S. Shahidinejad et al., a distinction is made between 'short' parking periods (less than about half an hour), 'medium' parking periods (less than about three hours) and 'long' parking periods (longer than three hours). Based on this consideration, a probability density function $P_P(\Delta t_P)$ can be obtained.

In total, seven parameters are needed to describe a PEV daily travel profile:

- the initial charge of the battery of PEV – C_S;
- travel version A versus travel version B (see Fig. 3.11);
- the duration of parking stop – Δt_P;
- the time of the beginning of the daily travel – t_S;
- the total length of the daily travel – l,
- the length of the first trip – l_1,

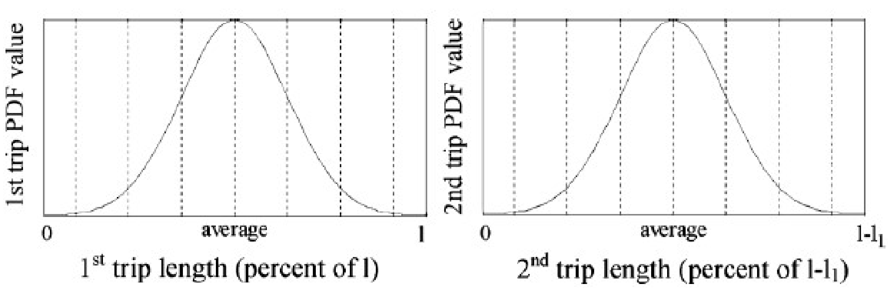


Fig. 3.13. Probability density functions of daily travel parts (source: Rezaee et al.).

- the length of the second trip – l_2.

For Δt_P, t_S, l, l_1 and l_2 the probability density functions have been described. This allows to obtain probability $P(i)$ for a specific stochastic case (abbreviated as SC) with number *i*. More precisely,

$$P(i) = 0.5(P_S(t_S)P_P(\Delta t_P)P_T(l)P_{T1}(l_1)P_{T2}(l_2)). \tag{3.11}$$

In order to be able to take the product of the probabilities, the correlation between the random variables must be sufficiently small allowing to assume there is no correlation at all. Coefficient 0.5 indicates the distinction between the travel version A and travel version B (assuming half of the vehicles realise travel version A and the other half realises travel version B).

In case N, stochastic cases are considered implying i ranging from 1 to N,

$$\sum_{i=1}^{N} P(i) = 1. \tag{3.12}$$

3.3.6. Power transfer function

The power transfer function PTF describes the daily power transaction of a PEV during 24 hours. A typical PTF is visualised in Fig. 3.14.

Actually, power transfers with the electrical grid are considered. Between t_{hi} and t_{ho}, the battery is charged at home. The charging power is fixed and equals P_H which is always negative (power is extracted from the grid to load the battery). Notice also the total energy transfer ΔE_H. Between t_{pi} and t_{po}, power exchange occurs at the parking lot. In the case visualised in Fig. 3.14, P_P is positive, i.e., discharging of the battery occurs (but also a negative P_P is possible in case of charging the battery). Notice also the total energy transfer ΔE_P. Outside these two time intervals, no battery charging or discharging occurs.

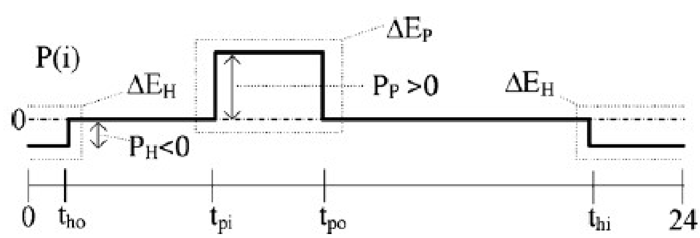


Fig. 3.14. Power transfer function of stochastic case SCi (source: Rezaee et al.).

The power transfer function $PTF(i)$ corresponds with stochastic case SCi and has probability $P(i)$. By adding the impact of all $PTF(i)$, taking the probabilities into consideration, the so-called average daily power transaction variation (ADPTV) is obtained. More precisely,

$$ADPTV = \sum_{i=1}^{N} P(i) \cdot PTF(i). \tag{3.13}$$

In case the parking lot has no distributed generation units (e.g. photovoltaic panels) and there are also no separate energy storage devices at the parking lot, then the power evolution obtained as ADPTV needs to be added to the existing average daily load profile. Actually, the number of PEVs N_V must be taken into consideration. The additional power evolution equals $ADPTV \cdot N_V$.

3.3.7. Calculation of the ADPTV for four basic cases

When considering the ADPTVs, four basic cases can be considered:

- Case 1: BEV is used with a battery charger at home.
- Case 2: PHEV is used with a battery charger at home.
- Case 3: BEV is used without a battery charger at home.
- Case 4: PHEV is used without a battery charger at home.

In the paper by Rezaee et al. computer code has been used to generate $480 \cdot 10^3$ stochastic cases (SC) for BEVs and PHEVs separately. At home, slow charging occurs (starting from a SOC of 20 %, e.g., 4 hours are needed to fully charge the battery). At the parking lot, fast charging or discharging occurs (e.g., 1 hour is needed to fully charge the battery). By adding all ADPTVs, an estimation of the impact on the grid is obtained. A distinction has been made between the power profile exchanged at the parking lot (Fig. 3.15) and the power profile exchanged at home (Fig. 3.16).

Figure 3.15 makes a distinction between the four basic cases. When having a BEV with a battery charger at home (Case 1), the vehicle starts at home with a fully charged battery. The capacity of the battery is rather large, which implies the battery can be partially discharged at the parking lot (giving a positive ADPTV). During the night (at home), energy is extracted from the grid (when the general power consumption is low) and stored in the battery. During the day (at the parking lot), this stored energy is partially extracted from the battery, which is useful to charge other batteries (or the energy can be injected into the grid

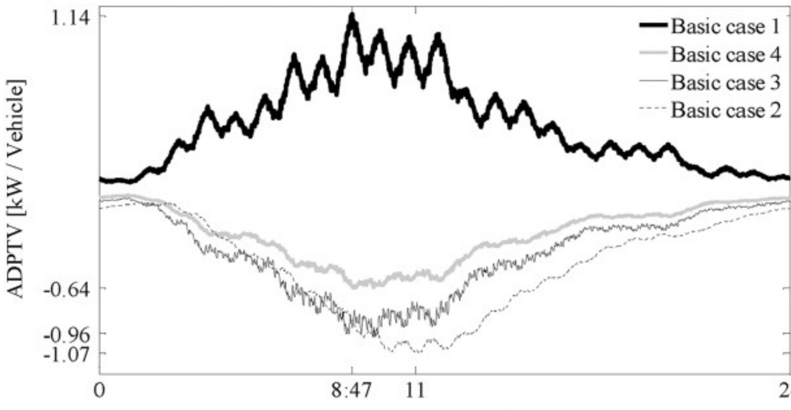


Fig. 3.15. ADPTV of a PEV at a parking lot (source: Rezaee et al.).

when the general power consumption is high).

When considering Cases 3 and 4, no battery charging occurs during the night at home. This implies the batteries must be charged at the parking lot during the day. This implies negative ADPTVs. Since the battery of a BEV is larger than the battery of a PHEV, the values of the ADPTV are more negative for Case 3.

Figure 3.16 only considers Basic case 1 and Basic case 2, since only these two cases consider a battery charger at home. Since the battery of a BEV is larger than the battery of a PHEV, the values of the ADPTV are more negative for Case 1 which considers a BEV.

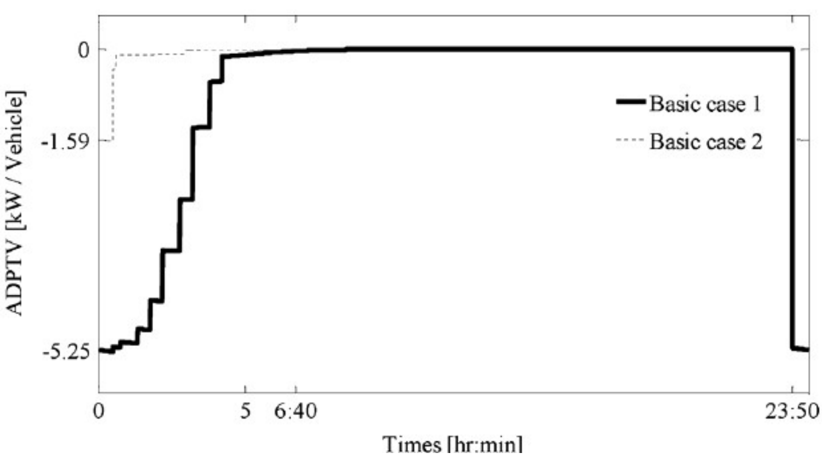


Fig. 3.16. ADPTV of a PEV at home (source: Rezaee et al.). ***3.3.8.***

Impact on the grid

Finally, the impact on the electrical grid will be considered. In the paper by Rezaee et al., the electrical grid visualised in Fig. 3.17 has been considered (see also http://sites.ieee.org/pes-dsacom/ and http://sites.ieee.org/pes-testfeeders/resources/). The grid has a voltage level of 24.9 kV and contains 34 buses. Three parking lots are considered and they are connected with the grid at nodes 5, 15 and 28 (which is a choice).

The base load of the grid (without the impact of the PEVs) is given in Fig. 3.18. Notice the limited power consumption during the night. A peak consumption occurs around 11 am, and the largest peak consumption occurs in the evening between 8 pm and 9 pm. Somewhat less power is consumed during the afternoon.

The impact of PEVs must be added to the base load in Fig. 3.18. It is necessary not only to have knowledge about the shape of the additonal load but also important to know the penetration rate of PEVs. A realistic estimate of the penetration rate can be obtained by studying the behaviour of the citizens.

In the paper by Rezaee et al., the situation in the United States is considered with an average electrical power demand of 1.5kW/year for a private household (which corresponds to approximately 13.1 MWh/year). On average, a household in the United States has 1.87 vehicles (situation in 2009).

When considering Fig. 3.18, with an average power consumption of 830 kW, actually 553 households are considered. This implies the use of 1034 verhicles.

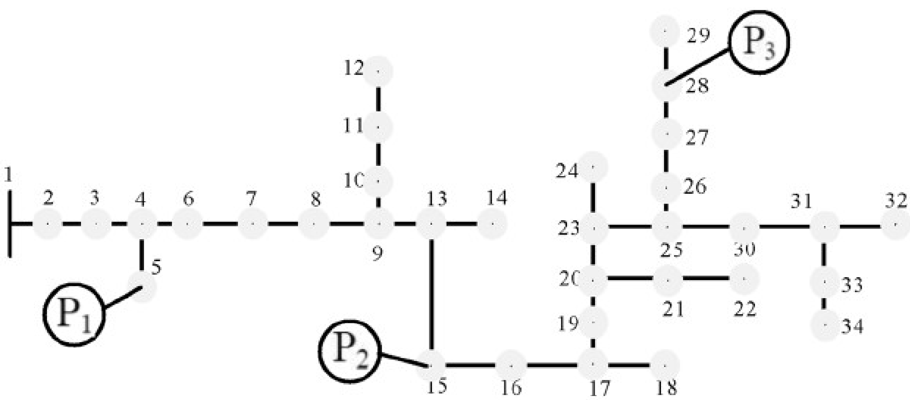


Fig. 3.17. Electrical test grid (source: Rezaee et al.).

In case a household possesses a maximum of one PEV, this corresponds to a penetration rate of PEV of almost 50 %. The impact on the electrical grid will be studied in case of different penetration rates, i.e., 0 % (no PEVs are used), 8 %, 22 % and, finally, 50 %.

A distinction can be made between four scenarios. For instance, in Scenario 1, 25 % of the vehicles are BEVs and the owners have a battery charger at home (Basic case 1); 25 % of the vehicles are PHEVs and the owners have a battery charger at home (Basic case 2); 25 % of the vehicles are BEVs, but the owners have no battery charger at home (Basic case 3); 25 % of the vehicles are PHEVs , but the owners have no battery charger at home (Basic case 4). In a similar way, the other scenarios are considered in Table 2.

3.3.8.1. Voltage drops

Depending on the scenario, different power profiles are needed giving other current levels in the grid and implying other voltage drops. Figure 3.19 visualises the voltage profile (at bus 29) in case of Scenario 1. In case of 0 % penetration, only the base load of Fig. 3.18 accounts for voltage drops. Notice that the two power peaks account for higher voltage drops implying a lower remaining voltage level. As the penetration rate increases (more BEV and PHEV are used needing power from the grid), more additional power is required from the grid. Higher current levels account for larger voltage drops implying a lower remaining voltage level.

When considering Scenario 1, voltage drops occur mainly during the night. About 50 % of the vehicles are charged during the night, implying additional power consumption. During the day, the voltage drops are lower, since part of the cars are being discharged (V2P mode) while parked in the parking lot. This implies that the

Table 3.2. PEV Penetration Rate Distribution Scenarios

Distribution scenario	Basic case 1	Basic case 2	Basic case 3	Basic case 4
Scenario 1	25 %	25 %	25 %	25 %
Scenario 2	0 %	0 %	50 %	50 %
Scenario 3	11 %	22 %	22 %	44 %
Scenario 4	22 %	44 %	11 %	22 %

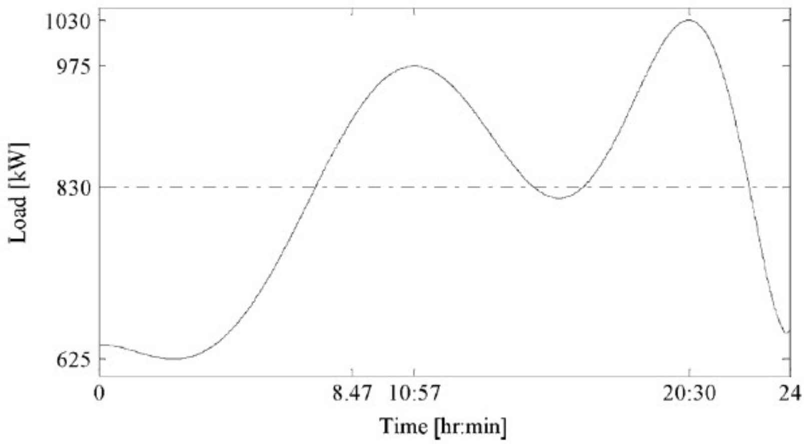


Fig. 3.18. Base load of the grid without the impact of the PEVs (source: Rezaee et al.).

power needed to charge the batteries in the parking lots (partially) comes from other discharging batteries. The additional load for the grid is limited.

Figure 3.20 visualises the voltage profile at the same bus 29 in the case of Scenario 2. There are differences when comparing the results in Figs. 3.19 and 3.20. Neither Basic case 1 nor Basic case 2 are used, which implies that no home charging occurs during the night (there is no home charging at all). The cars are all charged at parking lots (quite often during the day), which implies higher voltage drops during the day. The resulting bus voltage is lower.

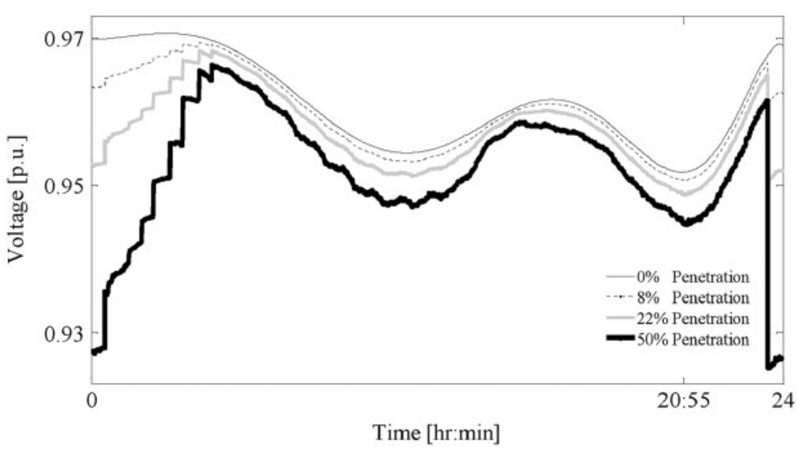


Fig. 3.19. Voltage profile at bus 29 in case of Scenario 1 (source: Rezaee et al.).

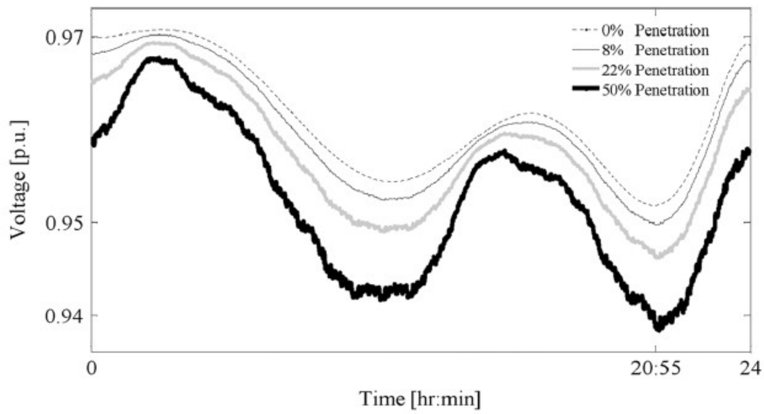


Fig. 3.20. Voltage profile at bus 29 in case of Scenario 2 (source: Rezaee et al.).

Figure 3.21 visualises the voltage profile at the same bus 29 in the case of Scenario 3. The results of Scenario 3 are somewhat similar to the results of Scenario 2. Indeed, when considering Scenario 3, about 66 % of the vehicles have no battery charging opportunities at home (a lot of charging occurs in parking lots, mainly during the day). Notice, however, that charging of batteries occurs also during the night, implying that voltage drops during the night.

Figure 3.22 visualises the voltage profile at the same bus 29 in the case of Scenario 4. The results of Scenario 4 are somewhat similar to the results of Scenario 1. Indeed, when considering Scenario 4, about 66 % of the vehicles have battery charging opportunities at home. Quite a lot of battery charging occurs during the night at home, implying a quite large voltage drop. During the day, when parked at a parking lot, a lot of vehicles operate in the V2P mode which

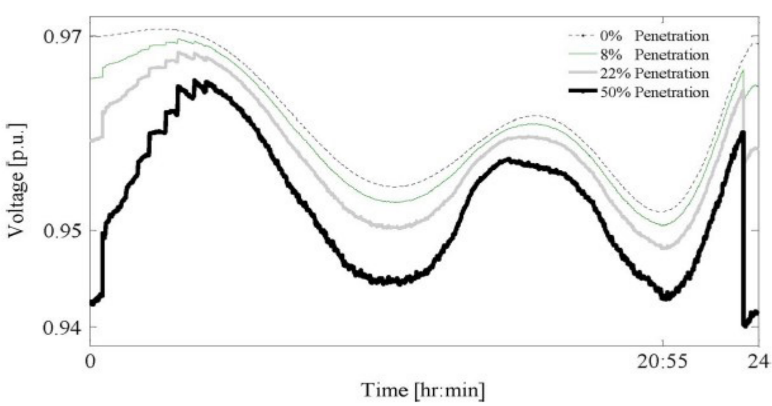


Fig. 3.21. Voltage profile at bus 29 in case of Scenario 3 (source: Rezaee et al.).

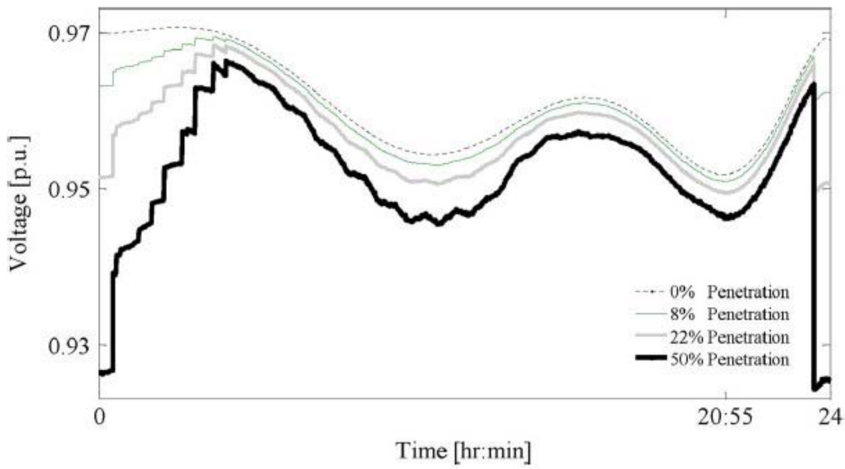


Fig. 3.22. Voltage profile at bus 29 in case of Scenario 4 (source: Rezaee et al.).

provides local power to charge other vehicles without having too much impact on the public electrical grid.

3.3.8.2. Grid power losses

The grid currents are determined by the power base load and the additional load due to the PEVs. These grid currents not only account for voltage drops but also for power losses. Indeed, Joule losses arise, since the currents are flowing in the resistive parts of the grid impedances. Table 4 gives an overview of the grid power losses (rated in per unit or %), where

- a distinction has been made between the four scenarios,
- a distinction has been made depending on the penetration level.

As the penetration rate of the PEVs increases, the losses increase in all scenarios. Actually the differences between the considered scenarios are rather limited.

Although a lot of papers study the (possible) impact of PEVs, BEVs, and PHEVs on the electrical grid, the two papers discussed above give a decent idea of the challenges faced by the grid operators. The introduction of a large amount of PEVs, BEVs and HPEVs generally increase the consumed electrical power implying the following:

- an increased load for all grid components (e.g. transformers and feeders);

Table 3.4. Overview of the Grid Power Losses.

Penetration Level	Scenario 1	Scenario 2	Scenario 3	Scenario 4
0 %	2.458	2.458	2.458	2.458
8 %	2.637	2.651	2.639	2.641
22 %	2.972	2.998	2.957	2.986
50 %	3.876	3.821	3.764	3.874

- an increased voltage drop across the grid impedances;
- increased Joule losses in, e.g., the feeders of the grid.

The impact of this increased number of electrical loads (car batteries) can be reduced and/or handled by introducing serveral ways of intelligence. Coordinated loading of batteries can simply include a planned delay of the loading process, i.e., loading batteries during the night when the base load of the grid is rather small. Introducing electronic power interfaces which are able to charge and discharge the batteries is also useful. During the night the base load of the grid is rather small and a number of batteries are charged which are partially discharged (in a smart way) during the day when the base load of the grid is larger. By discharging a number of batteries during the day, it is possible to charge other batteries which really need the charging process (and are not able to postpone the charging process) without causing an overload of the electrical grid.

Bibliography

Z. Darabi, M. Ferdowski, Aggregated Impact of Plug-in Hybrid Electric Vehicles on Electricity Demand Profile, IEEE Transactions on Sustainable Energy, vol. 2, no. 4, October 2011, pp. 501–508.

H. Ramadan, A. Ali, C. Farkas, Assessment of Plug-in Electric Vehicles Charging Impacts on Residential Low Voltage Distribution Grid in Hungary, 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), 2018, pp. 105–109.

S. Rezaee, E. Farjah, B. Khorramdel, Probabilistic Analysis of Plug-In Electric Vehicles Impact on Electrical Grid Through Homes and Parking Lots, IEEE Transactions on Sustainable Energy, vol. 4, no. 4, October 2013, pp. 1024–1033.

S. Shahidinejad, S. Filizadeh, E. Bibeau, Profile of Charging Load on the Grid Due to Plug-in Vehicles, IEEE Transactions on Smart Grid, vol. 3, no. 1, March 2012, pp. 135–141.

Chapter 4:

Energy-saving technologies in transport

Andrii Hnatov,
Shchasiana Arhun
Yuriy Borodenko
Oleksandr Dziubenko
Nataliia Rudenko
Kateryna Danylenko

Kharkiv National Automobile and Highway University

4.1. Approaches for automobile transport development

The key tasks of improving automobile transport are to make better its fuel, economic and environmental parameters. Fuel, in a broad sense, is defined as energy carriers bringing the powerplant of a car into action.

Potential energy carriers are disposed in the Earth's interior and on its surface, in the atmospheric air, and even in space. Moreover, some of them exist in various substances and properties. For example, hydrogen in various chemical compounds is found both underground (hydrocarbons), on earth (biomass) and in the atmosphere.

The effectiveness of using an energy product, on the one hand, is estimated by its energy intensity, on the other hand, by the costs to obtain it (transportation) and the quality (efficiency) of transforming into a consumer form of energy. In this case special emphasis is on renewable and unlimited types of energy environment. It should be understood, however, that any product obtained as a result of the phase transition process of a substance will ultimately have a lower energy density than a product obtained as a result of the process based on chemical reactions. In turn, the products obtained as a result of chemical reactions have an energy density lower than that of substances undergoing nuclear transformations.

Leading car manufacturers are researching and developing various types of concept cars. As this takes place, there are three approaches to develop powerplants operating on alternative energy sources (Fig. 4.1).

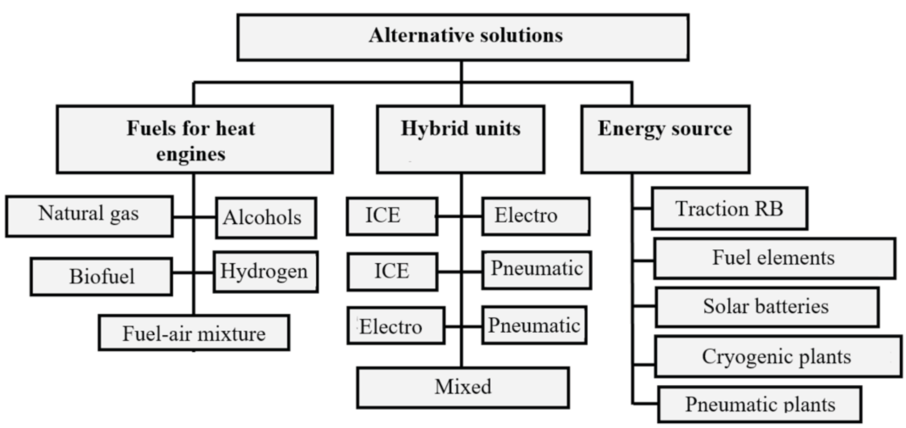


Fig. 4.1. Approaches for automobile powerplant development.

The first approach is to adapt (convert) the design of heat internal-combustion engines (ICE) for alternative types of liquid or gaseous fuels. The advantage of this approach is the minimal engine conversion and fuel production costs. The effectiveness of using alternative fuels, in this case, is evaluated by several parameters, distinguished by quantitative performance characteristics:

- environmental safety (density of harmful substances);
- production activity costs (supplies, mining, processing, preparation);
- fuel power capacity (calorific value);
- quality of use (engine efficiency);
- engine converting and additional equipment costs;
- infrastructure facility costs.

The second approach involves the use of alternative motors powered by unconventional energy sources. Examples of such concepts are the latest developments in motor vehicles with electrochemical galvanic (see Section 4.2), pneumatic (see Section 7.5), solar, nitrogen, hydrogen and even nuclear energy sources. The limitation of this approach can be considered as significant costs for developing powerplants and energy sources. It is specified that the indicators of specific energy and production costs of most alternative energy source materials are significantly inferior to traditional gasoline and diesel fuel. So, for example, to obtain the energy generated by 3.7 liters of fuel material, it is necessary to use 400 kilogram of batteries (fuel calorific value of 0.085 liters is equivalent to the electrical energy consumption of 1 kWh).

A compromise option is a combination of these approaches in order to build hybrid vehicles. The characteristics of hybrid powerplants (HPP) are the following: main and auxiliary motors, energy source (energy storage reservoir), energy converters, recuperation system at braking. The electronic power distribution system controls the HPP with standard powerplant parameters (see Section 4.3).

It should be emphasized that the use of ICE and an electric drive in the HPP configuration cannot be considered the only viable option. If certain vehicle requirements are in agreement with the operating conditions, it may be quite appropriate, for example, to use a pneumatic drive, where the pneumatic unit is the main or auxiliary motor (see Section 7.6).

Let us dwell in more detail on each aspect of the alternative options and characterize the innovations from the perspective of economic efficiency and environmental safety.

For cars with heat motors, natural hydrocarbon gas ranks 3rd after gasoline and diesel fuel, which is kept in special high pressure tanks in a liquefied or compressed state (tank-stored fuel volume).

Natural gas sources in the Earth's interior are various geological locations. There are traditional (gas deposits, oil-dissolved gas) and unconventional (coal, bitumen, shale) sources of natural gas. The production method and the volume of natural gas deposits largely determine its raw material price [31].

The main part of natural gas is made up of methane (CH_4) from 70 to 98 %. The gas may contain heavier hydrocarbons (methane homologues): ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), pentane (C_5H_{12}).

To fuel a car (after processing raw materials), compressed (compressed under a pressure of 20 MPa) methane and liquefied (under a pressure of 1.5 MPa) propane (propane-butane mixture) are used.

The general strengths of gas fuel are:

- lower price in relation to gasoline;
- environmental safety;
- increase in mileage on one volume of tank-stored fuel;
- less wear-out of converted engine components.

The main weaknesses of gas fuel are:

- relative lack of infrastructure facilities;
- reduction in power and improved wear-out of non-converted engine components.

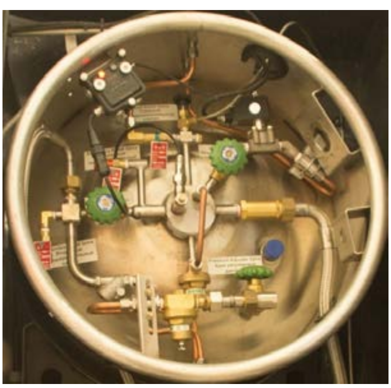
Different densities and pressures of tank-stored methane and propane largely explain the advantages and disadvantages of their use as automotive fuel. Compressed gas equipment (CGE) for propane is 70 % cheaper than for methane, but the fueling cost is higher. In such a case, the weight of high pressure air tanks with propane is several times lighter than the ones with compressed methane, and the reservoir fluid capacity of liquefied propane with the equal volume of the high pressure air tank provides three times more mileage than the same capacity of methane. The loss in power of ICE running on gasoline, respectively on propane, is equal to 5 %, and on methane makes 20 %. The environmental propane parameters are worse than the ones of methane. Methane is less explosive, the equipment for preparing and filling methane is simpler than for the propane, but at the same time, the network of methane filling stations is less developed.

It ought to be noted that currently there is a tendency to use liquefied methane in automobile transport [32]. This technology, on the one hand, makes it possible to boost the fuel energy density, on the other hand, it involves significant energy consumption, since liquefied methane is obtained by cooling the gas up to 160 °C below zero. Moreover, its storage requires the special cryogenic (thermostated) reservoirs at filling stations and high pressure tanks on board a car (Fig. 4.2).

To supply fuel to ICE cylinders, a two-stage reduction in pressure is needed in the evaporator and the speed reducing drive. The use of liquid methane as an ecological motor fuel is justified in the conditions of its operational use at fixed mileage. So, for example, a tank with a capacity of 320 liters (120 kg of methane) provides the power supply for a bus about 400 km of distance run, and a truck tractor with a 700-liter tank is able to cover more than 700 km of the route.

Alcohol-based fuels (ethanol, methanol, butanol) and their mixtures are characterized by high calorific values [31].

Ethanol (ethyl, grain alcohol) is produced by fermentation and distillation of grain products. Ethanol, like other alcohols, can be blended with gasoline



a

b

Fig. 4.2. Cryogenic high pressure tanks for liquid methane: a – MAZ 203945 bus (end); b – IVECO Stralis truck tractor.

to produce fuels with a higher octane number and lower emissions than pure gasoline. If the percentage of ethanol in a mixture with fossil fuels is at least 85 %, then such a mixture is considered to be an alternative fuel. Blends with low ethanol concentrations are considered additive-added fossil fuels.

Bioethanol is ethyl alcohol obtained by processing plant materials (biofuel). Mixed with gasoline or diesel, it can be applied in most vehicles with both converted and non-converted ICEs. From the environmental point of view, the use of bioethanol makes sense, since the carbon dioxide releasing during the combustion in the engine is compensated by the gases absorbed while producing processes. The key undesirable effect of bioethanol is its increased consumption.

Methanol (wood, methyl alcohol) can be used as an alternative fuel in vehicles with a converted ICE (a mixture of methanol (85 %) and gasoline (15 %)). In current times, methanol has limited use in transport. However, in the future, it can be applied as a source of hydrogen for the fuel cell operation.

Ammonia (spirit of ammonia) is a compound of hydrogen and nitrogen. Ammonia has a low energy density (half lower than gasoline) but is cheaper than other alcohols to produce. It is used as a fuel for piston motors and also in fuel cells. Ammonia has no carbon, so it produces zero harmful carbon dioxide emissions. Its application restrictions are related to the safety of ammonia storage and transportation. Motor transport uses mixed fuel, which contains 70 % of ammonia and only 30 % of gasoline.

Biodiesel fuel (fats) is a fuel derived from vegetable oils (rapeseed), animal fats or microalgae lipids (recoverable organic elements). The industrial production of biodiesel is more expensive than the production of diesel from oil. The substance is safe, biodegradable and cuts down the amount of air pollutants (solid impurities, carbon monoxide, hydrocarbons). It is used both in its pure form (for converted ICEs) and as an additive to hydrocarbon diesel fuel (for non-converted ICEs).

Hydrogen fuel (an alternative to hydrocarbon fuel) has a calorific value three times higher than that of gasoline and does not emit harmful substances in the act of combustion. This fuel is a recoverable product, as it is generated from water and, after the combustion, turns into water condensate. However, the hydrogen production is much more expensive than the gasoline production (due to the energy costs). In addition, its storage under pressure is explosive. The use of hydrogen as a monofuel in an unconverted engine reduces its power up to 70 %. In a converted ICE, the power can be boosted up to 117 %, but at the same time, nitrogen oxides are exhausted and make shorter the motor life. A compromise solution would be to mix conventional fuels with hydrogen to improve the flammability of lean mixtures.

Multicomponent fuel-air mixtures (fuel of P series) contain ethanol, gas condensate liquid and an auxiliary solvent obtained from biomass. They have a high octane number and are used in motor vehicles with a universal fuel system in a pure form or mixed with gasoline in any ratio. It should be noted that all types of alternative combustible fuels are non-renewable energy sources.

As for the third approach, we will give a brief description of the powerplants of automobile conceptual prototypes with alternative energy sources.

Electrical energy is used for motor vehicles powered by traction rechargeable batteries (TRB) or run on fuel cells. In the first case, while developing an electric vehicle, a number of specific issues arise related to ensuring the required TRB capacity (weight, dimensions, voltage rate) and the electric drive power (electric motor, transmission); the TRB location on board a car (mass centre); the use of system for cooling energy recovery of electrical units; maintaining the operational powerplant characteristics (acceleration response time, TRB charging time).

The main limiting factor in the development of motor vehicles with autonomous electric traction is the low energy consumption of traction rechargeable batteries (TRB). For example, the best lithium-ion TRBs have a specific energy of up to 0.4 MJ/kg, while for the gasoline ones this indicator makes about 40 MJ/kg. Even taking into account the threefold difference in efficiency and the possibility of braking energy recuperation, it takes about a ton of rechargeable batteries to replace a 40-liter gas tank.

Thus, the drawbacks of using an electric drive include: high cost and limited service life of TRBs (up to 7 years), the need for long-term charging, and a limited resource of autonomous movement. Whilst the operating costs of maintaining an electric vehicle are lower than for a car with ICE [33].

In a *fuel cell*, an electrochemical reaction occurs between hydrogen and oxygen, which is the reverse of electrolytic process with the participation of catalytic agents. As a result of this reaction, an electrical potential difference is formed between the cell electrodes. The reaction is accompanied by the release of heat (without ignition) and water. To generate the electricity in a car, the fuel tank contains hydrogen, and oxygen is supplied from the air. Currently, many leading car manufacturers are developing vehicles with fuel units. So, for example, a hydrogen car designed on the basis of a Mirai (Toyota) front-wheel drive electric car with a 112 kW electric motor is equipped with a fuel unit with a 5 kg supply of hydrogen in high pressure tanks (120 liters at a pressure of 70 MPa), which provides a drive range of up to 400 km (Fig. 4.3).

It takes 3 minutes to fill such a car with hydrogen. The power unit contains

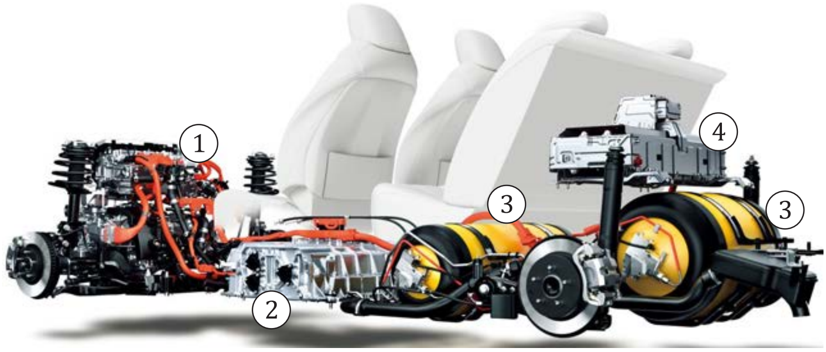


Fig. 4.3. Composition of Toyota Mirai hydrogen car chassis: 1 – electric motor; 2 – block of fuel cells; 3 – hydrogen high pressure tanks; 4 – electric battery.

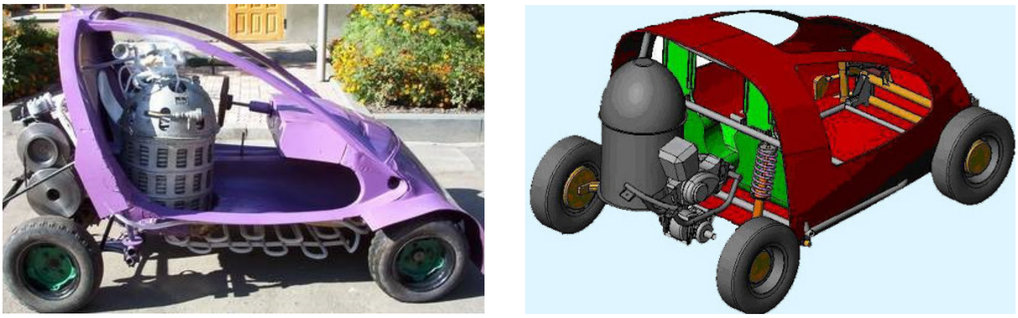
370 DC hydrogen fuel cells and an amplifier-invertor (up to 650 V). The maximum car speed reaches up to 175 km/h. For storing the energy storage in the fuel unit and the regenerative braking, a 21 kWh nickel-metal hydride TRB is applied [34].

The *energy stored in liquid nitrogen* (78 % of air), as well as in the compressed air (atmospheric recoverable resources), can be used to run a car. In these conditions, two methods of converting energy are considered, namely, direct conversion using a compressed air motor and a two-step conversion with the electrical generating unit and an electric motor (series hybrids).

The *cryogenic powerplant* [35] consists of a cryogenic tank (1), a gasifier (2), an air heat exchanger (3) and a compressed air motor (4) (Fig. 4.4).

The pneumatic drive of wheels (Fig. 4.4, a) or a generator (6) (Fig. 4.4 b) is run by the piston (4) or the vane (rotary) compressed air motor (5). A HPP with an electric drive (7) includes a buffer energy tank for putting into action the braking energy recovery system. Cryomobiles with an electric drive are in many parameters comparable to electric plug-in hybrids. In addition to the abovementioned, the following advantages can be highlighted:

- available cheap non-explosive and non-combustible raw materials;
- the best environmental parameters at the disposal and restoration of thermal high pressure tanks as compared to electric rechargeable batteries;
- fast charging versus electric rechargeable batteries;
- use of by-product heat from an ICE (in hybrids with ICEs).



a b

Fig. 4.4. Designs of cryomobile: a – with a pneumatic drive; b – with an electric drive.

The disadvantages of cryogenic powerplants are caused by several factors:

- high energy consumption of producing liquid nitrogen (up to 1 kWh/l) and additional equipment costs (thermostated high pressure tank, gasifier, heat exchanger);
- low energy density of liquid nitrogen in comparison with hydrocarbon fuels;
- negative impact of the by-product of nitrogen production from the atmospheric air (liquid oxygen) on the surrounding organics;
- danger of direct contact with human body (freeze burn);
- destructurization of some structural materials under the impact of low temperatures;
- manifold freezing of heat flow;
- lack of fueling infrastructure.

The factors mentioned above lead to a growth in motor vehicle mass, a complication in design, a drop in efficiency and a markup in powerplant cost.

Solar energy, in fact, is not a limited product of the energetic space environment. The core issue of mobile solar energy sources is the limited surface of a vehicle for obtaining sufficient electrical power by means of solar cells, their efficiency is about 20 %. An electric vehicle with solar cells assembled into a battery also includes a storage battery.

By modern scientific and practical achievements, it has been proved that a conventional passenger solar car can travel a couple of tens of kilometers on the energy generated by the solar battery during daylight hours. In this case, for moving a vehicle, a solar panel with a capacity of several hundred watts, an electric motor with a capacity of about 2 kW and a storage battery with the comparable capacity are required. However, the low mileage and the insufficient drive power with solar batteries do not provide completely consumer demands for on-road performance and the operational parameters of a modern car. Therefore, the further development of solar energy in cars shifted to the second approach of development, that is the hybridization of the electrical energy source. In fact, a solar battery has become an additional means of replenishing the energy of TRB electric vehicles.

At the present time, some prototypes have been tested and even small-scale versions of electric vehicles recharging from the 'free' sun have been presented [36]. Due to their specificities, the designs of such electric vehicles are outwardly different from the cars with traditional bodies (Fig. 4.5).

An electric vehicle with solar batteries Eclectic produced by the French company Venturi (Fig. 4.5 a) provides a mileage without recharging the storage (traction) battery up to 50 km at an average speed of 50 km/h. Solar panels with an area of 2.5 m^2 (300 W) can accumulate enough energy during a bright day for a run of up to 10 km. If necessary, a nickel-metal hydride TRB with a capacity of 9 kWh is charged from a standard socket connector within 5 hours. In addition to solar energy, Eclectic can use the wind energy (mileage up to 15 km/day). The conceptual prototype has an 11 kW asynchronous electric motor with 45 Nm of torque. The weight of a car makes 400 kg, and the price starts from $ 20,000.

The next model of the conceptual prototype Venturi Astrolab (Fig. 4.5 b)

a b

Fig. 4.5. Designs of solar electric vehicle: a – Eclectic; b – Astrolab.

reached the commercial level with the improved characteristics: its powerplant power is equal to 16 kW (torque at 50 N·m), the maximum speed is 120 km/h, the mileage is up to 110 km. In these circumstances, the area of photovoltaic cells with a total power of 600 W boosted up to 3.6 m^2, and the TRB capacity grew up to 7 kWh (with the weight of 110 kg). The total weight of a car remained unchanged, but the price jumped up to € 92,000.

Further improvement of solar electric cars is in the direction of increasing the efficiency of solar panels and cutting down the TRB capacity at a given powerplant power. Developers of competing companies, in the near future, expect from their electric concepts to achieve up to 20,000 km of run per year, only on solar energy. If this occurs, the on-road vehicle characteristics will meet the requirements of a wide variety of cars.

To the discussed above energy sources on vehicles, in the distant future, it is possible to add *nuclear reactors* operating on atomic thorium. The energy density of such fuel allows a vehicle to be operated only at 'starting fueling' (8 grams per million kilometers of run) [37].

And finally, *electromagnetic fields* (geomagnetic, ground, near-earth and cosmic emitters) in the distant future can be implied as an energy product to activate mobile objects. When this occurs, filling stations can be transformed into distributed infrastructure objects in the form of a complex of antenna resonators-receivers and energy translators. The latter will transmit energy in the form of electromagnetic oscillations of a fixed (carrier) frequency on board a vehicle while driving. Thus, the frequency rationing of 'contactless charge' by a wave energy carrier will minimize the cost of a resonator-receiver on board a vehicle.

4.2. Component parts and structure of an automobile electric drive

In the general case, the electric drive structure (ED) of a car can be represented by variant components of its main functional elements (Fig. 4.6).

Primary source of electrical energy in electric vehicles is the *traction rechargeable battery* (TRB). It consists of many individual batteries connected in series (Fig. 4.7).

At the TRB output, a voltage of 300...700 V is formed with a capacity of approximately 60 kWh, comparable with the traction motor power (up to 100 kW). In electric vehicles, lithium-ion batteries are most commonly used. The average service life of such TRBs makes up 8 years, the battery cycle life

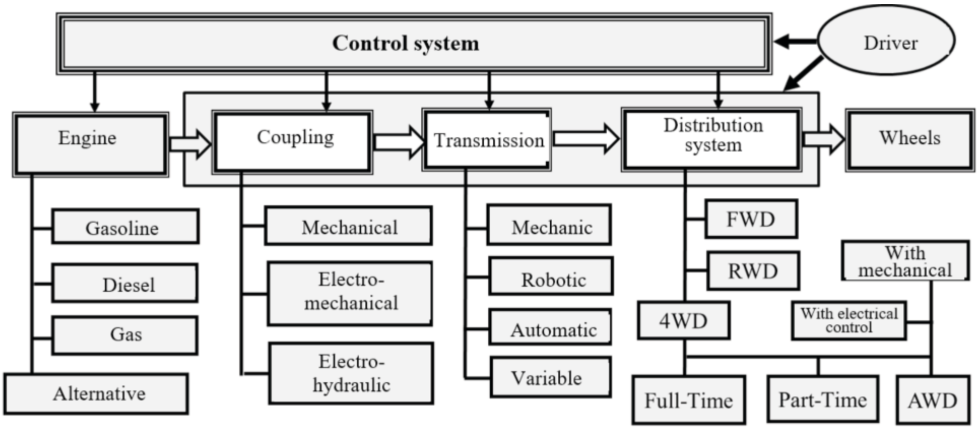


Fig. 4.6. Structural parts of car electric drive.

is several thousand cycles. In this regard, such TRBs have an increased specific energy capacity (0.8...2.6 kWh/kg), low self-discharge and improved charging characteristics [38].

The main limitation of lithium-ion batteries is a narrow operating temperature range (25...45 °C) and its high cost. Low temperatures cause a loss in capacity; high temperatures lead to the structure destruction. Among other things, the TRB designs include temperature sensors and a forced cooling system as well. An important role, in this case, plays the proper TRB operation which ensures: standard battery charging and discharging; reasonable operation in winter and summer periods; competent choice of a load mode; timely determination of the

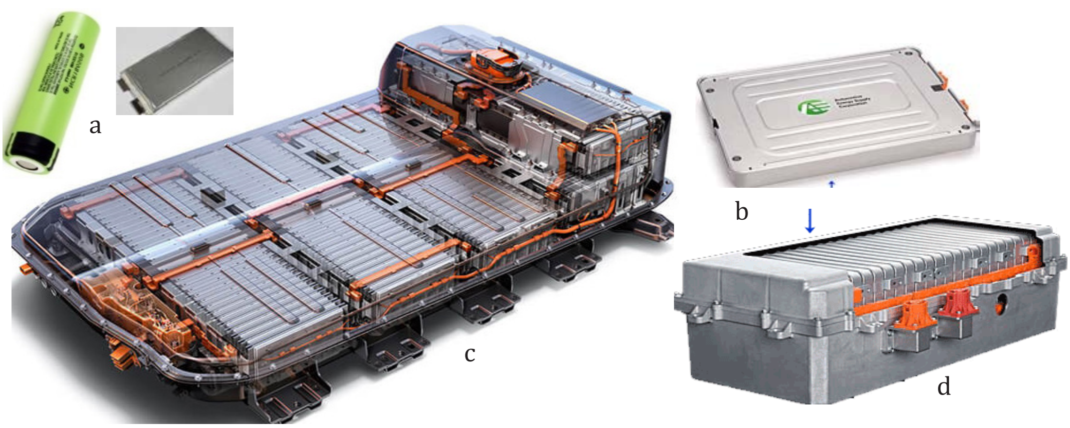


Fig. 4.7. Structure of traction rechargeable battery: a – battery cells; b – battery module; c – battery pack; d – Chevrolet Volt EV traction battery in the assembled condition.

capacity loss of individual batteries [39].

Battery charging from an external industrial electrical power network is guaranteed due to the presence of an on-board *charger* in the design.

A *generator* is used in hybrid units and is powered by an ICE. It serves as an autonomous source of electricity to charge a TRB or to power a traction motor. In regenerative systems, the generator is powered by the braking energy of a vehicle.

A high-value *capacitor* is commonly applied as a buffer energy storage in hybrid units or recuperative systems. In comparison with a TRB, the capacitor is characterized by high dynamics of charge-discharge processes.

The use of electric traction has become a priority area for the development of a high-quality motor vehicle structure due to a number of advantages of electric motors over ICEs in transport modes:

- high efficiency rate (up to 95 %);
- high specific power and low price;
- easy-to-use and easy to work with;
- eco-friendly while operating;
- reliability and durability;
- air cooling;
- minimum transmission weight (does not require a gearbox);
- possibility to recuperate the braking energy in the generator mode;
- maximum indicators of torque M and power P in a wide range of rotation rates.

In electric vehicles, various types of electric motors can be used as traction motors, which have the property of reversibility (the ability to work as a motor and a generator as well). The name of an electric motor type contains several words determining its qualitative characteristics (attributes):

- current type (AC, DC) and the number of phases (single-phased, three-phased, multi-phased) of the supply voltage;

- method of transmitting the energy to a rotor (collector, brushless, inductor, reactive);
- principle of AC motor construction (synchronous, asynchronous with squirrel-cage or phase rotor);
- method of supplying the voltage (network, valve);
- special properties of the rotor **configuration** (with built-in permanent magnets, with external **insertion** of permanent magnets, with **field magnetizing coil**, reactive, inductive) and motor designs (**framework,** flanged cylindrical, disk);
- method of torque transmission (axial with an inner rotor, hub with an outer rotor) and the armature winding location (on the rotor, on the stator).

The latter features are essential for wheel-motor systems. Let us characterize an electric motor with limited and special purposes.

Inductor synchronous machines (motors and generators) have a traditional stator design with distributed poles of an armature polyphase winding, connected by the 'triangle' or 'polygon' scheme. Field winding is also located on the stator and forms an inductor with machine iron (a device for exciting a magnetic flow). Due to the absence of an electrical circuit on the rotor, such machines belong to the group of non-contact rotor devices. The design of a 4-phase inductor machine is shown in Fig. 4.8.

In a synchronous motor, the excitation from permanent magnets placed on the rotor can be used (there is no excitation winding). Such motors have a number of benefits:

- no brush-collector unit (longer life cycle, reliability, operating rotation rates);
- wide range of supply voltages, significant over-torque is allowed;
- dynamic of high moment;
- feasible speed adjustment without sacrificing torque at low speeds or controlling the power at high speeds;

a b c

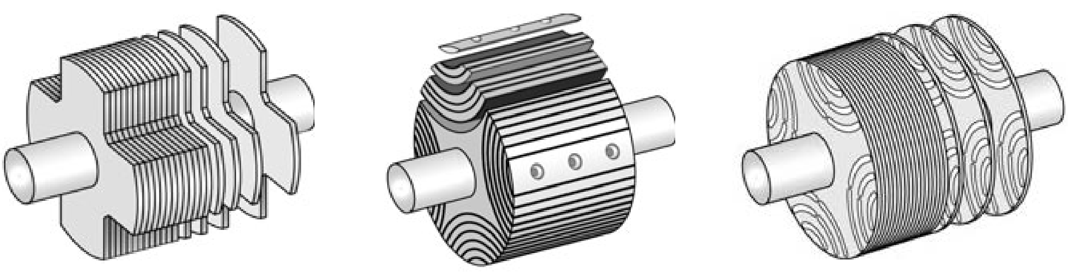
Fig. 4.8. Design of induction machine: a – layout diagram; b – rotor design; c – stator design.

- efficiency over 90 %;
- minimal losses at idle;
- low weight-size indicators.

A *synchronous reluctance motor* (SRM) has a stator of traditional pole structure with a distributed winding and does not have an excitation winding or permanent magnets. Torque occurs in a jet engine due to the inequality of magnetic conductivities along the transverse and longitudinal rotor axes. To obtain a reactive field, various rotor designs are applied (Fig. 4.9).

The rotor surface with apparent poles imitates the design of an inductor motor rotor, and therefore it is called a reluctance motor (inductor-reactive). The strong points of jet engines are the following:

- senso drive;
- simple and reliable rotor design without magnets and windings;



a b c

Fig. 4.9. Rotor design of synchronous reluctance motor: a – with projecting poles; b – axially laminated; c – cross-laminated.

- low heating (no currents in a rotor);
- low electric motor price;
- low moment of rotor inertia (energy saving on the acceleration);
- wide range of rotation regulation (frequency regulator).

The weak points of jet engines include a low power ratio and the need to use a frequency converter while operating in the high-speed range.

DC valve motors, in contrast to direct-connected motors (from the industrial network), have external current commutation supplying the voltage by DC/AC converters – inverters. The brushless valve motor system, apart from a multi-phase AC machine (synchronous or asynchronous), additionally includes a coordinate converter, an inverter and a position sensor.

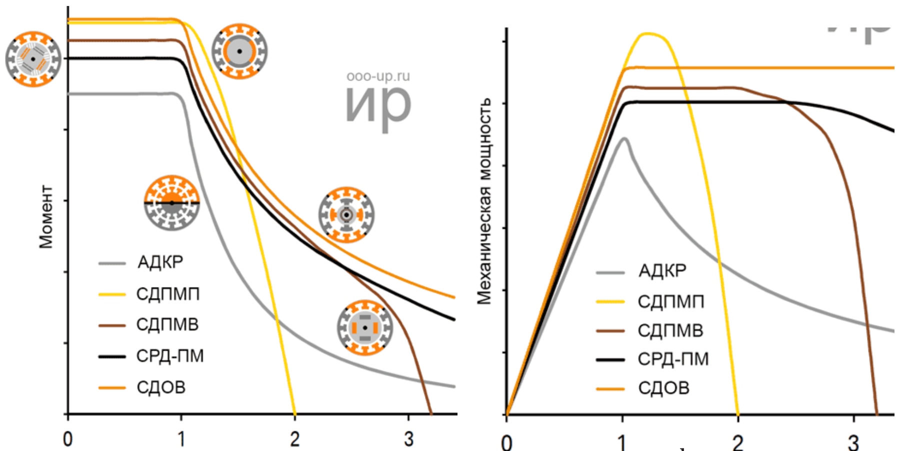
A valve inductor motor (VIM) combines the properties of both an electric machine and an integrated system of controlled electric drive. As a system of controlled electric drive, a VIM makes it possible to control this process in accordance with the characteristics of a particular load: to regulate the speed, torque and its power.

Valve electric motors of the following types are used as traction electric motors in vehicles:

- squirrel-cage induction motor, SCIM;
- surface magnet synchronous motor, SMSM;
- interior permanent magnet synchronous motor, IPMSM;
- permanent-magnet reluctance motor, PMRM;
- magnetizing coil field synchronous motor, MCFSM.

The comparative characteristics of the listed above types of valve motors with a limited stator current are shown in Fig. 4.10.

According to the given performance characteristics, electric motors of PMRM and MCFSM types are the most suitable for operating as traction units on a car. Such motors ensure high efficiency over a wide operating range [40].



a b

Fig. 4.10. Mechanical characteristics of electric motors of different types: a – moment; b – power

In contrast to an electric motor (EM) of general use with **installation** on a flat **framework**, original designs of machine bodies for an automotive drive have been developed (Fig. 4.11).

The Yasa Motors company has presented a disc unit with a weight of 25 kg capable of delivering up to 650 Nm of torque. The design originality has been obtained by the use of flat pole technology.

Combining an electric motor and a part of the transmission into one module (electric transmission) led to creating universal modules of various purposes (Fig. 4.12).

Such modules can be used as the main drive for an electric vehicle or as an auxiliary drive in a hybrid car by adding an electric drive to one of the axles. An

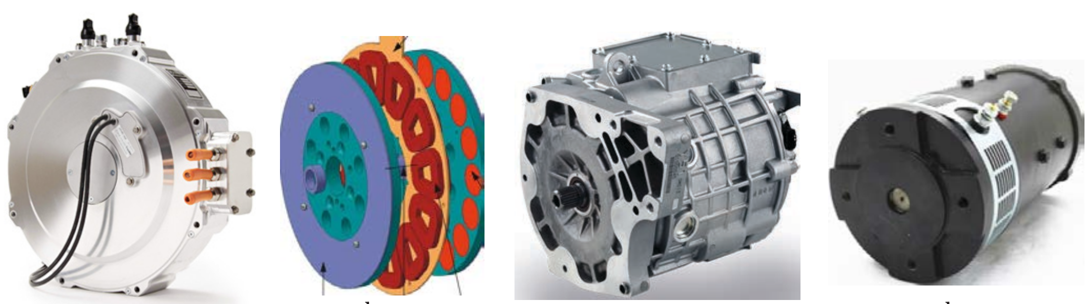

a b c d

Fig. 4.11. Traction motor designs: a – Yasa Motors disk motor; b – structure of flat poles; c – flanged; d – cylindrical.

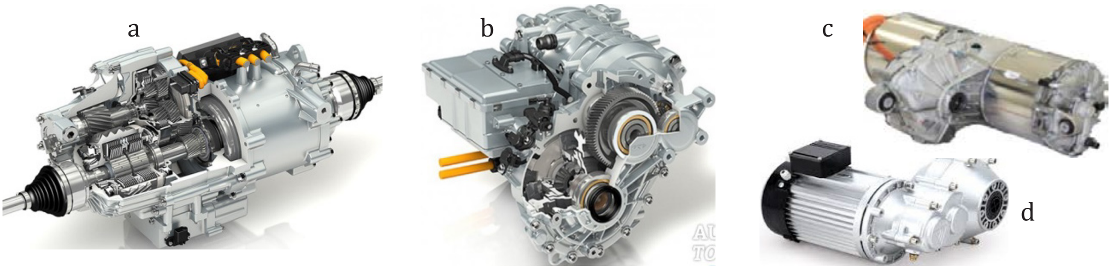


Fig. 4.12. Industrial designs of electric drive modules: a – Driveline rear axle; b – eDrive central unit; c – driving axle; d – cardan or axial Renault Zoe installation.

example of applying a three-phase asynchronous unit is a Chevrolet Volt hybrid vehicle. A car with a three-phase synchronous motor i-MiEV from Mitsubishi is exclusively an electric-type [41].

The second approach in implementing an electric vehicle drive is the use of motor-wheels technology, with a number of benefits such as:

- no cumbersome transmission;
- excellent drive dynamics (torque up to 700 Nm at low revolutions);
- vehicle maneuverability due to the independent control of each wheel;
- simplified braking energy recovery system;
- ensuring active traffic safety (implementation of algorithms for electronic braking systems, dynamic vehicle stability, driver assistance).

Along with this, there are significant downsides presented further:

- excess weight of mechanisms placed inside the rim cuts down comfort and controllability, improves the wear-out and boots the efficiency of the **suspension**;
- high air-tightness requirements;
- significant costs to repair a motor-in-wheel or its replacement.

Motor-in-wheel systems of various types differ in the principle of construction and the rate of integration (Fig. 4.13).

So, for example, in the design of Voltage sports car, produced by the Michelin

Fig. 4.13. Motor-in-wheel design options: a – active wheel-suspension of a discrete Active Wheel structure; b – stepping integrated electric drive E-Wheel Drive; c – disk electric motor; d – Duyunov's wheel-in-motor; e – Shkondin's wheel-in-motor.

company, Active Wheel traction motor is used (Fig. 4.13 a). It weighs only 7 kg, which makes it possible to achieve an acceptable wheel weight of 11 kg. The basis of Active Wheel is a lightweight aluminum frame, connected to the body subframe by a simple rigid lever. The connection is made movable so that a wheel can rotate.

All Active Wheel elements are attached to the inner frame surface, and the rim itself is fixed on a flat disc-shaped hub. Its braking mechanism consists of a rotating disc and calipers with electromagnetic actuators. While braking, the traction motor operates as a generator creating electrical energy for TRB charging. The suspension consists of a steel spring and electric shock absorbers. The motor controlling shock absorbers are also responsible for turning a wheel. Thanks to the large turning angle, an electric vehicle is much more maneuverable than conventional cars.

Schaeffler company presented the E-Wheel Drive (electric all-wheel drive), a tractive electric hub drive, developed in the cooperation with Ford (Fig 4.13 b). All required functional elements – electric motor, power electronics and controller, as well as brake unit, are located inside the wheel disk. Two electric motors mounted in the rear wheel arches of a Ford vehicle have a capacity of up to 40 kW each (80 kW and 66 kW) and develop torque of up to 700 Nm.

a b c

Fig. 4.14. Duyunov's motor-in-wheel design: a – turning on the windings according to the 'Slavyanka' scheme; b – magnetic system design; c – exterior view.

Motor-in-wheel of Duyunov's model is designed on the basis of an asynchronous motor, where permanent magnets are not used (Fig. 4.13 d). This design is developed on the grounds of the innovative technologies of combined windings. The advantages are the following: low cost, excellent **coasting**, low noise level, long service life, increased reliability. Motor windings are combined into a six-phase motor according to the 'Slavyanka' scheme (a combination of 'star' and 'triangle'), its phases are connected to the network in three stages (Fig. 4.14 a).

Designing originality of Duyunov's motor lies in the fact that the motor stator (1) is inside, and the squirrel-cage rotor (2) rotates outside (Fig. 4.14, b). The information from the angular stator position sensor (encoder) and the temperature sensor are used to control an EM. Sensors (3) are located on the stator, and the incremental encoder disk is on the rotor.

The strengths of the original EM in comparison with conventional asynchronous electric motors are: high efficiency in the load range of 25...150 % (electricity consumption reduction by 15...50 %); growth in maximum torque by 10...100 %; increase in starting torque by 20...50 %; reduction of starting currents by 2 times; noise level lowering by 6...7 dBA; cut down in temperature of winding heating [42].

In the *Shkondin's motor-in-wheel*, the stator is also located inside a rotor (Fig. 4.13 e). On a stator there are 11 pairs of magnets with the alternating polarity, and on a rotor there are 6 electromagnets with U-shaped cores, which form 12 rotor poles (Fig 4.15 a, b).

Thus, in the Shkondin's hub motor, armature windings are placed on the rotor and field magnets are on the stator. When a rotor rotates, the current of different polarity is fed to its windings via a collector-and-brush assembly unit.

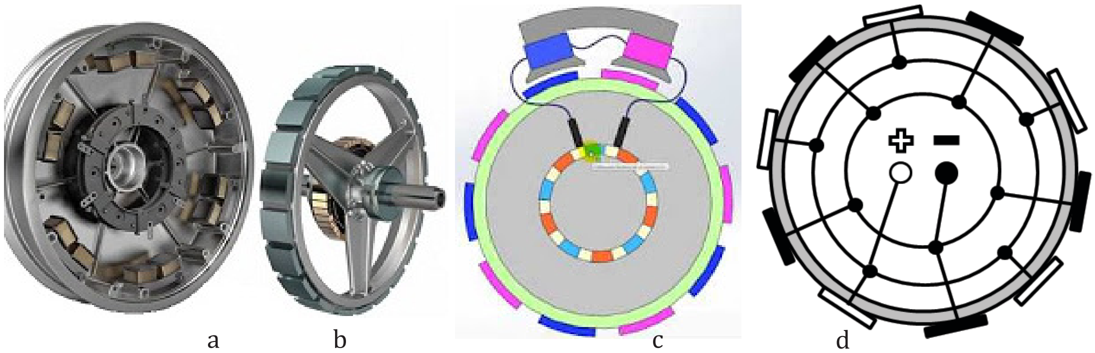

a b c d

Fig. 4.15. Design of Shkondin's motor-in-wheel: a – rotor; b – stator; c – phasing of the poles; d – connection of stator collector plates.

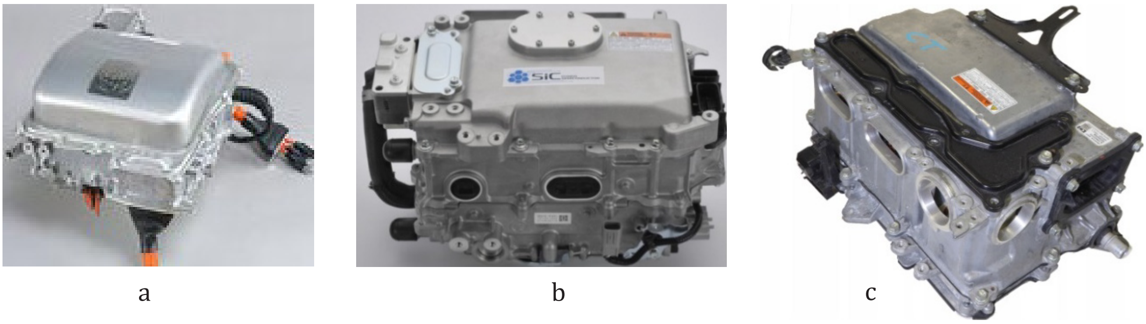
Phasing of rotor pole coils and stator magnetic poles is performed in such a way that at certain moments in the coinciding poles of the stator and a rotor, the winding power supply is turned off, and in the remaining (non-coinciding) poles the attraction is replaced by its repulsion (Fig. 4.15 c, d). As a result, torque is generated by the so-called Shkondin's electro-mechanical trigger.

In addition, a trigger partially recuperates the energy from the windings that are not involved in the torque creation (the coincident position of the poles) and operates as a generator. The main strengths are high efficiency, simple design and relatively low cost. The weaknesses include poor thermal conditions, difficulty in adjusting the rotation rate, presence of a mechanical commutating equipment [43]. The last drawback is eliminated in the collectorless motor type, where armature windings are mounted on the internal stator, and the pole magnets are placed on the external rotor. With that, current commutating in stator windings is performed by electronic keys (hardware drivers) on the basis of the information received from rotor sensors (RS).

To adjust source voltage parameters (TRB, generator) and consumers (electric motors) of electrical energy, various electronic devices are applied, combined in the design of AC adapter modules [44] (Fig. 4.16).

Some AC adapters (secondary voltage sources) in the structure of indicated modules have different purposes (conversion function) and a circuit design [45], [46] (Fig. 4.17).

DC voltage (Fig. 4.17 a) is boosted by current commutating in the L reactor. Induced voltage pulses with double amplitude at the output of variable torque (VT) key are integrated by the capacity of smoothing C filter, keeping a DC voltage at the output of the converter. A resistor at the output of the R converter provides



a b c

Fig. 4.16. Exterior of AC adapter modules in vehicles: a – Toyota Prius II; b – Camry Hybrid SiC; c – Lexus CT 200 h.

the C discharge when the rate of regulated voltage drops.

DC voltage (Fig. 4.17 b) is cut down by boosting the reactive resistance of inductive L element, with the help of rising the current commutating frequency by VT key. A resistor at the input of the R converter adjusts the converter resistance with the power of input voltage source.

To convert a constant voltage into a three-phase voltage system, a **six-shoulder key bridge** powers a traction electric motor (Fig. 4.17 c). Series key application according to a certain algorithm allows to obtain the required

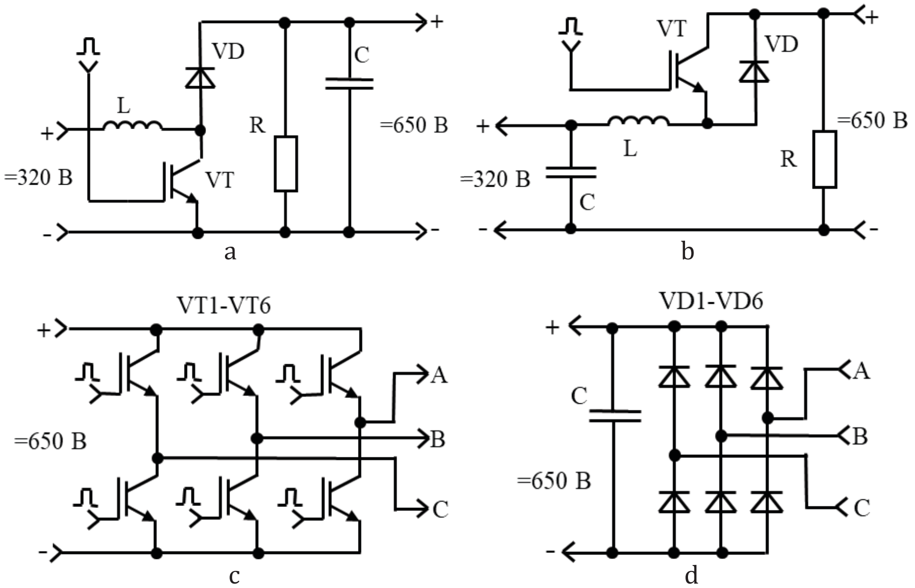


Fig. 4.17. Electrical schematic diagrams of voltage converters: a – up DC/DC converter; b – down DC/DC converter; c – DC/AC converter (inverter); d – AC/DC

sequencing of voltage pulses for each phase. MOSFET or IGBT transistors are usually applied as inverter power keys, less often four-layer switches. Voltage reverse conversion (from three-phase to constant) is performed by the three-phase rectifier according to the Lavrent'ev's scheme (Fig. 4.17 d).

Reversible electrical machines convert electrical energy into torque and reverse conversion. The most effective converters, in this regard, are three-phase valve machines, characterized by high specific indicators.

The *control system* of electronic drive performs several functions:

- controlling the energy used;
- controlling the break energy recuperation;
- assessing the charge rate;
- controlling the dynamic motion;
- providing the necessary motion mode;
- controlling the traction;
- controlling the voltage.

4.3. Application of valve electric machines

The principle requirements for the ED control system correspond to vehicle motion modes:

- smooth rotation change;
- automatic control at set speed;
- smooth control of traction and braking torque;
- automatic maximum torque and power restriction;
- limitation of charging current while regenerative braking;
- ability to move in a coasting mode with subsequent smooth acceleration or electric braking.

At the functional level, valve electric motor systems (VEMS) consist of an electric machine (controlled **object**) with a coordinate sensor (feedback), a voltage frequency converter (control system) and power circuits (actuators). Based on this, valve control systems can be distinguished by general classification features:

- type of a controlled motor (direct current motor (DCM), synchronous, asynchronous);
- current type of primary voltage power source;
- method of control (scalar, vector);
- type of feedback sensor (rotor position, current in phases).

On board a car, a primary voltage power source is a DC TRB, therefore the use of DCM requires minimal costs for implementing a control system (direct or adaptive pulse width modulation (PWM) regulation). However, due to a low power density and its reliability (the presence of a collector-and-brush assembly unit), DCMs are significantly inferior to AC machines. Valve systems of synchronous and asynchronous motors are also distinguished according to their connection voltage, that is, direct current BLDC (BrushLess DC) and alternating current PMAC (Permanent Magnet AC). Naturally, valve systems of the first group are used in a vehicle.

The control method for AC valve machines determines the operation algorithm of a system frequency converter [47].

The *scalar control method* consists in keeping a constant ratio between the amplitude values and the connection voltage frequency in the entire operating rotation range. At the same time, the current strength in windings, and therefore the torque on a shaft, remain unchanged. Such a system is simple to implement, but has the following drawbacks:

- impossible to implement a sensorless control system by the asynchronous motor (AM) with a sharp increase in load;
- a sensor system of rotation control under load for a synchronous motor has low accuracy and can completely go out of synchronism;
- impossible to simultaneously control both torque and the rotation rate.

Vector control, unlike scalar control, controls independently and noninertialy the rotation rate and torque on a motor shaft. In this case, in addition to the noted parameters, it is also necessary to control the phase of connection voltage (to control a value and a space vector angle) [48]. To obtain a rotating magnetic field, the transfer from voltage parameters to phase currents parameters depending on motor load by a coordinate converter is required. Rotating coordinate system of phase currents creates the maximum motor torque when the phase flow vector is shifted to the rotor flow vector by 90°. Vector control is performed in several ways:

- field-oriented control;
- torque control with space-vector flow modulation;
- direct torque control with a switch-on table;
- adaptive direct torque control;
- direct torque control with pulse width modulation;
- direct self-government.

The implementation of the algorithm for controlling brushless direct current electric motor (BDCEM) from the constant voltage network, requires information on the current position and frequency of machine rotor rotation (feedback). This information is obtained on the basis of signals from sensors of **rotor angle** or the instant current value in machine phases.

Analog and digital sensors of various construction principles and design are used as rotor sensors (RS) (Fig. 4.18).

Three-phase AC tachogenerators with excitation from permanent magnets (Fig. 4.18 a) are installed in the structure of a synchronous machine. The electromechanical sensor signal simulates phase voltages supplying to a synchronous machine and includes the information about a rotor position (instant voltage values) and the motor rotational speed (voltage frequency) during its stationary rotation. Its disadvantages are: no signal (information about the rotor initial position) at the moment of **starting**; available only for synchronous motors; signal inadequacy under dynamic loads; the requirement to convert a signal into a digital form and retrofit the basic machine design.

A *transformer-type* synchronizing signal sensor (resolvers) on a stator has three windings (an excitation and two receiving windings). The receiving (secondary) coils are shifted relative to each other by 90 electrical degrees

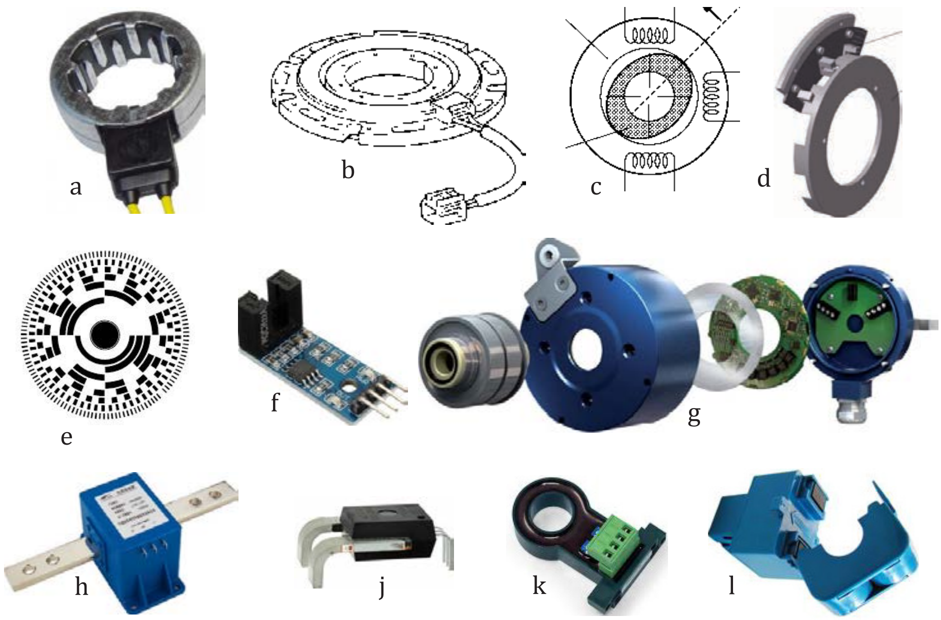


Fig. 4.18. Design and **structure** of sensors of brushless direct current electric motor system: a – generator type; b, c – transformer type (resolvers); d – Hall **assemblage**; e – coding disk of encoder; f – encoder measuring part; g – encoder assembly; h–l – current sensors of various designs.

(Fig. 4.18 c). An ellipsoidal core is mounted on a rotor sensor. AC voltage of carrier frequency is supplied to a primary excitation coil, which is transformed into receiving windings with different transmission (transformation) ratios depending on the rotor (core) position. When a rotor rotates, signals in receiving windings are modulated in amplitude with a shift of 90 electrical degrees. Thus, on the basis of two sinusoidal signals, not only a current position is identified, but also the direction of rotor rotation. The main advantage of a modulating transformer over a generator-type sensor is the presence of an information signal in a stationary (starting) state of traction motor.

*Hall **plate element*** consists of a set of digital Hall sensors placed on a stator with certain angular offset and a rotor disk with shielding plates or permanent magnets (Fig. 4.18 d). When the disk rotates, a digital signal is created in the sensor, its placement corresponds to the angular position of traction motor rotor. The main advantage is receiving a signal in a digital form at the rotor rotation and in its start state.

Encoders (optical, magnetic, capacitive) contain a coding disk on the rotor (Fig. 4.18 e) and a measuring unit on the stator (Fig. 4.18 f). Usually, the design of a sensor has a complete form (Fig 4.18 g). The main advantage is getting a signal in the form of a binary code with high resolution (in dynamics and statics), when this occurs the disadvantage is only its high cost.

Contactless current sensors designed on the basis of linear Hall microcircuits have a galvanic (Fig. 4.18 j, l) or remote (Fig. 4.18 m, n) connection. Unlike those considered, current sensors are not placed in the design of an electric machine, but are mounted in the power electronics unit (sensorless systems). Special sensors have a built-in analog-digital converter (ADC). Information from phase current sensors can significantly simplify the algorithm for controlling currents in stator windings of an asynchronous motor.

For example, let us consider a block diagram of the valve synchronous motor system, which consists of a synchronous motor (SM) with a coordinate RS, an electrical machine controller (EMC) and power circuits. RS signal also includes information about the rotor rotation frequency of rotation velocity sensor (RVS) (Fig. 4.19).

The BDCEM speed rate is set by a signal from the accelerator sensor AS (electronic pseudo-pedal). At the moment of the start of movement or in a stationary speed rate, the coordinate converter calculates optimal phase-time parameters of the phase currents based on the information received from the RVS sensor and the phase converter. The phase converter calculates coordinates on the grounds of instant current values in two motor phases. The converter control unit, according to the obtained data, generates a sequence and pulse period-to-pulse (PWM **controller**) duration ratio to control converter power switches. Converter power circuits commute the current of SM phase winding with the power supply voltage. A DC/+DC up converter adjusts the rated TRB voltage with the operating voltage of a motor. A PID controller with feedback stabilizes the selected speed mode by the signal from the rotation frequency sensor (RFS).

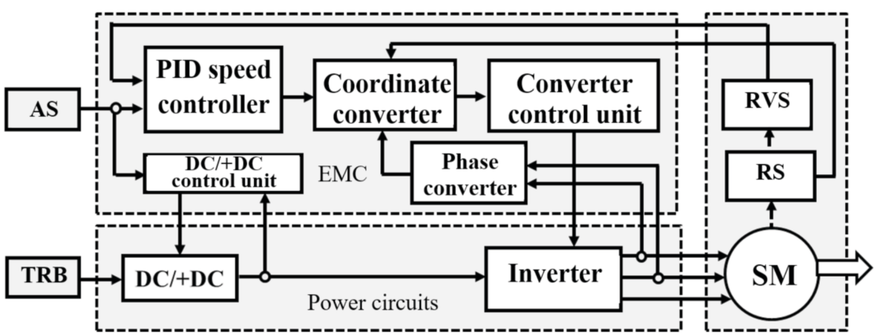


Fig. 4.19. Block diagram of valve synchronous machine system.

In dynamic modes, the change in rotation speed under a given motor load can be achieved in two ways – by changing the supply voltage at the DC/+DC output or by changing the average current value with the help of PWM control in the converter control unit. In the first case, the frequency rate or the **control of pulse relative duration** are adjusted by the DC/+DC converter key (see Fig. 4.17 a), in the second one, the **pulse relative duration** of **control signals** is performed by inverter keys (see Fig. 4.17 c).

A control system of a brushless DC asynchronous motor is designed in a similar way, but taking into account the fact that the magnetic field vector of a rotor (**flux linkage**) is determined on the grounds of the characteristics of stator rotating field and the distributed parameters of short-circuited rotor windings. All things considered, an additional recalculation of coordinate values is carried out by the derived functions of phase currents.

The reversibility of synchronous machine is used to charge a TRB with the regenerative braking energy. In this case, an ED operates in modes of motor 'M' and generator 'G', activated by pressing accelerator and brake pedals, respectively (Fig. 4.20).

In a motor mode, the power is transmitted in the manner described above (full arrows). In a generator mode (dotted arrows), a three-phase rectifier (see Fig. 4.17 d) and a DC/-DC down converter are used (see Fig. 4.17 b). The EMC receives the information on the rotor speed *n* in the generator and the charging current *i* rate in the TRB battery. On the ground of these signals, the EMC regulates the transmission coefficient of DC/–DC converter by changing the frequency of control pulses f_2, and limiting the charge current to an acceptable rate. In real systems, the temperature of power element is monitored to determine their electrical load and the TRB charge ratio.

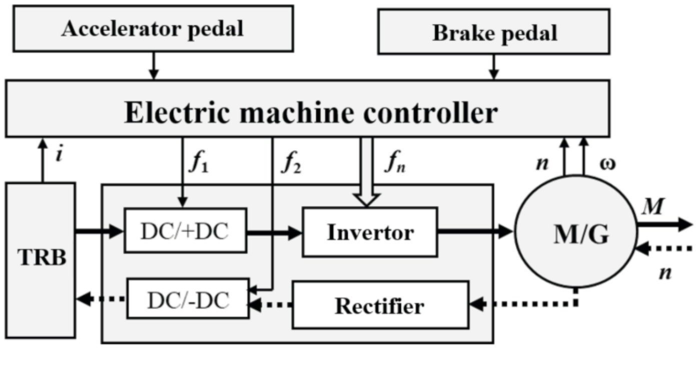


Fig. 4.20. Functional diagram of an electric drive with a reversible machine.

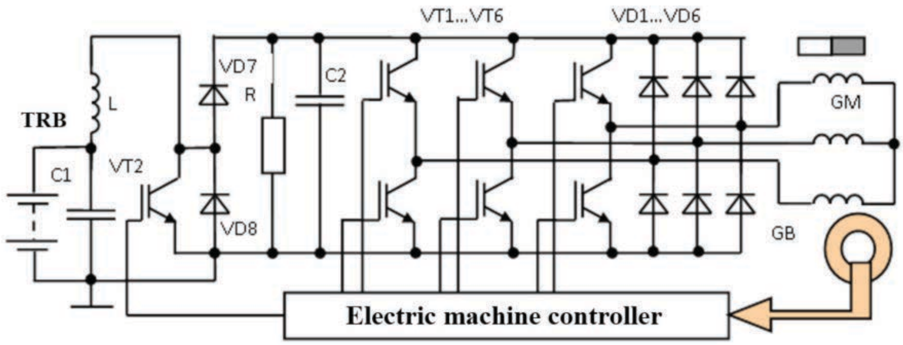


Fig. 4.21. Electric diagram of power circuits of a reversible drive.

Synthesis of individual voltage converters of various purposes (see Fig. 4.17) allows them to be combined into a block of power circuits (Fig. 4.21).

The diagram (Fig. 4.21) indicates the following elements: VT1 ... VT6 are transistor **switches** of an inverter; VD1…VD6 are rectifier diodes; VT7 is a **switch transistor of DC/–DC down-converter**; VT8 is a **switch transistor of DC/+DC up-converter**; VD7 is a DC/+DC rectifier diode; VD8 is a VT8 protection diode against pulse overvoltage; C1 is a smoothing filter; C2 is an integrating capacity; L is a throttle reactor; R is a discharge resistance.

4.4. Structure of electric hybrid power plants

An electric hybrid in the full sense of the word is understood as a car driven by an ICE and a traction electric motor. At the present stage, electro-hybrids are classified according to several characteristics (Fig. 4.22).

Depending on what role an electric machine plays in the power plant, hybrids are divided into assisted hybrids, mild and full hybrids. In the first type, an electric vehicle implements Start-Stop and braking energy recovery systems, in the second one, an electric motor serves as an ICE assistant in the acceleration and speed-up modes, in the third one, an electric drive provides a limited autonomous mode of movement by the electric traction.

Assisted-hybrids usually have a resident voltage of 12 V, and an electric machine (generator) has low power. In such systems, a reversible electrical machine of dual purpose (**dinostarter**) can be applied.

A mild hybrid capacity and an electric motor of limited power, which starts working when a car accelerates along with the operation of the main ICE. Mild

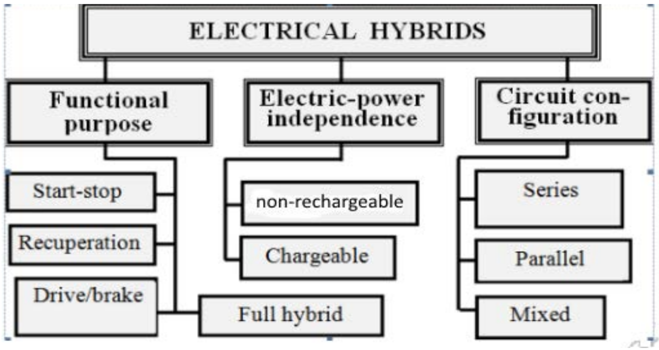


Fig. 4.22. General classification of electric hybrids.

hybrids contain electric drive systems with over (36 V, 42 V) or high (300–700 V) voltage. At the same time, the **composition** of hybrid drive provides for a parallel or **series combining** of alternative engine torques (Fig. 4.23).

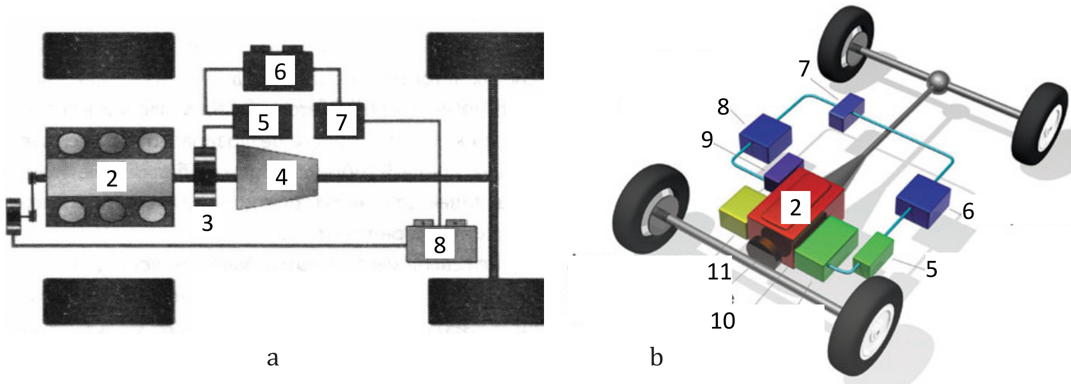
In Fig. 4.23 let us indicate the following parts of mild hybrids: 1 – generator (12V); 2 – ICE; 3 – electric motor; 4 – automatic transmission; 5 – module of power electronics; 6 – TRB; 7 – DC/DC-converter; 8 – rechargeable battery; 9 – starter; 10 – reversible electrical machine; 11 – electromagnetic clutch.

A *full hybrid* is characterized by the high drive voltage and the significant traction motor power. For these reasons, the system has a powerful high-voltage TRB. In full hybrids, TRB charging devices from external sources increase the autonomous run on electric traction. Such cars have received the name of a plug-in hybrid. A compromise scheme of active Start-Stop system is possible, where the electric traction is applied only in the mode of short-term movements at low speed (driving in a traffic jam) with the use of electric starter. The main disadvantage of this solution is the significant power consumption of a starter electrical motor in starting characteristics. This can be solved by using electric brushless DC or pneumatic starter machines. In the latter case, a car has the maximum efficiency and torque at minimum vehicle speeds.

From the perspective of the structure and the method of transmitting power to car wheels, there are three configurations of hybrid systems, namely, series, parallel and mixed (combined).

A *series hybrid* is a **configuration** where an ICE, operating in the maximum fuel economy mode, sets in motion only a generator (Fig. 4.24 a).

The electricity generated by the G generator by means of ICE with a fuel tank is fed by the **voltage switch converter** (VSC) either to the traction electric

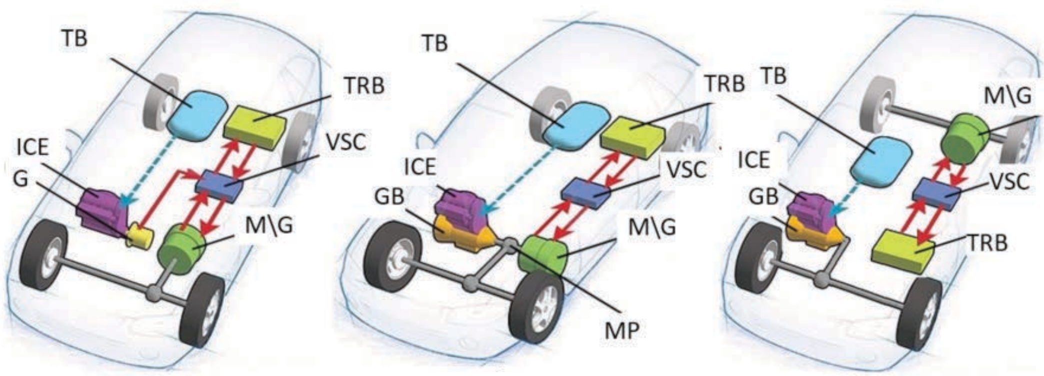


a b

Fig. 4.23. **Compositions** of mild hybrids: a – parallel; b – **series.**

motor (reversible machine M\G), or to the TRB. The electric motor transmits the required torque to wheels by the transmission, and in the braking mode it works as a generator, recharging a TRB (recuperation mode of braking energy). Such hybrids are marked with the abbreviation REEV (range-extended electric vehicle) or EREV (extended-range electric vehicle). The strong points of this scheme are the ICE constant operation in the most economical mode, ease of control and the absence of a complex transmission. Among the minuses there should be noted: low efficiency of mechanism for transferring energy from the ICE to driving wheels of a car; impossibility to achieve the power combining of alternative motor drives; inability to move while completely discharging a rechargeable battery (it takes certain time to charge from a running ICE).

Parallel configuration provides for the mechanical connection of ICE with a



a b c

Fig. 4.24. Schemes for switching on the HPPs of front-wheel drive vehicles: a – **series**; b – parallel; c – mixed.

GB and a traction M\G electric motor with driving wheels of a car (Fig. 4.24 b). An electric motor, though, is powered by TRB. Since torque from an input shaft can be directly transmitted to wheels, the efficiency of this hybrid is higher. However, in regards to fuel and environmental performance characteristics, the parallel circuit is inferior to the **series** one (the ICE cannot operate constantly in the economy mode). It should be emphasized that in certain **compositions** of parallel circuit, an electric motor can be connected directly to an ICE crankshaft or by means of the clutch to the vehicle transmission.

A ***series**-parallel* (mixed) configuration combines the advantages of both abovementioned schemes. Depending on the driving conditions, a car can move either by electric traction, by ICE, or in the hybrid mode. In these circumstances, alternative motors can be connected to one or different wheel axles. In the first case, a hybrid gearbox with a planetary timing gear [49] is used. For the second option (Fig. 4.24 c), a special transmission is not required, but the possibility of charging TRB from the ICE is excluded. In both options, electricity generated in the regenerative mode by M/G machine is fed to the TRB recharging. The mixed hybrid is also used in the all-wheel drive models. In such a case, front axle wheels are driven by the combined drive, as in front-wheel drive models, and a certain traction electrical motor is used to drive rear axle wheels.

The practical experience of car makers shows that further improvement of electric hybrids involves the implementation of:

- overvoltage supply and two-level power supply system;
- several reversible brushless DC electric **machines**;
- high-voltage TRB with power-consuming **structures**;
- modern components of power electronics.

For example, let us take a closer look at the functional structure of an electrical part of HPP with one reversible electrical **machine**. The central unit of HPP control system is a high voltage circuit controller (VCC). The controller distributes power to power plants in the installation, ensuring the optimal use of the power of electric drive and ICE to obtain the required torque on car wheels while accelerating and the energy recuperation while braking. The electromechanical part of power plant consists of hybrid gearbox (HGB), voltage converter unit (VCU) and high-voltage battery unit (HVBU) (Fig. 4.25).

Hybrid gearbox is like a double-acting electromechanical power transmission. Brushless electric motor, as part of HGB, operates in the modes (**statuses**) of

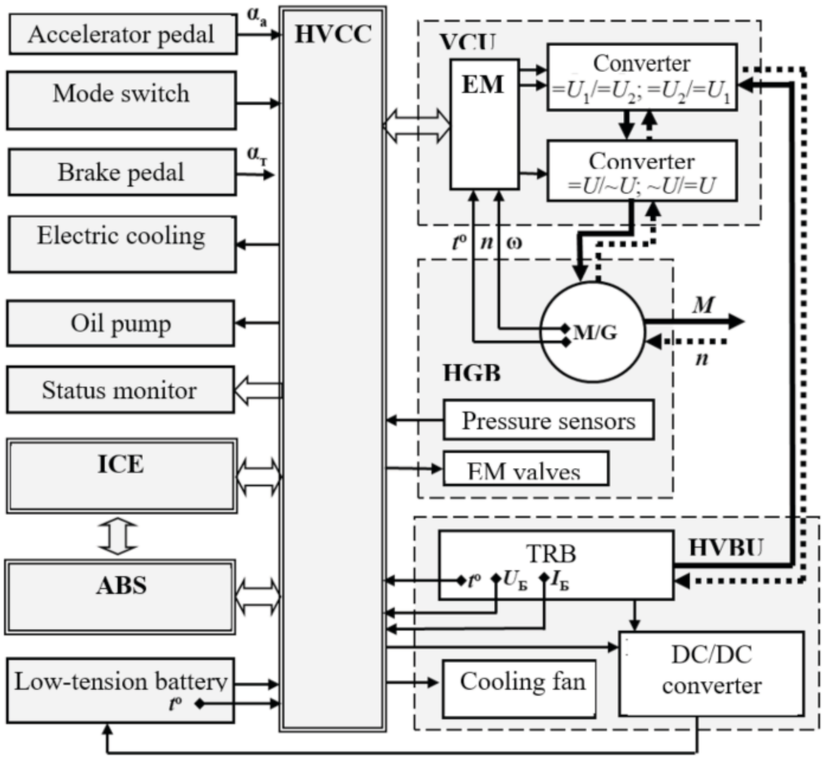


Fig. 4.25. Block diagram of HPT control system.

engine ‘M’ or generator ‘G’, depending on the driver's actions and various road situations.

In the ‘M’ mode, electric power is supplied to the electric motor from the traction battery. If this occurs, a DC/DC overvoltage converter ($=U_1/=U_2$) and a three-phase inverter ($=U/\sim U$) are used in the VCU. Operating modes of voltage converters (voltage rate, current values and inverter switching frequency) and, accordingly, of electric motor (torque and revolutions per minute) are determined by signals from accelerator pedal sensors α_a and the angle of electric machine shaft ω (fed to the EMC).

In the process of the energy recovery in the ‘G’ mode, the electrical power of TRB charge is back transformed into the mechanical power of rotor rotation in the electric generator. In this status, a three-phase rectifier ($\sim U/=U$) and an undervoltage DC/DC converter ($=U_2/=U_1$) are involved in the power circuit of a drive. The TRB charge mode at each moment while braking is calculated in HVCC on the ground of the information received from the ABS sensors and the brake pedal α_g.

If the driving mode of a car and the driver's actions activate the hybrid drive, the HVCC controller determines the driving mode of a car and controls the ICE and traction motor moments, taking into account the charging rate of high-voltage battery (voltage U_B and current I_B in the TRB charge circuit) and its temperature t^o. The difference between the required driving power and the ICE power is compensated by electric motor power.

Additionally, the information on the temperature of electrical machine, inverter and rectifier can be used, which indirectly characterizes the operational and technical condition of these units. In such event, the HVCC controls the fan of TRB cooling system and the charging of low-voltage battery (from the TRB by the down DC/DC converter), considering its temperature.

A HPP with two electric machines may improve driving, energy and environmental performance characteristics in Toyota Prius cars, it consists of: ICE, electric starter-generator, traction electrical motor, and infinitely variable transmission (IVT) [50]. The ICE traction power is divided into two streams by the planetary mechanism. One of the output shafts of HGB is connected to an electric motor and wheels, and the other one – to a generator. Both motors (or each separately) rotate driving wheels by a speed-reducing drive (Fig. 4.26 a).

Figure 4.26 depicts: 1 – ICE with clutch; 2, 3 – electrical machines MG1, MG2; 4 – sun gear; 5 – pinion carrier; 6 – crown gear; 7 – front axle of differential chain drive; 8 – planetary power splitter; 9 – two-stage planetary speed reducing drive with two hydraulic brakes; 10 – valve case; 11 – mechanically driven oil pump; 12 – electrically driven oil pump.

Solenoid valves of hydraulic box control the working fluid pressure and switch the lines activating the drives of planetary gear brake mechanisms.

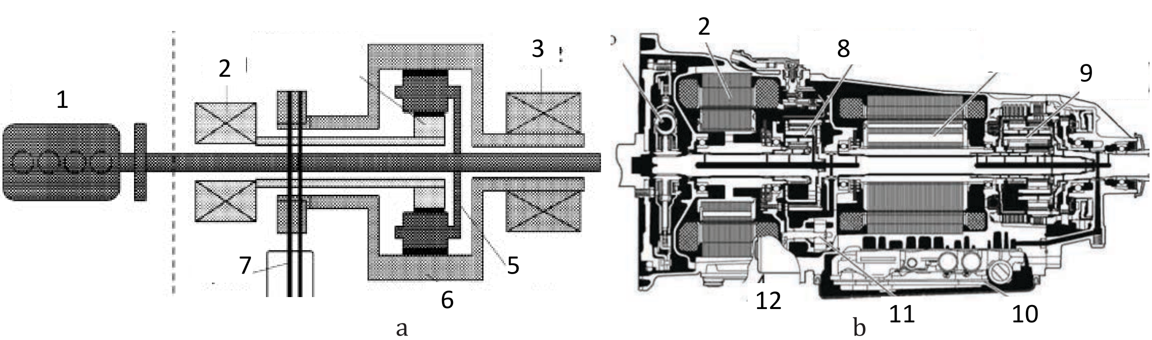


Fig. 4.26. Hybrid power plant with two electric machines: a – kinematic diagram of front axle drive; b – design of electromechanical transmission GS450h with rear axle drive.

A two-stage planetary speed reducing drive with hydroelectric control is used in transmission with a rear axle drive (Fig. 4.26 b). In this case, the automation, with the help of braking mechanisms, changes the transmission ratio of MG2 electric machine connection. Torque from MG2 is transmitted via the rear sun gear to the rear wheel drive.

The pressure in the hydraulic system of HGB is provided by a mechanically driven oil pump. The electrically driven pump duplicates the function of a mechanical pump when the ICE is not running. An additional electric pump makes it possible to circulate liquid to cool MG1 and MG2 according to the predetermined algorithm. The following sensors are applied to control the HGB operation: rotation frequency of input shaft; temperature of coolant liquid; pressure (actuation of drive gears).

A mode switch (selector) allows to choose one of the modes of transmission control, namely, dynamic, regular hybrid or winter hybrid. The purpose and operation of remaining components of HPP control system is similar to the system with one machine (see Fig. 4.25). The algorithms for functioning HPP (traction control, regenerative braking, power distribution) of a system with two reversible electric machines are designed in accordance with their statuses in the operating modes. At the same time, the control strategy is only to follow certain rules:

- the lower the car speed, the more the driving force from an electric motor is involved;
- while braking, the priority is given to the regenerative system;
- efficient use of the power capacity of power plants according to the redistribution of energy flows between them.

The implementation of power distribution algorithm in transport modes is as follows. A vehicle is driven off by MG2 electric motor **linkage**. If the TRB charge is not enough for this, then a MG1 starter is turned on to run an ICE. Further, MG1 operates in the generator mode from the ICE, charging a TRB if necessary. In the hybrid driving mode, the tractive effort from the ICE by planetary mechanisms is transmitted partly to a wheel drive and partly to MG1 for powering an engine MG2 to create additional tractive effort. While accelerating, a TRB supplies energy to MG2 for getting additional tractive effort with the ICE power. When braking downhill, the ICE is turned off, the kinetic energy of a car by the MG2 is accumulated into the TRB. In this case the MG2 engine is used for reversing.

In the THS II power plant, DC voltage converters are implemented in an

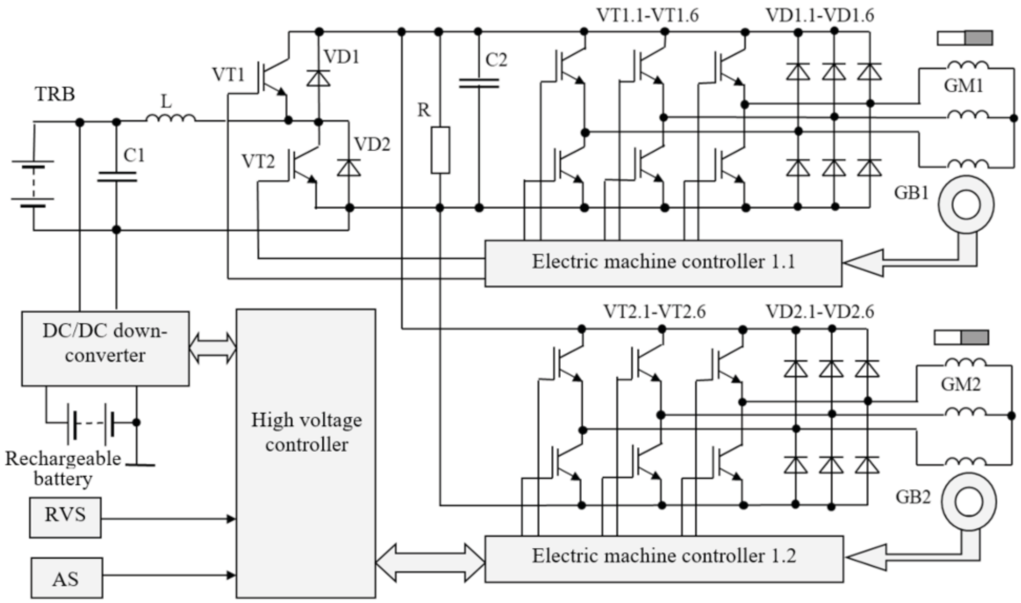


Fig. 4.27. Diagram of power circuits in an electric drive of HPP with two electric machines.

integrated power module (IPM) based on transistors of a combined structure (IGBT transistors) VT1, VT2 [50] (Fig. 4.27).

Both MG1 and MG2 are connected via the identical inverter and rectifier circuits. A HVCC controller receives commands from the operating controls (accelerator sensor AS, rotation velocity sensor RVS, etc.), setting the modes of electric machines (electrical machine controller EMC) and rechargeable battery charge.

In the mode of DC constant overvoltage, commuting the current in the reactor L by the VT2 key, when VT1 is locked under the TRB voltage of 288 V, leads to induce the voltage pulses of double amplitude on the VT1 collector. In addition, the C2 capacitor is charged up to the amplitude rate and controls the constant voltage of 650 V at the output of converter.

In the mode of DC constant undervoltage, commuting the reverse current in the reactor L with the VT1 key when VT2 is locked under the rectified generator voltage of 650 V, tends to induce the resistance of L reactor and a drop in average current value in the TRB charge circuit. At the same time, the C1 capacitor is charged up to the constant voltage rate of 288 V at TRB terminals.

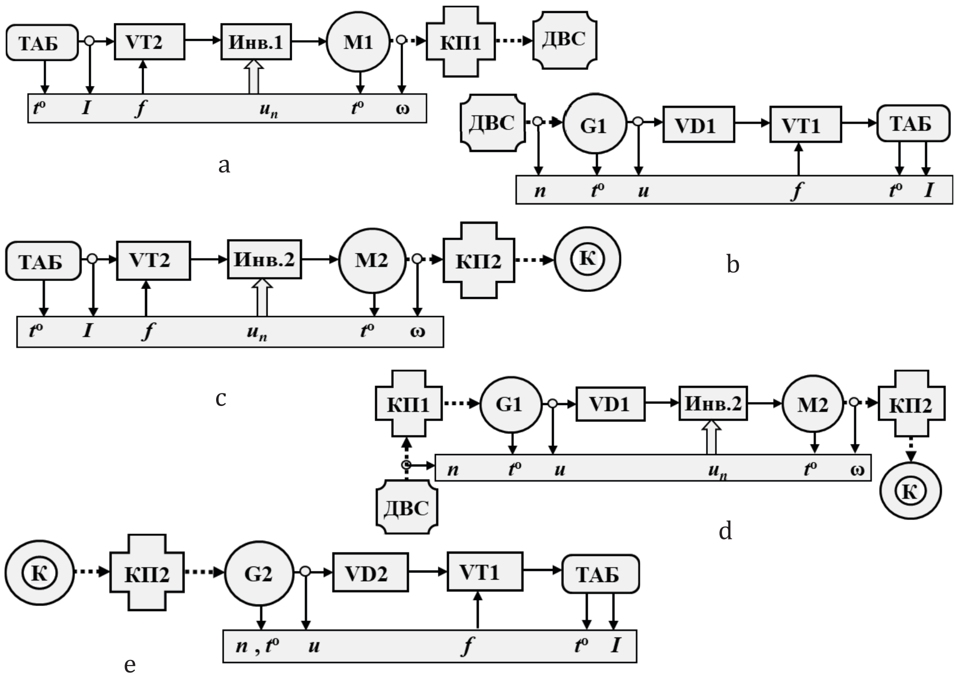


Fig. 4.28. Diagrams of power transmission circuits by HPP mode statuses: a – ICE starting; b – high-voltage battery charging from an ICE; c – electric drive; d – series operation of ICE and electric motor; e – recuperative battery charging.

The information bus passes the information on the power state of voltage sources (TRB charging, generator speed rotation) and the specified driving mode of a car (accelerator pedal position and mode selector) from the HVCC controller to the electric machines controller of EMC. Corresponding pulses for controlling the keys of DC voltage converters and inverters are received by signal circuits from the EMC.

Rear wheels and a MG2 electric machine are connected by the cardan shaft and are not separated by the clutch. To break the transmission of torque to rear wheels when the selector is in neutral, all power transistors controlling the operation of electric cars are turned off. As a result, MG1 and MG2 do not induce braking response to transmission.

Analyzing the HPP functioning in the transport modes, it is possible to arrange units creating the power transmission circuits by five transport modes of HPP operation (Fig. 4.28).

Diagrams in Fig. 4.28 indicate mechanical segments of power transmission: W – wheel, GB1/GB2 – speed reducing/increasing transmission. Electric circuits:

TRB – traction battery; VT – keys of DC voltage converters, VD – rectifiers; 'Inv.' – inverters. Measuring circuits of HVCC and EMC controllers receive the following information from the power circuit elements: n – revolution rate; t^o – unit temperature; u_n – inverter control pulses: u – output voltage; f – controlling pulse rate; I – average current value in the battery charging circuit; ω – the angular position of rotors in the electrical machine. Additionally, the information on the temperature of inverters and rectifiers can be used, which characterizes the operational and technical (indirectly) states of these elements.

4.5. Using a pneumatic drive

Compressed air machines are much cheaper than their electric or hydrogen equivalents with approximately equal performance parameters (sizes, speed, remaining mileage interval). In addition, compressed air cylinders have a longer service life than chemical accumulators of electricity.

The strong points of a pneumatic drive are the following [51]:

- eco-friendly (no harmful emissions, and oxygen is not burned);
- low cost of drive and 'fuel';
- ability to recharge a car at home;
- possibility of using an energy recuperator.

The key operational strength of pneumatic motors is to obtain the maximum torque and efficiency while starting off (no air consumption). While accelerating, these performance characteristics drop significantly.

The weak points of a pneumatic drive include:

- low efficiency (5–7 %) in comparison with an ICE (18–20 %);
- need for an external heat exchanger;
- low driving performance of pneumatic vehicles (remaining mileage interval, power);
- low density of energy carrier (specific energy intensity);
- pressure escaping in the structure of pneumatic equipment;

- large-sized balloon is required for air filling.

As a comparison, the air at a pressure of 30 MPa has an energy density of about 50 kWh per liter, and a conventional gasoline engine has the energy density of 9411 kWh per liter. That is, gasoline as a fuel is almost 200 times more efficient.

Designs of most rotary pneumatic machines are capable of operating as a motor and a compressor as well. Such reversible machines are usually called pneumatic units. Pneumatic motors are distinguished by the nature of working body motion: rotary rotation machines (motors); linear (pneumatic cylinders); reversal; special [52].

Rotary motors suitable for the MV drive are distinguished by general classification criteria (Fig. 4.29).

In mud pneumatic motors, mechanical work is performed as a result of compressed air expansion in the cylinders of piston machine (the potential energy of compressed air is used). In air turbine engines, the air flow affects the turbine blades (the kinetic energy is used). In the rotary structures of pneumatic motors, due to the partial overlap of cavity volume, a combined conversion of compressed air energy occurs.

Rotary (vane) pneumatic engine has a plate design of a rotor (1), eccentrically placed in the cylindrical cavity of the stator (2), which has the inlet (3) and outlet (4) ports (Fig. 4.30).

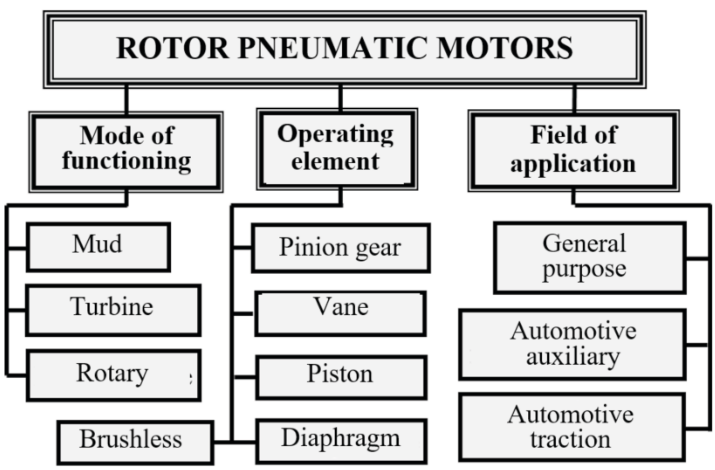


Fig. 4.29. Classification structure of pneumatic motors.

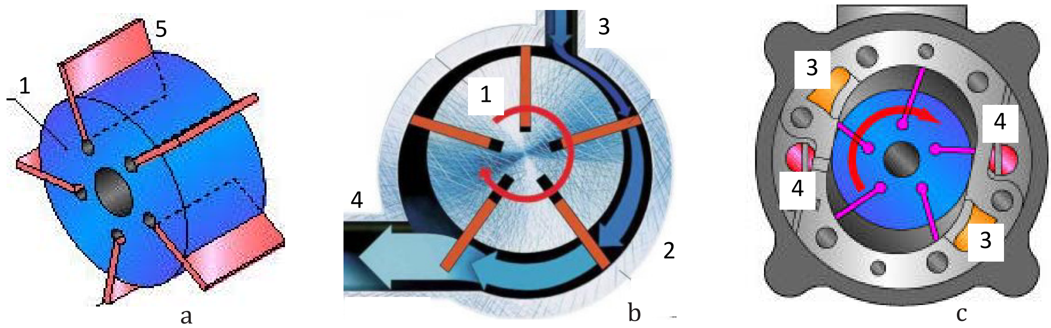


Fig. 4.30. Design of rotary motors of general purpose: a – rotor; b – single-chamber motor; c – two-chamber motor.

Plates (blades) (5) mounted on the cylindrical rotor (1) have a spring-loaded structure and form rotating chambers in the stator cavity, in which the pressurized air supplying via the inlet channel (3) expands. The pressing force of plates is determined by the rotor speed (centrifugal forces). To increase the power density in such motors, double expansion chambers with two inlet (3) and two outlet (4) ports are used (Fig. 4.30 c). Rotary motors can be manufactured in both reversible and non-reversible configurations. Reversible rotary units have equal flow cross-sections of inlet and outlet ports.

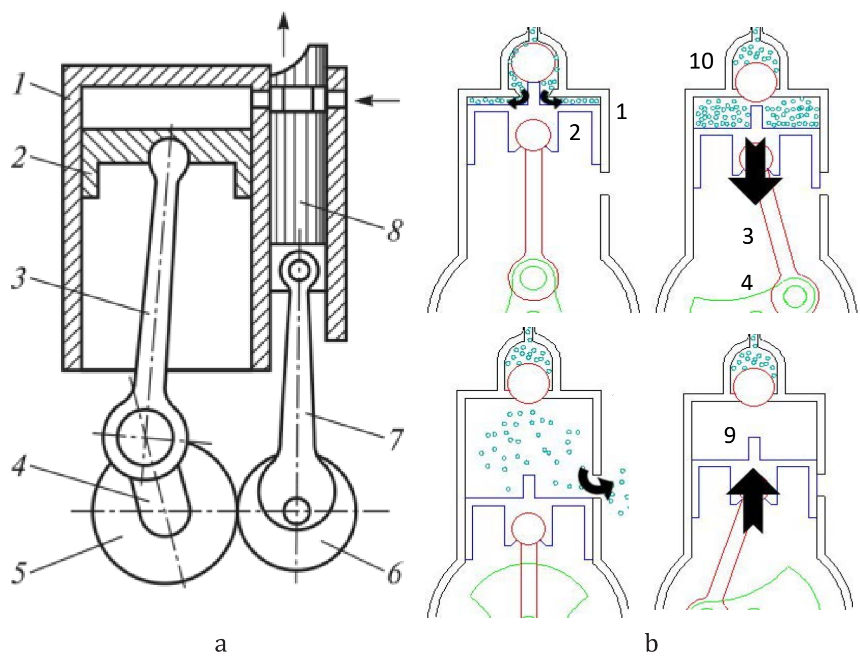


Fig. 4.31. Design of piston pneumatic motor: a – with a spool valve; b – with ball valve.

The main benefit of rotary motors is their relatively low weight, size and cost. The drawbacks include the lack of a secured working gap sealing and, as a result, a drop in efficiency at low rotational rate when starting a motor, as well as noise emission while operating and also relatively rapid plate wearing.

Piston pneumatic engines, similar to ICEs, have a cylinder-piston group (1), (2), a crank-and-rod mechanism (3), and a crankshaft (4) (Fig. 4.31).

The compressed air is supplied to the working compressed cylinder by the distributing spool valve motion (8), which is driven by the output shaft via the gears (5), (6) and a connecting rod (7) (Fig. 4.31 a). A distributor is designed in such a way that the cylinder cavity is uncoupled approximately at (5/8) of the full stroke piston length from the intake port. After cutting off the air, the piston moves due to the air expansion in the enclosed volume. With the return of the piston stroke, a spool valve is vented through the working cavity to the atmosphere. At the moment when the piston is at a certain distance from the end of the stroke, the spool valve closes the outlet port and as the piston moves further, the remaining air is compressed. Thus, the piston motor operates with partial expansion of compressed air and partial compression on back stroke.

In piston motors, air dispensers of two types are used, namely, spool valve and shaft valve. Multi-cylinder pneumatic motors are designed similarly to ICEs. Motors with double-acting pistons have the distinctive design where the compressed air is supplied to both sides of a piston.

In the second scheme, a piston (2) with the rod (9) forces on the ball valve (10), supplying the compressed air into the cylinder chamber (Fig. 4.31 b). Air expansion makes the piston move. After the air is released, the piston returns to its initial position, due to the energy of inertial mass of flywheel and crankshaft counterweights (as in the ICE).

The use of spring return instead of flywheel masses tends to lower the weight-size parameters of engine design. However, in this case, spring compression losses occur, and the rotation uniformity of output shaft drops. Piston PMs, in comparison with rotary ones, have lower air escaping, allow overloading, and changing the rate of cylinder chamber filling.

In the *diaphragm* PM, the compressed air is supplied to the working chambers by means of the distributor (1), and an elastic diaphragm (4), connected to the rod (3) with the spring (5) operates as a double-sided piston (Fig. 4.32).

Compressed air, passing the spool valve necks (2) into the working chamber, moves a coupling rod by pressing on the diaphragm (4). At the same time, the

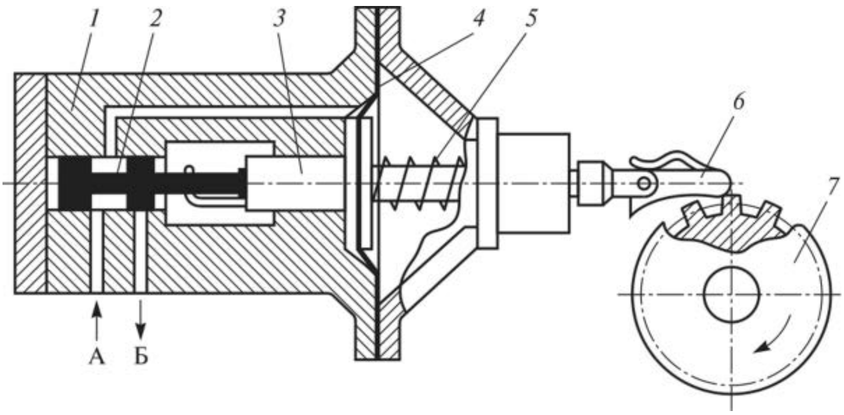


Fig. 4.32. Diaphragm of diaphragm pneumatic motor.

spool valve (2) moves and compresses the spring. At the end of the diaphragm stroke, the spool valve shuts off the air supply port 'A' and opens the outlet port 'B', the air exhaust occurs. The diaphragm is returned to its initial position by the spring (5), switching the spool valve to the compressed air supply. The cycle repeats, the reciprocating movement of rod (3) is converted into rotation by the hooking pushrod (6) and the tooth gear (7).

Diaphragm motors are characterized by high torque at low speed. They do not require high precision fit of movable elements and lubricants. The rate of diaphragm elasticity determines the valve rod stroke.

Gear PMs have lower power-to-weight ratio, higher compressed air escaping and lower efficiency. Propeller and turbine pneumatic motors are characterized by high rotation speeds with minimal torque and significant air flow under load.

Pneumatic motors of general purpose as well as compressors are widely used in range of tools. On a car, auxiliary pneumatic units are implied as compressors and pneumatic drives of different types and purposes. So, for example, pneumatic starters of various designs are used to start an ICE (Fig. 4.33).

To obtain sufficient torque on the drive gear, multi-stage and planetary speed-reducing drives operate in pneumatic starters. *Automotive traction motors* are presented as converted and of original designs.

To ensure the PM reverse and reversible modes of pneumatic unit, uncontrolled valves and valves with mechanical, pneumatic or electrical control are implemented in the supplying lines of pneumatic systems. At the same time, valves are also classified according to a number of common features:

a b c

Fig. 4.33. Exterior of pneumatic starters: a, b – rotary type; c – turbine type.

- purpose (shut-off, reverse, regulating, safety, switching, distributing);
- designs of locking and regulating element (ball, cone, disk, sloop valve);
- method of control (direct action, pilot, combined action, bi-stable);
- number of locking positions (one-way, two-way);
- number of connected inputs (from two to five).

For valves with electric control, the design of back mechanism includes for one of two initial (de-energized) states of shut-off and control element, i.e., normally closed or normally open. The examples of valve devices used in mechatronic systems are shown in Fig. 4.34.

In Fig. 4.34, 'P' is the power line (input); 'A' and 'B' are user lines (output); 'R' and 'S' are the reset lines.

In *direct-acting shut-off valves*, a piston plunger (traction relay armature) (1), under the spring return (2) and the pressure above the piston plunger, directly affects the shut off valve (3), blocking the output line 'A' (Fig. 4.34 a). When the voltage is supplied to the valve winding, a piston plunger is drawn into the coil body (4) and frees the port 'A' passage. The concussion spring (5) adjusts the shut off valve and minimizes shock loads of valve gear.

The advantage of a direct-acting valve is the simplicity of design and the ability to switch at zero pressure change. The drawback is the limited value of commuted pressure (determined by the electromagnetic unit power). This noted

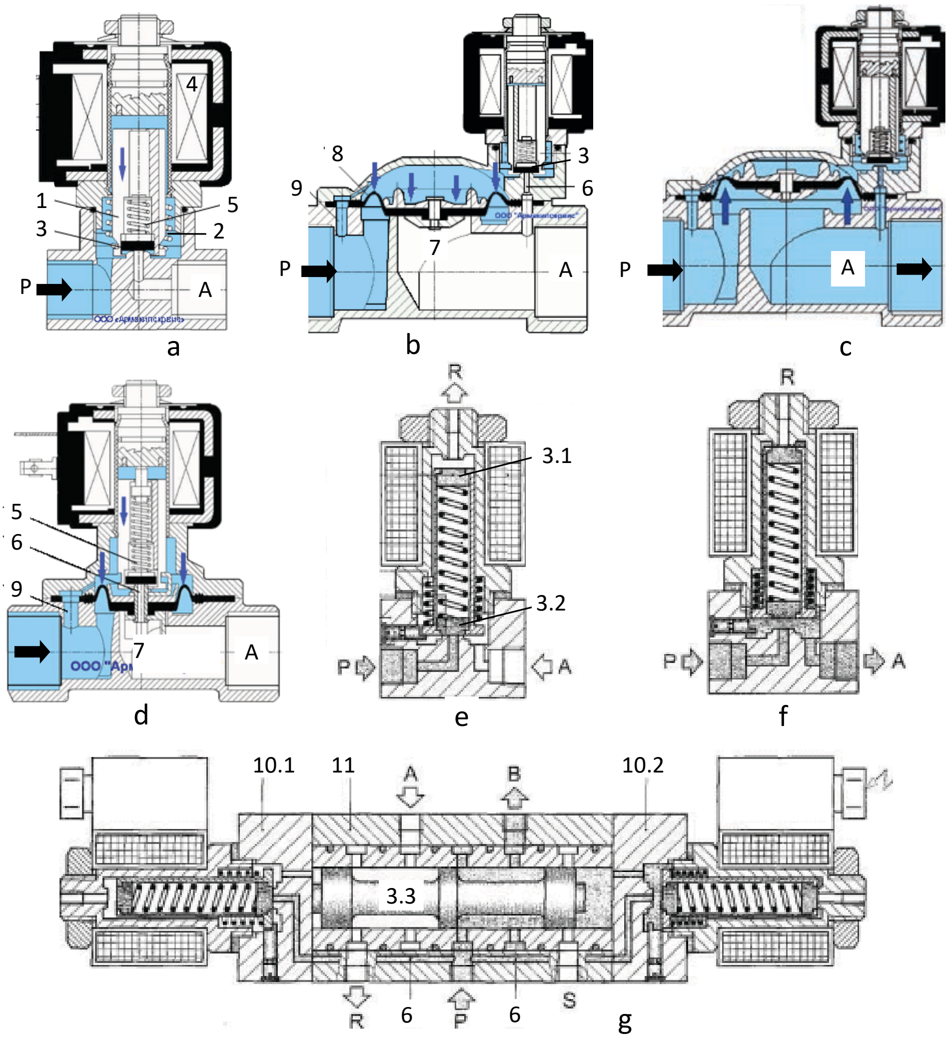


Fig. 4.34. Design of electrically operated pneumatic valve: a – of direct action normally closed; b – pilot normally closed; c – pilot normally open; d – combined action normally closed; e, f – two-way distributor with axial valves; g – bi-stable dispenser with slide valves.

weak point is overcome in indirect valves with a pilot (amplifying) port (Fig. 4.34 b, c).

In such design, the solenoid valve controls the pilot port (6), and the main line section is closed by a diaphragm or a piston valve of a larger area (7). In the initial state, the pilot port is closed by the shut off **valve** (3), the pressure above the diaphragm (8) by the bypass port (9) presses the valve (7) to the support saddle, blocking the use line ‘A’. When power is supplied to the winding, the **plunger,** when a shut off valve goes up, is opening the pilot port (6). At the same time, the pressure above the diaphragm (piston) drops, the resulting force acting on the diaphragm moves it up and opens the main line section (Fig. 4.34 b).

A *normally open pilot valve* has a similar design and works in the reverse sequence (Fig. 4.34 c). The differences lie in the configuration of back mechanism of shut off valve (retraction changes to push-out). Thus, valve switching is performed by the electromagnetic force and the pressure in the line, which makes it possible to significantly reduce the specific power of electromagnetic unit. Along with the possibility of commutating high pressures, for the pilot valve operation a certain pressure drop is required between lines ‘P’ and ‘A’.

A compromise solution for the pilot control is the design of a *combined-action valve*, when a piston plunger operates with both the shut off valve of pilot port and the diaphragm valve of the main bypass section (Fig. 4.34 d). When the shut off valve is closed, the pressure affects the diaphragm by the bypass port (6), forcing the main section valve to the support saddle. When the pilot port (6) is opened, the diaphragm valve (7) moves up and the pressure is supplied to line ‘A’.

The *distributor with axial valves* has two shut off valves (3.1 and 3.2), overlapping lines ‘P’ and ‘R’ under the electromagnetic drive control (Fig. 4.34 e, f). In the lower armature position (plunger), when a coil is de-energized, line ‘A’ interconnects with line ‘R’ by slots on the armature outer surface. When the voltage is supplied to the coil, the armature overcomes the spring return force and moves up. At the same time, port ‘R’ is closed and port ‘P’ is opened for supplying compressed air to line ‘A’.

A *bi-stable distributive valve* can be in one of two possible switching positions any given duration, since there are no elements (return springs) in its design that clearly determine the position of shut-off and control elements. For actuating bi-stable mechanisms, it is enough to send a short-term (pulse) control signal (low power consumption while controlling).

A *bi-stable distributive valve with the electro pneumatic control* is a combination of two pilot electrically controlled distributive valves (10) and the basic distributive valve (11) with two-way pneumatic control (Fig. 4.34 g). Compressed air is supplied to the pilot valves by the ports (6), which are connected to supply line ‘P’. When voltage is supplied to one of the coils, the corresponding pilot distributive valve is actuated feeding the compressed air into the pneumatic compressed cylinder of sloop valve (3.3). The spool valve motion leads to the pressure commutation along the lines.

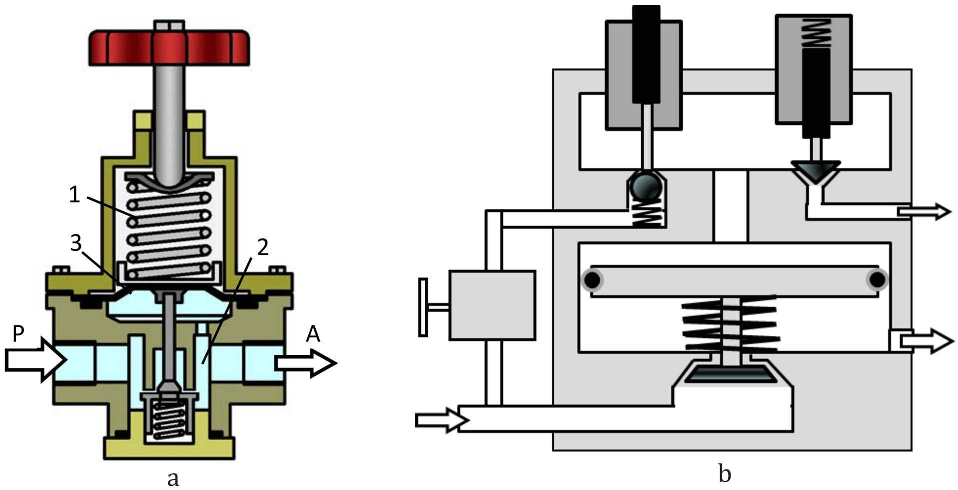


Fig. 4.35. Pressure control units: a – adjuster; b – modulator.

Torque and pneumatic motor revolution are controlled by changing the air consumption or its pressure at the inlet in the expansion chambers. These parameters are changed by adjusters (speed-reducing drives) and pressure modulators (Fig. 4.35).

A *pressure relief valve* (PRV) is mounted after the compressed air cylinder and allows to set and control a setting pressure value in the supply line at the mechanical level. In this case, the reference pressure value at the PRV output is set by the spring (1), and the pressure feedback is obtained by means of the port (2) and the diaphragm (3) (Fig. 4.35 a).

In the mechatronic system of pneumatic drive, the key control PM element is the pressure modulator (PM) with electro valves (Fig. 4.35 b).

A pressure modulator works in the following manner. The inlet pressure (port 'P') enters use line 'A' by the proportional valve (4), its position is determined by the pressure ratio above the piston (5) and below the piston (6), taking into account the spring return action. To change the pressure above the piston, the control pressure is supplied to the chamber with the inlet (7) and the outlet (8) and normally closed solenoid valves, its maximum value is set by PRV valve.

Pressure boost in line 'A' occurs when valve 7 is open (pulse width modulation (PWM) control), and valve 8 is closed. Controlling the pressure in port 5 and line 'A' goes on when valves 7 and 8 are closed. The pressure in port 5 and line 'A' drops when valve 7 is closed by releasing pressure into the atmosphere by valve 8 (PWM control).

Industrial samples of pneumatic systems ABS and EBS (brake pressure modulator, proportional accelerator valve, double solenoid valve, rear axle modulator, full-trailer brake control valve) can be used as a pressure modulator.

Let us consider options for constructing pneumatic drives according to the PM design complexity, starting with the conversion of piston motor ICEs and ending with configurations of the original design.

Pneumatic motor with a mechanical air distributor, known as a Nikolai Pustynsky's engine, is created on the basis of an ICE operating on a two-stroke cycle, with the design modification of gas distribution mechanism (GDM). An air distributor via cylinders has the form of a sealed cylindrical housing, where a rotor with air ports rotates inside without zero clearance. Rotor inlet ports are in communication with the pressure source, and the outlet ports are vented to the atmosphere. The stator cavity of dispenser is connected with the inlet and outlet ports of ICE cylinders by means of the holes in the distributor housing. The holes of air ports on the generating surfaces of rotor and PE stator cavity are phased in such a way that when a rotor rotates, the gas distribution of air flows occurs according to the working algorithm. The distributor rotor is driven in the same way as camshaft drives in a base ICE. Thus, the ICE conversion involves replacing a standard camshaft drive with the original air distributor.

A pneumatic drive consists of: an engine with the inlet (6) and the outlet (7) ports of cylinders; a cylinder with compressed air and a filling valve (1); a pressure regulator (2); a pressure modulator (PM); control system components (Fig. 4.36).

Engine operating mode is determined by its torque load and the air pressure value at the inlet to the distributive valve. The electronic part of mechatronic system controls pressure in vehicle operating modes by the electronic accelerator pedal (accelerator sensor) and the pressure modulator (PM). Modulators of pneumatic brake systems of industrial designs with two solenoid valves (4 and 5) can be used as a pressure regulating valve. The pressure sensors (P) control the operating state of pneumatic transmission in the air-pressure lines. The engine speed rate is controlled by the crankshaft speed sensor (*n)*, the vehicle speed is controlled by the sensor (*V)*. On the basis of signals from the sensors, the electronic control unit (ECU) generates signals to control pressure modulator valves. An electric pressure cut-off valve (3) turns off the pressure supply in the automatic mode.

A similar approach, but with the use of electrically controlled gas distribution valves, can additionally provide for a brake energy recovery circuit (Fig. 4.37).

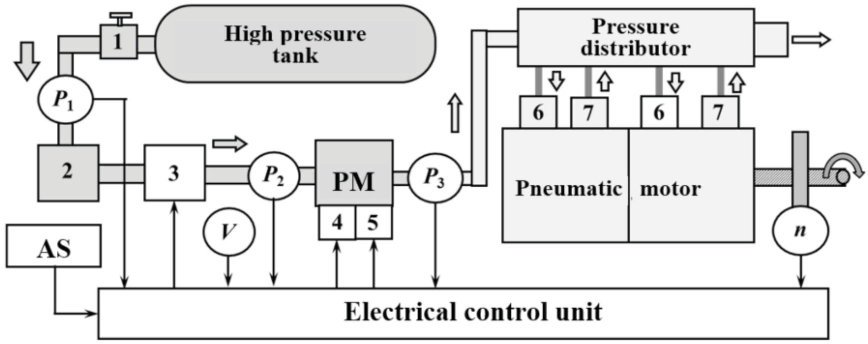


Fig. 4.36. Pneumatic drive circuit with the converted motor.

In the figure, solid arrows indicate the direction of control system signals (sensors and actuators), wide arrows indicate the gas flows. In the engine of this design, there is no rigid mechanical connection between the crankshaft (pistons) and GDM valves. The ECU (10), according to the information on the angle of crankshaft drive (17), controls the cycle of switching GDM valves (8 and 9) in the motor and compressor modes. Moreover, to start and control the working process in the motor, it is necessary to know the piston position in the cylinders not only while crankshaft is rotating, but also in the initial static state. For this purpose, encoders are used as a sensor for identifying the angle and the revolution rate (17) (see Fig. 4.18).

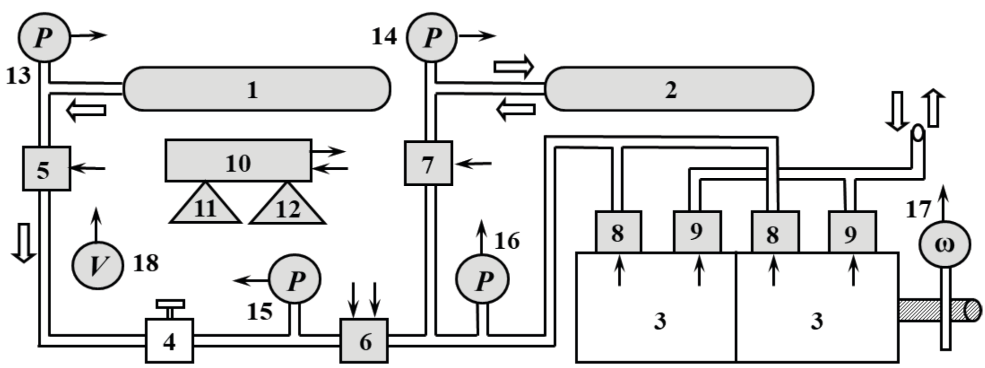


Fig. 4.37. Pneumatic transmission diagram with the recuperation circuit: 1 – high pressure cylinder; 2 – pneumatic batteries; 3 – cylinders of pneumatic unit; 4 – pressure reducing valve; 5, 7 – cut-off electrovalves; 6 – pressure modulator; 8, 9 – inlet and outlet electric GDM valves; 10 – ECU; 11, 12 – accelerator and brake pedals; 13–16 – pressure sensors; 17 – crankshaft sensor; 18 – vehicle speed sensor.

When an actuator operates as a motor with the cylinder (1), valve 5 is open and valve 7 is closed. According to the signal from the accelerator pedal sensor 11, the ECU regulates the pressure in the line of intake valves (8) by means of modulator valves (6). At the same time, the specified mode is controlled by digital proportional-integral-derivative (PID) controllers with the pressure feedback (sensor 16) and the revolution rate (sensor 17). Stopping the engine after reducing the pressure by regulator 6 is duplicated by the overlapping valve 5.

A signal from the brake pedal sensor 12 changes the drive into a compressor. In this case, regulator 6 cuts off the air supply from cylinder 1 and opens valve 7. The ECU switches the GDM valve cycle on the grounding of signals from sensors 17 and 12. While reducing the compressor pressure (sensor 16), the charging port of battery 2 is closed by valve 7. The ability to continue the battery 2 charging in the braking mode is estimated by the ECU according to the pressure ratio of 14 and 16.

Drive start-off and acceleration with the help of the accumulator energy 2 is limited by the accumulated pressure 14 and is activated by accelerator 11 (valve 7 is opened).

The weak point of converted engines with the direct air supply is the cooling process when the air in cylinders expands in the power stroke.

Guy Nègre's piston motors contain the thermodynamic effects to boost the gas energy due to their heating while compressing [53]. The design of the first generation Nègre's motor is based on two equivalent ICE cylinders; one of them operates as a compressor, and the other one works like a driving machine. The intermediate chamber with controlled valves is created between the gas ports of cylinders to expand the compressed air supplying from the outside. Strokes of Nègre's pneumatic motor operation, in contrast to an ICE, are divided into two cylinders. In the compression cylinder 1, the inlet and compression processes occur, and in the decompression cylinder 2, the processes of working stroke and release go on (Fig. 4.38 a).

The line with the air injector 5 is connected from the high-pressure cylinder 4 to the intermediate chamber cavity 3. The first two cycles are fullfilled in the usual way (Fig. 4.38 a, b). At the end of the compression stroke, the air charge by valve 8 enters the intermediate chamber under compression pressure while heat releasing (Fig. 4.38 c). When cylinder piston 1 is at top dead centre, chamber 3 is closed and the air is injected into it out of the cylinder under high pressure, resulting in both the pressure and temperature rise (Fig. 4.38 c). Further, the idling occurs when valves 6 and 7 are open (not shown in the figure). After injector 5 is closed, the piston is moved under the pressure of valve 9 expanding in the charge cylinder of chamber 3 (Fig. 4.38 d). In the reverse stroke of cylinder piston 2, the air is released into the atmosphere by valve 7 (Fig. 4.38 a). After that, the considered five strokes of working cycle with three stages of compressed air expansion are repeated.

The actuation of intermediate chamber valves and the air injector is adjusted under electronic control for the working process of Nègre's engine in operating modes.

Using the cylinders of equal volume in the first generation of Nègre's engines makes it possible to cut the converting ICE cost while their manufacturing. Along with this advantage, the key drawback is in the engine cooling and the energy loss due to the air expansion. However, compared to pneumatic motors, where the air is supplied directly to cylinders, the three-stage expansion lowers the negative cooling effect.

The second generation of Nègre's engines includes a pair of cylinders with different volumes in one block (Fig. 4.39).

The working cycle of pneumatic motor is as follows. Compressed air from cylinder 1, by valves 2, 4 and the inner cavity of intermediate chamber 3 enters a small cylinder chamber 5 and while expanding it converts the pressure energy

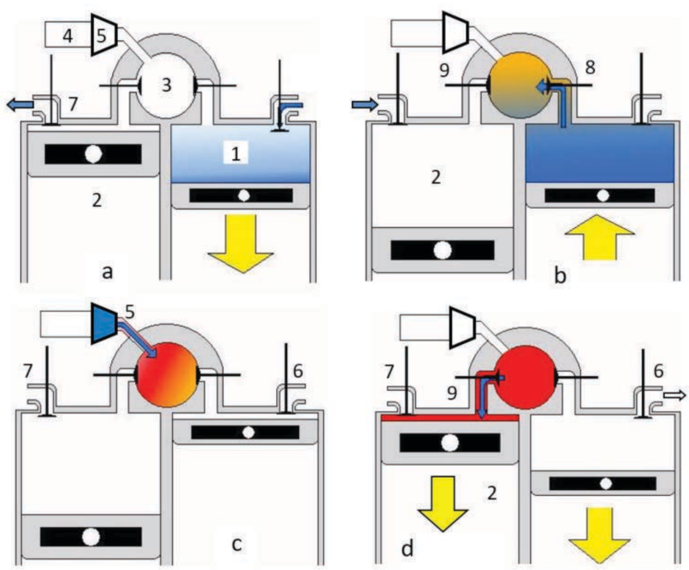


Fig. 4.38. Operation process of Guy Nègre's pneumatic motor: a – intake/exhaust phase; b – compression phase; c – expansion phase; d – working stroke.

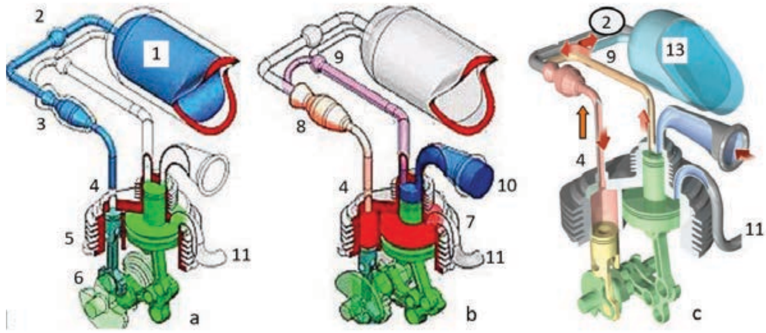


Fig. 4.39. Working process in the second generation of Guy Nègre's engine: a – air inlet; b – air charge expansion; c – air circulation in the compressor mode.

into the crankshaft rotation 6 (Fig. 4.39 a). Then, valves 2 and 4 close, the air in cylinder 6 is compressed by the piston up to a pressure of 20 kgf/cm^2 with its diabetic heating up to a temperature of 400 ° C. Then valve 4 opens and the hot air charge expands and goes into the chamber of the second (large) cylinder 7 via the heater outer cavity 8 and the switching valve 10 (Fig. 4.39 b). At this point, the compression stroke of atmospheric air rate that has previously entered the inlet port 10 ends in a larger cylinder. The obtained air mixture heats up and expands, increasing the pressure in the large cylinder, the piston moves and rotates the crankshaft (working stroke). In this case, the exhaust air is released into the atmosphere by the exhaust pipe 11.

The second generation of motors has the original design of crank-and-rod mechanism (7), which allows the piston of decompression cylinder to pass the top dead point with delay and the subsequent downward acceleration while the steady crankshaft rotates. At the same time, the connecting rod system allows the piston to be kept at top dead point in the 70° cycle sector and provides more efficient cylinder filling with air and the energy use while expanding.

The pneumatic unit is reversible, so that the cylinder can be refilled from the auxiliary electric motor. In this case, the exhaust pipe (12) and the valve (2) are overlapped, the valve (4) remains open, and the switching valve (10) commutes the air supply to the cylinder by the filling line in accordance with the compression strokes of each cylinder (Fig. 4.39 c). To cut the pressure loss caused by the cylinder cooling, it can be thermostated using the outer shell (13) with the heated air.

Thus, the second generation design of Nègre's engines drops two of the three expansion stages and, therefore, it is simpler and more reliable. However, the manufacturing costs are higher in comparison with the first generation engine due to the originality of the motor main units. The considered algorithms for functioning the second generation of Nègre's engine are carried out with the help of the electrovalves under electronic control.

The initial project of MDI (Motor Development International) with a Guy Nègre's engine could run not only on compressed air, but also on natural gas, gasoline and diesel. In further developments, in order to increase the drive power, the engine designs of this class began to be assembled into multi-chamber units, built according to the scheme of a symmetrical piston V-shaped engine with the cylinder alignment of 180 ° (Fig. 4.40).

At present, on the basis of Nègre's engines, manufacturers of foreign companies have created small-class pneumatic vehicles of various purposes: a multi-purpose family – Family, a pickup and a minivan – Multi, a taxi – City, a city – Mini (Fig. 4.41).

A two-seater Air Pod weighing 220 kg, equipped with a 4 kW pneumatic motor, is capable of accelerating up to 75 km/h (Fig. 4.41 a) and has a travel reserve up to 200 km. The machine is controlled by a joystick. To ensure smooth operation and optimize the energy consumption, the Air Pod motor uses a simple electromagnetic distribution system that controls the air flow into the motor. In addition, the system of supplying the air into the motor is equipped with a dynamic speed reducing drive of variable volume.

City car Mini CAT (Compressed Air Technology) with a plastic body weighing about 550 kg, is designed to carry three people at a speed of up to 110 km/h (Fig. 4.41 b). The 800 cc four-cylinder motor with the 30 MPa cylinder develops the power of about 18 kW. at 4000 min-1 and provides a mileage of 150 km. Exhaust cold air while driving can be used in the air conditioning system. The motor is also designed to run on conventional fuel with the switch over on-demand.

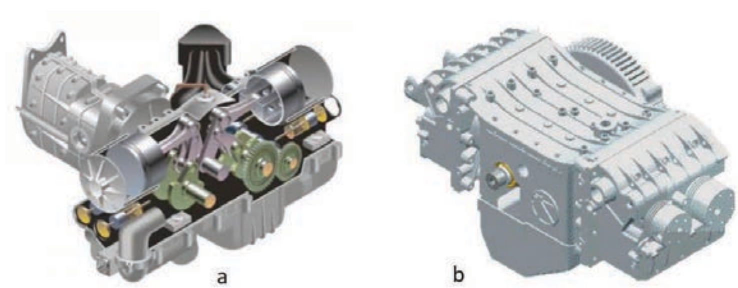


Fig. 4.40. Designs of Guy Nègre's pneumatic motor: a – with four chambers; b – with eight chambers.

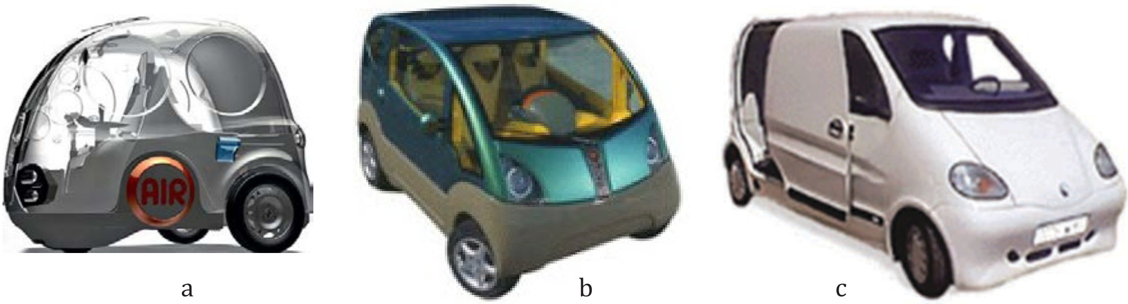


Fig. 4.41. Pneumatic vehicles with a Guy Nègre's engine: a – three-wheeled MDI Air Pod; b – triple Tata Mini CAT; c – cargo MDI Multicar.

Cargo single Multicar speeds up to 100 km/h. One filling is enough for 200 km run. It is able to ship to 300 kg of cargo while its own weight makes 210 kg (Fig. 4.41 c).

Series One Cat pneumatic car contains up to 300 liters of compressed air under a pressure of 30 MPa in the compressed cylinders, speeds up over 64 km/h and provides a range of up to 210 km. The cargo model of One Cat with a weight of 210 kg is able to ship up to 300 kg of cargo. The braking system in the cars of this class recuperats up to 13 % of the compressed air energy.

Due to the low specific storage density, pneumatic vehicles have found wide application as in-plant, airfield and municipal transport for public service.

The unconventional combined pneumatic motor Angelo Di Pietro has the advanced features of rotary and piston designs, as it has several expansion chambers (1) and separated by blades (2) [54] (Fig. 4.42 a).

Stator blades, spring-loaded (3) on the axis of their unit, can be placed into the slotted openings, while ensuring a firm adherence to the rotor ring (4). The rotor ring (4), clamped from the outside by blades (2), mounts inside on special rollers (5), fixed by their axes on the shaft rotor (6). Air pressure is sequentially supplied to the chambers formed by stator petals, causing the volumetric chamber expansion, which leads to the eccentric rotation of the rotor ring (4) and the axial rotation of rotor shaft (6).

Air is supplied and discharged to the chambers by the mechanical distributor. The distributor cage (7) with nine inlet and four outlet ports is mounted in the stator back cover (8) (Fig. 4.42 c). A cone-shaped distributor rotor, placed in the cage (7), is fixed to the rotor shaft (6) and sealed by the shut-off cover with the air supply connection. When the motor rotates, the container under the shut-

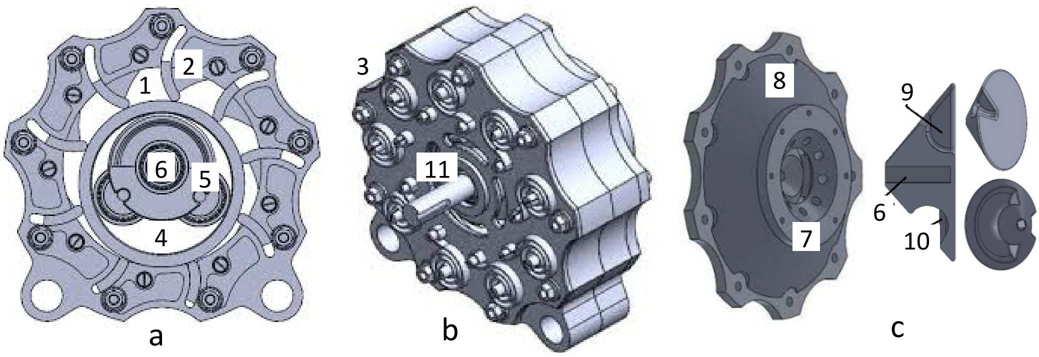


Fig. 4.42. Design of pneumatic motor Angelo Di Pietro with nine chambers: a – view without covers; b – assembled, front view; c – assembling of air distributor in the rear cover.

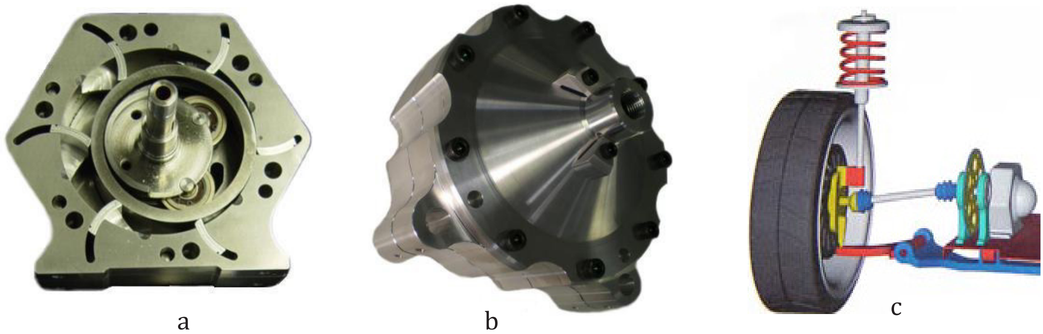
off cover is sequentially connected to the inlet ports by the dispensing cavity of distributor rotor (9). The air, in turn, is released from the chambers (1) by the outlet cavity of rotor (10) and the outlet ports of the distributor holder, which are connected to the ventilation openings of the motor front cover (11) by the inner cavity of rotor ring (4). Thus, the series change in the chamber volume (1) rotates a motor rotor under air pressure. At the same time, the speed and torque of the motor are easily controlled by adjusting the volume or the pressure of air supplying to the motor. These functions are performed electronically under control of pressure modulators.

The Angelo Di Pietro engine has several advantages over other pneumatic motor designs:

- light weight and design simplicity;
- no noise and vibrations;
- control of the rotation rate and torque;
- can be mounted directly on the motor vehicle wheels.

Due to the listed advantages, in pneumatic vehicles built on the basis of Di Pietro engine, there is practically no need for a transmission. At the same time, the driving characteristics of a vehicle are determined by the number of active motor chambers and the pneumatic potential of cylinders (Fig. 4.43).

For example, let us consider the characteristics of pneumatic vehicles with Di



a b c

Fig. 4.43. Angelo Di Pietro's pneumatic motors: a – with six chambers; b – with nine chambers; c – as part of a power suspension

Pietro's pneumatic motors of industrial designs for various purposes [55] (Fig. 4.44).

The O2 Pursuit motorcycle has an 11 kg engine, powered by air from the driving cylinder (Fig. 4.44 a). Remaining mileage interval makes up to 100 km and the maximum speed is up to 140 km/h.

The Gator cargo carrier designed by the Engineair company is the first compressed air car in Australia to enter the commercial market (Fig. 4.44 b). Its load-carrying capacity is up to 500 kg, the volume of air cylinders is equal to 105 liters, the mileage at one filling is up to 16 km. Notably, its refueling takes a few minutes. Such cars for golf clubs speeds up to 50 km/h, and under normal operating conditions, one air refueling provides transporting within an hour.

Note that the power units for driving a car by the energy of compressed air, according to the operation principle, can be pneumo-mechanical or pneumo-hydraulic systems. In the second case, liquid is used as a transmission link in the system. This solution results in eliminating the loss in pressure energy due to escaping and parasitic volumes of supply pipelines (to increase the drive efficiency).

A pneumo-hydraulic drive consists of a reversible hydraulic machine (hydraulic unit), which, in the pump mode, pumps liquid from the reservoir to the pressure accumulator with gas pressure shock absorber, thereby providing the boost in the compressed gas energy. In the motor mode, torque on the shaft in the hydraulic unit is created by the pressure of liquid pressed out of the hydraulic accumulator by the expanding gas in the pressure shock absorber. The mode of the hydraulic unit is switched by electric valves under electronic control. In this case, information from the pressure sensors in the *gas pressure shock absorber*

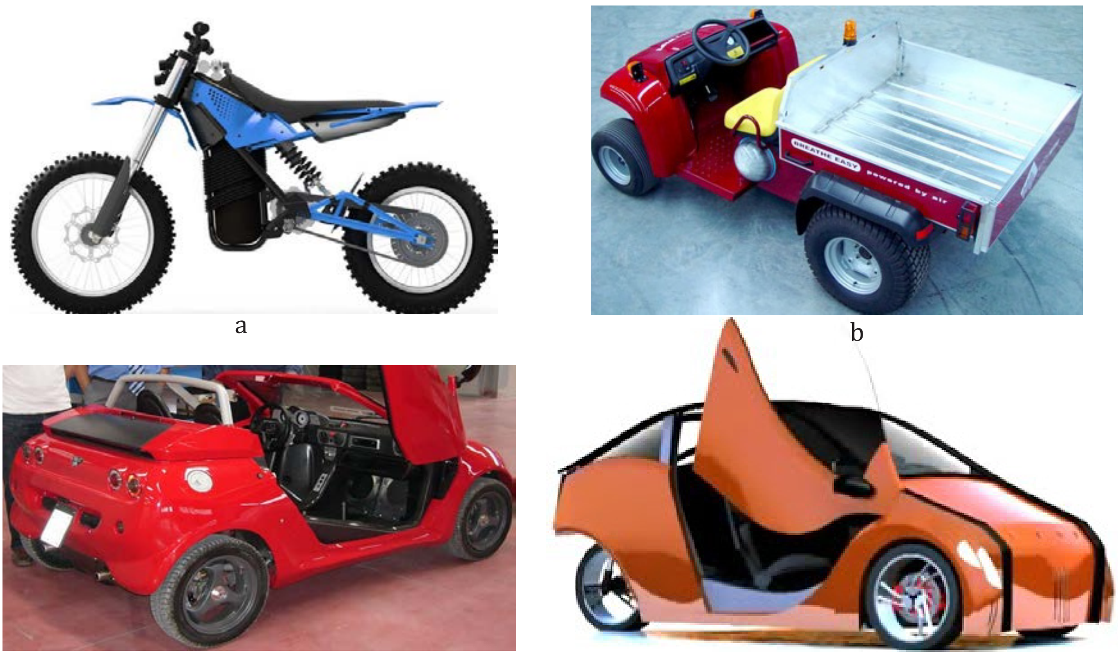


Fig. 4.44. Engineair's pneumatic transport: a – O2 Pursuit motorcycle; b – Gator's cargo conveyor; c – two-seater car; d – triple **conceptual prototype.**

and the level (volume) of the working fluid in the compressed air reservoir is used.

Thus, the energy resources of hydraulic system are limited by the volume of working fluid and the gas pressure in the hydro-accumulator **supporting** chamber. Therefore, the use of hydraulics is acceptable only for braking energy recovery systems or while working as part of hybrid power plants.

4.6. Compoponents of pneumatic hybrid power plants

Let us consider the main approaches for creating pneumatic hybrids connected with the use of:

- air-hybrid motors;
- converted ICEs as pneumatic engines in the HPP;
- pneumatic motors of general purpose or original designs in the HPP;

- hydraulic transmissions as part of pneumatic HPPs.

Air-hybrid engines (AHE) are built on the basis of an ICE with an additional air-charge gas distribution system [56]. There are AHE schemes with alternate (series) and simultaneous (parallel) conversion of thermal fuel energy and the compressed air energy in engine cylinders.

Motors designed by the series technology have a compressor to initially generate the pressure in the compressed cylinder. The compressed air cylinder can be filled by the external or built-in compressor. The built-in compressor is driven from the AHE in the ICE mode or from the car wheels in the regenerative braking mode. The compressed air stored in the compressed cylinder is pressurized in the cylinders for the efficient fuel combustion. The possibility of pneumatic restarting makes it possible to cut down the engine idling and thereby improve the fuel, economic and environmental performance of a vehicle.

In the conceptual prototypes of AHE with the parallel operation of fuel (1) and pneumatic (2) cylinders, equal cylinders or cylinders of an adapted volume are used (Fig. 4.45).

Initially, AHE air cylinders operate as a compressor driven by fuel cylinders. Excess pressure, in this case, is accumulated in the compressed cylinder (3) with its subsequent use or for pressurizing the air into fuel cylinders (1) or for the pneumatic drive from the air cylinders (2) as a compressed air motor. At the same time, the operating mode of the pneumatic accumulator is controlled by the distribution valves (4). Alternatively, one of the cylinders of a 4 or 3 cylinder ICEs operating in a two or four stroke cycle (AHE technology – Air Hybrid Engine) can be used for a pneumatic drive. If this occurs, the unit is capable of operating in several modes:

- power output in the ICE mode;
- power output in the pneumatic motor mode;
- joint work of air and fuel cylinders;
- pressurization of air into the fuel cylinders;
- charging the air cylinder by the power of fuel cylinders;
- charging the air cylinder while regenerative braking;
- restarting the ICE by compressed air;

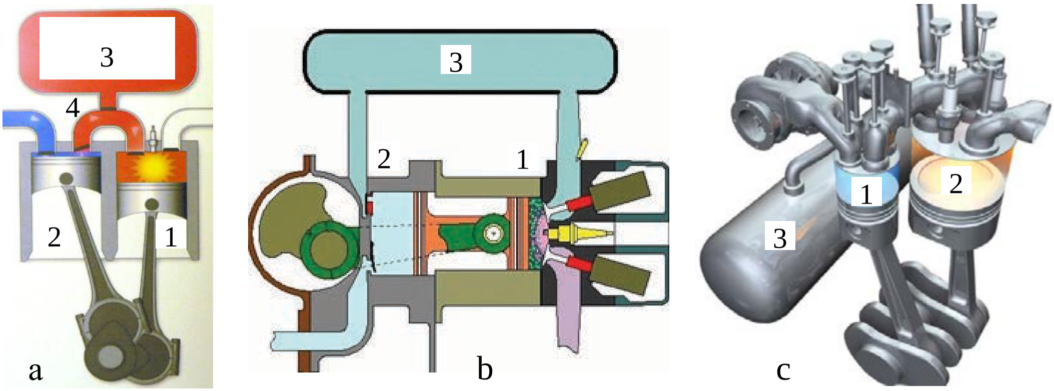


Fig.. 4.45. Designs of air-hybrid engines with parallel energy conversion: a – energy conversion scheme; b – scheme with cylinders of equal volume; c – design with cylinders of adapted volume.

- recuperation of thermal engine power in the compressed cylinder.

Thus, by analogy with electric hybrids, sequential AHE, depending on the implemented modes, can be classified as microhybrids or full non-chargeable hybrids.

AHE control algorithms are implemented using the electronic control system (ECS). In these circumstances, the monitoring system must control the operating parameters inherent in both the pneumatic drive and the ICE. The complete list of sensors and the composition of ECS actuators is determined by the number of functions and the flexibility of the HPP control algorithm.

Pneumatic hybrid power plants, as in the case with electric hybrids, are classified according to general characteristics (Fig. 4.46).

In the general case, a HPP with a pneumatic drive, except an ICE, can include various power plants:

- pneumatic unit (PU), functioning as a motor and a compressor, or separate functional units;
- pneumatic low pressure accumulator (PA);
- pneumatic high pressure tank (PT);

- heat exchanger (HE);
- electric machine functioning as a motor and a generator (electrical machine) or separate functional units;
- traction rechargeable battery (TRB).

The energy exchange between the listed units of the HPP and the 'K' car wheels can be presented by generalized structure (Fig. 4.47).

The power distributed between the power plants has a different nature of energy, parameterized by temperature t^o, current I and pressure P. The mechanical power of rotation at a given frequency is determined by torque M. Bidirectional arrows in the diagram in Fig. 4.47 indicate the possibility of implementing the energy recovery systems. The power plants and car wheels are connected mechanically by the transmission.

From the perspective of *energy autonomy*, hybrids with a pneumatic drive, similar to electric ones, can be divided into two types – externally charged hybrids (ECHs) and non-chargeable hybrids (NHRs).

The choice of the composition of the pneumatic HPP with regard to the autonomy is based on the energy ratio of the pneumatic drive for a given driving cycle. At the same time, a variation of the power plant parameters in the pneumatic unit (pressure level, accumulator volume, efficiency of pneumatic unit), on the one hand, controls the dynamics of the vehicle movement due to the energy of the alternative engine, and on the other hand, overcomes the starting

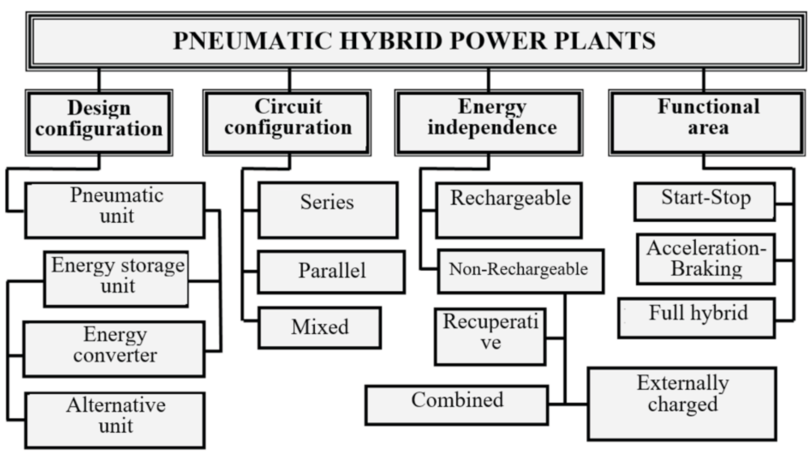


Fig. 4.46. Classification structure of pneumatic HPPs.

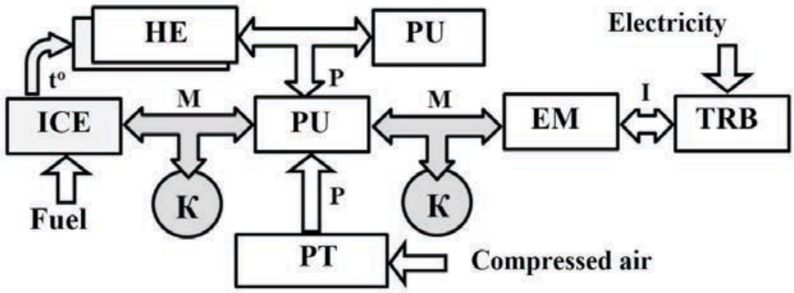


Fig. 4.47. Generalized structure of energy exchange in the HPP with pneumatic unit

moment and provides a given cycle movement by the pneumatic drive energy.

According to the generalized structure, it is possible to build various HPP compositions using PU, where alternative engines operate as the main or auxiliary power plant. The engine utilization rate is usually measured by the amount of energy consumed in the driving cycle. So, various percent performance characteristics are considered as engine power: miles of operation in a given mode; amount and cost of the energy carrier spent on the mileage.

Thus, the rationality of using one of HPP compositions is defined by the operating conditions and vehicle movement modes. So, for example, an electro-pneumatic composition has the strong points of a pneumatic drive when starting and accelerating a car (minimum pressure consumption and maximum torque) and an electric drive at speed rate (minimum current consumption) with significant mileage (high energy capacity of the battery). At the same time, it must be understood that the charged electric drive operates as the main one, and the recuperative pneumatic drive works as an auxiliary one. A three-component HPP can also be considered as a two-stage pneumatic system for starting a powerful ICE, where the primary pressure is compressed in the pneumatic rechargeable battery due to the electric drive energy [57].

The functionality of the HPP is determined by the purpose of the pneumatic drive system (start-stop, acceleration-bracking, full hybrid) and the flexibility of control algorithms (list of sensors, method of control, method of implementing control actions).

Start-Stop systems (SSs) are implemented in the AHE structure (see above) or as pneumatic starters for automatically restarting an ICE (see Fig. 4.33).

Acceleration-Braking systems (ABs) function by recuperating the energy of compressed air in the pneumatic accumulator when vehicle wheels are braked and then accelerated. For this purpose, the accumulator circulates as a gas spring.

The classification difference between *full hybrids* (FH) and AB systems is rather conditional, as FHs have an additional source of energy for charging the battery in the form of an alternative traction motor (for non-rechargeable ones) or the external station (for rechargeable ones). Apart from that, the battery additionally functions as a receiver.

The configuration defines the type of hybrid. So, for example, for pneumatic hybrids with ICEs, called EHPV (engine hybrid pneumatic vehicle), by analogy with electric hybrids, it is customary to consider three schemes (configurations) of power transmission from the power plant units to car wheels.

In the *series scheme* SH (series hybrid), a compressor loads an ICE. In this case, the most economical operating mode of the ICE is chosen. Compressed air from the compressor enters the pneumatic accumulator and then passes the speed reducing drive to the pneumatic motor, which transmits torque to the wheels. While braking, the compressor is driven by the car wheels, providing the accumulation of kinetic energy. The advantage of this scheme is the relative ease of the power plant control, the absence of special transmission units, the possibility of using low-power ICEs in the economical modes. The disadvantage of this option is low efficiency.

In the PH *parallel scheme* (parallel hybrid), an ICE and a pneumatic motor are connected to the driving wheels by the transmission. At the moment of regenerative braking, the pneumatic unit operates as a compressor and provides the air recharging into the pneumatic accumulator. This configuration results in greater efficiency. For the HPP parallel scheme, the complexity of transmission and control system, also a wider range of necessary operating modes of ICE, the need to refuel high-pressure tanks from external compressors are required.

In the *mixed scheme* of MH (mixed hybrid), the drive of driving wheels can functions by means of: pneumatic unit; ICE; pneumatic unit and ICE connected in parallel. In this case, the energy of compressed air in the pneumatic accumulator can be replenished from the outside (for charged ones), from the ICE by a compressor (for non-charging ones) or from the wheels while regenerative braking. The key element of hybrid power plant is a power distributor, which redistributes the power flows between the vehicle's chassis, main engine, auxiliary engine and energy recovery circuit.

Thus, a formal description of a pneumatic hybrid can be represented by an abbreviation: the energy autonomy of a vehicle (EC/NR); the system purpose

(SS/AB/FN); the composition (EHPV/EHHV/PHEV); and the power plant configuration (S/P/M). For example, the abbreviation NR-FN-EHPV-P should be understood as a non-rechargeable full hybrid with an internal combustion engine that works in parallel with an pneumatic motor. In addition, the abbreviation can indicate the features of the power plant units (design of engines, batteries, heat exchangers, mechanical transmission).

Let us consider several examples of pneumatic HPP cars presented in the form of structural diagrams, projects, experimental samples and industrial products.

In the Acceleration-Breaking system, a receiver operates as a pressure accumulator. The use of a reversible vane pneumatic unit in recuperative modes requires the complication of the pneumatic system circuit. This is due to the fact that the reversibility of most rotary machines is characterized by a multidirectional rotor rotation (in the engine mode in one direction, and in the compressor mode in the opposite direction). The driven shaft of the vehicle's transmission, as a pneumatic unit drive, rotates in one direction during the power take-off from the wheels, and during the transmission of torque to them. This can be solved with the help of mechanical (reversible controlled speed reducing drive) or pneumatic (valve) decoupling. The second option is preferable from the point of weight-size parameters, unification of components, cost minimization for further developing the initial transmission of a car, flexibility in the implementation of control actions.

Mechanical reverse valves and electrically controlled spool pressure distributors serve for valve decoupling (Fig. 4.48).

The spool valve (1), in contrast to the axial shut off valve, does not experience the pressure resistance and makes it possible to cut the rigidity of the pull-back spring (2), and therefore the power of electromagnetic unit (3) for their control (Fig. 4.48 a). Holes (4) are vented to the atmosphere, excluding the air resistance in the enclosed volume.

In the bi-stable valve, the spool is driven in both directions by the electromagnetic units (3) (Fig. 4.48 b). The absence of a reset spring and the pulse control can significantly narrow down the power and energy consumption of electromagnetic units, as well as speed up the valve response.

In the *mixed non-rechargeable EHPV-hybrid*, the pneumatic unit is connected to the output shaft of the gearbox (GB) by the drive mechanism (1) (Fig. 4.49).

The heavy arrows in the scheme in Fig. 4.49 indicate the air flow direction in the lines for the compressor (solid) and motor (dotted) modes, thin lines – the

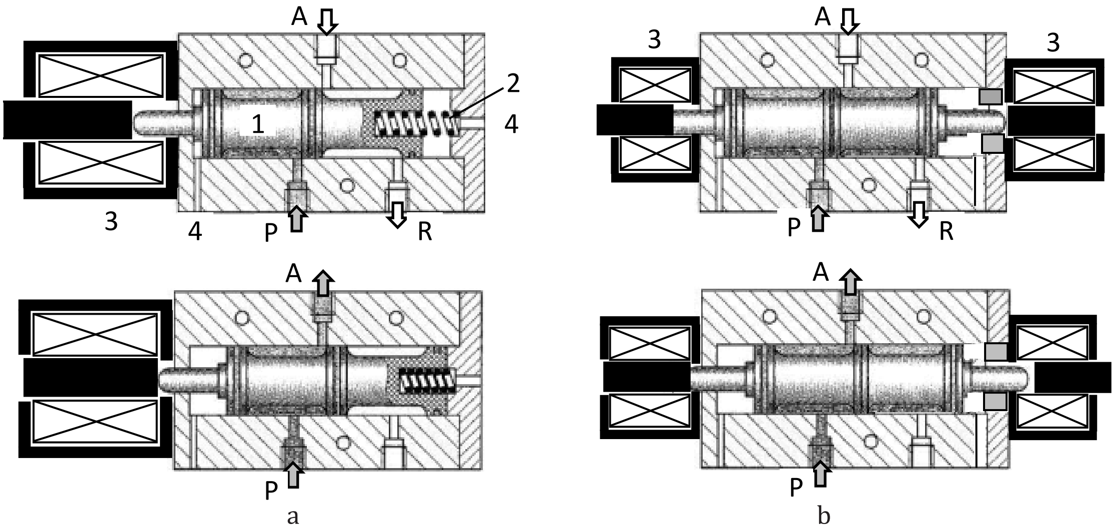


Fig. 4.48. Design of spool pressure distributor: a – with a reset spring; b – bi-stable.

actuator control signals.

The drive mechanism (1) can operate as uncontrolled (permanently connected summing differential) or controlled (disconnected). In the second case, it is disconnected by the electromagnetic clutch as part of the differential (1) (see Sections 2.2 and 2.4).

While operating as a pneumatic motor, the shut-off valve (9) is open, and the

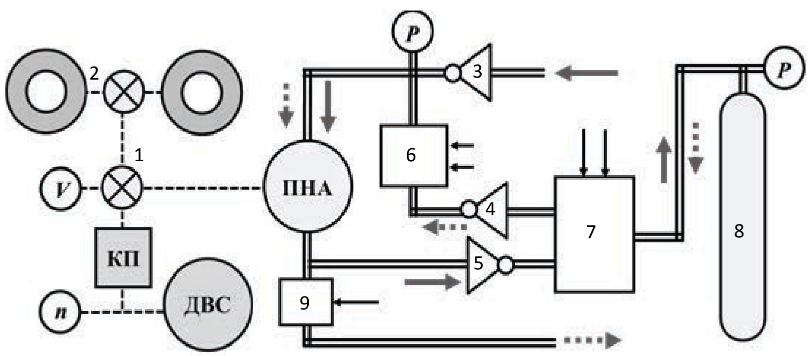


Fig. 4.49. Scheme of non-rechargeable pneumatic hybrid: 1 – drive mechanism; 2 – driving axle; 3 ... 5 – reverse valves; 6 – pressure modulator; 7 – spool pressure distributor; 8 – compressed air cylinder; 9 – shut-off valve.

distributor (7) (see Fig. 4.48) provides the air flow from the cylinder (8). At the same time, the modulator (6) (see Fig. 4.35 b) controls the air pressure at the inlet of the pneumatic unit, and, consequently, the torque on its shaft, at a given loading on the drive of the driving axle (2).

The compressor mode is activated either by the ICE (refueling the compressed cylinder (8)) or by the car wheels (regenerative braking). When the valve (9) is closed, the distributor (7) is switched to the air charging of the compressed cylinder (8). When the compressed cylinder is charged to the pressure limit value, the valve (9) opens the compressor outlet, eliminating the compression losses in the pneumatic unit (the ECO valve function). In the case of using the controlled drive (1), the compressor shaft is disconnected by the electromagnetic clutch on the ground of the ECU command.

Thus, the mixed hybrid configuration includes five operating modes: ICE drive; pneumatic drive; hybrid drive; regenerative braking; air charging of the compressed cylinder by the ICE energy.

The mechatronic system of mixed hybrid is controlled in the semi-automatic or automatic modes. The primary information for starting the optimal control algorithms comes to the ECU from the position sensors: mode selector; accelerator pedal; brake pedal; gearbox selector. Additionally, the vehicle state of the drive system (pressure in the cylinder (8) and line *P* of air supply; vehicle speed (*V)*; revolutions per minute (*n)* and the ICE load) is identified by the information from sensors.

As it has been already noted above, the compressed air, while discharging, lowers its temperature, and the gas cooling in the enclosed space leads to a drop in pressure. To control the specific energy consumption of the thermodynamic system (to reduce losses in cooling) of the pneumatic drive in the HPP, the side thermal energy of the ICE is used. The conductor of the ICE thermal energy can be either the liquid of the engine cooling system or the exhaust gases. In addition, cooling systems of transmission units, electric machines and traction batteries can be considered as potential heat sources. Thus, the use of heat exchanger in the hybrid pneumatic drive system (see Fig. 4.47) tends to increase the energy resource of the HPP as a whole and lower the gradient temperature loads of its elements. It should be noted that the use of heat exchanger is justified for pneumatic systems operating in the exhaustible mode under significant pressure (from high-pressure cylinders), which is typical for rechargeable pneumatic hybrids [57].

Let us consider the application of a heat exchanger using the example of a full chargeable hybrid for the parallel EHPV circuit, where a converted ICE operates

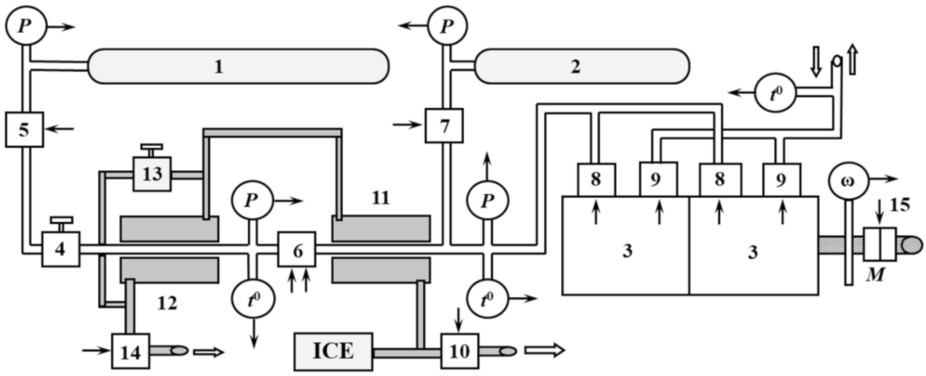


Fig. 4.50. Pneumatic transmission scheme of HPT with a heat exchanger: 1 – high pressure cylinder; 2 – pneumatic accumulator; 3 – cylinders of pneumatic unit; 4, 13 – pressure reducing regulators; 5, 7 – electrovalves; 6 – pressure modulator; 8, 9 – inlet and outlet electric timing valves; 10, 14 – electric shut off valves; 11, 12 – heat exchanger sections; 15 – electromagnetic clutch.

as a pneumatic unit (see Fig. 4.37). The pneumatic transmission of the HPP additionally includes a two-stage heat exchange system between the compressed air supply line and the exhaust tract of the thermal ICE (Fig. 4.50).

The heat exchanger consists of two sections, namely, the initial (11) and final (12) air heating, to heat pneumatic line sections, where the air expands. Moreover, the first, hotter section, operates in the section of lower pressure (after the modulator (6)), and the second one in the expansion section after the speed reducing drive (4). The pressure (*P)* and the air temperature (t^o) in the sections are controlled by the appropriate sensors. Electric shut off valves (10 and 14) control the air heating rate in the exhaust system. The speed reducing drive (13) enables adjusting of the intensity of heat exchange of the second section with its effective volume.

The controlled clutch mechanism (15) connects the pneumatic unit shaft to the mechanical transmission of a vehicle to transmit torque (*M*) while operating as a pneumatic drive, regenerative braking and joint operation of engines [58].

A *series non-rechargeable hybrid* EHPV [59] has a vane pneumatic motor, so the thermal energy of fuel combustion in the ICE cylinders results in heat exchange (Fig. 4.51).

The air is pumped into the pneumatic cylinder (2) by a pressure of up to 2 MPa by the crankcase compression of ICE, created by the piston at the pressure in the working stroke of cylinder (6) up to about 4 MPa. The air charge passing

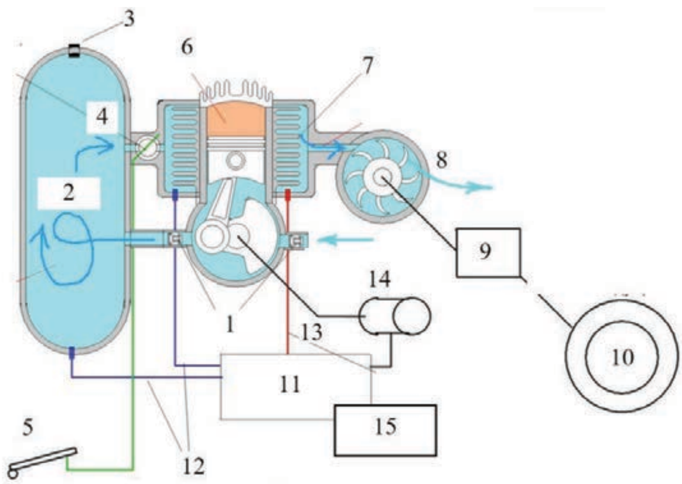


Fig. 4.51. Mechatronic system scheme of a series HPP: 1 – jets; 2 – receiver; 3 – safety valve; 4 – pneumatic motor power regulator; 5 – accelerator pedal; 6 – ICE combustion chamber; 7 – heat exchanger jacket; 8 – pneumatic motor; 9 – pressure reducing regulator; 10 – driving wheel; 11 – ECU HPP; 12 – pressure sensors; 13 – temperature sensor; 14 – starter; 15 – ECU ICE.

the cooling jacket of the ICE (7) boosts the pressure by the temperature up to 3 MPa and enters the inlet (8) of pneumatic motor. The rotor of the latter transmits torque by the pressure reducing regulator (9) to the driving car wheels (10).

The vehicle speed is changed by means of the electrically controlled pneumatic throttle (4) and the electronic pedal (5). The electronic unit (11) controls the pressure (12) in the compressed cylinder and the ICE jacket, as well as the air temperature to control the start of the ICE from the electric starter (14). At the same time, the pressure is periodically pumped at the optimal engine rpm speed (depending on the sensor indices).

The pneumatic motor does not break while free-wheeling. The ICE fuel consumption at a speed of 2000 min-1 for pumping the air of 800 l/min is 0.5 liters. The air consumption to keep a torque of 360 Nm at a frequency of 6000 rpm makes 400 l/min.

HPPs with a *hydraulic transmission* EHHV (engine hydraulic hybrid vehicle) are of greatest practical interest. The energy source in such pneumo-hydraulic systems is a hydraulic accumulator; it presents a compressed cylinder with two chambers separated by an elastic shell, diaphragm or piston. One chamber is closed and filled with gas (gas **back-up**). The liquid (usually transmission) is supplied to the second chamber, and it pressures out the gas chamber volume, increasing the gas pressure (usually nitrogen) in it.

A *parallel non-rechargeable hybrid* EHHV contains four drive modes [60] (Fig. 4.52).

Figure 4.52 depicts: 1 – ICE; 2 – gearbox; 3 – hydraulic unit; 4 – cardan transmission; 5 – hydraulic accumulator; 6 – working fluid reservoir. Compressed air tanks of reservoir (6) and accumulator (5) have a similar design, but with a difference in the pressure of the gas back-up by dozens of times. This allows to adjust the liquid volumes in the compressed air tanks in the static system (Fig. 4.52 a). To build an EHHV drive by the mixed scheme of power transmission to the wheels, a more complex mechanical transmission of a conventional car or additional hydraulic units are used.

The strength of pneumo-hydraulic systems in comparison with the pneumatic ones is their closed cycle and the high efficiency of a hydraulic unit. A significant drawback limiting the use of EHHV hybrids is the required volume of the working fluid and, accordingly, its weight. To this should be added the dimensions and weight of compressed cylinder in the accumulator and the reservoir.

The composition of *mixed non-rechargeable hybrid* EHHV may include several reversible hydraulic units, for making better the fuel, economic and environmental performance characteristics of a vehicle [61]. To implement a series configuration in the alternative drive options, an additional hydraulic unit driven by an ICE is used (Fig. 4.53).

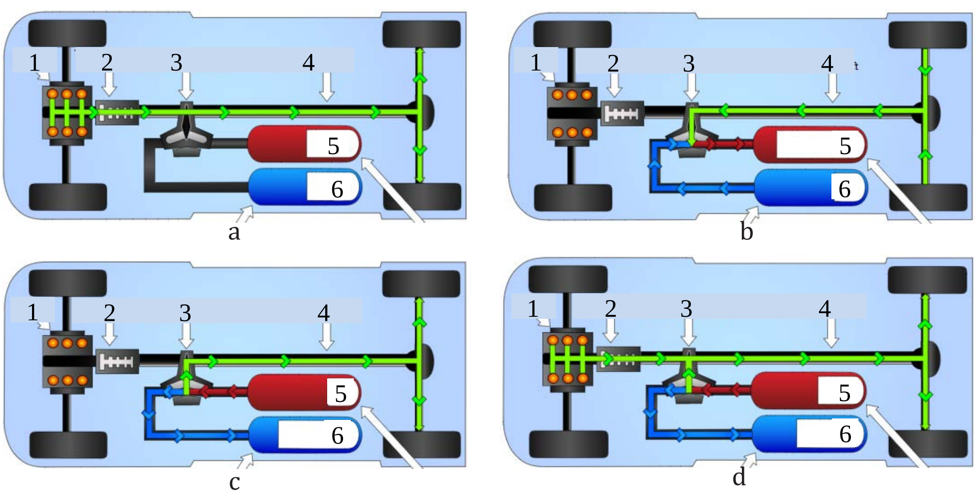
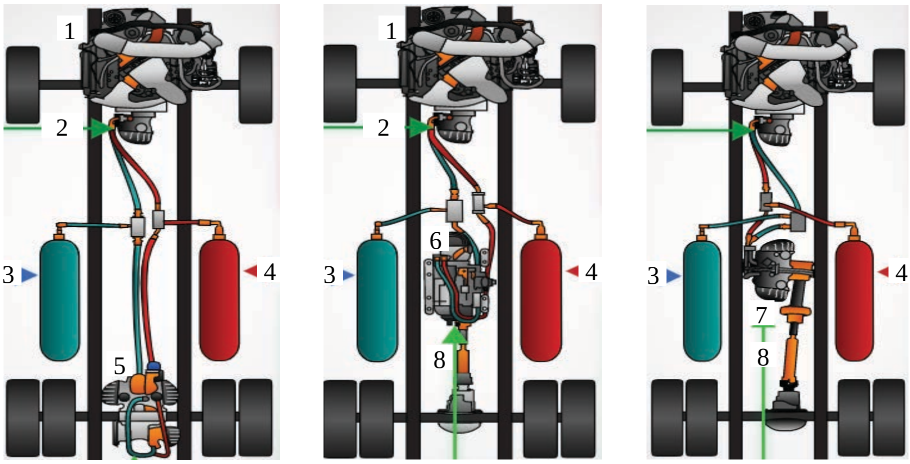


Fig. 4.52. Power transmission scheme in the following modes: a – ICE drive; b – regenerative braking; c – hydraulic motor drive; d – hybrid drive.



a b c

Fig. 4.53. Alternative configurations of EHHV actuators: a – with an integrated rear axle; b – with integrated gearbox; c – with a separate hydraulic unit.

Figure 4.53 shows: 1 – ICE; 2 – hydraulic unit with an ICE drive; 3 – reservoir with the working fluid of low pressure; 4 – hydroaccumulator of high pressure; 5 – rear axle with two hydraulic units; 6 – gearbox with a hydraulic unit; 7 – double hydraulic unit; 8 – cardan transmission. All hydraulic units are activated by general compressed cylinders (3 and 4). The hydraulic unit (2) pumps the pressure into the hydraulic system from the ICE energy, starts-up and accelerates the car by the front wheel drive. Hydraulic units of the rear axle (5) provide a regenerative mode for each wheel. The multiplicator of a four-speed gearbox with a computer switching without a clutch is placed in the hydraulic unit drive (6) and it also provides the regenerative braking and torque boost to the rear wheels. A double hydraulic unit (7) of a separate design, fixed to the vehicle frame, functions in the same way.

French company PSA Peugeot Citroen presented a *mixed non-rechargeable front-wheel drive hybrid* EHHV, mounted in Citroen C4 Cactus Air Flow 2L and Peugeot 208 Hybrid Air 2L produced on an industrial scale [62]. The hybrid power plant (1) combines a 2-liter ICE and a hydraulic drive with compressed cylinders of a hydraulic accumulator (2) and a fluid reservoir (3) (Fig. 4.54).

The hydraulic unit (4) is connected to the ICE shaft (5) by the planetary transmission (6), located in the automatic gearbox (Fig. 4.50 d). In the hybrid mode, the total torque from the power plants (4 and 5) is transmitted to the front wheel shafts by the axle drive. The operation of regenerative braking and pneumatic traction drive is carried out as described above (see Fig. 4.52), only for the front axle.

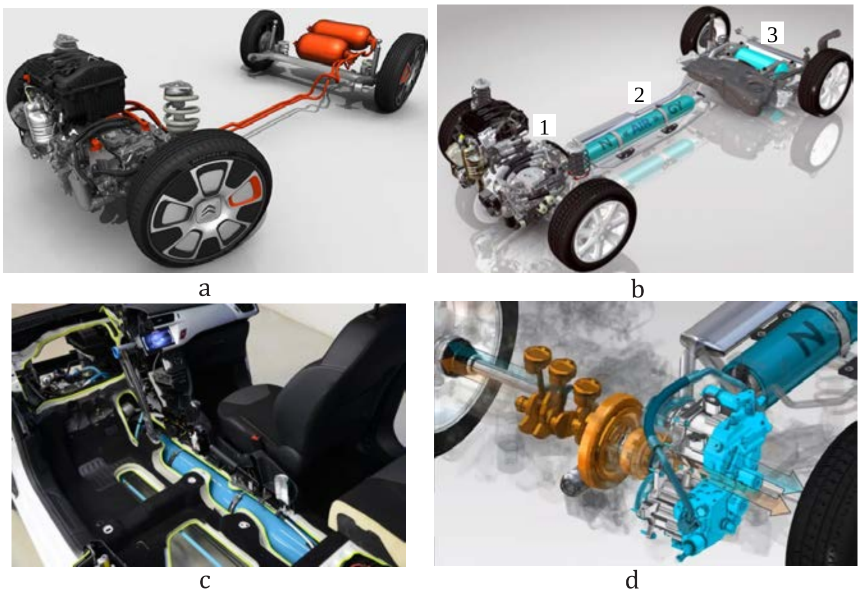


Fig. 4.54. Industrial designs of hybrid drive EHHV: a – structural configuration of air flow elements; b – structural configuration of hybrid air elements; c – hydro-accumulator layout; d – power unit design.

The compressed air energy in a fully charged battery provides a range of several kilometers at speeds up to 70 km/h. The hybrid mileage makes 1300 km with a full tank and the air charging. The weight of hydraulic transmission units does not exceed 100 kg. Hybrid units in the urban cycle save the fuel consumption up to 45 %.

Pneumatic charging hybrids, called PHEV (pneumatic hybrid electric vehicle), are equipped with high pressure cylinders (up to 300 bar) with a filling unit (1), a speed reducing drive (2) and a receiver of reduced pressure (3) [63]. The PHEV power plant includes a pneumatic motor (4) and an electric machine (5) powered by the traction battery (6) (Fig. 4.55).

PHEV powerplants from different manufacturers have their own performance characteristics. So, for example, the power plant of Energine *parallel hybrid* from South Korea contains a pneumatic converted engine (4) and a reversible electric machine (5); their shafts are interconnected by the V-belt transmission (7) (Fig. 4.55 a). The HPT control system of a hierarchical structure distributes the power between the power plants depending on the vehicle driving conditions. In the second hierarchy level, the ECU (8) controls the electric machine (5) while operating as an engine and a generator (recuperation of braking energy into the rechargeable battery charge).

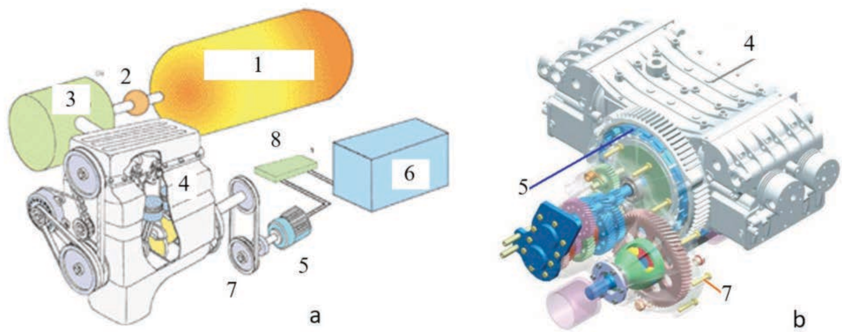


Fig. 4.55. Pneumo-electric power plants of PHEV vehicles: a – designed by Energine; b – CAT machines.

When a car is moving at a constant speed over 20 km/h, the key thing is an electric motor, as a vehicle starts, speeds and climbs uphills by means of the pneumatic traction. The Energine's hybrid accelerates and keeps the maximum speed of 120 km/h for an hour using only the pneumatic traction. This car weighs 1260 kg, including an aluminum body of 400 kg, both engines of 60 kg, a 100-liter air cylinder of 40 kg. Such cars have Ni-MH traction batteries with an operating voltage of 280 V.

The power plant, built on the basis of the Nègre's four-cylinder pneumatic engine, is mounted by MDI on the CAT series vehicles [64]. The HPP composition includes a reversible pneumatic unit (4), an integrated reversible AC electric machine (5) as part of the transmission (7) (Fig. 4.55 b). The HPP control system, in addition to the described functions, charges compressed cylinders not only from the compressor station, but also from the household power supply using the pneumatic unit (4) while operating as a compressor driven by the electric machine (5), working as a motor.

In conclusion, it should be emphasized that the options for creating a HPP are unlimited by the examples studied above. Technological progress tends to use alternative energy sources and new methods of its transformation, improving the power plant structure in a car, as well as the methods and ways of controlling the operational modes of vehicles.

References

1. Жавнер В. Л., Смирнов А. Б. Мехатронные системы: учеб. Пособие. СПБ.: Изд-во Политехн. ун-та, 2011. 131 с.

2. Применение мехатронных и робототехнических систем на транспорте. 2019. URL: https://helpiks.org/4-75856.html.

3. Соснин Д. А. Автотроника. Электрооборудование и системы бортовой автоматики современных легковых автомобилей: Учебн. пособие. М.: СОЛОН-Р, 2001. 272 с.

4. Дентон Т. Автомобильная электроника; пер. с англ. Александрова В.М. М.: НТ Пресс, 2008. 576 с.

5. Данов Б. А. Електронні системи управління іноземних автомобілів. М.: Гаряча лінія-Телеком, 2002. 204с.

6. Бороденко Ю. М., Дзюбенко О.А., Биков О.М. Діагностика мехатронних систем автомобіля. Підручник. Харків: ХНАДУ, 2016. 320 с.

7. Соснин Д. А., Яковлев В.Ф. Новейшие автомобильные электронные системы. М.: СОЛОН-Пресс, 2005. 240 с.

8. Литвиненко В. В. Автомобильные датчики, реле и переключатели. Краткий справочник. М.: ЗАО КЖИ «За рулем», 2004. 176 с.

9. Системы мониторинга. 2019. URL: https://studopedia.su/7_1705_sistemi-monitoringa.html

10. Деменков Н. П. Нечеткое управление в технических системах: учеб. пособие. М.: Изд-во МГТУ им. Н. Э. Баумана, 2005. 198 с.

11. Искусственная нейронная сеть. 2019. URL: https://ru.wikipedia.org/wiki/

12. Искусственные нейронные сети (ИНС). 2019. URL: https://www.it.ua/ru/knowledge-base/technology-innovation/iskusstvennye-nejronnye-seti-ins.

13. Управление транспортным средством с помощью нейронной сети. 2019. URL: https://habr.com/ru/post/101789

14. Барский А. Б. Нейронные сети: распознавание, управление, принятие решений. М.: Финансы и статистика, 2004. 176 с.

15. Сигеру Омату, Марзуки Халид, Рубия Юсоф. Нейроуправление и его приложения. 2-е изд. М.: ИПРЖР, 2000. 272 с.

16. Полный привод: особенности, плюсы и минусы конструкций. 2019. URL:

https://5koleso.ru/articles/garazh/polnyy-privod-osobennosti-plyusy-i-minusy-konstrukciy

17. Виды полного привода автомобилей с конкретными примерами. 2019. URL: http://autoepoch.ru/avtoazbuka/polnyj-privod-avtomobilya.html

18. Электромагнитная муфта. 2019. URL: https://stankiexpert.ru/tehnologicheskaya-osnastka/zapchasti/ehlektromagnitnaya-mufta.html

19. Что такое Haldex, как он работает и чем отличается 1-5 Gen. 2019. URL: https://www.drive2.ru/b/1488033/

20. Система полного привода xDrive. 2019. URL: http://systemsauto.ru/transmission/xDrive.html

21. Электронное сцепление. 2019. URL: http://systemsauto.ru/coupling/eClutch.html

22. Электронная педаль сцепления. 2019. URL: https://znanieavto.ru/mufta/elektronnoe-sceplenie.html

23. Роботизированная коробка лада AMT. 2019. URL: https://otoba.ru/transmissii/vaz/2182.html

24. Блокирующийся дифференциал elocker eaton. 2019. URL: https://b2motor.com/index.php?route=information/articles&articles_id=16

25. Блокируемые дифференциалы. 2019. URL: https://5koleso.ru/content/blokiruemye-differencialy

26. Активный дифференциал или управление тягой. 2019. URL: https://automobilproject.pp.ua/aktivnyj-differencial-ili-upravlenie-tyagoj/

27. Система A.D.D. описание работы системы. 2019. URL: https://www.drive2.ru/l/473028869918556343/

28. Межосевой дифференциал с изменяемой степенью блокировки DCCD. 2019. URL: https://autodata.ru/article/all/mezhosevoy_differentsial_s_izmenyaemoy_stepenyu_blokirovki_dccd/

29. Полный привод SH-AWD - уникальное решение от Honda. 2019. URL: https://pikabu.ru/story/polnyiy_privod_shawd__unikalnoe_reshenie_ot_honda_4332705

30. Актив, пассив, гибрид? 2019. URL: https://www.mitsubishi-motors.ru/pressauto/aktiv_passiv_gibrid/

31 Альтернативные виды топлива. URL: https://xn--80aaf3axghkd.xn--p1ai/alternativnye-vidy-topliva

32. ГазоМАЗы: сжиженный метан и автобусы для Челябинска. URL: https://autoreview.ru/articles/gruzoviki-i-avtobusy/gazomazy-szhizhennyy-metan-i-avtobusy-dlya-chelyabinska

33. Четыре главных вопроса перспективы массового использования электромобилей. 2019. URL: https://hevcars.com.ua/reviews/4-pregrady-na-puti-razvitiya-uspeha-elektromobiley/

34. Автомобиль на водороде. Пора ли прощаться с бензином? https://habr.com/ru/company/toshibarus/blog/430372/

35. Разработка первой в Украине демонстрационной модели экологически чистого автомобиля с криогенной силовой установкой. http://www.holod-konsultant.ru/articles/nizk/kriogennii_avtomobil.htm

36. Авто на солнечных батареях. Виды и перспективы. URLhttps://amastercar.ru/blog/solncemobili-mif-ili-realnost.html

37. Атомные автомобили: прошлое и будущее. URL: https://www.mirf.ru/science/atomnye-avtomobili

38. Аргун Щ. В. Визначення найбільш ефективних тягових джерел струму для електромобілів. Автомобильный транспорт. Сб. Научных трудов. Харьков: ХНАДУ, 2017. Вып. № 41. С. 11 – 22.

39. Аккумуляторные батареи для электромобилей – все, что вы хотели знать. 2019. https://hevcars.com.ua/reviews/akkumulyatornyie-batarei-dlya-elektromobiley/

40. Электродвигатели. 2019. URL: https://engineering-solutions.ru/motorcontrol/motor/

41. Электромобиль-конструктор: моторы и привод от GKN. 2019. URL: https://e-move.com.ua/elektromobil-konstruktor-motory-i-privod-ot-gkn/

42. Мотор-колесо Дуюнова. 2019. URL: http://fb.ru/article/341644/motor-koleso-duyunova-opisanie-ustroystvo-shema-i-otzyivyi

43. Мотор-колесо Шкондина. 2019. URL: http://www.electra.com.ua/istoricheskie-fakty/454-motor-koleso-shkondina.html

44. Инвертор: что это и как работает. Принцип работы гибридного Lexus (Toyota). 2019. URL: https://club-lexus.ru/faq/index.php?article=724

45. Преобразователи напряжения. Виды и устройство. Работа. URL: https://electrosam.ru/glavnaja/jelektrooborudovanie/jelektropitanie/preobrazovateli-napriazheniia/

46. DC/DC-преобразователи: принципы работы и уникальные решения Maxim Integrated. https://www.compel.ru/lib/134297

47. Управление вентильным электродвигателем. 2019. URL: http://electricalschool.info/elprivod/1910-upravlenie-ventilnym-jelektrodvigatelem.html

48. Векторное управление двигателем. 2019. URL: https://engineering-solutions.ru/motorcontrol/vector/

49. Классификация гибридных систем. 2019. URL: https://ustroistvo-avtomobilya.ru/dvigatel/klassifikatsiya-gibridnyh-sistem/

50. Бажинов О. В., Смирнов О.П., Сєріков С. А., Гнатов А. В., Колєсніков А. В. Гібридні автомобілі. Харків, ХНАДУ, 2008. 327 с.

51. Плюсы и минусы автомобилей на сжатом воздухе. 2019. URL: https://autogeek.com.ua/plyusyi-i-minusyi-avtomobiley-na-szhatom-vozduhe/

52. Пневматические двигатели, конструкция и параметры. 2019. URL: https://studref.com/481952/tehnika/pnevmaticheskie_dvigateli_konstruktsiya_parametry

53. Выкачивание воздуха автомобиля. https://spectrum.ieee.org/energy/environment/deflating-the-air-car

54. Jaroslaw Zwierzchowski. Design type air engine Di Pietro. 2019. URL: https://www.epj-conferences.org/articles/epjconf/pdf/2017/12/epjconf_efm2017_02149.pdf

55. Air cars. 2019. URL: https://sites.google.com/site/rijdenopperslucht/home/menu-en/engineair-en

56. Hybrid Opposite Piston Engine – HOPE & Portable Range Extender. 2019. URL: http://www.hybrid-engine-hope.com/other_competitive_solutions (дата обращения 16.01.2020).

57. Бороденко Ю. Н. Концепции построения пневматических гибридов. // Автомобильный транспорт. Сб. Научных трудов. – Харьков: ХНАДУ, 2020. – Вып. № 46 – С. 1–8.

58. Бажинов А. В., Бороденко Ю. Н., Сериков С. А., Дзюбенко А. А. Система управления пневмотрансмиссией гибридного автомобиля. Вісник східноукраънського національного університету імені Володимира Даля. Луганськ: СНУ, №6(112), 2007. – С. 11 – 14.

59. Пневмопривод. Автомобили, работающие на сжатом воздухе. 2020. URL: https://oborudow.ru/transmissiya/ pnevmoprivod-avtomobili-rabotayushchie-na-szhatom-vozduhe-airpod/ (дата обращения 16.01.2020).

60. How Parallel Hydraulic Hybrid Vehicles Work. 2019. https://archive.epa.gov/otaq/technology/web/html/how-it-works-parallel.html (дата обращения 16.01.2020).

61. Hydraulic Hybrid Vehicle Layouts. https://archive.epa.gov/otaq/technology/web/html/prototype-vehicles.html

62. Hybrid Air – воздушный гибрид от Peugeot Citroën. 2019. URL: https://www.facepla.net/the-news/eco-transportation-mnu/3180-hybrid-air-peugeot-citroen.html (дата обращения 16.01.2020).

63. Pneumatic Hybrid power leading the future. 2020. URL: https://pneumaticservovalve.wordpress.com/2015/03/26/ (дата обращения 16.01.2020).

64. Воздушные автомобили. 2020. URL: https://www.metodolog.ru/node/230

Chapter 5:

Mathematical aspects of multimodal transportation

Jelena Pecerska
Nadezhda Kunicina
Anatolijs Zabasta
Leonids Ribickis
Andrejs Romanovs

Riga Technical University

The goal of this chapter is to give a concise introduction into some mathematical aspects in the context of multimodal transportation.

The basic mathematical background is assumed, as well as the basic knowledge in system modelling, probability theory, mathematical statistics and information technology.

The chapter is divided into six sections which can be read consecutively.

5.1. Modes of transportation

By the general definition of transportation from Merriam-Webster the *transportation* is

1 : an act, process, or instance of transporting or being transported;

2 a: means of conveyance or travel from one place to another.

b: public conveyance of passengers or goods especially as a commercial enterprise

There are also other interpretations of the term, but for the purposes of this chapter we accept this general definition.

The conventional European transportation modes are mentioned at the *Eurostats* site. Freight transport in the European Union (EU), covering the transport modes, is road, rail, air, maritime and inland waterways [1].

The APICS Dictionary [2] provides the following insight into multimodality: multimodal solutions – transportation plans that involve multiple means of transportation and coordinate the physical and information requirements.

The best known definition of multimodal transportation is adopted from the United Nations Convention on International Multimodal Transport of Goods published in 1980. "'International multimodal transport' means the carriage of goods by at least two different modes of transport on the basis of a multimodal transport contract from a place in one country at which the goods are taken in charge by the multimodal transport operator to a place designated for delivery situated in a different country". The multimodal transport nowadays is not necessarily only international, so the definition is acceptable for the purpose of this chapter.

The transportation process is under consideration in the frame of transportation systems of different scales – global, state, region, or others. It is even more precise to discuss the network of transportation systems, as soon as it is practically not possible to distinguish a real independent transportation system.

The unique transportation case of freights or passengers may be provided in combined or multimodal mode. In multimodal transportation process various means of transportation are involved. A passenger planning a tour to Stockholm from Riga is actually planning a combined transportation, e.g., by train to Riga Central Station, by tram or taxi to Riga Passenger Terminal, by ferry to Stockholm Passenger Terminal, and by bus or taxi to the hotel in the city centre. One trip from home to the hotel of destination is implemented as a combination of several transportation modes.

There is a typical mistake of using the term *intermodal* instead of *multimodal*. The question is – is this process a multimodal or intermodal transportation? The answer is very simple. There is a difference in terms of service provider. If you are planning your tour individually and buying all the tickets by yourself, you may identify your transportation as intermodal. If you are enjoying the tour operator service – just fix the start and the destination points of the trip – you get a single voucher for your trip from the tour operator, mentioning where, when, how and by whom you are transported at particular meeting points, then this trip looks like multimodal. The same approach helps to distinguish the freight transportation mode – a single service provider for a particular transportation case is in the case of multimodal transportation. If the customer is planning the changes in transportation modes and making the choice of route, service provider, etc., intermodal transportation is under consideration. The subject of this chapter is multimodal transportation related mathematical issues.

The problems in the area of analysis of transportation systems are of various nature. Modern analysis approaches suppose implementation of various types of mathematical models. The models represent processes, activities, demand, alternative analysis, traffic assignment, traffic flows. The traffic flow models may differ by level of detail. The particular entities under consideration are typical for micro models. The high level of abstraction and consideration of flows rather than particular entities is the feature of macro models. And finally, the combination of detailed description of some process features with the generalised approach to flow simulation are represented in mezo models [3]. In the frame of these models, the set of variables describes the transportation systems, providing the information about system state evolution over time. The model-based analysis approach in transportation area is relevant for any type of transportation, not only for the multimodal transportation. Thus, these models are mentioned also further throughout the chapter.

Multimodal freight transportation has obvious advantages:

- flexibility
- any type of container
- route selection based on the customer's wishes
- the ability to redirect during delivery
- delivery 'from door to door'
- monitoring at any stage of transportation
- no need to search for multiple carriers
- the authority transfer in organization of a complex process to the carrier and others

The main advantage of multimodal transportation is the ability to take the benefit from each transportation mode. This may be not only low cost, but also acceptable delivery time and reliability, etc. These and other reasons explain the fact that multimodal transportation is the most modern type of delivery used in international transportation.

There are no principle disadvantages in multimodal transportation. However, there is a higher risk of damage or loss of goods due to the use of various types of transport, as well as to extra loading and unloading operations. The solution is in the choice of a reliable carrier.

5.2. Mathematical notations in descriptive multimodal transportation models

Transportation process, as research object, can be described in terms of discrete events. In discrete-event approach only the points in time at which the state of the system changes are considered.

In systems, which are analysed as discrete-event systems (DES), the operation of a system is interpreted as a chronological sequence of events. The events are the instant changes of system state. The events occur at particular, irregular time. The general event types associated with the multimodal transportation process comprise transportation start, finish, transportation mode change and other type of events if necessary. The transportation process is associated with its entity – the transportation unit (TU). Every passenger, parcel or freight unit are considered as TU. In the paradigm of DES the following multimodal transportation process definition could be provided: Multimodal transportation process is the chronological sequence of events, each event is the state change and is managed by TU or manages the TU.

The objects of transportation process determine the special features of a process and natural units of measurement. The transportation process may be a single unique process or regular repeated process. The regular repeated process may have both constant and variable parameters.

The transportation process may be interpreted as a material flow. Material flows are inherent components of diverse systems. Material flows ensure both the interaction of system elements and system links with the environment. Thereby most of complex systems may be considered and analysed as material flow handling systems. The previous years publications apply such terms as 'flow system', 'material flow system', 'material handling' or 'material flow handling system'. They state that the guarantee of effective interaction between the flow and manufacturing or distribution systems is an important task. The volume of the costs, related to control of the material flows between the elements of production or distribution system, is evaluated from 13 to 30 percent of total production or distribution costs. The concept of the flow system is wider than the concept of material flow system. The material flow system concept can be applied to describe the transportation systems as well. Taking into account the above mentioned, we could formulate the alternative definition of transportation system – a technically controlled material flow system, created to move the objects of material flows. Almost all up-to-date simulation software provide tools and components for material flow analysis, description, incorporation into simulations, obtaining and visualization of appropriate simulation results [4].

The specification of transportation process could be made in the form of informal description, table, formalism, graphic scheme, or algorithm. The type of the specification is determined by the purposes of application. Transportation process specification as a table is an elementary transportation flow model. Table 5.1 shows an example of such specification.

This tabular model can also be named a protocol of transportation process events. The protocol is based on the approach proposed for the material flow protocol [5]. The table may also be designed in a more compact form, however, for the purpose of this section it is provided in the detailed form.

Here i is an ordinal number of the considered TU_i, $i = \overline{1, M}$. In this example variable M is a maximum of considered TU, j is the number of the event type, $j = \overline{1, K}$, where E_1 is transportation start event, EK – transportation finish event, and other events may be associated with transportation mode change or other relevant events in the transportation process of the freight under consideration. K-2 is the number of all possible events in the transportation process. The event time (tij) is a variable, corresponding to the occurrence of the j-th type of event related to the i-th TU.

The transportation process P, described in Table 1, is a structure

$$P = \langle T, E, F \rangle, \tag{5.1}$$

where

T – a set of particular events times;

E – a discrete set of events types;

F – a discrete set of TU_i, $i = 1, M$.

The protocol can be augmented with the relevant information about the TU, such as weight, volume, transportation and storage conditions and other properties.

The protocol can be visualized as time diagram, as shown in Fig. 5.1.

In Fig. 5.1 the transportation process events are shown as vertical line segments of different height corresponding to particular TU. However, it is possible to use the height notation to introduce the volume, weight or container type of the TU. The time moments are according to the Table 1. It is possible to see the irregular time periods between particular events related to a concrete TU. Figure 5.1 provides Δt_{1i} intervals between events of the TU_1 process. The data may be used for further simulations and as simulation validation data.

Creating transportation process specifications, the best results can be achieved by combination of different types of description. A universal transportation process description method or type could hardly be defined. Certain types of description, such as formal and graphic, are more applicable in the area of mathematical modelling.

The transportation process protocol is not the only form of the formal description proposed by researchers. This approach corresponds to the Event Logs that are obtained from the observations. As the event log is a record of

Table 5.1. Protocol of Transportation Process Events.

The ordinal number of a transportation unit in a process (i)	The event type (Ej)	The event time (t_{ij})					
1	E_1	t_{11}					
1	E_2		t_{21}				
…	…			…			
1	E_j				t_{j1}		
…	…					…	
1	E_K						t_{K1}
2	E_2	t_{12}	t_{22}		t_{j2}		t_{K2}
…							
I	t_i	t_{1i}	t_{2i}		t_{ji}		t_{Ki}
…							
M	t_M	t_{1M}	t_{2M}		t_{jM}		t_{KM}

events, the correspondence between the process protocols and event logs is very close [6].

This form of process description can be used as a conceptual model for the simulation of the process. The process mining is also based on the event log. Process mining uses event data to extract process-related information. The process mining is introduced as a technique to provide means for process improvement in various application domains. “Process mining is an emerging discipline providing comprehensive sets of tools to provide fact-based insights and to support process improvements. This new discipline builds on process model-driven approaches and data mining” [6]. As the most challenging task of the process mining the automatic generation of process models, based on the event log, is highlighted. The event log form of process description can be used for input data preparation for mathematical modelling. For example, extracting the time moments from the log or protocol provide the data for finding *interarrival* time distribution between freights. These interarrival times or time intervals are calculated as

$$\tau_j = t_{j+1,1} - t_{j,1}, j = \overline{1, M}.$$

The variable τ (probability) distribution is relevant to simulate incoming flow of freight objects.

The section sets out general requirements to the formal description of multimodal transportation process, which should fit the specific requirements. The qualitative description develops the basis for the analysis and formalization of transportation process at a conceptualization stage of simulation model.

5.2.1. Objectives of multimodal transportation

The effectiveness of multimodal transport is that it uses the main advantages of each type of transport: costs, speed, accuracy, and environmental impact.

Multimodal transport has many advantages that correspond to the objectives of multimodal transportation in comparison with conventional unimodal transportation or intermodal transportation. These general objectives are:

- Total costs reduction
- Delivery time reduction
- Origin and destination points all over the world
- Security

The objectives of multimodal transportation do not differ in principle from

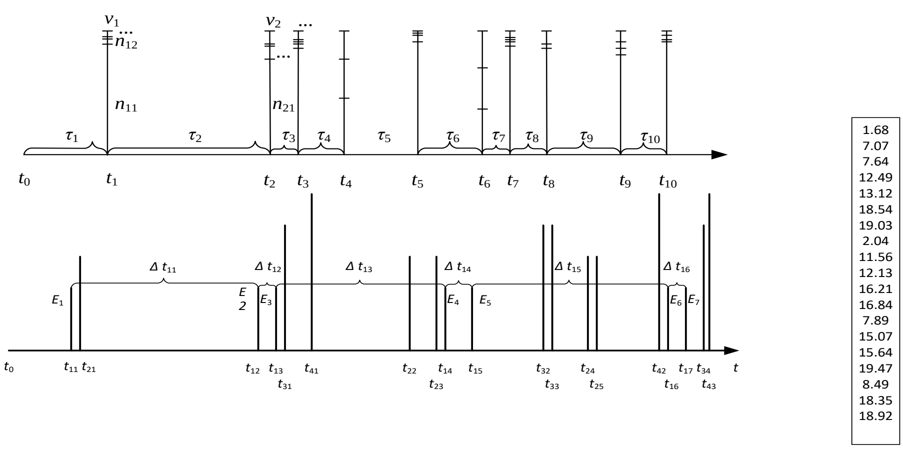


Fig. 5.1. Transportation process events over time.

the goals of any other types of transportation. These objectives can be considered from the point of view of optimizing the transportation process considering various goal functions.

5.2.2. Optimisation tasks in multimodal transportation

Since the goal of multimodal transportation of goods and passengers is to achieve the best possible results, it is quite logical that various optimization problems are considered in this area.

Optimisation in general is a mode of action, providing the choice from all possible options of resource usage to obtain the best results. This not exactly academic definition is emphasising the resource usage. A more specific mathematical formulation may be – the procedures or algorithms of finding the extreme value of a goal function involved in this choice. As soon as the goal function is mentioned, the choice procedure should be formalised.

There are also organisational or infrastructure-based ways of optimisation of multimodal transportation [7]:

- Application of a centralized system for the import and export of goods; development of an integrated network of transport, warehouses and logistic hubs; creation of multimodal logistics enterprises. As soon as these activities are infrastructure-based the investments are necessary.

- Creation of multimodal corridors and regional transport and logistics systems; expansion of the set of transport and forwarding services; modernization of warehouses and hubs; restructuring of transport enterprises. Both investments and infrastructure changes are assumed.

- Determination of rational areas of transport application or equal distances; calculation of the economic effect of the selected transport mode. Calculations become significantly more complicated. Formal definition of the transport route needs to be actualised.

- Freight flow forecasting. High possibility of forecast errors, as soon as the transportation parameters are dynamic.

Bottleneck identification that is concerned with the analysis of infrastructure resource plans [8].

Bottleneck identification is concerned with optimisation of intermodal and multimodal transportation chains, timely and concurrent use of resources, transaction analysis, and multicriteria approach.

Coming back to the mathematical aspect, we assume that the optimisation tasks are solved with optimisation models. Within the scope of this section it is not realistic to provide a complete review of the types of optimisation models, however, we introduce some basic concepts.

The first idea is about the optimisation models. Not only purely analytical models are used, other types of mathematical models are also suitable. The widely used mathematical optimisation models are algorithmic or simulation ones. In this case the optimisation algorithms do not give 'precise' optimal solution, as analytical tasks do. Simulation optimisation usually gives the approximate, or best possible solution. Simulation-based optimisation is used in different types of transportation tasks. There is no unique recommended optimisation technique. The review [9] provides the summary of optimisation algorithms that are relevant to simulation optimisation of complex systems and is based on the Winter Simulation Conference publications over the long period of time.

Despite the fact that the applied optimisation tasks belong to completely different areas, they have a common form. The main stages of optimisation task formulation are the construction of the goal function, the definition of the set or the space of feasible solutions, and the set of constraints of control variables. The solution is obtained by changing the control variables, taking into account the constraints, and looking for the extreme value of the goal function.

The tasks are more complex in case of multi-criteria or multi-objective optimisation, when multiple goals are formulated. The typical fault formulation of the optimisation task is "to find the solution providing maximum profit by using minimum of resources". The error is in the task formulation – to find the optimum of two contradicting values. The correct task formulation may be "to find the solution providing maximum profit by using a pre-defined amount of resources". There is a possibility to formulate an alternative task as well.

What are the specific features of optimisation tasks in multimodal transportation area? First of all, it is necessary to consider the main performance indicators of the multimodal transportation process.

The performance of a process or organization can be defined in different ways. Typically, three dimensions of performance are identified: time, cost, and quality [6].

Other authors provide more detailed approach to the factors that may be taken into account to evaluate the performance of the transportation process

[10]. The factors that may be introduced into performance evaluation are: route length, environmental Impact, time window compliance, and other factors. Most of these detailed factors can be expressed in terms of time, cost and quality.

For optimisation it is absolutely necessary that the goal function is measurable.

The formulation of a simple optimisation model of a multimodal transportation process is based on the process description and transportation goals.

Example 1

The example is based on the case described in [11]. The transportation process under consideration is a process of one particular freight transportation from manufacturing facility to the regional distribution centre. There are various potential modes of transportation. The efficiency of the selected scenario is evaluated by the following particular criteria:

- Transportation costs and costs of additional operations with cargo (reloading, insurance, customs), *TC*, currency units.
- Pipeline inventory costs, *PC*, currency units.
- Risk of sudden transportation disruption costs (calculated based on the history of the transportation or expert evaluation), *DC*, currency units.
- Transportation time or lead time, *L* , days.

The current problem has a countable number of alternatives or alternative scenarios, so we have a discrete problem. The alternatives are formulated explicitly. The goal of the scenario evaluation is to minimize the values of all the partial criteria. Table 5.2 shows the calculated values of individual criteria of the scenarios.

The first action that can be done after the simplified analysis of alternatives, is the decision not to use Scenario E because all its partial criteria values are worse than of Scenario F.

Analysis of total costs shows that Scenario D has the lowest total costs:

$C_d = TC + PC + DC = 17583.8 = \min (C_a, C_b, C_c, C_d, C_f)$.

Thus, the goal function, including only costs, has minimum value for this scenario.

But in this case we have more than one criteria; and the second one is the transportation time and the functions that are providing the common 2-dimensional vector criterion:

$F = (C,L).$

Criterion F has values in the space of 2-dimensional vectors. All existing valid alternatives (a, b, c, d, and f) can be described in terms of possible decision vector, for example, alternative a has the criteria vector $F_a = (C_a, L_a)$.

In our case, the choice in the set of alternatives is mathematically equivalent to the choice in the set of vectors. All the definitions and results can be formulated both in terms of alternatives and in terms of vectors. It is always possible to make the transition from one form of presentation to another. A problem formulation includes a set of valid alternatives (a, b, c, d, and f) and a vector criterion F, and now our example problem is formulated as a multi-criteria optimization problem.

The feasible alternative may be chosen based on the Pareto optimality principle. The Pareto optimal vector for multi-criteria alternative is the one that has the higher (or lower) values of at least one of the criteria in comparison with other vectors. Looking back at the example problem, we can see that there is a constraint or constraints missing for the second criterion – transportation time L. Scenario D has the best cost value C_d, but the transportation time is the longest of all feasible solutions. If there are no constraints for this criterion, alternative d is the solution. If the constraint is introduced, for example,

$L \leq 3,$

then the Pareto optimal solution is alternative f.

In conclusion of the example, we note that the introduction of additional choice criteria increases the dimension of the problem. However, high-dimensional problems can also be solved using the Pareto optimality principle, or by other methods. If it is possible to build a simulation model of the process, it is possible to use plethora of optimization methods. Algorithms, heuristics and iterative methods are applicable, and the researchers have a wide choice.

In some cases multi-criteria methods are used that do not require mathematical models. The fundamental complexity of multi-criteria choice problems is that it is impossible to determine a priori what is called the best solution. If the number of alternatives is small, the selection is made using outranking methods [12]. It is also possible to use the preference for making the optimal choice. For advanced readers and deeper understanding of the issue, further reading is recommended, e.g., [13].

Table 5.2. The Partial Criteria Values for Alternative Scenarios.

Scenario	*TC*	*PC*	*DC*	*L*
A	19974	9.24	1.5	2.2
B	21158	10.5	1.5	1.8
C	20293	8.4	1.4	2.1
D	17541	42	0.8	3.7
E	20977	85	1.0	4.1
F	19147	84	1.1	2.7

A modern trend in the transportation process quality control is the evaluation and control of environmental impact, as a part of overall business process quality management [14]. The most significant environmental impact associated with transportation is the production of greenhouse gases. The largest source of greenhouse gas emissions from human activities is from burning fossil fuels for electricity, heat, and transportation. Detailed emissions calculation on single freight or even parcel may be performed and included into the optimisation goal function. Most publications recommend to calculate the performance indicators for transportation process in particular, i.e., not taking into account other aspects of a business process.

European Standard UNI EN 16258: 2013 defines a common methodology for the calculation and declaration of energy consumption and greenhouse gas (GHG) emissions related to any transport service (freight, passengers or both). It is possible to provide detailed distance-based calculations of energy consumption and relevant greenhouse gas emissions for any transportation mode [15].

Multi-criteria is an integral feature of most real selection problems and requires special analysis methods. The decision-maker needs to study the Pareto principle application, which plays an important role in decision-making, as well as the theory of the relative importance of criteria, goal programming and analytic hierarchy process. The complexity of multimodal transportation processes suggests decision makers in logistics to work with efficient solutions, mainly to capture different aspects. Thus, we recommend to study multimodal transportation processes with a specific focus on multi-criteria optimisation.

5.2.3. Mathematical models in multimodal transportation performance analysis

For the purposes of this chapter we assume the minimal acquaintance with system modelling and appropriate terminology. The term *model* is interpreted as any type of object replacing the research object. This replacement should be useful, i.e., the model provides the information about the object. The type of the model is related to

- object properties (structure, layout, etc.),
- application area (teaching, training, etc.),
- formalisation level (descriptive, mathematical, graphical, simulation, etc.),
- implementation goal (control, forecast, identification, etc.),
- time factor (static, dynamic),
- randomness (deterministic, stochastic).

One can find more detailed approach to model classification as soon as there are lots of interpretations of model concept, types of models, modelling approaches and model application areas. This is true for multimodal transportation as well.

Deterministic models

The deterministic models suppose no random factors under consideration. Known parameter values, rules, thus the behaviour of the model is fully predictable under known initial conditions. Models of this type are used for all types of technical-engineering calculations.

Deterministic optimisation model of cost function is introduced in the previous section. The questions related to the evaluation of costs are not covered in the example. We assume that the carriers are providing information on tariffs and transportation times, thus the costs can be calculated for each alternative. The optimisation model is fully deterministic.

However, in relation to the multimodal transportation we can introduce other deterministic data-based models that are useful for alternative analysis.

To introduce a formal problem statement for route assignment, the road network topology is taken into account. The network description includes origin and destination nodes, intermediate nodes that are relevant to the specific transportation mode, nodes where transportation mode can be changed. The nodes are described with their coordinates, available connections to other nodes, and the length of these connections. It is possible to create a formal network description for the previous example using the graph theory.

We call a graph several nodes (vertexes), some pairs of them are linked by lines (edges). Within the scope of this chapter we consider non-empty graphs. The graph is called connected if each vertex is joined to any other with lines. A loop is a path along the edges of the graph that starts and ends at the same vertex. Also, a graph is called weighted if each edge corresponds to a number (weight). There cannot be two edges connecting the same vertices [16].

In this case only two paths have intermediate nodes, where transportation mode is changed (Fig. 5.2).

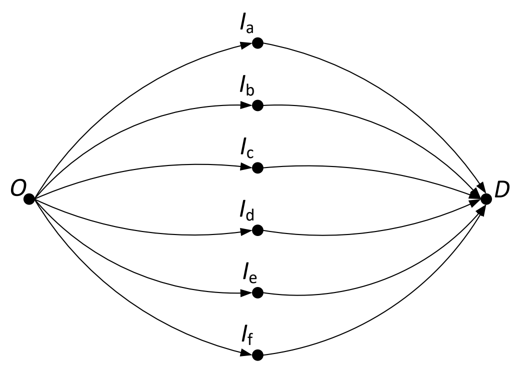


Fig. 5.2. The graph of the example transportation process.

To use the graph theory for the formal description of the transportation task, we need to introduce intermediate nodes for alternatives a, b, and c. These nodes are necessary because in graph theory the nodes can be joined with a single line only. As our process alternatives a, b, and c are unimodal, the introduction of the fictive nodes supposes that the weight of the lines joining the origin and destination points should be assigned to two lines instead of a single line. So the graph used for the formal description is simple and the distances are determined in terms of transportation times. The nodes and links in the transportation network correspond to the vertices and edges of graph, and the distances, or travel times are interpreted as weights of the edges in graph theory. Weight is a numerical value assigned to an edge of a graph.

The multimodal transportation alternative choice problem can be considered with the help of non-empty, final, directed, and weighted, graph G with the set of 8 vertices V, 16 edges E and a set of weights W:

$G = (V, E)$,

$V = \{O, I_a, I_b, I_c, I_d, I_e, I_f, D\}$,

$E = \{\{O, I_a\}, \{I_a, D\}, \{O, I_b\}, \{I_b, D\}, \{O, I_c\}, \{I_c, D\}, \{O, I_d\}, \{I_d, D\}, \{O, I_e\}, \{I_e, D\}, \{O, I_f\}, \{I_f, D\}\}$,

$W = \{2.2, 0, 1.8, 0, 2.1, 0, 2.1, 1.6, 1.5, 2.6, 1.1, 1.6\}$.

The weights of graph edges correspond to the amount of effort needed to travel from one vertex to another. In our example the weights are travel times. In other applications the weights may be distances, fuel consumption, environmental impact and other labels.

In some logistics tasks the edges, vertices and weights are interpreted in a different way. The interpretation is based on the type of the problem to be solved.

There is a task that was not discussed within the example, but it is supposed that the best or shortest route of each alternative is considered. The shortest route to finding the problem solution can be by using a deterministic data-based model.

The best-known deterministic problem in transportation process analysis is a problem of route assignment. Knowing origin and destination points, various routes are analysed. The choice of the feasible route is possible considering travel time, costs, and environmental impact. This type of models is not specific for multimodal transportation, routing tasks are relevant to all types of transportation.

Finding the shortest route is a vital task and is used almost everywhere, not only in transportation. The road transportation routes themselves are evaluated using map and road-based information. These routes may be evaluated by using the graph theory and relevant algorithms. The shortest route is searched between two specified nodes in the graph. We use the term 'shortest' using distances between nodes, but we can formulate the best route finding in terms of fuel consumption, environmental impact, time or other factors of interest.

The most frequently used route optimisation algorithms are:

- Dijkstra algorithm
- Floyd–Warshall algorithm
- Bellman–Ford algorithm

- Brute force search algorithm.

These algorithms are easily executed with a small number of nodes in the graph. As their number increases, the task of finding the shortest path becomes more complex.

The deterministic model can be transformed into stochastic one if deterministic parameters or factor values become random. In our example it is supposed that transportation times are known as constants. Nevertheless, it is obvious that transportation time is subject to uncertainties. The same happens when we do not have perfect real-time information about cost, then the route choice decisions have uncertain factors.

Stochastic models

Most real life problems and tasks include variability. The variability or uncertainty may be of different nature. Some factors may be described as random variables. These factors are introduced into models by using probability functions. The values of such factors are generated by using probability functions and random numbers and are unpredictable. The methods and approaches for data collection, processing and generation are widely used for various types of stochastic simulations. Weather conditions, traffic flow density, and customer order arrivals are examples of unpredictable variabilities that can be described and introduced into models by using probability functions or sampling from standard statistical distributions.

Some types of variabilities are not really unpredictable. The examples of these type of variabilities are scheduled events, such as times of switching the direction of traffic lanes. These types of variability may be introduced into the model with the help of schedule or event log.

The main types of stochastic models are simulation models and Monte-Carlo simulations.

Simulation models are mathematical, algorithmic and dynamic models created using software. These models simulate the processes in real systems and may be created in the frame of diverse approaches. The most known and advanced simulation approaches are considering the following types of systems:

- Discrete state and continuous time (or discrete-event)

- Discrete time and continuous state (system dynamics)

- Discrete rate (combining the discrete-event time advance technique with system dynamics flow intensities)

- Discrete time and state (cell automata)

There are lots of software products on the market that can satisfy the researchers' needs in the area of transportation. However, there are specific software features that can best fit some specific problem. The Institute for Operations Research and the Management Sciences provides a biannual software survey, accompanied by analytical article on the topic [17].

Discrete-event system simulation (DESS) provides the advanced possibilities to simulate processes at micro level of detail. The particular freights, vehicles, lanes and other resources are simulated as entities of interest. Models of this type can be created using data extracted from event logs. The processes of a real system are simulated based on the typical events of this system. The system process is a superposition of all entities processes during a particular time period. Figure 5.3 provides the illustration of the system process.

It is clear from the illustration that even simple processes being put into a common system process are providing a comprehensive common process. The analysis of such a process is based on the concept of resource utilisation by entity and entity delays. Queues and vehicles also are interpreted as multiple-content resources. Modern DESS software provides various types of reports for analysis purposes.

However, it is important to understand what the expected types of results provided by DESS are. As simulation is based on resource utilisation and delays, the generic types of simulation results are:

- resource and the whole system utilisation in terms of contents;

- resource and the whole system utilisation terms of time;

- entity time in resource, entity time in system.

As simulation is mostly used for stochastic systems analysis, the models include variability, and these results are stochastic as well. So when we discuss resource utilisation we suppose that statistical data are obtained.

Resource utilisation in terms of contents shows the average number of entities in this resource. This value is calculated using information about how long the entity was there in the resource, or as weighted average. Usually maximum and minimum values are provided, frequency statistics may be available, as well as the time diagram of entities in the resource.

Resource utilisation in terms of time is estimated as the ratio of time spent working to the time of observation. The average entity time in resource is the same as the average resource operation time. This time is evaluated as time spent working per all entities that used the resource during the time of observation, or as an arithmetic average. Usually utilisation is analysed in more detail: used, blocked or not working because of working schedule.

Typical simulation results time diagrams are shown in Fig. 5.4. The time diagrams of the resource and queue contents of the process are shown. In all diagrams the vertical axis denotes the content volume, and horizontal axis denotes the time value.

After discussing the typical results considered in DESS, we can conclude that for multimodal transportation process simulation, this simulation type is rather appropriate. The level of detail allows to introduce transportation mode change events for consideration, as well as to simulate relevant vehicles and container types, loading and unloading operations and various delays. The results of simulation can support the decision about organisation of multimodal process based on total transportation time, resource utilisation and extended with appropriate technical-economical calculations.

If we return back to our simple example of transportation mode choice from 6 alternatives, we can make it more realistic and useful if we introduce variability. The most realistic way to consider variability is to use statistical information about transportation time.

Example 2

For the example we use alternative f of Example 1 – multimodal transportation by rail and ship using a 20-feet container. The alternative transportation time is considered as random variable L_f and consists of three variables: transportation time by rail t_r, transportation mode change time t_c , and transportation time by ship t_s:

$$L_f = t_r + t_c + t_s \, .$$

In previous example, transportation mode change time t_c is considered as a part of transportation time by ship t_s. For the purpose of the current example it is considered as a particular time. These variable times are random, and we have only guesses about minimum and maximum values of these variables, for

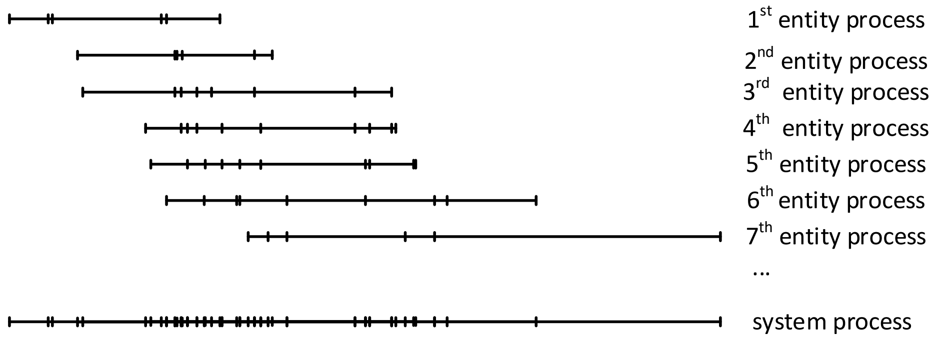


Fig. 5.3. The superposition time diagram of particular entities processes in a system.

example,

$t_r = 1.1 \pm 0.11$; $t_c = 0.4 \pm 0.04$; $t_s = 1.2 \pm 0.12$.

The cases where the real type of random variable is unknown are quite common. In such cases some assumptions about the variables are accepted. For this example the uniform distribution is assumed as a guess about the variable times t_r, t_c, and t_s. The next step is the analysis of the random variable L_f that can be performed with Monte Carlo sampling. As a result the statistical distribution of the random variable L_f is obtained and the risk of the late delivery may be evaluated. The Monte Carlo sampling results are shown in Fig. 5.5.

The Monte Carlo sampling provides the way to evaluate the probability that the total transportation time is longer than expected. For the example the average L_f value is 2.6997, while the probability evaluation of L_f >=2.77 is 0.2490.

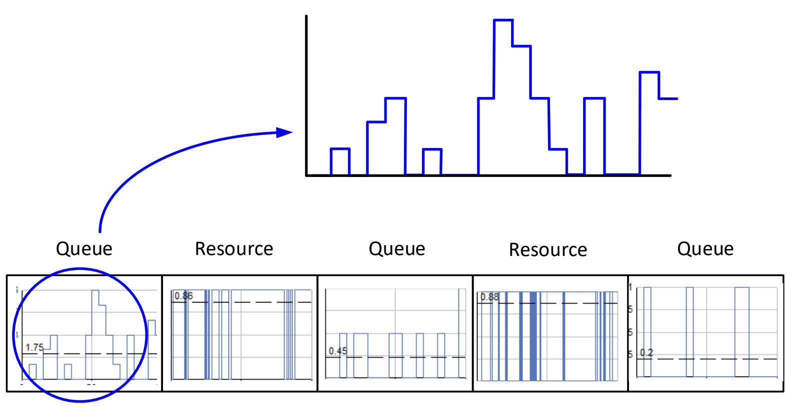


Fig. 5.4. Simulation results time diagrams of DESS software.

The assumptions about the random variables under consideration may be refined when additional information or more observations are obtained about these transportation times and their variability. However, introduction of a stochastic model of the transportation time provides additional utility to the model of transportation time.

The simple example demonstrates the achievement of taking taking into consideration introduction of the stochastic aspect. The main advantage of the stochastic model is that it fully reflects the assumptions made. Stochastic models make it possible to make analytical conclusions in the conditions when deterministic calculations are impossible or insufficient.

5.2.4. Multimodal transportation and the layer model

At the end of the twentieth century, Schoemaker [18] proposed a framework for transportation model analysis. A basic model was formulated for passenger transportation. The basic model consists of three layers: Activities, Transport services and Traffic services. The layers are interconnected and may be interpreted as a transportation logistics system. The Activities layer provides the demand for transportation and receives information about potential supply of demanded services. The intermediate layer – Transport services layer – receives the demand for transportation from Activities and information about potential supply from Traffic services. The Transport services layer provides the demand for transportation aids for Traffic services and information about potential supply of demanded services for Activities.

Transport services may be provided in various modes, including multimodal transportation. The quality of the service depends on the available traffic services. The framework is mostly based on the public transport analysis. The multimodal transportation in the layer model is provided by the transport service integrator. The transport service itself is considered as a combination of service components and means of transport. The most important conclusion of the research is emphasising the role of transport service provider (integrator), as well as considering the activities and functions of transport service provider [18].

The basic layer model is developed and updated by other researchers in more recent times [19].

It should be mentioned that the term 'layer model' in multimodal transportation has also other interpretations, e.g., the topology structure and the different layers of multimodal transport network introducing the layers of road, bike, foot, train and bus networks of the passenger multimodal transportation [20].

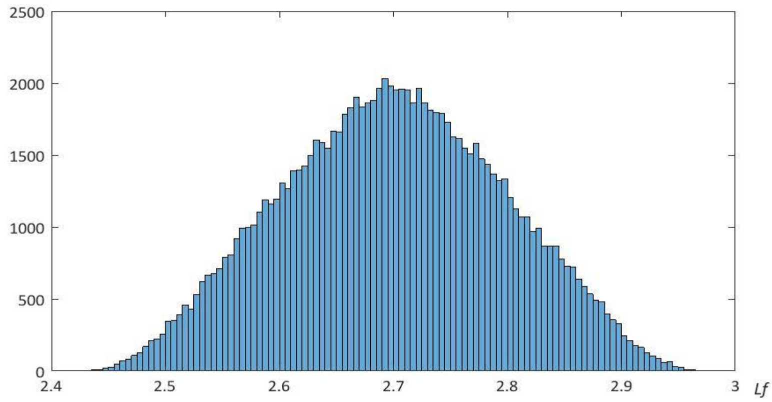


Fig. 5.5. The transportation time L_f histogram from 100000 samples.

There is also a more traditional layer models interpretation related to smart city architecture as layered architecture in transportation context [21].

5.2.5. The future of multimodal mobility

The trend of the development of multimodal transportation is the successful interaction between various transportation modes. The maximum effect of multimodal transportation can be obtained only by improving the overall performance of all stages of the transportation process, where all the stages are working as a single system. The new and smart services, the improvement of the existing and related services, the increased attention to environmental issues are organisational trends. The rapid evolution of the mathematical approaches in transportation organisation tends to make the multimodal transportation more available and preferable.

5.2.6. Transportation travel demand models

Transportation travel demand models (Fig. 5.6) imitate traveller's behaviour during the trip by carrying out a variety of activities and using predefined transportation infrastructure, with the objective to

- forecast the future transportation demand;
- identify potential bottlenecks in the transportation system infrastructure;
- take a decision about the efficiency of transportation solutions before its implementation;

- perform the feasibility studies for investments;
- identify various transportation system characteristics (transportation intensity, intersection geometry, available functionality of the area (social, educational, retail, etc.)) and social-economic measures impact on traveller's choice during the trip.

Transportation travel demand models are expressed in different forms and ways [3], [9], [11]. Some transportation travel demand models represent several interrelated aspects of regional travel behaviour, defining how individuals are involved, where and how transportation activities take place and how individuals get to those activities. Other demand models address more limited tasks, e.g., trips related to freight transportation or analysis of single transportation corridors and street crossings [6]. There are aggregated and disaggregated (or behaviour) transportation travel demand models. In the aggregate models, transportation demand is determined based on average indicators, e.g. for a public transportation demand model it could be several residential and commercial buildings, average transportation costs, etc. In disaggregate demand models, transportation demand is determined based on traveller (individual) level or household level. And both demand models are mainly trip-based travel demand models. The developed transportation travel demand model should be easily adaptable to different territories, cost-effective, developed in a predictable period of time and able to evaluate the behaviour of individual travellers during the trip [11].

5.2.7. Trip generation

Trip generation determines the number of trips attracted by new site or objects (e.g. shopping centre, school, cultural institution, parking lot, etc.). Generated trips are all trips entering or leaving a new object or site, have not crossed research area before the object or site development. Household structure, household income, number of vehicles in households, employment density and other socio-economic indicators are used to determine the number of trips. Indicators such as the length or duration of the trip are not taken into account for the generated trips determination. Generated trips consist of origin and destination trips. Origin trips determine the number of trips per household, for each purpose of the trip, while destination trips determine the number of trips linked to each location. Each type of trips has its determination methods, but as a result the number of origin trips must match the number of destination trips [20].

The cross-classification method and linear regression equations are commonly used to determine the number of generated trips [14]. The cross-classification method identifies specific socio-economic groups with common trip generation characteristics and is used to identify trips at strategic and tactical levels.

The linear regression method is widely used to determine the trips based on historical data in traffic flow impact analysis. Trip generation coefficients or trip generation regression equations are used to determine the number of trips. The decision to use an equation or coefficient depends on the available data quality.

Linear regression equations

The linear regression method determines the number of trips generated by the definite area (dependent variable) as a function of the independent variables (1.1).

$$Y = a + \beta x + \varepsilon, \tag{5.2}$$

where Y is dependent variable; a, β is model parameter; x is independent parameter. The estimated regression equation is used to determine the generated trips, where the generated trips are expressed as the number of trips per X unit,

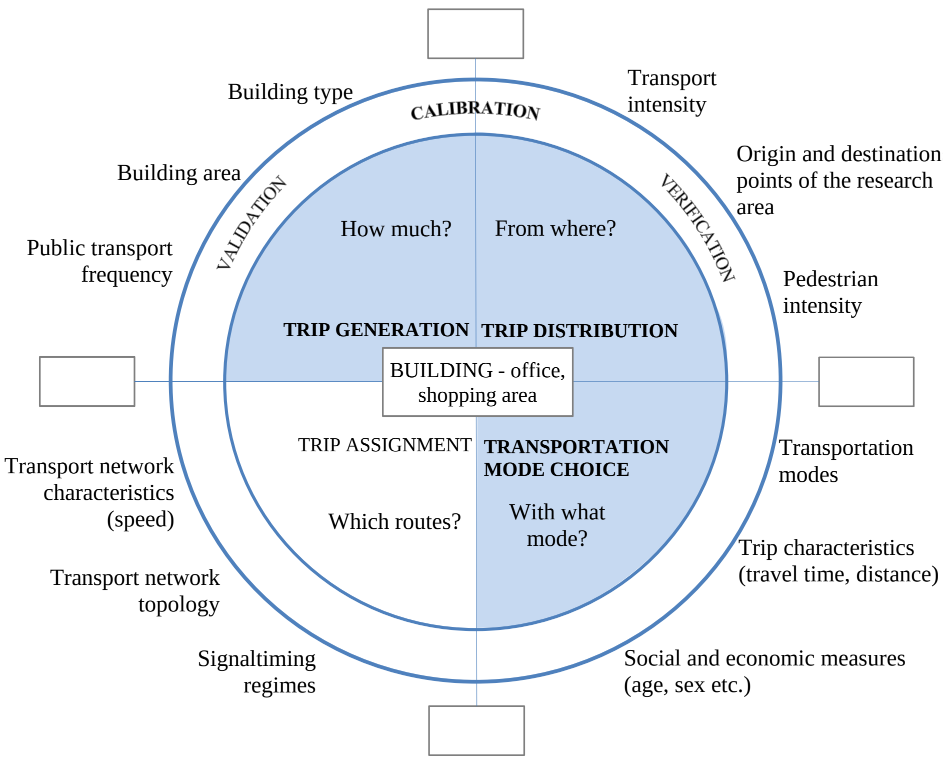


Fig. 5.6. Traditional trip-based transport travel demand model.

where X is a factor that is characterized by the type of land use, such as gross leasable area or number of apartments (1.2), (Fig. 5.7).

$$Y^* = a + b_1x_1 + b_2x_2 + \cdots + b_nx_n, \tag{5.3}$$

where Y^* is the number of generated trips to the household; $x_1, x_2, \dots, x_n$ are population, density, number of dwellings, area, etc.; and $b_1, b_2, \dots, b_n$ are regression coefficients.

There are several information systems for determining the generated trips based on regression equations and average rates:

- TRICS (Trip Rate Information Computer System) [41]
- NZTPD (New Zealand Trip Rate and Parking Database) and RTA (Roads and Traffic Authority of New South Wales) [42]
- ITE Trip Generation [20]

Previous mentioned information systems determine the number of generated trips based on object functionality, location and analysis periods. Some information systems take into account the availability of transport infrastructure around an object. Factors impacting the determination of the generated trips:

- Different transportation intensities between cities, regions and countries.
- Various demographic indicators, living conditions and economic indicators.
- Driver's behaviour (travel speed, the interval between vehicles).

Neural network

Neural networks are widely used as a data analysis methodology in the context of trip generation estimation [2], [19]. Studies by the authors of [16], [39] show that the application of the neural networks provides similar results for determining the number of generated trips compared to regression methods. For example, Worksite Trip Reduction Model [27] neural networks are used to predict the number of trips (related to jobs) when changes are planned in Worksite Demand Management Programs. Linear regression was used as a standard to test the quality of the neural network solution, as well as to analyse variables and determine significant ones. The application of the neural networks showed similar results in comparison with linear regression; however, for neural networks [27] overlearning of the training data set was observed.

5.2.8. Trip distribution

In the trip distribution origin and destination, transportation intensities together with generated trips are converted into origin-destination matrices by using modelling methods. The result of trip distribution is the origin-destination matrix that defines the trips from the origin points to the destination points for all origin-destination matrices pairs. Traditional origin-destination matrices determination (O-D matrix) is based on surveys and interviews, which is an expensive and challenging task (Fig. 5.7). Statistical and dynamic methods, such as the growth factor/Fratar method, the gravity method, the network equilibrium method, the intervening-opportunities method and the distribution-assignment method are used to determine the origin-destination matrices [1], [8], [18] from the collected transportation intensities and to classify them according to the following characteristics of the transportation network [15]:

- Simple transportation network without possibility to choose the routes within transportation network.

- Transportation network without congestion and with a possibility to

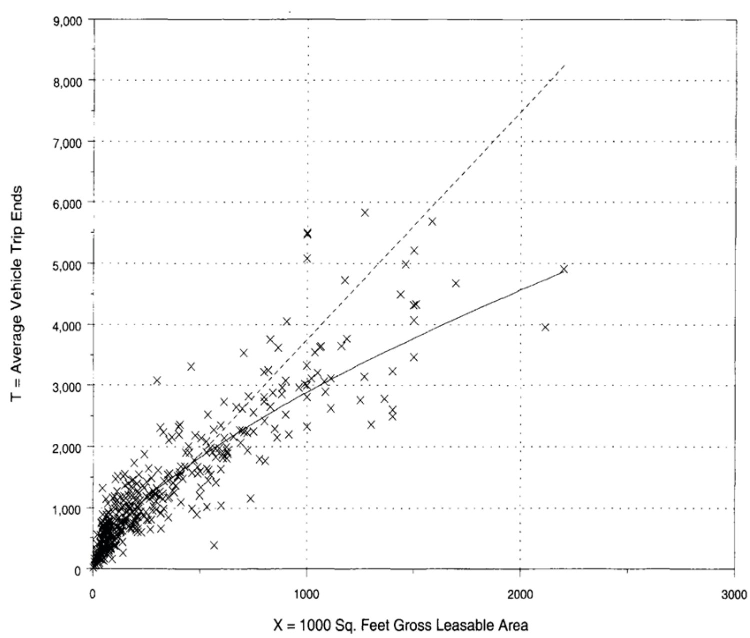


Fig. 5.7. Generated trip determination for retail.

choose the routes.

- Common transportation network with congestion and with a possibility to choose the routes.

The Fratar method proportionally converts the collected trips data to the origin-destination matrix. The method is used to determine O-D matrices for a small simple street network without traveller's choice of route. The disadvantages of the Fratar method are the inability to reflect the trips changes between O-D pairs, changes in traffic costs, such as changes in travel time due to congestions, or changes in transport network structure, such as changes in street intersection geometry. The gravity method is one of the first to be used to determine the O-D matrix from collected transportation data. Later, the authors in [36] generalized the gravity method, following the author's [45] suggestions to combine the gravity and intervening-opportunities methods into one. The next stage in the O-D matrix development was the entropy maximization and information minimization techniques [49], which were used after Wilson's work in 1970. Entropy maximization technique together with other statistical methods – generalized least squares method [12], [23], [29], maximum likelihood method [35], and Bayesian inference method [35], [23] were used to determine the OD matrix for an uncongested street network with proportional road assignment and route selection. For a congested transport network where transport demand exceeds the capacity, the assumption of proportional road assignment with road selection no longer works, as road selection and O-D matrices become interdependent [5]. The Fixed-point and Bi-level approaches can be used to include a path selection process for determining the O-D matrix for a congested street network [10].

Gravity model

The gravity method is based on the assumption that the number of trips between the two origin and destination zones is proportional to the size of each zone and inversely proportional to the distance between the two zones. The gravity method states that trips generated in the i-th zone will be distributed to other j zones according to the relative attractiveness and relative availability of each j zone [13]. The following iterative steps for trip distribution with the gravity method:

- create an O-D matrix with lower costs between any O-D pairs. The cost may be expressed in terms of distance, travel time, waiting time, stopping time, etc.;

- determine the cost function parameters, which represent the trip distance characteristics according to the trip purpose;

- calculate the cost function matrix values for each trip purpose;
- convert origin and destination traffic intensities to the origin-destination matrix; in this way, the initial O-D matrix is determined by trip purpose.

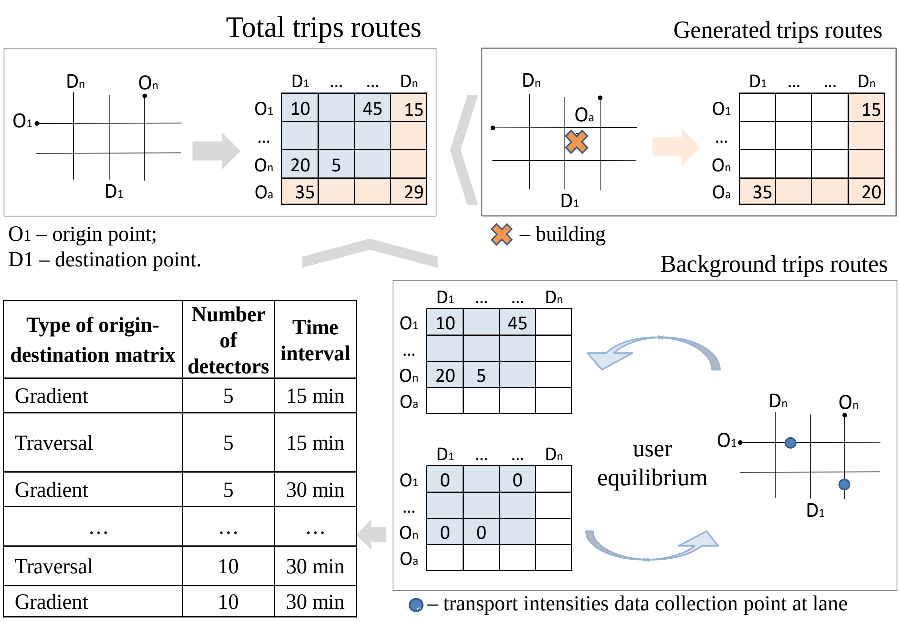
Introduce restrictions in the iterative process. The number of trips in the O-D matrix must be horizontally equal to the number of trips in origin and the number of trips in the O-D matrix vertically must be equal to the number of trips in the destination.

Trip distribution by gravity method [13] is performed according to formula (5.4):

$$T_{ij} = a_i O_i b_j D_j f(c_{ij}), \tag{5.4}$$

where

T_{ij} – trips generated in i zone and attracted in j zone;



Type of origin-destination matrix	Number of detectors	Time interval
Gradient	5	15 min
Traversal	5	15 min
Gradient	5	30 min
…	…	…
Traversal	10	30 min
Gradient	10	30 min

Fig. 5.7. Example of trip distribution determination for a building object.

a_i – balancing factor for trips arranged horizontally in the O-D matrix (origin zone restriction);

O_i – generated trips for i zone;

D_j – attracted trips for j zone;

b_j – balancing factor for trips arranged vertically in the O-D matrix (destination zone restriction);

$f(c_{ij})$ – cost function c_{ij} for trip from origin i zone to destination j zone.

There are several cost functions [13] for the gravity method extension, for example, gravity method with exponential cost function (5.5), capacity cost function (5.6), combined cost function (5.7).

$$f(c_{ij}) = exp\,\{-yc_{ij}\}, \tag{5.5}$$

$$f(c_{ij}) = c_{ij}^{-n}, \tag{5.6}$$

$$f(c_{ij}) = c_{ij}^{-n} exp\,\{-yc_{ij}\}. \tag{5.7}$$

The use of the gravity method requires an extensive calibration of the O-D matrix determination and a multi-step iterative process.

Bi-level approach

The Bi-level approach is one of the commonly used in the trip distribution for congested transportation systems. The approach was proposed by Spiess (1990), and at the upper level, the origin-destination matrix is adjusted based on the initial O-D matrix obtained from the lower level and from the observed transportation intensities (5.14). At a lower level, the trip assignment is performed (the problem of balancing is addressed). The estimation of the O-D matrix begins with the initial O-D matrix estimation, followed by an iterative transport assignment process (with dynamic user equilibrium) and matrix adjustment (with a Kalman filter) until convergence is achieved and a customized origin-destination matrix is obtained.

$$MinF(v(g),\hat{v}) = \frac{1}{2}\{\textstyle\sum_{a \in A}[v(g)_a - \hat{v}_a]^2\}, \tag{5.8}$$

where

$v(g)_a$ – estimated transportation intensity on link a at lower level;

MinF(v(g),v^) – target function (distance between observed and estimated transportation intensity);

g – adjusted O-D matrix;

A – section subset, sections with transportation intensities;

$\hat{v}_a$ – measured transportation intensities on a link.

5.2.9. Mode choice

Mode choice is one of the transportation system components that assesses the usability of different transportation modes (eg., cars, public transport, pedestrians, cyclists, etc.) by describing each available transportation mode and each user type. For example, public transport is chosen by travellers if there are no other options (for example, travellers are too young or old to drive, the social or physical aspects of the traveller do not allow driving) or by comparing several alternatives (e.g., walking vs driving) concluded that this was the most acceptable mode. Trips could involve one or more modes depending on the traveller's destination.

As a result of the transportation mode choice, the origin-destination matrix is developed, in which the number of travellers' transport trips from one zone to another by transport modes available in the research, is determined. Transportation mode choice includes sample selection and appropriate mode choice model selection in traveller's trip behaviour simulation context. Traditional transportation mode choice methods and models, for example, multinomial logit model [7], [17], discriminant analysis, are based on random utility maximization principle and do not include the traveller's transportation mode preferences during the trip. In order machine learning methods for transportation mode choice give more flexible structure ratio between mode categories than the traditional logit models, and do not require prior knowledge about the initial specific model data structure.

Learning tree

The application of the decision tree algorithms in tour-based and activity-based transport travel demand models has shown good performance results [21], [47]. The use of decision trees in transportation mode selection is related to the method's ability to represent events by providing comprehensive information on the transport system parameters that influence mode selection, as well as requiring less input for data calibration compared to discrete choice models. For example, the use of decision trees to select the 'working' destination mode has shown that it is useful in cases where it is important to identify the reasons why travellers choose a particular mode of transport or to identify variables for transport policy-making [46]. The use of a decision tree in the choice of freight transport mode showed good performance, in case there were no complex interactions between the variables [33].

There are a significant number of algorithms that implement decision trees, of which the C4.5 algorithm and the CART (Classification and Regression Tree) algorithm has gained the widest applicability and popularity. Algorithm C4.5 [30] is an improved version of the ID3 algorithm, which adds the ability to work with missing data, improves the tree trimming criterion, and offers a mechanism for retrieving and simplifying the rule and decision tree structure without loss of recognition quality. Algorithm C4.5 uses a systematic approach based on attribute entropy calculations to create a decision tree. The entropy of the attributes is calculated according to formula (5.9):

$$H(C \mid A_i) = \sum_{j=1}^{M_j} p(a_{k,j})[-\sum_{i=1}^{N} p(c_i \mid a_{k,j}) \log_2 p(c_i \mid a_{k,j})], \quad (5.9)$$

where

H(C|Ai) – i-th attribute entropy;

p(ak,j) – probability that attribute Ak has j-th value;

p(ci|ak,j) – probability that results belong to i-th class if attribute Ak has j-th value;

Mk – attribute Aj values j = 1, 2,...Mk;

N – number of partitions i = 1, 2,...N;

K – number of attributes k = 1, 2,...K.

The example of C4.5 decision tree algorithm results for transportation mode selection is shown in Fig. 5.8.

The entropy is calculated for each node, and the nodes are identified by the entropy value, so the tree root has the strongest parameter (with the lowest entropy value), and then in descending order to the stage where the ‘branches’ end with the solution (e.g., car, public transport, pedestrian, cyclist). The C4.5 algorithm allows to create decision trees with an unlimited number of descendants of nodes, however, unlike the CART algorithm, it cannot work with continuous values of targets, thus solving only classification tasks.

5.3. Trip assignment

Trip assignment is the last step in the transportation travel demand model development. The goal of trip assignment is to identify congested routes, determine the travel times, costs and routes within pairs of the origin-destinations matrix. As a result of the trip assignment, transport trips have been assigned into certain directions, modes and routes [20]. When trips are assigned along the street lanes, it is important to ensure the shortest possible travel time between origin and destination, taking into account the transport situation – congestion, traffic intensities, level of service at intersections and other factors impact on the transportation system.

A trip assignment is performed manually or with simulation models, which facilitates the calculation in cases when it is necessary to check several trip assignment scenarios. There are several methods of trip assignment:

- All-or-nothing assignment. It is assumed that there is no congestion in the street network and that the identified routes are determined by the shortest travel time, and the traffic flows are distributed over certain routes. The method is easy to use, does not require high costs, but does not take into account congestion and the capacity of intersections when describing the choice of most travellers, which will result in an unrealistic transport situation per hour of peak traffic in the urban context.

- User Equilibrium assignment. It is a deterministic user equilibrium method that balances traffic volumes and street network's level of service costs, which, in turn, are related to the analysed street sections. The level of service includes many components, such as travel time, travel costs, flow stability. When choosing the user Equilibrium assignment method, it is considered that the user of the transport system is aware of the transport situation, knows the required travel time on alternative routes and, based on this information, chooses the travel route. Mostly the user of the transport system does not have complete information about the transport situation on all roads, then for more information, the use of ITS systems can help.

- Stochastic trip assignment. Stochastic traffic assignment methods assume that several roads may have the same travel time between the origin and destination, and multinomial logistic models are used to determine and predict the behaviour of a transport system participant.

- Dynamic trip assignment. Dynamic assignment, unlike the previous methods mentioned above, allows assessing the choice of route, which

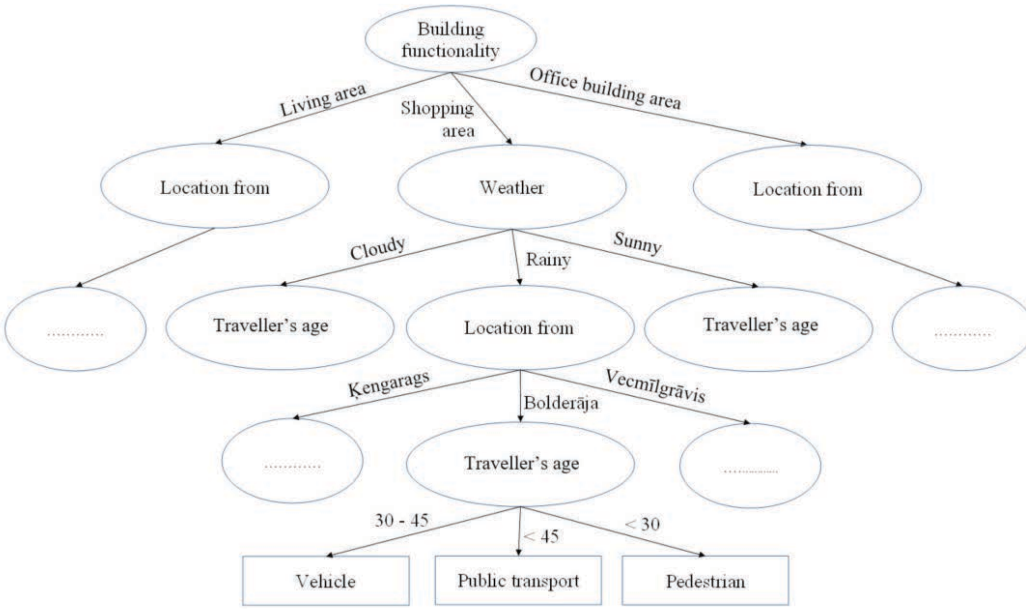


Fig. 5.8. Example of C4.5 decision tree algorithm results.

means that it is possible to change and adjust routes during the trip, taking into account the transport situation on the street network, such as congestion.

- There are several transport simulations tools that provide static and dynamic trip assignment, such as EMME / 2, CUBE, TMODEL, TransCAD, VISSIM / VISSUM, Aimsun, and more. Dynamic assignment implemented in the integrated transport simulation software has been used to adjust transport routes during the trip and to analyse various transport experiments in transport system over-congested situations.

References

1. "Freight transport statistics - modal split," 14 August 2020. [Online]. Available: https://ec.europa.eu/eurostat/statistics-explained/pdfscache/1142.pdf.

2. P. Paul and J. B. Atwater, APICS Dictionary. Fifteenth edition, APICS, 2016.

3. J. Barceló, Ed., Fundamentals of Traffic Simulation, Springer, 2010.

4. S. Robinson, Simulation. The Practice of Model Development and Use. 2nd

Edition, Red Globe Press, 2014.

5. M. Schenk, J. Tolujew и T. Reggelin, «Mesoskopische Simulation von Flusssystemen: algorithmisch steuern und analytisch berechnen» In Beiträge zu einer Theorie der Logistik, P. Nyhuis, Ред., 2008, pp. 463-485.

6. W. M. van der Aalst, Process Mining. Discovery, Conformance and Enhancement of Business Processes, Springer, 2011.

7. Е. Деев и С. Н. Корнилов, «Методика оптимизации мультимодальных контейнерных перевозок,» Современные проблемы транспортного комплекса России. №1(6), pp. 17–20, 2015.

8. D. P. Möller, Introduction to Transportation Analysis, Modeling and Simulation. Computational Foundations and Multimodal Applications, Springer, 2014.

9. B. Zvirgzdiņa and J. Tolujevs, "SIMULATION-BASED METAHEURISTIC OPTIMIZATION OF LOGISTICS SYSTEMS," in Proceedings of the 12th International Conference Reliability and Statistics in Transportation and Communication, Riga, 2012.

10. J. Udomwannakhet, P. Vajarodaya, S. Manicho, K. Kaewfak, J. B. Ruiz and V. Ammarapala, «A Review of Multimodal Transportation Optimization Model,» In 5th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, Bangkok, Thailand, 2018.

11. С. М. Гедрис, Г. И. Оникийчук и К. И. Гедрис, «ВЫБОР СПОСОБА ТРАНСПОРТИРОВКИ ГРУЗОВ КАК ЗАДАЧА МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ» 2016.

12. А. В. Лотов и И. И. Поспелова, Конспект лекций по теории и методам многокритериальной оптимизации, МГУ, 2014.

13. M. Caramia and P. Dell'Olmo, Multi-objective Management in Freight Logistics. Increasing Capacity, Service Level and Safety with Optimization Algorithms, Springer, 2008.

14. "An official website of the European Union," 2014.

15. M. Schmied and W. Knörr, "Calculating GHG emissions for freight forwarding and logistics services in accordance with EN 16258. Terms, Methods, Examples," European Association for Forwarding, Transport, Logistics and Customs Services (CLECAT), 2012.

16. R. Diestel, Graph Theory, Graduate Texts in Mathematics, Springer, 2017.

17. J. J. Swain, "2019 Simulation Software Survey. A half century of simulation product progress . . . and limitless possibilities going forward," ORMS Today, 7 October 2019.

18. T. J. H. Schoemaker, K. Koolstra and P. H. L. Bovy, "Traffic in the 21st century – a scenario analysis for the traffic market in 2030," in The infrastructure playing field in 2030, Delft, University Press, 1990, pp. 175–194.

19. R. van Nes, Design of multimodal transport networks. A hierarchical approach. TRAIL-Thesis Series T2002/5, Delft University Press, 2002.

20. K. Sossoe, Modeling of multimodal transportation systems of large networks. Automatic Control. Engineering, Paris: Université Paris-Est, 2017.

21. L. Deka and C. Mashrur, Transportation Cyber-Physical Systems, Elsevier, 2019.

Chapter 6:

Specifics of applying the project approach to the development of cyber-physical systems for clean transportation

Natalia Morkun
Iryna Zavsiehdashnia

Kryvyi Rih National University

6.1. Specifics of cyber-physical systems development and implementation projects for clean transportation

6.1.1. Project key definitions

Projects and programs are of great importance for different types of entrepreneurship. Due to projects implementation, enterprises can increase their efficiency and revenues, reduce costs, and optimize their business processes in significant way. The results are most tangible in high-tech areas of activity.

For a long time, the term 'project' was referred to the field of construction and architecture and also was defined as a set of design documents. The term described the process of products creation just in the field of material production. At present days, this term has acquired a new stable meaning, associated with the spread of ideas about modern management methods.

Residential and public buildings construction, industrial facilities construction, complex of computer programs development, new technologies and techniques creation, flight to the moon, reforming an existing or creating a new organization, holding an international symposium, preparing a performance, introducing a new specialty in universities, country regions development – are vivid examples of project activities. Recently, innovative scientific and technical projects, related to the development and implementation of cyber-physical systems in various areas, have been classified as the most complicated types of projects.

So what is 'a project'? What do these different approaches have in common?

There is no single generally recognized term of 'a project' in the references, but several definitions of this term can be given [1]:

- 'A project' is a separate enterprise with specific objectives that often include time, cost and quality requirements.
- 'A project' is a temporary measure, organized with a purpose to create certain product or service.
- 'A project' means a set of interrelated activities that are organized with a purpose to achieve, within a given period of time and within a specified budget, tasks with clearly defined objectives.
- 'A project' is a one-time, unique activity or a set of actions. The result of this activity is an achievement of clearly defined objectives for a certain period of time.

6.1.2. The main features of projects

Thus, it is clear from the above definitions that all projects have certain common features. We can distinguish the following features [2]:

- *Obligatory purpose*. Any project must accurately identify its main aim and objectives. We might even note that if there is no aim, then there is no project! A visual presentation of a project aim and its written fixation immediately launch two interrelated processes: the narrowing of purpose-setting as soon as the planning process begins. Just the actual achievement of the purpose determines the completion of the project.

- *Changes*. Project implementation always leads to changes in the subject area in which it is organized. Project implementation is considered to be a purposeful transfer from the existing to some desired state in the subject area.

- *Limited time*. Any project has predetermined start and end dates. Finishing procedure of the project takes place either in case all project objectives are achieved; or it is clear that the main purpose of the project cannot be achieved for some reasons, or there is no more need to achieve the aim of the project.

- *Limited resources*. The amount of resources (finance, staff, equipment, materials) allocated to the project is closely related to the budget and is always finite.

- *Unrepeatable, unique results*. At the end of the project, unique results are produced, which can be represented in the form of products or services. This feature applies to the entire project as a whole.

- *Sequential development*. It means that the development will be implemented in accordance with certain stages and steps. Typically, the content of the project is generally formulated at the early stages of the project and subsequently detailed and refined as the project team gains a clearer and more complete understanding of the project's goals and desired outcomes.

- *Complexity and differentiation*. the complexity of the project means the need to take into account all internal and external factors that directly or indirectly affect the outcome of the project. At the same time, each project has a well-defined scope of its subject area and should be separated from other projects.

- *Specific project organization*. Most large projects cannot be carried out within existing organizational structures and require the creation of a specific organizational structure for the duration of the project, while for small and simple projects the creation of a special organization is not required or even justified.

6.1.3. Project success and failure factors

There are many factors that contribute to the successful completion of a project, but there are also many that potentially hinder this.

Sufficient time and resources, especially on urgent projects, and well-defined goals contribute, especially if the project team has experience with unclear goals or misunderstandings that arise from confusion in the early stages of the project.

Based on their research, Sam Elbeik and Mark Thomas pointed out ten factors that have been identified by managers of international organizations as critical to the project success [3]:

- realistic and clear outcomes;
- precise planning and control;
- high qualification of the project manager;
- good administrative support;
- sufficient time and resources;
- fulfillment of their obligations by all participants;
- wide involvement of consumers (of the project results);
- effective communications;
- good organization and structure of the project;
- possibility to stop the project implementation.

Elbeik and Thomas also reviewed reports on several projects and found the following common shortcomings [3]:

- the team had doubts about the goals of the project;

- the team was not sure what should be done;
- by the end of the project the goals were only partially achieved;
- works were performed later than set in the schedule;
- the planned budget was exceeded.

All of these shortcomings have led to the failure or termination of many projects. Problems related to the requirements of potential users of the results of the project emerged frequently: either these users were not asked at all about their needs, or training was not provided for those who had to work in a new way.

If the project is supposed to lead to the successful implementation of a new way of working, the issues of employee training need to be considered in a proper way.

6.1.4. Specifics of cyber-physical systems development and implementation projects for clean transportation

Cyber-physical systems are a rather complex concept. To date, they have not received an unambiguous and generally accepted definition, since these systems are at the intersection of several areas at once and, depending on the implementation, are capable of affecting a variety of aspects of activities.

Their main common feature is a very close interaction between computational processes and physical processes. Therefore, we can say that a cyber-physical system is a complex system of computational and physical elements that constantly receives data from the environment and uses them to further optimize the control processes. Cyber-physical systems were created based on the integration of a number of previous technologies. The main technologies for supporting cyber-physical systems include Internet of Things (IoT) technology, embedded systems, Ubiquitous computing, and special technologies for network exchange [4].

In the current paper, a more extended definition of cyber-physical systems designing project is represented – this is a separate category of innovative IT projects associated with the assessment, selection, development or modernization, adaptation, customization, configuration, implementation, and integration of cyber-physical systems in a particular industry.

The main outcome of cyber-physical systems development projects is to increase the efficiency of management of all aspects of the activity of any object by using intelligent internal network modelling.

As mentioned above, the main principle of cyber-physical systems operation is the significant relationship between their physical and computational elements. The 'brain' of the system in the form of artificial intelligence and other technologies receives data from sensors in the real world, analyses this data and uses it to further control physical elements. Thanks to this interaction, the cyber-physical system is able to work effectively in changing conditions, as an analogue of the human body or a modern company that analyses the market situation in order to develop the product that is in demand exactly right now. Moreover, the cycle 'manage4ment – data acquisition – data processing – management' should produce positive results and create new value every time the system works well [4].

Regarding the transport industry, there are 2 types of cyber-physical systems: internal systems inside a mobile object and external systems that unite a complex of mobile objects, for example, a metropolis traffic system. Systems of the first type solve problems of separate object management in a difficult dynamically changing situation. An example is autonomous transport management systems, which receive real-time information from road users and road infrastructure to avoid accidents and select the optimal route to the destination, taking into account the situation on the road. Systems of the second type solve the problem of traffic control in the system of traffic flows.

One of the brightest projects in the area of 'clean transportation' is the development of electric vehicles Tesla with a functional system 'Autopilot', which is based on neural networks.

6.1.5. Clean Transportation cyber-physical systems project designing peculiarities

1. High innovation level and technological complexity. The projects implement modern information technologies, automation tools, and specific equipment. So, this factor requires involvement of highly qualified specialists in the project implementation processes.

2. Project effectiveness may not always be expressed directly in monetary terms. The effectiveness of the use of cyber-physical systems is measured by such 'soft' indicators as the speed and quality of decision-making, the level of automation of managing large amounts of data, etc. You can apply indicators that are not directly related to monetary, for example, decreasing the number of traffic accidents as a result of the use of intelligent systems or reducing the time spent in traffic jams [5], [6].

3. Projects for creating cyber-physical systems are always associated with high risk. Risks may relate to deadlines, exceeding planned labour intensity, and the planned results for these types of projects are especially high. These projects are characterized by high intensity combined with deep granularity of network schedules and iterative work executions. Usually, detailing of labour resources is required for a specific performer. Non-labour resources and materials are tracked much less frequently.

4. Most cyber-physical systems projects require colossal budgets. In large companies, the scale of project activities in this IT field is measured in millions of dollars, and, moreover, the implementation of new projects is ongoing. The development of cyber-physical infrastructure in growing companies requires large and regular investments. Large budgets, in turn, mean a greater level of responsibility and, accordingly, a greater level of competence of those people who manage these projects.

5. Specific distinctions at the level of ideology of the customer and the contractor.

The customers are, as usual, the business representatives, and the contractors are IT specialists, researchers. This aspect determines the existence of certain difficulties in identifying the requirements, project expectations, as well as technical specifications. Concerning IT projects, in most cases the management of development and implementation processes is not carried out by the business leader but transferred to the IT specialists. As a result, communication conflicts, differences of expectations, requirements and results are possible between them. The way to avoid such a 'misunderstanding' between business and IT has to use methods of identifying and recording project outcomes expectations, as well as to collect all project requirements at all levels of users properly.

In current chapter, the project management process will be considered on the conditional example of the development of cyber-physical information system in the format of an electric vehicle autopilot system. According to the project management methodology, there are certain stable processes like calendar planning, budget planning, etc. They do not depend on the type of project. But there are certain aspects that are important and the specifics of the project should be taken into account. Particular attention will be paid to the processes of substantiation of the concept of cyber-physical systems development projects and quality management planning of such type of projects because these issues are unique and innovative.

The creation of a cyber-physical system begins from the moment of the first negotiations between the Customer and the potential Contractor and may never end, since good and useful systems are constantly being improved and developed.

6.1.6. The main stages of a cyber-physical system development

6.1.6.1. Preliminary stage

At this stage, it is necessary to realize the main goals and objectives of the future project. To this end, key representatives of the Customer and the Contractor organize meetings at which they discuss the concept of a cyber-physical system, key technical points, terms and volumes of work performed, as well as cost and sources of funding. The result of the preliminary stage, in addition to the agreed terms of the future contract, should be the first and most fundamental project document – the Project Statute.

The Project Statute defines the following fundamental points related to the development and implementation of the cyber-physical system [5]:

- A brief description of the project, goals and objectives of creating a cyber-physical system.
- General description of the work content.
- Project boundaries: terms, budget, list of automation objects.
- Product description: list of supplied hardware and software, type and number of licenses, etc.
- The organizational structure of the project: the list and roles of the project team members on the part of the Contractor and the Customer, their responsibilities and duties, the project document management system.
- Main stages of development and implementation of the cyber-physical system, an enlarged schedule for their implementation.
- The most significant risks of non-fulfillment of project obligations, as well as ways of risks minimization.

In other words, the Project Statute is precisely the document that is developed by the project manager together with the main members of the project team, approved by the management of the Contractor and the Customer and should not be adjusted during the entire time of creating the cyber-physical system.

The completion of the preliminary stage can be considered the moment when a contract for services for the development and implementation of the cyber-physical system is signed and the project charter is approved.

6.1.6.2. Requirements capture

Up to this point, all the key figures – the project participants – have been identified, and nothing prevents the project team from starting to gather and approve the requirements for the future cyber-physical system. Contractor's representatives communicate with the future users and administrators of the system, as well as with their managers. In the course of the survey, not only the requirements and wishes for the implemented solution are systematized, but also the documentation is analyzed, which should become the source of the initial data of the system, or the formation of which should be automated as a result.

The result of this stage should be the appearance of technical specifications for the development and implementation of the cyber-physical system. The terms of reference should be based on the terms of the contract and the requirements set out in the Statute of the project and contain the following sections:

- Purpose and goals of the system.
- Description of the system object and the main automated processes.
- System requirements: structure requirements, functions (tasks) solved by the system; requirements for technical and organizational support, requirements for reliability, safety, etc. [6]
- Composition and content of work on the creation of a cyber-physical system.
- Procedures for control and acceptance of work results.
- Requirements for the scope of work on the preparation of the automation object for launching the cyber-physical system into operation.
- Requirements for the composition of design and user documentation.

The approval of the Terms of Reference by the Customer means the completion of the stage of requirements capture. In some cases, the Customer already has a technical assignment (included in the tender documentation) before the start of work on the project. In this case, the results of the survey and requirements capture are recorded in particular technical tasks, which detail and concretize the general requirements for the cyber-physical system presented in the initial Terms of Reference.

6.1.6.3. Design stage

At this stage, all scenarios related to the development and implementation of a cyber-physical system are designed in detail by the efforts of the Contractor. This is implemented in accordance with the conditions of the information environment (system landscape) of the Customer, as well as in accordance with the requirements for the integration of the system being created with other software products already available and operated by the Customer. The result of the design stage should be the design of the following sections of the technical (conceptual) project [4]:

- architecture of the cyber-physical system;
- description of the structures of the cyber-physical storage (database);
- design solutions, presented by a detailed description of automation scenarios for all affected by the implementation of the system in business processes or technological processes;
- scenarios for integration of the created cyber-physical system with external software products;
- sources of initial data and options for the initial content of the system;
- the concept of differentiating access rights to the data based on user roles, which determine their permissions;
- end-users' training concept.

6.1.6.4. Implementation stage

This is the stage of implementation of all requirements for the cyber-physical system which has been set out in the Terms of Reference and Technical Design. At this stage, the Project team develops all the necessary software components, creates database structures, installs, configures and tests all components of the cyber-physical system on its basis, simulates integration scenarios, and so on. The completion of the implementation stage is confirmed by the appearance of such project documents as the Guide for installing and configuring a cyber-physical system, a program and methodology for testing the system, as well as a database template and a register of all completed software.

6.1.6.5 Preparing a cyber-physical system for implementation

All tasks of this stage are carried out on the basis of the equipment or the Customer's enterprise and include: the installation and configuration of all system components in the Customer's information environment, preliminary testing, development of user documentation, user training, loading initial data, testing the system in accordance with the program and test methodology and other preparatory tasks.

A procedure for the cyber-physical system implementation must be developed and approved by the time all preparatory work is completed. The Project Statute, in particular, should define users and their roles in the system in accordance with their job responsibilities.

6.1.6.6. Pilot industrial implementation

This is the last stage of the development and initial implementation of the cyber-physical system process. The task of the stage is to conduct a trial operation of the system successfully for a certain time, and the goal is to confirm that the cyber-physical system has been created. This will determine the level of achievement of the planned project result.

At this stage, users start operating the cyber-physical system in accordance with the regulations developed at the previous stage. During the pilot implementation, errors are recorded and the necessary modifications are agreed. The Contractor eliminates errors, makes improvements and generates a protocol on the successful completion of pilot operation, provided that the system begins to function in accordance with all the requirements presented to it earlier.

As a rule, the contract for the design of the cyber-physical system ends with the end of the pilot implementation period. The system goes into full operation mode. And the Contractor, if the Customer is interested in this, makes a separate agreement for its support for the period established by the terms of the agreement.

6.1.6.7. Maintenance and further development of the cyber-physical system.

As a result of operation, it may be discovered that some requirements for the created cyber-physical system contain inaccuracies and therefore require a different formulation or additions, and the system itself needs to be improved. Not every Customer company has staff who are able to introduce all changes into the operation of the system dictated by the real situation independently. Therefore, they make a separate agreement with the Contractor for the maintenance of the cyber-physical system.

The cyber-physical system Users start communication with the support service staff, who accept requests for improving the functionality and eliminating defects, submitting requests for work and periodically informing users about the status of their request. The list of possible improvements as well as the procedure for processing applications are determined by the terms of the contract. If there is a need for work that does not fit into the essence of the support contract, the Customer and the Contractor sign a separate agreement. This type of agreement is referred to works on modernization and development of the operated cyber-physical system.

The main difficulties in the implementation of cyber-physical systems in clean transport systems:

- insufficient formalization of management processes at the enterprise;
- lack of full understanding of the mechanisms of implementation of decisions and how performers work;
- the need to reorganize the enterprise in accordance with the implemented cyber-physical system;
- the need to change the technology of the business process;
- complexity of training process;
- the need to form a qualified project team, which includes employees of the enterprise and one of the high-ranking managers of the enterprise, interested in the implementation of the system.

Factors for a successful implementation of a cyber-physical system:

- participation of company executives in implementation;
- availability and compliance with the implementation plan;
- managers set clear goals and requirements for the project;
- participation in the implementation of the Customer company's specialists;

- quality of the cyber-physical system and the team of the solution provider;
- re-engineering of business processes prior to the introduction of the cyber-physical system;
- the enterprise has a developed strategy for the development of directions for the use of cyber-physical systems.

6.2. Cyber-physical system implementation project life cycle concept

Each project, from the emergence of an idea to its full completion, goes through a series of successive phases of development, the full set of which composes its life cycle.

The life cycle of a project is a set of sequential phases, the number and composition of which is determined by the needs of the project management organization or organizations participating in the project. The life cycle of a project has a specific start and end point that can be tied to a timeline.

The project life cycle concept is one of the central concepts used in project management. Life cycle is essential in order to

- determine the beginning and end of the project, as well as to determine its duration;
- form the structure of the project and establish a list of works;
- determine the dynamics of expenses and employment of personnel;
- establish the milestones of the project, which are used for better control and management.

6.2.1. Structure of the project life cycle

The life cycle of a project can be divided into phases, phases – into milestones and stages.

There is no universal approach to dividing a project into phases, the main thing is that such a division exhibits some milestones, the passage of which marks the achievement of one or more project results and provides additional information for assessing possible directions of its development.

The main general phases of the project are:

- Project concept. The main content of this phase is the pre-project analysis, that is, the preliminary formulation of the project objectives.
- Project development. First of all, this is the analysis of the project and the search for contractors, contract tenders, main tenders.
- Project implementation includes detailed design, creation of the project object (cyber-physical systems), trial operation.
- Completion of the project. The main task of this phase is to put the project object into operation.

6.2.2. Life cycle of cyber-physical systems development and implementation projects

Project managers break the project life cycle into stages in different ways. More often, in cyber-physical systems development and implementation projects the following stages are distinguished [1]:

- awareness of the need for a cyber-physical system,
- formulation of requirements,
- system modelling,
- equipment selection,
- programming,
- testing,
- operational support.

6.2.3. Model of a 6-stage cyber-physical system project life cycle

However, the most traditional is the division of the project into six major stages: definition of content (concept), planning, modelling, implementation, close-out and evaluation (in accordance with the classical approach).

Each stage ends with the achievement of certain intermediate results:

- Stage 1 – defining the scope of the project – ends with the submission of an agreed summary of the project, which, however, can be adjusted in the next stages.

- Stage 2 – project planning – ends with the presentation of the project plan; it can also be subject to various adjustments in the next stages.

- Stage 3 – modeling of a cyber-physical system – ends with the development of a mathematical model of the cyber-physical system.

- Stage 4 – project implementation – ends with the receipt of the initial project results.

- Stage 5 – close-out of the project – ends with the execution of reports on the project implementation.

- Stage 6 – project evaluation – ends with the assessment of the implementation of key project indicators.

In many organizations, the evaluation is carried out at the end of each phase of the project to verify that the project is on track. As a result of such an assessment, decisions can be made to change plans and use additional measures to improve performance or even terminate the project.

Figure 6.1 shows that the project life cycle is not linear. During the life cycle, there can be a return to the previously passed stages for making various adjustments and improvements. Moreover, in the process of project implementation, ideas about its purpose may change. This is especially true for the projects of creating cyber-physical systems, including clean transportation. It is typical for such type of projects to set the solution of existing problems as the goal, rather than material tangible outputs. Projects are not carried out in a vacuum, but in rapidly changing contexts. Thus, the impact of the changing environment on the project life cycle must be managed. Flexibility is one of the key attributes of successful project management.

6.2.4. 8-stage project management model

Another approach to representing the life cycle is based on the classic six-step model developed by Elbeik and Thomas [3]. The eight-stage project management model also assumes the division of the process into stages, but unlike the previous model, it allows the simultaneous passage of these stages (Fig. 6.2). In particular, according to this model, communications occur throughout the entire period of the project. It is also assumed that team building, leadership and motivation begins immediately after the project scope is defined and continues until the project is completed.

At the stage of defining the content, a comprehensive discussion of the project is carried out with all stakeholders and the main objectives of the project are determined. At this stage, costs and terms are also estimated and, most often, an assessment of the possibility of implementing the project is carried out. The stage is considered completed after the project summary has been written and agreed.

At the second stage, a preliminary project plan is developed. Planning is an ongoing process, as the plan is the basis for reviewing the progress of the project and making corrections when necessary.

The third stage begins with determining the composition of the team of employees who will participate in the project. Project team members are usually involved in the development of the plan and often make a significant contribution based on their expertise and experience. Team building, leadership and motivation continue until the end of the project.

The fourth and fifth stages combine work on the creation of the cyber-physical system – description and mathematical modeling, as well as direct development.

After the preliminary project plan has been drawn up, work activities begin to obtain the project results. Execution occurs at the stage called in this model

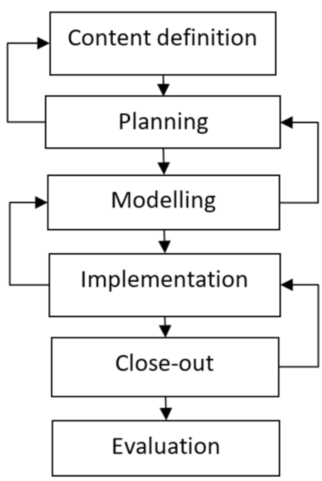


Fig. 6.1. Project life cycle of CybPhys system.

a 'management stage'. At this stage, the actual results of the tasks are tracked and compared with the planned results to assess the progress of the project. Monitoring is very important to ensure that the set goals are achieved on time and within the agreed budget. In this stage, regular reviews of the plan are usually carried out to identify and solve problems and make necessary adjustments.

Communication should occur continuously both within the project team and between the project team and other stakeholders, including all people contributing to the intended outputs. Communication is partly carried out in accordance with formal reporting procedures, but mostly informal communication.

The last stage is the evaluation and completion of the project. Evaluation is carried out in order to find out whether all the intended outputs of the project have been achieved. Evaluation is also important because it collects useful information about the processes involved in the implementation of the project and learned lessons that can be used in the future.

The project completion process must be managed to ensure:

- completion of all unfinished tasks;
- termination of all project-related work;
- accounting for all resources, including those remaining unused and subject to transfer to other departments of the organization or sale.

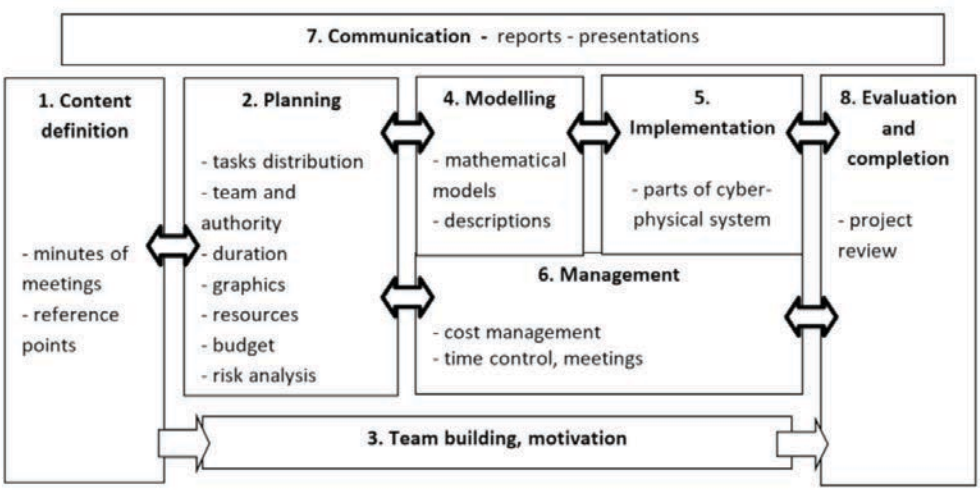


Fig 6.2. Eight-stage model of cyber-physic system project management

Note that neither the project life cycle model nor the 8-stage project management model is a complete project planning scheme; projects are diverse and must be planned in different ways. However, both models help to understand the following important things:

- each of these models can serve as a useful model for developing a structured approach to project management;

- both models are implicitly based on the assumption that the project can be built to a predetermined sequence of actions.

In practice, however, projects often go through the planning, implementation, and evaluation phases many times. Moreover, many projects start in the conditions of absence of the necessary information, which becomes available later and causes changes in the basic assumptions of the project.

It is necessary to represent the project planning as an ongoing process, not a one-time activity, the results of which can be used without changes until the completion of the project.

The life cycle idea is associated with the idea of birth, growth, maturity, aging and death. We are talking about the 'life' of a project, meaning this sequence of life stages. As the project develops, it acquires its own history, influences the way how the project participants work and what they consider possible. Numerous previous decisions and actions taken limit flexibility. When a project begins to deliver the expected results, a vague sense of approaching the end may well arise. If the team enjoyed participating in the project, completing the project can create a sense of loss. However, if a project is successful, it will definitely leave behind something valuable and lasting.

6.3. Project concept definition for the development of cyber-physical systems for clean transportation

6.3.1. The essence of the project concept

Defining a project concept is the first phase of a project's life cycle. The phase includes a preliminary survey and collection of initial data for the project; identifying the need for project results; development and comparative assessment of alternatives; presentation of the concept, its approbation and examination; concept approval.

Projects arise with the aim of meeting human needs. The need appears and is recognized, and the management determines whether this need has to be satisfied. If the need is recognized as significant, the project is organized in order to satisfy this need. Thus, needs are the main driving force behind projects. This defining aspect of needs makes them very important for project management. Needs affect the entire content of the project. If at the very beginning the management does not understand the needs and their implications, if the needs are not correctly formulated or if they are mistakenly accepted to meet the inappropriate needs, then the project has a bad start and it is safe to say that it will be unsuccessful.

6.3.2 Three phases of determining the need for a project. Reasons for initiating projects of cyber-physical systems for clean transportation

There are three phases of identifying needs [2]:

- The emergence of needs – all sectors of the organization are constantly faced with the emergence of new needs in the environment that changes. Needs can arise both inside and outside the organization. Stakeholders should try to anticipate needs ahead of time in order to respond proactively.

- Needs recognition is an awareness of the needs that lead organizations to leverage existing data and experiences both internally and externally and to gather additional information for discussion with stakeholders and to assess the impact of changes that occur. From the point of view of project management, the main task in this phase is the transformation of emerging needs into goals that will begin to determine the results of the project.

- Formulation of needs – in this phase, the understanding of new needs is clarified through a more accurate description of their characteristic features. This allows managers to determine the most appropriate ways to meet the needs, provides an opportunity for stakeholders to participate in the development process, and leads to a clear statement of what must be done or provided to meet the need, that is, the formulation of the project.

Insufficient precise definition of needs can cause the following project problems:

- fuzzy goals;

- unrealistically wide scale;

- solution of incorrectly posed problems;
- conflicting goals of change for people, systems and organizations;
- loss of time on solving tasks that are not included in the scope of responsibilities of the staff, may be optional or impracticable.

6.3.3. The structure of the project concept for the development of cyber-physical systems for clean transportation

The concept of the cyber-physical systems project for clean transportation should include [2]:

- Purpose of the project.
- Main objectives describing the criteria for quality and success.
- Description of the planned results (requirements for cyber-physical system).
- List of project stakeholders and assessment of their interests.
- Assessment of costs and benefits of the project.
- Main risks of the project.

The purpose of the project is a broad concept and can be correlated with the mission and values of the organization, while the purpose of the project defines more precisely what they seek to achieve by implementing the project, and how to determine its success.

The SMART principle is often applied to project aims [1]. The aim of the project should be:

- Specific – project managers have to imagine what they want to achieve through the project implementation clearly.
- Measurable – criteria should be developed to measure the degree of achievement of the goal.
- Achievable – managers have to be confident of the possibility of achieving the goal in the existing environment and with available resources.

- Realistic – no one should try to achieve the impossible.
- Time-bound – the timeframe for achieving the goal should be determined by real needs.

Clarity of purpose is important in order to clearly understand what exactly needs to be done. It will also establish whether the goals will be achieved or not. If the project manager has set clear goals at the initial stage of the project, then he has a certain system of views on the final result. Based on the goals set, the project is structured so that it can be effectively monitored and managed.

Description of the desired project outcomes for cyber-physical systems development projects are the requirements for the system itself.

The cyber-physical system must meet the following requirements [7]:

- Possibility of flexible system expansion.
- Ability to work together with various software products: software, software for information support of subject areas; communication software; collaborative software; software for operational analysis of information and decision support; project management software and other supporting products.
- The system must be multifunctional and meet the needs of all users.
- Modular principle of system construction with operatively independent functional blocks.
- The system must be able to migrate from platform to platform.
- Support for distributed information processing technologies.
- Ensuring security using various methods of control and differentiation of access to information resources.
- Support for multilevel electronic archiving technologies.
- High performance characteristics (ease of administration, ergonomics, user-friendly interface, etc.).

The aim and description of the results of the project implementation must be coordinated with the key indicators: budget, time, quality. These indicators must be balanced for successful project management. A successfully completed project is completed on time, within budget and meets all quality requirements.

Each person who is in the organization or outside of it, and at the same time has or may have a legitimate interest in the project and its results, is an interested person – a stakeholder. It is necessary to identify these people or groups in order to meet their expectations and manage the developmental impact of the project they are likely to want to do.

Project participants and stakeholders are the main element of its structure, so they ensure the implementation of the project concept. Depending on the type of project, on its scale and complexity, as well as on the structure of the life cycle, from one to several dozen organizations can take part in the implementation of the project. Each of them has its own functions, the degree of participation in the project and the measure of responsibility for its fate. All these organizations or individuals, depending on the functions they perform, are usually united into specific groups (categories) of project participants.

The project must have an initiator or author of the main idea.

The main project participant is the customer – the future owner and / or user of the project results. The customer can be either an individual or a legal entity. At the same time, the customer can be either one single organization or several that have combined efforts, interests and capital to implement the project and use its results.

An equally important role belongs to the investor – the organization or the individual who invests in the project.

In projects for the development and implementation of cyber-physical systems, as a rule, an external consultant organization is involved. This company performs the design work or directly develops the cyber-physical system.

The rights to use the scientific and technical achievements necessary in the project are provided (usually on commercial terms) by the organization – the licensor, which can be a legal entity or individual – the owner of licenses and know-how.

A special place in the implementation of the project is occupied by the project manager (manager). This is a legal entity or individual to whom the customer and/or investor delegate the authority to manage the project: planning, monitoring and coordinating the work of other participants.

The project manager leads the team. Project team is a specific organizational structure created for the duration of a project in order to effectively achieve its goals.

An important part of the project concept is the allocation of necessary resources and the assessment of expected benefits of the project. At the conceptual stage, the costs and benefits are not financially assessed. It is only necessary to single out and classify them, as well as to determine the sources of project financing.

In this chapter, we will consider the main classification of the costs and benefits of a cyber-physical system implementation project.

For cyber-physical systems implementation projects, costs fall into two main categories:

- development costs – costs that arise between the start of the project and the moment when the information system is put into commercial operation;

- operating expenses – expenses associated with the operation and maintenance of the corporate information system.

As for benefits, in project management they are divided into:

- certain benefits – direct, visible and measurable benefits, of course based on cash flows entering the organization or not going through the organization;

- uncertain benefits – indirect, uncertain, and less easily financially measurable benefits.

Sources of financing for the implementation of cyber-physical systems:

- Own funds – own capital, funds of own funds – depreciation, repair funds, production development fund, reserve fund, etc.

- Borrowed funds – bank loans, investments of external stakeholders. When using borrowed funds, it is necessary to take into account the fee for the use of funds (bank interest, interest of the borrower, share in the profit from the project).

An important element of project management is risk management.

Project risk is an uncertain event or condition that, if it occurs, has a positive or negative impact on at least one of the project objectives, such as timing, cost, content, or quality.

The ability to manage risks is very important in the project management process, as it allows management to make more informed estimates of the time and effort required to implement the project. This involves assessing the risks associated with the fact that something unplanned may happen, for example, a possible delay in work schedule due to the illness of workers or the lack of the necessary materials at the right time. Risks in a project can be both internal, that is, arising within the project, and external, arising from the project environment.

The main general risk categories are:

- Material risks associated with the possibility of loss or damage to information, equipment due to an accident, fire or natural disaster.

- Technical risks that arise when systems do not work or do not work well enough to achieve the expected results.

- HR risks arising when key employees are unable to participate in theproject, for example due to illness, career changes or a strike.

- Socio-political risks arising, for example, when the project loses support as a result of a change in government cabinet, changes in the policy of the organization's top management, or protests from the public, the media, service users or staff.

- Legal risks arising when there is a threat of legal action due to the fact that some aspects of the project may be considered illegal, or due to the possibility of claims for compensation for damage to external organizations in the event of unsuccessful progress.

The main risks of a cyber-physical systems implementation project are as follows:

- Moral unreadiness of the company for implementation is the main risk that stands in the way of successful completion of the project. If the need for implementation is not realized by the management of the company and the heads of the most important departments, the project will be perceived as imposed from above. Naturally, no one will take the implementation process seriously.

- Violation of deadlines. The risk is closely related to the previous one: the lack of real support for the project from the management, misunderstanding of goals by employees, poor management within the company cause non-fulfillment or untimely fulfillment of project tasks. As a result, deadlines are missed. In addition to these reasons, the deadline may be the result of an illiterately developed project or the adoption of populist decisions like 'Hand over by November 7' to please someone's interests, as well as underfunding of the project.

- Budget overruns. There are times when, after the approval of the project documentation, various additional requirements and desires arise. This must be suppressed at the root; all superstructures should be left for the future. It should be remembered that 20 % of the capabilities of the corporate information system provide 80 % of its effectiveness. Overruns are also possible with an illiterate budget. No one denies the need to minimize project costs, but in everything you need to know when to stop.

- Lack of resources. At the other extreme, when approving the project, the management signs the estimate, rather than taking care of the sources of funds. Such a project is doomed to failure: no equipment supplier or consultant will work for free for the fact that the company did not take care of financing.

- The information obtained as a result of the formulation of the project concept is further used to assess the investment attractiveness of the project, in planning project actions and in making project decisions.

6.3.4 The main factors of success and failure of cyber-physical systems implementation projects for clean transportation

At the end, we present a set of factors that are prerequisites for the success and failure of projects for the implementation of cyber-physical systems.

Factors of successful cyber-physical systems project implementation:

- participation of top managers of the company in implementation;

- availability and adherence to the implementation plan;

- managers have clear goals and requirements for the project;

- involvement of the customer's specialists in the project implementation;

- high quality of the cyber-physical system itself and the team of the solution provider;
- reengineering of business processes before implementation;
- the company has a developed strategy;
- getting a quick and tangible project result.

The main reasons of failures of cyber-physical systems project implementation:

- inattention of the company's management to the project;
- lack of clearly defined project goals;
- informalization of processes in the company;
- the company's unwillingness to changes;
- instability of legislation;
- corruption in companies;
- low qualification of personnel in the company;
- insufficient project funding.

6.4. Quality management of cyber-physical systems implementation projects in clean transportation

6.4.1. The concept of project quality

The quality of a project is a holistic set of its characteristics related to its ability to meet specific needs. It is necessary to distinguish such conceptually different concepts as quality and sort.

Quality as a characteristic of the output or result of a project is "the degree to which a set of inherent characteristics meet the requirements" (ISO 9000). A sort as a constructive intent is a category assigned to the delivered results that have the same functional purpose, but different technical characteristics.

The project manager as well as the project team are responsible for reaching tradeoffs to achieve the required levels of both quality and sort. A quality level that does not meet quality requirements is always a problem, and a low sort may not be a problem. For example:

- There may be no problem if the developed cyber-physical system is of a low sort (has a limited number of functions) and at the same time of high quality (there are no obvious defects, there is a written user manual). In this case, the product meets the general purpose of use.

- The problem arises when a high-sort software product (has many functions) is of poor quality (there are many defects, poorly organized user documentation). In fact, the set of its high-sort functions turns out to be ineffective and / or insufficient due to low quality.

Quality management (within the framework of project management) is a system of methods, tools and activities aimed at meeting the requirements of project participants and stakeholders for the quality of the project itself and its products.

Modern approaches to quality management strive to minimize deviations and achieve results that meet specific requirements. These approaches highlight the importance of the following points [7]:

- Customer satisfaction. Understanding, evaluating, defining and managing customer requirements in a way that satisfies customer expectations. This requires a combination of compliance (the project must create what it was initiated for) and usability (the cyber-physical system must meet real needs).

- Prevention is more important than inspections. Quality should be planned, designed and built into the project management system, not just inspected during project implementation. The cost of preventing errors is generally significantly lower than the cost of correcting them once discovered through inspection or through the use of the cyber-physical system created.

- Continuous improvement. The plan-do-check-act (PDCA) cycle – the model described by Shewhart and refined by Deming – is the basis for quality improvement. In addition, quality improvement initiatives such as Total Quality Management (TQM), Six Sigma, and the combined use of Six Sigma and Lean Six Sigma can improve project management and also the quality of the project product.

- Management responsibility. The involvement of all members of the project team is essential for success. However, the project manager retains, within the scope of quality responsibility, the responsibility to provide the appropriate resources to the appropriate extent.

- Cost of quality (COQ). Quality Cost is the total cost of working to ensure compliance and resolve non-compliance. Quality assurance costs can be incurred throughout the life of a project. Moreover, issues related to post-project COQ should be addressed in the program management and portfolio management process, for example, project, program and portfolio management offices should apply appropriate analysis methods, templates and ways of allocating funds for this purpose.

Project quality management includes the processes and actions of the performing organization that define the quality policy, objectives and responsibilities so that the project meets the needs for which it was implemented.

Project quality management uses policies and procedures to implement the organization's quality management system in the context of the project and, where necessary, supports actions to continually improve the processes being undertaken by the performing organization. Project quality management seeks to ensure that project requirements, including product requirements, are met and verified.

The main elements of project quality are as follows:

- quality of the project product as the degree of achievement of the set goals;

- quality of project development and planning;

- quality of work performed on the project;

- quality of the resources that are involved in the implementation of the project.

They are all closely interconnected, the absence of one of them will not allow you to get a quality project.

6.4.2. Quality Management stages of cyber-physical systems implementation projects

Project quality management is carried out throughout the entire project life cycle. Figure 6.3 represents the stages of project quality management.

Concept stage. At this stage, the policy and strategy are defined to ensure the quality of the product being developed that meets the expected customer needs.

The concept contains the following sections:

- quality policy and strategy;
- general requirements and principles of quality assurance;
- standards, norms and rules;
- integration of quality assurance functions;
- requirements for the quality management system.

Planning stage. During the quality planning stage, standards are defined that should be used to ensure that the project content meets expectations of the project participants. Quality planning involves both identifying these standards and finding ways to implement them. The main tasks of the planning stage are:

- determination of quality assessment indicators;
- definition of technical specifications;
- description of quality management procedures;
- drawing up a list of controlled objects;
- choice of methods and means of quality assessment;
- description of connections with other processes;
- development of a quality management plan.

Organizing stage. The stage of organizing quality control requires the creation of the necessary and sufficient organizational, technical, financial and other conditions to ensure the fulfillment of the requirements for the quality of the project and its products and the possibility of meeting them.

Control stage. Quality control means determining whether the project results meet the quality standards and the reasons for the violation of such compliance.

Adjustment and *Analysis stages*. The quality control stage requires regular review of the progress of the project in order to establish actual compliance with the previously defined requirements:

- comparing actual project results to requirements;
- analysis of the quality progress in the project during its life cycle;
- formation of list of deviations;
- corrective actions;
- documenting changes.

Close-out stage. At the completion stage, the project team performs a summary quality assessment of the project results and completes the receipt of the results; the customer also draws up lists of quality claims; the team resolves conflicts and disputes, prepares documentation, analyzes experience and learned lessons on quality management.

The main quality assurance processes for all types of project are quality planning, quality assurance and quality control.

6.4.3. Project quality planning process

Quality planning is the process of determining which quality standards apply to a given project and how to meet them in the most efficient way.

The quality management plan includes activities that ensure the quality of the project results. One of the main components of the quality management plan for

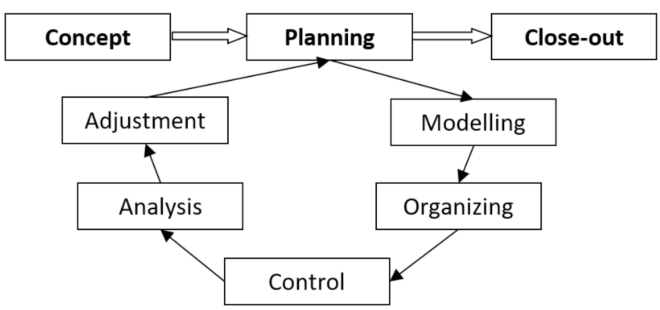


Fig. 6.3. Quality management stages of cyber-physical system implementation project .

cyber-physical systems implementation projects is the plan for testing models and systems.

The quality plan should determine how the project will ensure the quality of work from the standpoint of the organizational structure, resources, methodological support. At the quality planning stage, it is recommended to develop documents regulating the actions for controlling the quality of project management (reporting form for project implementation, project monitoring questionnaires) and quality management procedures, for example, quality control of project results, quality control of project documents, approval procedure for project documents, preparation and implementation of project control. To control the quality of the project documents, the quality plan should define a list of persons who approve each project document, the timing and form of their approval.

Many specific terms are introduced in projects for the development of cyber-physical systems, therefore it is necessary to include the development and approval of a project glossary in the project quality control plan. The project glossary is a structured list of all project terms and definitions, as well as abbreviations used, with a brief description of their content.

Quality planning begins with the definition of project quality objectives, policies and standards related to the project scope. Then the actions and responsibilities of the team members are defined, the implementation of which is necessary to achieve the goals and adhere to the standards.

The result of quality planning is presented in the form of quality assurance plans and management processes that ensure the implementation of these plans, and is achieved through synchronization with the main (content, schedule, cost) and auxiliary (risk and team) planning processes.

The task of quality planning tools is to make project management processes predictable. To plan the quality of the project, it is recommended to use the following methods:

- Quality assurance program. It is an action plan that ensures compliance of the actual quality of the project with the planned quality. The development of the program begins with the preparation of initial information, which includes the company's policies and procedures in the field of quality, customer requirements, description of the project content, etc. The basis for creating a quality program is the structure of the project. Quality standards are set to measure the Customer's expectations. Standards can be international, national or corporate.

Once the quality standards are set, it is necessary to determine the tasks, the solution of which will ensure compliance with the standards, then the responsibility for the implementation of the planned work and the timing of their implementation.

- Cost-benefit analysis. The purpose of the method is to maintain the required ratio between income and expenses in the project. Project quality assurance, of course, leads to additional costs, therefore, for each proposed method of quality assurance it is necessary to analyze the profitability ratio.

- Benchmarking involves comparing an existing or planned project with other projects in order to develop ideas to improve the quality of project implementation.

- Experiment planning is a statistical method that allows to determine the factors that affect certain variables of a product or process.

The result of planning is the development and approval of a quality management plan.

Quality management plan is a description of how the project team will carry out work and tasks in the field of quality. Depending on the needs of the project, the quality management plan can be very detailed or generalized. Quality assurance measures should be developed at the beginning of the project and carried out on the basis of independent expert assessments.

6.4.4. Project quality assurance process

Quality Assurance is the process of performing planned, systematic quality operations to ensure that all specified processes are performed and that the project meets the established quality requirements.

The quality assurance process includes methods for continually improving the quality of future projects. The knowledge and experience in quality assurance gained in the current project should be used in the preparation of quality assurance plans for the following projects.

To ensure quality, the project team needs:

- Information on work performance – information (about the status of delivery results, about the necessary corrective actions, as well as progress reports) used in the audit, quality peer review and process analysis.

- Change requests approved that should contain changes in working methods, product requirements, quality requirements, content and schedule. The approved changes are checked for their impact on the quality management plan.

- Quality checklists (quality metrics) – instructions for the reviewer. The listed items should be meaningful enough so that if the checklist gets overwhelmed, it will be used.

The results of the project quality assurance process are:

- Requested changes are intended to carry out special measures to improve the efficiency of rules, procedures and processes in the performing organization.

- Recommended corrective actions – actions that have been developed as a result of quality assurance measures are recommended for immediate implementation.

- Updated quality standards to be used in the future in the quality control process.

- Updated management plan.

6.4.5. Project quality control process

Quality control is a process that includes tracking the intermediate results of a project, determining their compliance with accepted standards, and developing actions to eliminate the causes of deviations from the standard. The quantitative assessment of quality control is carried out on the basis of statistical analysis and probability theory.

The following methods and means are used to carry out quality control:

- Causal diagram helps to identify possible causes affecting the quality of the product or a process in the project.

- Control charts are designed to determine the stability of the process and the predictability of its development. Control charts are a graphical display of the interaction of process variables during the process and provide an answer to the question whether the process variables are within the established boundaries.

- Pareto chart is a special type of histogram, sorted by the frequency of occurrence of an event, reflecting the number of detected defects, then, problems are a consequence of causes related to a certain set or category.

- Forecast scheme reflects the history and pattern of changes. It is a line graph showing the data entry points arranged on the graph in the order of their occurrence. The forecast diagram gives an idea of the trends of the process over time. The trends are also analysed using these charts. Trend analysis is often used to monitor schedule performance and project costs.

- Statistical samples are part of the controlled products which allow to make a conclusion about all products of this type in the project. Correct sampling often helps to reduce quality control costs.

- Inspection includes processes such as testing initiated to determine that the project results meet acceptable requirements and standards. Distinguish between testing both individual business processes and their totality (integration testing). Test scripts are developed for testing.

Note that for the quality control of the developed cyber-physical system, a summary table of test scenarios is being developed.

The QC results of a project are the results of QC activities communicated through feedback to the QA department:

- updated quality management plan;

- recommended corrective actions – specific actions that are triggered by the results of project quality control operations;

- recommended preventive actions – special measures to prevent the occurrence of conditions under which the project processes can go beyond the established parameters;

- defect correction is recommended – suggestions for eliminating defects. You can use the defect log to generate a set of recommendations for correcting defects;

- updated project management plan.

6.5. Model for express-assessment of the decision to close problem projects in a crisis condition

In current conditions, the project-oriented approach is one of the most effective approaches for business organization. Companies are implementing investment projects of various scopes in order to develop their activities, including in the field of cyber-physical systems. The success of projects largely depends on the correct investment policy. This policy is always associated with costs and risks. But with correct calculation the costs will certainly pay off and the company will reach a new level of development. But in the case of an incorrect investment policy of the company, strategic mistakes, or as a result of external factors, a situation may arise of the need to close the project.

2020 was distinguished by global scale crisis phenomena: the global pandemic of COVID-19, an economic crisis, currency fluctuations, etc. All these events in the aggregate had a negative impact on the project activities of most companies. At the moment, a large number of companies found themselves at different stages of their investment projects, and many of them found themselves in a situation of inability to continue the project. At the same time, in some investors, companies the project was at the initial stage, they had already done most of the work, and one of them was at the stage of launching an investment project, when the crisis phenomena of force majeure took place. It is worth noting that in such a situation, the unquestioning closure of the project is not always the right decision.

In order to decide whether to continue or not to continue an investment project, the manager needs to weigh the pros and cons, as well as analyse the risks and opportunities of the external environment.

This subsection describes an express method for choosing the optimal solution associated with stopping a project in terms of minimizing costs.

There are key options for minimizing losses in case of impossibility to implement the project in full during the crisis period:

- freeze the project for a while, wait for the conditions to change;
- reorganize the project – adjust the goal, reduce the requirements for the result;
- sell the project, subject to the availability of interested investors;
- close the project.

This method is based on a point-rating assessment of factors that affect the implementation of the project. The choice of the rating format is due to the fact that these models are the simplest in practical application and their results are quite simple to interpret.

The developed model takes into account the following indicators:

- R1 – the level of spent funds at the time of project evaluation – the actual percentage of the total planned investments for the project, which were spent at the time of the need to make decisions – [0; 100] points. The higher the value of the indicator, the more difficult it is to make the decision just to close the project. An increase in the indicator point leads to an increase in the rate.

- R2 – expert assessment of the probability of obtaining the planned amount of net present value (NPV) from the project. New conditions may not allow the customer of the project to receive the planned amount of profit. Estimated in the range of [0; 100] points, which corresponds to the level of probability in percentage. The lower the value of the indicator, the faster it is necessary to make the decision on the completion of the project.

- R3 – expert assessment of the possibility to increase investment in the project. In most cases, the new conditions require an increase in either the volume of investment or operating costs. Projects for which it is possible to obtain additional funding may have a time lag and do not require quick closure. The indicator is assessed by experts, the range of values is [0; 100], where 0 means that it is absolutely impossible to increase investments, 100 – there is a fundamental real opportunity to receive full amount of additional financing.

- R4 – assessment of interest in current project from potential investors or competitors. It characterizes the opportunity (as an assessment of the probability) for the customer of the project to sell it to more influential or wealthy groups, thus compensating for the cost of resources and possibly obtaining additional profit. The indicator is assessed by experts, the range of values is [0; 100], where 0 means that in modern conditions it is impossible to find a buyer or a new investor for the project, 100 – there is a real offer to sell the project at an acceptable price for the investor or customer.

We accept the hypothesis that the influence of each factor is not equivalent for making a decision, thus, each of the factors should have a different specific coefficient in the model.

The specific coefficients of the model were obtained by an expert method by interviewing 10 IT project managers of Ukrainian companies. The task of the experts was to distribute 10 points among the proposed. Thus, the model looks as follows [9]:

$$R = 0.3 \cdot R1 + 0.27 \cdot R2 + 0.24 \cdot R3 + 0.19 \cdot R4.$$

As a result of the assessment, the project can receive from 0 to 100 points.

If the evaluation value falls within the range [0; 30] , it is recommended to make a decision on the early closure of the project; if [31; 50], it is recommended to consider the issue of selling the project to another investor, or to reorganize the project; if [51; 80] , it is recommended to take a pause and wait for changes in the external conditions of the project; if [81; 100], it is recommended to continue the project.

Note that this model can be used when assessing a project portfolio in order to rank (prioritize) projects in terms of cost optimization during a crisis period.

Conclusion: Closing a project is far from the only and not always the most effective way to complete a project in a crisis. The paper presents a model for the express assessment of a decision on the management of problem projects, which can become an effective decision-making tool for a project manager in deciding the fate of a project. This model can be expanded by adding other factors, depending on the specifics of the project.

References

1. PMBOK® Guide – Sixth Edition (2017).

2. Morkun N., Zavsiehdashnia I. Management of IT projects in the conditions of a mining and metallurgical complex. - Kryviy Rih: Publishing Center of KNU, 2017. - 238 p. [in Ukrainian].

3. Elbeik, S. & Thomas, M. (1998). Project Skills, Oxford, Butterworth-Heinemann.

4. Lyovin B., Tsvetkov V. Cybernetics and physical systems for transport management // World of transport and transportation, Vol. 16, Iss. 2, pp. 138–145 (2018).

5. M. Švéda, O. Ryšavý. Dependable cyber-physical systems networking: An approach for real-time, software intensive systems. IF AC Proceedings

Volumes, 46 (28) (2013), pp. 116-119

6. Broy, M.; Cengarle, M.; Geisberger, E. (2012) Cyber-Physical Systems: Imminent Challenges. Monterey Workshop.

7. Gang Xiong, Fenghua Zhu, Xiwei Liu, Xisong Dong, Wuling Huang, Songhang Chen, and Kai Zhao. Cyber-physical-social System in Intelligent Transportation. IEEE/CAA Journal of Automatica Sinica, Vol. 2, No. 3, July 2015.

8. Sun, Di-Hua & Li, Yongfu & Liu, Wei-Ning & Zhao, Min & Liao, Xiao-Yong. (2013). Research summary on transportation cyber physical systems and the challenging technologies. Zhongguo Gonglu Xuebao/China Journal of Highway and Transport. 26. 144–154.

9. Morkun N., Zavsiehdashnia I. Model for rapid assessment of the decision to close problematic investment projects in a crisis. European project culture in Ukraine: state, problems, prospects. 2020. – pp. 137–139. [in Ukrainian].

Chapter 7:

Intelligent information technologies and systems in transport

Sergey Ruban
Vitaliy Tron

Kryvyi Rih National University

7.1. Fundamentals of intelligent systems in transport

The use of modern information technologies in transport is an effective means of increasing the efficiency of the functioning of transport systems, improving the economic performance of transport enterprises, and increasing their competitiveness. At the same time, obtaining an effect from informatization processes requires a clear understanding of the features of transport systems and the fullest possible consideration of these features when choosing and using information technologies in the design process of specialized information systems.

Currently, the tasks of creating information systems and automation systems with intelligent properties are urgent. These properties include the ability to 1) learn, which involves the formation of new knowledge in the form of various kinds of mathematical models; 2) classify – to distribute objects and impacts automatically; 3) adapt – to change work algorithms in accordance with the changing conditions and parameters of the functioning environment.

The greatest impact in the field of technical regulation of intelligent transport systems (ITS) have the following world standardization systems: International Organization of Standartization (ISO), European Committee for Standartization (CEN).

The Technical Committee of the International Organization of Standartization – ISO/TC 204 – Intelligent Transport Systems is responsible for the development of standards of information, communication and control systems for land transport, travel information, traffic management, public transport, commercial transport, emergency services and commercial services in the sphere of intelligent transport systems [1]. The Committee includes 19 working groups, in particular in the following areas:

- big data and artificial intelligence,
- ITS database technology,
- integrated transport information, management and control,
- communications, cooperative systems.

The Technical Committee of the European Committee for Standartization – CEN/TC 278 – Intelligent Transport Systems leads the preparation of standards in the field of intelligent transport systems in Europe and also serves as a platform for European stakeholders to exchange knowledge, information, best practices and experiences in the field of ITS [2]. The Committee includes 9 working groups, in particular in the following areas:

- cooperative ITS,
- mobility integration,
- traffic and traveller information,
- public transport,
- road traffic data.

Standard “ISO 21217:2020(en) Intelligent transport systems – Station and communication architecture” [3] was prepared by Technical Committee ISO/TC 204, Intelligent transport systems. It provides the intelligent transport systems (ITS) station and communication reference architecture that is referenced in a family of deliverables from standard development organizations (SDOs) for cooperative intelligent transport systems (C-ITS). In accordance with the standard, applying information and communication technologies allows to improve surface transportation in terms of safety, efficiency, comfort, and sustainability.

There is a large number of communication services in the area of intelligent transport systems [4], [5] characterized by diverging requirements. To support this, ISO 21217:2020 justifies the need to combine multiple access technologies and communication protocols with distinct performance characteristics (Fig. 7.1) [3]: communication range, available bandwidth, end-to-end transmission delay, quality of service, security, etc.

Combining multiple access technologies and communication protocols requires a common approach to the way communications and data are securely managed (Fig. 7.2) [3].

Similarly to the ISO Open Systems Interconnection (OSI) 7-layer architecture, the ITS station architecture is divided into three independent communication layers with ITS *Applications entity* (road safety, traffic efficiency, etc.) on top of this hierarchy:

- access layer,

- networking and transport layer,
- facilities layer.

To support communications and applications, there are several additional cross-layer entities in charge of the management and security activities.

Management cross-layer entity is responsible for:

- regulatory management,
- cross-layer management,
- station management,
- application management.

Security cross-layer entity is responsible for management information base:

- firewall and intrusion management,
- authentication, authorization, profile management,
- identity, crypto-key and certificate managment.

Details of the functional elements of the ITS station architecture are specified in a set of related standards [3]. For instance, ISO 24102-1:2018 "*Intelligent transport systems – ITS station management*" in which local ITS station

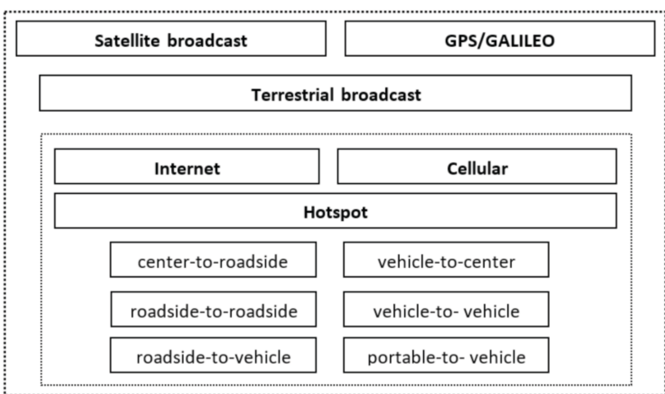


Fig. 7.1. Examples of ITS communications.

management protocols are specified by means of management processes, and data that are exchanged between the station management entity, applications entity security entity and the various communication protocol layers.

The implementation of this ITS station architecture (Fig. 7.2) is referred to as an “ITS station unit” (ITS-SU) [3]. The functionalities available in an ITS-SU can be implemented in one or multiple physical units, referred to as “ITS station communication units” (ITS-SCUs). The various ITS-SCUs of one single ITS-SU may even be split over a large geographical area, e.g., along a motorway several tens of kilometres in length.

ITS-SUs conformant with this document may be deployed in various environments, including vehicles of any kind (vehicle ITS station), on the roadside infrastructure (roadside ITS station), in data centres (central ITS station), or in nomadic devices (personal ITS station) [3].

ITS station units are designed for supporting the secure provision of the C-ITS services and secure allocation of resources with prioritized access. Security means covering the two essential operational modes [3]: establishment and maintenance of secure session; authentication of the broadcast message sender.

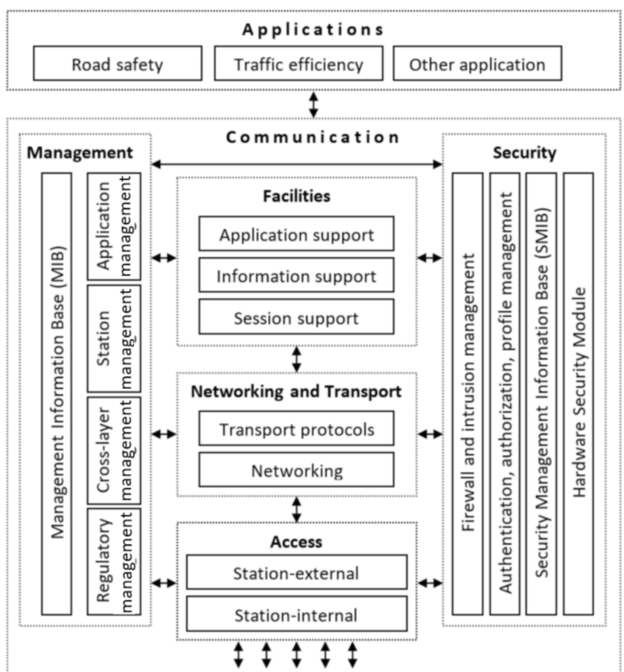


Fig. 7.2. ITS station reference architecture.

Supporting multiple access technologies and communication protocols, also referred to as ‘hybrid communications’, is a design principle of the ITS station architecture. The ITS station architecture is thus specified with no pre-defined mandatory communication technologies. It can support any type of existing and forthcoming technology.

Presently, specifications have been developed to support a number of access technologies, for example [3]:

- all kinds of cellular access technologies (e.g., specified at 3GPP with profile standards from other SDOs tailoring them to the ITS station reference architecture);
- satellite communications;
- other technologies such as infrared, millimetre wave (ultra wideband communications), vehicular Wi-Fi (ITS-G5/US-DSRC/ITS-M5: all profiles of IEEE 802.11 OCB), and optical light communications;

and several flavours of communication protocol suites:

- GeoNetworking/Basic Transport Protocol from ETSI;
- FNTP from ISO;
- WSMP from IEEE; and
- the suite of IPv6 protocols from IETF with supporting specifications from ISO.

The ITS station architecture actually combines [3] localized communications and networked communications. Localized communications, that is, communications to nearby stations without involving networking from a source station through nodes of a network to a final destination station – also referred to as ‘ad-hoc communications’.

Unlike many legacy applications, the choice of the access technology and communication protocol can be made transparent to the applications. This is achieved through a number of functionalities across the ITS station architecture in support of hybrid communications (Fig. 7.3).

Before transmitting data, applications provide their communication requirements to the management entity of the ITS-SU for each type of communication flow:

- amount of data,
- level of security,
- level of priority,
- end-to-end transmission delay, etc.

The management entity maintains such elements of information as:

- existing capabilities of the ITS-SU and their status,
- local regulation enforcing the use of a specific communication profile,
- characteristics and load of available radio technologies,
- current load of the ITS-SU, etc.

Based on the communication requirement and the current view of the management, the uppermost relevant communication profile (uniquely identified by an ITS-S communication profile identifier) is selected and ITS station resources are securely committed for identified communication flow.

ITS station architecture serves as a reference for numerous C-ITS services developed around the world, and more particularly, in Europe. Early deployments of C-ITS services conforming to ITS station architecture have been initiated in Europe under the framework of the C-Roads [6] and InterCor initiatives [7] supported by the European Commission.

The C-Roads Platform [6] allows to join together authorities and road operators to harmonise the deployment activities of cooperative intelligent transport systems across Europe. Also, members of the single pilot activities as well as of the C-Roads-Working Groups actively contribute to the work of the EU-C-ITS-Platform.

The overall goal of the InterCor project is to achieve safer, more efficient and more convenient mobility of people and goods [7]. The project aims to enable vehicles and related road infrastructure to communicate data through cellular, ITS G5, or a combination of both networks on road corridors.

National pilot deployments are underway all across Europe, for instance,

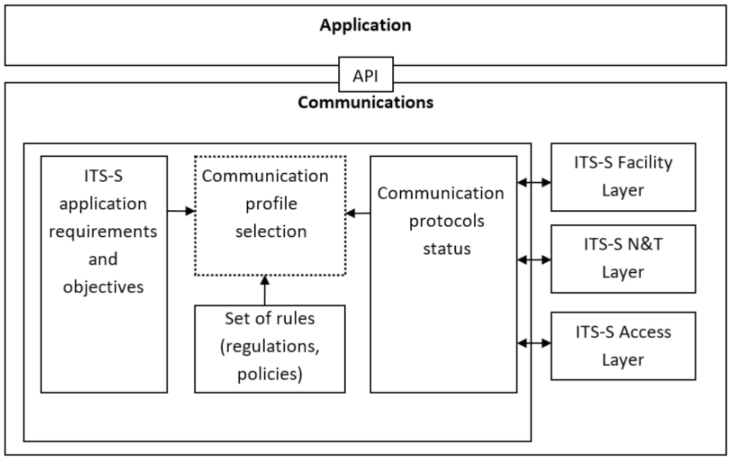


Fig. 7.3. Architecture of communication profile and path selection.

SCOOP, NordicWay.

SCOOP project (France) [8] main objectives are as follows: improving road safety and the safety of road operating agents; making traffic management more efficient and contributing to the reduction of emissions; optimizing infrastructure management costs, making vehicles fit for the future and developing new services. Vehicles are equipped with sensors to detect events such as a slippery road, an emergency brake, etc. and with on-board units to transmit the information to vehicles behind (V2V) and to the road operator (V2I) through road side units. The road operator can also transmit information (roadworks, etc.) to the vehicles through their on board units (I2V).

NordicWay projects (Finland, Norway, Sweden, Denmark) [9] are aimed to enable vehicles, infrastructure and network operators to communicate safety hazards and other information from roads in the Nordic countries between different stakeholders.

All of these experiences, gained through early deployments, demonstrate that it is not possible to provide the same level of services to all vehicles in all locations. The type of service and the performance of the service depends on national decisions, local road environment, density of population, density of vehicles equipped, cellular coverage, and numerous other factors. In addition, and importantly, the roadside infrastructure equipment and vehicles have a life expectancy that far exceeds the innovation cycle of new radio and communication technologies. Equipment at the roadside and in vehicles is, therefore, likely to have to accommodate new communication technologies during its lifetime.

7.2. Vehicular ad-hoc networks

Advances in technology for the production of smart devices based on microprocessors, miniaturization of electronic components and increase of their functionality have affected the automotive industry. Modern vehicles are equipped with a significant number of intelligent sensors, actuators, built-in computers and communication devices that implement the ability to exchange information in various data transmission networks [10]. A huge amount of data received by devices for monitoring the state of a vehicle makes it possible not only to solve the problem of increasing the efficiency of the functioning of an individual vehicle, but also of the transport system as a whole. For example, information about the current position and speed of vehicles makes it possible to solve the problem of optimizing the route according to various criteria, predicting and preventing various emergency situations, for example, violations of the traffic schedule, congestion, and as a result, reduce fuel costs, reduce the impact of transport on the environment, etc. The accumulation of information on the technical and economic performance of motor vehicles in various operating conditions makes it possible to analyse the efficiency of the design of individual units and the vehicle as a whole. Thus, the use of intelligent sensors, information and communication technologies in vehicles has made it possible to implement intelligent transport systems (ITS), the main goals of which are to improve the safety and efficiency of road traffic, as well as to provide information and entertainment services. [11].

As noted in [11], from a technical point of view, intelligent transport systems are based on network technologies, in particular on wireless self-organizing networks, in which vehicles can interact with each other without a significant need for external infrastructure and centralized control. In this context, transport networks are similar to mobile ad-hoc networks (MANET) and wireless sensor networks [12], but at the same time they have a number of distinctive features, which include [11], [13]:

- dynamic topology associated with frequent changes in the speed and direction of movement of vehicles;
- intermittent connectivity – connection between communicating devices can break at any time due to dynamic topology;
- presence of mobility patterns – unlike MANET, vehicles move according to certain patterns (depending on the topology of road networks of

geographic areas, road conditions, the presence of traffic lights and speed limit zones), which helps in creating routing patterns for transport networks;

- traffic density – varies from low to high, depending on the geographical area (city, highway, countryside) and on the time factor (peak hours, seasonal congestion);

- unlimited battery power and storage;

- heterogeneity – it is considered that the nodes of transport networks can have different characteristics and capabilities (communication ranges, capabilities and categories of detection).

The issues of finding effective ways to organize network interaction between vehicles, as well as between vehicles and infrastructure, are one of the most

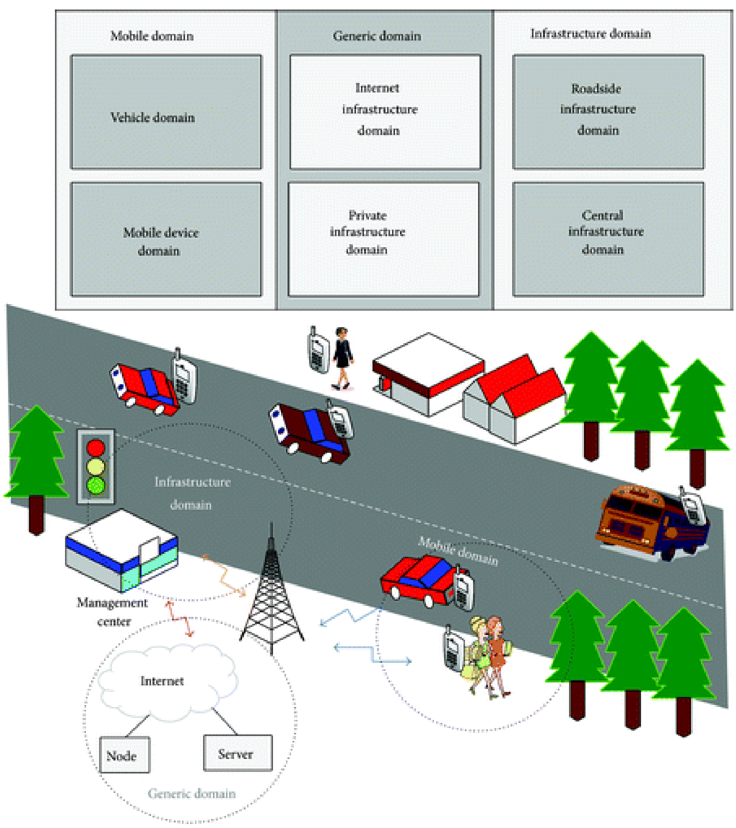


Fig. 7.4. VANET system domains [14].

difficult issues in the modern network industry and have formed a separate area of research, which is known as vehicular networking. To define this direction, several terms are used today, in particular vehicular ad-hoc network (VANET), inter-vehicle communication (IVC), Car-2-X (C2X), or vehicle-2-X (V2X) [12].

According to [14], entities in VANET can be divided into 3 domains (Fig. 7.4) which form the VANET architecture:

1. Mobile domain consists of two parts – the vehicle domain and the mobile device domain. The vehicle domain consists of entities that are constantly moving – buses, cars, trucks, etc. The mobile device domain includes various portable devices such as mobile phones, tablets, laptops.

2. Infrastructure domain consists of a roadside infrastructure domain, which includes roadside infrastructure objects (traffic lights, lighting poles, etc.), as well as a central infrastructure domain, which includes management centres (Traffic Management Centres, TMCs, Vehicle Management Centres VMCs, etc).

3. Generic domain consists of Internet infrastructure and private infrastructure (various nodes, servers and other computing resources).

It is worth noting that there is an alternative vision of the VANET architecture. Thus, the CAR-2-CAR communication consortium(C2C-CC), which is one of the main driving forces for the development of communication technologies for motor vehicles in Europe, published its 'manifesto' in 2007. In this document the architecture of the CAR-2-X Communication System was presented, explaining the key components and how they interact, the main protocols and interfaces [15].

The draft reference architecture of the C2C Communication System is shown in Fig. 7.5 and comprises the following domains:

1. In-vehicle domain consists of one on-board Unit (OBU) and one or more application units (AUs). An AU is a specialized device (such as an integrated sensor or passenger's handheld device) that can run one or multiple applications that use the OBU's communication capabilities. AUs are permanently connected to OBU via wired or wireless connection [11].

2. The ad-hoc domain, or vehicular ad-hoc network (VANET), is composed of vehicles equipped with OBUs and stationary units along the road, termed road-side units (RSUs). OBUs can communicate with each other using

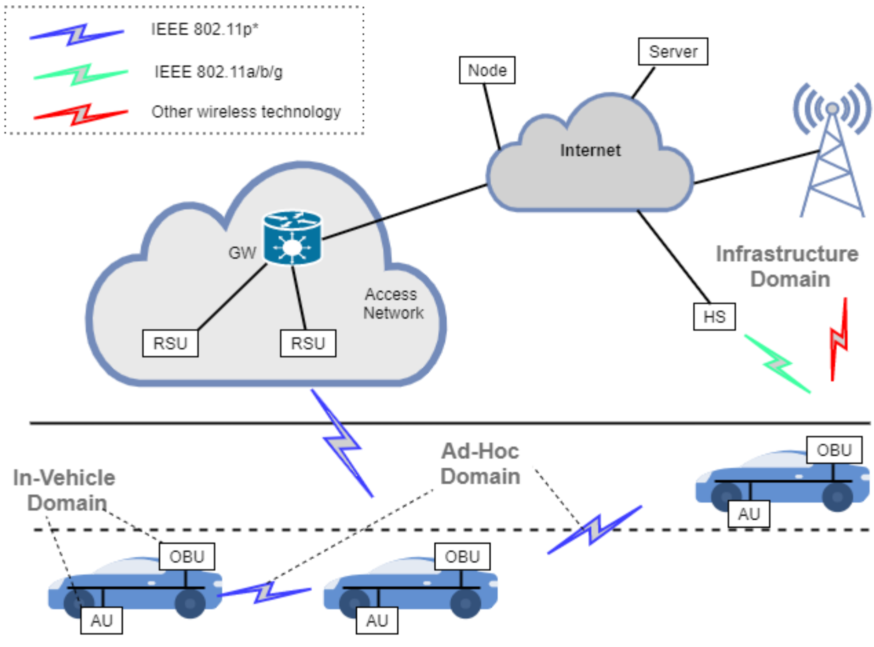


Fig. 7.5. Reference architecture of the C2C Communication System [6]: AU – appplication unit; OBU – on-board unit; RSU – roadside unit; GW – gateway; HS – hot spot.

short-range wireless protocols, primarily solving the problem of traffic safety. In the absence of a direct connection between OBUs, dedicated routing protocols implement multi-hop communication. RSUs are designed primarily to improve road safety by expanding an ad-hoc network that is provided by connecting RSUs to the infrastructure network. Thus, the OBUs can be considered as mobile ad-hoc network nodes and RSUs – as static nodes. OBU can access the Internet via RSU connected to infrastructure, public commercial or private wireless hotspots (HSs) to communicate with Internet nodes or servers.

3. The infrastructure domain consists of HS and RSU. If neither RSU nor HS provides Internet access, OBUs can use the communication capabilities of cellular radio networks (GSM, GPRS, UMTS, HSDPA, WiMax, 4G).

The communication types in VANET can be divided into the following categories, which are closely related to the previously described components [14]:

1. In-vehicle communication – communication in a vehicle between an on-

board unit (OBU) of a vehicle and its control units (AU) [11], designed to determine the characteristics of the vehicle and the driving mode (in particular, to determine the driver's fatigue and drowsiness), which is of particular importance to ensure safety.

2. Vehicle-to-vehicle (V2V) wireless communications between vehicles via their OBUs, in order to provide drivers with a platform for exchanging data and warning messages.

3. Vehicle-to-infrastructure (V2I) is implemented through bi-directional wireless communication between the on-board units of vehicles and connected to the RSU network infrastructure and is designed to receive real-time updates on traffic conditions, weather conditions, etc.

4. Infrastructure-to-infrastructure (I2I) communications between RSUs enable extending the coverage of the network [11].

5. Vehicle-to-broadband cloud (V2B) communications between vehicles and broadband cloud via wireless broadband technologies such as 3G/4G.

Figure 7.6 describes the key functions of each type of VANET communication [14]–[16].

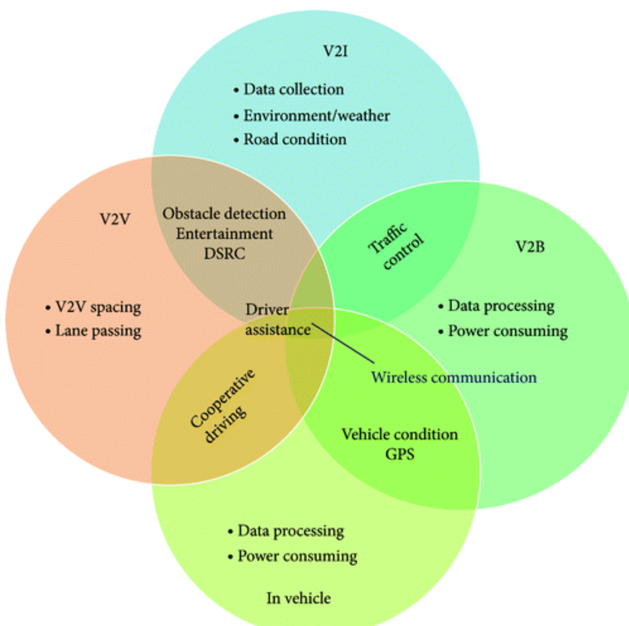


Fig. 7.6. Key functions of different VANET communication types [14].

References

1. ISO/TC 204 Intelligent transport systems. URL: https://www.iso.org/committee/54706.html

2. CEN/TC 278 – Intelligent transport systems URL: https://standards.cen.eu/dyn/www/f?p=204:7:0::::FSP_ORG_ID:6259&cs=1EA16FFFE1883E02CD366E9E7EADFA6F7

3. ISO 21217:2020(en) Intelligent transport systems — Station and communication architecture. URL: https://www.iso.org/obp ui/#iso:std:iso:21217:ed-3:v1:en

4. Ali M. Standards of Communications in the Intelligent Transport Systems. In book: Autonomous Vehicles Intelligent Transport Systems and Smart Technologies. Nova Publishers. 2014. DOI: 10.13140/2.1.2055.1040.

5. Zhang H., Lu X. Vehicle communication network in intelligent transportation system based on Internet of Things. *Computer Communications*. 2020. Vol. 160. pp. 799–806. DOI: 10.1016/j.comcom.2020.03.041.

6. C-ROADS – The Platform of Harmonised C-ITS Deployment in Europe. URL: https://www.c-roads.eu/platform.html

7. InterCor (Interoperable Corridors). URL: https://intercor-project.eu/

8. Project SCOOP. URL: http://www.scoop.developpement-durable.gouv.fr/en/

9. NordicWay. URL: https://www.nordicway.net/

10. Ashokkumar K., Sam Baron, Arshadprabhu R., Britto. Cloud based intelligent transport system. *Procedia Computer Science*. 2015. Vol. 50. P. 58 – 63. URL: https://daneshyari.com/article/preview/489833.pdf

11. Yeferny T., Hamad S.. Vehicular Ad-hoc Networks: Architecture, Applications and Challenges. *IJCSNS International Journal of Computer Science and Network Security*, Vol. 20, No. 2, February 2020. ArXiv, abs/2101.04539. URL: https://arxiv.org/ftp/arxiv/papers/2101/2101.04539.pdf

12. Dressler Falko, Hartenstein Hannes, Altintas Onur, Tonguz Ozan K. Inter-Vehicle Communication – Quo Vadis. *IEEE Communications Magazine (COMMAG)*, vol. 52 (6), June 2014. pp. 170–177. URL: https://www2.tkn.tu-berlin.de/bib/dressler2014inter-vehicle/dressler2014inter-vehicle.pdf

13. Tomar Ravi, Prateek Manish, Sastry G. H. Vehicular Adhoc Network (VANET) - An Introduction. *International Journal of Control Theory and Applications*, International Science Press 2016, 9 (18), pp. 8883–8888. ffhal-01496806. URL: https://hal.archives-ouvertes.fr/hal-01496806/document.

14. Liang W. Li Z. Zhang H, Wang S. Bie R. Vehicular Ad Hoc Networks: Architectures, Research Issues, Methodologies, Challenges, and Trends. *International Journal of Distributed Sensor Networks*. August 2015. doi:10.1155/2015/745303. URL: https://journals.sagepub.com/doi/10.1155/2015/745303.

15. Baldessari R., Bodekker B., Deegener M., Festag A., Franz W., Kellum C. C., Kosch T., Kovacs A., Lenardi M., Menig C. et al. Car-2-car communication consortium manifesto. 2007. URL: https://elib.dlr.de/48380/1/C2C-CC_manifesto_v1.1.pdf.

16. Faezipour M., Saeed A., Nourani M., Addepalli S. Progress and Challenges in Intelligent Vehicle-Area-Networks. Communications of the ACM, Vol. 55, No. 2, pp. 90–100. DOI: https://doi.org/10.1145/2076450.2076470

Chapter 8:

Intelligent transport systems

Joan Peuteman

KU Leuven

When discussing the future of the transport systems, people often focus on the use of hybrid and electric cars. When dealing with aircraft and ships, people often focus on the use of the so-called more electric aircraft and more electric ships. In general, the use of electric trains and electric trams is mature technology. And indeed, important changes are expected in the (near) future in an attempt to make the transportation sector greener (e.g., reducing harmful exhaust gasses and reducing CO_2 emissions).

People often forget other evolutions (or revolutions) in the transportation sector like the development of Intelligent Transport Systems ITS. By integrating electronic and ICT technology in cars, trains, trams, aircrafts or ships, a number of goals can be achieved. Especially when focusing on cars and trucks, engineers try

- to make the traffic on the road safer,
- to reduce traffic jams,
- to reduce the energy consumption while driving,
- to combine the use of cars and trucks with the use of trains, trams, ships, aircraft.

8.1. Industry 4.0 and smart cities

8.1.1. The industrial revolutions

The present state of the industry is the result of a number of important revolutions which occured in the past. The first industrial revolution started at the end of the 18th century. Due to the rise of steam engines (where mainly coal has been used as the primary energy source), a large scale mechanization of the industry appeared.

The second industrial revolution started at the end of the 19th century with the rise of electrical applications (in combination with the rise of the gas and oil industry). Linked with the second industrial revolution is not only the use of a new energy source like electricity but also the advent of new communication technologies like the telegraph and the telephone.

The third industrial revolution started at the end of the sixties or the beginning of the seventies of the previous century. Due to the rise of electronics (the use of transistors, microcontrollers and microprocessors) and the increased use of computers and telecommunication technology, new automation applications conquered the industry. Especially the use of programmable logic controllers (PLC) had a large impact on the industrial automation.

The fourth industrial revolution actually builds upon the third industrial revolution. This fourth industrial revolution and the term 'industry 4.0' are often used interchangeably. Others consider the term 'industry 4.0' as the subset of the fourth industrial revolution which deals with the industry. Anyhow, it is a technical revolution which takes place in the beginning of the third millenium and is driven by the availability of the Internet.

Due to the availability of the internet, industrial processes and machines are able to communicate and interact with each other in real time. A number of design principles are very important:

- Machines, devices, sensors and also people are able to communicate and interact with each other using the Internet of Things (IoT) or the Internet of People (IoP).

- Large amounts of useful data is available to the operators, which allows them to make correct decisions.

- Humans are supported by assistance systems which not only provide but also visualise the (large amount of) data. This allows to make correct decisions in a short period of time, i.e., it is possible to react sufficiently fast.

- Machines, devices, cyber physical systems are able to make decentralized decisions on their own. This increased autonomy of the machines implies only a limited number of decisions, problems, conflicts that need to be sent to a higher level.

The use of the new technologies supporting 'industry 4.0' applications in the industry are also useful in a much broader range. For instance, the use of real time communication possibilities (including the use of mobile devices), the use of (smart) sensors, location technologies, decent human-machine interfaces, and computing and cloud storage also impact the society outside the traditional industry. Here, the rise of 'smart' cities is a good example.

8.1.2. Smart cities

It is a worldwide evolution that an increasing number of people are living in the cities. It is a challenge to provide all these people with homes, schools, libraries, hospitals, decent transport systems, water supply, electrical energy, decent waste management, etc. Actually a large number of definitions exist for the 'smart city' concept, but by using electronic technologies, (smart) sensors, actuators, software, ICT technology, (big) data, internet connections new applications arise. A large number of devices, buildings, vehicles and people need to communicate with each other (and the government) to provide, collect and use large amounts of data to make intelligent decisions.

Sometimes, six important aspects of a smart city are considered:

- Smart economy: Entrepreneurship, innovation and a flexible labour market can be stimulated.

- Smart people: A decent educational system providing people with social networks which participate in the public life.

- Smart governance: A governance which stimulates the participation of the entire population and which operates in a transparent way is needed. The governance provides the required public and social services.

- Smart environment: It is important that the use of energy, water, food and other natural resources is restricted to a minimum. Decent waste management is needed to reduce environmental problems.

- Smart living: People need decent homes, and the city must be safe. Moreover, the governance must provide decent healthcare, cultural activities and stimulants to social cohesion.

- Smart traffic and traffic infrastructure: By reducing traffic jams, by providing intelligent parking infrastructure, and by integrating assistance in case of a (car) accident smart mobility solutions arise.

Based on the evolutions and the approaches sketched above, the development of decent 'Intelligent Transport Systems' is embedded in a broad technical, economical and social context. In the next paragraphs, we will focus on the possibilities and technical aspects of these 'Intelligent Transport Systems'.

8.2. Intelligent transport systems: global view

When dealing with 'Intelligent Transport Systems', two major goals arise. First, dealing with the increasing congestion of the roads, which is a major problem in a lot of industrialised countries. Decent road infrastructure, including a sufficient number of roads with a sufficient number of lanes, is needed but will not be able to solve all problems. The congestion of the roads, including traffic jams, accounts for a tremendous loss of time causing a severe economic loss (loss of expensive working hours, loss of fuel, increased risk of accidents, etc.). Improving the road safety for all road users is the second major goal. To reach these goals, a wide range of combined approaches is needed.

8.2.1. Reducing the congestion of the roads

There exist a number of ways to reduce the congestion of the roads. Supplied with correct and real-time traffic information the (potential) drivers are able to choose the most appropriate route to the destination. Alternatively, the potential driver can be informed about alternative choices, like using public transport (tram, bus, train, etc.). The traffic information needs to include correct information concerning the infrastructure works on the roads (including repair and maintenance works).

In order to reduce the congestion of the roads, car-pooling is also an option which can be supported by ICT based tools.

In case the driver looks for a parking lot, real-time information concerning the available nearby parking lots allows to avoid useless and inefficient search for a parking lot. Inefficient searches for parking lots account for an important loss of time (for the driver and the passengers), account for energy losses (useless fuel consumption), and have a large and disturbing impact on the traffic load. Notice that real-time information concerning the most nearby parking lot requires

- real-time knowledge of the position of the car (GPS system) which needs a parking lot;
- real-time communication between the car and a data centre which needs to have real-time information about the free parking lots.

Information concerning available parking lots is not only useful for private cars, it is also useful for trucks. Finding a decent parking lot is mandatory for the truck drivers in order to be able to respect the drive and rest times imposed by law in a lot of countries.

8.2.2. Improving the road safety

In order to improve the safety of all vehicles on the road, several approaches can be combined. A modern car already has a lot of mechatronic devices which improve the safety of the car.

- Airbag system which contains an accelerometer: In case of a car accident, a large (negative) acceleration will be measured. This measurement data is used to conclude that an accident happens, and the airbag will be opened to protect the car driver and the passengers.

- Anti-lock breaking system (ABS): The ABS system operates by preventing the wheels of the car from locking up while braking (e.g., during an emergency braking). Achieving a smaller stopping distance (although not always realistic) is a good property but it is especially important to improve the steering control of the car.

By integrating real-time communications between nearby cars (and possibly other road users like cyclists), accidents can be avoided. Suppose a car is driving behind another car. Suppose the first car performs a sudden braking because the driver tries to avoid hitting a pedestrian (it can be a child crossing the road). In case the first car sends a message to the second car behind him, it is possible to warn the driver of the second car as fast as possible. In case the driver of the second car does not respond sufficiently fast, his car can perform an automatic emergency brake.

The presence of nearby but invisible cars (e.g., invisible due to an obstacle) can be detected, which allows avoiding collisions. Especially when the visibility is low due to fog, detecting nearby cars is very useful. Real-time communication between the car and traffic infastructure can be useful to prevent the car driver from errors. For instance, when the car reaches the red traffic light and the car does not slow down, it is useful to warn the car driver. Possibly, the car performs an automatic brake due to the red traffic light.

In case an accident still happens (which can, e.g., be detected by the activation of the airbag system), the car automatically sends an emergency call to the emergency services. The emergency services have information about the location of the accident and the type of car (or cars) involved in the accident. Possibly personal communication by the car driver (or passenger) and the emergency services provides extra information. But also without such personal communication the emergency services are able to send help (e.g., an ambulance) much faster, which can be lifesaving (time reductions up to 50 % are considered to be realistic). The implementation of such an automatic communication with the emergency services requires collaboration of all actors: car manufacturers, telecom providers, and emergency services.

In case of an accident, it is useful to send information not only to the emergency services. By sending real-time information to nearby car drivers, additional accidents, including chain collisions, can be avoided.

8.3. Wireless communication

In order to reach the goals described in the previous sections, a lot of communication is needed. A distinction can be made between radio-frequent communication and optical communication.

- Radio frequent communication (using electromagnetic waves) can use several frequency bands, but in the present appplications a distinction can be made between the VHF band (Very High Frequency band: frequencies from 30 to 300 MHz corresponding with wavelengths of 10 m to 1 m), the UHF band (Ultra High Frequency Band: frequencies from 300 to 3000 MHz corresponding to wavelengths of 1 m to 10 cm) and the SHF band (Super High Frequency Band: frequencies from 3 GHz to 30 GHz corresponding with wavelengts of 10 cm to 1 cm). Wireless communication standards like ZigBee, Bluetooth, Wi-Fi, UWB and CALM are important options.

- Optical communication often uses near-infrared light, i.e., infrared light with wavelengths approximately ranging from 700 nm to 1400 nm (frequencies between 215 THz and 430 THz). The wavelength range of near-infrared light depends on the technical and scientific field of expertise, but it is the part of the infrared region which is nearest to the visible light region (wavelengths approximately ranging from 620 nm to 380 nm corresponding to frequencies ranging from 400 THz to 789 THz).

When dealing with wireless communication in the transport sector, a distinction can be made between transactions among vehicles (V2V – vehicle-to-vehicle communication), transactions between vehicles and infrastructure (V2I – vehicle-to-infrastructure communication), and transactions between vehicles and hand held devices (V2D – vehicle-to-hand-held-device communication). Additionally, also communication is needed between several sensors and components inside one single vehicle, i.e., in-vehicle communication.

8.3.1. Radio frequent wireless communication

In order to introduce intelligence in the transport sector, wireless sensor networks are really needed. The nodes of the wireless sensor networks contain fixed nodes (stationary nodes originating mainly from infrastructure) and mobile nodes (moving nodes attached with vehicles). The wireless sensor network is time-dependent, since the nodes attached with the vehicles not only move, they also enter and leave the network.

To realise radio frequent wireless communication, several technologies are available, and each technology has a number of properties, advantages and disadvantages. When comparing Zigbee, Bluetooth, Wi-Fi, UWB and CALM, a number of conclusions can be made (inspired by the paper by K. Selvarajah et al.).

Zigbee:

- Useful to realise in-vehicle, vehicle-to-vehicle and vehicle-to-infrastructure communication.
- A frequency of 2.4 GHz is used (also other communication frequencies exist).
- A network range up to 100 m is realistic.
- A bandwidth of 250 kbps is realistic, which is small in comparison with the other technologies.
- In general the power consumption is low, and it is also a low cost solution.
- When comparing with Bluetooth, the Zigbee technology can accomodate a larger number of devices.

Bluetooth:

- Useful to realise in-vehicle communication.
- Similar to Zigbee, a frequency of 2.4 GHz is used.
- A network range up to 70 m is realistic.
- A bandwidth of 12 Mbps is realistic, which is higher than the bandwidth obtained when using Zigbee.
- In general, more power is consumed in comparison to Zigbee.

Wi-Fi:

- Useful to realise vehicle-to-vehicle and vehicle-to-infrastructure communication.
- Similar to Zigbee and Bluetooth, a frequency of 2.4 GHz is used.

- A network range up to 100 m is realistic.
- A bandwidth of 54 Mbps is realistic, which is higher than the bandwidths obtained when using Zigbee or Bluetooth.
- In general, Wi-Fi consumes more power than Bluetooth (which comsumes more power than Zigbee).

UWB:

- Ultra Wide Band is useful to realise in-vehicle, vehicle-to-vehicle and vehicle-to-infrastructure communication.
- A frequency of 3.1 GHz is used.
- A more limited network range of only 20 m is obtained.
- A large bandwidth of 1000 Mbps is used, i.e., it is useful to realise multimedia networking.

CALM:

- Useful to realise vehicle-to-vehicle and vehicle-to-infrastructure communication.
- A frequency of 5.8 GHz is used.
- A larger network range of 1000 m is obtained.
- A bandwidth of 54 Mbps is realistic (comparable with Wi-Fi).
- The CALM technology consumes quite a lot of power.

8.3.2. Radio frequent wireless communication on the road

Research projects, performed by the academic world and the industry, are gathering experience concerning radio frequent wireless communication on the road. The paper by K. Selvarajah et al. (see the reference) desribes experiences gathered by the so-called ASTRA project. In the ASTRA project, a number of Crossbow MicaZ modules are used to allow radio communition using Zigbee. Such a module contains several components like an Atmel processor, memories, batteries. The modules are also equipped with an antenna to realize wireless Zigbee communication and an expansion connector to connect with sensor boards (which can contain sensors used to measure light intensity, temperature, barometric pressure, acceleration).

Experiments (see Fig. 8.1) have been performed in Newcastle, close to the Central Station, i.e., in real life environment where radio wave reflections occur due to buildings and a broad range of objects. Figure 8.1 shows a simplified view of the experiment. A number of vehicles are driving along the road (in different directions), and their modules are mobile like the vehicle (indicated by an encircled M in Fig. 8.1). There are also a number of fixed modules (e.g., attached to a bus stop), which are indicated by an encircled F in Fig. 8.1.

The experiment in Newcastle allowed to test vehicle-to-vehicle communication (communication between the first moving module visualised by an encircled M and the second moving module). The experiment allowed to test vehicle-to-infrastructure communication (communication between a moving module and a fixed module visualised by an encircled F).

Due to the limited communication range of Zigbee and the larger distances between the fixed modules, it is unlikely that the fixed modules are always able to communicate with each other directly. The fixed modules need to communicate indirectly with each other using the moving modules while they are sufficiently close to the desired fixed module. Also communication between the moving modules is possible when they are sufficienly close to each other.

The ability of modules to communicate with each other does not depend only on the distance. Radio reflections and interference can have an important impact on the link quality during the communication. The impact of a number of parameters on the link quality has been investigated.

- When increasing the power level of the transmitter, in general, the range increases and the number of lost data packets decreases (but increasing the power level can also stimulate undesired reflections).

- The impact of the weather is limited, but an increase of the range is noticed at lower temperatures.

- When using thicker materials to pack, more absorbtion of the radio waves occurs. This implies a lower signal strength and a smaller range.

- Reflections due to buildings and other objects can have a positive or a negative impact on the performance of the radio communication.

- The communication range can increase (and the data packet loss can decrease) by increasing the height of the radio modules.

- The orientation of the antennae has an impact on the recieved signal strength and on the data packet loss.

8.3.3. Combining radio frequency wireless communication and near-infrared optical communication

Developing a reliable vehicle-to-vehicle communication or a reliable vehicle-to- infrastructure communication is not that easy. By combining radio frequent wireless communication and near-infrared optical communication a much higher reliability can be obtained.

Figure 8.2 shows a situation where a road sign detection system intends to increase the safety of the car driver and passengers. In a traditional situation, the driver is expected to see the road sign without any technical assistance. In Fig. 8.2, the road sign is equipped with a radio frequent broadcaster and a near infrared optical emitter. The car is equipped with a radio frequency receiver and a near-infrared optical receiver. This allows to warn the driver of the presence of the road sign. Indeed, the information received by the radio frequent receiver and / or the near infrared optical receiver is processed by an on-board computer which allows to warn the driver (e.g., by visualising the road sign on the dashboard of the car). This warning is useful in case the driver does not see the road sign due to, e.g., bad weather conditions, objects blocking the view, distraction of the driver, etc.

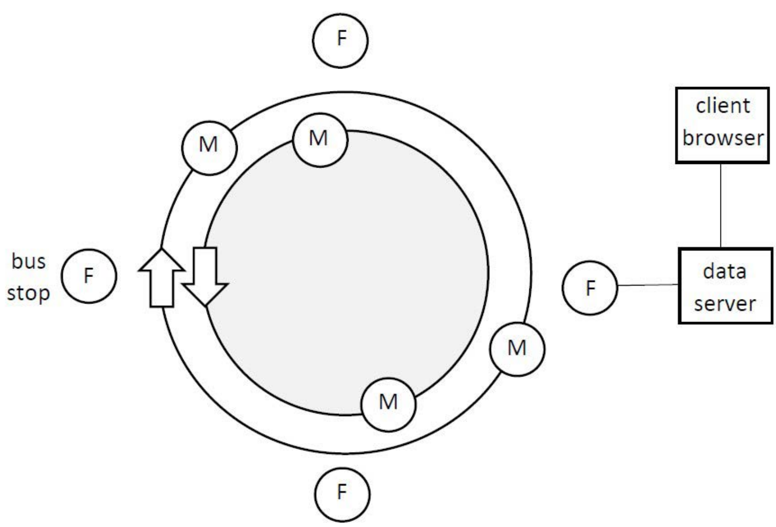


Fig. 8.1. Experimental setup in the ASTRA project.

Quite often, the near-infrared receiver picks up the signal prior to the radio frequency receiver. In such a situation, the radio frequency signal is a confirmation of the information picked up before by the near-infrared reciever. In some cases (e.g., in case of a bend), the radio frequency signal is received prior to the near-infrared signal. In some other cases (e.g., in case of a visual obstacle), possibly only the radio frequency signal is received.

Possibly, the near-infrared signal has a larger range than the radio frequency signal. But, e.g., in case of really sunny weather, the radio frequency signal often has a larger range. It is also interesting to observe that, in general, the near-infrared communication performs better when the car is travelling at high speed. The radio frequency communication generally performs better when the car is travelling at low speed. All these observations show that both technologies complement each other.

Both technologies complement each other. In case of radio frequency communication, it is difficult to detect the location and the direction of the detected object. In case of near-infrared communication, it is easier to detect the direction of the detected object (but the communication is susceptible to visual obstacles).

It is also important to avoid false detections of a warning. In case of an optical detection sensor, possibly a false detection occurs due to unappropriate weather conditions. For instance, in case of really sunny weather, reflections on the road can erroneously be identified as an object (e.g., a road sign).

In case of radio frequency communication, a false detection can occur due to a car travelling in the opposite direction. In case of radio frequency communication, it is difficult to detect the direction / position of the broadcaster. Several examples of undesired behaviours exist. When a car, driving in the opposite direction, communicates about a braking action, no action is needed. But due to the reception of the braking message, possibly a dangerous and unnecessary braking reaction will follow. The combination of different communication technologies helps to avoid false detections.

8.4. Broad range of applications

When dealing with intelligent transport systems, the number of applications is practically unlimited. A number of applications have already been discussed, including some technical background. In the present paragraph, some other applications will be briefly discussed.

8.4.1. Vehicle-to-infrastructure communication by emergency vehicles

In case of an emergency, emergency vehicles (ambulances, police cars, fire trucks, etc.) need to reach their destination as fast as possible. To reach that goal, they are allowed to neglect red traffic lights, which is actually quite dangerous. In case an emergency vehicle sends a signal to the nearby traffic lights, the traffic lights give priority to that emergency vehicle by giving a green signal to that emergency vehicle.

8.4.2. Transportation of dangerous goods

When dealing with safety related aspects of the road traffic, the transport of dangerous goods using trucks needs attention. Trucks are used to transport flammable liquids, gases (compressed, liquefied or dissolved under pressure), corrosive materials, etc. It can be good practice to monitor and manage in real time trucks transporting these dangerous goods.

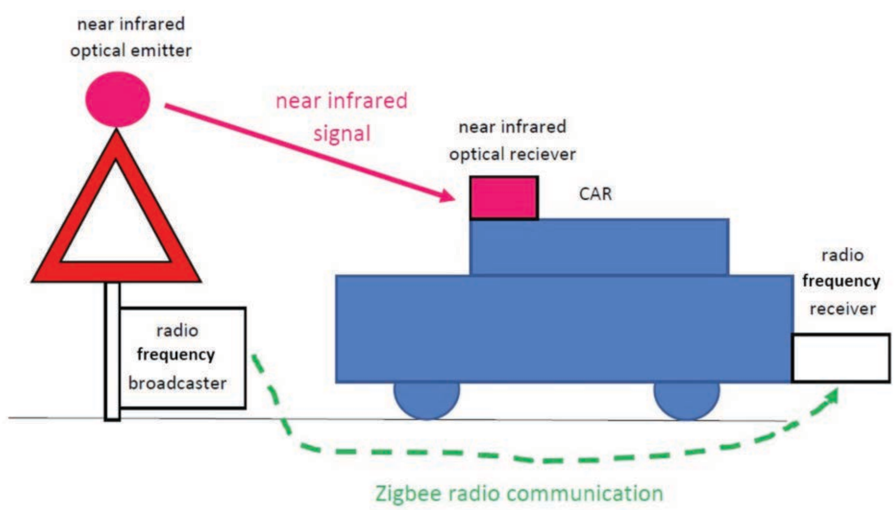


Fig. 8.2. Infrastructure-to-vehicle communication

A Transport Integrated Platform (TIP) typically consists of four subsytems: an on board unit, a transmisson system, a database, and a GIS based application.

The on board unit contains, mainly on the trailer of the truck, a number of sensors. These sensors measure important parameters giving information concerning the transported load (e.g., temperature, pressure, etc.). The measurement data originating from the sensors are collected and processed. The truck also contains a GPS antenna to have information concerning the location of the truck. Information is also available from the odometer but also, e.g., the speed and the direction of the truck is known.

The truck has a GPRS transmitter/receiver which allows to transmit real-time data (or information is transmitted, e.g,. every 5 minutes) to a remote server. Actually, different technologies can be used to transmit the data, since signal loss can occur due to GSM uncovering in some areas (e.g., satellite communication is a possibility).

The transmitted data is collected in a database. Efficient and reliable storage of the data is needed. The data must be available for diagnostic activities. Geographical Information System (GIS) analyses the data and displays them in a graphical way. The graphic interface allows scalability and allows an operator to retrieve the information. For instance, a control room allows a real-time monitoring of the trucks. The control room is able to communicate with the truck drivers during day and night. The control room contains tools for managing alarms and anomalies which are detected during the transportation.

Bibliography

B. Arief, A. von Arnim, TRACKSS Approach to Improving Road Safety through Sensors Collaboration on Vehicle and in Infrastructure, Proceedings of the 2nd IEEE International Symposium on Wireless Vehicular Communications (WiVeC 2008), Calgary, Canada, September 21–22, 2008.

M. Benza, C. Bersani, M. D'Incà, C. Roncoli, R. Sacile, A. Trotta, D. Pizzorni, S. Briata, R. Ridolfi, Intelligent Transport Systems (ITS) applications on dangerous good transport on road in Italy, 7th International Conference on System of Systems Engineering (SoSE), Genova, Italy, July 16–19, 2012.

E. Felici (author), D. Pieters (red), Intelligent Transport Systems (ITS), Agentschap NL – Ministerie van Economische zaken (retrieved from https://www.rvo.nl/sites/default/files/Intelligent%20Transport%20Systems.pdf).

Intelligent Transport Systems (ITS) for Road, Innovation and Networks Executive Agency (INEAA) of the European Commission (retrieved from https://ec.europa.eu/inea/sites/inea/files/2016_cef_h2020_its_trifold_web.pdf).

K. Selvarajah, B. Arief, A. Tully, P. Blythe, Deploying Wireless Sensor Devices in Intelligent Transportation System Applications, Intelligent Transportation Systems, Dr. Ahmed Abdel-Rahim (Ed.), ISBN: 978-953-51-0347-9, InTech, Available from http://www.intechopen.com/books/intelligent-transportation-systems/deploying-wireless-sensor-devices-in-intelligent-transportation-system-applications.

A. Tully, Pervasive tagging, sensors and data collection: a science and technology review for the foresight project on intelligent infrastructure systems, IET Intelligent Transport Systems, vol. 153, pp. 129–146, 2006.

A. von Arnim, M. Perrollaz, A. Bertrand, J. Ehrlich, Vehicle Identification Using Near Infrared Vision and Applications to Cooperative Perception, Proceedings of the 2007 Intelligent Vehicles Symposium, pp. 290–295, Istanbul,Turkey, June 13–15, 2007.

Bibliography: websites

Crossbow MICAz Wireless system: http://www.openautomation.net/uploadsproductos/micaz_datasheet.pdf

Industry 4.0: https://en.wikipedia.org/wiki/Industry_4.0

Smart Cities: https://en.wikipedia.org/wiki/Smart_city

Smart Cities: https://3bplus.nl/wat-een-smart-city-slimme-stad-een-introductie and https://3bplus.nl/technologie-een-smart-city-sensoren-iot-en-big-data

The four industrial revolutions: https://www.sentryo.net/the-4-industrial-revolutions/

Chapter 9:

EMC related aspects of cyber-physical systems in cars

Joan Peuteman

KU Leuven

9.1 Electromagnetic compatibility

EMC is the abbreviation for electromagnetic compatibility. Two electronic devices are electromagnetically compatible when the operation of the first device does not have an impact on the operation of the second device and vice versa.

EMI is the abbreviation for electromagnetic interference. EMI occurs when two electronic devices are operating in close proximity to each other and when the operation of one device has a (negative) effect on the operation of the other device. Examples of EMI are:

- The use of a mobile phone might disturb the electronic devices in an airplane needed by the pilots. Due to this reason, the use of mobile phones (and other electronic equipment) is forbidden during the takeoff and the landing of an airplane.

- The use of a mobile phone might disturb the operation of several electronic devices in a hospital (needed to save the life of a patient). Due to this reason, it is often forbidden or people are at least discouraged to use a mobile phone in a hospital.

- The use of a mobile phone or other electronic equipment might disturb the operation of electronic systems controlling the behaviour of the airbag, central locking system, audio systems, ABS system or other devices of a car.

In an airplane, it is realistic to forbid and avoid the use of mobile phones and other electronic devices of the passengers during the takeoff and the landing. Due to the rise of mobile phones and other mobile electronic devices, it becomes harder to avoid the use of mobile electronic devices by patients and visitors in a hospital. Especially in a car or a bus, it is really impossible to avoid the use of mobile electronic devices by the drivers and the passengers. This implies that avoiding EMI is very important when designing cyber-physical systems for cars, i.e., the electronic devices of the drivers and the passengers must not to disturb the proper working of the electronic devices (also the cyber-physical pystems) integrated into the car. Moreover, it is important that the electronic devices (also the cyber-physical systems) integrated in the car by the car manufacturer do not disturb each other. Application of the EMC philosophy is needed to reach these goals.

"You will not disturb and you will not be disturbed" is the main principle inside the engineering discipline EMC. EMC is often defined as follows (International Electrotechnical Vocabulary):

"The ability of a device, equipment or system to function satisfactory in its electromagnetic environment without introducting intolerable electromagnetic disturbance to anything in that environment."

Due to technical evolutions, without precautions it becomes harder to satisfy the EMC philosophy. In a lot of devices / applications, the number of electronic parts has increased steadily and this evolution will continue in the future. For instance, a car (also a bus or a truck) contains a large amount of electronic parts (including sensors, electronic control units, actuators, etc.). Each electronic part (possibly) causes an emission of disturbances, and care must be taken in order to prevent failure of other parts.

Due to technical evolution, people more and more depend on electronic applications to guarantee their safety. This evolution is also visible when designing cars, autonomous cars (driverless cars, self-driving cars), automated guided vehicles, etc. Especially in such situations, it is important that the electronic devices (and the cyber-physical systems) have no EMI related problems.

9.1.1. Emission and immunity

On the one hand, more and more electronic devices intend to emit disturbances. The so-called emission level increases. For instance, the traditional incandescent lamp only accounts for a very limited emission of disturbances. Notice, however, the rise of power electronics, clocked processors, and also mobile devices requiring wireless communication, which accounts for much more electromagnetic emissions.

On the other hand, electronic devices are more and more susceptible to disturbances, i.e., their so-called immunity level has decreased. Indeed, the transition from vacuum tubes to semiconductor transistors, from 5V technology to 3V technology (or even lower voltage levels), etc. imply that electronic devices are more and more vulnerable to disturbances.

Figure 9.1 visualises the natural increase of emission level in the last century. The same Fig. 9.1 also visualises the natural decrease of the immunity level in the last century. In case the emissions level exceeds the immunity level, one device will disturb the proper operation of another device. Thus, EMI related problems arise.

The evolution visualised in Fig. 9.1 illustrates that the emission level increases as a function of time and that the immunity level decreases as a function of time. At a certain instant of time, both curves intersect and (undesired) EMI occurs. In the EU, the EMC-Directive intends to avoid an intersection of the curves depicted in Fig. 9.1 by

- keeping the emission level sufficiently low,
- keeping the immunity level sufficiently high.

The EMC-directives intend to realise the situation visualised in Fig. 9.2. The emission level is not allowed to exceed the immunity level, and even a gap is needed between both levels. Legislations in other parts of the world (e.g. FCC in the US) tend to achieve the same goal.

Until now, we mainly emphasized that a first device is not allowed to disturb the proper operation of a second device. This is correct, but a device is also not allowed to disturb its own operation, i.e., also the so-called 'intra-system EMC' is very important.

9.1.2. Emission

When talking about emission, it is important to make a distinction between conducted emission and radiated emission. An electronic device can emit disturbances using conductors (conducted emission), i.e. undesired currents are flowing in the conductors causing voltage drops. When the device is fed by the electrical power grid, a distinction can be made between

- harmonics (low frequency, multiples of the 50 Hz grid frequency),
- high-frequency disturbances having frequencies higher than 9 kHz,
- changes in the voltage level of the power grid and flicker.

When considering a car, the electronic devices are generally fed by a battery providing a DC voltage. Cyber-physical systems (electronic devices, electrical loads) can also emit high frequent disturbances using the conductors which feed all loads with the DC voltage of the battery. Due to changes in the power consumption, i.e., by changes in the consumed current, the supply voltage level changes. These changes in the voltage level can also have an impact on the operation of other devices.

An electronic device generally also accounts for radiated emission. The conductors and other components behave as antennas (transmitting aerial) causing radiated emission in the environment/atmosphere. Electromagnetic waves are emitted, and these electromagnetic waves have an impact on other conductors and components which behave as receiving antennas. Due to the voltages generated by these receiving antennas, the proper operation of electronic devices can be disturbed.

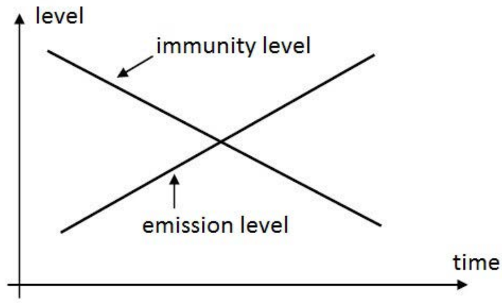


Fig. 9.1. Natural evolution of emission and immunity levels.

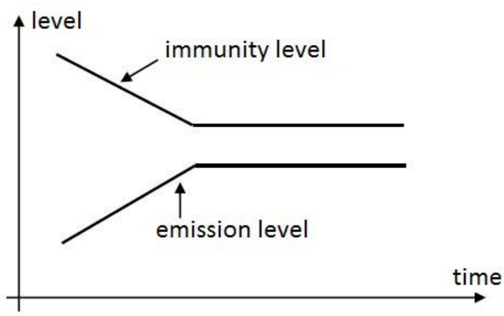


Fig. 9.2. EMC-directives limit emission and immunity levels.

9.1.3. Immunity

When considering immunity, it is also important to make a distinction between disturbances which propagate using conductors (immunity against conducted disturbances) and disturbances which propagate using the atmosphere (immunity against radiated disturbances).

When considering, e.g., the battery power supply for the electronic devices in a car, immunity against changing supply voltage levels is important (e.g., voltage dips can occur). Also immunity against incident electromagnetic waves, magnetic fields and electrostatic fields is important.

Electromagnetic immunity indicates to what extent an electronic device (cyber-physical system) is able to withstand the influence of disturbances. In case the proper operation of the device is disturbed, there are a number of possibilities. Possibly, the proper operation is temporarily disturbed and the device behaves normally once the disturbances have disappeared.

The situation can be worse. Possibly, the device will not behave normally once the disturbances have disappeared. The device only behaves normally after resetting or switching the device off and on. The user must interact to restore the normal behaviour.

The worst situation occurs when the disturbances damage the device permanently, i.e., a normal reset or switching the device off and on does not solve the problem. In such a situation, the device must be repaired or replaced.

Another criterion is the question to what extent the undesired behaviour of the device harms the safety of people (e.g., consider a medical device in a hospital but also a safety critical device in a car). Does the EMI related problem cause material damage and what are the consequences of this damage?

9.2. Automotive electronics

The number of electronic devices and components in a car is increasing fast, which implies ever new EMC related challenges. A car contains power supply units, electronic control units (ECU), connections to sensors and actuators, data communication networks, etc.

When focusing on the communication needs, a distinction can be made between wired and wireless communication. Communication between different devices in one single car, communication between the car and road side infrastructure and also communication between different cars are needed.

Electronic control units are almost everywhere in a modern car. The engine control unit is a classical example, but ECUs are also used to control transmissions, airbags, ABS systems, cruise control systems, electric power steering systems, audio systems, mirror adjustment systems, recharging systems for hybrid / electric cars, etc.

The number of applications is ever increasing, we restrict ourselves to a limited number of examples.

- Auto start/stop system: A number of sensors (e.g., speed sensors, steering angle sensors, etc.) determine whether the engine can be shut down in order to save fuel and to reduce harmful exhaust gases.

- Hill-hold control: A tilt sensor detects the tilt of the car when standing on a slope. The wheels are kept clamped for a number of seconds after the driver has released the brake pedal.

- Rear park assist system: The system is useful while parking the car and engaging the reverse gear. By using ultrasonic sensors on the rear bumper, the driver is warned when other cars or objects are too close to the car.

9.2.1. Electronic control unit and EMC related aspects

Figure 9.3 visualises a block diagram of an electronic control unit. Notice the power supply which is fed by a DC voltage source (e.g., originating from the battery) which can be highly volatile. This non constant DC voltage is converted into different voltage levels. Changes in the original DC voltage level must not impact the final voltage levels (i.e. they are constant). To change the voltage level, linear voltage controllers or switched DC/DC converters are possible. Although switched DC/DC converters (SMPS – switching mode power supply) have a higher efficiency and they also provide the possibility to boost the voltage level, they can be an important source of EMC related problems due to the switching of the semiconductors in the converter.

Common voltage levels needed for the digital circuits are 5 V, 3.3 V, 2.4 V, 1.8 V or even lower. When considering analog circuits, other voltage levels are more common. When considering sensors, voltages of 5 V or 12 V are quite common. When considering actuators, often higher voltage levels are needed (e.g., some fuel injectors need a voltage up to 200 V, which is much higher than the original supply voltage).

Notice the presence of a number of sensor inputs interfacing with digital and analog sensors. When considering a modern car, the range of sensors is very broad: speed sensors, acceleration sensors, tilt sensors, rain sensors, temperature sensors, seat belt sensors, ultrasonic sensors to detect objects, etc. Multiplexers are needed to select the required sensors. Analog to digital converters (ADCs) are needed to convert signals originating from the analog sensors to digital signals which can be processed.

Power drivers are needed to drive a large range of actuators. Especially when considering the power drivers, care must be taken to drain off the heat losses. The power drivers also need much attention concerning EMC (especially from the point of view of emission requirements). Especially when high currents are switched, large disturbances might be generated.

Figure 9.3 also visualises a number of transceivers which allow the electronic control unit to communicate with other devices. Several communication protocols are possible, but the use of the CAN bus or FlexRay are common choices when dealing with electronic devices in cars (FlexRay is designed to be more reliable and faster than the more traditional CAN bus).

9.2.2. Clock signals

When considering an electronic control unit as visualised in Fig. 9.3, notice the presence of a clock needed to clock the microcontroller (belonging to the digital core) and other digital components. An ideal clock signal is a rectangular signal which often has a duty cycle of 50 % . In real life a clock signal merely has as trapezoidal shape as visualised in Fig. 9.4. Notice the rise time t_r and the fall time t_f (an ideal clock signal has rise and fall times which equal zero).

If the timing of the rising edge or the timing of the falling edge of the clock signal is not accurate, jitter occurs. Due to this jitter phenomenon, clocked digital hardware will act too early of too late. Figure 9.5 visualises a practical example where a processor reads data from a data bus at the positive edge of the clock. In case the ideal clock signal is used, the processor samples valid data from the data bus. In case jitter occurs, the processor samples the data too early, i.e., data is sampled before it is valid. Electromagnetic interference can impair jitter, which implies a higher risk that digital hardware will act too early or too late.

Additionally, electromagnetic interference can cause voltage spikes in the clock signal. These spikes can be misinterpreted as additional clock signal edges. Sometimes different integrated circuits (processors) are hit in a different way, by spikes implying a distortion of the global operation of the entire system.

9.2.3. Analog sensors

Proper working of an electronic control unit, as visualised in Fig. 3, depends on information originating from analog and digital sensors. Automotive microcontrollers contain analog to digital converters which convert the analog information originating from analog sensors to digital values. These digital values are stored in registers allowing to process the information. Quite often, the number of analog inputs is larger than the number of analog to digital converters, implying the need for analog multiplexers to select the appropriate sensor input.

When considering analog sensors, a distinction can be made between resistive sensors, capacitive sensors and inductive sensors. Especially resistive sensors are frequently used in the automotive industry. Resistive sensors convert a physical value (e.g., temperature, magnetic field strength, light intensity, etc.) into a resistive value.

Notice in Fig. 6 the resistive sensor providing resistance R_S, which depends on the physical value, which needs to be measured. Notice resistance R_2 connected

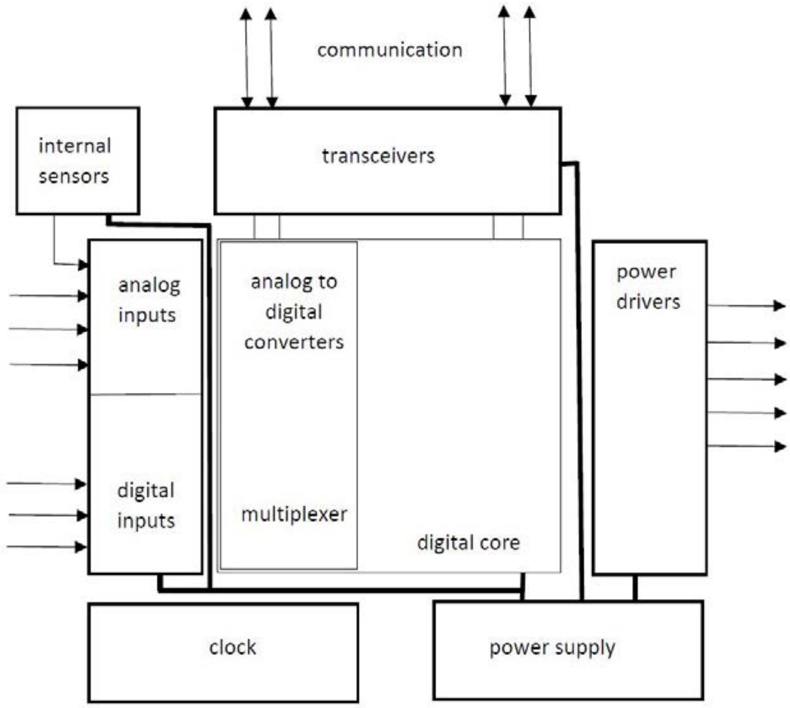


Fig. 9.3. Block diagram of an electronic control unit.

with the power supply and the resistive sensor R_S. A voltage divider obtained by R_2 and R_S implies a voltage across R_S which depends on the measured physical value. No real current is flowing in resistor R_1, since R_1 only intends to protect against voltages higher than the power supply. Notice also capacitor C which diverts high-frequent disturbances (also burst pulses are diverted). The controller contains an ADC converting the input voltage across R_S in a digital value which can be stored. Notice also the presence of additonal protection circuits (indicated by the black rectangle in Fig. 9.6).

Capacitive sensors are less common when considering car electronics. Some engines have oil quality sensors, where the oil quality has an impact on the permittivity. By measuring the capacitor value, the permittivity is obtained which gives information about the oil quality. Capacitive humidity sensors are also used implying, e.g., the possibility to detect rain. Acceleration sensors can also contain capacitive micro-mechanical sensors (e.g., to be used in an airbag to detect a car crash).

Inductive sensors are, just like capacitive sensors, not that common when

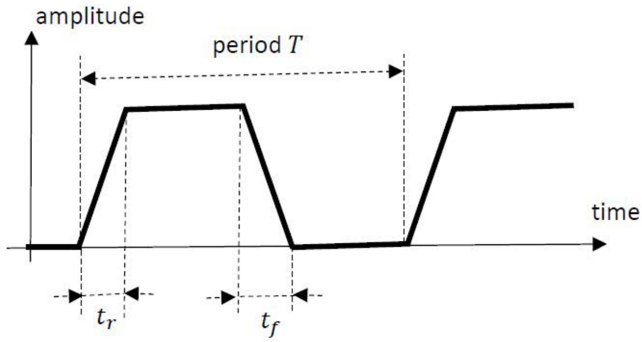


Fig. 9.4. Clock signal.

considering car electronics. Inductive sensors are used to measure, e.g., the motor speed or the wheel speed. Using a permanent magnet and an induction coil, voltage pulses are induced. By counting these voltage pulses, information about the speed is obtained.

9.2.4. Digital sensors

Digital inputs are actually switches which are open or closed. By replacing the resistive sensor R_S in Fig. 6 by a switch, a digital input is obtained. Typical examples of digital switches in a car are brake pedal switches or door contact switches. From a broader point of view, bus interfaces are also digital outputs or inputs.

9.2.5. Power drivers

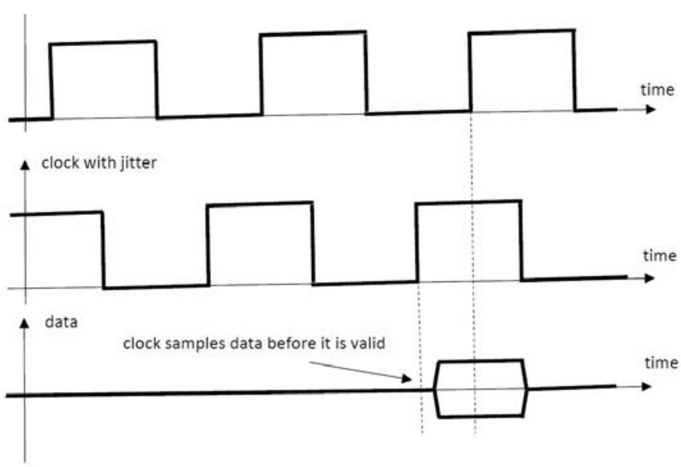


Fig. 9.5. The impact of jitter when reading data.

Based on information originating from the input sensors and the communication with other devices, actuators are switched on or off by the electronic control unit using a transistor. Bipolar transistors can be used, but MOSFETs (metal-oxide-semiconductor field-effect transistors) are more common. In case of high-voltage applications, mainly IGBTs (insulated gate bipolar transistors) are used.

In case the transistor conducts (operates in saturation mode, i.e., switch S in Fig. 9.7 is closed), the supply voltage is available across the actuator which is an electrical load. In case the transistor does not conduct (operates in cut off mode, i.e., switch S is open), no current is flowing in the load.

By opening and closing switch S with an appropriate duty cycle, a pulse width modulated voltage (PWM voltage) appears across the actuator (the load in Fig. 9.7). By changing the duty cycle, the mean value of the voltage can be controlled.

9.2.6. Transceivers

It is important that electronic control units are able to communicate with other electronic control units using a bus system. In general, data can be sent in both directions implying the need of transceivers (combining a transmitter and a receiver).

When considering a car, the communication buses are operating in a harsh envirmonment, i.e., the bus systems face a lot of disturbances. For instance, a CAN bus (controller area network) is often used to communicate between

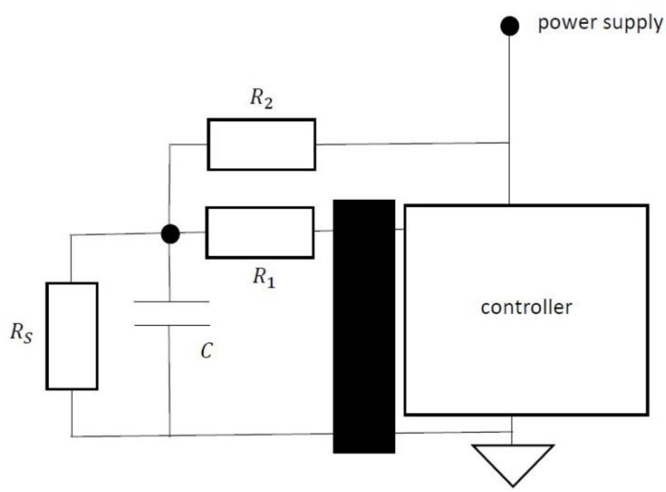


Fig. 9.6. Analog input obtained by a resistive sensor.

automotive ECUs. In the next section, the main properties of the CAN bus will be considered. When considering this CAN bus, attention goes to the reliability of the communication and the robustness against electromagnetic disturbances. Alternatives for the CAN bus are the use of a MOST (media oriented systems transport) network or FlexRay.

9.3. The CAN bus

As already mentioned, the CAN bus is often used for data communication in the automotive industry. Attention goes to the reliability of the communication and the robustness against electromagnetic disturbances.

In general, data exchange between, e.g., automotive ECUs is not based on analog signals. Analog signals are too sensitive to electromagnetic interference (EMI). Digital signals are used to transmit data. As long as the EMI amplitudes are rather small, the information in the digital signals remains unchanged. As the EMI amplitudes increase, the bit error ratio BER increases. The bit error ratio BER is the ratio between the number of corrupted bits and the total number of bits.

In case the EMI amplitudes are rather small, the BER increases slowly as the EMI amplitudes increase. But when the EMI amplitudes exceed a threshold, the BER increases quickly. Finally, no communication is possible anymore.

Bits are transmitted in adjacent groups called messages, packets or data frames. As the BER increases, an increasing number of frames become erroneous. The majority of the protocols try to detect erroneous frames or even try to correct these erroneous frames. In case an erroneous frame is detected, the frame can

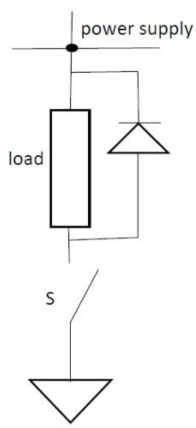


Fig. 9.7. Power driver with actuator.

be retransmitted or the data in that frame is simply omitted. For instance, the CAN bus shuts down communication nodes which provide too many erroneaous frames.

9.3.1. Reducing EMI: electrostatic coupling

In order to reduce EMI related problems, a lot of serial buses use an UTP cable (unshielded twisted pair), a coaxial cable, or a STP cable (shielded twisted pair). In order to understand the use of such types of cables, we consider the origin of electrostatic and magnetic coupling. We also study approaches to make the communication less vulnerable to disturbances.

Consider the situation visualised in Fig. 9.8. A voltage source $e(t)$ sends a signal to an amplifier or a receiver using a conductor (conductor 1). Analog and digital signals can be considered. In the proximity of conductor 1, other electric or electronic devices account for disturbances. The origin of these disturbances is modelled by the voltage source $e_D(t)$, load R_D and a conductor (conductor 2). Due to the presence of a parasitic capacitor C, the voltage source $e_D(t)$ has an impact on conductor 1. Actually, the input of the amplifier / receiver is obtained by applying the superposition theorem with $e(t)$ and $e_D(t)$.

In general, the parasitic capacitor C is small. In case $e_D(t)$ contains steep slopes corresponding to high frequent components, the capacitor has for these high frequencies a low impedance, implying that these high frequent components of $e_D(t)$ are also applied at the input of the amplifier/receiver. When considering low frequent components, the small parasitic capacitor C behaves as an open circuit, implying that these low frequencies have no impact on the amplifier/receiver.

Figure 9.9 visualises how the impact of $e_D(t)$ on conductor 1 and on the input of the amplifier/receiver can be reduced drastically. When having a shielding around conductor 1, the parasitic capacitor C is situated between conductor 2 and that shielding. It is important that the shielding around conductor 1 has a proper grounding.

The behaviour of the configuration visualised in Fig. 9.9 is visualised in Fig. 9.10. Notice again the voltage source $e(t)$ which is connected with the input of the amplifier/receiver by conductor 1. Due to the grounded shielding, the parasitic capacitor C connects $e_D(t)$ with the ground. This implies that the impact of the high frequent components of $e_D(t)$ on conductor 1 and the input of the amplifier/receiver is avoided.

9.3.2. Reducing EMI: magnetic coupling

Consider the situation visualised in Fig. 11. Voltage source $e(t)$ sends a signal to an amplifier or a receiver using conductor 1. In the proximity of conductor 1, other electric or electronic devices account for disturbances. The origin of these disturbances is modelled by the voltage source $e_D(t)$, load R_D and conductor 2, where current $i_D(t)$ is flowing. Due to the presence of mutual inductance M, current $i_D(t)$ has an impact on conductor 1. Actually, voltage

$$M \frac{d\, i_D(t)}{d\, t}$$

is induced in conductor 1, as visualised in Fig. 9.12.

Notice in Fig. 9.12 that the induced voltage is connected in series with the original voltage $e(t)$ and applied at the input of the amplifier/receiver. Especially the high frequent components in $i_D(t)$ have an impact, since in general the mutual inductance M is small.

The mutual inductance M can be reduced by having a larger distance between conductor 1 and conductor 2. The impact of the mutual inductance can also be reduced by using two conductors between voltage source $e(t)$ and the input of the amplifier / receiver. It is very important to twist these two conductors (giving a twisted pair), as visualised in Fig. 9.13. Notice that the voltage

$$M \frac{d\, i_D(t)}{d\, t}$$

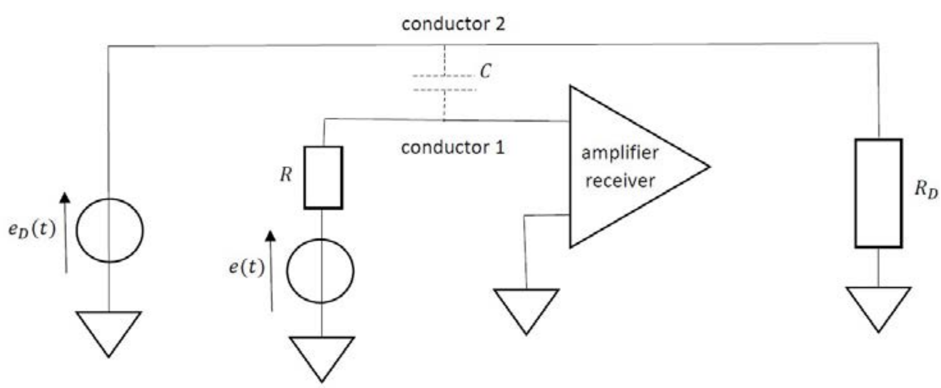


Fig. 9.8. EMI based on electrostatic coupling.

is induced twice, but taking the polarities into account, both voltages cancel each other. Actually the voltages are induced in each loop, and in each loop they cancel each other.

Figure 9.13 visualises the use of an UTP cable (unshielded twisted pair) which is often used in a CAN bus. Voltage $e(t)$ is the voltage generated by the transmitter and applied on the cable / bus. Notice the receiver at the other side of the CAN bus.

In case EMI based on electrostatic coupling and magnetic coupling needs to be avoided, a STP cable (shielded twisted pair) can be used. Due to the shielding, electrostatic coupling will be eliminated, and due to the twisting, magnetic coupling will be eliminated.

9.3.3. Physical layer of a CAN bus

A CAN bus is a serial bus where EMI is reduced by using an UTP cable (unshielded twisted pair) and by using differentially driven lines, as visualised in Fig. 9.14 for a low-speed CAN bus and Fig. 9.15 for a high-speed CAN bus. Physically, the bus contains a CAN-high (*CAN_H*) wire and a CAN-low (*CAN_L*) wire.

Figure 9.14 visualises the physical layer of the low-speed CAN bus. An idle voltage corresponds to 0 V on *CAN_H* and 5 V on *CAN_L*. In case of a logical 0, the *CAN_H* is pulled up to more than 3.6 V and the CAN_L is pulled down to less than 1.4 V which implies a differential voltage of more than 2.2 V. In case of a logical 1,0 V on *CAN_H* and 5 V on *CAN_L* is chosen in order to have a differential voltage of –5 V. In case of the low-speed CAN bus, the maximum data rate equals 125 kbit/s.

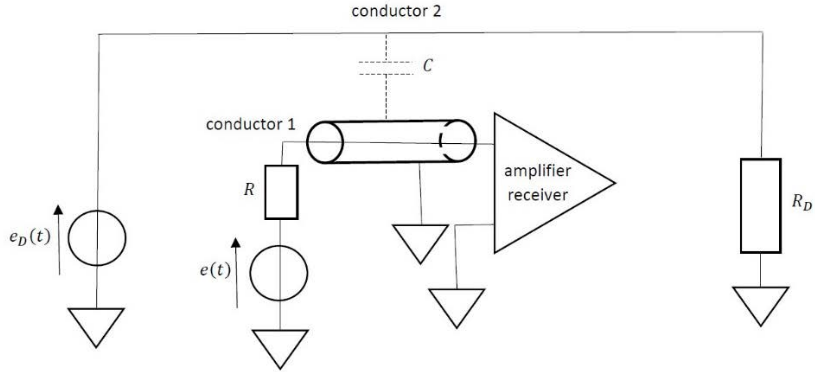


Figure. 9.9. Reducing EMI based on electrostatic coupling.

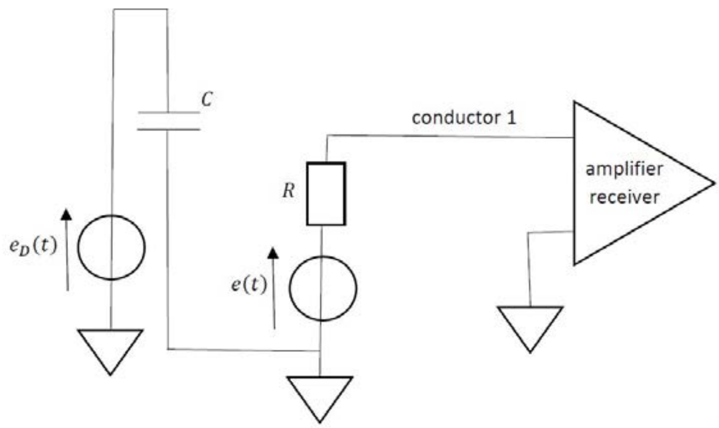


Fig. 9.10. Reducing EMI based on electrostatic coupling.

Figure 9.15 visualises the physical layer of the high-speed CAN bus. An idle voltage corresponds to 2.5 V on CAN_H and 2.5 V on *CAN_L*. In case of a logical 0, *CAN_H* is pulled up to more than 3.5 V and *CAN_L* is pulled down to less than 1.5 V which implies a differential voltage of more than 2 V. In case of a logical 1, a 2.5 V on CAN_H and a 2.5 V on *CAN_L* is chosen in order to have a differential voltage of 0 V. In case of the high-speed CAN bus, the maximum data rate equals 1 Mbit/s (but 500 kbit/s is more typical).

Notice that the EMI of the CAN bus is reduced by using an UTP cable, which reduces the impact of magnetic coupling. By using differentially driven lines, the impact of electrostatic coupling is reduced (i.e. normally no shielding is needed, no STP cable is needed). In a few cases with high data rates (used for multimedia applications), a STP cable can be used.

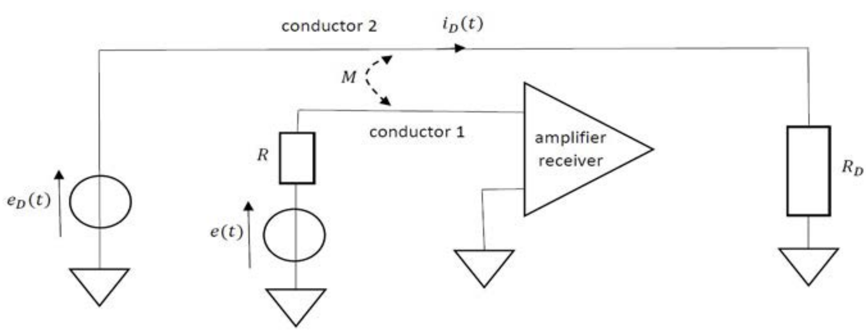


Fig. 9.11. EMI based on magnetic coupling.

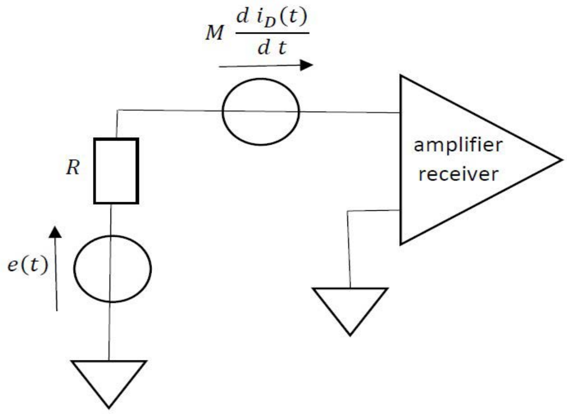


Fig. 9.12. EMI based on magnetic coupling.

Notice that in Fig. 9.14 (low-speed CAN bus) and in Fig. 9.15 (high-speed CAN bus) differential voltages are used, which are available between the twisted lines of, e.g., the UTP cable. Since a differential voltage is used, the impact of electrostatic coupling is eliminated (or reduced considerably). As visualised in Fig. 9.16, twice the same voltage has been induced having no impact on the differential voltage. Figure 9.16 visualises the situation of a high-speed CAN bus, but an identical situation occurs in the case of a low-speed CAN bus.

9.3.4. Transmitter, receiver and termination of a CAN bus

Figure 9.17 visualises a simplified version of a CAN bus containing transmitter *Tx* and receiver *Rx*. In case the transmitter transmits a zero (0), both MOSFETs will conduct. The upper MOSFET in Fig. 9.17 is a P-channel enhancement MOSFET. In

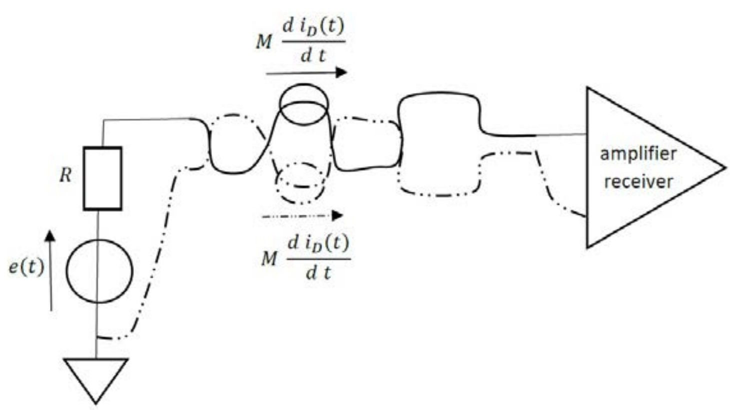


Fig. 9.13. Reducing EMI based on magnetic coupling.

case a low voltage is applied to gate G, a conducting P-channel is obtained, i.e., the MOSFET will conduct and apply a high voltage to the *CAN_H* wire. Due to the inverter, a high voltage is applied to gate G of the lower MOSFET, which is an N-channel enhancement MOSFET. The lower MOSFET will conduct and apply a low voltage to the *CAN_L* wire. As visualised in Figure 9.15 (high-speed CAN bus), this implies indeed a zero (0) is applied to the CAN bus.

In case the transmitter transmits logic 1, the MOSFETs will not conduct. The high voltage applied to gate G of the upper MOSFET does not create a conducting channel, i.e., the upper MOSFET will not conduct. Due to the inverter, a low voltage will be applied to gate G of the lower MOSFET, which does not create a conducting channel, i.e., the lower MOSFET will not conduct. Since both MOSFETs do not conduct, the *CAN_H* and *CAN_L* wires both have a voltage level of 2.5 V which corresponds to logic 1 in Fig. 9.15.

The applied signal propagates through the CAN bus to reach receiver *Rx* where the differential voltage is sent to a comparator. When receiving logic 0, the voltage level of *CAN_H* is higher than the voltage level of *CAN_L* implying a high voltage at the output of the comparator. Due to the inverter after the comparator, finally logic 0 is obtained at *Rx*.

Notice also the presence of termination resistors in order to avoid reflections, since the CAN bus behaves as a transmission line. Termination resistors are available at both most distant points of the CAN bus. The characteristic impedance of this transmission line is about 120 Ω for a differential mode signal.

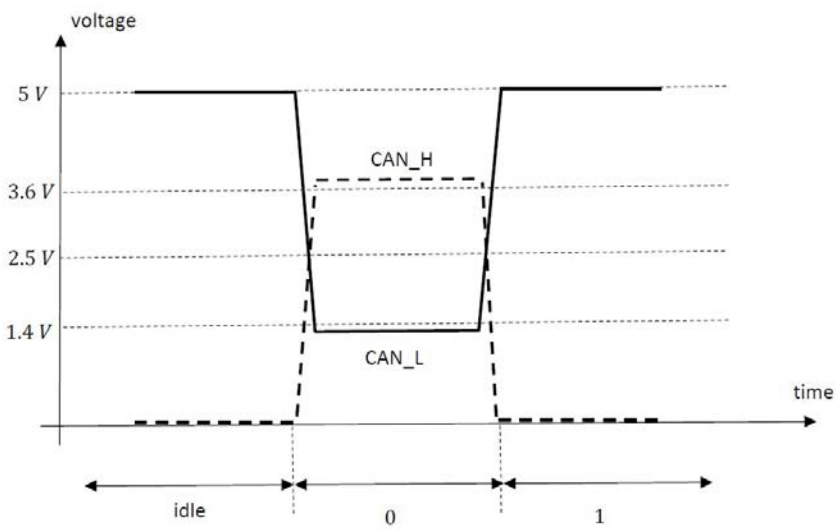


Fig 9.14. Voltage levels of a low-speed CAN bus.

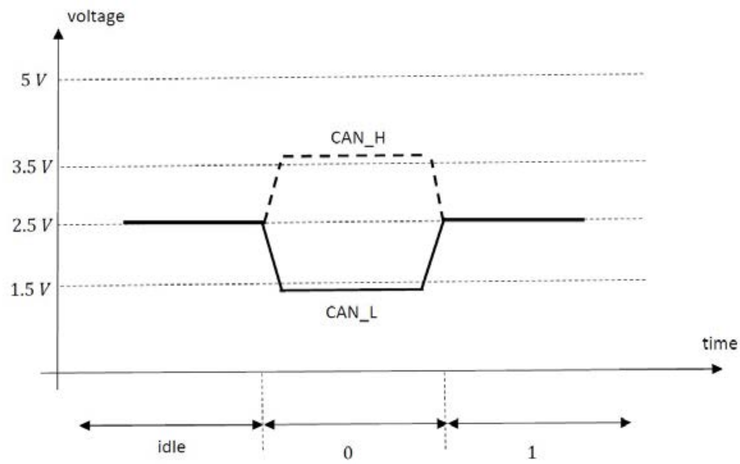


Fig. 9.15. Voltage levels of a high-speed CAN bus.

Termination is obtaind by using two resistors R = 60 Ω at both distant points.

9.3.5. Data frames of a CAN bus

A discussion of the entire working principle of the CAN bus goes beyond the scope of the present text. But it is important to know that a number of bits are combined in a frame. Although there exist four types of CAN-frames, we will restrict ourselves to data frames intended to transmit data from transmitter to receiver.

A frame starts with a single start-of-frame bit which equals logic 0. Logic 0 is dominant over logic 1. Indeed, consider the CAN bus in Fig. 9.17 and assume two transmitters trying to transmit a bit at the same time. If one of them transmits logic 0 and the other one transmits logic 1, finally logic 0 will be applied to the CAN bus. The transmitter which tried to transmit logic 1 is able to detect the dominant 0 of the other transmitter. This dominance property is also crucial when the identifiers in the frame determine which frame has the highest priority.

A standard frame contains an identifier with a length of 11 bits. A lower identifier implies the frame has a higher priority (based on the dominance principle of logic 0). The identifier is followed by a remote transmission request bit. This bit equals 0 in the case of a data frame. The remote transmission request bit is followed by the identifier extension bit which is logic 0 in case an identifier of 11 bits has been used. The identifier extension bit is followed by a reserved bit which equals logic 0.

The reserved bit is followed by the data length code. The data length code

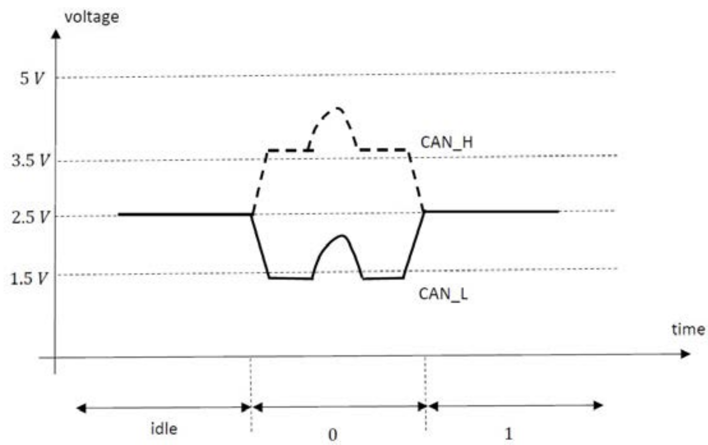


Fig. 9.16. Voltage levels of the high-speed CAN bus in the case of electrostatic coupling.

needs 4 bits and mentions the number of data bytes in the data frame (4 bits are needed to have between 0 and 8 bytes). The data field finally contains the data bytes which need to be transmitted. The data field has a length between 0 and 8 bytes, as indicated by the data length code.

After the data field, a CRC code (cyclic redundancy check) with a length of 15 bits has been included. This CRC code allows to detect errors in the data due to EMI related problems. A reliable data transmission is stimulated not only by using UTP cable and differentially driven lines but also by using this CRC code.

The data frame ends with a CRC delimiter (1 single bit), an ACK slot (1 single bit), an ACK delimiter (1 single bit) and, finally, an EOF (end-of-frame) having a length of 7 bits.

As already mentioned, immunity against EMI related problems is stimulated by using UTP cable, a differentially driven bus and a CRC code. In case still something is going wrong, an ECU can warn other ECUs by using special error detection frames.

9.4. EMC design hints

When designing electronic devices, cyber-physical systems (and electronic control units in particular) in cars, it is a good attitude to respect a number of EMC design hints in order to reduce the emission levels and to increase the immunity levels.

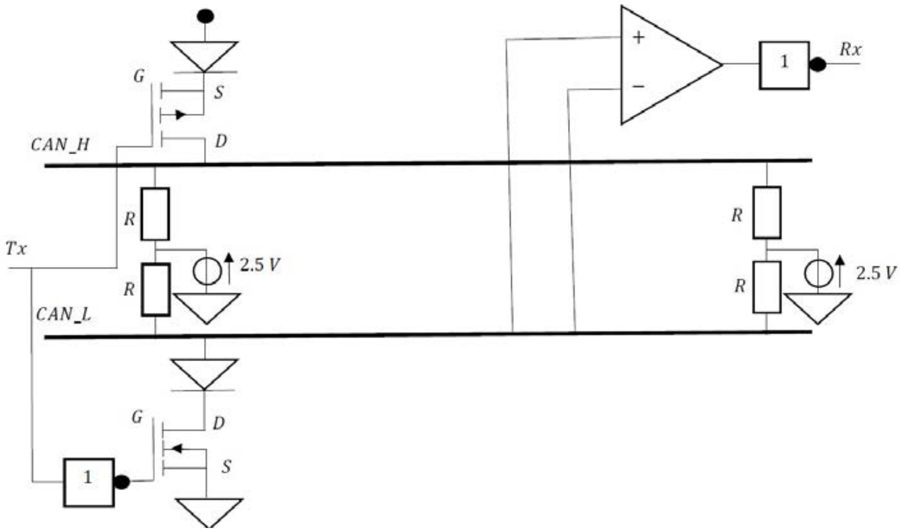


Fig. 9.17. CAN bus with transmitter and receiver.

9.4.1. A number of hints when designing a printed circuit board and electronic circuits

When designing a PCB (printed circuit board), it is important to have a low ground impedance and a low power impedance. A full plane is preferred above a single trace. A star-shaped power distribution network is preferred above a bus-shaped power distribution network. It is a good practice to have separate analog and digital parts.

It is important to keep loop areas as small as possible. Each conductor behaves as a loop in combination with its return conductor. When designing a PCB, the loop area is minimized by having a return trace which is close to the conductor. Remember a UTP cable, twisting the conductors is also an approach which reduces the loop areas as much as possible. By reducing the loop area, smaller voltages are induced, since the time-varying magnetic flux in the loop becomes smaller.

When using capacitors, it is a good practice to keep the capacitor wires as short as possible. The capacitor wires behave as inductors in series (approximately 1nH/mm) with the actual capacitor. This implies the behaviour of a series RLC resonant circuit is obtained. When using surface mounted capacitors (SMDs), a smaller inductive behaviour is obtained. In the case of a surface mounted capacitor, it is important to avoid stub traces to the capacitor, since they also behave as inductors in series.

When using inductors, it is a good practice to avoid long wires in parallel. A parasitic capacitance is obtained between the inductor and the wire, which implies a parallel resonant circuit is obtained. Actually, the use of inductors is often avoided. In general, inductors generate a magnetic stray field which couples into the neighbourhood. Only inductors with a closed magnetic loop (toroidal inductors) do not generate a magnetic stray field. These toroidal inductors are expensive. Moreover, inductors are, in general, quite expensive and they account for a large volume and weight.

In order to reduce EMI, it is a good practice to keep a sufficiently large distance. By increasing the distance between components and conductors, the parasitic capacitance (as, e.g., in Fig. 9.8) decreases. By increasing the distance between components and conductors, the mutual inductance (as, e.g., in Fig. 9.11) decreases.

It is a good practice to keep subsystems separated from each other. Subsystems are kept separated from each other by using shieldings, and filters and by keeping a distance between the subsystems.

When designing electronic circuits, it is a good practice to have knowledge of the frequencies which occur in the system. Some frequencies are desired but others are undesired. It is important to avoid overlaps in the frequency spectrums. If, e.g., a specific frequency is used to realise data communication and there is a disturbance having the same frequency component, then EMI will occur if the disturbance is too large. In case the frequency used to realise the data communication differs from the disturbing frequency, normally no EMI will occur.

When considering Figs. 9.8 and 9.11, the data communication is based on electrical signals. By replacing copper conductors with electrical signals by an optical communication channel, EMI can be avoided. Although optical communication is really robust with respect to electromagnetic disturbances, there are also a number of drawbacks: installation problems, lack of standardisation, high cost. Also in the case of optical communication, EMC related aspects of the transmitter and receiver must be taken into consideration.

9.4.2. Reducing the impact of electromagnetic waves in the far field

When considering an electronic control unit in a car, it can have a closed metal case. The metal case helps to conduct heat, but it also reflects impinging electromagnetic waves in the far field. Far field shielding is limited due to openings in the metal case. When having an electromagnetic wave with higher frequencies (i.e. having a smaller wavelength), the electromagnetic wave only needs smaller openings to enter. Especially the connectors account for openings in the metal case. Important also is the fact that the metal case generally consists of two halves which are joined together. When these halves do not overlap, a narrow but long slot / opening in the metal case is obtained allowing electromagetic waves to enter the metal case and disturb the behaviour of, e.g., the electronic control unit.

A closed metal case can be replaced by a plastic case, but a thin metal layer can help to reflect the impinging electromagnetic waves in the far field.

Conclusions

When considering cyber-physical systems in a car, a broad range of applications arise. In order to obtain a robust and reliable operation, EMC related aspects must be taken into consideration in order to obtain a sufficiently large immunity against a broad range of electromagnetic disturbances.

Bibliography

H. Balzer, P. Fischer, M. Lutz, J. Kappert, EMC: Immuniteitstesten en meetmethoden, De Muiderkring, Weesp, 1993 (in Dutch).

K. Borgeest, EMC and Functional Safety of Automotive Electronics, IET, Herts – United Kingdom, 2018, ISBN 978-1-78561-409-5.

J.J. Goedbloed, Elektromagnetische compatibiliteit, analyse en onderdrukking van stoorproblemen, Kluwer Technische Boeken B. V., Deventer – Antwerpen, 1990 (in Dutch).

H. Johnson, M. Graham, High-Speed Digital Design: A Handbook of Black Magic, Prentice Hall, Upper Saddle River, New Jersey, 1993, ISBN 0-13-395724-1.

A. S. Tanenbaum, Computer Networks, Prentice Hall, Upper Saddle River, New Jersey, 2003.

L. Tihanyi, Electromagnetic Compatibility in Power Electronics, IEEE Press, J. K. Eckert & Company Inc., Butterworth-Heinemann Ltd., Oxford, 1995.

References: websites

Clock signals and jitter: https://ilsiamerica.com/technical-information/mems-oscillator-terms-and-application-notes/clock-jitter-definitions-and-measurement-methods/

The CAN bus: https://en.wikipedia.org/wiki/CAN_bus

Chapter 10:

Microsimulation of road traffic cyber-physical systems

Vladimir Sistuk
Yurii Monastyrskyi

Kryvyi Rih National University

10.1. Transportation cyber-physical systems

10.1.1. Classification and definitions of road transport cyber-physical systems

There are multiple definitions of cyber-physical system (CPS). In conventional interpretation, CPS can broadly be defined as a combination of physical component and embedded software to monitor the progress of a specific process [5]. CPS understands the various types of large-scale infrastructure that have a large amount of physical information and components and have a broad impact on our society as they are deployed. It reinforces the requirements to understand the process of integrating physical components and information based on the use of complex computations, communications and management [7]. The task of improving the efficiency and safety of road transport has prompted research on the phenomenon of CPS, which has, as noted above, the physical, informational and social component.

At the same time, as far as cyber-physical traffic systems are concerned, the human is the main participant in every process. The behaviour of a driver or pedestrian ultimately determines changes in urban mobility and thus involves infrastructural changes and traffic management. A transport system is not a simple combination of physical and information components, its control cannot be boiled down to control of such physical processes as fluid or gas flow [7]. The driving style and mobility of the population are also extremely important for an adequate description of the interaction of information technology and behavioural component. It is the reproduction of the nature of this interaction that becomes the cornerstone for modelling CPS in transport (TCPS), which will be discussed in detail in the following paragraphs of this chapter. At the same time, the timeliness and quality of information have a significant impact on the behaviour of road users, especially when it comes to information flow. For the cyber-physical traffic system, it is important not only to implement traffic control as a preventive, informative and regulatory function, but also to control road users' behaviour.

The CPS in transport is a part of the vehicle control, the travel route, the organization process. The CPS consists of a human-machine interface that enables to keep feedback using the sensor equipment and influence the physical process with the actuators, evaluate and store the data. These systems are widely utilized in transportation processes. Compared to conventional transport systems, the TCPS aims to achieve greater efficiency and reliability by enhancing the role of feedback in the systems of interaction between the virtual and real world.

There is a classification of the TCPS, which takes into consideration the division into three types of systems as follows: infrastructure-based TCPS, vehicle-infrastructure coordinated TCPS and vehicle-based TCPS [7].

For infrastructure-based TCPS, which aims to control road traffic in real-time, the physical components are traffic lights, road cameras, electronic control units, traffic control centre, and the computation component is specialized software. The data flow is a function of prediction, monitoring, warning.

In the case of the coordinated TCPS for vehicle infrastructure, the physical component consists of vehicles with navigational equipment, control units and road signs using wireless communication technology. Such systems aim to manage green time, to warn about dangerous sections of the road, and the occurrence of congestion. An important requirement for the information component is the speed of data exchange.

Vehicle-based TCPS consists of sensors and embedded devices, car system actuators, and embedded software in electronic control units. Their work may include identifying blind spots or approaching the vehicle.

Another approach to the classification of TCPS is cyber-physical traffic control system allocation in addition to infrastructure-based TCPS and vehicle-infrastructure coordinated TCPS.

The information process for the cyber-physical traffic control system is a model for describing traffic control and calculating the optimal algorithm for controlling road behaviour. The main task of such a system is to increase the efficiency of traffic management and road safety.

The development of a CPS in transport includes the use of computer modelling, optimization, real-time control, computer networks and the issue of cybersecurity.

Effective traffic control involves the analysis of traffic organization and design of the information system. Modelling of the TCPS provides a mechanism for integrating information technology into the transportation process, determining passenger flows, traffic intensity and supporting modern communication, computational (intelligent analysis) and cybernetic technologies. In this case, the road user's behaviour can be determined both by calculation and by real-time tracking sensors. Data on the behaviour of road users can be used to configure traffic control devices or implement traffic control modes.

Spatial planning (geographic information systems) is used to transfer the traffic management system to cyberspace, for example, to represent a traffic model. In essence, transport modelling of TCPS is the use of specialized geographic information systems. This chapter is devoted to such systems.

10.1.2. Real-time traffic control systems

A real-time control system is a control system that receives the data, processes them and returns the results adequately [5]. It is important for modelling that the simulation time coincides with real-time. In control systems, real-time is determined by the amount of delay, which must be minimized. Technically, this condition is realized by the requirement to the controller (control unit), the response speed of which must be sufficient to generate a control signal, given the limited response time of the system, measured in milliseconds or microseconds.

Intelligent traffic lights serve as a real-time CPS traffic control system. The signal monitoring system uses a set of sensors to detect traffic density and adjusts traffic lights via artificial intelligence to smart traffic lights control. Employing advanced sensors and control technology, signals from intelligent traffic lights communicate with other intelligent traffic lights and vehicles to handle traffic in a more efficient way [5]. Vehicles are tracked by surveillance cameras or transport sensors. Changes to the traffic signal timing are carried out in real time.

One of the advanced software products that are part of the TCPS is the visual programming language FLOW from DataFromSky. The solution is designed to convert video information to real-time motion data. The developers offer three product options: Traffic Enterprise, Traffic Embedded and Traffic Camera.

Video recordings obtained from unmanned aerial vehicles and city cameras are subject to processing. The FLOW computer vision technology is designed for use in the design of a smart city, in the processes of traffic monitoring, traffic management and parking, and traffic safety. As noted by the developers of the technology [1], it can be argued that the surveillance camera turns into a smart sensor.

The platform is universal for a variety of devices, from a camera with an artificial intelligence processor to a personal computer.

According to the developers, to work with the processed video, it is enough to remove the camera and set the task according to the type of data required. The technology automatically recognizes objects, classifies them into 8 categories and analyzes traffic in detail: driving time and speed. The program modules allow monitoring (spatial, temporal and attributive filters) and control (the connection between traffic data with traffic light objects); assess safety (forecasting of road traffic accidents); and understand (different visualization options). Trajectory data is stored for at least 14 days.

Figure 10.1 shows the Traffic Embedded specifications for embedded devices.

10.2. Road traffic microscopic simulation

10.2.1. Transport modelling for traffic control systems cybernization

In open sources [7] cybernization of the physical component is considered as an approach to the design and modelling of CPS based on the deployment of software for this component. The complexity of modelling TCPS is that the model of the traffic control system should include comprehensive information about the traffic behaviour.

The approach to the type of traffic information involves its distribution into syntactic, semantic and pragmatic (Table 10.1). The first type of information is digital data in electronic control units, the second one is designed for travellers and related to traffic management, the third one characterizes the behaviour of drivers and pedestrians. Table 10.1 shows a detailed description of information flows in the CPS traffic system.

Thus, the concept of TCPS provides a constant impact on the behaviour of road users. Modernization of the traffic control system is based on its computation. At the stage of TCPS design, the computation acts as a method of its research; at the stage of operation of TCPS, it is a traffic control data processing. When modelling

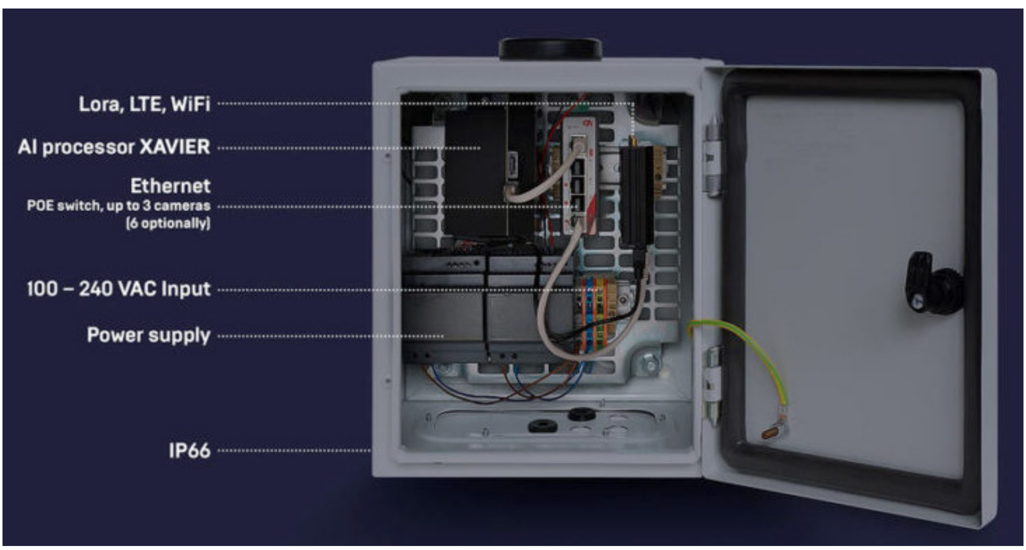


Fig. 10.1. The specifications of the Traffic Embedded.

Table 10.1. Information Flows in TCPS.

Type of information	Carrier of information	Information form	Description
Syntactic (digital data)	Embedded electronic units, cameras, transport detectors	Video footages, maps, traffic data	The record of traffic data system acquisition system, traffic operation video, traffic accidents scene
Semantic (traffic management data)	Means of traffic management	Traffic signs, sound, light, voice	Road name, warning information, radio broadcasting, traffic lights, vehicle turn signals
Pragmatic (road users' behaviour data)	Road user (traveller)	Obtained traffic information that influences behaviour	Traffic management data, traffic information that is correctly recognized, information that has an impact on the behaviour of road users, messages from road users

the traffic control systems, researchers are faced with the transport system computation problem as well as with road users' behaviour representation problem. Among other things, the role of the human factor is decisive.

10.2.2. Classification of methods of traffic simulation

With the increasing complexity of transport networks, microscopic simulation became one of the main assessment tools for optimizing various traffic management systems. The results of traffic simulation serve as a basis for the development of TCPS. Road network simulation enables analysis and subsequent approval of the most efficient (in terms of cost, traffic safety, capacity and other factors) engineering solution.

Traffic flow simulation allows solving of the following issues:

- To forecast passenger and automobile flows on street, road networks of the country, region or a certain city.

- To analyze in detail changes of passenger/automobile flows at the realization of various decisions associated with a change of town-planning or transport infrastructure.

- To form optimal modes of traffic signalization at various objects of the road network.

- To develop an order of construction of objects of town-planning and transport infrastructure.

- To optimize the work of urban transport [9].

- To summarize the information about the capabilities of able simulation models, able simulation models are capable to do the following:

    - to assess the level of possible hazard at the designed object making decisions on traffic management and control;

    - to calculate traffic intensity for different categories of road users, the level of traffic congestion on the road network elements, the traffic conditions parameters;

    - to determine parameters of the level of service of specific areas or objects (first of all, the time of vehicle departure to the area or to the object from the specified points);

    - to determine parameters for the calculation of economic expenditures for vehicles and pedestrian traffic, as well as for environmental impact assessment of automobiles.

The transport flow description determines the type of transport model. Thus, the traffic flow can be represented by macro parameters (speed, density, intensity) as well as at the micro-level considering patterns of movement of individual vehicles or little groups of them.

Also, in light of the traffic flow data acquiring method, the models are divided as follows:

- Analytical models, which are based on theoretical and empirical relationships between the parameters of the transport flow and transport infrastructure (such as software VISUM, TRANSCAD, CUBE).

- Simulation models that seek to describe and reproduce in time the processes of traffic flows and their interaction with the transport infrastructure (software VISSIM, AIMSUN, PARAMICS, CORSIM).

Thus, from the considered positions the traffic flow models can be divided into analytical macromodels, analytical micromodels, simulation macromodels, and simulation micromodels.

Simulation micromodels reproduce the dynamics of traffic flows given each vehicle's behaviour and traffic conditions.

The level of detail and the accuracy class of the model should correspond to the purpose for which the model is intended: the transport model, depending on the scale of the tasks, can cover the area from the whole region to a single intersection at the lower level. Depending on the required degree of detail of the description of flows and the accuracy of the received parameters, the corresponding class of model is chosen.

Purpose, of the practical side and solution of problems, appears in models only as a part of specialized software packages. Thus, it is possible to classify by purpose only ready-made software for modelling traffic flows. Viewed in this way, software products are divided into those used for [9]:

- preliminary (sketch) planning, including pre-network methods;
- strategic planning (forecasting);
- tactical planning and management (macro- and mesomodels);
- planning of certain types of transport (freight, taxi, etc.);
- detailed analysis of flow motion (micro- and mesomodeling);
- optimization of parameters of coordinated traffic light regulation, including in automated traffic control systems;
- calculation of geometric parameters and control parameters at individual intersections;
- analysis of pedestrian traffic;
- assessments of the level of traffic safety.

It ought to be noted that recently there has been a trend to combine tools to solve multiple tasks in a single software package or an interconnected family of products from one developer. There are also many software modules for connecting different products for various purposes to simplify the process of entering and exchanging data. Thus, it is difficult to assign many software products to one of these categories by purpose.

In the simulation, the dynamic processes of the system and the original are replaced by processes simulated by the model algorithm with the same ratios of durations, logical and time sequences as in the real system. Both in the process of simulation, the operation of the system under study and during the experiment with the original, certain events and states are recorded, which then calculate the necessary characteristics of the quality of the system (delays, average speed, number of stops, congestion level). Each experiment with the model is based on a random number that determines the further development of events in the model. Statistically significant results can be obtained only by averaging the results over several replications (runs) of the model. The number of runs is determined following the experimental design theory.

An important feature of this method is that simulation models allow testing of an unbuilt object, simulate various possible scenarios of its operation and conduct a series of experiments related to various emergency situations, while checking the stability of the object in similar situations.

The structure of the simulation micromodel can be represented through the interaction of the elements of the DCRE system (driver-car-road-environment).

The driver. This element of the system contains the above models of behavior when driving in a flow (car-following, changing lanes, choosing a gap in the flow) and also takes into account the driver’s reaction time, desired speed, degree of speed limit and other behavioral characteristics.

Car. The car element is described mainly by external dimensions (length and width), traction and speed performance, belonging to any class, passenger capacity.

Road. This element of the system describes all the parameters of infrastructure, organization and traffic management. These include geometric parameters of the DCRE (number and width of lanes, turning radii, slopes), traffic management parameters (lane directions, maneuvers inhibit, speed limits, dedicated lanes), control parameters (traffic light control modes, parameters of automated traffic control systems).

Environment. The concept of the environment in the simulation includes the impact of weather conditions on traffic flows in terms of visibility restriction and reducing the adhesion of the pavement. These factors lead to a decrease in speed and increase the distance between cars.

Thus, microsimulation software describes the traffic flow's movement at the level of interaction of the elements of the DCRE system.

To sum up, let us determine the factors under which the use of microsimulation will be justified at the stage of variant design and analysis of the effectiveness of decision-making on traffic management. These include:

- availability of transport intersections with a complex configuration;
- network congestion level is 0.8 and more;
- sensitivity of the system to the redirection of traffic flows in time and space;
- presence of traffic signalization with calling phases, prioritized pass or green time;
- areas where consideration must be given to the interaction between individual vehicles (zones of crossing and merging flows);
- close to the intersection and the area where the length of the queue often exceeds the length of the distance between stops;
- the need to analyze the strategies of automated traffic control systems;
- the need to analyze the interaction of vehicles and pedestrian or bicycle flows;
- the need to analyze the movement of categories of vehicles with different behaviour;
- the need to analyze decisions on the dynamic management of the scheme of traffic organization (no-parking zone, reverse traffic);
- the need to analyze the conditions of priority access to public transport;
- the need for visualization of the movement of the flows.

10.2.2. The problem of initial data collection

All the initial data needed to develop a simulation model can be integrated into four blocks [9]:

- Transport information – data to build and calibrate the model; data for model validation; data for the future.
- Infrastructure information – geometric parameters; characteristics of traffic management; control parameters, including indication of subsystems of intelligent transport systems; parameters for the future.
- Explicit data for certain software.

Let us detail the data for each group.

Transport information for model development includes the following parameters:

- The values of intensity of traffic flows in the space intervals and in the directions of traffic at the intersections of the considered section of the transport network that are obtained by measurements and usually broken down into intervals of 5–15 minutes.
- Flow composition by types of vehicles.
- The values of pedestrian flow intensity.
- The value of the intensity of bicycle flows.
- *Calibration parameters*
    - Values of actual saturation flow at signalized intersections entrances.
    - The average values and variance of traffic speed on control sites.
    - Travel time between control points.
    - Queue lengths in front of stop lines.

Transport information and calibration parameters are collected for each analysis period. The number of analysis periods is determined based on the project objectives.

Capacity and saturation flows are key parameters to calibration, as they determine when a traffic delay situation occurs.

A more accurate estimate of delays at intersections can be made, for example, by measuring the number of stopped cars and the duration of their downtime. Methods of full-scale determination of delays are also described in the literature [4].

Data for model validation also consist of traffic intensities, traffic speeds, queue lengths; they must be collected regardless of the initial project data in another period. To validate the basic model, it is advisable to use data corresponding to maximum traffic intensity [6].

Data for the future include forecast values of traffic intensity. Such data can be obtained from various sources (macromodels, growth factor method).

Infrastructure data. Geometric parameters can be obtained from the available design documentation, satellite or aerial photography, as well as by direct measurements.

Geometric parameters include the number of lanes (considering their actual use); lane width; the length of the section of road; the length of the extension sections; radii of curvative (considering their influence on speed); longitudinal slopes; parameters of sidewalks and bicycle paths; cross-section configurations (including pedestrian crossings and bicycle path crossings).

Characteristics of traffic management consist of prohibitions of certain manoeuvres, distribution of traffic directions in lanes, availability of dedicated lanes (for certain modes of transport), local speed limits; the presence of road inequalities (including a sleeping policeman), presence or prohibition of parking along the carriageway, restriction of manoeuvres associated with lane change and the presence of traffic bans on certain types of vehicles.

Control parameters contain peripheral equipment locations (traffic lights, transport detectors); traffic light control modes (number and succession of phases, timing, main and intermediate cycles); adaptive control parameters; location and type of transport detectors; parameters of coordinated management; algorithms of variable information signs control systems operation; algorithms of information and navigation systems operation; other algorithms of operation of main and network automated control systems.

Infrastructure parameters over the longer-term consist of data on expected changes in the structure of the transport network, the scheme of traffic management, the configuration of intersections and control systems.

In addition to the above-mentioned parameters, information for vehicles behaviour description can be used, as well as data which are specific to each software.

Such data include the length of cars (in the form of distribution or average value and variances); desired speed distribution; average and maximum values of acceleration and deceleration; exhaust gas emissions by types of vehicles and traffic modes; specific parameters of driver behaviour models adjustment (car-following, lane change).

Thus, the given classification of initial data allows to qualitatively estimate possibilities of traffic simulation software.

To prepare a complete set of correct traffic data needed to develop the model, the source data is collected in the field.

It has been known that before collecting the initial data, it is necessary to determine the scope of a particular project [9]. The scope of the project area depends on the 'zone of influence' of the surrounding transport network. The zone of influence is the study area, and the surrounding transport network that affects operations in the study area (including consideration of future network development criteria) may vary from a single intersection to a section of the network that covers a single urban area. This approach allows to implement the real behaviour of vehicle and pedestrian traffic in the project area. The zone of influence may be greater than the minimum boundaries of the study area. It is important to understand the performance of the transport infrastructure within the proposed project.

The quality and accuracy of the original data play a crucial role in the reliability of the simulation results. Consider the types of input data for this stage as follows:

- required types of source data,
- selected measurement technologies,
- location of measurement sites,
- time interval of measurements,
- other data sources.

The output must be a complete set of necessary source data, collected in compliance with all requirements for the established accuracy.

Analysis of the methodical literature on transport modelling shows [6], [9] that the period of survey must meet four criteria:

- Interval of observations should be at least 10 minutes.
- The number of vehicles in the most congested areas of the network and the rate of growth of traffic may slow down, but the number of modelled vehicles should continue to increase for the same period.
- The estimated travel time of vehicles from one end of the network to the other must be equal to or twice as long as the simulated time.
- The simulated traffic queue length should repeat the actual observations at this time of day.

Also, the survey time for microsimulation study must cover the period from the occurrence of congestion to their dispersion.

It is a good practice to use traffic data and field observations to determine the appropriate peak period of traffic load in the study area. Most often, the hourly traffic intensity is adjusted upwards (it is affected by peak hours) to reproduce the 15-minute period during which it reaches its maximum value.

It is appropriate to make sure that the traffic conditions on the days of the survey are the same and there is no influence of such factors as weather conditions, repairs, road accidents, mass events.

It must be also said that the study should be undertaken in the most representative periods, avoiding holidays, school holidays, vacation periods, and other atypical periods to minimize the variance in the obtained values.

10.2.3. The initial data input and base model development

Common use of simulation software usually includes the following aspects [9]:

- Study of transport corridors on highways to determine the efficiency of the transport system.
- Advanced studies of motorways, including management issues such as reversible traffic, speed control, traffic control when leaving the main route,

route planning.

- Development and analysis of highway management strategies, including high-speed sections at different levels, and calculation of harmful effects on the environment at the stages of road construction.

- Study of transport corridors on highways with unregulated and traffic light-regulated intersections.

- Analysis of alternative strategies of updated and adaptive management for local networks.

- Detailed testing of control logic and analysis of traffic management effectiveness.

- Choice of priority schemes of traffic lights for public transport within the framework of multimodal research.

- Guiding traffic of public transport, such as light rail transport, trams and buses, including analysis of the operation and capacity of bus and tram terminals.

- Detailed analysis of vehicles speed during maneuvers with the restricted view.

- Assessment of alternative options for traffic management on highways and in urban environment.

It is of importance to note that the mathematical model is required in any traffic simulator to represent the transportation supply system while simulating the technological and organizational aspects of physical transportation supply. Second, a demand model must be developed to simulate the demand of people and vehicles using the supply mechanism. Unlike macroscopic transportation models, traffic control must be meticulously modelled depending on supply and demand. As a result, the simulator has three main building blocks, plus one more that produces the effects of each simulation run.

The first block consists of road and railway networks, as well as signposts and parking facilities. To model the physical roads and tracks, this block is needed. As the starting (origin) and ending (destination) points of trips, public transportation stops and parking lots are needed. They are included in the first block because they are physical and stationary network elements.

The second block specifies the technical characteristics of a vehicle as well as traffic flow specifications. Origin-destination matrices or traffic generation at link entries are used to characterize traffic. This block contains the assignment model and route flow descriptions. Within this block, public transportation lines are described as a series of links and stops.

Many of the elements needed to control traffic are included in the traffic control block. This block contains definitions for four-way stops, priority rules with gap approval, and traffic signal control options. Since a signal post with signal heads is part of the infrastructure block, signalization, which includes signal settings definitions and actuated control, is part of the traffic control block.

During a traffic-flow simulation, vehicles (block 2) can trigger detectors (block 1), which will influence vehicle-actuated signal control (block 3); therefore, all three blocks are constantly active during the simulation, with interdependencies between them [6].

The fourth block, which is responsible for all types of data output, is a little different.

Without a feedback loop, the evaluation block processes the data generated by the first three blocks. During the simulation, output can be produced in the form of animated vehicles and traffic control states, or statistical data on detector calls and vehicle states presented in dialogue boxes. The majority of measures of effectiveness (MOEs) are generated during the simulation, stored, and filed at the conclusion of each simulation.

Let us consider the elements of each block in more detail.

Graphs with nodes at intersections and connections on-road segments are widely used to reflect road networks. If two or more links merge, links cross each other, one link splits into two or more links, or a road segment's characteristics change, nodes are required.

Microsimulation software does not require an explicit description of nodes for the sake of added versatility. Connectors, on the other hand, model the functionality of joining, crossing, and separating two connections.

Infrastructure objects are added to segments and connections in the required lane. Since infrastructure objects do not have a physical length, they are fixed at a particular coordinate. The main infrastructure facilities include speed limit signs, yield signs, traffic lights. Additionally, for this group you can specify transport detectors, public transport stops, parking lots, speed limits.

The attributes provided by the vehicle in the modelling of private transport are divided into those that are mandatory and additional. The mandatory attributes are as follows: type of the vehicle; length of the vehicle or vehicle length distribution; as a function of speed, distributions of technical and desired acceleration and deceleration rates; maximum speed or distribution of maximum speeds; width of the vehicle. The additional attributes are as follows: colour and 3D model, or colour and 3D model distribution; vehicle weight or vehicle weight distribution; vehicle emission class; vehicle owning costs.

From the software manuals (first of all, PTV VISSIM) we will define [6] that vehicles are generated stochastically at link entries or stops. Data input flows are described separately for different periods. The time difference x between two successive vehicles will obey the exponential distribution with mean $1/\lambda$, as the number of departures in a given time interval [0,t] follows the Poisson distribution (lack of aftereffect) with mean λt, where λ is vehicles per hour (VPH). The probability of time gap x between two sequentially generated vehicles can be calculated using formula

$$f(x,\lambda) = \lambda e^{-\lambda x}.$$

All of the technological features of private transport apply to public transport vehicles, with a few exceptions. A public transportation line is made up of buses, trams, or light rail vehicles that follow a schedule and serve a set of public transportation stops. The dwell time is determined by the time of the direct stop and the duration of passenger service at the stop. It is calculated as follows:

Simulated arrival at next stop + dwell time + max (0 (start time + departure time offset − simulated arrival + dwell time); slack time fraction.

The timetable for two consecutive stops determines the departure time offset. If set below 1, the slack time fraction accounts for early starts. The public transport can wait until the scheduled departure time is reached if the scheduled departure time is later than the amount of the arrival and dwell times. When the slack time fraction or departure time offset is zero, the vehicle can only depart based on traffic conditions and dwell time.

For both uncontrolled and signalized intersections, traffic control is conducted. The priority rules of the uncontrolled intersection apply in the first instance. In the simulation, the following scenarios are considered:

- Uncontrolled intersections where traffic must yield to right-hand traffic.
- Uncontrolled intersections where traffic on the terminating road must

yield to traffic on the continuing road.

- Intersections of two-way stoplights and all-way stoplights.
- Cloverleaf intersection, where vehicles approaching the cloverleaf intersection must yield to traffic already present.
- Merging zones, where traffic entering from a ramp must yield to traffic on the main road.
- Semi-compatible movements (permitted turns) at signalized intersections include right and left turns that conflict with parallel pedestrian movements or left turns that are parallel to opposing through movements.
- The movement of buses with their designation, which leaves the stop in the forward direction.

Traffic lights are a component of the infrastructure unit, while traffic light indicators are a component of the traffic control unit. Several traffic lights with the same timing are part of the same signal group, the smallest control element in the traffic control unit.

There are various options for estimating the delay, the length of the queue, the number of stops at the intersection to adjust the time parameters of traffic lights. Microsimulation is used, inter alia, as a tool to control the feedback between the traffic light timing and to account for the stochastic impact of vehicle arrivals at intersections to activate and adaptively control traffic. Traffic simulation software contains a programming language with a graphical interface for determining the program of traffic light control.

After the trial simulation has shown the compliance of the developed model with the real conditions, the simulation model is considered ready to form a list of results in the form of a report. The collection of simulation results is based on the activation of sensors and counters. Thus, in micromodels used counter '*Vehicle travel time*' is set by the segment on which the duration of the vehicle will be recorded. '*Queue counter*', designed to record the length of the jam and the time of its origin, is installed in places where it is most likely to occur.

Figure 10.2 shows the common algorithm of simulation model development.

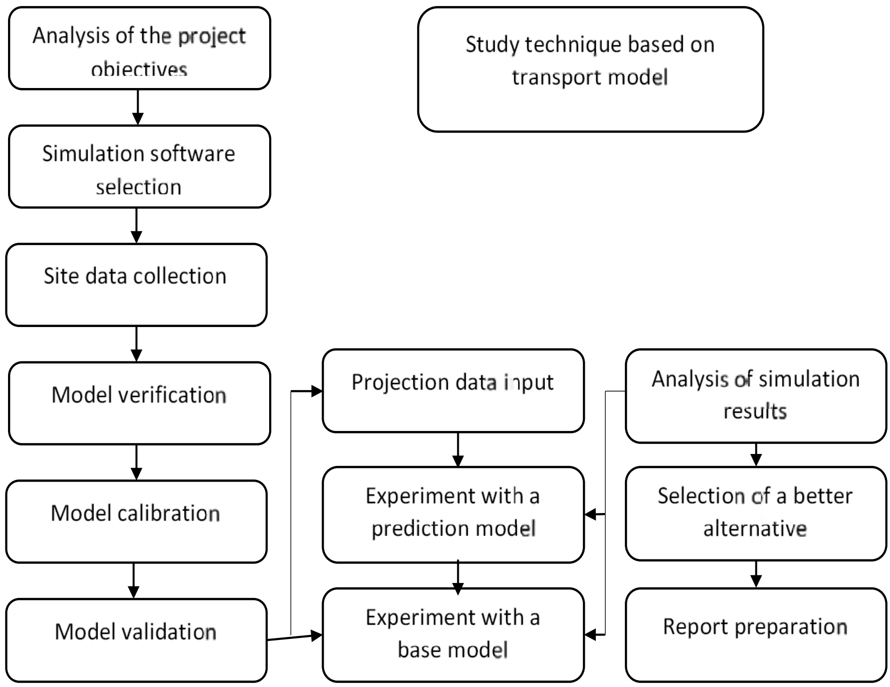


Fig. 10.2. Road traffic microsimulation technique.

10.2.4. Driver behaviour model

Driver behavior patterns, which determine how drivers make decisions in terms of lane choice, car-following, and route choosing, are an important component of any model of microscopic simualtion.

Traffic simulators are built around a model of driver behavior on the road, which includes things like following the leader in the column and changing lanes. These models' mathematical apparatus was developed in the last century. Thus, for the PTV VISSIM simulator, the Wiedemann 74 (urban traffic) and Wiedemann 99 (highway) models were developed in 1974 and 1999, respectively. It [6] assumes that the driver can be in one of four states.

- *Free movement.* The driver tries to reach and maintain his desired speed, but in reality, the speed cannot be maintained at a constant rate.

- *Approach.* The process of adjusting the driver's speed to match the lower speed of the vehicle ahead. As he approaches, the driver brakes so that the difference in speed between the two cars is zero by the time he gets to a safe distance from the car in front.

- *Following*. The driver follows the car in front without accelerating or decelerating, maintaining a constant safe distance. The speed difference oscillates around zero.

- *Braking*. The use of medium or strong braking if the distance between vehicles becomes less safe.

The last-mentioned state can happen if the vehicle in front suddenly changes speed, or if a third vehicle changes direction in front of the driver. Acceleration is described for each mode as a result of the speed, the difference in speed between two consecutive vehicles, and individual characteristics of the driver and the vehicle.

When the driver reaches a certain barrier, which can be described as a combination of the difference in speed and distance between two vehicles, he changes his behaviour from one state to another. A small difference in speed, for example, can only be tolerated over short distances, whereas large differences in speed force oncoming drivers to react much sooner. The ability to estimate the difference in speed and distance among drivers varies, as does the desired speed and safe distance. Due to the combination of psychological aspects of physiological limitations, this model is called the psychophysiological model of following.

Following consists of lane choice, lane change, and one-line driving. When approaching an intersection, the driver selects the lanes with the best road conditions. The road situation is determined in three stages. At first, a decision is made about whether or not to change lanes, which always happens when the state of interaction between other vehicles (the mode of movement in the model) differs from free traffic. Then, examine the adjacent lanes to determine which has the best road situation, either free traffic or more time before the collision. If one of the adjacent lanes provides a better situation along the way, the final check is whether it is possible to change the lane while keeping the vehicle on the opposite side in mind, which is modelled as accepting the gap between the vehicles.

However, lane choice is often determined by mandatory lane changes for desired turns at intersections. Each connecting segment in the development of the transportation network has two distances assigned to it: the distance at which the lane changes and the distance of the emergency stop. The first distance, which typically ranges from 100 m to 500 m, characterizes the moment of recognizing the presence of a connecting segment. The distance of the emergency stop, in turn, is the area where the driver will stop if reaching the lane to go to the connecting segment is impossible. However, the connection is only for certain lanes; for example, a right turn on a three-lane highway will be connected only with the extreme right lane, forcing the vehicle to be exclusively in this lane if a right turn is required.

The algorithm for changing the vehicle's lane is based on decision-making in the case of an unforced or mandatory change of lanes, taking the gap between the vehicles into account. The vehicle's driver is prepared to admit that he is forcing the lagging car to slow down in the right lane. The value of this braking is determined by the program's calibration settings, and for mandatory lane changes, it is determined by the distance to the emergency stop position of the next connecting section, which completes the lane change, i.e., the driver becomes very close to the emergency stop. Permissible deceleration values for vehicles lagging in traffic and for the vehicle itself are driver behaviour model parameters that can be determined selectively for pairs of segments and vehicle types.

The driver chooses the lane's lateral position based on a simple principle: the lateral position should correspond the most time before the collision. To determine this position, the driver divides the available road width into virtual lanes that are located on the right and left sides of the vehicle in front, as well as some lateral safe distance. The safe distance that the driver wishes to maintain when overtaking another vehicle is determined by the type of this vehicle as well as the speed of the vehicle overtaking. Because the simulation programs provide a linear dependence of the distance on the vehicle's speed, the user can determine the minimum safe distance at very low speeds.

10.2.5. Microsimulation model calibration

Previously, the analysis of the results of microscopic simulations of traffic flows was based on the default parameter values, which led to incorrect conclusions and significant discrepances with the data collected in kind. To address this issue, researchers perform a calibration procedure for the simulation transport model [6], which entails adjusting the program's parameters to achieve reliability and accuracy of the results through a conditional relationship between the simulated interchange parameters and traffic data collected in the field. There have been numerous attempts by researchers and experts to develop procedures for successful calibration and verification of simulation model parameters. Calibration of the simulation model is an iterative process that involves adjusting the model's input parameters to obtain analysis results that can reasonably represent the current observed traffic conditions. There are two main methods of calibration: option search (manual calibration) and system approach (automated calibration).

The general procedure for calibrating a simulation model is typically as follows [9]:

1) identifying the model parameters that need to be adjusted or calibrated;

2) choosing the required measures of effectiveness (MOEs) and data for validation;

3) counting the number of simulations required to achieve a 95 percent confidence interval with a 5 percent absolute error for all MOEs;

4) establishment of validation criteria and objectives;

5) modification of the selected parameters as long as the obtained simulation results closely resemble real-world road conditions or meet the validation objectives.

To achieve an acceptable level of accuracy, the calibration parameters must be adjusted consistently. Numerous input parameters describe the geometry of the network, traffic demand, general configuration, traffic organization, traffic flow and vehicle characteristics, driver behaviour, and route selection strategies in the simulation model. Typically, these input parameters have a strong inverse relationship with the mechanism of influence of other parameters. As a result, adjusting several parameters to solve one problem can easily lead to other problems elsewhere in the simulation. The primary calibration efforts are concentrated on the driver's behaviour and vehicle parameters. Driver behaviour parameters have a direct impact on vehicle interaction, traffic regulation, changing the saturation flow rate on highways, and providing different driving styles for traffic flow, such as aggressive and passive actions.

The simulated values for the selected performance indicators are compared to the observed values for the same indicators during the verification stage. The verification process determines how closely the model matches the actual traffic conditions. In this account, the following MOEs are typically used to test the VISSIM model [2]:

- Transport load using diagrams of transport load for 'peak time'.

- Travel time using the 'floating car' method, travel time during 'peak times' on several routes.

- Queue length using the maximum queue length for intersection approaches, which is the number of vehicles in the queue at the start of the green traffic light. It is determined in two stages: the current queue length is measured each time step (1 minute), and the maximum value for each time interval is calculated from these values (15 minutes).

Because measuring microscopic parameters directly is difficult, most calibration methods use locally determined macroscopic flow parameters as measures of effectiveness (MOEs) to calibrate the microscopic parameters of driver behavior.

The average travel time, for example, is an aggregate indicator that determines the state of traffic flow. Calibration methods assume that the set of microscopic parameters is optimal, resulting in the least amount of estimation error between modelling and field observations for certain macroscopic parameters.

To calibrate the VISSIM model, GEH statistics (Geoffrey E. Havers) [9] are commonly used, which are a modified $\chi 2$ statistic that takes into account both absolute and relative errors, defined as:

$$GEH = \sqrt{\frac{2\cdot(M_i - O_i)^2}{M_i - O_i}}, \tag{10.1}$$

where Mi and Oi, respectively, are simulated and observed transport load (veh/h) on the i-th lane if the GEH-value is less than 5; this indicates that the simulated and observed indicators are sufficiently correlated.

The validation criteria are summarized further.

1. The simulated lane load must differ from the observed load by 15 % for flows ranging from 700 to 2700 veh/h, from 100 veh/h for flows less than 700 veh/h, or from 400 veh/hour for flows greater than 2700 veh/hour. These targets must be met in 85 % of cases.

2. The sum of the simulated flows per lane is within 5 % of the total of all lanes' actual flows.

3. For the traffic flow on a single lane, GEH statistics should be less than 5 in 85 % of cases.

4. GEH statistics with a number less than 4 characterize the sum of all lanes in the lanes.

5. The simulated travel time must differ by 15 % (or 1 minute) from the observed travel time for at least 85 % of the routes.

6. The simulated queue length must differ from the observed queue by 20 % (or 12 vehicles).

However, numerous attempts to use GEH statistics have yielded inconclusive results. Researchers have developed a low calibration procedure for simulation models in this regard.

Most of the studies conducted to date can be identified as belonging to the issues of highway model calibration, intersections, and intersections of various levels (Fig. 10.3).

On the other hand, great emphasis is placed on optimization methods and algorithms for selecting the necessary calibration parameters, as well as on the calibration procedure itself. Let us take a quick look at the differences

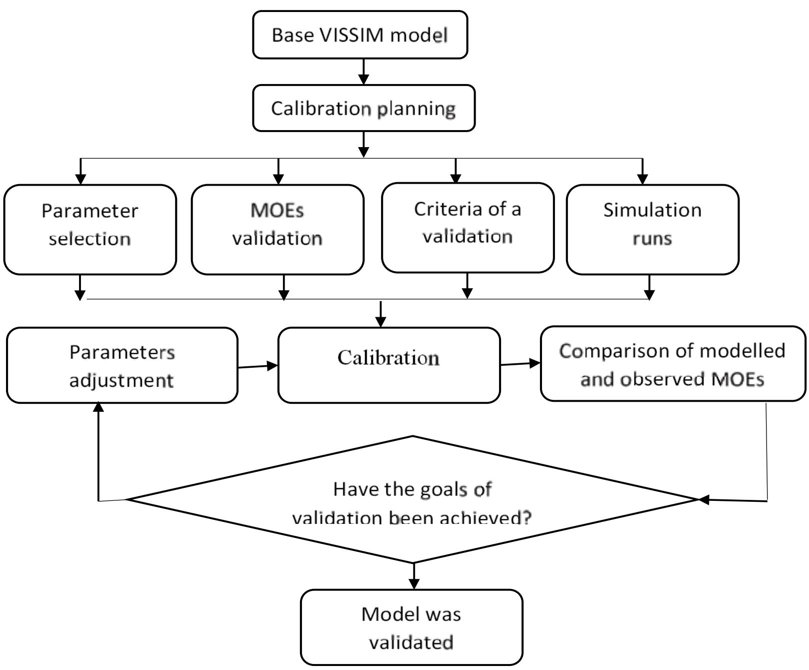


Fig. 10.3. Calibration of the simulation model in general.

and similarities of the methods that use the optimization procedure and the adaptability function and contain roughly the same elements.

10.2.6. Calibration of a microsimulation model based on computer vision technology

The reproduction of traffic conditions and flows based on the configuration of the road intersection is known as simulation. The VISSIM program's simulation results include travel time on each route and the length of the formed queue of vehicles. These indicators, however, cannot serve as a decision-making guide until the driver's interchange pattern and behaviour are properly adjusted, reducing the uncertainty of the transport process. As previously stated, model calibration is critical because the driver behaviour varies significantly depending on location and traffic conditions (time of day, weather, etc.). Local characteristics and traffic conditions are rarely represented by the simulation software's default settings. As previously stated, numerous methods of calibration of the simulation model's parameters have been proposed as a result of the growing popularity of microscopic modelling and the importance of calibration. Due to the difficulty of directly measuring microscopic parameters, most of these methods use locally determined macroscopic flow parameters to calibrate the microscopic parameters of driver behaviour [6].

Total indicators, such as travel time, determine the flow's state. If the deviations in the values of the total simulated indicators and the field observation data are small, the chosen set of microscopic calibration parameters and the appropriate method of carrying out this procedure are adequate. This hypothesis enabled traffic researchers to narrow the scope of their search for optimal calibration methods. The set of microscopic parameters, however, does not always reproduce local conditions. This statement is especially relevant in urban traffic, where the number of influencing factors increases by an order of magnitude. Under certain conditions, one of the sources of reliable data can be the results of automated video traffic analysis, which will primarily reduce the number of iterations because data will be generated not stochastically but by analyzing a specific video frame with artificial intelligence, greatly simplifying the overall calibration procedure. It is critical to test this hypothesis because no such study has been conducted to date.

DataFromSky Viewer [1] is a software solution for automating the processing of traffic video fragments with the functions of viewing, creating, and editing tracking data based on artificial intelligence technology. The tracked vehicle is represented in this program as either an object with an immediate position at any time or a specific trajectory in time-space.

Each tracking object is given an identification number (ID), and the parameters of speed, acceleration/deceleration, and time spent at the intersection are calculated. For each tracking object, a speed and acceleration graph is generated. Vehicle trajectories, current workspace designation configurations, results of correspondence matrices analysis (OD-matrices – origin-destination), road safety, and time intervals between two consecutive vehicles (Headway) are all visualized.

The platform's primary functions are as follows [1]:

- Creation of new tracking log files via the video geo-registration procedure and manual marking of the vehicle trajectory (*Annotation Configuration*).
- Examining available trace protocol files ('*.tlg*') alongside analyzed video sequences.
- Modifying existing trace protocol files, such as adding new tracks, adjusting, and deleting contained tracks.
- Exporting vehicle tracking data to CSV files for further analysis of the vehicle trajectory.
- Calculating and export of data for parameters such as '*Time-to-Gate'* and '*Time-to- Follow'*.
- Creating video visualizations of marked or detected vehicle trajectories.

The elements that comprise the video's working space are as follows:

- *Lanes* – areas within which vehicles move.
- Counter (*Gate*) – a virtual line that captures intersecting vehicles, can be directional or selective depending on classification, and can be represented as an input, output, or other.
- *Reference points* – frame geo-links.
- *Analysis nodes* – virtual objects that connect other notations.
- *Traffic region* – areas which are used to determine the presence and speed of objects.
- *Action region* – area, which is used to determine the presence and speed

of vehicles

- *Anonymization Region* – area which is used to conceal a portion of the workspace

To create counters, lanes and hidden areas, areas of motion, and action in DataFromSky Viewer, you must configure the appropriate frame designation (*Manage Annotation Configuration*).

The main results of traffic analysis are as follows:

- Calculating transport loads on meters.
- Categorizing all tracking objects (including cyclists and pedestrians).
- Calculation of OD-matrices in certain directions – their colour designation.
- Measurement of speed and acceleration at a specific point of the workspace, displaying the colour designation and creating heat maps of the workspace.
- Analysis of traffic safety based on surrogate indicators (*Time-To-Collision*), post-encroachment time (*PET*), emergency braking (*Heavy Braking*).

Based on our own research, we conclude that the results of traffic video post-processing can also be used to calibrate the simulation model. The following parameters are available for DataFromSky intelligent video data analysis software:

- tracking log protocol;
- OD-matrix (origin-destination) by the number of counted recognition objects and the duration of individual routes;
- statistics of recognition object passage through virtual lines in the workspace (counters – Gates), as determined by the notation configuration;
- data on the composition of traffic flows;
- statistical report on movement dynamics, including minimum, average,

and maximum values of speeds, accelerations, and decelerations for each recognized object.

We present a study of the roundabout in the city of Kryvyi Rih as an example of the use of data obtained by the technology of intelligent analysis of traffic videos to calibrate the simulation model.

The roundabout is in the heart of Kryvyi Rih (Horkoho Square). The footage was obtained from a video surveillance camera in the office building at 37 Metalurhiv Avenue (Fig. 10.4). The frame rate of the original video is 25 fps, the number of frames is 154 818, the resolution of the video stream is 1 280 by 720 pixels, which corresponds to the technical conditions of the software developer DataFromSky Viewer. The duration of the video is 1 hour 45 minutes.

To create gates, lanes, and hidden zones, motion, and action areas in DataFromSky Viewer, you must configure the appropriate frame designation (*manage annotation configuration*). Figure 10.4 shows the workspace designations created in Datafromsky Viewer. Workspace designation configuration, which includes 17 lanes, 12 counters and an action area.

The symbol configuration serves as the foundation for calculating the OD-matrices of traffic flows at the intersection.

The origin-destination (OD) matrix depicts the fundamental parameters of a specific route from the Entry Gate to the Exit Gate (number of vehicles, average driving time, minimum and maximum time, standard deviation).

Fig. 10.4. Annotation configuration in DataFromSky Viewer.

Table 10.2. Designation of Entry and Exit Gates and Traffic Flow Directions.

The number and designation of gates	The type of gate	Traffic flow direction
Gate 1	Entry	Entrance from Myru Avenue to the roundabout
Gate 2	Exit	Departure from the roundabout to Metalurhiv Avenue
Gate 3	Exit	Departure from the roundabout to Myru Avenue
Gate 4	Entry	Entrance from Haharina Avenue to the roundabout
Gate 5	Exit	Departure from the roundabout to Haharina Avenue
Gate 6	Entry	Entrance from Metalurhiv Avenue to the roundabout
Gate 7	Exit	Departure from the roundabout to Volhohradska Street
Gate 8	Entry	Entrance from Volhohradska Street to the roundabout

The origin-destination matrices can be used to calculate the average hourly traffic in areas that require special attention.

Instead of the full video duration, a time interval of one hour was used to calculate the hourly traffic. The final result is the OD-matrix shown in Fig. 10.5.

The highest traffic volumes are for the routes from Entry Gate 8 to Exit Gate 2 (from Volhohradska Street to Metalurhiv Avenue), as well as to the Exit Gate 3

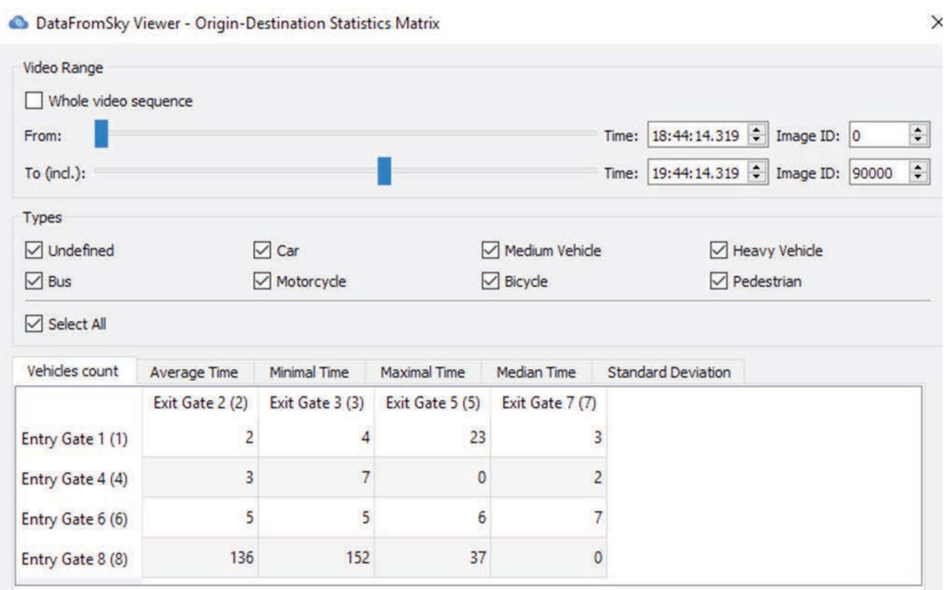


Fig. 10.5. OD-matrix by the number of vehicles on the route for 1 hour.

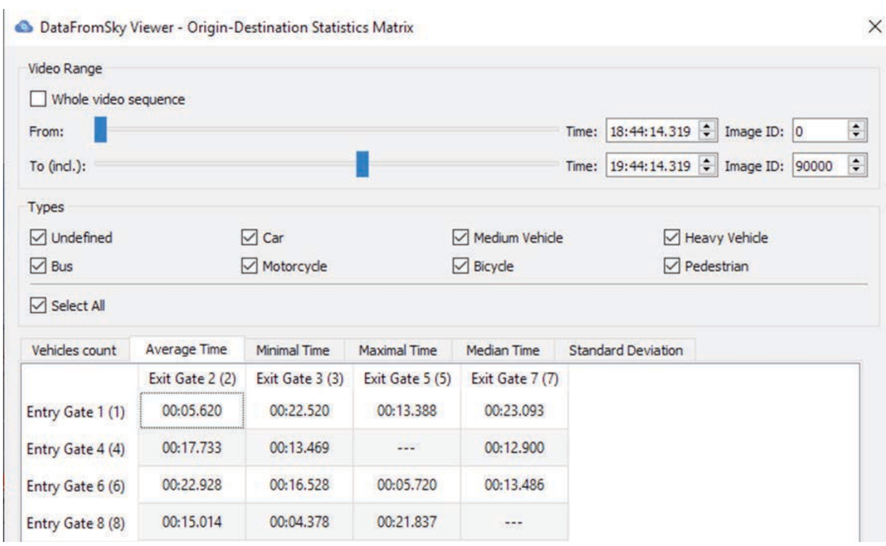
(from Volhohradska Street to Myru Avenue), and Exit Gate 5 (from Volhohradska Street to Gagarin Avenue).

The following routes are distinguished based on the average travel time on the route (Fig. 10.6):

- From Entry Gate 6 to Exit Gate 2 (turn around the intersection from Metalurhiv Avenue).
- From Entry Gate 8 to Exit Gate 5 (traffic from Volhohradska Street to Haharina Avenue).
- From Entry Gate 1 to Exit Gate 3 (turn around the intersection from Myru Avenue).
- From Entry Gate 1 to Exit Gate 7 (traffic from Myra Avenue to Volhohradska Street).

The *Traffic Analysis Report* from the DataFromSky program is the input traffic volume (Vehicle Inputs) [1] for microsimulation in VISSIM, according to which the total number of vehicles through Entry Gate 1, Gate 4, Gate 6, Gate 8 is received as a transport load in one hour.

The traffic dynamics parameters determined by the Datafromsky Viewer are:



	Exit Gate 2 (2)	Exit Gate 3 (3)	Exit Gate 5 (5)	Exit Gate 7 (7)
Entry Gate 1 (1)	00:05.620	00:22.520	00:13.388	00:23.093
Entry Gate 4 (4)	00:17.733	00:13.469	---	00:12.900
Entry Gate 6 (6)	00:22.928	00:16.528	00:05.720	00:13.486
Entry Gate 8 (8)	00:15.014	00:04.378	00:21.837	---

Fig. 10.6. OD-matrix by average travel time, s.

- the average acceleration / deceleration value to adjust the behavioral model of lane change in VISSIM;
- the distribution of average speed.

According to the results of processing 6 143 trajectories, the average longitudinal deceleration is $-0.47\ m/s^2$, and the maximum longitudinal acceleration is $6.11\ m/s^2$. These values are used in the VISSIM Lane Change Behavior model.

The frequency histogram is created at the beginning of the process to establish the law of distribution of the average speed of movement from the Traffic Analysis Report (Fig. 10.7).

The speed distribution (interval series) is approximately normal. A technique based on the establishment of frequency and accumulative weights was used to test this hypothesis. As a result, an empirical distribution function is produced (Fig. 10.8). The resulting curve, as shown in Fig. 10.8, corresponds to the function of normal distribution.

Using Gauss' formula, we calculate the density of the probability distribution for the investigated velocity indicator (Fig. 10.9).

When calibrating the VISSIM model, we will use the normal distribution

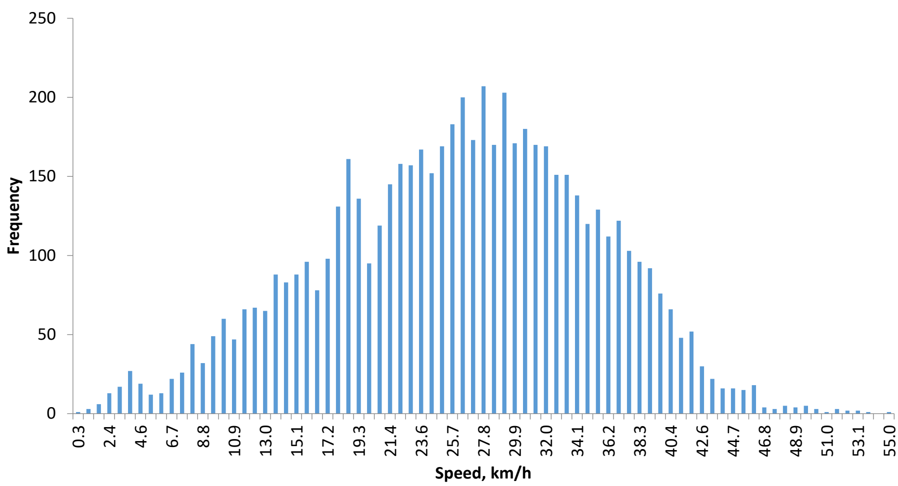


Fig. 10.7. Frequency histogram for average speed based on DataFromSky Viewer report.

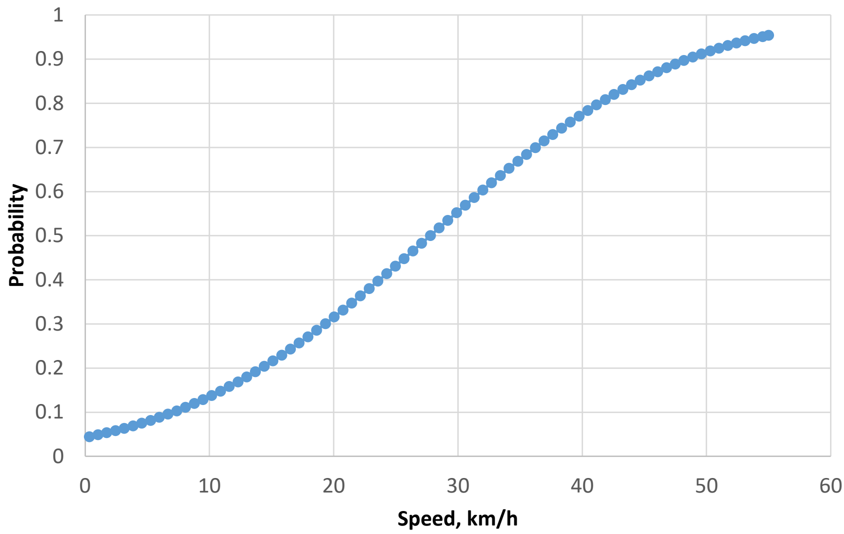


Fig. 10.8. Empirical curve of the normal distribution function

function curve for practical reasons.

The Traffic Analysis Report also obtained the incoming traffic flow parameters for each of the gates for 1 hour of video (Table 10.3).

The intersection model was created in PTV VISSIM software considering

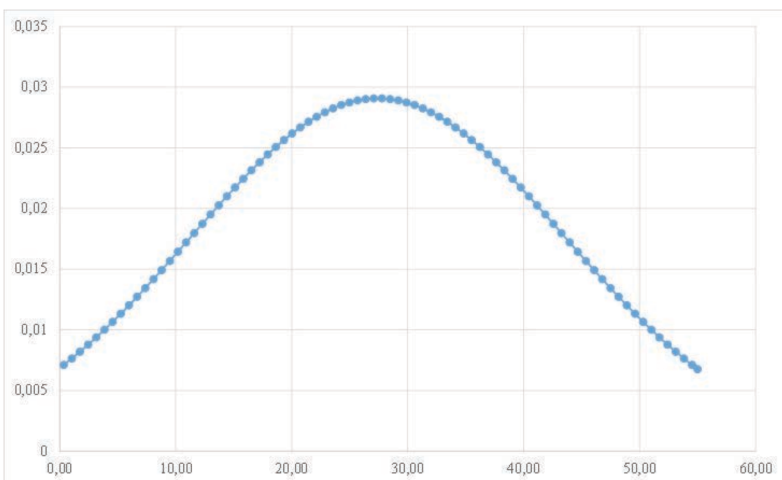


Fig. 10.9. Density of probability of speed distribution.

Table 10.3. Distribution of Vehicle Types in Input Traffic Volumes.

	Gate 1	**Gate 4**	**Gate 6**	**Gate 8**
Cars, ratio	0.88	0.98	0.99	0.89
Trucks, ratio	0.01	0.01	0.00	0.02
Buses, ratio	0.10	0.01	0.01	0.09
Bicyclists, ratio	0.00	0.00	0.00	0.00

the features of three program structural blocks: infrastructure objects (number of lanes, intersection configuration, lane width); traffic organization (zone conflicts); and road traffic (incoming flows and vehicle routes) [6].

The model's input traffic flows are determined by the total number of counted vehicles that pass through the input counter (Fig. 10.10).

The created simulation model is subject to a calibration procedure based on the results of DataFromSky's video motion analysis.

The flow rate is adjusted in the first stage of calibration.

There are two kinds of speed: desired speed and actual speed. When driving,

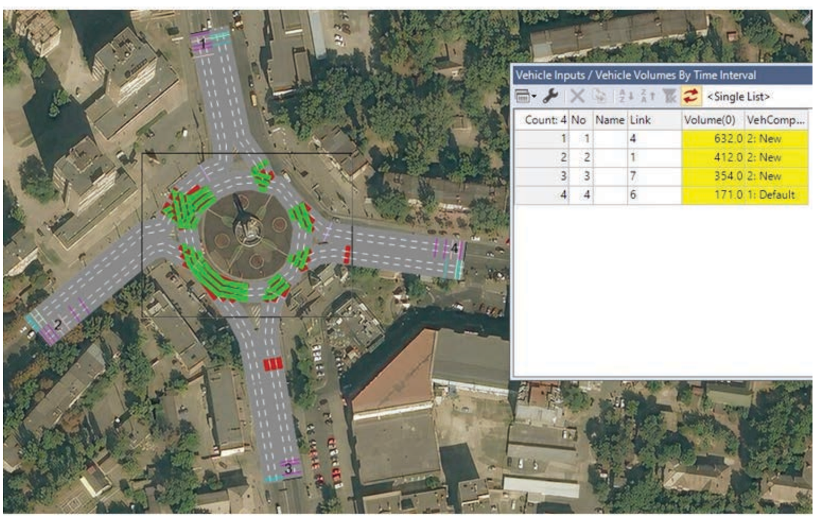


Fig. 10.10. Traffic intensity at the entrance of routes (veh/hour).

certain factors will always determine the maximum speed that can be maintained. They will be determined by the roadway structure and terrain conditions (i.e., horizontal and vertical curves of the road), the level of aggressive driving, and the interruption of the flow of surrounding vehicles (i.e., congestion, traffic signals, stop signs). In the absence of these constraints, it would be possible to travel at a speed determined solely by the roadway structure and the level of aggressive driving. The 'desired speed', also known as free-flow speed, is the fastest you can travel in this state. In reality, maintaining the 'desired speed' is impossible due to a variety of constraints, such as traffic signals and slow-moving vehicles. The adjusted speed is referred to as the 'actual speed' in this case [6].

In VISSIM PTV, only 'desired speed' can be entered, whereas 'actual speed' can be measured as MOEs after simulation.

To access the desired speed distribution list, navigate to 'Base Data> Distributions> Desired Speed'. After that, the Desired Speed Distributions window will appear. The window may include a list of previously desired speed distributions that are already in use in the default program [6].

By clicking the green 'Add' button or right-clicking in the window of the desired distribution speeds, a new distribution is created, which opens the 'Desired speed distribution' window, which defines:

- Distribution number
- Distribution name
- Minimum desired speed
- Maximum desired speed

A distribution number is a one-of-a-kind identifier assigned to each probability distribution. When you use the number that best describes the distribution, it becomes simple to identify each event probability distribution. For example, if you intend to use a specific distribution for a specific speed limit (say, 40 km/h), you can enter number 40.

PTV Vissim allows you to further determine the distribution using the appropriate schedule by specifying a minimum and maximum desired user-defined speed. The distribution graph's axes are as follows: the x-axis represents the desired speed point, and the y-axis represents the probability of the selected desired speed point.

We use the data from the report Datafromsky Viewer as the minimum and maximum values of the speed, and as a graph – the empirical curve of the normal distribution function, as shown in Fig. 10.7.

Figure 10.11 depicts the created graph of the distribution of the average desired speed for Horkoho Square.

When a vehicle enters the network, it is assigned a fractil value, which is used to set the appropriate speed value from the desired probability distribution when the vehicle passes through the specific points or areas of deceleration. This value is stored in the *Desired Speed Fractile* attribute and will remain constant for the vehicle throughout the simulation start. For example, if the fractil value is 45 %, the vehicle will always be assigned a distribution value of the desired speed of 45 %. If the value is set to 100 %, the vehicle will always have the highest level of speed distribution.

In addition to the speed distribution, the traffic composition is subject to calibration through the *Vehicle Compositions* function (Fig. 10.12) according to the Traffic Analysis with DataFromSky Viewer for each input counter.

The lane change model also sets the threshold values of decelerations (Fig.

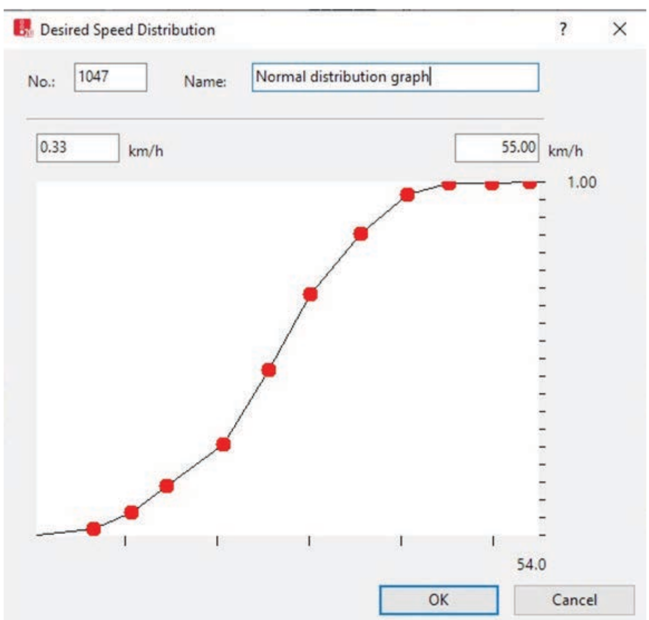
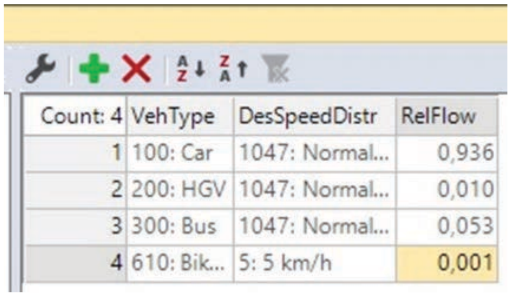


Fig. 10.11. Desired speed distribution graph.



Count: 4	VehType	DesSpeedDistr	RelFlow
1	100: Car	1047: Normal...	0,936
2	200: HGV	1047: Normal...	0,010
3	300: Bus	1047: Normal...	0,053
4	610: Bik...	5: 5 km/h	0,001

Fig. 10.12. Calibration of the composition of the traffic flow.

10.13).

Thus, the data from DataFromSky Viewer in the form of OD-matrices served as the hourly traffic volume input for VISSIM software. The DataFromSky Viewer statistical report on traffic dynamics was used to calibrate the VISSIM simulation model, with the calibration parameters being:

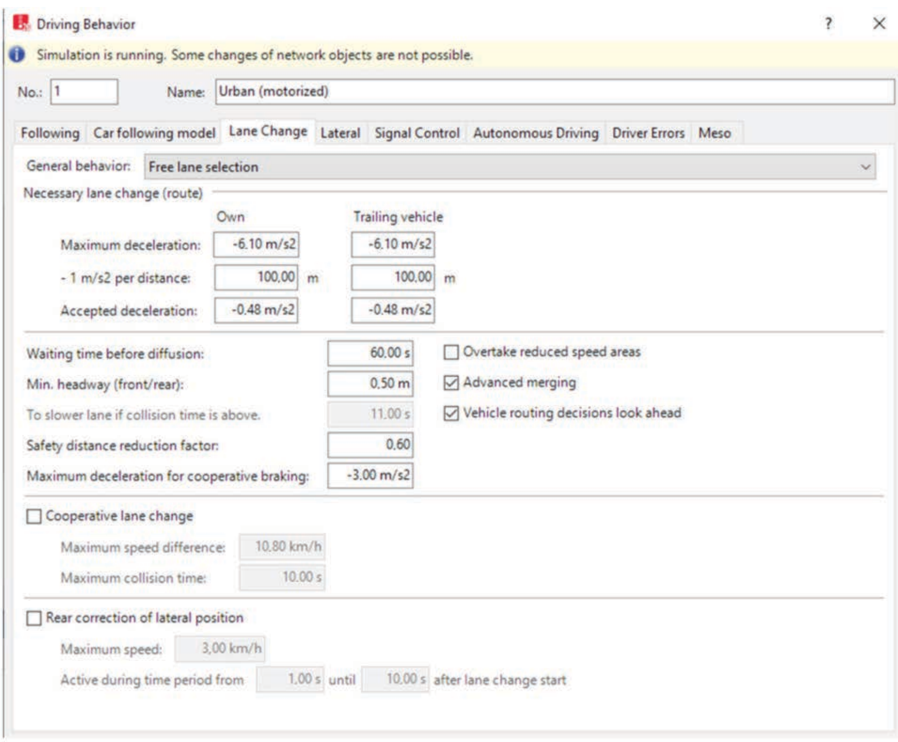


Fig. 10.13. The Lane change model calibration.

- the value of the average acceleration / deceleration to adjust the behavioural model of a lane change in VISSIM;
- the probability distribution of average speed;
- information on the composition of the traffic flow.

10.2.7. Simulation results analysis

Road traffic simulation, particularly at the micro-level, is largely a creative endeavour that includes a significant number of assumptions and hypotheses accepted by the specialist during model development. Each such specialist decision must be convincingly substantiated and documented. For the possibility of external verification, all stages of model development should be described in as much detail as possible in the reporting materials. It is necessary to develop requirements for the form and content of the report on simulation results for the convenience of model verification by any of the stakeholders [9].

In traffic management projects, it is advisable to make out a report on simulation results specifying the following content:

- The report number and date
- A foundation for modelling
- The name of the street / road / road network / local object
- The length of the street / road / road network / local object
- Research schedule
- The person responsible for the research practice
- Survey findings
- Report on the development of a model of the current situation, including calibration and validation
- A list of researched traffic organization measures
- A design of simulation experiment
- Simulation results

- Modifications to the proposed measures
- Conclusions based on simulation results
- Additional information

The sections of the report listed below can provide the following information:

- Description of the problem to be investigated
- Justification for the use of modeling
- Description of the project's role and place
- Formulation of project goals
- Description of the design area
- Justification for the modelling method chosen
- Determining the list of modes of transport considered in the model
- Compiling a list of source data and a description of the process by which they were obtained
- Description of the options and scenarios to be investigated
- Substantiation and selection of measures of effectiveness (MOEs)
- Assessment of the need for additional data processing
- Description of the source of obtaining the O-D matrix
- Justification for dividing the matrix into periods
- Description of the process of finalization of O-D matrices
- Description of the model development process
- Description of the method of checking the correctness of the input data
- Selection of indicators for calibration

- Documenting the model calibration process
- Documenting the model validation process
- Support for the method of determining the forecast level of demand
- Description of the process of analyzing the model's sensitivity
- Description of the experimental design
- Description of the need to adjust traffic organization in the perspective period's model
- Analysis of experimental results (graphs, diagrams, tables)
- Additional data processing
- Description of the scripts of the reporting videos
- Conclusions and recommendations based on simulation results

To visualize the simulation results, videos with a simulation of traffic flow can be used. To record a video, it is best to replicate the model with the measures that are closest to the average values. Video scripts are chosen at the discretion of specialists based on the project's goals, objectives, and features. It is recommended that such scenarios be pre-coordinated with stakeholders [5].

10.3. Assessment of the level of service (LOS) of the cyber-physical traffic system based on the HCM method

10.3.1. Partial evaluation criteria

The large number of different tasks and situations encountered when working with the TCPS inspires the idea of evaluating it using a set of partial criteria. In recent years, numerous criteria for assessing the quality of TCPS have been considered in publications [5], [7]. Partial criteria for evaluating TCPS can be classified by types of traffic and system elements. Table 10.4 [3] summarizes this classification.

It is believed that traffic intensity is the determining factor in assessing the operation of TCPS, which is included, for example, in the calculation of phase coefficients of traffic lights and, as a result, the duration of traffic lights and the

Table 10.4. Criteria for Assessing Traffic Management on Individual Elements of the Road Network.

Traffic type	**Element of the road network**	**Evaluation criterion**
Vehicles	**1**	**2**
	Section of a road	Capacity Speed
	Flying junction	Capacity
	Roundabouts	Capacity Average delay Total delay Queue length
	Signalized intersections	Capacity Average delay Total delay Queue length Percent of stopped vehicles
	Uncontrolled intersections	Capacity Average delay Total delay Queue length Percentage of stopped vehicles
	Road network	Capacity Total delay Travel time Number of stops (when using the network)
Pedestrians	Sidewalks	Capacity Flow density Speed
	Zebra crosswalk	Average delay Queue length
	Signalised crossing	Capacity Average delay Queue length

permit signal for a specific direction [4].

In calculations, the physical value of intensity is not used, and it is reduced to the conditional car using the corresponding reduction coefficients Kr. Furthermore, the reduction is carried out according to different criteria in different countries [8].

The saturation flow is another determining parameter in the design and analysis of the TCPS. This is the maximum flow of vehicles from the infinitely long queue in front of the stop line that moves to the traffic light and passes through the intersection during this signal. It also reflects the effect of traffic disruptions during traffic light control on its behaviour [3]. All known methods of measuring saturation flow take a long time. As a result, when designing or improving the control mode, the values of the ideal saturation flow and its adjustment coefficients, which account for local conditions, are used.

The current method of calculating the saturation flow is quite simple [8] because the basic value of the saturation flow is determined taking the carriageway width into account using the following empirical formula:

$$S = 525Bcw , \tag{10.1}$$

where Bcw is the width of the carriageway of a certain direction in which the traffic flows in the appropriate phase, m.

The saturation flow is adjusted by appropriate correction factors in the presence of the longitudinal slope of the road, the radii of its roundings in the plan, and the distribution of vehicles in the directions of movement. One additional factor considers other parameters that influence traffic conditions.

To assess the effectiveness of changes to the TCPS, several quantitative parameters (quality control criteria) were used. These include, among other things, average and total delay, driving time, number of stops, average speed, queue length (average and maximum), congestion time, probability of passing the first traffic light signal, throughput, control symmetry (relative to driving directions or conflicting flows), emissions, traffic noise, fuel consumption, and so on.

It is worth noting that the average delay of the vehicle before the intersection is one of the most commonly used criteria in optimizing traffic management. The average vehicle delay duration has been widely used as a criterion for optimizing control at a specific intersection.

It has been established that the average delay is closely related to indicators such as traffic intensity, queue length, total delay, and control mode parameters. One of the most acceptable indicators for managing the network in the state of saturated flows, when it is necessary to reduce the probability of congestion, is the queue length.

10.3.2. Integral criterion for evaluating the transport network

The level of service criterion (Level of Service, or LOS) [3] adopted from queuing theory can be used as an integral criterion for assessing the TCPS. The main characteristics of the queuing system (the length of the queue at a given point in time, the length of the period during which the n-th request is waiting for service, the average length of stay of the application in the system) occasionally necessitate complex calculations. As a result, the concept of using a simplified indicator, namely the load factor, to assess traffic conditions arose:

$$k = N/P, \tag{3.2}$$

where N denotes the intensity of the requirements and P denotes the intensity of the service requirements.

The criterion was also chosen for the sake of clarity for a wide audience. The basis of the gradation of LOS (Table 10.4) is the load factor (volume-to-capacity ratio).

In the United States, the level of service has become the primary criterion for assessing transportation networks and traffic management, and it has been included in regulations, most notably in the guidelines for assessing road capacity Highway Capacity Manual (hereinafter HCM2000) [3]. This criterion is not currently used in Ukraine's road construction standards or industry standards, but it is being considered by design institutions with the involvement

Table 10.4. Graduation of Level of Service [3].

LOS assessment	**Value**	**Traffic conditions**
A	≤0.1	Free flow
B	≥0.1	Steady flow
C	≥0.3	Steady flow
D	≥0.7	Flow close to unstable
E	≥1.0	Unstable flow

Table 10.5. Road Network Elements and Types of Traffic Considered in HCM [3].

Road network element	**LOS parametres**
Interrupted traffic (signalized and uncontrolled intersections where vehicles are delayed)	Cars
Urban streets	Speed
Signalized intersections	Delay
Unsignalized intersections	Same
Roundabouts	–
Ramps and its connections	Delay
Continuous movement	
Two-lane freeways	Speed, time share of car-following
Multi-lane freeways	Density
Zones of intersection of flows	÷
Passenger transportation on the route	Speed
Pedestrians	–
Bicyclists	Space, delay
Urban highways	Same

of transport modelling experts.

The latest edition of HCM 2000 [3] proposes LOS indicators for most elements of the TCPS (Table 10.5). Naturally, for each type of traffic (transport, pedestrians) and each type of road network elements (races, intersections, sidewalks, pedestrian crossings), a certain indicator and the appropriate method of its determination are used. Moreover, the criteria used as indicators of the level of service can be used in other types of assessments – environmental and economic (Table 10.6).

One of the most commonly used qualities of service criteria for queuing systems is average delay or duration of service. Naturally, the average delay is traditionally used to assess the level of service provided by vehicles at a signalized intersection (see Table 10.7).

A fundamentally important area of LOS upgrade is considered to be [3] the creation of assessment methods (multimodal LOS) – a comprehensive assessment of LOS, taking into account the joint movement of different users (road transport, passenger transport, bicyclists and pedestrians). Because different types of users interact in the space of an urban street, it is critical to understand how changing the LOS of one infrastructure object affects the LOS of another.

10.3.3. Determination of the level of service (LOS) of cyberphysical traffic systems (TCPS) by means of microsimulation

Table 10.8 shows a comparison of the criteria for evaluating TCPS, which can be obtained by the HCM method and microsimulation in PTV VISSIM. So, the value of traffic flow density, delay and volume to capacity ratio are computational parameters that cannot be defined directly through a simulation model analysis.

The volume to capacity (v/c) ratio is not calculated through microsimulation in VISSIM due to the stochastic nature of this value. It is necessary to argue in favour of the fact that the capacity of road sections is the result of deterministic approach. Although the capacity could be determined in any moment of simulation time, this value could not be the same in other moments in the simulation process while complying with the components that define the capacity variance.

Optionally VISSIM can determine the value of vehicles delay but this value is not HCM compliant. Properly calibrated microscopic simulation models will produce delays that more accurately reflect field operations related to the given network geometry, multimodal volumes, and control strategies than deterministic equation based on methods like those included in HCM (see Table 10.8).

In HCM, delay is considered as control delay and vehicles stop. It is determined as follows:

$$d = d_1(PF) + d_2 + d_3, \tag{10.3}$$

where

d1 – control delay, veh/s;

PF – adjustment factor for the traffic light;

Table 10.6. LOS Indicators and their Relationship to Other Types of Assessments in HCM [3].

Road network element and traffic types	**Criteria used to determine LOS**	**Possibilities of using criteria for other evaluations**		
		Air condition	**Level of noise**	**Economic assessment**
Urban streets	Speed	+	+	+
	Travel time	+	-	+
	Delay	+	-	+
Controlled (signalized) intersections	Delay	+	-	+
	Volume -to- capacity ratio	+	-	+
Uncontrolled intersection	Delay	+	-	+
	Queue length	+	+	+
	Capacity ratio	+	-	+
Pedestrian traffic	Space	-	-	-
	Delay	-	-	-
	Speed	-	-	+
	Volume-to-capacity ratio	-	-	+
Two-lane freeways	Percentage of Car-following	-	-	-
	Speed	+	+	+
Multi-lane freeways	Traffic flow density	-	-	-
	Speed	+	+	+
	Volume-to-capacity ratio	+	-	+
Sections of highways	Traffic flow density	-	-	-
	Speed	-	-	+
	Delay	+	+	+
	Travel time	-	-	+
Model section of a road	Traffic flow density	-	-	-
	Speed	+	+	+
	Volume-to-capacity ratio	-	-	-

Road network element and traffic types	Criteria used to determine LOS	Possibilities of using criteria for other evaluations		
		Air condition	Level of noise	Economic assessment
Intersections of traffic flows	Traffic flow density	-	-	-
	Speed at an intersection's area	+	+	+
	Speed out of an intersection's area	+	+	+
Ramps	Traffic flow density	-	-	-
	Speed	+	+	+
Hill's junctions	Delay	+	-	+
Public transport	Movement interval	+	+	+
	Daily worktime	+	+	+
	Passenger compartment workload	+	+	+
	Reliability	+	+	+

d2 – additional delay for casual arrivals and queuing saturation which are adjusted according to the analysis period and a type of traffic light;

d3 – start delay of queuing which causes subsequent delay of all analyzed vehicles.

The first and second values are defined as follows:

Table. 10.7. LOS for Signalized Intersections Based on the Value of the Average Delay [3].

LOS	Traffic conditions	Time, s
A	There is no or minimal delay	≤5
B	The regulation cycle is only slightly longer than usual, and there is good coordination	5.1–15
C	The length of the adjustment cycle has increased, owing to good coordination	15.1–25
D	The regulation cycle lasts a long time, and there is good coordination	25.1–40
E	The regulation cycle lasts a long time, and coordination is poor	40.1–60
F	Most drivers find traffic conditions to be unacceptable; the intensity on the approaches exceeds the capacity of the intersection	>60

$$d_1 = \frac{0,5 \cdot \tilde{N} \cdot (1 - g/\tilde{N})^2}{1 - \left[g/\tilde{N} \cdot \min(1, X) \right]}, \tag{10.4}$$

$$d_2 = 900 \cdot T \cdot \left[(X-1) + \sqrt{(X-1)^2 + \frac{8 \cdot k \cdot I \cdot X}{c \cdot T}} \right], \tag{10.5}$$

where

T – duration of the analysis period, h;

C – cycle length, s;

k – delay factor that depends on traffic light settings;

I – filtration adjustment factor;

c – capacity of a signal group (veh/ h);

X – saturation rate.

Saturation rate is

$$X_c = \frac{Y_c \cdot C}{C - L}. \tag{10.6}$$

The capacity is defined with retaliation to adjustment saturation flow for the appropriate lane:

$$c = s_i \cdot \frac{g}{C}, \tag{10.7}$$

where g is green time signal, s; C is control time, s; si is saturation flow, veh/h.

In HCM the set of adjustment factors is presented for an accounting additional impacts on the basic value of saturation flow:

$$s = s_0 \cdot N \cdot f_w \cdot f_{HW} \cdot f_g \cdot f_p \cdot f_{bb} \cdot f_a \cdot f_{LU} \cdot f_{LT} \cdot f_{RT} \cdot f_{Lpb} \cdot f_{Rpb}, \tag{10.8}$$

where

s_0 – base saturation flow, (vh/h/ln);

N – number of lanes, N;

Table 10.8. The Fusion of TCPS Evaluation Criteria which can be Obtained from HCM Technique and PTV VISSIM Simulation Model.

Evaluation criteria for defining LOS	HCM technique	VISSIM simulation model
Speed	+	+
Travel time	+	+
Queue length	+	+
Travel time	+	+
Flow density	+	-
Vehicle delay	+	+
Pedestrian delay	+	+
Volume of capacity	+	-
Public transport schedule	+	-
Vehicles emissions	-	+

f_w – lane width adjustment factor;

f_{HW} – heavy-vehicle adjustment factor;

f_g – grade adjustment factor;

f_p – parking adjustment factor;

f_{bb} – bus blockage adjustment factor;

f_a – area type adjustment factor;

f_{LU} – lane utilization adjustment factor;

f_{LT} – left-turn adjustment factor;

f_{RT} – right-turn adjustment factor;

f_{Lpb} – left-turn ped / bike adjustment factor;

f_{Rpb} – right-turn ped / bike adjustment factor.

Thus, capacity ratio and saturation flow are the main factors for additional delay definition. LOS is defined just from the additional delay.

Ultimately, the travel time has the main influence on the average traffic speed ST:

$$ST = T_r + d. \tag{10.9}$$

In order to automate the assessment of LOS for TCPS, we used the appropriate program code for VISSIM. For this purpose, the parameters from VISSIM were used, namely, the value of delay and average control delay, and average traffic speed. For the node LOS calculation, the user defined attributes (UDAs) were used in VISSIM. There were four UDAs:

- Attribute LOS to calculate the average node LOS
- Attribute WorstLOS to determine the worst LOS of all time intervals and simulation runs
- Attribute WorstMovLOS to calculate the worst traffic LOS
- Attribute NodeLabel to show the results in one label for current, previous interval and worst throughout all simulation runs and time intervals

The values for UDAs were accepted according to the HCM table for signalized intersections [8]. The programming code for the first UDA is as follows:

```
IF([TOTRES\VEHDELAY(…ALL)]≤10; "A";

IF([TOTRES\VEHDELAY(…ALL)]≤20; "B";

IF([TOTRES\VEHDELAY(…ALL)]≤35; "C";

IF([TOTRES\VEHDELAY(…ALL)]≤55; "D";

IF([TOTRES\VEHDELAY(…ALL)]≤80; "E"; "F"))))).
```

To calculate the worst node LOS the following code was used:

```
IF(NUMTOSTR([TOTRES\VEHDELAY(MAX, MAX, ALL)])=";";

IF([TOTRES\VEHDELAY(MAX, MAX, ALL)]≤10; "A";

IF([TOTRES\VEHDELAY(MAX,MAX, ALL)]≤20; "B";

IF([TOTRES\VEHDELAY(MAX, MAX, ALL)]≤35; "C";

IF([TOTRES\VEHDELAY(MAX, MAX, ALL)]≤55; "D";
```

IF([TOTRES\VEHDELAY(MAX, MAX, ALL)]≤80; “E”; “F”)))))).

For estimation of the worst movement LOS we made use of the code:

IF(NUMTOSTR([MAX:MOVEMENTS\VEHDELAY(MAX, MAX, ALL)])=”;”;

IF([MAX:MOVEMENTS \VEHDELAY(MAX, MAX, ALL)]≤10; “A”;

IF([MAX:MOVEMENTS \VEHDELAY(MAX,MAX, ALL)]≤20; “B”;

IF([MAX:MOVEMENTS \VEHDELAY(MAX, MAX, ALL)]≤35; “C”;

IF([MAX:MOVEMENTS \VEHDELAY(MAX, MAX, ALL)]≤55; “D”;

IF([MAX:MOVEMENTS \VEHDELAY(MAX, MAX, ALL)]≤80; “E”; “F”)))))).

To show the node label we used the code:

“Current node LOS:”=[LOS, CURRENT, CURRENT];

“Last interval LOS:”=[LOS,CURRENT, LAST];

“Worst node LOS:”=[WORSTLOS];

“Worst movements LOS:”=[WORSTMOVLOS].

An example of the use of intersection LOS determining by means of micro-simulation is showed further.

The Department of Road Safety of the National Patrol Police of Ukraine in the city of Kryvyi Rih has set the task to investigate the possibility of queuing on the bridge in Bykova Street if the traffic lights at the intersection of Bykova and Ivana Avramenka streets are installed.

Figure 10.14 shows the proposed timing program and traffic lights locations (Project CB-04/1-18-EH).

The VISSIM-simulation results show that queuing is absent for the unsignalized intersection option. It is also proved by the data obtained during investigation of traffic flows on the interchanging.

If traffic signalization is installed, the average queue length of 84.4 m before the intersection will be on the ramp from Bykova Street, 38.1 m before T-shape intersection at the same street at the ramp from the bridge, 47.3 m from the ramp on the side of Symonova Street.

LOS determination is obtained considering the average delay at the intersection, delay upstream the traffic light, and the average traffic speed. The comparison of the obtained parameters with the same ones determined via HCM technique is shown in Table 10.9.

The maximum deviations between the traffic management performance indicators in the intersection are obtained for such criteria as the flow density ratio and average delay. They are 35.1 % and 21.0 %, respectively. HCM technique and VISSIM-model made such data, as the average travel speed and the average delay differ by 11.0 % and 20.0 %, respectively. The worst node LOS is F. It was obtained upon indication of flow density for the two assessment methods. This implies that the average speed at the interchanging is 25–33 % of the speed in freeway conditions (FFS) for this urban street type, and extensive delays and queuing are observed at the intersection. The VISSIM model versus HCM technique showed more similar scenario if the decision on signal lights installation is made. Based upon the average travel speed, which is dependent on the running speed and the amount of control delay incurred at signalized intersections, the interchanging

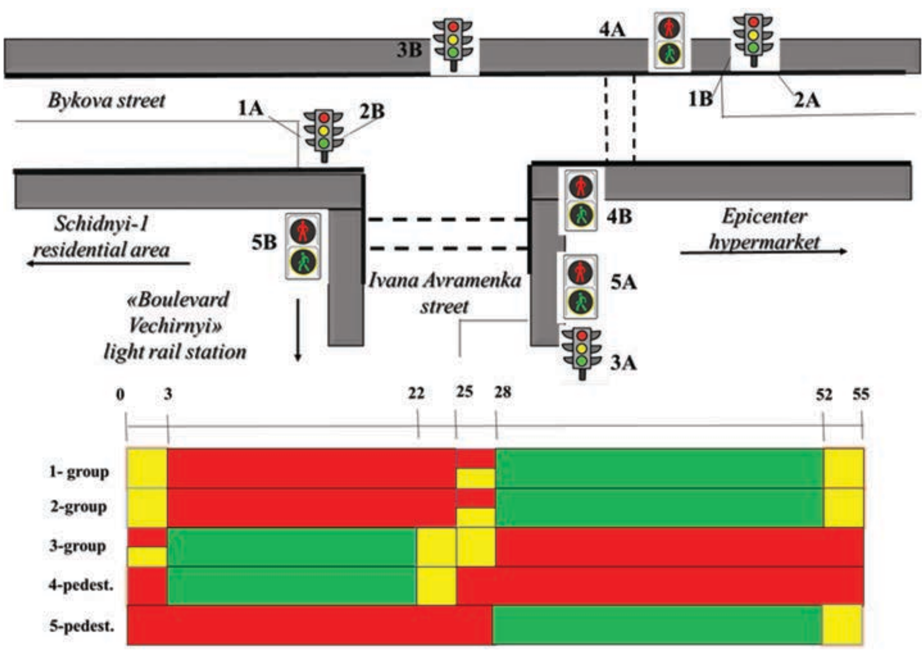


Fig. 10.14. The proposed timing program and traffic lights locations (Project CB-04/1-18-EH).

Table 10.9. The Results of LOS Determination

Parameter	HCM	VISSIM	Deviation, %	LOS criteria	
				HCM	VISSIM
Average travel speed, km/h	28.1	24.57	11.0 %	C	C
Traffic flow density, auto/km	280.1	430.8	35.1 %	F	F
Average delay, s	37.1	47.0	21.0 %	D	D
Control delay (ACD), s	34.1	27.3	20.0 %	C	C

was assigned LOS C. This value describes stable operations; however, ability to maneuver and change lanes in midblock locations may be more restricted than at LOS B, and longer queues, adverse signal coordination, or both may contribute to lower average travel speed of about 50 % of FFS for the street class.

Due to the results of HCM technique and VISSIM model application, it has been determined that according to the average delay and ACD, the interchanging has got LOS D. This implies that small increases in traffic flow may cause substantial increases in delay and decreases in travel speed. Node LOS D occurs due to adverse signal progression, inappropriate signal timing, high traffic volumes, or a combination of these factors. The average travel speed is about 40 % of FFS.

10.3.4. LOS and MOEs determination using traffic video analysis data

Assessment of MOEs and LOS of intersection based on TCPS simulation can comprise the following steps (see Fig. 10.15):

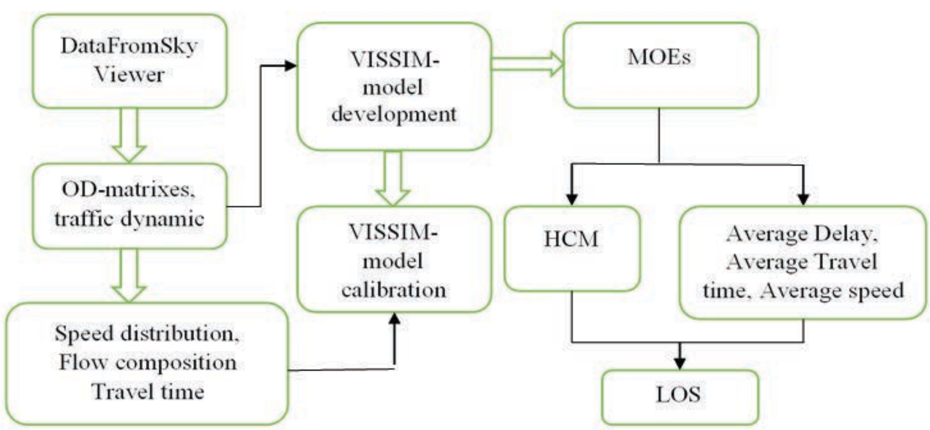


Fig. 10.15. Assessment of MOEs and LOS by means of TCPS simulation.

- traffic video analysis in DataFromSky Viewer software;
- statistical analysis of traffic dynamic and OD-matrixes obtained in DataFromSky Viewer;
- development of microsimulation model of intersection in VISSIM by means of video analysis results;
- calibration of VISSIM-based model by utilizing vehicle speed distribution, traffic flow composition, vehicle travel time;
- comparison of OD-matrixes from DataFromSky Viewer with simulation results in VISSIM;
- determination of MOEs (average delay, average speed, queue length, average travel time);
- LOS calculation based on the obtained MOEs and via HCM-technique.

The implementation of this algorithm is demonstrated for Horkoho Square in Kryvyi Rih. In Subsection 10.2.5, the calibration of a simulation model of this intersection using traffic video processing technology and artificial intelligence is discussed. As a result, we proceed immediately to the final two stages of the presented algorithm (Fig.e 10.15), which are as follows:

- Average travel time
- Average queue length
- Average control delay

Simulation results showed that the average travel time for specific traffic routes ranged from 15 to 38 s, which corresponds to OD-matrices for route time passage. The total travel time is 9 661 s.

The average queue length at the intersection was 1.70 m, with a maximum value of 17.8 m. The queue may form in the traffic direction from Metalurhiv Avenue to Haharina Avenue (from Gate 4 to Gate 7, see Table 10.3).

The average delay associated with the flow interruption (DelayAvg) was 1.27 s. The total stop time (DelayStopTot) was 84 s. The total delay of vehicles in the network was 654 s. The specific delay, which is determined by the total time of all delays, divided by the sum of the number of vehicles in the network and the number of vehicles that have reached their destination, (*DelayStopAvg*) was 0.22 s.

The number of stops (excluding public transport and parking) was 63. The specific number of stops (the total number of stops divided by the sum of the number of vehicles in the network and the number of vehicles that have reached their destination) was 0.19.

The average speed (total distance divided by the total time in the network) at the intersection was 32.7 km/h, which is 16.2 % more than the average speed from the results of traffic video analysis (27.4 km/h).

The number of vehicles that passed to the destination was 311. However, 119 vehicles were not counted among the incoming traffic loads.

The next step in the analysis is to determine the LOS at the investigated intersection. The above-mentioned traffic efficiency indicators will be used for analysis. To calculate the LOS in VISSIM, we will use the user's first attribute settings, specifically the average value of LOS, taking into account the intersection delay and average speed.

Table 10.10 shows the comparison of LOS evaluation using two techniques.

The deviation in the indicators of delay at the intersection, as determined by two methods, was 30.6 %, while the deviation in the average speed was 10.7 %.

The minimum LOS of the roundabout was in terms of average speed. It was B obtained from HCM-technique [3] and TCPS simulation results. This means that the average speed is 70 % of the speed in the free flow conditions (FFS) for this class of streets (45 km/h), and there are minor obstacles to traffic movement and maneuvers.

LOS A is assigned by two methods in terms of the value of transport delay. This indicates a free flow at the intersection, the average speed is 90 % of the speed in the free flow conditions (FFS) for this class of streets. There are also free conditions for maneuvers.

Table 10.10. Summary on LOS Determination Results.

Indicator	**HCM-technique**	**VISSIM-model**	**Deviation, %**	**LOS-criterion**	
				HCM	**VISSIM**
Average speed, km/h	29.2	32.7	10.7 %	**B**	**B**
Average delay, s	1.83	1.27	30.6 %	**A**	**A**

Conclusions

The interaction of road users with each other and the transportation infrastructure is at the heart of the cyber-physical road system (TCPS), as well as other complex systems. The nature of such interaction determines the classification of TCPS. The purpose of the information flow circulating in the system is critical. For example, if we are talking about warning information, then its implementation is usually carried out using traffic organization. For the transport infrastructure it is important to have such type of information as the spatial parameters of the transport system. Work with this type of information is entrusted to geographic information systems (GIS). When modelling traffic, driver behaviour and the parameters of its settings determine the interaction of road users. Transport microscopic modelling software allows to accurately reproduce the driver's behaviour in a computer simulation model. However, these solutions are typically used to automate decision support before the scenario being implemented in the field. The range of cyber-physical systems, the work of which is aimed at improving the efficiency of real-time traffic control, is currently limited to individual software products and embedded devices.

When evaluating TCPS for practical purposes, a set of partial criteria is used, such as traffic intensity, traffic delay, traffic queue length, section capacity, saturation flow, travel speed. It is also possible to apply the integrated service level criterion (LOS-criterion) of the network and its elements. It can be calculated manually or through micro-simulation, and the results will differ from one another.

A promising area is assessing of the effectiveness of the TCPS using the intelligent traffic video analysis, microsimulation software and determination of partial (MOEs) and integrated performance indicator of traffic management (LOS). The initial indicators of the traffic video analysis (first of all, the traffic intensity and the traffic flow composition) can act as parameters for calibrating the simulation model.

This is for the first time hat such an algorithm for evaluating the efficiency of the TCPS is presented.

References

1. DataFromSky TrafficCamera [Accessed online: https://datafromsky.com/trafficcamera/

2. Dusan Jolovic, Aleksandar Stevanovic, Soheil Sajjadi, Peter T. Martin, 2016. Assessment of Level-Of-Service for Freeway Segments Using HCM and Microsimulation Methods, ISEHP 2016. International Symposium on Enhancing Highway Performance, Transportation Research Procedia, 15 (2016) 403–416.

3. Highway Capacity Manual, 2016. 24– 40.

4. Karl Wunderlich, Meenakshy Vasudevan, and Peiwei Wang, 2019. Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software. U.S. Department of Transportation Federal Highway Administration. 15–25.

5. Lipika Deka Mashrur Chowdhury, 2018. Transportation Cyber-Physical Systems 1st ed. Elsevier. 81–83.

6. Protocol for Vissim Simulation, C. Mai, C. McDaniel-Wilson, D. Noval, et al. [Accessed in 2014, http://www.oregon.gov/ODOT/TD/TP/APM/AddC.pdf]

7. Shi Jianjuna, Wu Xub, Guan Jizhenc, Chen Yangzhoua, 2013. The analysis of traffic control cyber-physical systems, 13th COTA International Conference of Transportation Professionals (CICTP 2013), Procedia - Social and Behavioral Sciences, 96 (2013) 2487 – 2496.

8. Volodymyr Sistuk, Yurii Monastyrskyi, 2019. Comparative study of VISSIM and HCM technique LOS determination as exemplified by T-shape and partial cloverleaf interchanges, ICCPT: Current Problems of Transport: Proceedings of the 1st International Scientific Conference, Ternopil. Scientific Publishing House “SciView”. 11-21.

9. Guidelines for the development and implementation of traffic-management actions. The use of software products for mathematical modeling of traffic flows in assessing the effectiveness of traffic organization design solutions [In Russian], 2017. pp. 40–56.