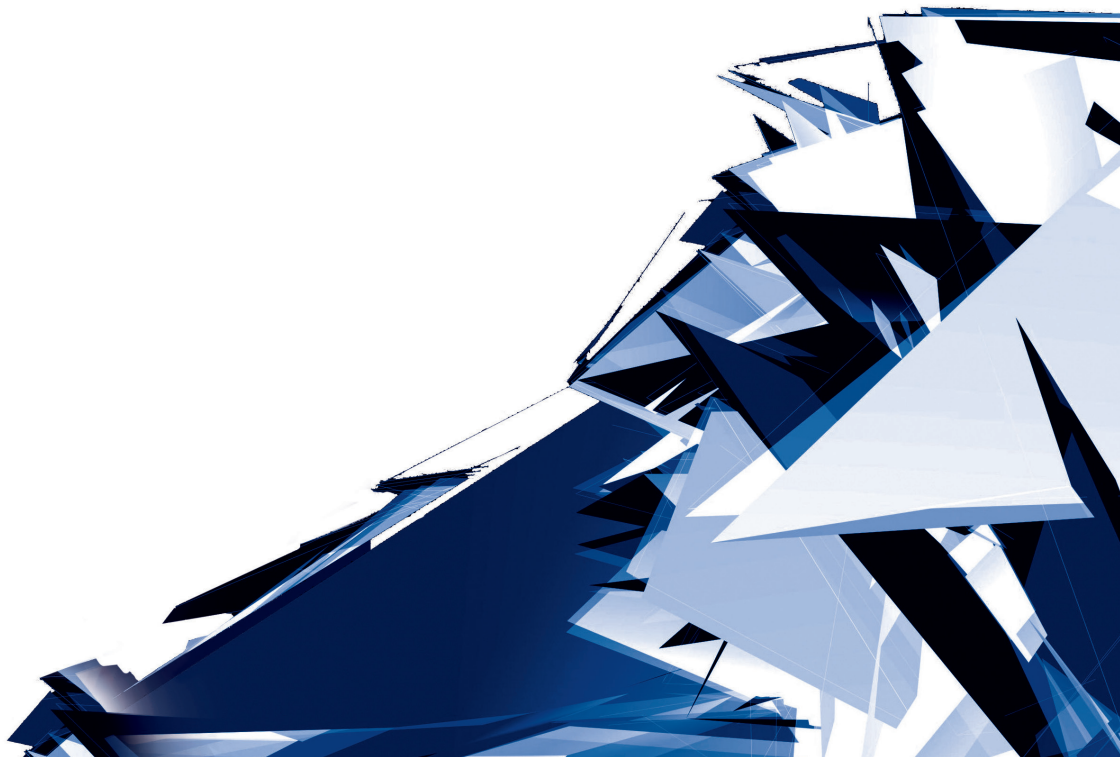
Technical Transactions

Czasopismo Techniczne

Volume 2

Year 2017 (114)



www.ejournals.eu/Czasopismo-Techniczne
www.technicaltransactions.com
www.czasopismotechniczne.pl

ISSN 0011-4561

Basic version of each Technical Transactions magazine is its online version
Pierwotną wersją każdego zeszytu Czasopisma Technicznego jest jego wersja online

Contents

ARCHITECTURE AND URBAN PLANNING:

Nikos A. Salingaros

Eight city types and their interactions: the "eight-fold" model 5

CHEMISTRY:

Barbara Michorczyk, Elżbieta Hędrzak

Study of oxidative coupling of methane integrated with CO oxidation 71

ENVIRONMENTAL ENGINERING:

Michał Grodecki

Numerical modelling of gabion joints 83

Adam Kasprzak, Paweł Popielski

Comparison of temperature and displacements of a gravity section and a buttress section illustrated with the example of concrete dam in Rożnów 91

Justyna Kwaśny, Wojciech Balcerzak

Characteristics of selected methods for the synthesis of nanometric zirconium oxide – critical review 105

Dawid Łątka

Stress state laboratory verification in masonry structures according to the flat jack method 119

Michael Poehler, Paul Uwe Thamsen

Design of a decentralised intelligent network for five wet pit pumping stations 129

MECHANICS:

Zygmunt Domagała, Krzysztof Kędzia

Analysis, modelling and verification of the phenomena occurring in a hydraulic prop during dynamic load 139

Karolina Głogowska, Janusz W. Sikora

The production and properties of micro-extruded low-density polyethylene microtubes 155

Marek S. Kozień, Dariusz Smolarski

Comparison of results of stress cycle counting by the direct spectral method and the Fu-Cebon method for bi-modal type stress processes 167

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
ARCHITECTURE
DOI: 10.4467/2353737XCT.17.016.6209

Nikos A. Salingaros (salingar@gmail.com)
Department of Mathematics, The University of Texas at San Antonio, USA

Eight city types and their interactions: the "eight-fold" model

Osiem typów miast i ich interakcje: model „ośmioaspektowy"

Abstract

A model for understanding the city in terms of eight characteristic "city types" is proposed. The most human cities consist of adaptive city types that act in a congruent manner. Any city can be analyzed as a particular mixture of these eight types. Competing city types combine and interact in different ways, and users feel the result as an essential quality of the environment. Some types either add to, or cancel and destroy each other, whereas others can juxtapose without interacting. This eight-fold model of city types helps us to predict the success or failure of distinct urban regions in promoting urban life. It also suggests how to repair declining or non-existent pedestrian activity, and how architectural projects could affect the city adversely or positively. One section of this paper is devoted to techniques for designing urban spaces that invite human engagement, and another to designing a campus.

Keywords: Planning, urban morphology, urban design, theories of urbanism, biophilia, complexity, fractals, networks

Streszczenie

Zaproponowano model rozumienia miasta w kategoriach ośmiu charakterystycznych „typów miast". Najbardziej uczłowieczone miasta składają się z adaptacyjnych typów, które funkcjonują w sposób spójny. Każde miasto można traktować jako szczególną mieszankę tych ośmiu typów. Konkurujące typy miast łączą się ze sobą i współdziałają na różne sposoby, a rezultat tych interakcji jest odbierany przez użytkowników jako integralna cecha otoczenia. Niektóre typy uzupełniają się nawzajem lub wzajemnie się znoszą bądź niszczą. Z kolei inne typy funkcjonują obok siebie, bez wchodzenia w interakcje. Ośmioaspektowy model typów miast pozwala przewidzieć sukces bądź niepowodzenie poszczególnych regionów miejskich w promowaniu życia miejskiego. Stanowi on również wskazówkę co do tego, jak można odbudować malejący lub nieistniejący ruch pieszych oraz czy dane projekty architektoniczne będą miały negatywny czy też pozytywny wpływ na miasto. Jedna z części tego artykułu jest poświęcona technikom projektowania przestrzeni miejskich zachęcających użytkowników do interakcji, a druga dotyczy projektowania kampusu.

Słowa kluczowe: planowanie, morfologia miasta, urbanistyka, teorie urbanistyczne, biofilia, złożoność, fraktale, sieci

Part 1: Eight city types

Introduction

A healthy city helps us to live and enjoy our lives fully, and that quality is determined by the urban structure.

Concepts mostly new to urban planning such as biophilia, complexity, fractals, and networks offer a better way of designing a city. The basis for these principles is theoretical, yet they lead directly to real-world applications. I'm addressing the results to a wider audience of practicing design professionals who wish to make their product better. I provide theoretical backing for new methods that work, and also identify what doesn't. The eight-fold model has been derived from scientific inquiry and experiment. Urban fabric designed using the eight-fold model will foster a more human environment than what has been implemented in the past 70 years.

There exists an enormous market for good design and human-scale urban spaces, and, if given a choice, people will choose those. The proposed alternatives to the standard methods of design can be applied at little or no additional financial cost. Most developers already know that they can be more successful with good design than with bad. It's simply a matter of understanding what is "good" versus what is "trendy". In the case of government projects, where the benefit is meant to serve the inhabitants, these alternative design methods guarantee a more human result. Politicians who align themselves behind an innovative human-scale methodology better serve the interest of their constituents.

The forces behind the insipid global uniformization occurring everywhere combine pure ideology with a ruthless profit motive, and are not in the least interested in local identity and culture. Left to itself, this agenda imposes a homogeneous urban style. Cities or countries that resist this and attempt to assert their heritage and tradition are too often branded as "backward" in the international press, which is content to follow destructive fashions, and is thus complicit with efforts to erase local characteristics. My objective is to provide new tools for urbanists the world over wishing to protect their city from this transformation.

But there are occasional dangers of being led astray by standard practices, because the pragmatic recommendations presented here compete head-on with accepted typologies and solutions currently in use. A forward-looking architect understands the political and economic advantages of an innovative way of design thinking. This architect will be able to communicate these new proposals to politicians and developers, which will serve their collective interest. The media also play a key role, but they tend to take their cue from the architectural community, which is why it's so important to convince the architects by providing proof.

Implementing urban innovations holds the greatest hope for the future of the world's cities. I offer an alternative that is better for users and for the city as a whole, yet there is a lot of confusion on this point. Persons who have the power to push for positive change have become too used to building cities in a certain way. After decades of being told that building

with conventional typologies was the smart way to make cities, it will require sustained effort to appeal to the common sense and basic intuition of decision makers. But once the paradigm is exposed, and examples of what is bad about modern cities are revealed, they will immediately recognize the advantages of a new method.

An abstract yet practical decomposition

The complexity of a city – for good or bad – is better understood by decomposing it into equally complex components.

To gain a better understanding of the complexity of urban form and function, I propose eight abstract city types. Each one of these "pure" city types is described in terms of its mechanisms, morphology, and how it arises from construction and societal forces. The eight abstract "types" of cities correlate with distinct design methods and associated design philosophies. Some city types add to help one another, whereas others cancel and destroy each other. The actual built city – as observed and experienced – results from interactions among the eight different types. Cultural, social, political, economic, and geographical factors influence this interaction, creating the highly complex system that is the city.

TABLE: THE EIGHT CITY TYPES
NOURISHING-PHYSICAL
FRACTAL
NETWORK
SPONTANEOUS SELF-BUILT
VIRTUAL
DEVELOPER
ANTI-NETWORK
INHUMAN

I am not identifying eight theoretical types of city, but conjecture that every city is some mixture of these eight city types. The combinatoric possibilities of mixing distinct city types in various proportions makes possible an infinite variety of cities with different characteristics. A city also varies its mixture of city types in different places and regions, giving a heterogeneous quality of place. Moving a short distance exposes a user to an entirely different mixture of city types. Several older cities have managed to retain their living qualities, especially in rings around the historical center where the positive mixture of city types results in a healthy and attractive living environment.

At the opposite extreme, homogeneous cities have a poor mixture of city types, or consist predominantly of types that are not adaptive. A functionally homogeneous city that mixes unhealthy city types remains the same over large distances: no displacement suffices to escape from a deadening environment. For this reason, mono-functional zoning kills the life in a city by creating single-use, separated functional regions, each of them with a homogenous

destined use. The pedestrian realm is erased, leading to a homogeneous city with no room for the human scale and meaningful human activity.

This is not a personal opinion, but a mathematical result. A living city unites several component subsystems that depend upon each other. Physically separating the subsystems into abstract entities identified with a single function (instead of partitioning into the eight city types presented here) destroys the overall system. Planners have become aware of this condition, but many have yet to fully understand why this is so. It is because a system rigidly divided into single-use subsets (i.e. business, residential, light industrial zoning, etc.) loses all of the interlinked complexity that gives it its perceived living properties. Simplistic zones containing compartmentalized city functions are not system pieces working inside the whole, because the emergent properties due to their interactions have disappeared (see "The Law of Requisite Variety and the Built Environment").

Living cities have a healthy mixture of uses that extend both in space, and in time. On the short term, the city also serves as an attractor of different users over different times of day, and in different seasons of the year. It is impossible, therefore, to "plan" the city for a unique segment of the population. Even if it is successful for its intended aim, the urban fabric is in danger of remaining unused during those times that the niche population group does not wish to use it. This gap in urban rhythms of occupation leads to urban pathologies.

In addition to environments that change as one moves around, a given place varies its composition of underlying city types over time. Over the long term, this evolution is due to both changing urban morphology, and to the restructuring of urban functions. The temporal development of urban form for different uses and as reacting to urban forces maintains the living city, although a wrong turn can degrade it.

Even though the eight-fold model presented here is original and distinct from previous urban design theories, several researchers have long studied city form as it is directly experienced. One such related school is Urban Morphology. (An entry to that discipline is the book by Serge Salat, *Les Villes et Les Formes*, [21]; the interesting recent paper by Rémi Louf and Marc Barthelemy, *A typology of street patterns* [17]; and our intersection with that approach, *Urban Nuclei and the Geometry of Streets: the Emergent Neighborhoods Model*, [40]) My own work evolves directly from that of Christopher Alexander [1–3].

What each city type does

The eight city types evolved together at different times, but our age has forgotten what each type contributes to the city.

Cities were originally created by people extending their own biology and mind into the built environment [2]. Human functions that demand close proximity with others led to clustering and living in urban densities. To accommodate those functions in a manner that is not dysfunctional or oppressive, city form and the shaping of both interior and exterior spaces evolved into what we see today in traditional settlements. Living urban patterns are the result of flows (pedestrian movement, bicycle and vehicular transport, mass transit, etc.)

and the partial enclosure of urban space by buildings, but not the other way around. Looking at how a city works in relation to its inhabitants, rather than the image it produces, gives us a better way to understand and manage urban places.

The resulting urban conceptions – both positive and negative – are represented in the eight-fold model of abstract city types. They can be conveniently plotted together as an octet, shown in the following figure:

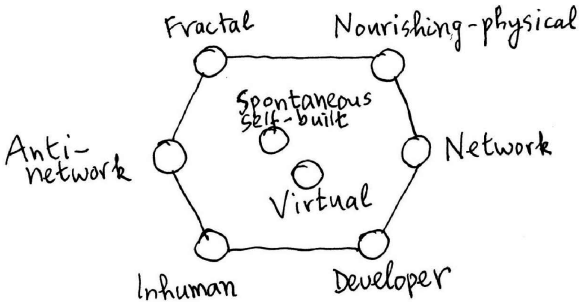


Fig. 1. The eight-fold model of city types

Once you know what to look for, the presence of each city type can be easily recognized while experiencing a piece of urban fabric in person. Knowing the different forces that generate the different city types helps to identify them. Ultimately, the objective is practical: reshaping existing cities by combining the healthy city types will guarantee a more human environment.

The industrialized world experienced a drastic discontinuity after the beginning of the 20th Century. Industrial models of city form were imposed on urban structure. Curiously, no experiments were ever undertaken to verify whether these non-evolved industrial-modernist typologies were in any way optimally adapted to human needs, or even minimally acceptable for the users. Today, it is idealized abstractions of "images of technology" rather than technology itself that impact the city's shape.

We face several obstacles in implementing the eight-fold model to re-build cities. Modern development, given its need for (and impetus from) capital investment, is more often fueled by profit that from satisfying people's needs. Wherever construction is profit-driven without strong and knowledgeable administrative authority in place, forces other than maximizing profit are neglected. It is very easy to override (through political influence and corruption) government regulations destined to protect the human user. Local authorities, whose job it is to protect humans from unsafe and unhealthy environments, are often unable to understand, in a qualitative sense, how different types of development lead to totally different end results. And thus those authorities continue to permit bad environments to be constructed.

Type 1. The Nourishing-physical city

The architect's deepest moral responsibility is to provide a healthy environment for people, where all design components contribute to human wellbeing.

Human beings respond viscerally to forms, colors, surfaces, and spaces. Those reactions are built into our neurological system. Unless we are forced to ignore our basic instinctive response to the immediate environment, we respond according to how our body was designed by evolution. This predisposition has the fundamental consequence that ALL human beings respond in the same way to the elements of forms and spaces in their environment: however, some persons override their instinctive body reactions in order to interpret the city differently in terms of some learned response. (The universality of human responses is established by the work of Edward O. Wilson [24], who disproved assumptions by sociologists claiming isolating cultural preferences).

A traditional city (e.g. the historic center of Krakow) was built by humans for humans to maximize a physical sense of wellbeing that sustains, nourishes, and heals. Urban spaces are connected to flows (e.g. pedestrian, bicycle, vehicular, and public transport) and they cradle human life by providing this psychologically-welcoming environment. People built cities according to their emotions and needs, and functionality that threatened those criteria (buildings or street design that create anxiety in the pedestrian) did not automatically override our neurological signals from the built form. Emotional feedback was valued just as much as mechanical efficiency, therefore, urbanism developed in a balanced tension between mechanical functionality and human feelings. By following its psychological responses, humankind evolved ideal design solutions found today in historical cities: a combination of urban structures on all scales that create the welcoming human environments sought after by both residents and tourists. This fact explains why people from all over the world enjoy the humanity and life found in the best-preserved pieces of traditional urban fabric, and will pay to experience it. Global tourism is a trillion-dollar industry.

Understanding how and why physical built form can give back nourishment to a person comes in part from Biophilia. This is the notion that our body responds positively to the presence of biological forms in our immediate environment (due to Edward O. Wilson: [14]). Those biological forms include both living beings as well as their representations; hence the importance of figurative representational art throughout the ages. The positive effect is also triggered by ordered abstract designs representing the mathematical patterns found in biological systems. Abstract complex patterns such as those found throughout Islamic Art and embodied in the ornamental traditions of all human societies have a strongly healing effect. Their purpose is to foster a greater sense of wellbeing by combating anxious impulses in our mind (see my booklet *Biophilia and Healing Environments* [29]).

The local (close-range) 1characteristics of Biophilia are responsible for the immediate response of our body to the built environment. Elements of Biophilia also include our perception of space up to large distances, the presence of water, and the quality of natural light on the scene. Those factors have to do with the angle of view opening up to larger scales. For this reason, it is essential to establish coordination among a variety of different structures/

elements on all urban scales. A Nourishing-physical city depends upon a perceivable coherent morphology on all scales.

Positive connectivity with the immediate environment goes beyond biophilia. The best cases urge our connection viscerally, and we feel that we inhabit a healing environment. This positive connection occurs through geometry and coherence: it's what we experience in a great historical religious building. Spaces, overhangs, walls, and surfaces can be either psychologically hostile or welcoming; we could feel either threatened or safe from traffic or other pedestrians, etc. Any urban or architectural element that induces anxiety degrades our experience of the environment because we cannot suppress neurological signals of alarm. Consequently, we retreat from the public into a private realm.

Type 2. The Fractal city

An obsession with "design purity" and formalism removes the life-sustaining qualities of the built environment.

A "fractal" reveals complex structure at every scale of magnification. Examples in nature include a cauliflower, a fern leaf, and the mammalian lung, whereas artificial fractals are found in computer graphics. The point to note is that even artificial fractals look very "natural". In addition to showing a hierarchy of complex structure at all different levels of scale, coherent fractals exhibit self-similarity: any portion magnified by a fixed scaling factor will have similarities with the unmagnified structure. Pure mathematical fractals look identical when magnified by the scaling factor specific to that fractal. Cities were first interpreted as fractals by Michael Batty and his research associates [8; 36: Chapter 6].

Fractal structures depart from smoothness either by perforation, or by folding and accretion (two very different mechanisms). Perforation creates walls that have holes in them for doors and windows. The anti-fractal smooth, sheer wall is pierced, even if it is a glass curtain wall. Living urban interfaces were very often perforated, blocking vehicles while allowing pedestrians to flow freely (e.g. bollards, arcades, etc.). Perforation allows pedestrian flows to diffuse across such a semi-permeable barrier, like organic membranes that stop one element while letting another one pass through. A fractal also arises from folding a wall into meandering curves or extruded corners. Accretions add smaller scales to the fractal. Such urban boundaries with cusps for activity pockets abound in historical cities.

Living cities are fractal because mechanisms exist and work on every scale, from 1 cm to 10 km. This makes sense when we consider how human activity takes place on many different levels of scale. Our intimate living spaces contain the scales 15 cm to 3 m, whereas our tactile and visual senses look for ordered structure to define living environments below that: 1 mm to 15 cm. Those are the scales of traditional ornament and patterns found in natural materials such as wood and stone. On the urban scale, distinct human activities create distinct and well-defined urban structures to accommodate them, ranging from 3 m up to the size of the entire city. Elements of the urban fabric containing ornamental, architectural, and urban scales must be individually created, then made coherent with each other. The process of creating coherence coordinates flows and organizes urban space and its boundaries to support those flows.

The notion of fractal connectivity has philosophical implications. The large-scale universe is known to be connected to its small-scale details. Physics acts by recursion across scales. Fractal connectivity occurs in nature and in traditional and vernacular architectures, which have semi-permeable boundaries within semi-permeable boundaries (i.e. nothing is abruptly separated, and structure exists on all scales). But fractal connectivity is absent from industrial 20th century architectural and urban forms. In a profound sense, those 20th century non-fractal structures are anti-natural, and are perceived that way (see my book with Michael Mehaffy, *Design for a Living Planet* [31]).

An essential geometrical feature of a fractal is its inverse-power distribution of sizes: it has only a few large pieces, several of intermediate size, but very many small ones. The smaller the pieces, the more of them there are. The universal power-law distribution requires that "the number of components in a system is inversely proportional to their size". This distribution of sizes is obeyed by most natural systems (e.g. DNA, lungs, blood vessels, nerves, etc.), as well as by complex artificial systems (e.g. the world-wide web, electrical power grids, etc.). The abstract design process that gave rise to over-scaled repetitive urban and architectural forms since the Second World War is clearly the geometrical opposite. The repetitive forms of post-war design are fixed at a single scale, eliminating the human spectrum of smaller scales in the fractal distribution.

Traditional urban fabric is fractal because it shows variation and structure on many smaller scales: it is perforated; it folds and interweaves; it has extrusions and encloses courtyards; it is permeable through arcades; the roads have different capacity; buildings have openings and windows and tectonic subdivisions, etc. Eliminating the spatial fractality characteristic of living cities was done for an imagined "industrial efficiency". For example, strictly rectangular block buildings without smaller-scale subdivisions are not fractal. A grid of same-size roads without lower-scale branching is not fractal.

People who build for themselves will naturally create fractal expressions for the simple reason that human creative thought works simultaneously on many different scales. Human physiology is essentially fractal (see *Pavements as Embodiments of Meaning for a Fractal Mind*, co-authored with Terry M. Mikiten and Hing-Sing Yu, Chapter 7 in my book *A Theory of Architecture* [27]). Self-built or informal settlements have a fractal structure, but only for the small and intermediate scales. What is missing from most informal urban settlements are the larger scales. Traditional cities include those because they were designed with interventions by a higher organizational authority. But when a city is entirely formed by bottom-up forces, the larger organizational scales are missing.

Fractal structure occurs not only in the spatial, but also in the temporal dimension: a city works on time scales of 1 sec to 10 years and more (e.g. the center of Rome). At least it should, because human activities encompass the entire spectrum of time periods. Urban spatial and temporal fractalities mimic natural and biological structures and rhythms. But this essential quality was disrupted with the coming of industrial-modernist urbanism. The imposition of large-scale forms and single-use (mono-functional) zoning squeezed the rich spectrum of human actions into rigid time periods, seriously limiting the users' quality of life. This also affects education and learning in an institutional model, because children are much more sensitive to the absence of fractality from their environment.

Type 3. The Network city

We need to reverse the conception of the city as a collection of forms, to determine instead how those forms encourage human movement.

The city is a connecting mechanism, exactly like an organism connects its internal parts. A city's original reason for existence was to enable intimate contact among persons at close distances. Life is defined by networks, not buildings, hence a living city works because of its many overlapping networks of flows. Unless pedestrian flows in the city are encouraged, and are protected by means of intelligent urban furniture and the careful design of crossover points, they will cease.

The connective network of a city consists of overlapping yet competing networks of distinct character: pedestrian, bicycle, local automobile and truck, faster through roads, long-distance vehicular, light rail, etc. Each of these networks has different strength and requires a distinct geometry, infrastructure, and topology. For example, the pedestrian network mandates straight-line connections, because humans naturally minimize bodily effort. Cars, on the other hand, can easily take a more circuitous route to their destination. Yet we see the reverse rules applied in 20th century planning: making roads wide and straight for cars, while forcing pedestrian paths to wind around features meant for the car city, or eliminating them altogether.

It's trivially easy just to plan straight roads that maximize vehicular speed, but much more difficult to design the complex interwoven connections of the pedestrian realm and its proper feeding by the road network. The design of interfaces ideally protects the weaker flow at nodes where distinct networks cross. Otherwise, weak flows are cut, leaving only the physically stronger flows. The same phenomenon occurs when increasing road width discourages small-scale and low-speed traffic (e.g. bicycles and slow-moving cars), and favors only faster and heavier vehicular traffic.

In places like Tokyo and New Delhi the pedestrian flow is not especially protected by urban design. But the population density is so large that pedestrian flow continues even though it is threatened by vehicular traffic. What is occurring is that the pedestrian flow network has comparable strength to the vehicular flow network, so the two exist in a competitive yet balanced equilibrium. True, in those settings, the pedestrian is exposed to an enormous amount of stress! But what kills the pedestrian network in US cities is the enormous distance between nodes, something that is actively planned for to accommodate only automobile traffic.

Urban planners cater to developer's needs (who in turn are driven by profit and competition) to do a simple formal design – a "one size fits all" lacking small-scale negotiations – that by its very nature negates the necessary complexity that develops naturally over time. If society does not oblige planners to implement both pedestrian and vehicular flows correctly, as components of a larger network, then planning professionals will simply ignore the difficult part of urban design. We have to approach the problem as distinct networks, with the added complication of ensuring their proper interface. Urban planning needs to solve far more complex problems than professionals are used to dealing with (see [34]).

A city planned uniquely for fast automotive traffic is not a living city. We experience the Network city in traditional settlements where flows occur spontaneously (at a temporal scale appropriate to humans); flows compete and overlap at different points in the urban fabric, yet manage to co-exist over large portions without destroying each other. It is essential to protect the weak flows (i.e. pedestrians) from stronger flows (i.e. automobiles and trucks) by means of intentional structures such as bollards, raised pavements, arcades and colonnades; all elements of traditional urban structure that were eliminated by industrial modernism.

There has been important pioneering work on understanding cities in terms of their flows and networks, instead of their buildings and physical massing by Christopher Alexander [1], Jan Gehl [11], and Kevin Lynch [18]. My own work follows from that [36, Chapter 1: *Theory of the Urban Web*]. This network approach forces us to revise long-standing misconceptions about urban causality. Namely, erecting a massive building will not necessarily make it useful, nor will it automatically connect it to the surrounding urban life. That will depend almost exclusively upon the networks the new building sets up, and whether those succeed in connecting to existing networks.

The same reasoning applies to newly-created urban spaces. So many cities fell into the trap of commissioning an urban space, which was approved on the basis of an image or drawing (i.e. a formal rendering). While such an urban space may be visually attractive, it will actually be used only if it creates a complex network. Urban flows, in turn, depend upon several critical factors: a rather wide swath of the surrounding region; whether flows already exist there; and if the new insertion is capable of connecting to them and channeling them through the urban space (The Network city is treated in [36, Chapter 1: *Theory of the Urban Web"* & Chapter 6: *Connecting the Fractal City*].

Type 4. The Spontaneous self-built city

The human brain has intrinsic design skills that were applied throughout history to build cities, without the need for architects or planners.

The world is facing a population problem where over 1 billion (out of 4 billion) city-dwellers live in slum conditions. The spontaneous forms erected by self-built settlements, unfettered by formal design, mimic biological growth. This phenomenon is positive as far as including Biophilia into the city structure, but self-built urbanism could have serious problems (e.g. the favelas in Rio de Janeiro; Dharavi in Mumbai). A single family builds its own dwelling, where the people use available, local, and scrap materials. Those materials tend to be manufactured rather than natural, and are not naturally adapted to impromptu building; they might collapse with severe weather conditions. Yet here is certainly the most efficient manifestation of sustainability. Energy use is minimal, for the simple reason that none is available, while centrally-generated energy sources are pirated to support life in the informal settlements.

Vernacular or indigenous settlements house most of the world's population, but only a fraction of them are "slums". They are however despised by academic architectural culture. Genuine slums do have problems. In addition to often-terrible living conditions because of poverty and the lack of health and other essential facilities, there is usually

no infrastructure; there is minimal access and poor connectivity because of the absence of transverse large-scale roads. This last point has to do with incomplete fractality: informal settlements are built by the family, not by any higher organized entity, and therefore lack ordering on the larger scale. This turns into a major deficiency, because network connectivity is compromised.

Another problem is that there is often no land ownership in illegal settlements. This legal obstacle prevents a family from upgrading its house, as was always done during the historical evolution of the city, because it's not worth investing effort in something that can be taken away at any moment. The slum thus remains perpetually in a decayed state, with no incentive at upgrading by its residents because they are not "owners". This is what happens in places like Brazil. Contrast this with informal settlements in countries where land deeds have been awarded by the government: this occurs with the informal settlements in Turkey (in expectation of votes, of course!). We observe a slow but striking evolution towards better-constructed and more permanent urban fabric.

The process of upgrading buildings over time eventually generated historical cities as we know them. Many if not most settlements began informally in this manner. Another set of cities was initially centrally planned as military or colonial settlements: Roman camps that later evolved into European cities, and the Laws of the Indies used by the Spanish to found numerous cities in Latin America. Those are still recognizable today with a more-or-less rectangular Hippodamean grid at their core. A few cities remain formally planned since their founding.

If the regulatory/financial power that shapes today's cities understood the process generating the Spontaneous Self-built city, then we could solve one of the pressing questions of humankind. All over the world, informal settlements are growing out of control. This is a natural process, as much as centralized government is threatened by it. The solution proposed here is to "go with the flow" and insert infrastructure to create an informal city with an acceptable quality of life. This can only be done by acknowledging the forces driving the Spontaneous Self-built city, working with it and not trying to destroy it and replace it with the industrial-modernist model (My book *P2P Urbanism* [37] develops strategies for upgrading informal settlements).

Type 5. The Virtual city

Virtual connectivity gives support for the nourishing old-fashioned city, and makes the "futuristic-looking" city irrelevant.

Information and communications technologies enable global connectivity. A person can live anywhere, and is able to connect to the world. To inhabit the built environment today means living part of one's life inside the Virtual city. Nevertheless, life in the Virtual city is detached from urban geometry. This separation would appear to validate industrial-modernist living environments – such as apartments in high-rises on the periphery. The virtual world accessible through a screen was initially constructed from metaphors taken from the physical world, although it has its own framework and special rules. This transfer of physical into virtual worlds faces the major danger of bringing about techno-seclusion, a problem that is a consequence of design.

Many people with laptops work in emotionally-comfortable public environments. People choose (from among an infinite number of alternatives) those places having the human qualities of traditional urban fabric, and feel the need to be in close proximity and to share physical space with other humans (i.e. in cafés). This is the primal force of urbanization, and those frequented places are decidedly urban. Anyone in that situation can decide where to work from, yet they crave biophilic, architectural, and urban qualities that we identify with traditional city places.

Our built environment should have the goal of emphasizing essential human qualities through positive visceral emotions. Otherwise, people will conveniently withdraw to the emotionally-welcoming environment of the Virtual city, leaving an emotionally-dead industrial-modernist city (Inhuman city). The Virtual city incorporates "human" characteristics as a main reason for why it works (and also abuses them). Abandoning the Nourishing-physical city opens up the potential for human abuse within the Virtual city. Along with all the wonderful opportunities it affords, the Virtual city also includes a dystopia of manipulative social media, a vicious jungle full of ravenous predators. Anyone can pretend to be what they are not, bypassing our built-in system of verification that works only when we encounter strangers face-to-face in the physical city.

Why do the media enthusiastically promote fashionable design trends linked together with the virtual city? They are not related. We are masters of technology and can often work from our laptop in any physical setting we find. Our body reacts negatively to psychologically hostile places, regardless of all the media hype about their famous architect. We avoid them. Current societal trends are indisputably hostile to the millennial urban environment. Few people seem to have noticed this apocalyptic danger threatening the physical city.

Type 6. The Developer city

The marketplace has failed, for the most part, to select life-enhancing architectural and urban typologies.

This city type is not based on morphology, but includes the forces that drive urban construction. In many places around the world, the government or commercial developers predominantly decide what gets built. (I am excluding many private residences.) The implemented models could be entirely arbitrary: the principal consideration is that they were built previously and can therefore be copied. It is enough that they make a lot of money for the developers and construction industry. Developers do analyze what works or not, otherwise their work is less successful, hence less profitable. If the public, through ignorance or manipulation, accepts minimally-satisfactory typologies, then that's what developers build as the easiest solution.

The business of construction is an economic engine for industry. Building on speculation, not immediate needs, makes enormous profits through marketing. The formula is the following:

1. Find a cheap method of building that is efficient for the supporting industries of construction, finance, regulations, and permitting;

2. Make a deal with the government authorities to allow such construction without examining the human consequences. Developers don't necessarily get away with this, but do so often enough to make the ploy worthwhile;
3. Construct as a speculative financial venture;
4. Use advertizing to sell the units, or sell the entire building wholesale to the Government and let it worry about occupancy issues.

This model makes social housing into an excellent profit scheme for the developer (but often creates a nightmare world for the eventual residents).

There are various types of developers implementing different models. One builds suburban subdivisions at the city's edge on cheaper agricultural land. This is assembly-line construction, built on speculation. It is cost-effective because each step of the process is well-known. The permitting process involving identical units is highly simplified, hence advantageous. There are no potential surprises because the developer thinks that the product is what the market wants. In fact, this is not the case. The developer, by insisting on standardization, shapes the market by offering only the same type of suburban house. Even if there are several developers, all of them tend to build the same thing; hence there is really no market choice.

Another model followed on already built (and possibly degraded) sites is to destroy and replace, instead of repairing existing urban fabric. Vastly more profit is made from complete re-building rather than upgrading existing structures. But at the same time, with developer-driven speculative and government construction done in the cheapest possible manner, there often remains no alternative to tearing down existing structures completely before rebuilding. This practice perpetuates a hugely unsustainable model. The model turns destructive when it is applied to demolish solid buildings from the past.

When you tear down pre-industrial homes you lose a greater variety of scales. Most new residential construction uses modular paradigms, which forces uniform sizes everywhere. Trying to save money, the psychologically comfortable 11 and 12 foot (3.35 to 3.66 meter) ceilings and unusual, attractive, and restful spaces in earlier homes are forgotten; a thing of the past. Low-ceilinged contemporary homes are oddly scaled next to human beings, and compared to the older homes that adapted to human spatial sensibilities. The standardized rectangular footprint in new tract houses also obsessively conforms to a simplistic modular form.

The user feels the physiological and psychological effects of the Developer city, which range widely. Developers are not primarily concerned with adaptation to human needs, but only to make the largest profit in the shortest possible time. Developers can employ healthy and tested typologies, but also discredited inhuman ones. It's no surprise why developers build an anti-fractal city (in both the city center and suburban sprawl); that is due to ignorance. The media and academia worsen the situation by perversely condemning fractal urban models while praising the anti-fractal ones.

Type 7. The Anti-network city

Ideology and special interests killed an essential dimension of the living city by eliminating pedestrian connections.

Are there eight or six city types? The reality of type plus anti-type comes from elementary particle physics, where we observe particle and antiparticle pairs having opposite properties. Because this concept is obviously not standard in urban thinking, I have included the anti-types as separate entries in the eight-fold classification. Yet logically and practically, it suffices to know the characteristics of a particular city type, then to do the contrary in order to generate the anti-type. Hence the two pairs of annihilating opposites: Network/Anti-network cities, and Nourishing-physical/Inhuman cities.

Two city types describe deficient complex systems that negate essential characteristics. Understanding why this negative action happens is essential to repairing existing urban fabric. People need to physically connect to urban nodes, those places where a critical threshold of pathways cross, and where points of interchange are situated. Nodes include work, school, food stores, retail, administration, and church; everywhere a person needs to be physically present in daily life. In the Network city, the nodes are placed so as to be physically accessible, and the connective networks link all these nodes together in a way that makes life more a pleasure than a burden. Design of new urban fabric, or the repair of existing one, starts with the nodes and then facilitates path formation.

The Anti-network city situates essential nodes outside the range of pedestrian movement. With one stroke, a vast number of connections in the Network city are severed (because nobody can walk to a large number of destinations), thus making access problematic. In order for it to function at all, the Anti-network city consumes enormous amounts of fossil fuel. Why was a deliberate and huge step backwards in urban connectivity implemented by so many cities following World War II? The answer is known and documented. Planning laws were re-written under pressure from automobile and petroleum companies. They aimed to privilege the automobile for all daily movement, and post-war planning succeeded in that.

The apparently innocuous idea of introducing the automobile to the dense city (so convenient for our daily lives) back in the 1920s, momentarily displaced us but permanently changed the very structure of the urban fabric to privilege vehicles, which then permanently displaced us. This is actually the main cause of many issues explained in this paper.

Here are the origins of the four-way conflict between opposite pairs of urban structural qualities:

1. Pedestrians versus motor vehicles;
1. The conflict between human-scale versus inhuman-scale urban fabric;
3. Intricate urban fabric adapting to pedestrian movement versus monolithic designs adapting to vehicles;
4. Fractal urban fabric with superimposed network versus non-fractal urban fabric that is anti-network.

The Anti-network city fundamentally negates what the Network city achieves, in terms of human wellbeing. Still, the Anti-network city represents the evolution of a distinct city type

as the result of a sequence of key decisions taken by citizens, planners, and politicians. It did not arise randomly, but was carefully planned for one particular benefit that most everyone wanted at the time.

Type 8. The Inhuman city

Design based on machine fetish and a fanatical hatred of the past created many of the standard typologies in use today.

I come to the most problematic of the eight city types: the Inhuman city. Could architects and planners deliberately design a type of city that is hostile to human beings? Unfortunately, yes. To understand this, it is helpful to consider the social narrative that the contemporary world is built upon.

In the 1920s, it was assumed without reflection that the "city of the future" bore absolutely no resemblance to the traditional city. The world's majority housing stock of vernacular/indigenous buildings was summarily dismissed as inappropriate for our times. The problem is that this way of thinking is never diagnosed as pathological. Consequently, a set of typologies whose principal objective was to replace and undo traditional ones is endlessly praised and replicated. Protests from select individuals who perceive these industrial-modernist environments to be inhuman are countered by the same propagandistic message of "progress", "innovation", and "sacrifices for a better future" that drove anti-traditional typologies to be implemented in the first place.

It is useful to consider an analogous biological condition to describe this phenomenon. In an autoimmune disease, the body's system for attacking invading pathogens turns instead to destroying the healthy body. Even though it evolved to protect the body from foreign invaders, our immune system can be tricked into turning against the body itself. A similar thing happened to cities.

Design rules adopted after World-War II had a drastic effect on the quality of life in terms of what could be absorbed from the environment (and distinct from amenities available through technology). The living urban environment found in traditional cities (those without major problems) was no longer identified as the "body" of the city that we desired, but as something alien that must be removed. This unfortunate switching came about from an ideological-political dismissal of the past and everything associated with it (after World War I).

Readers from outside the discipline might think that this is a crazy situation. But those from within the discipline will recognize the obstacles to changing the *status quo* that supports rows of repeated concrete block housing, ill-defined windswept open space, glass skyscrapers in all climates, cookie-cutter dormitory sprawl houses, suburban subdivisions accessible only via a single collector road, etc. Those typologies are deeply ingrained in architectural culture and are supported by an explanatory framework (The ideology behind the Inhuman City is discussed in Chapter 9: *Geometrical Fundamentalism*, written with Michael Mehaffy, of my book *A Theory of Architecture* [27]).

The actual city we live in

People can liberate their deepest feelings to sense a city, and justify those visceral responses using the eight-fold model.

By revealing many of the mechanisms that govern urban form, the eight-fold model provides a handy diagnostic tool. It helps us to predict how well distinct city regions work, based on which city types are found there. We need this knowledge for deciding what should be added or changed in order to make a place function better. It also identifies what to preserve from an otherwise mistaken upgrading. Without this understanding, disasters occur through urban interventions that make no sense except as a superficial visual design. So many radical re-structurings destroy (out of ignorance of urban processes) what is already there and is working perfectly.

We live in a wide variety of cities, each of which contains a different mix of city types. The city type mixtures create a visible recognizable morphology in different neighborhoods. Distinct combinations of the eight city types have drastically different consequences for the quality of city life.

Experiencing a particular piece of urban fabric, an observer asks: "what qualities are present to make this a lively city?" Or, conversely, "what qualities are present, and which ones are lacking so that I experience this city as lacking human life (not to be confused with impressive buildings)?" One can look for spatial structure on different scales (Fractal city); temporal activities of different periods that are encouraged by the urban morphology (Fractal city); the presence of plants and other human beings (Biophilia/Nourishing-physical city) and ornamentation and the use of natural materials that produce the same positive connective effect (Biophilia/Nourishing-physical city); the freedom of choice to move about using a variety of transport channels: pedestrian, public transport, private vehicles (Network city), and so on.

The opposite situation – when we sense the city fabric to be without life – also lends itself to analysis in terms of the eight city types. We look for the causes of unease and see sheer, smooth, and forbidding walls, either windowless concrete, metal, or curtain-wall glass (Inhuman city); the elimination of fractal qualities and the removal of color (Inhuman city); highly restricted possibilities of movement (Anti-network city); giant buildings (monolithic office towers that occupy an entire city block) or urban plazas that feel oppressive to those who experience them (Developer city); housing blocks that no child feels love for and cannot find a spot with psychological comfort to play in (Developer city), etc. As explained later, the Developer city could choose to add with the Nourishing-physical city by following rules for creating parks, pedestrian trails, playgrounds, and urban spaces.

This reasoning raises an alarm about introducing abstract sculptural or oversize buildings without biophilic, fractal, and symmetry qualities into the urban fabric, something that is welcomed enthusiastically by today's architectural culture. But even a tiny injection of pathological city types can ruin a living city. The vast majority of world architecture up until our times had biophilic qualities (combining fractals, symmetries, scaling, geometrical coherence, etc.). Deliberately opposing precedent, more recent buildings mimic the formal (i.e. image-based) aspects of non-living environments such as wind-sculpted rocks, blocks of ice, which have nothing in common with living structures.

Part 2: City type interactions

Diagnosis followed by healing interventions

The eight-fold model is used to diagnose morphological urban pathologies, and offers the framework for finding the appropriate healing treatment

We can follow steps to diagnose and repair diseased urban fabric. The value of the eight-fold model lies in defining the distinct city types so that they – or their absence – can be recognized in an actual city. Unhealthy urban fabric can be diagnosed to determine whether and which unhealthy city type is present. If it is, then long-term planning steps have to be set up to eventually remove it. The corollary, diagnosing the absence of a healthy city type, prompts a similar planning effort to implement mechanisms that will build up the missing city type over time.

At the same time, many of the urban pathologies we face today are the result of combining political ideology with commercial forces. Present-day planning practice has picked up unhealthy and unreasonable taboos that act against the life of a city. As a consequence, some healthy city types are banned, and are thus not found as integral components of the contemporary city. I suggest how to add those back to create a living environment. Once explained in these clear terms, the solution is easy to recognize and implement; yet we are fighting an ignorance of the mechanisms responsible for urban morphology that led to those misunderstandings.

Of special interest is discovering that adding certain city types in combinations that were long forbidden (e.g. Nourishing-physical + Spontaneous Self-built city) helps us to create a healing environment. Digging into the original reasoning behind such condemnations exposes their poor logic, and opens the door to revision. Many healthy combinations of city types in the eight-fold model are now ignored and forgotten, because a simplistic non-organic model of the city has been applied for so long as to become standard.

The complexity of a model for cities cannot be too low. Simplistic (one-dimensional) models are both useless and dangerous. According to Ross Ashby's "Law of Requisite Variety", a model of a complex system itself has to have some minimum threshold of complexity (see my article: *The Law of Requisite Variety and the Built Environment* [38]). Take the analogy with a map of a territory. Maps are necessary for representation and navigation, but too simple a map misses many crucial features. On the other hand, a map that includes too much information is also useless because it is too complicated. Jorge Luis Borges described the extreme case in his essay *On exactitude in science* (reproduced in my book with Michael Mehaffy, *Design for a Living Planet* [31]).

When communicating to decision-makers and politicians, however, an urbanist has to simplify the message drastically. Politicians may be intelligent, but they can have competing interests, and their lexicon is quite different from that of an architect or urbanist. Therefore, advice must be offered to decision-makers in simple and compact form. The eight-fold model makes the components of a living city easy to explain. Urbanists must constantly monitor what politicians are actually implementing, to correct for mistakes due to short-term thinking.

MUTUAL ANNIHILATION

A healthy future for humankind can only come about by stopping design ideology and special interests from destroying living environments.

Certain combinations of city types generate a living city, whereas other combinations define a sick city. Interactions among city types are like mixing cooking ingredients. Some combinations enhance each other, whereas others don't match or instead cancel each other. Compatible city types add and overlap to create living urban fabric. The measure by which to judge the result is how well users are able to conduct their daily tasks in a pleasant and especially a healing environment. This is not an aesthetic or intellectual judgment – a sick city wastes enormous quantities of energy to run, it spoils people's everyday life, and is worst for children.

Juxtaposition and mutual annihilation depend upon local circumstances. In practice, opposite city types can co-exist next to each other in close proximity. But obviously they can never support each other. The forces that make conflicting city types work, and the forces they themselves generate, will always clash on the ground and thus strain the functionality and geometry of the city. This observation does not refer to a healthy diversity: that's due to entirely different yet complementary urban elements coming together. Here, instead, we have strict opposition and not union.

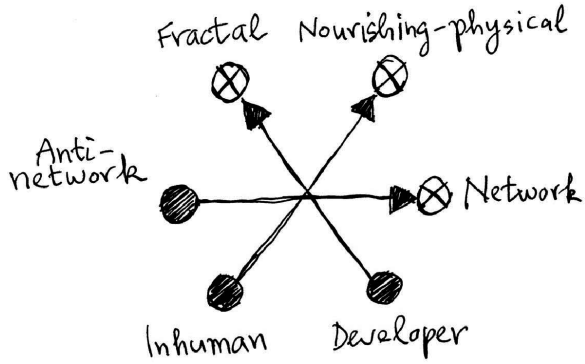


Fig. 2. Opposite city types annihilate

The idea of two city types annihilating each other helps to explain some dynamics of urban morphology. Forces that generate mutually incompatible city types cannot act on the same point: either one city type or its opposite will be built there. No overlapping mixture or compromise solution is then possible. The eight-fold model explains why dramatic discontinuities are generated, such as slums adjoining exclusive high-rise communities separated by a high fence. There is complete urban contrast on either side of the separating boundary.

Adjoining but incompatible city types create an urban discontinuity. For example, the Spontaneous Self-built city is able to co-exist next to upper-class high-rises. An

impenetrable border arises as a natural insulating phenomenon. Morphological urban equilibrium is achieved only because of this urban boundary. Tearing down the wall destroys one city type – the Spontaneous Self-built city – to replace it with the Developer city. The developer colludes with the state to forcibly evict the residents from the slum, and then erects more luxury high-rises on the vacated land. The alternative can also happen after the wall is taken down: local crime invades the high-rises, forcing the upper and middle classes to move out. The Spontaneous Self-built city takes over the Developer city.

Opposing forces acting within the regions from the two sides protect the distinct city types from each other. The impenetrable wall is simply the geometrical expression of where opposing forces meet in a balanced tension, like the militarized border between two hostile countries. Government and powerful commercial forces that run the "official" city oppose the small businesses and organized crime that run the informal settlement. The border delineates where one set of forces transitions into the other. As long as these two sets of forces have comparable strength, the border remains more-or-less stable. Otherwise, one of the city types takes over and displaces the other.

This phenomenon of encroachment and annihilation extends to contrasted situations other than the forced demolition of slums. Driven by greed, perfectly functional traditional neighborhoods the world over have been invaded and destroyed. A Nourishing-physical city is replaced by the Developer city employing industrial-modernist typologies. The replacement looks, feels, and works very differently from what it replaced. Most often, the capacity for urban life is drastically reduced. The victim settlement could not garner enough political support to oppose the combined power of the developer acting with local government against the current residents' interests.

Cooperation and addition

Addition is a basic operation defined only between objects with common qualities; otherwise things remain separated.

The interactions among city types can be of several kinds:

1. True addition – superposition: every building, path, and portion of urban space add to create a harmonious whole;
2. Coexistence – juxtaposition: several buildings that, because of their geometry and floor plan, relate minimally to one-another;
3. No interaction – isolation: a building disconnected and unrelated to the urban fabric it is situated in;
4. Destruction – mutual annihilation: one building, small or large, is sufficiently non-adaptive to make the surrounding urban space hostile and unpleasant to use.

Looking at the city as an organism clarifies its functioning subsystems. Several independent structural frameworks – the eight city types – have to cooperate, balancing their mutual competition for resources (such as ground space). When all the subsystems act in unison, they link and depend upon each other. Each distinct city type contributes a complementary

functionality to the city. Without the healthy city types present and interacting positively, not only do we experience a city as severely limited, but it has sustainability problems as well.

There are deep mathematical reasons why some architectural and urban elements can or cannot combine to provide a better urban function and experience. The general rule for the addition of city types is: "Only elements that share systemic complexity with each other can join their systems together into a larger whole". (An analogy is Velcro that will not attach to a smooth surface). Therefore, the process of addition – in what way they come together – is just as important as the city types present on the ground, because they can add, or simply juxtapose, or even annihilate each other.

Healthy city types

Good city design has to apply only city types that create a healthy human environment, and not their opposites.

It is worthwhile listing those city types that contribute more-or-less to a healthy city for the majority of users.

TABLE: FIVE HEALTHY CITY TYPES
NOURISHING-PHYSICAL CITY
NETWORK CITY
FRACTAL CITY
SPONTANEOUS SELF-BUILT CITY – only in part
VIRTUAL CITY – only under certain circumstances

These city types link strongly to natural and biological structure, and are thus better for us (by acting positively on our physiology and psychology). A city combining them correctly is perceived as being "alive". Structural aspects of adaptive cities evolved through human beings seeking optimal feedback during normal use. An important consequence of this common evolved relationship to adaptive geometry is that healthy city types can also link to each other. Once we understand this additive process, then we can dispense with the types altogether, and understand the city directly in terms of its living structure [3].

For millennia, human beings used their innate sense of form-generation to construct their environment. New practical developments of city form such as introduced by new materials and evolving industrial needs, but which conflicted with biological structure, were never allowed to take over. City formation thus followed compatible biological rules up until the advent of industrial modernism. Informal settlements continue this process today. Intelligent urbanists therefore study both marginal and traditional settlements for insight into the positive effects of adaptive, self-adjusting, and spontaneous generative processes.

Unfortunately, the five healthy city types listed above are misinterpreted because they necessarily "look" traditional. Ever since architectural culture became image-based (the defining characteristic of industrial modernism), designers have been strictly taught to

judge all forms of design by how closely they resembled images from the industrial-modernist canon. With an obsessive focus on the visual appearance of supposed "innovation", modernist design practitioners naturally develop a built-in aversion to anything remotely resembling traditional urban fabric. Healthy city types are avoided as a matter of prior conditioning, without anybody being aware that this is occurring. This terrible prejudice has evolved and softened over time, so that neo-traditional architects and planners today create wonderful human environments. But their work often goes unrecognized by the mainstream.

The Virtual city can add nicely to the Nourishing-physical, Fractal, and Network cities to help with flow. But it could also detach us from the physical city. The Virtual city can fit into either a healthy or an unhealthy city, so its presence cannot be used as the basis of judgment. This important point is discussed in a separate section, below.

How adaptive city geometry cradles life

The good city types add nicely because they share common qualities that contribute to human health and wellbeing.

The term "nourishing" means that the immediate environment can provide nourishment to our mind and body through information and sensations. This largely subconscious process comes from emotions triggered by coordinated function and geometry, and is just as necessary for our daily life as food and vitamins. Its absence, on the other hand, is believed to lead to stress and eventually to long-term degradation of our health and wellbeing.

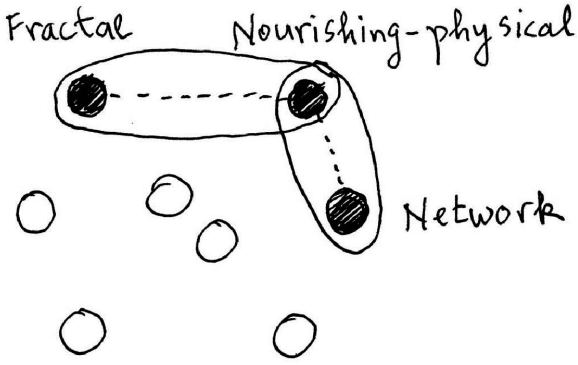


Fig. 3. Adding healthy city types

Here is a brief summary of how healthy city types combine. The Nourishing-physical city includes characteristics of the Fractal + Network city, and adds its own spatial properties. These three city types combine into a successful city, in which they exist as overlays in the same space and time. The Fractal city makes possible a nourishing biological response to architectural and urban forms that trigger our biological memory. Coherent urban combinations in the spatial layout are crucial, yet functioning urban fabric is also dynamic.

Ease of movement comes from an extensive complex web of connections on all scales that is the Network city (see my chapter: *Beauty, Life and the Geometry of the Environment* [28]).

A city that is loved by its inhabitants combines healthy city types, and provides the setting to enjoy the different stages of life and raise one's children. The resulting geometry could have an infinite variety of specific implementations, all of which have the framework for nurturing living structure. In the best cases – the ones most loved by people, not necessarily by architects – the geometry cradles and supports everyday human functions. It is, of course, possible to live in a city made of architects' dreams, which could be exciting, but the experience would be abstract or one-dimensional, and could over time leave the inhabitants unhealthy and unsettled.

Part 3: Rules for urban space

Why design "hard" plazas?

Recent urban spaces that feel hostile and unwelcoming express political anger that turns against their prospective users.

Design rules for creating usable and welcoming urban spaces can be learned, with some effort, from studying historical urban spaces that still work to attract users. Many good examples from around the world are full of people during many hours of the day. What happened is that the opposite design rules have been consistently implemented in post-war planning. But those rules, being more about the ideology of modernization than about human wellbeing, are prescriptions for keeping people away! A large number of urban plazas remain empty (except for stray dogs and vagrants), even those awarded with architectural prizes.

Nowadays, urban planners are not even aware of what urban space really means, and why it is such an essential ingredient of a living city. Nor is anyone else aware of this in the chain of the regulatory system that oversees urban interventions. Rules for getting successful urban spaces built, and for regulating them, must be conveyed to those public sector workers who are responsible for them.

Pioneering work to determine which urban squares are actually used was performed by Christopher Alexander [4] and by William Whyte [23]. Jane Jacobs described the spatial complexity of the living city [12]. Several urbanists are now beginning to implement the true principles of urban space design and function. This, despite the mainstream insistence on "design-by-image", which doesn't bother to understand the socio-geometric forces that give rise to form, nor to work out the forces that the structure will generate if built in a different location.

Because so many of today's designers tend to be exclusively visual, they miss all the other factors. It is essential to understand how good urban fabric reveals itself from human movement and reactions. There is a tremendous complexity to the emotional and perceptive processes that guarantee the use of urban space.

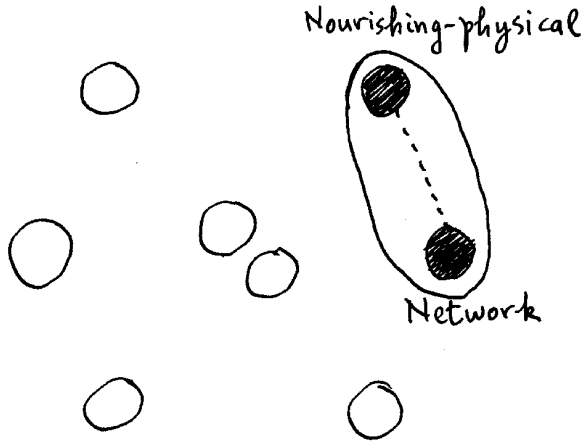


Fig. 4. An urban space is a key node in the Network city, but its functionality depends upon it also belonging to the Nourishing-physical city

A new plaza inserted into an older living city can continue to be fed by the existing Network city. But inserted into a new city, it's dead space. Why? Urban space requires the sum Nourishing-physical + Network city in order to work, otherwise people will stay away (see [32, 36]). It must be realized that only evolved socio-geometric design solutions can guarantee the attractiveness and use of urban space. Any new construction that is conceived in isolation – as a stand-alone item – has not evolved, and consequently cannot effectively marry to existing flows. Rejecting traditional patterns of human use only leads to a sterile environment. Barcelona made this mistake in letting design ideology create many "hard" plazas that do not collect paths.

As is detailed below, a "hard" plaza can work when it is a transit space, i.e. just another very wide pedestrian street. This presupposes attractive pedestrian destinations all around the plaza, so that paths essentially cut across the plaza as a convenience. Piazza San Marco in Venice is of this category. Because of its size, Piazza Navona in Rome is mostly a transit space, yet it also includes attractive destinations with its three fountains. You don't need to add anything else. But urbanists can destroy a transit plaza by inserting useless obstacles in an effort to make it "interesting". Such accoutrements include an abstract sculpture, or pools of water placed unintelligently so that they cut across the pedestrian paths.

A "soft" urban plaza, such as the older La Rambla strip in Barcelona, is characterized by fractal qualities, biophilia, connective networks, etc. There are bushes, trees, old-fashioned benches, lamps with detail, and other street furniture, umbrellas and canopies, ornamented 19th Century kiosks, so that the ensemble is highly fractal and biophilic. The pavement is varied, and the biophilic effect is multiplied several-fold by the flowers and fruits presented for sale. This is not merely a romantic idea or pretty picture; it is an essential enhancement of the quality of place through biophilia and the fractal hierarchy of scales. Most important, La Rambla is "fed" by dense pedestrian urban fabric along both sides.

A "hard" plaza, by contrast, consists of a plain slab pavement, strict rectangular geometry, no trees, no kiosks, and is either starkly empty, or it may contain abstract sculptures, severe and uncomfortable "design" benches, and lamps boasting an industrial-minimalist look. Those add absolutely no biophilic qualities to the experienced space. New plazas also tend to be situated in the wrong places, so that the existing path structure does not serve to feed users into and across the space. Those plazas are conceived as giant abstract sculptures themselves, obeying no design rules other their architect's personal artistic whim.

Why are Barcelona's new plazas uncompromisingly "hard"? My friends from Barcelona explain that those designs expressed pent-up sentiments that were freed by the ending of the Franco dictatorship. Socio-political forces included frustration, reaction to oppression, the urge to provide public platforms for expressing the new freedom, etc. They assure me that today, with a totally changed socio-political dynamic, those plazas would be designed and built within a much "softer" typology. Even so, nobody has yet thought to upgrade those unused plazas using traditional solutions to create a more human environment. The "hardness" of the plazas has been forever linked to the political sentiments of that instant in time, and to revise the geometry would be seen as rejecting the historical change – a ridiculous notion, yet deeply felt.

Creating attractive urban space

An attractive urban space envelops its users and provides a feeling of reassurance while being there.

The criterion for success relies on observations of use over time. Christopher Alexander and co-workers extracted socio-geometric design "Patterns" from the best – most human – architectural and urban environments, which are central to this investigation (some design patterns are listed later in this paper) [4]. Those Patterns can then be applied to design new environments, and to diagnose and repair urban spaces that repel rather than attract users (For a summary of the design rules, see *Geometry and Life of Urban Space*, written with Pietro Pagliardini [32]; and also Volume 3 of Alexander's *The Nature of Order* [3]).

The main characteristic of successful, usable urban space is that it define a giant outdoor room open to the sky. It is necessary to surround the open space by welcoming façades, perforations and folding of the built fabric, and consumer activities. Users are attracted to the texture, tectonic balance, composition, color, and ornamentation of the building façades bounding an urban space. But contemporary building fronts that follow the industrial-minimalist aesthetic fail to provide this "welcoming" attraction. Without biophilic façades, even the best "designed" urban space will never attract users to linger in its interior.

A network of linked urban spaces is necessary for a city to be alive in the sense of encouraging positive human activity and interaction. Urban spaces define the nodes of the pedestrian network, and other transportation networks should add to (but not destroy) these principal channels of pedestrian circulation. Focusing on the life in open spaces is contrary to current architectural trends. Contemporary design focuses on "signature" buildings, which are

formally abstract, while leaving the adjoining/surrounding space to chance. That approach misunderstands (ignores) how living cities actually function. Unfortunately, these "stand-alone" buildings happen to have gained the center stage for the media and the public, at least for the time being.

The importance of building façades

The structures surrounding an urban space – both in their architecture and in their situation – are a major factor determining its use.

This is more an architectural question than an urban one. We build living urban fabric through architecture that surrounds and defines urban space, thus architecture and urbanism are inseparable. Geometrical coherence acts to channel flows on many different scales. The perception of the urban plaza as a harmonious whole depends very strongly on certain mathematical properties of the surrounding building façades (among other criteria, of course). Ordered complexity shown on a building's front is created by mimicking the structural rules giving rise to life forms, and thus expressive of life itself. Surrounding façades exhibit the following features:

TABLE: DESIGN ELEMENTS FOR BUILDINGS FRONTING URBAN SPACE.

1. Ordered structure on a hierarchy of decreasing scales, from the largest size down to the microstructure;
2. Sophisticated fractal patterns (patterns within patterns), including those generated by recursion and Cellular Automata;
3. Ordered complexity, in which many different patterns on smaller scales are coordinated through symmetries to produce a coherent whole [3, 27];
4. Scaling symmetry, where the different scales are related to each other by magnification (a characteristic of fractals);
5. Traditional patterns such as reflectional, translational, and rotational symmetries superimposed in a coherent manner;
6. The vertical symmetry axis emphasized, because our body evolved in gravity and connects psychologically to the vertical;
7. Avoidance of extensive horizontal or diagonal elements on buildings, since those give rise to feelings of tectonic instability, hence anxiety. Arches are fine, because they are reflectionally symmetric across a vertical axis;
8. The presence of color, both interesting in itself in every occurrence, but also obeying a large-scale color harmony. Colors reminiscent of death (e.g. grey concrete, black or brown surfaces) and colorless surfaces upon which the eye cannot focus (e.g. transparent or translucent glass curtain walls, reflective metal) are negative, whereas welcoming colors reminiscent of our natural environment, flowers, fruit (e.g. both rich and pastel colors that humans find psychologically nourishing) are positive.

These general criteria attract human beings to approach them, and to enjoy experiencing them subconsciously from every distance. Since our sensory system has evolved to cope with gravity, and is set up to recognize biological forms with vertical symmetry, skewed forms generate alarm and physiological distress. Unless there is a vertical axis of reflectional symmetry, a person could experience nausea caused by the inner ear's mechanism for vertical orientation. Any symmetry axis is fine on a floor pavement, but an explicit or implicit vertical axis on a façade or entrance is essential for sensing stability. Our reaction of alarm at unbalanced diagonal forms cannot be learned or changed.

Permeability to pedestrian flow

A functioning urban space is a node concentrating pedestrian paths from the surrounding region; otherwise, it is a giant sculpture only fit to look at.

The urban plaza needs to be highly permeable to pedestrian flow. Anything in the plaza that is likely to attract users is of secondary importance. Even a statue of General José Olivaro (Hero of the Revolution) is not enough! From the mechanism of biophilia, that's better than some abstract art or contemporary sculpture. Amusements and a play area do attract families with children. Trees and a shaded canopy serve those who wish to rest for a moment. The surrounding paths bring pedestrian flow to cross the plaza, and the street furniture ought to accommodate users who are channeled to walk to the plaza. Use depends critically upon the pedestrian activity in several surrounding blocks, as it depends equally upon the street and sidewalk design that permits easy pedestrian access to the plaza.

People will use a plaza situated in the Network city at a point where multiple flows cross. If it is the only open space within a large region, that will actually bring people to the plaza; but even then, a "hard" design and hostile urban furniture will drive people to stay out or detour around it. Dreary, unused contemporary plazas are observed the world over.

Through its placement and physical design, an urban space should encourage people in a hurry to cross it (3 min) instead of taking a parallel external path. This process corresponds to "catchment" of local pedestrian flow, diverting it to feed the plaza. While they are traversing the space, people's attention should be drawn momentarily yet repeatedly to architectural details in the surrounding façades (2 sec), and to possible greenery in the square. Other people must be attracted to stroll at a more leisurely pace (10 min), and some to sit down and relax (15–30 min). Families with young children should feel welcome to stay (30–60 min). The way of achieving this is through a complex adaptive design that accommodates all human spatio-temporal scales.

To guarantee the "feeding" of the urban space, mixed-use buildings three blocks deep surrounding the plaza have to supply potential users: this span correlates with a 5 minute walk. Some of those pedestrians will naturally walk by the plaza, and, if the environment and path structure are welcoming, people will choose to cross the urban space. Once there, a percentage of those users might decide to linger. There is a distribution of time periods for different users, or even for the same user on different occasions: to stay for

anywhere from 1 minute to 1 hour (I describe how the Network city feeds urban plazas in my chapter with Pietro Pagliardini and Sergio Porta: *Geospatial Analysis and Living Urban Geometry* [33]).

The urban space is protected from encroachment by parked cars and vehicular traffic. Utilize wide and raised sidewalks, arcades, bollards, etc. to protect the pedestrian, direct the traffic, and keep cars outside the pedestrian realm. We could provide tangential vehicular flow to "feed" the plaza, but at the same time make it impossible for cars to enter and take it over. Restrict vehicular flow to one or two sides maximum, otherwise an urban space entirely surrounded by roads is effectively cut off. The key concept here is to plan for both access and transit for pedestrians, but access and very restricted transit for vehicular traffic (see [32]).

Alexandrine Patterns that define urban space

Christopher Alexander's design patterns offer rules for designing urban spaces that invite users.

Living space envelops and nourishes us. This primal, biological sense of space goes far beyond strict mechanical utility. A new approach to designing urban spaces that is freed from often-irrelevant architectural accretions can help to bring our cities back to life. Living urban spaces are the "neural nodes" of the city, connecting the flows that bring it to life. This toolbox is what architects have long sought, but which many have paradoxically rejected. Empirical facts encoded in socio-geometric patterns lead us to understand the elusive properties of "living" spaces, which exist on a much deeper level than we are used to thinking about when we design.

There are many disappointing examples of architects who studied traditional urban fabric in great detail (i.e. Barcelona, Paris, Rome), yet failed to understand the information structure of the older buildings. When those individuals applied what they learned, it was only to produce a caricature of what they had seen. For this reason, it is essential to learn the correct systemic rules under which living urban fabric is synthesized.

Some living patterns selected from Christopher Alexander's *A Pattern Language* [4] reveal key elements that can help us in the design of urban spaces (for more details of the pattern method, see [16] and Chapters 8 & 9 of my book *Principles of Urban Structure* [36]). The following pattern summaries are my own, and they focus on the spatial aspects of open spaces. The reader is urged to consult the original, lengthier version of each numbered pattern, which includes research material giving detailed arguments and/or scientific validation for the patterns.

TABLE: FIVE PATTERNS FROM "A PATTERN LANGUAGE".

Pattern 60: ACCESSIBLE GREEN. People will only use green spaces when those are very close to where they live and work, accessible by a pedestrian path.

Pattern 61: SMALL PUBLIC SQUARES. Build public squares with a width of approximately 60 feet. Their length can vary. The walls enclosing the space, whether partially or wholly surrounding it, should make us feel as if we are in a large open public room.

Pattern 106: *POSITIVE* OUTDOOR SPACE. The built structures partially surrounding an outdoor space, be it rectangular or circular, must define, in its wall elements, a concave perimeter boundary, making the space itself convex overall.

Pattern 124: *ACTIVITY* POCKETS. The success of urban space depends on what can occur along its boundaries. A space will be lively only if there are pockets of activity all around its inner edges.

Pattern 171: TREE PLACES. Trees shape social places, so shape buildings around existing trees, and plant new trees to generate a usable, inviting urban space.

A space will be used if it is designed in a way that feels enveloping and reassuring on the subconscious level. This has nothing to do with fashionable "design". Regardless of what our brain is recalling about architectural culture, industrial style, design ideology, contemporary norms linked to progress, etc., our body is reacting the way it has evolved to do so. Our body will signal with either a fight or flight reaction (in urban spaces that are not welcoming) or, under the appropriate circumstances, it could tell us that staying and experiencing this particular environment is just what we need.

Irrational belief in the redemptive powers of the "industrial look"

Contemporary Art applied to urban space – such as abstract sculptures and "installations" – repel instead of attracting people to use the space.

The design tools presented in the eight-fold model and its supporting documentation prove useful in designing new urban spaces, and in reviving existing plazas that nobody uses. But it is not enough to know the correct rules for creating useful urban space; we also have to clear up some misconceptions. Design thinking that purports "a designed urban space with industrial-looking objects draws people to it" is irrational and disproven by the evidence. Yet dominant culture continues to link industrial-modernist images with economic prosperity, moral superiority, and progress. Re-examining those irrational beliefs is necessary before we can implement a new scientific approach to urban space design.

The industrial "image-based" design paradigm has been tested repeatedly all over the world in many distinct situations, and it has failed every time to satisfy intuitive human reactions. The reason is that our body evolved to react against, and be repelled by anti-fractal forms, shiny surfaces, and unnatural materials. And yet, many architects and urban designers have internalized the belief that industrial-modernist design is "the best", and they don't question that assumption. I'm not trying to make a joke out of this, but to underline instead how much hostility we encounter from colleagues who still design according to industrial modernism.

In the present context, applying the industrial-modernist "look" to re-make existing plazas threatens our built heritage. A local developer and contractor can lobby for a "renovation" of a historic plaza; they get to make a nice profit. Whenever ideologues have succeeded in convincing the local politicians to do this, the result is dead open space, a loss of magnificent

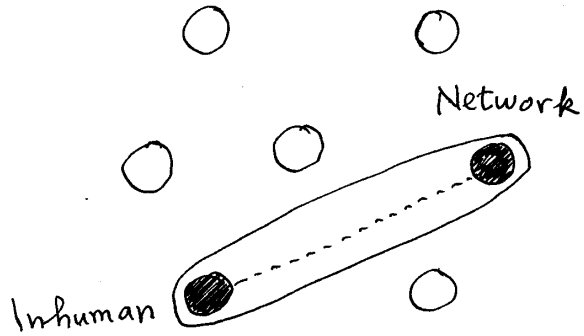


Fig. 5. "Installations" by a famous artist inserted into the Network city combine it with the Inhuman city. If the pedestrian flux is strong enough, users will ignore the negative emotions triggered by the installation; if the flux is weak, users will avoid the plaza altogether

century-old trees, etc. The world's most wonderful urban spaces are under threat from this misplaced image-based modernity, and many have already succumbed. Such acts of vandalism are allowed to occur because our society has been purposefully misled about what possesses "life" and what doesn't.

Part 4: How the Virtual city can either harm or help the Nourishing-physical city

Adding virtual city to physical city

The wonderful and liberating potential of information and communications technologies is best understood as an addition in the eight-fold model.

The Virtual city is a wonderful extension to the physical city. It instantly connects local to global nodes. The problem is that it can add itself to either the Nourishing-physical or to the Inhuman city. In the latter case, the opportunities afforded by the Virtual city can mask the trenchant problems present in the physical city. The opposite is also true: local problems and injustices can be brought to international attention via the Virtual city.

Modern society runs by organizing massive and continuous exchanges of information (see [31]). Our lives are improved when we superimpose the virtual universe consisting of interfaces and software onto the built environment consisting of buildings, streets, and urban spaces. This fit is best achieved through adaptive solutions, as represented by socio-geometric patterns [4]. Interpreting – and designing – the physical city using patterns makes the process of marrying it to the Virtual city effortless, since addition is then defined between two compatible epistemological frameworks (each one adapted to human sensibilities). But imposing image-based design that ignores rules for human interaction clashes with the software interface and results in a mismatch.

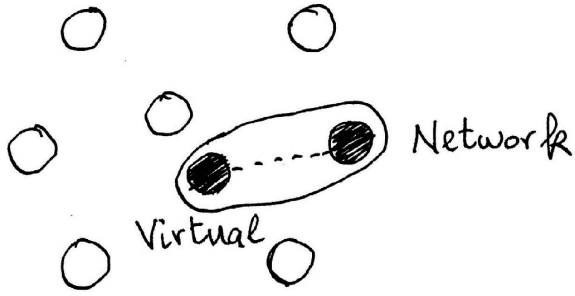


Fig. 6. The Virtual city adds to the Network city

The Virtual city is neither good nor bad in itself. The technology can work either to support wonderful adaptive human environments, or to make life possible in rather inhuman physical environments. Which one of these is favored depends upon cultural forces. The Virtual city is connected to the human scale, since it opens into the virtual world through a screen having human dimensions (devices from a few cm to 1 m in size). But its hidden design goes far beyond that because interface implementation uses traditional paradigms for interactions and information input to the human mind.

Unless the interface is intuitive, it will not be used. A ruthless selection towards the most accommodating experience and interface shapes the virtual city. Programs and websites that are in the least degree awkward to use – that do not adapt to our own biology – don't survive for long. Nor do those that consume our energy for irrelevant purposes. In a later section, I develop the reasons why the Virtual city functions, and when it doesn't, to satisfy human needs.

Let's consider the positive implications of the sum Virtual + Network city. Having virtual connectivity removes pressure from strong physical flux networks. People are no longer forced to move long distances for many of their essential tasks, nor to invest the energy to do so. The result is a self-adjusting flow that chooses where to move, and chooses the channel of communication, usually the one with the least effort and energy expenditure.

The Virtual city therefore frees us from many previously tedious displacements. One of the greatest successes of the Virtual city occurred when it coupled with the Spontaneous Self-built city in the slums of the developing world. When cell phones were introduced into those parts of the city, residents were immediately empowered to contribute to the local economy. Communications networks that were delayed for decades, or for which no government official ever planned, suddenly became a reality. The Grameen Bank, which gives microcredit to poor people in the developing world, immediately realized and applied this opportunity to generate entrepreneurial networks.

Emotional nourishment comes from the built environment through the coupling Nourishing-physical + Fractal city. The Virtual city repairs the skewed scale distribution of non-fractal modernist cities – which have a preponderance of long-distance connections – towards a more balanced spectrum of sizes. The presence of the Virtual city frees up infrastructure so that more attention can be paid to the closest physical connections. This encourages and enables intimate short-range contact with nature and with other human beings.

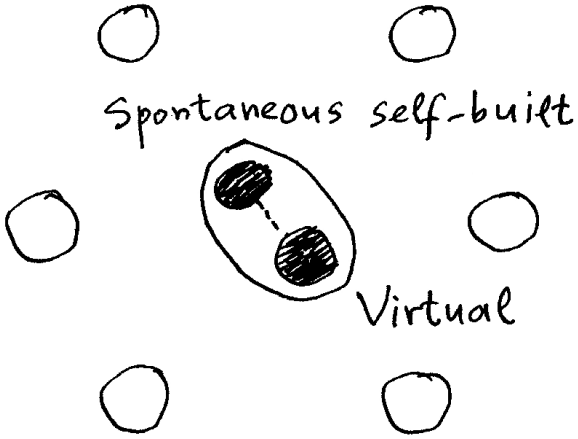


Fig. 7. The Spontaneous Self-built city has recently coupled with the Virtual city to connect into the larger city, and even the global economy

How the virtual city can make us Human Again

Instant connectivity with the world has surprisingly reinforced the most human, immediate, and local elements of the environment.

Restructuring the city for its nearest connections would represent the ideal situation envisaged by those motivated to create the Virtual city. People can experience the most wonderful settings of the Nourishing-physical city while being connected to the rest of the world. We now see many people working with their laptop computers in historic cafés and in human-friendly urban parks and plazas around the world. This makes it more urgent to preserve those welcoming traditional pieces of built urban fabric, and to try and create more of them in our environment.

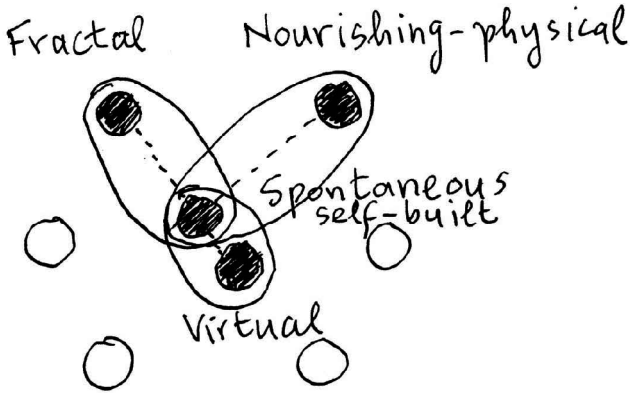


Fig. 8. The Virtual city adds to the physical city

Let's reverse the historical paradigm. What can we learn from the Virtual city that will help us design better buildings and cities? Above all, we need to channel design complexity so that the urban system offers a positive user experience in all sensory channels. Our goal in building should be to create healing environments: neither to impose some design ideology, nor to erect emotionally-hostile forms for short-term gain. The Virtual city is now positioned to either enhance, or to replace the physical city. What happens will depend upon society's priorities for the physical built environment. If an architect wishes to focus exclusively on form as image, that's up to the client to accept or not.

Components of the built environment were traditionally selected using adaptive criteria such as design patterns that produced emotionally-welcoming spaces and comforting surfaces, easy access for multiple channels of transport, and spaces that foster spontaneous human encounter. But the physical city now gives priority to the efficiency of vehicular movement, vertical stacking, and abstract design styles. Here it is necessary to distinguish one group of psychological reactions of excitement and fascination from sculptural forms, from the very different group of healing and nourishing emotions. Strange and menacing architectural forms fascinate small children, who experience a visceral thrill of anxiety. Such feelings are not healing in the long term, hence inappropriate for the living city (see my book *Anti-architecture and Deconstruction* [26]).

Comparing the interfaces of the physical and virtual cities, the city is degraded when new buildings express, through their geometry, the opposite of a "healing environment". Intrusive new signature buildings habitually ignore any surrounding historical buildings. Proposals for everything from new office buildings, to Art Museums, to concert halls may be visually imaginative, but with a decidedly transgressive quality in opposing both their surroundings and the history of the place. Yet disputing the long-term evolution of traditionally-scaled urban fabric does not help to create the Nourishing-physical city, but rather its opposite.

The industrial-modernist urban model was initially justified by the emotionally-loaded term "modernization" that continues into the Virtual city. But modernization is made possible by controlling – and often destroying – nature, whereas the Virtual city can actually save energy, and human-scale and natural environments. Science and technology permit us to save what is good about the environment with the help of information and communications technologies. We can continue to live only if we maintain an extremely delicate equilibrium with our environment.

But the Virtual city can also be misused. Academic architectural culture loves to pair Inhuman + Virtual city. But this is not a healthy coupling! Instead, it represents an excuse to keep converting the built environment into a giant abstract sculpture, which removes life from the physical city. Technology then becomes a diversion to draw attention away from non-adaptive physical design that creates a world of isolated individuals in a dead environment. This is not the fault of the Virtual city – which is a wonderful technological tool – but only its misapplication.

Part 5: Design for a campus

Welcoming open spaces

It is possible today to build learning institutions that offer a marvelous, life-enhancing environment for students, faculty, and staff.

The experience and imageability of any particular campus depend upon its spaces and perceivable organized detail. Those qualities are what the visitor remembers, and what the students, faculty, and staff experience every day. This result is not accidental or haphazard, but can be achieved by deliberately applying mathematical design guidelines. Those combine visually-oriented design with functionality. I list some of the most common mistakes below, so that knowing to avoid them will lead to a much improved campus design.

Many campuses built in the past several decades contain dysfunctional urban spaces. Those spaces do not invite, and in many cases actually prevent pedestrian use expected of an open plaza. The problems can be divided into two categories: (1) impediments to crossing the space, and (2) problems inherent in the surrounding structures.

Physical obstacles to traversing open space include continuous low walls for sitting that cut diagonal paths (but those low walls could be very effective when situated radially/transversely); badly-placed pools of water that do the same thing; misusing green in lawn that is out-of-bounds for people and which prevents direct paths; changes of ground level that cannot be easily negotiated; steps that prompt a pause and mental concentration in the user, which could have been eliminated; unnecessarily steep sloping ground, etc. All of these built features betray a lack of understanding of what mechanisms make an urban space function as such.

Paths become robust when reinforced by an adjoining edge ([36], Chapter 1: *Theory of the Urban Web*). Elements such as benches, low walls, lawn boundaries, and stairs need to run *next to and parallel to* potential paths, not across them. A sufficiently wide staircase encourages flow along its bottom step much more than transverse movement up-and-down the stairs.

The second set of problems concerns the buildings surrounding the open space. The ideal qualities here include compositionally rich and visually welcoming façades, such as found in highly-ordered information, fractal scaling, and multiple symmetry content of traditional buildings. One feels the desire to cross a plaza or open space when attracted by a visible, emotionally-welcoming goal on the other side (whereas minimalist concrete, bonded brick without patterns or features, and glass curtain-walls – none of which attract us emotionally – trigger the opposite effect). Another welcoming quality of the boundary is to be found in porticoes on one or more sides of the plaza. Such a protected space encourages pedestrian activity all around the boundary of the open space. Discontinuous arcades may look nice but, are, as a consequence, never used.

Alexander's Oregon patterns

Christopher Alexander derived design rules for the University of Oregon campus in 1975, and those rules are universal.

We can apply the Nourishing-physical + Fractal + Network city to design a campus that will contain all the positive qualities of our best-loved historical institutions. A college or university campus represents an urban microcosm, with its limited yet often extensive area and restricted mixture of uses. One needs different buildings for classrooms, research laboratories, libraries, student housing, cafeterias and student activities, sports, maintenance, administration, etc. The pedestrian realm is paramount, since students have to walk from building to building. Essential vehicular connections ideally go around or under the main network of pedestrian paths.

Christopher Alexander created a long-term planning strategy for the University of Oregon based on design patterns. Some of those patterns appear in his classic book *A Pattern Language* [4], whereas others are to be found only in the lesser-known *The Oregon Experiment* [6]. I recall some of those findings here, and explain how they apply to the eight-fold classification of city types. The pattern descriptions given below are my own summaries.

TABLE: ALEXANDRINE PATTERNS FOR DESIGNING A CAMPUS.

Oregon Pattern 2: OPEN UNIVERSITY. Do not isolate the university by surrounding it with a boundary; instead, interweave at least one side of the campus into an adjoining city, if that is possible.

Oregon Pattern 3: STUDENT HOUSING DISTRIBUTION. Locate some student housing within the center of the campus, with different percentages in regions as one moves away from the center. The first 500 m radius containing ¼ of the resident students; ¼ in a ring between 500 m and 800 m radius; and the rest outside 800 m.

Oregon Pattern 4: UNIVERSITY SHAPE AND DIAMETER. If possible, situate classrooms within a central core of ½ km radius, and non-class activities such as administration, sports centers, and research offices outside.

Oregon Pattern 5: LOCAL TRANSPORT AREA. Give priority to pedestrian flow in the central core of the campus, within a radius of ½ – 1 km. Vehicular traffic here must be made to go on slow and circuitous roads.

Oregon Pattern 12: FABRIC OF DEPARTMENTS. While each academic department ought to have a home base, it should be able to spread over into other buildings and interlock with other departments.

Implementing the Network city prevents cultural and social fragmentation, while the Fractal city helps to distribute forms on many different scales. The Network city emphasizes pedestrian paths forming a network of connected urban spaces, and protects those paths from encroachment by vehicular traffic. It also offers integral connectivity between the campus and the city outside. The special requirements of a campus give it even more urgent pedestrian needs. Every building needs vehicular access, but that must take second place to the pedestrian connectivity.

An obsession with mono-functional zoning often forces all student dormitories on a campus to be clustered together, while all administrative functions are housed in a single, imposing building, etc. Yet functional segregation does not produce an ideal learning environment, as it works against mixing and compactness.

The departmental pattern (*Oregon Pattern 12 given above*) points to a pragmatic approach that has a major influence on planning morphology. Whereas it is standard practice to segregate academic departments into separate buildings, that never works in practice. Suppose the "Chemistry Building" is funded and built. Yet by the time the Chemistry Department gets to move into its new offices and laboratories, it has either grown or shrunk in size, so it no longer perfectly fits the building. It is more practical to adopt the approach that no single building should be expected to contain a university department. Thus, it makes better sense to physically connect a building to adjoining buildings rather than have it standing apart.

Avoiding planned isolation

there is no practical reason to isolate a campus from the larger community, and that is only a holdover of single-use planning.

People perceive campuses with block buildings and hard open spaces as bleak, desolate, threatening, inhuman, and totalitarian. The human scale is missing. And yet this industrial style has shaped a majority of institutional construction for decades. It would appear that school administrators decided to industrialize education, and concluded that industrial-modernist architecture was most appropriate for the task. The campus becomes a piece of the Inhuman city in which buildings are placed too far apart to connect.

The planning habit of mono-functional zoning is also applied to unnecessarily separate a campus from a region of "normal city". This way of thinking is responsible for the "corporate campus" of major companies isolated in the woods, or at least far out in suburbia. But, while that setting has positive biophilic qualities, it is deliberately not part of the city. An even worse precedent is the misleadingly-named "office park", which is just a cluster of unrelated office buildings. Both of those urban typologies define a life separated from the rest of humankind.

Historical evidence points to the intentional isolation of workers from city life so that they could be totally controlled by the employer during the workday. The corporation tried to force employee allegiance by isolating them. In a similar vein, many people believe that social engineering was applied to High School and college campuses, implementing a fortress typology in order to better control rioting students. But this claim is unsupported: it just happens that architectural style coincided with typologies whose principal concern was security.

While the corporate campus was, at least in name, loosely copied from the traditional university campus, its urban model is the suburban shopping mall surrounded by vast areas of open parking. Everyone commutes by car. But now this typology has come full circle, with institutions of higher learning copying the isolated corporate campus and suburban office park.

"Walkabout" design with human sensors

A revolutionary method of direct human responses to imagined forms, performed on the actual site, reveals a vast amount of useful design information not otherwise available.

Design methods using emotional feedback from people have a lot in common with how the Spontaneous Self-built city arises. Slum dwellers do not follow building regulations, but are instead guided by their intuition and the physical limits of available materials, space, and topography. Incorporating aspects of that design freedom into conventional practice yields a method that adapts better to human feelings and sensibilities. I have proposed implementing this method to upgrade informal settlements and erect new self-built housing around the world (see [37]).

Given modern industrial materials and systems of construction, there is an economy to rectangular spaces in terms of standard materials, labor, and utility. Regular building codes have a very limiting effect on design, and act against individual negotiations with existing conditions. And yet, an intuitive method obviously worked for millennia. Ever since people have had to rely on architects and the building industry for so one century, they have forgotten or have suppressed their instinctive dwelling-making skills. If today's industrial-modernist paradigm is to be overcome, or at least modified to obtain a more human design, we need to re-awaken those timeless methods of design [2].

I'm going to delve into the design methodology known as collaborative, consensus, or participatory design. That approach involves eventual users in an essential manner in producing the design. I will need only one very specific component of the collaborative method, which makes design decisions on the basis of direct emotional feedback (an exploratory method for creating the Nourishing-physical + Fractal + Network city). An intuitive judgment based on the users' feelings and imagination is made before construction, giving birth to the design using only what exists already on the site.

The method is the following: choose a group of about five people, to include a child if children are going to use that place or live there. The group walks the grounds trying to imagine the proposed building fronts already standing; not in some predetermined form, but rather where a built wall and openings would feel best to reinforce those open spaces. The "walkabout" guarantees that urban spaces are well defined on a human scale and are connected by a network of pedestrian paths (Network city). For this process not to be ill-defined, the group needs some rules and guidelines of what is possible; and the group should include someone trained and knowledgeable in Alexandrine Patterns to guide the process. Decisions are reached by discussion and consensus.

Christopher Alexander suggests for the group to carry wooden stakes and poles with small flags on them [5]. Those are used to mark the paths, the boundaries of open spaces, and the footprint of the imagined buildings. Someone could hold a large Styrofoam panel and stand in particular spots so that the group can decide if that's the optimal position for a wall. If all goes well, then multiple factors such as solar orientation, adaptive use to wind flows, levelness of the land, and regard for natural elements on the site (trees, boulders, sharp drop-offs, steep hills, etc.) will be accommodated just by the sensory feedback.

After this design walkabout has been carried out on the actual grounds, and checked once again after the positions of other key elements have been decided, the plan is transferred to a measured drawing. "Cleaning up" the design so as to align directions and tidy up the geometry should be resisted, since that may invalidate the empirical discoveries of the group. This is the opposite of the standard procedure, in which everything down to the details is drawn in the office, and then built. In the conventional design approach, the users get to experience the final configuration after it is permanent; i.e. only after it is too late to make any adjustments, or even to correct major errors and omissions.

Alexander himself used this method to build a new high-school/college campus outside Tokyo [5]. Once the urban design and the architecture of each individual building had been determined, the construction of the campus was carried out via conventional methods. The resulting cluster of buildings and grounds show a degree of life that is essential for human engagement and wellbeing.

The exploratory design group should include persons who have a strong interest in using the built urban fabric after it's completed. It is recommended to have someone with sufficient technical knowledge to help provide structure to the decision-making process. Individuals participating in the "walkabout" should be encouraged to draw upon their human intuition and sense of place to guide them in their conclusions. This can be difficult at first, given the decades of industrial-modernist construction led by architects and professional builders, which distanced users from their instinctive sense of dwelling and place-making. The detachment was achieved by institutionalizing both design and construction.

Alexander's method puts our human sense of place ahead of industrial design practices, by promoting human intuition ahead of formal planning. Exploring the site, on foot, independently of existing paths and road structures (except for features that absolutely cannot be changed) helps to establish an optimal connected network of pedestrian paths linking urban spaces. At the same time, the exploratory process discovers how the pedestrian network should connect to internal and external vehicular networks.

The same method applies to diagnose already built urban fabric. An exploratory design group discovers and then maps those healthy places where it observes intense urban life, and which are deemed by their users to be vital. That quality is judged both by positive emotional feedback and by the density of pedestrian use. Such spots are marked as being protected from damage or encroachment by new projects. Yet those key healthy places could be architecturally modest objects, such as a tree, a wall, a corner, a small structure, etc., that conventional planning would not hesitate one second before eliminating.

Equally important is for the exploratory walkabout to identify pathological paths and places. If a place or pathway triggers psychological distress, there is something wrong with the geometry. The sensations could be a feeling of being oppressed; made anxious or threatened by the geometry or by something else; of being too exposed; ill-at-ease, etc. First identify those spots, and then think of possible restructuring and transformations to fix the problem – which is an emotional and/or intuitive reaction, not something that can be easily discovered from looking at a plan. If the new planning scheme requires that something be destroyed to erect a new building, then care should be taken to leave the healthy places alone while

sacrificing the unhealthy ones instead. This way of thinking can help repair the urban fabric by not allowing new construction in arbitrary locations, such as where someone thinks it's a good idea simply based on the plan.

How to build a fractal city through budget allocation

The geometrical notion of fractals combining components of different sizes translates into a funding formula that allows us to build all the sizes in an urban ensemble.

How do we optimally distribute the money to be spent on building the Fractal city? It has to be done using a fractal distribution of funding. Suppose that we have a central source that allocates different sums to specific projects, and where each project competes with the others for funding. This is the case with a university campus, since the majority of the budget comes from a single source, with the possible exception of specific donations for individual buildings (and even those often have to be "matched" by university funds). The administration has to argue for its projects' approval in front of the funding agency, its own coordinating board, or the government.

The conventional procurement method is rigidly anti-fractal because it concentrates on the largest projects: those need the most money, and not getting them approved carries the greatest risk. But that top-heavy mindset too often ignores the intermediate and small-scale projects. The budgetary thinking is that those can be accomplished by way of the university's general operating budget, or from discretionary funds found here and there. Yet that is seldom the case, and a systemic imbalance towards the largest scale remains to shape the built environment in undesirable ways.

A big project is easily presentable, hence an important marketing tool. The architect draws a pretty picture of the large stand-alone new building, which is used to convince the decision makers. The idea of a single structure and its striking image can be linked to expectations of how this new structure will make the University look like it is growing and thus successful, progressive, and modern. But the current system can create dead spaces in-between indifferent stand-alone new structures. It is much harder to use smaller, interlinked projects to market the university's value. Human psychology works against presenting an intricate, adaptive environment: *it has to be experienced in person because its life-affirming qualities do not show in a picture!*

The Fractal city suggests a better funding formula. Just as a fractal has components whose sizes obey an inverse-power distribution, we propose the same law to govern funding for projects according to cost/size. An inverse-power distribution is one where the number of objects in a system is inversely proportional to their size: there exist only a few large objects, several more of intermediate size, and very many smaller ones, increasing in number the smaller they get. Fractal funding would support only a few large projects, several of intermediate cost, and very many low-cost projects, in a balanced relationship that favors the lower-budget ones.

A simple means to apply a fractal distribution to the funding formula is to divide the total budget into equal portions; say five. Then assign each 1/5 portion of the budget equally among

a group of construction proposals having roughly the same cost/size. That will automatically guarantee that the smaller the projects are in terms of funding, the more of them will be approved. While we may never be able to systematically change the budgetary process, just getting this kind of thinking into the heads of the university planners as they work to prioritize projects might begin the process of seeing how fractal budgeting helps to create a greater equity in overall place-making within the campus. This revolutionary approach to budgeting is the best way to keep healthy urban fabric in repair. Most interventions and additions that can make a great deal of difference for the better are either of small or intermediate size. Those need to be done often. The largest projects, which the current system is skewed to privilege, are possible only every few years. The university sees these new buildings as visible proof that it is growing, and, while it may not display such a building in a student brochure, it feels satisfied with the news coverage. But those big projects are disastrous when they fail. Of course they make money for the builder, but that's not the point here.

Christopher Alexander first proposed this fractal funding formula in his long-term urban plan for the University of Oregon [6]. Alexander's result was based on his own original analysis, and came before the introduction of fractals into architectural theory. I explain why this inverse-power distribution is essential for the stability of all systems, as for example ecological systems (see [36]). There are really very deep justifications for this approach that have to do with the nature of complex systems. If by past precedent the formula for funding projects has become skewed towards the largest scale, we have to work to remedy this imbalance. How projects are funded is the key to creating more human-scale spaces and places.

The university campus as a microcosm of tradition

Institutions face a fierce opposition between living campus environments that look old-fashioned, and contemporary architectural expressions, which do not contribute to emotional and physical wellbeing.

Well-defined urban space is not merely an aesthetic option; it is a vital necessity to the campus experience on a human level. The most valued universities have prominent open spaces, not necessarily large, but always distinctive and very well defined. University open spaces work best of all, and are the most memorable, when flanked by historic buildings (i.e. particularly those with well-developed form languages in their designs). Those spaces frequently define the university's identity for the rest of the world.

Creating welcoming urban space depends upon building types. Many universities pride themselves on having buildings designed in contemporary styles placed prominently around campus, and newer additions seem to follow an institutional model of stand-alone buildings. Fashionable "contemporary" buildings are being built more and more with donations from wealthy donors (who expect their name to grace that building), but research shows that more traditional biophilic architecture lends itself better to a learning environment (Nourishing-physical city). Pre-modernist buildings provide, through their materials and designs, organized information that helps trigger a greater sense of wellbeing, which in turn

promotes greater participation and engagement on the part of students, faculty, and staff. To the contrary, industrial-modernist buildings emulate sensory deprived environments, which can create a degree of hidden anxiety that permeates the learning experience. It is harder to learn and retain information in stressful situations or environments. Parents expect their children to learn from traditional stores of knowledge, and, while innovation is expected and welcomed, it is not supposed to displace inherited knowledge. The traditional center of learning represents cultural inheritance, and that should also show in its buildings.

An informal survey of brochures put online to entice prospective students in the USA (and even more, to convince those students' paying parents) reveals that the vast majority features strictly traditional buildings. Those older buildings have an instinctive appeal because they link to stable and timeless values. While universities may indeed have industrial-modernist or alarmingly "contemporary" buildings on campus, those are not usually displayed in the brochures. Expensive private institutions, especially, employ psychological marketing techniques to justify the high expense of a university degree with their long-standing prestige. Those present their traditional campus structures instead of their more contemporary (abstract) structures, since people typically respond to the thrill of architectural transgression with alarm, and subconsciously sense that inherited knowledge is also being threatened (see [26]).

Two separate design problems are relevant to institutions of learning: (i) choosing an appropriate architecture for new buildings, and (ii) laying out the plan of the campus. The first question leads to a sort of schizophrenia, because parents tend to want traditional "reassuring" buildings, whereas the university is pushed by fashion trends to choose the opposite in new buildings. It would appear that the administration recognizes this conflict, preferring that the parents discover the alarming contemporary buildings on campus – representing transient ideas – only after their children start to attend classes at that institution.

The second problem creates a conflict between the need for additional buildings, and the necessity for all students to reach their classes within a 10-minute walk (the normal break between classes). These two demands are irreconcilable if the campus keeps expanding with singular new buildings, as most do. The solution is to implement an intelligent compactness and intricately folded complexity, such as I discuss here under the Fractal city. The opposite trend, which is to erect stand-alone industrial-modernist or "signature" buildings, negates compactness and useful urban spaces. For creating intelligent compactness and intricately folded complexity, traditional spatial solutions work best.

Institutions that have gambled with their endowments to erect gleaming new buildings by trendy architects are participating in a very expensive experiment. They invested in flashiness instead of reinforcing the spatial and urban qualities of the campus. They took a massive bet that those cutting-edge university buildings will draw in a new generation of paying students. A separate misconception is that cutting-edge research requires alien structures to house it, and thus universities erect flashy new buildings to draw in research dollars. Whether that occurs or not is a matter to be determined by future applicant statistics and number of grants. Nevertheless, partial results already hint that the experiment of innovation through fashionable but disruptive design is a dismal failure. Lists of "The ugliest campuses

in the USA" invariably include precisely those institutions whose buildings' design purposely panders to pseudo-intellectual pretentions that naturally oppose our biology (Inhuman city). Who wants to go to a University that is included in such a list?

A glowingly positive example is Christopher Alexander's High-School/College campus outside Tokyo, built in 1985 [5]. Alexander and his design team researched deeply into Japanese architectural culture to extract a form language appropriate for an institution of learning. The result is a modern campus that has comfortable, timeless qualities. Students, teachers, and parents love it. The only problem that arose was with the local construction companies, which had been expecting to build the usual concrete boxes.

Part 6: Learning from informal settlements

Solutions for and from the world's slums

Instead of flatly condemning self-built settlements, we can instead apply their design strategies to make planned cities more human.

The world's booming urban population is housing itself in vernacular/indigenous settlements. Both Christopher Alexander and John Turner urged a re-appraisal in how we conceive of self-built settlements [22, 37]. The Spontaneous Self-built city has an organic structure – a positive quality – like the Fractal city. However, urban problems include lack of infrastructure and higher-scale network connectivity. Those could be provided by the state, or by the residents helped by Non-Governmental Organizations (NGOs). But regulatory/financial power habitually does not consider organic, collaborative solutions. Instead, it pursues the Developer city paradigm where the government bulldozes indigenous settlements and erects social housing blocks in their stead. Or it forcibly evicts people and then sells the land for a huge profit. The world's poor cannot defend themselves against the financial power of a developer acting in collusion with government, which treats them in a feudal manner.

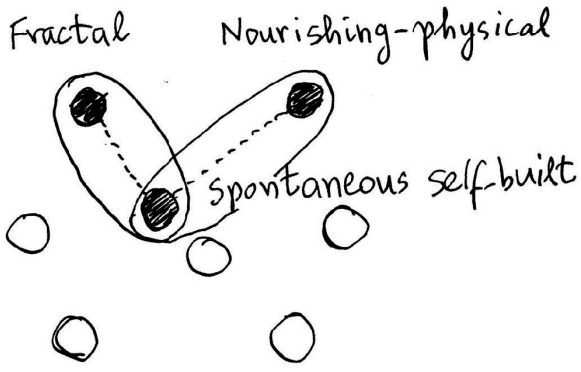


Fig. 9. Informal settlements evolve into living historical cities

The value of studying informal settlements becomes obvious when we look into the historical evolution of cities. Only a small fraction of those can be considered "slums". Some cities began as Spontaneous Self-built cities, acquiring infrastructure only later. Under favorable conditions of land ownership, settlements upgrade into better, permanent buildings, and we observe the same process taking place today. With the right conditions in place, the informal city could eventually evolve into the complex permeable fabric of the Nourishing-physical + Fractal + Network city.

This paper looks for solutions found in informal housing that could be beneficially applied to planned cities. The spontaneous nature of design when people build for their own survival and basic comfort is missing from most of today's cities (see [31]). Design based on feedback from man-made forms in the environment, like those that generate vernacular/indigenous cities, could never have led to the alienating places that have become the industrial-modernist standard around the world.

Instead of designing exclusively in the architectural office, under the influence of intellectualized design preconceptions, there is greater human benefit when the process is derived directly from the physical, visceral reality of place. Vernacular/indigenous architecture is our best guide to achieving this [2, 7]. In an earlier section, I presented an experimental, heuristic method, following Christopher Alexander. I wish to establish a fundamentally interactive approach that automatically links the Spontaneous Self-built + Nourishing-physical + Network cities.

To learn from the Spontaneous Self-built city, we need to outline the priorities of design under those special circumstances closest to basic survival. Purely contextual design is conceived with locally-available, low-cost materials that are cheaper to maintain long-term than imported products. Note that many cases exist where the building stock and layout are inherited from the planned city, so that we find skyscrapers (the Tower of David in Caracas) or suburban tract houses converted into slums. But I'm talking here only about the Spontaneous Self-built city.

TABLE: PRIORITIES IN THE SPONTANEOUS SELF-BUILT CITY

1) Focus is almost exclusively on the pedestrian realm and pedestrian connectivity;
2) Architecture utilizes manageable (softer) materials that can be shaped by non-industrial means. No large components of glass or steel are used, but mostly construction pieces on the human range of scales;
3) All built structures are on the human scale, with buildings typically no more than four storeys in height. This corresponds to Alexander's Pattern Number 21;
4) Optimized low-tech and passive energy use is achieved by means of insulation, indigenous construction methods, and solar and wind orientation;
5) Space is maximally used and is therefore at a premium; a large open space is a luxury;
6) Car access is included where possible, but this is not a priority for shaping the urban realm;
7) Available means of ornamentation are used, even if it is only a variety in surface texture or brightly-colored paint.

The actual materials and energy use are a primary concern for self-builders. Nevertheless, my interest here is with the morphological aspects, which I wish to utilize for designing more mainstream planned cities.

Many of the central design tenets of the planned city are irrelevant for self-builders. Those include formal planning grids; straight lines; monumental buildings and vistas, criteria of architectural "style" and fashionable "design", etc. In short, much of what characterizes a planned city is simply not a concern in creating an informal or vernacular settlement, which switches its priorities from those of the controlling class of society. Therefore, much of what is accepted as an indispensible part of "functional" design turns out to be gratuitous: materials used wastefully to project a certain industrial-modernist tectonic "look" that has nothing to do with human functions. (But human-scale ornament is essential and necessary).

It is in the vernacular/indigenous settlement that we find truly adaptive and sustainable design: passive solar energy use, indigenous construction, etc. This comparison reveals disturbing and stark contrasts between a population surviving on minimal or zero amounts of fossil energy, or on meager supplies of pirated electricity, versus glass skyscrapers brightly lit throughout the night in an ostentatious display of corporate and institutional hubris. The same can be said about the world's petroleum supply being squandered by commuters driving around suburban sprawl.

A key connective difference exists between desperate slums versus healthier indigenous cities: the former belong to the Anti-network city, whereas the latter belong to the Network city. Poor residents who are connected to the rest of the city can easily survive, find work, and commute to it. But life becomes desperate for those who are totally disconnected.

Forces set against informal settlements

City-building needs to be understood from the forces – both good and bad – that drive it, and which pit profit-making against owner-built housing.

The world's majority building activity has always been informal, and history documents a long and futile war against the ungoverned conditions of organic growth. The Spontaneous Self-built city is a solution, not a problem. It is built intuitively, following biologically-based and evolved rules. Today's industrial-modernist architectural culture, whose existence depends upon deliberate design and planning by a professional caste, sees this organic process as apostasy and even as an existential threat. At some point, the industrial-modernist design paradigm decreed that people have to live in industrial block mass-housing. This seems extreme for current conditions, yet the idea lives on to subconsciously shape projects today. That decision has been softened but never repudiated.

As the eight-fold model demonstrates, an adaptive approach to design can be annihilated by its opposite. We have that situation here, because the culture of contemporary architectural academia simply cannot accept and learn from vernacular/indigenous (i.e. place-specific and self-built) architecture. You can easily verify this, especially when architecture students are asked to design or build interventions in the slums of the developing world. Those projects – done out of a concern to help – turn out to be totally out of context and painfully disdainful

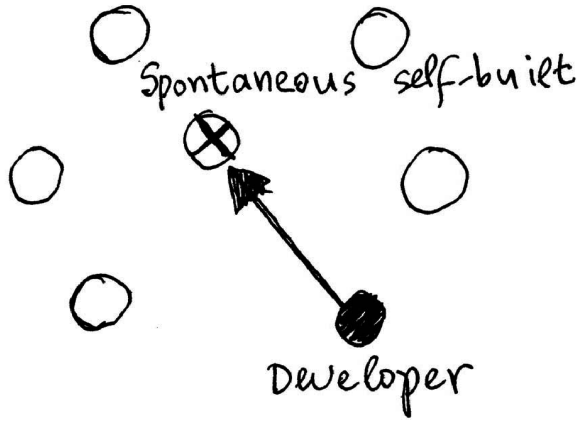


Fig. 10. Private developers and/or governments annihilate informal settlements

of the local building culture. They embody an opposite worldview and tend to stick out like alien invaders.

We continue to apply the industrial design paradigm for reasons that are now purely economical (profit-driven). Massive building projects, huge buildings, and blocks of social housing are a goldmine for those engaged in the banking, construction, financing, and real-estate industries. This is true in efficiently run democracies. Honest profits are made from the "cementification" of the environment, and the only disadvantage lies in the product, which is unsustainable and often inhuman. But the actual situation is far worse: in many political settings around the world, "cementification" is anything but honest, and is favored because it offers the opportunity for political influence and corruption.

The power game of who gets to build large projects, and what typology will be applied, is decided by factors outside architecture. There is no broad power base that can fight for a human built environment. Certainly not government. Moral, religious, and cultural authority is insignificant today compared to the huge money interests that drive cementification, not to mention when the financial backing of organized crime enters the picture. If only there were some comparable force to guarantee a level playing field; then we would see far more human-scale development. Limited human-scale development does still occur, whenever a small or medium-size commercial developer decides to adopt new traditional design typologies. When developers become aware of the market trends valuing more human places, they understand that this demand translates into greater profits.

Unhealthy isolation of slums

Inadequate living conditions in a marginal settlement are due as much to the lack of connectivity as to the built form of the urban fabric.

The problems of the slum include network isolation: it is not usually part of the Network city. Even in cases where no physical boundary has been erected to isolate a slum from

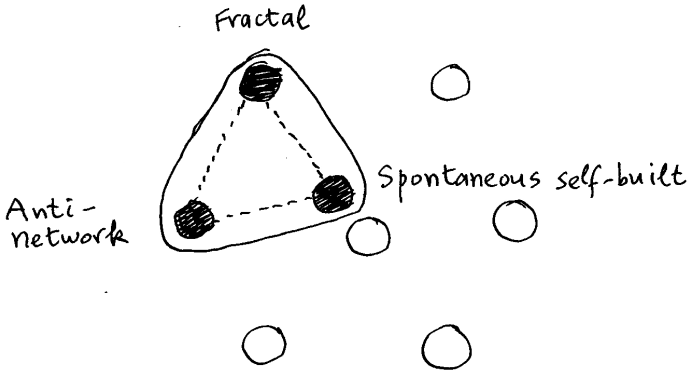


Fig. 11. A slum may possess positive fractal/biophilic qualities, but it is geographically and socially isolated from the rest of the city. The Anti-network city is not healthy, even when coupled with other healthy city types

the rest of the city, an inviolate psychological barrier exists on the ground, which defines the geographical boundary of the slum. As long as this boundary is respected, life activities continue on either side in parallel yet never mixing. But occasionally, one side decides to invade the other so as to expand its own territory, as was explained earlier.

Some slums have good pedestrian connections but very poor vehicular connections, hence they are isolated from private vehicular transport and from any pubic transport network. The medieval parts of Naples and many Arab-Islamic cities illustrate this isolation. This was precisely the problem that Baron Haussmann solved through his reconstruction of Paris, and Ildefons Cerdà did something similar in Barcelona. I don't agree with their heavy-handed interventions (driven, in the case of Haussmann, by military considerations) that destroyed healthy urban fabric, but at least their respective results became parts of the Nourishing-physical city.

Part 7: Identifying "bad" versus "good" city types

Problems with the Developer city

Developers need to move beyond post-war typologies that generate disconnecting environments, and begin to apply known and proven methods for creating living cities.

With the exception of private houses in some regions, most major construction efforts are designed and get built by either the government or by private developers. Both cases require the project to be reviewed and approved by one or two government agencies, whose fundamental purpose is human wellbeing (or life-safety). Unfortunately, the building review is performed by individuals who keep to local or national building codes (i.e. safety systems within a building). Planning review for larger-scale projects is too often predicated on

infrastructure requirements and strict adherence to zoning laws that limit pedestrian/vehicle standards. Projects only need to meet the minimum standards to be approved and built.

Decision-makers are typically unaware of any larger aspect of human wellbeing, such as those elements that directly affect the quality of urban space. That occurs because contemporary society has no clear conception of a living city, or it erroneously condemns elements of living cities as something undesirable from the distant past. There are simply too few people that understand that need for good urbanism. It does not help that the dominant architectural culture, for over a century, has rejected traditional solutions. Investors who wish to make a quick profit from construction most readily accept this prejudice, which does not require them to adopt a deeper and more responsible attitude towards human sensitivities.

Short-term turnover of enormous amounts of money can kill the life in a city through a "destroy and replace" strategy. Common people naively believe that the Government will protect us from such negligent practices by way of building regulations and minimum property standards. However, the financial forces are so powerful that building regulations are influenced to conform to the industrial model. And so, when a project follows these standardized regulations to the letter, the result is an industrial modernism that serves the industry and not the essential needs of human beings. Still, those legislated zoning and construction codes have remained essentially unchanged since the end of the Second World War.

In recent years, form-based codes (described in a later section) have begun to make their way into the city-planning lexicon. Unfortunately, without a deep understanding of the components of form-based codes, local city planners apply these principles with the same image-based sensibility as an industrial-modernist architect, thus making them ineffectual.

With the continued privileging of image-based (i.e. formal) design, a group of investors retains a starchitect to help sell an extremely expensive (i.e. profitable) project. It seldom has anything to do with the immediacy of place. A related publicity campaign promotes the project to the population at large, promising them a place on the world's stage and guaranteeing that their city or country will automatically be seen as "contemporary". This

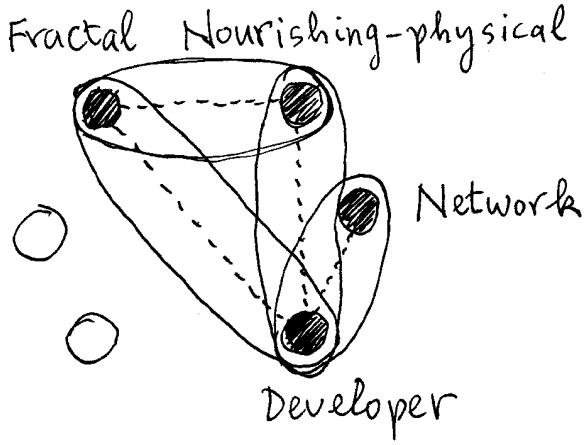


Fig. 12. A good mix occurs whenever a developer implements traditional design

public relations trick is usually enough to give the go-ahead based on the starchitect's media recognition. The reality is different: the new building or urban re-structuring could turn out to be mostly useless, as some cities have discovered after going bankrupt. The result is that a significant portion of a city's or nation's budget is wasted on a piece of fashionable sculpture: something adversarial to a nourishing built environment. Money that could be better spent on the city's infrastructure is instead recycled into the pockets of the developers and construction companies. The mercenary architect gets his/her share for collaborating in this deception.

It is most unusual to apply moral values to city form. Nevertheless, the eight-fold model has an explicit dimension of "bad" (city types towards the left and on the bottom of the eight-fold diagram) versus "good" (city types on the top and right of the eight-fold diagram), with all gradations in-between. The criterion for judgment is the physical and psychological wellbeing of all users. The design mainstream, however, uses five very different criteria for judgment:

1) A set of prototypes is applied to urban form because some famous person said to do that;
2) Zoning and planning laws (exceeding their original purpose of separating noxious industrial activities from other more benign uses) become stylistic dictates that separate the linked activities of a living city;
3) Poor solutions are perpetuated and typologies become standardized even after multiple implementations find them to be deficient in their human responses;
4) We privilege an abstract sculptural approach to city form that ignores how humans move through, and react to forms, spaces, and geometries.

The Developer city that chooses not to use new traditional typologies and codes is instantly recognized by its unnatural "look-and-feel". Even when buildings are made to look more natural and traditional, the result is often a poor imitation, because only superficial qualities are copied. No attempt is made to mimic the generative processes that grow a living city. Traditional rules for creating living structure have been followed poorly; but more often, the rules are totally ignored. This is because nobody understands the process, or feels that applying it is too time-consuming compared to building a standard ready-made box.

In an extractive global economy, such as we have today, industry pushes towards the largest scale. Commercial forces shape the same unsustainable and non-adaptive Inhuman city all over the world, with no concern for the civic realm or the individual. Not coincidentally, the same global companies and engineering firms are involved in far-flung projects. A handful of starchitects are its eager mercenaries. Financial backers like this business model because the projects are promoted using public relations images, and the starchitect's name comes in handy. Governments go along with this model, drawn into the promotion scheme by their own ego and the promise of international prestige.

Developers make money in a framework defined by architectural culture, market forces, and government regulations. They respond to opportunities and are constrained by legislation put in place by the dominant culture. The current situation favors industrial systems on the largest scale, run by fossil fuel, and disconnected from the urban fabric. Financing,

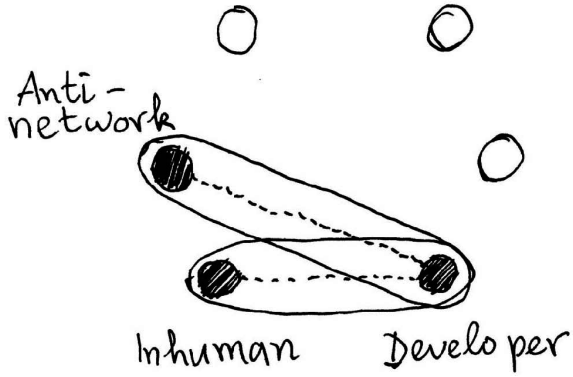


Fig. 13. A bad combination occurs when the Developer city implements pathological city types

insurance, leasing, management, and maintenance forces influence a developer to build either something healthy, or something unhealthy, usually for a comparable investment outlay. The decision of which one of these two ought to be built is made indirectly by society.

We should not lay exclusive blame on the developer or government for building the Inhuman city: it is society's fault for supporting "design by image" and the wrong urban codes. Developers will take advantage of the uneducated masses (including those working for the permitting system, who wish to shun responsibility). Society as a whole is uneducated about design and has been trained to accept and even prefer non-adaptive structures. The only thing we can do now is to educate the public on the dangers of continuing to degrade our environment in this manner. Maybe the arguments presented here will eventually convince actors in city construction of the advantages of creating a living city.

Singular expressions and the architect's dilemma

Getting a commission based on adaptive novelty has to rely on educating the public to value this instead of flashy and superficial novelty.

I am trying to empower every architect to adopt a step-by-step adaptive design method instead of the usual all-at-once gesture. To be successful, the feedback required for the interactive design operation presupposes a willingness to collaborate with factors other than pure individual artistic expression. But European and US culture privileges the singular, the flashy, and the obviously "new" instead of anything that resembles something copied from an old town. And people are not educated enough to recognize the difference between superficial copying and a design that has evolved to adapt to the conditions of a place.

To give morality and humanity back to the practice of architecture, the public needs to get beyond boosting the narcissistic starchitects. Those individuals have established the dangerous paradigm of selling "design that has to look like design", i.e. by having their work's value represented as a singular expression. By contrast, great architecture has a uniqueness that comes about from a synthesis of adaptive factors. Such creations are genuine works of art.

When designers are confronted by complexity, they feel they are not in control of the design process, and fear for their reputation. After understanding the mechanisms of systems and subsystems, architects have to convince the user of the long-term value of adaptive design. That's the opposite of the "flash in the pan" value of singular, high-profile projects that turn out to fail after only a short while, or which are impossible to maintain. Those architects of singular buildings manage to convince clients of the prestige value of non-adaptive designs: both groups are selling the ideas behind the design method. The solid scientific work referred to in this paper provides the underlying support system for adaptive design that will make its application possible.

Healthy fractal control of development

A developer who wishes to create living environments can apply guidelines derived from an analogy with fractal structure.

The ideas behind the Fractal city laid out in this paper can help in one of the most pressing problems challenging city building today. That is when the Developer city joins forces with the Anti-network city. The diagnosis of the problem is the following:

1) Existing traditional settlements are not respected, either in architectural language, local materials, or human-scale urban street layout;
2) Priority is unthinkingly given to the industrial-modernist alternatives, both in the architectural and urban arenas;
3) The local government has given up trying to protect its heritage, and just lets global financing do whatever it wishes;
4) Even though authorities and dominant culture may identify and protect monuments, they do not recognize as architectural heritage the value of more modest buildings and complex landscape responsible for urbanity.

These problems have many origins, but some of them can be identified with the loss of a fractal distribution in the funding, the physical components of a project, the design elements, etc. Recall that a mathematical fractal contains many different scales, with their distribution skewed towards the smaller scales. Loss of fractality is characterized by the elimination of the smaller scales to leave only the largest scale. (This change happens dramatically in an electrocardiogram just before the onset of a heart attack! The normally complex signal suddenly simplifies to a single frequency.) In the urban context, so-called liberalization of urban construction simply opened the door to a new colonialism – an aesthetic hegemony – by global firms. Existing legal impediments to destructive urban re-structuring, which restricted construction to using the local vernacular, were lifted.

Placing all emphasis on the largest scale destroys the smaller and intermediate scales in a complex system, and fractality is then lost. At the same time, systemic stability is also lost because the scaling distribution becomes unbalanced. Normally, the many small and intermediate scales support but also hold the fewer large scales in check. Changing the natural balance among distinct scales switches the dynamics of the system from equilibrium towards

unsustainable growth or collapse. The result today is to give free reign to real estate developers backed by immense global capital; allow speculative building that consumes green land; permit the destruction of perfectly sound historical urban fabric to be replaced with cheaply-built industrial-modernist blocks, etc.

What is clearly needed is a new set of planning instruments that implement local control based on human-scale urbanism. Those constraints could resurrect legislation that was applied in the past to maintain a dynamic balance, now unfortunately lost. Globally-uniform architecture lobbies inflict most of the damage, and those can only be controlled at the local level by strong government oversight. Powerful entities such as shopping-mall networks have been lobbying to bury existing planning and tax-based legislation that protects existing smaller-scale retail. Local business – comprising multiple smaller and intermediate-scale components in the urban fractal – will obviously be obliterated by new shopping malls going up everywhere. The government leaves the smaller actors defenseless, so those cannot join together to oppose the all-powerful external developer.

The decomposition of complex systems

Even though practitioners don't normally look at the underlying science of cities, they obtain widely different results according to how they conceive of a city's components.

A city is a highly complex system, as was first recognized by Christopher Alexander [1] and Jane Jacobs [12]. We can approximate the process of generating adaptive complexity as seen in historical urban fabric, but that is impossible if we wish to design that degree of complexity top-down and all-at-once.

I have developed the theory of urban system decomposition elsewhere (see [36, Chapters: 5 and 7]). A general technique for understanding any complex system is to decompose it into separate subsystems, in a theoretical exercise that is not an actual, physical decomposition. The present paper decomposes the complex living city into eight distinct types. Each city type represents an individual complex system. It's far easier to study the mechanisms/subsystems inherent in each of the types separately. After we understand the city types better, we can focus on combining them and interpreting the added (emergent) complexity resulting from their interactions.

The most widespread decomposition practiced today, however, is the worst of all: separating the city into mono-functional zones (industrial, commercial, residential, etc.) connected by high-use roads, and conceiving the city in terms of block buildings repeating monotonously on a plan. On the other hand, the best decomposition for accommodating the human scale is into subsystems such as the Network city (paths and roads connected on all scales, plus the network of linked urban spaces), and the Fractal city (the ensemble of all built structures existing in a balanced hierarchy of scales).

Declaring that the post-war program of urban re-structuring according to single uses was implemented with "good intentions" does not excuse the poor quality of the results. Unhealthy living environments that came from mixing industrial activities with human residential areas

gave rise to zoning standards, but did not demand total and unconditional separation to improve the situation. The problem of healthy versus unhealthy mixing is a complex one that requires a focused plan of attack. The cumulative damage that industrial modernism inflicted on the complex urban system includes:

1) A highly simplistic partitioning of the city into sub-parts (zones of a single function) that are themselves not functioning subsystems;
2) Physically separating the sub-parts, which kills the living system;
3) Selective suppression of several sub-parts that failed to satisfy ideological/political criteria;
4) Legislating codes that encouraged the developer and construction industries to implement those divisive changes on a massive scale.

We could alternatively decompose a city in terms of loosely-connected neighborhoods that individually contain essential urban life. Each neighborhood needs to be a viable Nourishing-physical + Fractal + Network city by itself (see [40]). A neighborhood is working if it has more internal than external connections; otherwise it is merely an assemblage of buildings (like a dormitory suburb) and not a subsystem. Those urban subsystems then connect into the larger-scale and more complex city (see [36]).

New traditional urbanism and "form-based" codes

A growing number of developers have discovered the life-enhancing qualities of traditional methods of design, and have successfully applied them in their projects.

In several places around the world, new traditional architectural and urban codes have recently been legislated into practice. "Form-based" codes constrain commercial developers to build urban fabric that obeys the human scale. Codes specify dimensional and relational constraints that reproduce older adaptive solutions specific to that location. Those solutions (derived from traditional urban forms) have been tried out and perfected over generations. Building according to such "form-based" codes generates a city that cradles life on all scales; hence it is reminiscent of older urban fabric (see [10, 15]; and my own chapters: [30, 36])

Form-based codes are necessary but not sufficient to guarantee a living environment. You need all of the multiple contributions of the eight-fold model. Places that have form-based codes as part of their urban plan still need to decide where the functional aspects should be concentrated. Both the positioning and dimensions of commercial nodes (e.g. retail areas adjacent to transit stations) need to be right for that region to function. Otherwise, visually-attractive schemes are a hit-and-miss situation with empty commercial spaces because those are too large for the area, or because they are not centrally located to draw from the largest number of pedestrians.

Adaptive environments necessarily bear a strong morphological resemblance to older cities, because the same basic urban genetic material is being re-used. Individuals who crave superficial innovation at any cost fight new traditional urban development out of ideology. Economic forces nevertheless keep the movement alive. As promised by New

Traditional Urbanism's promoters, the market value of the product is far higher than either conventional industrial modernism or suburban sprawl. Using form-based codes is much more lucrative to the commercial sector, which in turn can influence the local government to adopt them.

It is surprising how much good codes (when adopted by a city or region) can do to improve the living qualities of urban fabric over time. For example, a very simple observation by Christopher Alexander on the relative proportions of different areas according to designated use actually distinguishes living from dead city [3]. He measured a majority of today's urban regions to typically have the rough percentage distribution of areas devoted to pedestrians-green-buildings-cars as 2%–28%–23–47%. In contrast to this, Alexander found that living urban regions have a very different percentage distribution of 17%–29%–27%–27%. The area devoted to green is about the same, but the pedestrian/car ratio is vastly different. Therefore, even a broad requirement that brings the actual percentages closer to the second case will improve things.

Finally, "form-based" codes institutionalize the human scale in urban design and planning. This point is psychologically important because insurance companies and financial institutions are reluctant to insure or finance something new. They automatically support tower blocks and suburban sprawl because all their offices and agents have been doing this for decades. That mindset is fixed in a set procedure so that even when natural disasters wipe out monotonous office towers and vast areas of sprawl, the governing process does not permit them being rebuilt as humanly-adaptive compact city. An opportunity to finally reconfigure our cities for the better will be missed unless the governing authority perceives reconstructing in an adaptive manner to be advantageous and bureaucratically "safe".

Computing the complexity of city form

A theoretical lesson to generating an adaptive environment is that it necessarily comes from a sequence of design steps with feedback, but never all at once.

Criteria for judging the adaptability of cities are intimately linked to the complexity of urban structure. The traditional city reveals an incredible complexity that evolved from optimizing human use and adaptivity. A city that evolved to provide a "good" environment for its users built up a particular type of adaptive complexity. Understanding how to generate the correct, adaptive type of urban complexity will automatically lead to the preferred human-scale environment.

Can we reproduce in our industrial age the human living qualities of older, historical urban fabric, which evolved through multiple interventions over centuries? We should work to reproduce the positive human dimensions (i.e. feelings, sensations, belonging), and not simply the forms. We can learn to innovate with modern materials, means, and construction methods while focusing on human adaptivity. But industrial-modernist architectural and urban culture is mired in a deep belief in industrial production, and builds everything at once as a singular action. That turns out not to work without key adjustments.

A better alternative is to analyze the evolutionary development process as a sequence of adaptive computations. We can thus mimic the evolution of urban form towards adaptivity and sustainability before building it. The process is incremental, but there is no need to let it run for a lifetime. Overall forms evolve in computational time (see my chapter [39]). Most of the computations can be done virtually, with a few left open to implement adjustments during actual construction. Developers and builders rarely compute adaptations to the local climate, culture, and surroundings, and discourage the prospective owner from asking for any adaptive changes. There now exists an effective ban on plasticity of the design during the building process, since contractors demand exorbitant fees for changing anything during construction.

Evolved, computed complexity is mathematically distinct from any possible formal complexity imposed by planners on city form. An adaptive design results from a sequence of step-wise computations, each step interacting with what already exists on the ground and with the previous steps of the computation itself. Clearly, by bringing in all the countless factors and forces that shape the human environment, the resulting geometry will be highly complex. That complexity will not be random. Contrast this adaptive procedure to just building evenly-spaced blocks: that formal solution requires no computation at all!

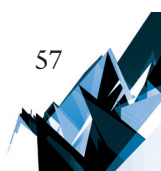
Suburban tract houses are not much better than repeating blocks as far as computation is concerned. Developers repeat some minimally-computed house, even though that model may be totally inappropriate to the existing setting. *It has been computed for a different place.* Most serious is that any scale larger than the individual house is not computed at all. There is neither interactive adaptation of the space between houses, nor of any house clustering and neighborhood configurations. Those spaces remain either simplistic or arbitrary. There is in fact no civic realm in suburban sprawl, as the scale of design is fixed at the single house. Suburbia's urban space adds up to a giant parking lot, which is what the streets eventually turn into.

Gaps in scaling where nothing is defined – for example, between the size of a single house and the entire suburban subdivision – negate any degree of urban coherence. Its skewed spectrum of scales condemns the simplistic geometry to become inhuman sprawl. Using a scaling rule developed by Alexander and myself [3, 36], Anssi Joutsiniemi classified urban forms in a very precise sequence of sizes. His work adapts an earlier hierarchy proposed by Constantinos Doxiadis, while confirming with incredible accuracy the scaling factor I had proposed as approximately $e \approx 2.7$ [13]. The point is that new projects that omit or erase the necessary smaller and intermediate scales create an Inhuman city.

Enriching the urban experience through "fractal loading"

A basic experiential human need is to be exposed to stimuli on many different spatial and time scales, which coordinate with each other in the same way as the components of a fractal.

Whenever the urban fabric and architectural environment promote interactions on a spectrum of time scales, then the phenomenon of "fractal loading" can occur. Activities



and tasks on different time scales superimpose to make a complex positive interaction that is more pleasurable than performing each task separately (see Chapter 7 of [36]). In the process of "fractal loading", the smaller scales load onto the larger carrier scale to create a fractal containing many scales together. The urban experience then resembles a symphonic combination where urban movement is tied to complex though nourishing information exchanges. A higher temporal scale – taking time to accomplish the main task – then also carries with it exchanges on many smaller temporal scales. For example, walking to an appointment along a street lined with interesting architectural elements and shop windows adds pleasure (from a multitude of short-term inputs) to a simple task.

A fractal in time composed of overlapping periods is a new concept to mainstream urbanism, which focuses primarily on the large-scale geometry, and tends to ignore time processes. Yet urban processes occur on a wide spectrum of temporal scales, from 1 sec to 50 years, with all the intermediate periods coordinated and supporting each other. Good urban planning does this, embedding flexibility to accommodate change in the longer time span.

After several decades of ignoring the temporal fractality necessary for vibrant human life, the intimate interaction between spatial and temporal fractalities is no longer understood. But unless the physical city is shaped according to an accommodating fractal spatial morphology, it cannot contain and encourage the actions that need to take place over the spectrum of temporal scales. It can never be truly alive. Businesses need to be complemented by residential and other uses to guarantee people coming and going at all times. By suppressing fractal structure, a city allows only movements and actions on a highly restricted range of periods. This is probably one of the most important, yet unknown, aspects of why traditional built environments work.

The city that privileges automobiles creates a non-fractal structure, because it predominantly contains lengths only appropriate to vehicular travel. The car city erases the livable human range of scales, so that the built environment broadcasts the visual message that pedestrians are out of place there. The older city centers of Barcelona, Istanbul, Marrakesh, and Vienna have an essentially fractal structure, having being built before anti-fractal urban design styles became popular. One of the key reasons that visitors find them so attractive is because they embody the nourishing effects of a Fractal city. Not only tourists but also local residents find those places highly enjoyable. They find the multiplicity of scales needed to fully engage with the environment.

Pathological city types

If ours is an intelligent society, it must finally recognize that repeated application of the same failed typologies will inevitably lead to the same bad result.

In this paper I'm trying to show the reader that most of what we dislike in today's city was intentionally conceived, and is an inevitable consequence of deliberate decisions taken by our society. Changing one part of the system while ignoring the others is not going to be effective. It is impossible even to contemplate reform until we understand how the dominant

construction/financial system shapes our cities by creating unhealthy city types. Two city types cause most of the damage in preventing an adaptive urban fabric and human-scale built environment:

TABLE: TWO PATHOLOGICAL CITY TYPES
ANTI-NETWORK CITY
INHUMAN CITY

The Anti-network city situates buildings according to a plan that looks neat from the air, but ignores how those buildings are connected. That's what happened with mono-functional zoning, which assumes that all connections are going to be long-distance ones. This extreme simplification of urban function implies a concomitant simplification of urban form. It was due to a conviction that the built environment should henceforth be designed intellectually, according to geometrical ordering that was presumably more "scientific". But the urban fabric cannot form living complexity under mono-functional zoning, detached stand-alone buildings, or neat geometric repetition. There is a simple solution to this: legislate mixed-use zoning and other elements of neo-traditional urbanism.

The Developer city is not included in the above table of "Pathological City Types". Despite the countless manifestations of Developer + Inhuman city built by commercial developers and governments, we do have examples of enlightened developers creating new traditional urban fabric that ranges from being acceptable, to highly attractive. The result depends on the codes and typologies that are adopted, which offers sufficient cause for optimism. A developer or government could decide to implement the design rules outlined in this paper to create a wonderful human environment in a Developer + Nourishing-physical city.

Designing for a total dependency on individual motor vehicles makes the isolating nature of everyday life unavoidable. The user lives in an environment that is socially limited compared to a Network city that includes pedestrian life. It is possible – indeed, one has no alternative to this – to drive to all functions and destinations, without interacting with any human being physically along the way. Human impatience and predisposition to use the easiest means of transport chooses to drive whenever possible. In the Anti-network city, people live in bubbles, with direct human interactions only when they get out of their car. Socializing is part of being human, but that component of life is now severely restrained by our built environment.

The Anti-network city turns over all connections to the automobile, which prevents the smaller network scales and signals the death of pedestrian connectivity. One of the original commercial purposes behind mono-functional zoning – to sell more cars – was introduced by Le Corbusier working for the auto manufacturer Gabriel Voisin. Mono-functional zoning was adopted by US car manufacturers, who bought up and then dismantled the existing tramways. (This is documented in the 1947 Federal lawsuit "United States versus National City Lines Inc.") Eliminating an efficient system of light rail transport forced everyone to depend upon private automobiles. This action boosted the car and petroleum industries towards total economic dominance.

People's love affair with their cars, and the tremendous personal freedom they confer, destroyed the pedestrian city. This is the key to what happened. Dominant culture aligned itself with the forces of the automotive and petroleum industries to drastically re-shape our cities. Everyone without exception thought that living in the countryside away from the city and its problems was a great idea, and supported it enthusiastically. We ourselves did that, and have no one else to blame.

The Inhuman city is implemented by a well-meaning government that is convinced by the spurious claims of economy represented by block housing. Monotonously repeating buildings provide the model for a significant portion of post-war construction. Whether state-sponsored or built by developers with the approval/support of local authorities, this model of living isolates individuals. It also wastes the space between the buildings, which is really unsuitable for human use (for geometrical reasons outlined in a previous section). The same holds true for endlessly repeating cookie-cutter houses in suburban sprawl: living qualities in the urban ensemble – outside one's door – are incredibly poor.

In the industrial-modernist dream, the new society supposedly needed an entirely new form of the city. Architects put people into monstrous blocks, and then repeated those blocks indefinitely. Coming easily from industrially-inspired Russian Constructivism (born in a culture where people were a class of things that simply needed to be housed), the clean lines and simplicity appealed to architects as the perfect new expression for a "modern" world. Wishful ideas about roof gardens, concrete playgrounds, and a commercial fourth storey were implemented but failed miserably because their architects misunderstood how people actually live and interact. Those failures, incredibly, became part of design canon. The basic problems here are mathematical: monotonous repetition; elimination of fractal scales; severing connections; reversing the definition of urban space, etc. We may attribute all those epic errors to an obsession with reversing traditional urban morphology.

Skyscrapers

Despite the truly enormous profits to be made from building vertically, the city suffers as a result of hyper-concentration.

Our cities are under serious threat from skyscraper proliferation. Living urban fabric is being destroyed by the insertion of skyscrapers (see [25]). The problem with skyscrapers is that they exploit but never attempt to work with the complex system that supports them. They are conceived on the basis of detached static geometry, imposing a single oversize scale that is the antithesis of our system conception of multiple cooperating scales and networks. Skyscrapers ignore mechanisms such as networks, connectivity, biophilia, and fractals on the human scale. Their single dimension of flow is vertical: straight up, leading nowhere and connecting to nothing on top.

In the eight-fold classification, skyscrapers belong to the Developer + Anti-network + Inhuman city, a bad combination. Skyscrapers are linked to huge profits for a few persons who build them, to the detriment of a city's life. They are a pure creation of the industrial-modernist

city, whose glittering utopian designs go back to the 1920s. Vertical stacking removes people from the urban web on the ground, thus severing social contact and preventing the random encounters that generate life in the city [12, 36]. A century of failed attempts has proved that social fabric cannot grow vertically after achieving optimal density (as in traditional settlements).

Perfectly sound high-density building stock of modest height in the city's center is demolished. Its buildings could have important historical interest because they could never be duplicated today: the workmanship no longer exists and it would be far too expensive. Current residents are either moved elsewhere, or promised a flat in the new skyscraper. The promise does not simply comprise a new and maybe bigger living space; it offers a futuristic world of technology, steel and glass, a new lifestyle linked to the latest wonders. But what the developer never admits is that all living and social connections are permanently severed. The former life on the ground can never be resurrected, despite all the wildly optimistic promises.

A skyscraper or cluster of skyscrapers concentrates density by several orders of magnitude. The rest of the city remains the same, even if it is functioning properly (although in many cases it is already severely strained). Existing networks, infrastructure, and distribution of urban spaces are unchanged but now become overloaded. How does the city accommodate the vastly increased network needs of the skyscrapers? It cannot and doesn't, because the added pressure overwhelms the existing system. The skyscraper is an architectural fantasy, created and built as if the rest of the city did not exist. Its drastic overloading effects are never balanced.

Skyscrapers generously provided with storeys devoted to parking are convenient, but are at the same time disconnected from the urban fabric. The inhabitants go from their condominium to their car in the basement parking and then out into the city. Thousands of people may live there, but seldom use the front door. All that design optimization in a skyscraper does is to stop people from participating in – and contributing to – a healthy urban fabric.

Yet architectural culture craves skyscrapers set in open spaces as symbols of prestige, progress, and prosperity. This idea has a powerful political and economic thrust, pursuing an ideology that imagines development occurring through futuristic and utopian images. The money behind skyscraper development rivals and oftentimes exceeds the power of local government to regulate it. Skyscrapers are an integral component of global finance, hence it makes little sense to criticize them using purely architectural or planning arguments.

Claims of energy efficiency tied to certain types of building certification are a nice marketing ploy for the skyscraper industry. As Michael Mehaffy and I have pointed out, glass-and-steel skyscrapers fail miserably in energy savings, despite what one reads in the media (see [31], Chapter 2: *Why Green Often Isn't*). Our analysis of why the industrial-modernist skyscraper typology is not resilient predates the eight-fold model of this paper, yet is consistent with the incompatibility between skyscrapers and the Network city. A super-tall building cannot be a resilient component of the city because it doesn't connect with the complex urban system in a mutually-beneficial manner.

Part 8: Reasons behind the Inhuman city

Designing the auto-dependent landscape

We demand the freedom of mobility offered by the car, but have turned a blind eye to the momentous consequences this has had upon architecture and urban morphology.

Much thought and intelligence was applied after the 1920s to optimize the city for rapid vehicular movement. Along with the multiplication of cars and trucks came the expansion of urban components needed by the auto-dependent landscape: gasoline stations, car parks, garages, car dealerships, car washes, strip malls, giant surface parking lots surrounding big-box stores and commercial malls, etc. All of these require a lot of space, and succeeded in eliminating the human being from the physical city to create an Inhuman city (which is no longer a walkable city). Generations of planners did not reflect on the possibility that the auto-dependent landscape would replace the Nourishing-physical city (see [30]).

Now we know better and recognize the harm in the results. The street grid and building plots, however, have been fixed for decades to favor the auto-dependent landscape. In many places around the world highways and open parking lots define urban morphology. The very difficult task to correct this in order to bring back the human scale has already begun. Massive urban re-structuring to re-route roads and cut up superblocks again into smaller plot sizes is underway in many cites.

Even though the auto-dependent landscape is self-perpetuating (build on cheap agricultural land just outside the city, which spawns more cars, etc.), commercial developers have discovered lately that human beings still prefer a human-scale environment. The tremendous success of retrofitting urban pedestrian zones that compete with indoor malls has reversed a decades-long trend. We are now poised to begin to build the Nourishing-physical city once again. Hopefully, cities in the developing world that are ready to bulldoze their nicest human-scale environments (copying dismal planning mistakes from 50 years ago) will learn from this experience and work to retain these places.

But the change in the way human beings interact with the built environment drove a drastic reversal of architectural design. Detail, ornament, and structural coherence on the human scale are not experienced from a car, and they became irrelevant. Speed dematerializes the world. What make the greatest impact now are large-scale forms and flashy, shiny structures to draw our attention from a distance as we drive by them. Formal planning encourages the industrial-modernist architectural style, and gave birth to the built urban fabric we inhabit today. Commercial advertising jumped up in scale from modest lettering to huge signs all over the landscape, creating a visual cacophony that competes for our momentary attention.

This is the fundamental and unexpected change triggered by the dominance of the auto-dependent landscape. Intimate contact between people and the built urban fabric was disregarded, because we are enclosed in our car interior for most of the time. Nourishing urban space is irrelevant in a city whose primary purpose is to facilitate fast vehicular flow. We

go from home to car to work or store, and back again without experiencing the Nourishing-physical city, so why take the trouble and expense to build it? We are fighting to implement something that has all the appearances of a nostalgic anachronism, *because we no longer need it while living inside our cars*.

A world hostile for children

By creating an environment in which children feel lost and threatened, our culture ranks with those in history that have been cruel to their children.

Researchers confirm the unhealthy qualities of the Inhuman city. The most harmful developmental effects occur with children. An industrial-minimalist world is no place for raising children [9]. From birth to their late teenage years when they finally have access to independent means of transport (i.e. young people can safely use public transport alone or drive a car), children are prisoners in their apartment or house. Children cannot connect to life in a fashionable contemporary "design" environment, and consequently suffer from a biophilia deficit [14]. Those who cannot complain (i.e. children and the elderly) are sacrificed to style.

The child's world focuses on a spectrum of sizes that is much smaller than that of an adult. The easiest place to see this is in active play spaces created by the children themselves. Children love to play in alcoves, under tables, in spaces within spaces, and in cubbies that correspond to their own physical size. The same fractal geometric qualities attract children to experience and enjoy spaces outside, and they did construct such habitats – for example, tree-houses – in past ages when children were allowed freely outside. Fractal articulations found in older urban fabric, disappear from a world of sheer walls and large abrupt structures. This deficiency tragically characterizes both interior and exterior spaces "designed" for children.

Children experience the world totally with their emotions, instinct, intuition, and senses. They are extremely sensitive to their environment. Unlike many adults, children have not (yet) been conditioned by abstractions to override their basic emotional responses. At the same time, the built environment has to physically protect them from dangers they have not yet learned to pay attention to. A city that is good for children should allow and encourage them to explore it, with conditions and restrictions. A house should join to its external urban fabric in a way that a child can safely explore its surroundings. The design profession faces a monumental task to achieve these requirements, because they are ruled out by auto-centric preferences.

The Inhuman city reveals a failure and neglect to build the Nourishing-physical city, but also a deliberate design strategy. Standard design solutions were implemented for so many years and nobody complained, or, if they did complain, the architectural and urban culture never listened to those outside their group. Both academics and practitioners are reluctant to change the way they design, and conformist cultures ignore those with a different way of thinking.

Many projects rely on tricks to get approved, yet neither the general public nor government authorities realize this; or they go along with the deception. We see gigantic building projects whose main aim is a cluster of skyscrapers in a totally unsustainable setting. (Where is the energy coming from? How do all these people connect to the rest of the city? Can the existing infrastructure handle the added strain?) Such projects are frequently coupled with a minor park presented in the media, with renderings showing smiling families and happy playing children. The real objective appears only as barely visible background on the renderings. Another trick is to have several stages of a towering project approved on the basis of the pedestrian realm and an attractive green park. Yet the first stage is always the skyscraper cluster. Once this is built, the subsequent stages are canceled, offering various excuses.

Architectural effects on the city

Architecture took a wrong turn when it adopted the machine paradigm based upon a total misunderstanding of both machines and human nature.

At the root of the contrast between the Inhuman and Nourishing-physical cities lies a program of formal conception whose supporting philosophy does not distinguish humans from machines. These two contrasting paradigms assume completely opposite architectures for their buildings and for the urban space networks those buildings reinforce. Working strictly within a formal design approach, it is very easy to remove biological qualities from our artifacts and from the built environment. People eventually get used to their absence, even though it's not good for them. When citizens lose the power to react to their environment, someone can build the Inhuman city without opposition (For a trenchant criticism of the industrial-modernist city see [19]).

An unhealthy fascination with "design purity" eliminates the Fractal city by removing everything but the largest scales, which are inadequate to define a human environment. Lacking traditional solutions for creating intermediate spaces and protective semi-permeable borders, a city becomes both deadening and dangerous because specific protective barriers are absent. Bollards, colonnades, and arcades are deemed to be "geometrically impure" because they introduce fractal structure at smaller scales. But that is precisely the point of fractal design: the presence of coordinated elements on all scales including the human scale.

Persons raised and educated to believe that what looks futuristic, industrial, and minimalist is "good" will come up with persuasive arguments for why such things should be built. This image-based design replaces older (but perfectly functional) urban fabric that only requires some regular repair. Or the exciting "look" of the superficially new and fashionable wins instead of a far more human, adaptable, and sustainable design during competitive selection. Decades of publicity about the claimed wonders of novel-looking industrial designs seduce politicians. Academics believe in the redemptive value of such designs, ignoring scientific evidence that those might be causing anxiety, psychological stress, and could be repelling people from urban spaces.

The fantasy of "modernization" by means of 1920's industrial-modernist images adds to the various forces shaping the city in a negative direction. Unfortunately, both the Inhuman city

and the Anti-network city are emotionally and perceptually linked to "progress", thus making them attractive and desired. Those two negative city types remove the human spatial and time scales. Inserting structures belonging to the Inhuman city into an existing Nourishing-physical city, in the mistaken belief of "modernization", destroys healthy components of the city in that location.

One would expect that research into this fundamental matter – which city types enhance human life, and which are unhealthy in the long term – would be the profession's priority. That's not the case, because formal industrial-modernist typologies are entrenched in the contemporary design process, leaving little possibility for reconsideration or revision. Architectural culture has invested all of its efforts over one century to try and justify those choices. It condemns healthy city types by saying that they look too "old-fashioned", while it praises unhealthy types that look exciting and technological, never considering the possibility of harmful effects on the users.

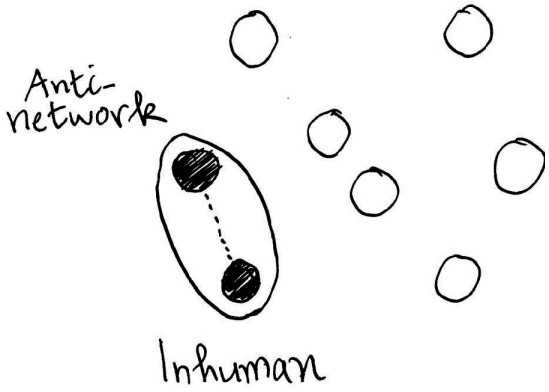


Fig. 14. Unhealthy city types add, because they share the same negative qualities

In an ironic twist, the powers that profit from building the Inhuman city have now stuck the label "sustainable" on their product. Like any deceptive marketing gimmick, this is part of the drive to continue what the industry has been doing for decades, which is to build poorly-adapted environments. But it is a clever ploy that deceives the public into believing that extremely expensive industrial materials can add up to a sustainable building. Quite the contrary, true sustainability occurs when energy wastage is minimized, both in extracting and transporting the building materials, and in the upkeep and running of the urban fabric (see [20]). Whenever you see a building with the "high-tech" look, don't believe its optimistic claims for sustainability (see [31]).

An inadequate education system

New architecture students taught science-based design information, such as is contained in the eight-fold model, would be the generation that can have the greatest effect.

What hope is there for more human (and genuinely sustainable) cities in the future? That depends entirely on the new generation of designers. Are schools teaching the tools that will help students understand how a city works, and how its morphology encourages life? We seem to be stuck in the sculptural paradigm, which is a hit-and-miss proposition. We continue to teach failed industrial-modernist urban typologies simply out of inertia and resistance to change. Because of their antiquated education, urban designers are influencing the way things are actually built less and less. In practice, market forces or the government do exactly what they want.

A major problem with today's design education is that it distances students from analytical thought processes, substituting them by so-called free creativity. Students are made to spend all of their time making miniature models and designs, without having developed any pragmatic basis for evaluating their models. That's simply not part of the current curriculum. It is now standard practice to propose a form, choose it in a competition, and then build it, without ever evaluating how it will affect users. Even though practicing architects are often contractually obligated to perform evaluations, students are misleadingly taught never to be concerned to evaluate a building or urban space before or after it is built.

There are in fact three fundamental criteria for judgment that students must learn to become more responsible practitioners:

1. A theoretical basis of which urban geometries work and which don't;
2. Methods for the mathematical analysis of urban form;
3. Techniques for experimental evaluation (before implementation) of how a design will probably affect the user after it is built.

A fascination with the strikingly "new" can have seriously negative consequences for both the profession and users. Those who enjoy this type of visual innovation don't care to investigate whether or not it provides a human living environment. To justify their personal aesthetic preference, academics invent "explanations" for monstrous, inhuman buildings and urban forms, and architectural critics promote them with well-written propaganda. To boost the novelty approach, exquisitely adapted vernacular/indigenous architecture is fervently and methodically debased. This practice is not only dishonest, but it has, over the years, eroded the honesty of the discipline itself (see [26]).

Nonsensical statements from individual architects justifying their own non-adaptive buildings and urban projects are taught as "architectural theory" to students. Architecture students, not unreasonably, trust their professors, even when their professors feed them marketing hype instead of tested knowledge. Consequently, the science of design has become progressively buried under a cloud of jargon and untruths. Absent the normal criteria of verification, discussions turn to cult authority and ideology for support. This has the deplorable effect of marginalizing the small practitioner, who is forced to conform to the wave of fashion, even though that's very poor architecture.

How do we fix this? My goal in this paper is to build living cities, not to reform irresponsible architectural education. We need massive curriculum re-structuring and a re-orientation of educational values. This process can only be market-driven; that is, the public has to demand human-scale cities, and then government planners, regulatory boards, developers, contractors, and construction companies will have to follow the popular demand. The last

institution to be affected by this change is going to be the schools, which might be obliged to abandon teaching useless and even harmful ideas, and replace them with tested results such as those presented in the eight-fold model.

Conclusion

We now have the design knowledge to build a wonderfully adaptive living environment, once we abandon prejudices and techniques that don't work well.

This paper introduced eight abstract city types and discussed their interaction and opposition in different contexts. Step towards controlling the forces actually responsible for the cities we live in were outlined. Even the most powerful forces shaping urban form could be directed towards generating a more human environment, once the majority of urban players understand the advantage of doing so. By contrast, those same forces can destroy city form (that is, create an inhuman and unhealthy city) if we judge a rendering only by its imageability, ignoring the possible consequences of the final result for human life.

I suggest a three-step path toward cities better adapted to human uses and sensibilities:

1. Research: discover scientific reasons for city form and urban processes;
2. Education: learn from facts and do not be misled either by ideology or by special interests;
3. Application: convince decision-makers to build human-scale cities and resist the allure of fashion.

The few recent examples of where this program was implemented successfully are all traditional. Those projects were commercially-driven, and turned out to make large profits for their investors. Small-scale developers have built the best examples. Traditional architectural forms were employed together with form-based urban codes, to great success. All that was needed was to extract the form-based codes from existing living urban fabric. After an initial reluctance of government permitting boards, those innovative projects went through. Much more resistance came from academic architectural culture, which mounted a massive effort to discredit neo-traditional developments. It pushed instead its own newly-disguised industrial modernism in a green setting, deceptively re-labeled as "Landscape Urbanism".

Finally, we need to confront the massive cultural forces that drive a city to conform to abstract images. That, in fact, is how cities and city regions are built whenever large money/power interests drive speculative construction. Recognizing those forces and re-directing them towards a more adaptive and healthy built environment is a matter of life and death for cities. And why should we consider opposing this trend? Because at the time of writing, mainstream urbanism is following an image-based and unscientific conception of the city, and the Inhuman city is unhealthy for its inhabitants! At the same time, such cities are unsustainable, and represent ticking time-bombs that will become unusable because they are too expensive to run.

It is a pleasure to thank Michael Imbimbo, Tomasz Jelenski, Agata Kantarek, Kenneth G. Masden, John Osten, and Adolf Sotoca for their helpful comments on early drafts of this paper.

References

A. Articles and books by other authors

[1] Alexander Ch., *A City is Not a Tree*, Architectural Forum, Vol. 122, No. 1, 1965, 58–61 and No. 2, 58–62; Reprinted in: *Design After Modernism*, John Thackara (Ed.), Thames and Hudson, London 1988, 67–84; http://www.rudi.net/books/200

[2] Alexander Ch., *The Timeless Way of Building*, Oxford University Press, New York 1979.

[3] Alexander Ch., *The Nature of Order, Books 1–4*, Center for Environmental Structure, Berkeley, California: *Book 1: The Phenomenon of Life*, 2001; *Book 2: The Process of Creating Life*, 2002; *Book 3: A Vision of a Living World*, 2005; *Book 4: The Luminous Ground*, 2004.

[4] Alexander Ch., Ishikawa S., Silverstein M., Jacobson M., Fiksdahl-King I., Angel S., *A Pattern Language*, Oxford University Press, New York 1977.

[5] Alexander Ch., Neis H-J., Alexander M.M. *The Battle for the Life and Beauty of the Earth*, Oxford University Press, New York 2012.

[6] Alexander Ch., Silverstein M., Angel S., Ishikawa S., Abrams D., *The Oregon Experiment*, Oxford University Press, New York 1975.

[7] Asquith L., Vellinga M. (Eds.), *Vernacular Architecture in the Twenty-First Century*, Taylor & Francis, Abingdon, Oxford 2006.

[8] Batty M., Longley P., *Fractal Cities*, Academic Press, London 1994.

[9] Day Ch., Midbjer A., *Environment and Children*, Elsevier, Oxford 2007.

[10] Duany A., Plater-Zyberk E., Speck J., *Suburban Nation*, North Point Press, New York 2000.

[11] Gehl J., *Life Between Buildings*, Van Nostrand Reinhold, New York 1987.

[12] Jacobs J., *The Death and Life of Great American Cities*, Vintage Books, New York 1961.

[13] Joutsiniemi A., *Becoming Metapolis – A Configurational Approach*, Datutop 32, Tempere University of Technology, Tampere, Finland 2010.

[14] Kellert S.R., Heerwagen J., Mador M. (Eds.), *Biophilic Design: The Theory, Science and Practice of Bringing Buildings to Life*, John Wiley, New York 2008.

[15] Krier L. *Architecture: Choice or Fate*, Andreas Papadakis, Windsor, England 1998; New edition, *The Architecture of Community*, Island Press, Washington 2009.

[16] Leitner H., *Pattern Theory*, Create Space, Amazon, 2015.

[17] Louf R. Barthelemy M., *A typology of street patterns*, Journal of the Royal Society Interface, 8 October 2014, http://rsif.royalsocietypublishing.org/content/11/101/20140924

[18] Lynch K., *The Image of the City*, MIT Press, Cambridge, Massachusetts 1960.

[19] Millais M., *Exploding the Myths of Modern Architecture*, Frances Lincoln Limited, London 2009.

[20] Mouzon S.A., *The Original Green*, Guild Foundation Press, Miami, Florida 2010.

[21] Salat S., *Les Villes et Les Formes*, Hermann, Paris 2011.

[22] Turner J.F.C., *Housing By People*, Marion Boyars, London 1976.

[23] Whyte W., *The Social Life of Small Urban Spaces*, The Conservation Foundation, Washington 1980.

[24] Wilson E.O., *On Human Nature*, Harvard University Press, Cambridge, Massachusetts 1978.

B. My own articles and books

Since this essay is a transcription of conference lecture notes, it contains mainly references to my own work that is freely accessible on the Internet. Those articles and books in turn contain exhaustive bibliographies and references to the numerous authors whose research contributes to a new, adaptive understanding of urban structure.

[25] *A Future Without Starchitects*, Chapter 4.3, [in:] *Kaleidoscope City: Reflections on Planning and London*, Manns J. (Ed.), RTPI Press, London 2014, 123-131, http://zeta.math.utsa.edu/~yxk833/Kaleidoscope-London.pdf

[26] *Anti-Architecture and Deconstruction*, OfftheCommonBooks/Sustasis Press, International Edition, *Vajra Publications*, 2014, http://zeta.math.utsa.edu/~yxk833/books.html

[27] *A Theory of Architecture*, OfftheCommonBooks/Sustasis Press, International Edition, *Vajra Publications*, 2006; reprinted with Index 2014, http://zeta.math.utsa.edu/~yxk833/books.html

[28] *Beauty, Life, and the Geometry of the Environment*, Chapter 2, [in:] *Reclaiming Beauty*, Volume I, Horvath A., Cuffe J.B. (Eds.), Ficino Press, Cork, Ireland, 2012, 63–103; Edited version of an essay from the *Athens Dialogues*, November 2010, http://zeta.math.utsa.edu/~yxk833/lifeandthegeometry.pdf

[29] *Biophilia and Healing Environments*, a 44-page booklet available from OfftheCommonBooks, 2015; also published free online by *Terrapin Bright Green LLC*, New York 2015, http://www.terrapinbrightgreen.com/report/biophilia-healing-environments

[30] *Compact City Replaces Sprawl*, Chapter, [in:] *Crossover: Architecture, Urbanism, Technology*, Graafland A., Kavanaugh L. (Eds.), 010 Publishers, Rotterdam 2006, 100–115, http://www.academia.edu/188203/Compact_City_Replaces_Sprawl

[31] *Design for a Living Planet*, co-authored with M.W. Mehaffy, US Edition, OfftheCommonBooks/Sustasis Press, International Edition, *Vajra Publications*, 2015; some chapters published online by *Metropolis*, http://www.metropolismag.com/Search/index.php?mod=CoreSearch&query=salingaros&sortby=datepublished&urlprefix=%2F

[32] *Geometry and Life of Urban Space*, co-authored with P. Pagliardini, Chapter, [in:] *Back to the Sense of the City, 11th Virtual City & Territory International Monograph Book*, Centre of Land Policy and Valuations (Centre de Política de Sòl i Valoracions), Barcelona 2016, 13–31, http://upcommons.upc.edu/bitstream/handle/2117/90890/CH00_CONTENTS%20INTRO_geometry.pdf

[33] *Geospatial Analysis and Living Urban Geometry*, co-authored with P. Pagliardini, S. Porta, Chapter 17, [in:] *Geospatial Analysis And Modeling of Urban Structure and Dynamics*, Jiang B., Yao X.A. (Eds.), Springer, New York 2010, 331–353, https://www.academia.edu/188975/Geospatial_Analysis_and_Living_Urban_Geometry

[34] *Growing Sustainable Suburbs: An Incremental Strategy for Reconstructing Sprawl*, co-authored with M. Mehaffy, L. Steil, Chapter 10.2, [in:] *New Urbanism & Beyond: Designing Cities for the Future*, Haas T. (Ed.), Rizzoli, New York 2008, 262–274, http://www.academia.edu/188207/Growing_Sustainable_Suburbs_An_Incremental_Strategy_for_Reconstructing_Sprawl

[35] *Neuroscience, the Natural Environment, and Building Design*, co-authored with K.G. Masden, Chapter 5, [in:] *Biophilic Design: The Theory, Science and Practice of Bringing Buildings to Life*, Kellert S.R., Heerwagen J., Mador M. (Eds.), John Wiley, New York 2008, https://www.academia.edu/188202/Neuroscience_the_Natural_Environment_and_Building_Design.

[36] *Principles of Urban Structure*, new US printing, OfftheCommonBooks/Sustasis Press; European Edition, *Techne Press*, Amsterdam 2005, 2014; Asian Edition, *Vajra Publications*, http://zeta.math.utsa.edu/~yxk833/books.html

[37] *P2P Urbanism*, book draft available for free download, 116 pages, 2011, http://zeta.math.utsa.edu/~yxk833/P2PURBANISM.pdf

[38] *The Law of Requisite Variety and the Built Environment,* Journal of Biourbanism, Vol. 4, Nos. 1&2, 2015 (published in 2016), 47–52, http://zeta.math.utsa.edu/~yxk833/TheLawofRequisiteVariety.pdf

[39] *Urbanism as Computation*, Chapter 13, [in:] *Complexity Theories Of Cities Have Come Of Age*, Portugali J., Meyer H., Stolk E., Tan E. (Eds.), Springer, Berlin 2012, 247–270, http://zeta.math.utsa.edu/~yxk833/urbanism-computation.pdf

[40] *Urban Nuclei and the Geometry of Streets: the Emergent Neighborhoods Model*, co-authored with M. Mehaffy, S. Porta, Y. Rofè, *Urban Design International*, Vol. 15, No. 1, 2010, 22–46, http://zeta.math.utsa.edu/~yxk833/UrbanNuclei.pdf

[41] Why Green Often Isn't, co-authored with M. Mehaffy, Chapter 2, [in:] *Design for a Living Planet*, 2013 (2015), US Edition, OfftheCommonBooks/Sustasis Press; International Edition, *Vajra Publications*; Chapter published online by *Metropolis*, http://www.metropolismag.com/Point-of-View/April-2013/Toward-Resilient-Architectures-2-Why-Green-Often-Isnt

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
CHEMISTRY
DOI: 10.4467/2353737XCT.17.017.6210

Barbara Michorczyk (bmichorczyk@chemia.pk.edu.pl)
Elżbieta Hędrzak
Institute of Organic Chemistry and Technology, Faculty of Chemical Engineering and Technology, Cracow University of Technology

Study of oxidative coupling of methane integrated with CO oxidation

Badanie procesu utleniającego sprzęgania metanu zintegrowanego z dopalaniem CO

Abstract

In this work, the process of OCM carried out over Mn-Na_2WO_4/SiO_2 integrated with selective oxidation over Ag/support was investigated. the effect of feed gas composition and OCM bed temperature as well as the position of Ag/support bed and additional oxygen injection before this bed were investigated. At optimal OCM conditions for the Mn-Na_2WO_4/SiO_2 catalyst (CH_4/O_2 = 3.75; V_{tot} = 77 cm^3/min; T = 780°C), the injection of additional 4 cm^3/min of oxygen into the bed of Ag/support (working at 250–300°C) leads to a preferential oxidation of CO to CO_2.

Keywords: oxidative coupling of methane, integration with oxidation, ethylene, silver catalyst

Streszczenie

W pracy zbadano proces OCM w obecności Mn-Na_2WO_4/SiO_2 zintegrowany w jednym reaktorze z procesem selektywnego utleniania CO prowadzonym na katalizatorze Ag/nośnik. Zbadano wpływ zmiany parametrów prowadzenia procesu OCM, takich jak skład surowca i temperatury oraz efekt dodatku tlenu nad złoże Ag/nośnik. Wykazano, że w optymalnej temperaturze pracy złoża OCM (CH_4/O_2 = 3,75; $V_{cal.}$ = 77 cm^3/min; T = 780°C) wprowadzenie dodatkowo 4 cm^3/min tlenu nad złoże Ag/nośnik (pracującego w temp. 250–300°C) prowadzi do preferencyjnego utleniania CO do CO_2.

Słowa kluczowe: utleniające sprzęganie metanu, integracja z utlenianiem, etylen, katalizatory srebrowe

1. Introduction

The direct conversion of methane to useful chemical intermediates, such as hydrocarbons (saturated, unsaturated and aromatic), methanol and formaldehyde, is one of the topics. Among all these direct processes, the closest one to commercialization is the oxidative coupling of methane (OCM) to ethylene and ethane [1, 2]. However, a still low per-pass conversion of methane (below 20–30%) in order to achieve high selectivity (70–80%) as well as stability of catalysts at high reaction temperatures (780–850°C) limits its industrial implementation [3].

An integration of OCM with other processes may improve the attractiveness of this process. Up to now, most attention was devoted to the integration of the OCM process with dry and steam reforming of methane [2, 4–9]. Several other possibilities for the integration of OCM with aromatization [10–12], benzene alkylation [13], pyrolysis [14], conversion to acetic acid and/or ethanol [15], synthesis Fischer-Tropscha [16], oxidative dehydrogenation of ethane [12, 17–19], methanation of CO_x [20, 21] were also explored.

An interesting issue can also be the integration of OCM with a selective oxidation of CO to CO_2. Such coupling of two exothermic reactions in one reactor (two catalytic beds in one reactor) can be attractive from the technological point of view. The combustion of CO, which is produced on an OCM bed as a by-product, may simplify the separation of the reaction mixture. The total conversion of CO to CO_2 should especially facilitate CH_4 recycling as per-pass conversion of methane that does not exceed 30%. In the case of the most promising catalysts, e.g. Mn-Na_2WO_4/SiO_2other by-products , such as CO_2 and water, it has no negative impact on the OCM process, so, in many cases, it is not required to clean the recycled methane from these gases, which can be used as a diluent [22–24]. However, in the case of integration in a single reactor, there are several problems due to the presence of many chemical compounds in the stream leaving the OCM catalyst bed, which e.g. may cause a deactivation of the CO oxidation catalyst. Moreover, hydrocarbon products and unreacted methane on a CO oxidation bed can be oxidized. Thus, the selection of a suitable catalyst is difficult because, usually, the catalysts active in the oxidation of CO also catalyze the total oxidation of hydrocarbons. An alternative solution can be the separation of ethylene and ethane from the reaction mixture before oxidation of CO to CO_2, but this solution requires an additional reactor for CO oxidation, which increases the amount of operations.

There are many catalysts active in the selective oxidation of CO to CO_2. A lot of noble metal-based catalysts, especially of the platinum group [25–29] and ones that are gold-based [29–33] as well as non-noble metal oxide-based catalysts (e.g. containing CoO, CuO and MnO_x) [28, 34–39] have been investigated. Unfortunately, most of these catalysts are active not only in the oxidation of CO to CO_2, but also in the oxidation of hydrocarbons [40]. For example, methane in the presence of a catalyst based on Pd can be oxidized even at around 300°C [41].

Our attention turned to silver-based catalysts. These catalysts have been recognized to show a relatively high activity in the oxidation of CO and a low activity in the oxidation of CH_4 [42, 43]. In this work, we have investigated the integration of the OCM process over Mn-Na_2WO_4/SiO_2 multicomponent catalyst with selective CO oxidation over Ag/

SiO_2 or Ag/Al_2O_3 in one reactor. the effects of the feed gas composition (CH_4/O_2 ratio, O_2 addition below second bed) and temperature on the product distribution and the conversion of substrates were carefully investigated.

2. Experimental

2.1. Catalyst preparation

The OCM catalyst ($Mn\text{-}Na_2WO_4/SiO_2$) was prepared by incipient wetness impregnation. Dry SiO_2 (Aldrich) was impregnated in two steps with an aqueous solution of $Mn(NO_3)_2 \times 4H_2O$ (POCh – Polish Reagents) and Na_2WO_4 (Aldrich), respectively. In a typical procedure, 1 g of silica was impregnated with 1 cm^3 of aqueous solutions containing the desirable content of each salt needed for obtaining 2 and 5 wt.% of Mn and Na_2WO_4, sequentially. After each impregnation, the materials were dried for 8 h at 120°C and calcined at 900°C for 8 h.

The catalysts of total oxidation (Ag/SiO_2 and Ag/Al_2O_3) were also prepared by impregnation. Dry SiO_2 (Aldrich) and $\gamma\text{-}Al_2O_3$ (Aldrich) supports were impregnated with aqueous solutions of $AgNO_3$ (Polish Chemical Reagents). The content of precursor in the solution was desirable for obtaining 15-wt.% of Ag. the thermal pretreatment procedures and conditions were the same as in the case of the OCM catalyst.

2.2. Catalytic performance

Catalytic tests were carried out in a flow-type tubular quartz reactor with the dimensions of (internal diameter) × (length) × (wall thickness) = 8 × 270 × 1 mm. Before the process, the catalyst placed in the reactor was heated in dry helium for 30 min at 800°C. The weight of the catalyst was 400 mg (grain size 0.2–0.3 mm) and the temperature of the process was in the range between 730 –800°C. The reactor was fed with the mixture of CH_4:O_2:He = 3.75 :1:2,8, and CH_4:O_2: He =2.5:1:4,8. The total volumetric flow rate was 77 cm^3/min in all runs. The reagents mixture was analyzed using the Agilent 6890 N gas chromatograph equipped with two columns (molecular sieve 5A for separation of CO and O_2 and Hayesep *Q* for separation of H_2, CO_2, H_2O, and hydrocarbons) and thermal conductivity detectors.

The conversion of methane (X_m) and selectivity to *i*-th product (S_i) were calculated according to the following formulas:

$$X_M = \frac{n_{M(inlet)} - n_{M(outlet)}}{n_{M(inlet)}} \cdot 100\% \qquad S_i = \frac{\sum a_i \cdot n_i}{n_{M(inlet)} - n_{M(outlet)}} \cdot 100\%$$

where:

$n_{M(inlet)}$, $n_{M(outlet)}$ – the numbers of methane moles in the inlet and the outlet of the reactor, respectively;

n_i – the number of moles of the i-th product in the outlet

a_i – the number of carbon atoms in the "*i*" product (ethane, ethene, propane and propene).

3. Results and discussion

3.1. Effect of additional bed position

Figure 1 illustrates the position of the oxidative coupling of methane ($Mn\text{-}Na_2WO_4/SiO_2$) and the CO oxidation catalyst beds in a tubular flow-type quartz reactor. the bed of CO oxidation is located below the OCM band.

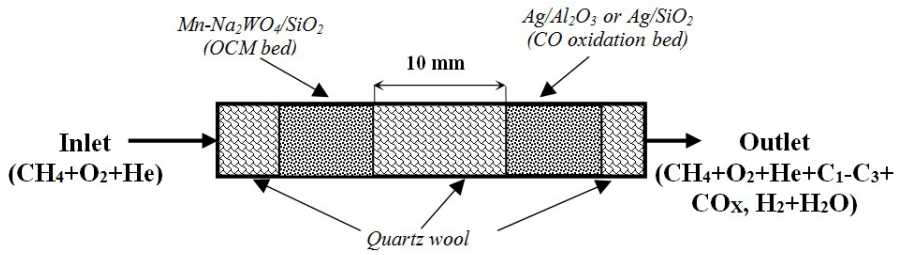


Fig. 1. Scheme illustrating $Mn\text{-}Na_2WO_4/SiO_2$ and Ag/support bed positions in the flow-type tubular reactor

Table 1 summarizes the catalytic results obtained in the presence and absence of an additional CO oxidation catalyst as a function of temperature. To clarify the effect of integration (in two separated beds), additional catalytic tests concerning the mixture of OCM and CO oxidation catalysts (in one bed) were investigated as well. A comparison of results obtained in the absence and presence of additional CO oxidation bed reveals that above 750°C the integration of OCM with CO oxidation in a separate bed exerts a small positive effect on methane conversion and selectivity to C_{2+}. In contrast, at 750°C and below this temperature, the integration exerts a significant negative influence on the product distribution. the dramatic difference in the catalytic behavior vs. temperature can be explained based on oxygen conversion over the OCM bed (in single OCM process). It is clear from table 1 that, above 750°C, oxygen reacts in 100% over the OCM bed; therefore, the second CO oxidation bed (Ag/SiO_2) has an insignificant effect on product distribution and methane conversion. the situation changes at 750 °C and 730 °C when unreacted oxygen appears in the mixture on the exit of the OCM bed. This oxygen reacts with CO and C_{2+} hydrocarbons over Ag/SiO_2 catalyst; therefore, the selectivity to CO and C_{2+} decrease significantly.

A similar explanation can be adopted for results obtained in the presence of mixed $Mn\text{-}Na_2WO_4/SiO_2$ and Ag/SiO_2 catalysts (denoted in Table 1 as $Mn\text{-}Na_2WO_4/SiO_2$ +Ag/SiO_2). Because Ag/SiO_2 catalyst is mixed with $Mn\text{-}Na_2WO_4/SiO_2$, it has permanent contact with oxygen at any investigated temperature; therefore, the selectivity to C_{2+} hydrocarbons in the presence of mixed catalysts (in one bed) is always lower than in the presence single $Mn\text{-}Na_2WO_4/SiO_2$ catalyst or in the integrated process carried out above 750°C.

The above results indicate that the reaction conditions optimal for the OCM process are not optimal for selective CO oxidation over Ag/SiO_2 catalyst. In the integrated process,

Table 1. Catalytic performance in oxidative coupling of methane process carried out in the presence and absence of additional oxidation catalyst bed

OCM catalyst (CO oxidation catalyst)	Set temperature of oven [°C]	Conversion of CH_4 [%]	Conversion of O_2 [%]	Selectivity to C_{2+} [%]	Selectivity to CO_2 [%]	Selectivity to CO [%]	Ratio of C_2H_4/C_2H_6 [mol/mol]	Temperature in OCM bed [°C][b]	Temperature in CO oxidation bed [°C][c]
Mn-Na_2WO_4/SiO_2	800	35,9	99,9	59,2	28,9	11,8	1,9	840–759	–
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2)		38,4	99,8	60,4	27,8	11,6	2,0	841–758	693–569
Mn-Na_2WO_4/SiO_2 + Ag/SiO_2		30,5	99,9	47,6	43,5	8,7	1,0	840–757	670–533
Mn-Na_2WO_4/SiO_2	780	35,8	99,9	59,8	31,1	8,9	1,8	821–742	–
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2)		37,4	98,4	60,5	30,2	9,1	1,9	822–740	673–550
Mn-Na_2WO_4/SiO_2 + Ag/SiO_2		30,3	99,9	41,4	48,8	9,6	0,7	825–738	654–520
Mn-Na_2WO_4/SiO_2	750	25,2	69,4	59,9	29,0	10,8	1,2	775–716	–
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2)		27,6	99,5	37,8	55,7	6,0	0,6	775–728	669–540
Mn-Na_2WO_4/SiO_2 + Ag/SiO_2		30,7	99,9	43,96	49,4	6,2	0,6	794–707	623–494
Mn-Na_2WO_4/SiO_2	730	13,4	38,3	55,7	32,3	11,8	0,6	748–694	
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2)		22,0	96,5	21,6	74,0	4,1	0,3	747–714	667–526
Mn-Na_2WO_4/SiO_2 + Ag/SiO_2		28,7	99,9	44,1	50,7	4,7	0,6	774–690	607–480

[a] Reaction conditions: Molar ratio of CH_4:O_2:He = 2,5:1: 2,8; V_{total} = 77 $cm^3 \cdot min^{-1}$, $m_{OCM\ cat.}$ = 400 mg, $m_{CO\ oxi.\ cat.}$ = 400 mg, Distance between catalyst beds 10 mm; Reaction time 2 h.

[b] Temperature measured on top and end of Mn-Na_2WO_4/SiO_2 bed (length of bed was 12 mm).

[c] Temperature measured on top and end of Ag/SiO_2 bed (length of bed was 12 mm).

the Ag/SiO_2 catalyst participates not only in CO oxidation, but also in the total oxidation of hydrocarbons. the range of temperature between 730–800°C is too high for selective CO oxidation.

Assuming initial catalytic results in the next parts of the experiments, we modified the conditions of the second CO oxidation bed by changing the positions of the Ag/SiO_2 bed (change of temperature) or by adding oxygen after the OCM bed (feed gas modification). the modifications are schematically presented in Figure 2.

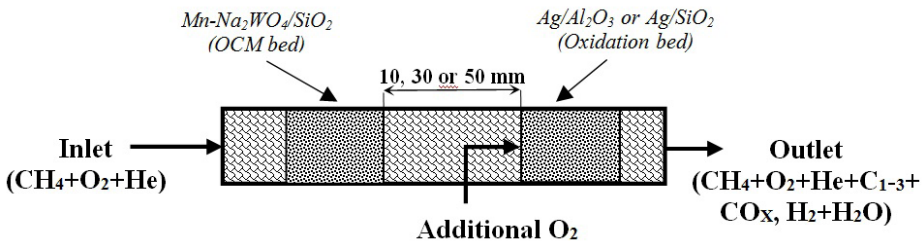


Fig. 2. Scheme illustrates Mn-Na2WO4/SiO2 and Ag/support bed positions in the flow-type tubular reactor as well as modifications in the bed position and feed gas composition injected into the second bed zone

The results obtained after the modifications are reported in Table 2. the process of OCM in the presence and absence of CO oxidation bed was carried out using a feed mixture with CH_4/O_2 = 3.8 molar ratio at two different temperatures. In selected cases, a silver catalyst supported with γ-Al_2O_3 was investigated as well. It is clear from Table 2 that the modification of the second bed position gives an insignificant effect when O_2 is converted in 100% over the OCM bed (e.g. at 780°C). The integration effect can be improved either by decreasing CO oxidation bed temperature or by O_2 co-feeding. In most cases, a change of the oxidation bed position (decrease of the temperature) leads to the preferential CO oxidation. Such a modification enhances the selectivity to C_{2+} hydrocarbons.

A similar effect of CO oxidation without a significant change in the selectivity to C_{2+} hydrocarbons is also observed in the case of Ag/Al_2O_3, which is an industrial catalyst for selective oxidation of ethene to ethylene oxide. It should be pointed out that, under the reaction conditions investigated in this work, any additional products of partial oxidation of ethene, ethane or methane are formed.

The best effect of integration is achieved when the OCM process is carried out at 780°C and the bed of CO oxidation is located at the end of the oven where the temperature drops to 230–250°C. In such conditions, the injection of additional 4 cm^3/min of oxygen leads to almost complete CO oxidation without C_{2+} selectivity changes. Nevertheless, even at such a low temperature, we did not observe any additional oxygenates in the product mixture, such as ethylene oxide or acetaldehyde. Further experiments concerning the modification of the oxidation catalyst composition and the optimization its work conditions as well as the influence of gaseous promoters are now in progress.

Table 2. Catalytic performance in oxidative coupling of methane process carried out in the presence and absence of additional CO oxidation bed

OCM catalyst (CO oxidation catalyst) $V_{add\ O2} = X$ [cm³/min][b]	Distance between beds [mm]	Temp. [°C]	Conversion of CH_4 [%]	Conversion of O_2 [%]	Selectivity to C_{2+} [%]	Selectivity to CO_2 [%]	Selectivity to CO [%]	Ratio of C_2H_4/C_2H_6 [mol/mol]	Temp. in OCM bed [°C][c]	Temp. in CO oxidation bed [°C][c]
Mn-Na_2WO_4/SiO_2	–	780	27.0	98.1	68.1	24.3	7.4	1.4	800–730	–
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2) $V_{add\ O2} = 0$	10		27.4	99.1	67.7	24.0	8.1	1.5	819–755	712–591
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2) $V_{add\ O2} = 0$	30		27.6	100	68.8	23.7	7.2	1.4	805–736	426–289
Mn-Na_2WO_4/SiO_2 (Ag/SiO_2) $V_{add\ O2} = 1$	30		27.9	100	63.3	31.8	4.6	1.4	805–735	435–291
Mn-Na_2WO_4/SiO_2 (Ag/Al_2O_3) $V_{add\ O2} = 0$	30		27.8	100	69.3	22.5	7.9	1.5	813–745	398–308
Mn-Na_2WO_4/SiO_2 (Ag/Al_2O_3) $V_{add\ O2} = 1$	30		28.2	100	65.4	32.3	1.98	1.5	813–744	398–307
Mn-Na_2WO_4/SiO_2 (Ag/Al_2O_3) $V_{add\ O2} = 4$	30		28.1	100	44.5	54.9	0.2	0.8	810–746	429–326
Mn-Na_2WO_4/SiO_2 (Ag/Al_2O_3) $V_{add\ O2} = 1$	50		27.8	100	68.9	22.9	8.0	1.5	813–753	271–222
Mn-Na_2WO_4/SiO_2 (Ag/Al_2O_3) $V_{add\ O2} = 2$	50		27.3	83.5	67.2	29.9	2.7	1.5	813–753	275–224
Mn-Na_2WO_4/SiO_2 (Ag/Al_2O_3) $V_{add\ O2} = 4$	50		27.4	78.7	61.8	37.6	0.2	1.4	–	–

Table 2

$Mn\text{-}Na_2WO_4/SiO_2$	–	750	16.9	62.3	63.7	27.26	8.72	0.8	761–703	–
$Mn\text{-}Na_2WO_4/SiO_2$ $(Ag/SiO_2)\ V_{add\ O2} = 0$	10		23.8	99.7	57.7	30.8	11.1	1.0	780–734	695–578
$Mn\text{-}Na_2WO_4/Si\ O_2$ $(Ag/SiO_2)\ V_{add\ O2} = 0$	30		18.6	82.2	59.6	39.9	0.0	0.7	768–706	415–283
$Mn\text{-}Na_2WO_4/SiO_2$ $(Ag/SiO_2)\ V_{add\ O2} = 1$	30		18.3	68.3	62.1	37.4	0.1	0.7	766–708	418–280
$Mn\text{-}Na_2WO_4/SiO_2$ $(Ag/Al_2O_3)\ V_{add\ O2} = 0$	30		21.5	99.9	55.7	43.7	0.0	0.7	774–719	396–305
$Mn\text{-}Na_2WO_4/SiO_2$ $(Ag/Al_2O_3)\ V_{add\ O2} = 0$	50		24.9	93.6	66.5	27.2	5.9	1.2	779–726	261–217
$Mn\text{-}Na_2WO_4/SiO_2$ $(Ag/Al_2O_3)\ V_{add\ O2} = 2$	50		24.8	76.3	66.4	32.0	1.2	1.2	778–726	263–219

[a] Reaction conditions: Molar ratio of CH_4:O_2:He = 3,8:1: 4,8; V_{total} = 77 $cm^3 \cdot min^{-1}$, $m_{OCM\ cat.}$ = 400 mg, $m_{CO\ oxi.\ cat.}$ = 400 mg; Reaction time = 2 h.
[b] *X* is 0, 1, 2 and 4 cm^3/min of additional O_2 injected before oxidation bed. [c] Description as in Table 1.

4. Conclusions

The integration of the OCM process with selective CO oxidation in a single reactor was investigated. It has been found that, under conditions suitable for the OCM process, controlling the selectivity in CO oxidation is very difficult. In most cases, CO and hydrocarbons are oxidized simultaneously over the oxidation bed. Preferential CO oxidation is possible only when the second oxidation bed is located at the end of the oven where the temperature is below 300°C and additional oxygen is injected before this bed.

References

[1] *Oxidative Coupling of Methane*, http://siluria.com/Technology/Oxidative_Coupling_of_Methane (access: 11.04.16).

[2] Godini H.R., Xiao S., Jašo S., Stünkel S., Salerno D., Son N.X., Song S., Wozny G., *Techno-economic analysis of integrating the methane oxidative coupling and methane reforming processes, Fuel* Processing Technology, vol. 106, 2013, 684–694.

[3] Michorczyk B., Ogonowski J., Michorczyk P., Węgrzyniak A., *Katalizatory dla procesu utleniającego sprzęgania metanu*, Przemysł Chemiczny, vol. 93, 2014, 1166–1173.

[4] Tiemersma T.P., Chaudhari A.S., Gallucci F., Kuipers J.A.M., van Sint Annaland M., *Integrated autothermal oxidative coupling and steam reforming of methane. Part 1: Design of a dual-function catalyst particle*, Chemical Engineering Science, vol. 82, 2012, 200–214.

[5] Tiemersma T.P., Chaudhari A.S., Gallucci F., Kuipers J.A.M., van Sint Annaland M., *Integrated autothermal oxidative coupling and steam reforming of methane. Part 2: Development of a packed bed membrane reactor with a dual function catalyst*, Chemical Engineering Science, vol. 82, 2012, 232–245.

[6] Tiemersma T.P., Kolkman T., Kuipers J.A.M., van Sint Annaland M., *A novel autothermal reactor concept for thermal coupling of the exothermic oxidative coupling and endothermic steam reforming of methane*, Chemical Engineering Journal, vol. 203, 2012, 223–230.

[7] Godini H.R., Xiao S., Kim M., Görke O., Song S., Wozny G., *Dual-membrane reactor for methane oxidative coupling and dry methane reforming: Reactor integration and process intensification*, Chemical Engineering and Processing, vol. 74, 2013, 153–164.

[8] Thybaut J. W., Marin G. B., Mirodatos C., Schuurman Y., van Veen A. C., Sadykov V. A., Pennemann H., Bellinghausen R., Mleczko L., *A Novel Technology for Natural Gas Conversion by Means of Integrated Oxidative Coupling and Dry Reforming of Methane*, Chemie Ingenieur Technik, vol. 86, 2014, 1855–1870.

[9] Godini H. R., Jaso S., Nghiem S. X., Görke O., Sadjadi S., Stünkel S., Song S., Simon U., Schomäcker R., Wozny G., *Miniplant-Scale Analysis of Oxidative Coupling of Methane Process*, Journal of Oil, Gas and Petrochemical Technology, vol. 2, 2015, 57–71.

[10] Skutil K., Taniewski M., *Some technological aspects of methane aromatization (direct and via oxidative coupling)*, Fuel Processing Technology, vol. 87, 2006, 511–521.

[11] Skutil K., Taniewski M., *Indirect methane aromatization via oxidative coupling, products separation and aromatization steps*, Fuel Processing Technology, vol. 88, 2007, 877–882.
[12] Qiu P., Lunsford J.H., Rosynek M.P., *Steady-state conversion of methane to aromatics in high yields using an integrated recycle reaction system*, Catalysis Letters, vol. 48, 1997, 11–15.
[13] Graf P.O., Lefferts L., *Reactive separation of ethylene from the effluent gas of methane oxidative coupling via alkylation of benzene to ethylbenzene on ZSM-5*, Chemical Engineering Science, vol. 64, 2009, 2773–2780.
[14] Czechowicz D., Skutil K., Tórz A., Taniewski M., *An integrated process of oxidative coupling of methane and pyrolysis of naphtha in a scaled-up unit*, Journal of Chemical Technology and Biotechnology, vol. 79, 2004, 182–186.
[15] Wensheng C., Grant P., *Process for producing acetic acid and/or ethanol by methane oxidation*, Patent. WO 2014/143865 A1, Pub.18.09.2014.
[16] Ghareghashi A., Ghader S., Hashemipour H., *Theoretical analysis of oxidative coupling of methane and Fischer Tropsch synthesis in two consecutive reactors: Comparison of fixed bed and membrane reactor*, Journal of Industrial and Engineering Chemistry, vol. 19, 2013, 1811–1826.
[17] Xu L., Xie S., Liu S., Lin L., Tian Z., Zhu A., *Combination of CH_4 oxidative coupling reaction with C_2H_6 oxidative dehydrogenation by CO_2 to C_2H_4*, Fuel, vol. 81, 2002, 1593–1597.
[18] Xu L., Xie S., Liu S., Lin L., Tian Z., Zhu A., *Combination of CH_4 oxidative coupling reaction with C_2H_6 oxidative dehydrogenation by CO_2 to C_2H_4*, Fuel, 81, 2002, 1593–1597.
[19] Michorczyk B., Suszyński K., Smoleń P., Hędrzak E., *Utleniające sprzęganie metanu zintegrowane w jednym reaktorze z odwodornieniem etanu do etenu*, Przemysł Chemiczny, vol. 95, 2016, 1936–1940.
[20] Rekoske J.E., *Oxidative coupling of methane with carbon conservation*, Uop Llc, Glenview. USA. Patent. US006096934A. Pub.1.08.2000.
[21] Kalakkunnath S,. *Oxidative Coupling of Methane to Ethylene by Siluria Process*, https://chemical.ihs.com/PEP/Public/Reports/Phase_2014/RW2014-07/ (access: 11.04.16).
[22] Shi J., Lu Y., Hu Ch., *Effect of CO_2 on the structural variation of $Na_2WO_4/Mn/SiO_2$ catalyst for oxidative coupling of methane to ethylene*, Journal of Energy Chemistry, vol. 24, 2015, 394–400.
[23] Litawa B., Michorczyk P., Ogonowski J,. *Influence of CO_2 on the catalytic performance of La_2O_3/CeO_2 and CaO/CeO_2 catalysts in the oxidative coupling of methane*, Polish Journal of Chemical Technology, vol 15, 2013, 22–26.
[24] Xu Y., Yu L., Cai C., Huang J., Guo X., *a study of the oxidative coupling of methane over $SrO\text{-}La_2O_3/CaO$ catalysts by using CO_2 as a probe*, Catal. Lett., vol. 35, 1995, 215–231.
[25] Kolts JH., Kukes S.G., *Catalytic oxidation of carbon monoxide*, USA. Patent. 4808394. Pub.28.02.1989.
[26] Kahlich M.J., Gasteiger A., Behm R.J., *Kinetics of the Selective CO Oxidation in H_2-Rich Gas on Pt/Al_2O_3*, Journal of Catalysis, vol. 171, 1997, 93–105.
[27] Mariño F., Descorme C., Duprez D., *Noble metal catalysts for the preferential oxidation of carbon monoxide in the presence of hydrogen (PROX)*, Applied Catalysis B: Environmental, vol. 54, 2004, 59–66.

[28] Oh S.H., Sinkevitch R.M., *Carbon Monoxide Removal from Hydrogen-Rich Fuel Cell Feedstreams by Selective Catalytic Oxidation*, Journal of Catalysis, vol. 142, 1993, 254–262.
[29] Avgouropoulos G., Ioannides T., Papadopoulou Ch., Batista J., Hocevar S., Matralis H.K., *A comparative study of Pt/γ-Al_2O_3, Au/α-$Fe_2O_3$and CuO–CeO_2 catalysts for the selective oxidation of carbon monoxide in excess hydrogen*, Catalysis Today, vol. 75, 2002, 157–167.
[30] Bethke G.K., Kung H.H., *Selective CO oxidation in a hydrogen-rich stream over Au/γ-Al_2O_3 catalysts*, Applied Catalysis A: General, vol. 194, 2000, 43–53.
[31] Grisel R.J.H., Nieuwenhuys B.E., *Selective Oxidation of CO, over Supported Au Catalysts*, Journal of Catalysis, vol. 199, 2001, 48–59.
[32] Bond G.C., Thompson D.T., *Gold-Catalysed Oxidation of Carbon Monoxide*, Gold Bulletin, vol. 33, 2000, 41–50.
[33] Sun X., Su H., Lin Q., Han Ch., Zheng Y., Sun L., Qi C., *Au/Cu–Fe–La–Al_2O_3: a highly active, selective and stable catalysts for preferential oxidation of carbon monoxide*, Applied Catalysis A: General, vol. 527, 2016, 19–29.
[34] Teng Y., Sakurai H., Ueda A., Kobayashi T., *Oxidative removal of co contained in hydrogen by using metal oxide catalysts*, International Journal of Hydrogen Energy, vol. 24, 1999, 355–358.
[35] Liu W., Flytzani-Stephanopoulos M., *Total Oxidation of Carbon Monoxide and Methane over Transition Metal Fluorite Oxide Composite Catalysts: I. Catalyst Composition and Activity*, Journal of Catalysis, vol. 153, 1995, 304–316.
[36] Liu W., Flytzani-Stephanopoulos M., *Total Oxidation of Carbon-Monoxide and Methane over Transition Metal Fluorite Oxide Composite Catalysts: II. Catalyst Characterization and Reaction-Kinetics*, Journal of Catalysis, vol. 153,1995, 317–332.
[37] Avgouropoulos G., Ioannides T., Matralis H. K ., Batista J., Hocevar S., *CuO–CeO_2 mixed oxide catalysts for the selective oxidation of carbon monoxide in excess hydrogen*, Catalysis Letters Vol. 73, 2001, 33–40.
[38] Hung-Kuan Lin H-K., Chiu H-Ch., Tsai H-Ch., Chien S-H., Wang Ch-B., *Synthesis, characterization and catalytic oxidation of carbon monoxide over cobalt oxide*, Catalysis Letters, vol. 88, 2003, 169–174.
[39] Wangcheng Z., Xinye Z., Yanglong G., Li W., Yun G., Guanzhong L., *Synthesis of mesoporous CeO2-MnOx binary oxides and their catalytic performances for CO oxidation*, Journal of Rare Earths, Vol. 32, 2014, 146–152.
[40] Li Z., Hound G.B., *a Review on Complete Oxidation of Methane at Low Temperatures*, Journal of Natural Gas Chemistry, vol. 12, 2003, 153–160.
[41] Stasinska B., *Katalityczne utlenianie metanu z powietrza wentylacyjnego kopalń*, https://www.researchgate.net/publication/267544218 (access: 11.04.16).
[42] Qu Z., Cheng M., Huang W., Bao X., *Formation of subsurface oxygen species and its high activity toward CO oxidation over silver catalysts*, Journal of Catalysis, vol. 229, 2005, 446–458.
[43] Imamura S., Yamada H., Utani K., *Combustion activity of Ag/CeO_2 composite catalyst*, Applied Catalysis A: General, vol. 192, 2000, 221–226.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.17.019.6212

Michał Grodecki (mgrode@pk.edu.pl)
Department of Geotechnics, Faculty of Environmental Engineering, Cracow University of Technology

Numerical modelling of gabion joints

Modelowanie numeryczne połączeń gabionów

Abstract

This paper presents results of a numerical analysis of the stability of a gabion retaining wall. The main objective of the paper is to identify how different methods of the modelling of gabion joints affect the stability of the structure.

Keywords: gabion, retaining wall, FEM, stability

Streszczenie

Praca przedstawia rezultaty analiz numerycznych stateczności muru oporowego z gabionów. Głównym celem jest ocena wpływu sposobu modelowania połączeń gabionów na stateczność muru.

Słowa kluczowe: gabion, mur oporowy, MES, stateczność

Symbols
ϕ – internal friction angle [deg]
γ – soil bulk density [kN/m^3]
c – cohesion [kPa]
SF – stability factor [–]

1. Introduction

The main subject of the investigation is the influence of gabion-to-gabion joints on the behaviour of a retaining wall. In general, there are two forces acting against relative displacements between gabions – friction between gabions and joint resistance. Friction between gabions is described by the friction coefficient. Joints are usually made of steel wire or clips; they are characterised by strength, usually given per 1m of the retaining wall.

Friction between gabions is treated as a special case of friction between the gabion and the soil in this work. According to [3], the friction coefficient between gabion and soil can be estimated from the following equation:

$$\alpha_{ds} \tan\varphi = \alpha_s \tan\delta + (1 - a_s)\tan\varphi \tag{1}$$

where:
$\alpha_d \tan\varphi$ – friction coefficient between gabion and soil;
δ – friction angle between steel and soil;
a_s – ratio of steel area to total gabion-subsoil joint area;
φ – friction angle of the subsoil.

If one choose a typical in Polish design practice assumption that there is no friction between soil and steel ($\delta = 0$) equation (3) can be simplified to:

$$\alpha_{ds} = 1 - a_s \tag{2}$$

Such an approach as that stated above is on the safe side as it leads to some underestimation of the friction forces between the gabions and the subsoil. In this work, $\delta = 0$ is used and the friction angle of the gabion filling is used to obtain an approximation of the friction coefficient between gabions.

Three different approaches could be used to model the joint behaviour. The first way of modelling assumes that connection is 'perfect' and that no interface elements between gabions are necessary. This approach does not allow modelling the failure of the wall caused by joint failure.

The second approach (more conservative) leads to the use of Coulomb-Mohr law for interface elements between gabions, where the friction angle would determine friction between gabions and the cohesion would estimate the joint strength. This approach allows modelling the failure of the wall due to joint failure (during sliding of gabions), but is not capable of modelling the joint resistance to a gap opening up between gabions (Fig. 1).

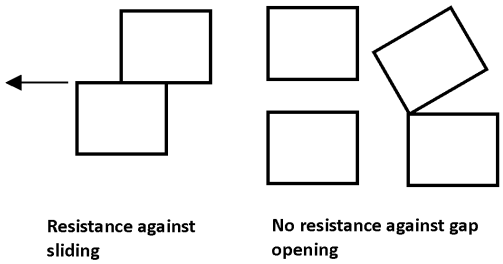


Fig. 1. Effect of the cohesion in the interface elements located between gabions

The required cohesion in interface elements can be calculated from the equilibrium equation of the joint capacity and part of the interface capacity is produced by cohesion:

$$c \cdot B = F_t \tag{3}$$

where:

c – cohesion in the interface element [kPa],

B – gabion width [m],

F_t – tensile strength of joints per 1 m of the wall [kN/m].

This leads to:

$$c = \frac{F_t}{B} \tag{4}$$

The third approach (the closest to reality) leads to the modelling of joints between gabions with the use of two truss elements arranged symmetrically which are perpendicular to the gabion surface (Fig. 2). Friction in the interface elements is then responsible for modelling the friction between gabions, the cohesion in interface elements is an approximation of the joint resistance against sliding (similarly to the second approach) and joints are responsible for modelling resistance to a gap opening up between gabions.

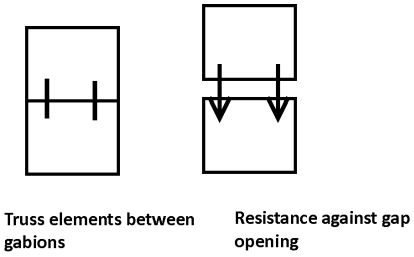


Fig. 2. Truss elements between gabions

Parameters of truss elements can be calculated from the following equations:

- equality of the joint capacity and ultimate load of truss elements:

$$F_t = f_{tl} \cdot A_l \cdot 2 / 1m \tag{5}$$

where:
f_{tl} – tensile strength of truss elements [kPa],
A_l – cross-sectional area of truss element [m^2].

- equality of axial stiffness of joints and truss elements:

$$A \cdot E \cdot 1m = A_l \cdot E_l \cdot 2 \tag{6}$$

where:
A – area of joints per 1m of wall [m^2/m],
E – Young's modulus of joints [GPa],
E_l – Young's modulus of the truss element [GPa].

From the above equations, the known properties of joints are F_t, A, E; the unknown properties of the truss elements are f_{tl}, A_l, E_l. As can be seen, there are two independent parameters of the truss element – ultimate load $f_{tl} \cdot A_l$ and axial stiffness $A_l \cdot E_l$, so one of the three required values (f_{tl}, A_l, E_l) could be adopted arbitrarily and two others calculated from above equations.

2. Numerical experiment

A stability analysis of a 4 m high model retaining wall was performed. Two variants of the wall were analysed – firstly, with the wall not buried in the subsoil; secondly, with the wall buried in the subsoil to half the height of a single gabion (Fig. 3).

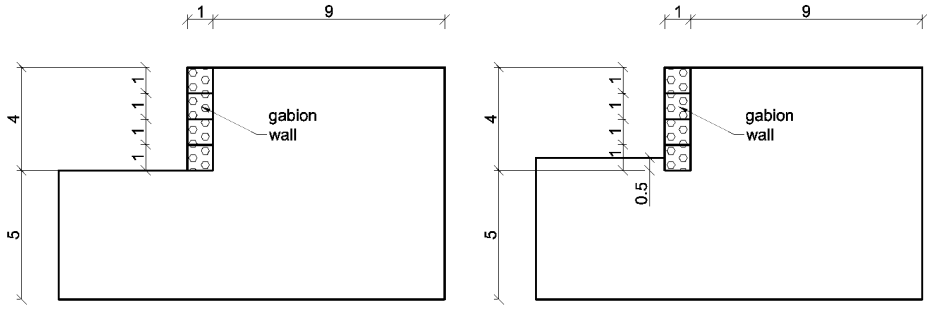


Fig. 3. Analysed object – not buried and buried in the subsoil (dimensions in metres)

A variety of retained soil and joint property combinations were used to obtain different modes of destruction and to judge in which situation the method of the modelling of joints affects the structure's stability. All numerical simulations were performed in plane strain conditions with the use of ZSoil Finite Element Method (FEM) software, described in detail in [5] and [7]. The stability factor SF was estimated using the c-f reduction method described in [6]. An interface element was used between the soil and the gabions. The filling and steel mesh properties were taken from [4]. The friction coefficients between the gabion and the subsoil, and between the gabions were calculated according to equation (3). The joints between the gabions were modelled using the different approaches described above.

A joint strength of F_t = 20.4 kN/m (minimum value according to ASTM A975 [1]) was used. An elasto-plastic Mohr-Coulomb model was used for retained the soil and the gabions. Additional cohesion for the gabions (which simulates the existence of the steel mesh) was calculated as described in [1] (strain at failure equal to 7% and tensile strength of the steel mesh equal to 20 kN/m were used as the input parameters of the mesh).

The obtained results show that for the wall not buried in the subsoil, the method of joint modelling does not affect the obtained stability factor values. However, for the wall buried in the subsoil to half of the height of one gabion, the situation is more complicated. For walls buried in the subsoil with low cohesion and high friction angle, the method of joint modelling does not have a significant influence on obtained SF values. However, for walls buried in the subsoil with low friction angle and high cohesion, such an influence can be easily seen. For such a wall, additional calculations considering 50% and 150% of minimum joint strength (according ASTM A975 [1]) were performed. The most important parameters of the retained soil and gabions used in such an analysis are presented in the Table 1.

Table 1. Parameters of the gabion and soil (with high cohesion and low friction angle) used in the analysis of the influence of joint strength on the stability of a structure

	c [kPa]	ϕ [°]	γ [kN/m^3]
Soil	35	5	20
Gabion	27	43	23
Interface elements between gabions	0 (Approach 1) 20.5 (Approach 2, 3)	40	–
Interface elements between gabions and soil	0	40	–

Truss elements (responsible for the modelling of a joint's resistance to gap opening) with a cross-sectional area equal to A = 0.51 cm^2 each and f_{tl} = 200 MPa (which gives F_t = 20.4 kN/m, as required by ASTM A975 [1]) were used.

An overview of the obtained results is presented in Table 2.

As can be clearly seen, SF increase due to the existence of joints between gabions is the greatest in Approach 3. Therefore, the influence of joint resistance to gap opening on wall stability is much more significant than the joint resistance against sliding. The total increase in the SF is equal to approx. 7.2% ((1.64–1.53)/1.53), the increase of the stabilising forces (1.64–1.53)/(1.53–1.00) is equal to approx. 20.7%. Such an influence is quite significant and could be taken into account in gabion wall design. However, the increase of the joint strength to 150% of the minimum strength required by ASTM A975 ([1]) does not have a significant influence on wall stability. The decrease of joint strength by approx. 50% results in some SF drops Therefore, according to the obtained results required by ASTM A975 ([1]), the joint strength is reasonable for walls with a typical height of about 4 m.

Table 2. SF values obtained from analysis of the influence of joint strength on the stability of a structure

	50% of required joint strength	100% of required joint strength	150% of required joint strength
Ideal connection (no interface elements between gabions)	1.54		
Approach 1 (only friction in contact elements)	1.53		
Approach 2 (friction and cohesion in contact elements)	1.55		
Approach 3 (friction and cohesion in contact elements + truss elements between gabions)	1.61	1.64	1.65

A typical stability loss mode (failure mechanism) of the wall buried in the subsoil is presented in Fig. 4.

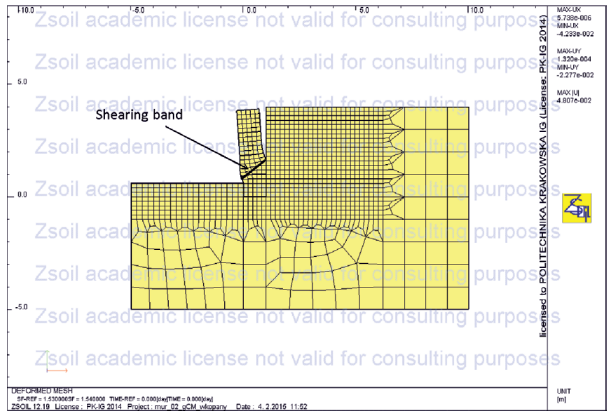


Fig. 4. Failure mode of the wall buried in the subsoil

3. Final remarks

The results of numerical simulations described above show that proper modelling of the joints between the gabions could affect the stability analysis results for the gabions buried in the soil with a small friction angle and high cohesion. For other situations, the influence of joints on a structure's stability is almost invisible. ASTM A975 ([1]) provides reasonable requirements for the strength of joints for typical walls with a height of approx. 4 m.

References

[1] *A 975 – 97 Standard Specification for Double–Twisted Hexagonal Mesh Gabions and Revet Mattresses (Metallic-Coated Steel Wire or Metallic-Coated Steel Wire With Poly(Vinyl Chloride) (PVC) Coating)*, American Society for testing and materials, 2011.

[2] Bathurst R.J., Rajagopal K., *Large-scale triaxial compression testing of geocell reinforced granular soils*, Geotechnical Testing Journal, Vol. 16, No. 3, 1993, 296–303.

[3] Bergado D.T., Youwai S., Teerawattanasuk C., Visudmedanukul P., *The interaction mechanism and behavior of hexagonal wire mesh reinforced embankment with silty sand backfill on soft clay*, Computers and Geotechnics, 30/2003, 517–534.

[4] Jayasree P.K., *Performance of gabion faced reinforced earth retaining walls*, PhD Thesis, Cochin University of Science and Technology, 2008.

[5] Podleś K., Truty A., Urbański A., *Analiza zagadnień geotechnicznych w systemie Z_SOIL*, X Jubileuszowa Konferencja Naukowa „Metody Numeryczne do Projektowania i Analizy Konstrukcji Hydrotechnicznych", Korbielów 1998, 100–108.

[6] Truty A., Urbański A., Grodecki M., Podleś K., *Komputerowe modele zagadnień osuwiskowych oraz ich zabezpieczeń*, Zeszyty Naukowo-Techniczne Stowarzyszenia Inżynierów i Techników Komunikacji Rzeczpospolitej Polskiej w Krakowie nr 88 (zeszyt 144), 2009.

[7] *Z_Soil.PC, Theoretical Manual*, ZACE Services Ltd., Lozanna 2000.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.17.020.6213

Adam Kasprzak
Paweł Popielski (pawel.popielski@is.pw.edu.pl)
Department of Hydraulic Engineering and Hydraulics, Faculty of Building Services, Hydro and Environmental Engineering, Warsaw University of Technology

Comparison of temperature and displacements of a gravity section and a buttress section illustrated with the example of concrete dam in Rożnów

Porównanie termiki i przemieszczeń sekcji ciężkiej i półciężkiej na przykładzie zapory betonowej w Rożnowie

Abstract

This paper presents the results and the process of numerical analysis of a selected section of the concrete dam in Rożnów. The calculations were carried out for two different variants: the gravity section and the buttress section. The effects of the geometry on the temperature distribution inside the section, and the effect of the temperature on displacements, were examined.

Keywords: numerical modelling, concrete dam, thermal analysis

Streszczenie

W artykule przedstawiono wyniki i proces analizy numerycznej dla wybranej sekcji zapory betonowej w Rożnowie. Obliczenia przeprowadzono dla dwóch różnych wariantów, dla sekcji ciężkiej i półciężkiej, sprawdzając wpływ geometrii na warunki temperaturowe, rozkład temperatury wewnątrz sekcji oraz ich wpływ na przemieszczenia.

Słowa kluczowe: modelowanie numeryczne, zapora betonowa, analiza termiki

1. Introduction

The concrete gravity dam in Rożnów was erected on the 80th kilometer of the Dunajec River, in a place where the river considerably changes its course, taking the shape of the so-called Rożnów Serpentine. The process of construction began in 1935 and finished in 1941, when filling the reservoir initiated. The Dam consists of 44 sections, seven of which are spillway sections and another five sections form the power station containing Kaplan hydroelectric sets and machine hall. The sections are 15 m wide, except for sections containing turbines (17 m wide) and sections near the left abutment (5 and 7.5 m wide). According to [1], the structure has been classified as a category I dam.

The structure is located on the Carpathian flysch. In the subsoil, there are sandstone layers with interlayers of shale. Sandstone nearby the location of the Dam is strongly cracked. In the initial project prepared by Professor Zbigniew Zmigrodzki from the Warsaw University of Technology, it was assumed to build straight axis buttress dam, consisting of T-shaped sections and U-shaped sections containing a power station. However, as ground works progressed and further geological researches of the subsoil were continued, four detachment faults were discovered and the project was changed. It was decided to curve the axis of the Dam near the left abutment and to alter the buttress sections to gravity sections.

To prevent water outflow from the reservoir, the ground was improved with two grouting screens reaching 30 m below the foundation of the Dam. By dint of this solution, subsoil rock was strengthened. Moreover, after finishing concrete works, subsidences were reduced.

2. Preparing of the numerical model

The subject of numerical modelling was the 18th section of the Dam. It is the highest of the typical sections. It is located in central part of the Dam, near to the so-called Rożnów Peninsula and borders on sections containing a hydroelectric power station. The calculations were carried out for the following variants: already built gravity section and initially designed buttress section. In addition, two different variants of concrete were considered: it was assumed to be an elastic substance in the first one and viscous-elastic in the other one, in which creep and relaxation were taken into account.

The first stage of the analysis was establishing the proper geometry of the model. The geometry of the gravity section was drawn on the basis of available archival materials. The final shape was verified by comparing it with laser scanning measurements [2]. The geometry of the buttress section was established according to the available publications containing initial project sketches. The geological layer system in the subsoil was simulated on the basis of archival studies, the data from articles concerning the foundation of the water power station [3] as well as the project of grout curtain elaborated by the Polish Geological Institute.

The system of geological layers in the subsoil is presented in Fig. 1. The models of the gravity section and the buttress section, along with classification of material properties, are presented in Fig. 2 and 3.

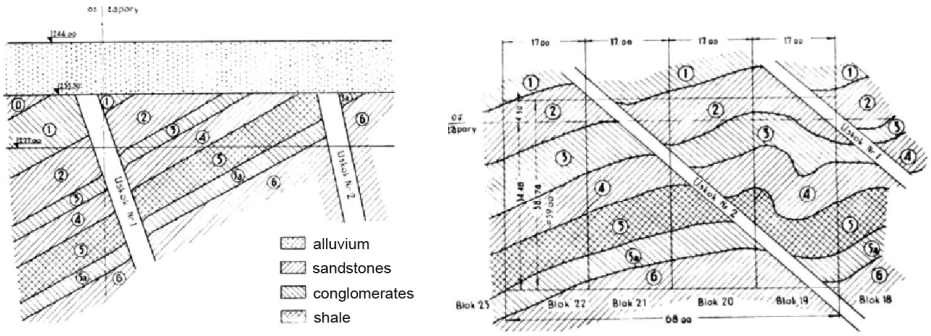


Fig. 1. The system of the geological layers near the 18th section of the Dam [3]

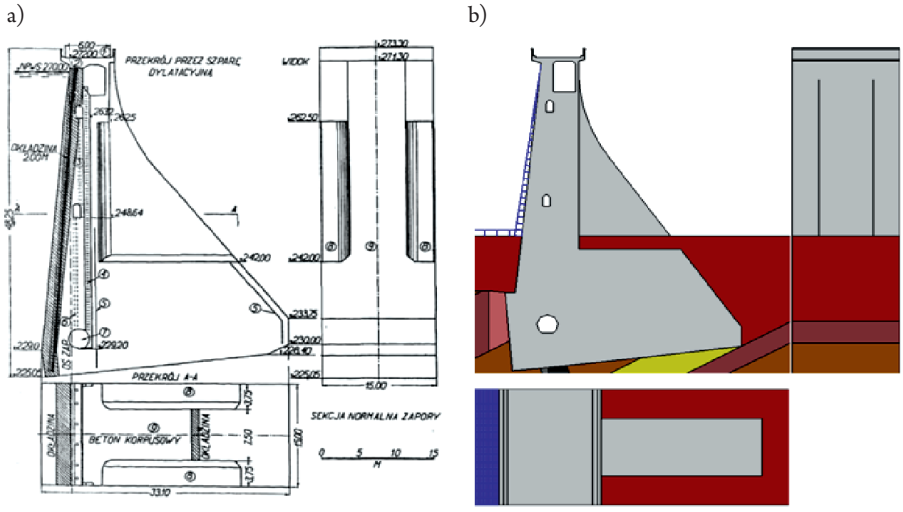


Fig. 2. The geometry of the buttress section: a) sketch [4], b) model generated by Z-Soil software [5]

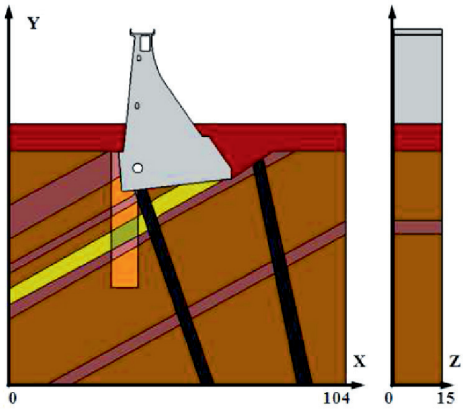


Fig. 3. The geometry and dimensions of the gravity section [5]

Values of the material properties used in calculations were assumed on the basis of researches carried out during the construction process of the Dam, as well as literature and standard data. Properties used in thermal analysis are given in Table 1 and properties used in calculations of displacements and stresses are shown in Table 2.

Table 1. Material properties used in thermal analysis [5]

No	Name	Symbol	Thermal conductivity	Volumetric thermal capacity	Thermal dilatancy
			[kN/d/C]	[kN/m²/C]	[1/C · 10⁻⁵]
1	Concrete	1	112.32	2058.00	1.13
2	Alluvium	2	121.00	1428.00	1
3	Sandstone I	3	232.40	2050.70	1
4	Sandstone II	4	232.40	2050.70	1
5	Conglomerate I	5	232.40	2059.02	1
6	Conglomerate II	6	232.40	2059.02	1
7	Shales I	7	99.36	2302.50	1
8	Shales II	8	99.36	2302.50	1
9	Discontinuity spacing I		232.40	2050.70	1
10	Discontinuity spacing II	10	232.40	2050.70	1

Table 2. Material properties used in mechanical analysis [5]

No	Name	Dead weight	Young's modulus	Poisson's ratio	Hydraulic conductivity
		[kN/m³]	[kPa]	[–]	[m/d]
1	Concrete	24.5	40 000 000	0.2459	$1 \cdot 10^{-9}$
2	Alluvium	17.0	80 000	0.20	10
3	Sandstone I	24.5	590 000	0.25	$4 \cdot 10^{-5}$
4	Sandstone II	24.5	590 000	0.25	$1 \cdot 10^{-9}$

5	Conglomerate I	24.6	330 000	0.25	$3 \cdot 10^{-6}$
6	Conglomerate II	24.6	330 000	0.25	$1 \cdot 10^{-9}$
7	Shales I	25.0	39 000	0.20	$9 \cdot 10^{-8}$
8	Shales II	25.0	39 000	0.20	$1 \cdot 10^{-9}$
9	Discontinuity spacing I	24.5	413000	0.25	10
10	Discontinuity spacing II	24.5	413000	0.25	$1 \cdot 10^{-9}$

To describe the creep of concrete, creep function of exponential type was assumed:

$$C(t,\tau) = A \times \left(1 - e^{-\frac{1}{B}(t-\tau)}\right) \tag{1}$$

Parameter A (the creep coefficient) is described by the following formula: $A = \frac{\phi}{E}$. The ϕ coefficient was estimated according to PN-EN-1992-1 Polish Standard. For humidity of RH = 80% (on the outside), dimensions of the element as at mid-height of the Dam, time of load imposing t_0 = 365 d and concrete strength C30/37 the result was ϕ = 1. Thus, it was designated that $A = 2.5 \cdot 10^{-8}$ [1/kPa]. For parameter B, time of retardation was assumed to be 33.3 d [6].

3. Thermal analysis

3.1. Boundary conditions

The problem of thermal (variable in time) was solved as the first one. Determined temperature fields and their changes in annual cycle of atmospheric conditions were implemented into the mechanical analysis as a load. For that purpose, the Fourier equation was solved:

$$\left(\lambda T_{,i}\right)_{,i} + \frac{\partial H}{\partial t} = c^* \frac{\partial T}{\partial t} \qquad \text{w } \Omega \tag{2}$$

where:
T – temperature [°C],
t – time [d],
λ – thermal conductivity [kN/d/C],
c^* – volumetric thermal capacity [kN/m^2/C],
H – source of the heat [kN/m^2].

Boundary conditions of three types were used in the analysis:

1) Boundary conditions of the first type consist in applying a known temperature to the surface. Boundary conditions of that type were applied to the upstream face and

to the surface of the reservoir bottom, where concrete and subsoil are in permanent contact with water, as well as around the model, where constant temperature of the ground (9°C) was applied on the boundaries;

2) Boundary conditions of the second type – convection.

$$-h\left(T-T_e\right)=\lambda\frac{\partial T}{\partial n} \tag{3}$$

where:

T_e – temperature of the surroundings [°C],

h – convection coefficient, based on PN-91/B-02020 Polish Standard h = 23 [W/m/K].

That boundary condition was applied to surfaces, which remain in contact with external air, this is: to the downstream face, to the upstream face above water surface and to the crest of the Dam. In case of the buttress dam, it was also applied to the lateral surfaces.

3) Boundary condition of third type – adiabatic:

$$0=\lambda\frac{\partial T}{\partial n} \tag{4}$$

That condition was applied to surfaces, where the heat flow induced by air movement is inconsiderable. The condition was applied to the lateral surfaces of the model, which are perpendicular to the axis of the Dam, to revision galleries and to external surfaces, where other boundary conditions were not applied.

Surfaces, to which boundary conditions were applied, are presented in the Fig. 4.

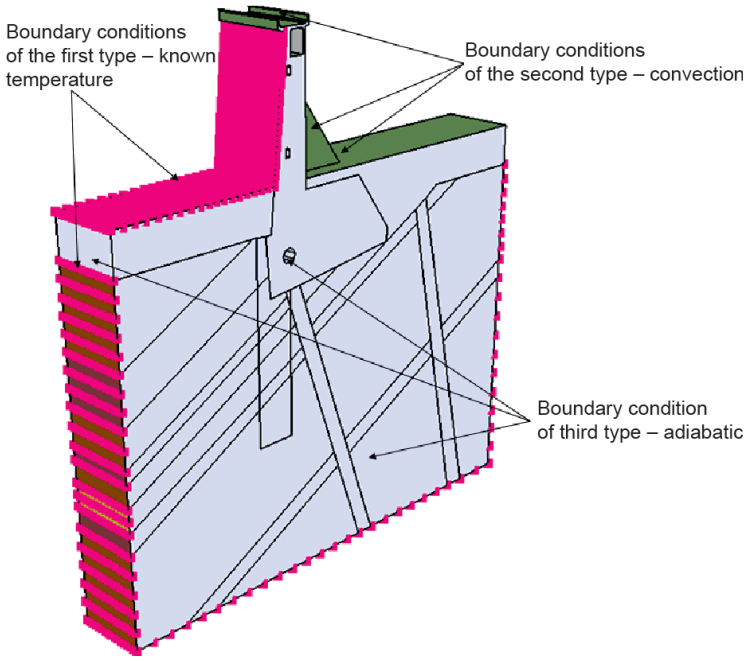


Fig. 4. Boundary conditions used in thermal analysis [5]

Calculations were conducted for a variable water level. The character of air and water temperature changes is sinuous and is described by the following formula:

$$T(t) = T_{sr}\left(1 - \frac{\Delta T}{T_{sr}} \cdot \cos\left(\frac{2\pi}{365}t\right)\right) \ [°C] \tag{5}$$

where:
$T(t)$ – temperature at a given instant of time t[°C],
T_{sr} – average temperature [°C],
Δt – temperature amplitude [°C].

Water temperature in the reservoir was assumed on the basis of thermal characteristics of the Rożnów Lake. Temperature distribution according to the depth is presented in Fig. 5. Temperature was implemented into the model in 5 points at every 5 m in depth. Temperature between them was linearly interpolated. The data used in calculations of temperature are given in Table 3 and in Fig. 6.

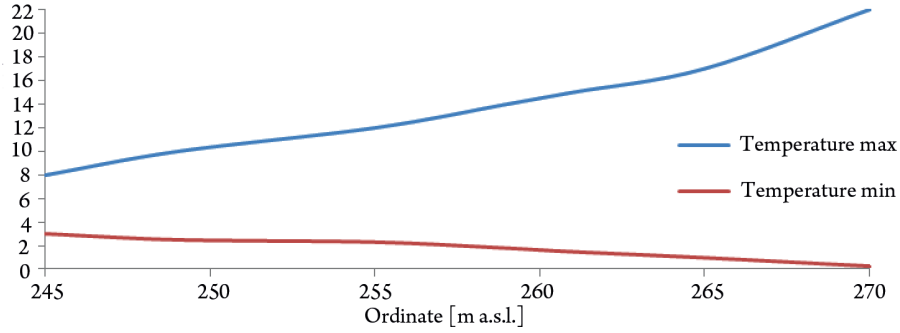


Fig. 5. Water temperature in the Rożnów Lake

Table 3. Data used in thermal analysis

	Symbol on Fig. 6	T_{max} [°C]	T_{min} [°C]	ΔT [°C]	T_{sr} [°C]
Air temperature		24	–6	30	9
Water temperature 245 m.a.s.l.		8	3	5	5.5
Water temperature 250 m.a.s.l.		10.5	2.5	8	6.5
Water temperature 255 m.a.s.l.		12	2.3	9.7	7.15
Water temperature 265 m a.s.l.		17	1	16	9
Water temperature 270 m.a.s.l.		22	0.3	21.7	11.15

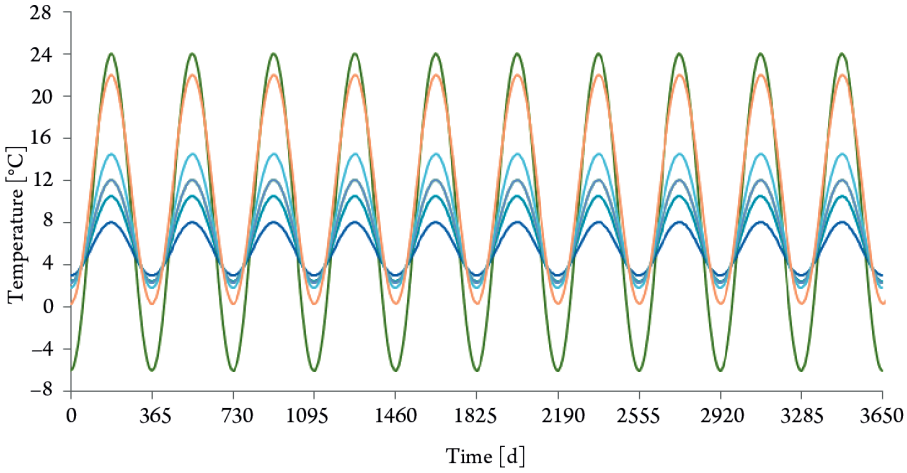


Fig. 6. Functions describing annual changes of water temperature

3.2. Results of thermal calculations

Analysis was carried out for 10 calculation series. Each of the series was 365 days long. A step of 5 days was set. Taking into account the transient character of the phenomena, the time of calculation must have been established. The differences between successive calculation series of identical boundary conditions are shown on Fig. 7. As can be noticed, the temperature in the model is stabilizing during initial cycles. The temperature difference exceeded 5°C after the first series. The difference between the consecutive cycles was decreasing more and more and was less than 0.5°C between series 5 and 4.

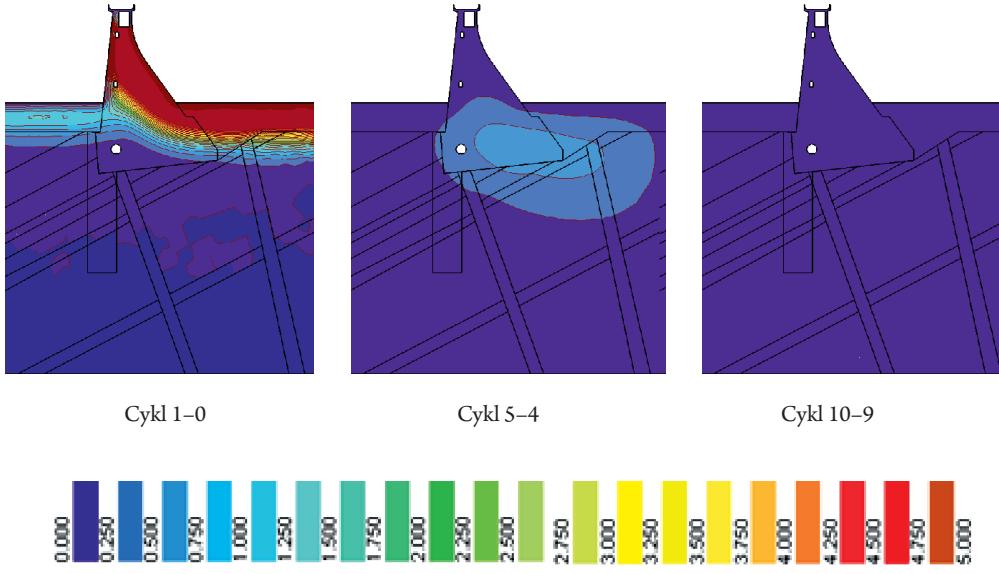


Fig. 7. Changes of temperature in consecutive calculating cycles

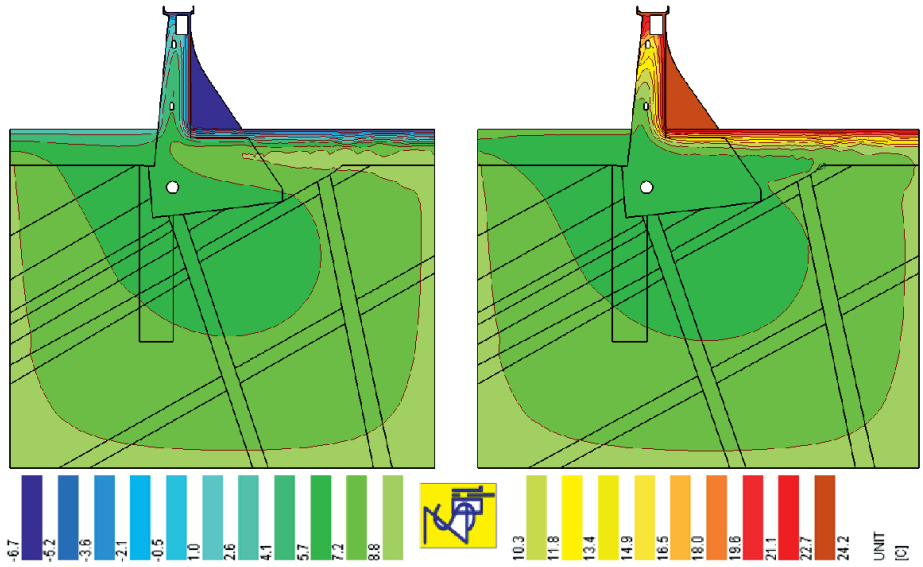


Fig. 8. Temperature distribution in the buttress section in summer and in winter

The temperature distribution inside the Dam and in the subsoil in the summer and in the winter is presented in Fig. 8. As can be seen, only near-surface layers of concrete and subsoil are susceptible to changes. Inside the structure, the temperature is constant and does not depend on weather conditions throughout the year. The temperature difference between the body of the dam and concrete surface layers induces thermal stresses. The exact temperature course throughout the year at selected points is shown in Fig. 9 and 10.

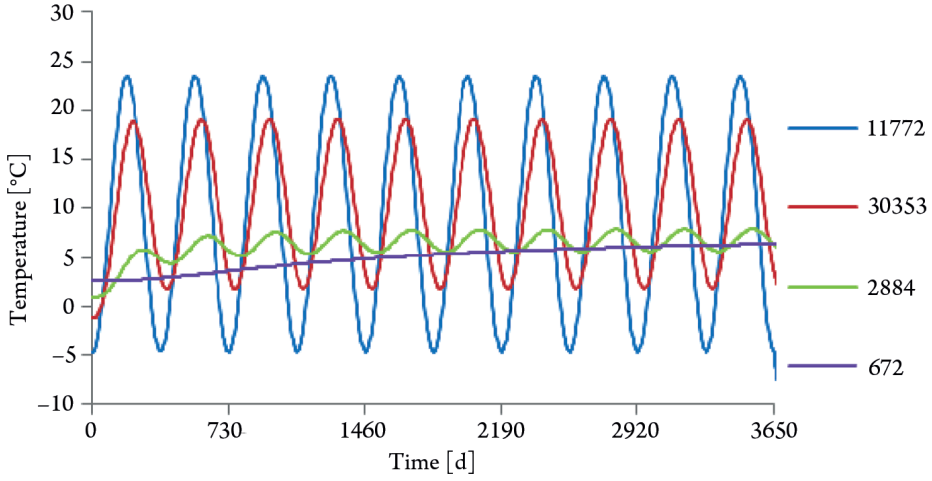


Fig. 9. Temperature changes in time at selected points of the buttress section

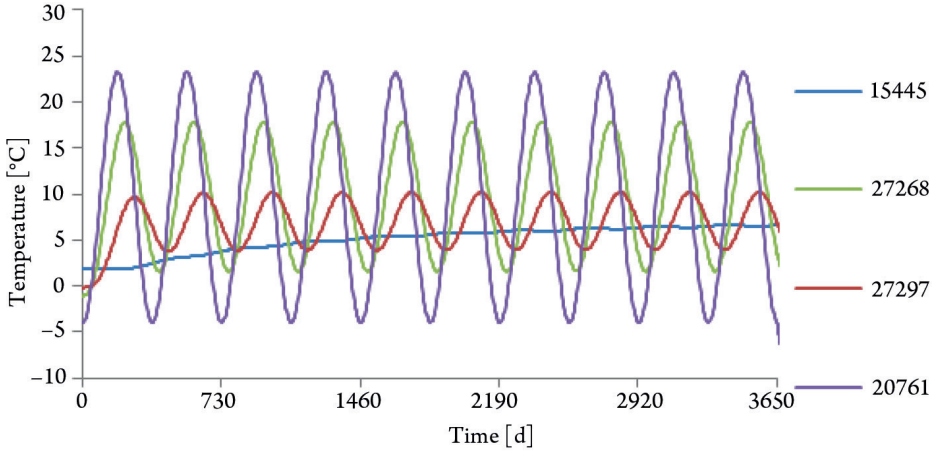


Fig. 10. Temperature changes in time at selected points of the gravity section

As the charts show, the near-surface layers of concrete react to temperature changes most quickly (points 11772 and 15445). The temperature increase in time at points located inside the Dam is slow. With time, increases diminish more and more. From the moment of stabilization of conditions, results of thermal analysis can be used for loading the mechanical model. The other points located inside the dam react with some delay. The highest temperature at points 30353 and 27268 occurs about 20 days later than at points situated on the surface, whereas at points in the second revision gallery (points 27297 and 2884), the delay is about 50 days. The inside of the buttress section heats up more quickly than the inside of the gravity section. The temperature in the buttress section is higher than in the gravity section during the summer and vice versa during the winter.

4. Analysis of displacements

4.1. Boundary conditions

The model used in the calculations of displacements was loaded with upstream and downstream (groundwater) water pressure. Upstream water level at 270 m.a.s.l. was assumed to be constant in time. Downstream water level ordinate was assumed to be 237.7 m.a.s.l. Moreover, hinged supports (applied to underneath of the model) and movable supports (allowing the subsoil and the structure to subside; they were applied to external vertical surfaces of the model) were used as boundary conditions.

Values of displacements obtained for elastic and viscous – elastic model were identical. During the summer, in the case of the buttress section as well as the gravity section, the crest of the Dam leans towards the upstream water, whereas during the winter, it leans towards the downstream water. Values obtained for the buttress section are almost two times higher.

Points localized inside the section do not displace. Results for selected points of the gravity section are presented in Fig. 12 and 13 and in Fig. 14 and 15 for the buttress section. (The points are located as in Fig. 10)

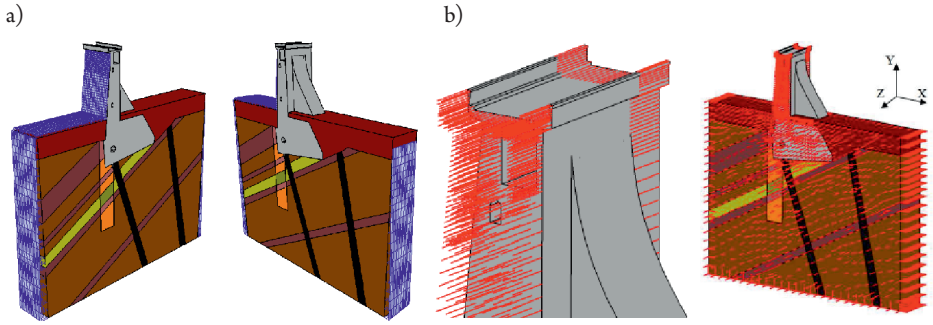


Fig. 11. Boundary conditions used in analysis of displacements [5]: a) Upstream and downstream water level, b) Supporting of the model

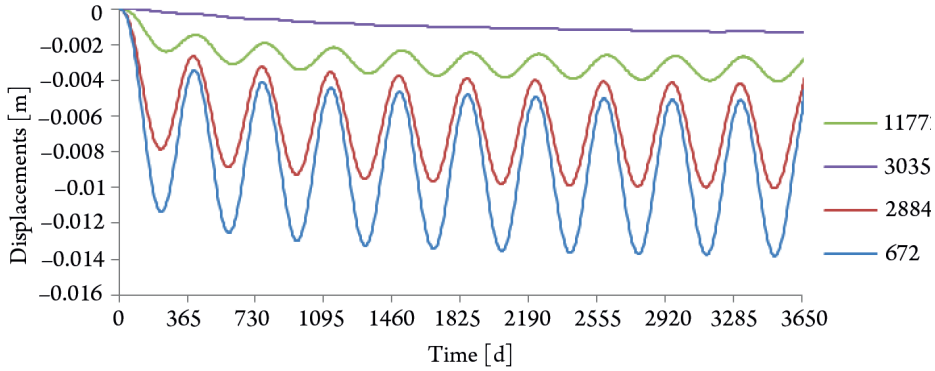


Fig. 12. Horizontal displacements of the points in the gravity section

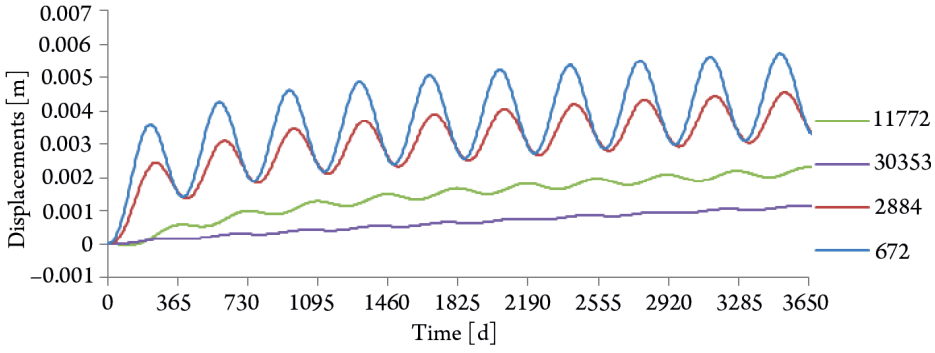


Fig. 13. Vertical displacements of the points in the gravity section

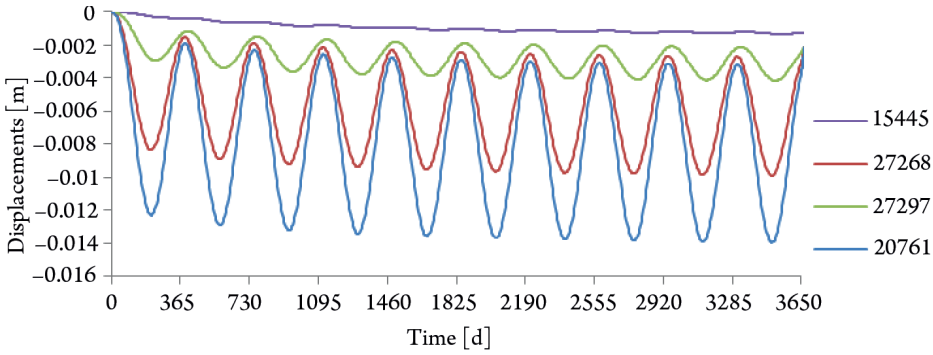


Fig. 14. Horizontal displacements of the points in the buttress section

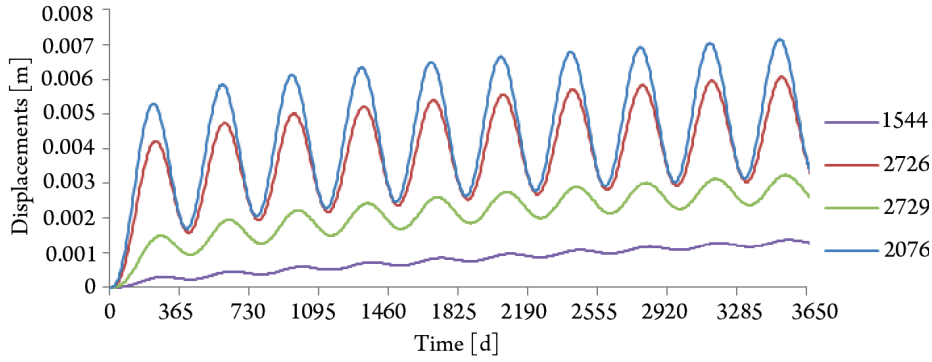


Fig. 15. Vertical displacements of the points in the buttress section

On the basis of the charts, it can be noticed that the increase of displacements is changing in time and it is diminishing in the subsequent years. Points located on the crest of the Dam are susceptible to the most significant displacements. The charts of displacements are of the same type as charts of temperature. In addition, it can be seen that points located near the concrete surface layers react to temperature changes much more quickly than points located deeper inside the section. The difference of horizontal displacement values between the summer and the winter is about 13 mm for the buttress section and about 7 mm for the gravity section.

5. Final conclusions

Temperature in the subsoil and in the inside of the gravity section and the buttress section is constant throughout the year. Near-surface layers are most susceptible to temperature changes. The difference between temperature of the inside of the Dam and external temperature is caused by displacements. The differences in temperature distribution of the gravity section and the buttress section are caused by differences of the geometry of the sections. The

displacement values obtained for the buttress section are more significant than the ones obtained for the gravity section. The displacements obtained for elastic and viscous-elastic model were identical. The Dam leans towards the upstream water in the summer and towards the downstream water in the winter. It is essential to take into account the effect of temperature changes on the values of displacements, while analyzing the monitoring of displacements in massive concrete hydrotechnical structures.

References

[1] Regulation by the Minister of Environment of 20th April 2007 On the technical requirements for hydrotechnical structures and their location. Journal of Laws of 2007, No 86, Item 579.

[2] Zaczek-Peplinska J., Popielski P., Kasprzak A. and others, *Development of large concrete object geometrical model based on terrestrial laser scanning*, Reports on Geodesy and Geoinformatics, Faculty of Geodesy and Geodetic Astronomy, Vol. 97, 2014, 91–102.

[3] Balcerski W., *Problemy statyczne fundowania zakładu wodno-elektrycznego w Rożnowie*, Gospodarka Wodna, 1938.

[4] Jarząbek S., *Rożnów: budowa zbiornika i zakładu wodno-elektrycznego do r. 1938*, Związek Polskich Fabryk Cementu, Warszawa 1938.

[5] Kasprzak A., *Numerical Model of the 18th Section of the Rożnów Dam*, Master thesis, Department of Hydraulic Engineering and Hydraulics, Faculty of Environmental Engineering, Warsaw University of Technology, 2014.

[6] Hrabowski W., Urbański A., *Trójwymiarowe modelowanie numeryczne i analiza insitu pól termicznych, filtracyjnych i mechanicznych w wybranej sekcji zapory betonowej w Zatoniu*, 10th Anniversary Conference of Dam Monitoring, IMiGW, Warszawa 2003.

[7] Tomik T., *Próba charakterystyki termicznej Zbiornika Rożnowskiego*, Prace Państwowego Instytutu Hydrologiczno-Meteorologicznego, vol. 96, Wyd. Komunikacji i Łączności, Warszawa 1969, 43–55.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.17.021.6214

Justyna Kwaśny (kwasny.justyna@gmail.com)
Wojciech Balcerzak
Institute of Water Supply and Environmental Protection, Department of Environmental Engineering, Cracow University of Technology

Characteristics of selected methods for the synthesis of nanometric zirconium oxide – critical review

Charakterystyka wybranych metod syntezy nanometrycznego tlenku cyrkonu – przegląd

Abstract

High chemical stability, resistance to changes in the pH, pressure and temperature meant that zirconium oxide is widely used in many fields. It is used in water treatment and waste water treatment processes, as well as air purification. In this paper, selected methods of nano-zirconia synthesis in liquid phase were characterized. These methods include, among others, the microemulsion method. Based on literature data, the advantages and difficulties associated with the use of each method are presented, in order to answer the question of which method of nanometric zirconium oxide synthesis in the liquid phase is the most advantageous. The authors also pointed out some directions of development for the discussed methods, which relate to, among others, solvent change and the use of additives in the form of polymers.

Keywords: nano zirconia, hydrothermal method, synthesis, microemulsion method

Streszczenie

Wysoka stabilność chemiczna, odporność na zmiany pH, ciśnienia i temperatury sprawiły, że tlenek cyrkonu znajduje szerokie zastosowanie w wielu dziedzinach. Jest stosowany w procesach uzdatniania wody i oczyszczania ścieków, a także w oczyszczaniu powietrza. W niniejszej pracy scharakteryzowano wybrane metody syntezy nanotlenku cyrkonu w fazie ciekłej. Wśród grupy tych metod wyróżnia się m.in. metodę mikroemulsji. Na podstawie danych literaturowych przedstawiono zalety i trudności związane ze stosowaniem poszczególnych metod, aby móc odpowiedzieć na pytanie, która metoda syntezy nanometrycznego tlenku cyrkonu, w fazie ciekłej, jest najkorzystniejsza. Autorzy wskazali również na pewne kierunki rozwojowe omówionych metod, które wiążą się m.in. ze zmianą polarności rozpuszczalnika i stosowaniem dodatków w formie polimerów.

Słowa kluczowe: nanotlenek cyrkonu, metoda hydrotermalna, synteza, metoda mikroemulsji

1. Introduction

Special properties of zirconia and its wide application mean that methods of its preparation are constantly being developed. The direction of this expansion results from the need to obtain materials with improved performance associated with the mechanical and thermal strength, and chemical stability.

Nano-sized zirconia can be obtained in both the gas and liquid phases, as well as from the solid. There are two general groups of methods for obtaining nanomaterials. These methods are "bottom-up" and "top down" [1–3]. "Bottom-up" methods rely on building a nanometric structure from below, that is, from the level of individual atoms or molecules. "Top down" methods, as methods from above, rely on grinding a solid on a micrometer structure to nanometer sizes 15). Methods of synthesis in the gaseous and liquid phases are among the "bottom-up" methods, and methods of synthesis in the solid state, in most cases, are classified as "top down" methods. Solid state synthesis, exemplified by mechanochemical synthesis, is based on the fact that chemical reactions are induced by mechanical energy. Whereby there is a reduction in the temperature of a chemical reaction [4–6].

2. Synthesis of nanometric zirconia in the vapor phase

Methods of synthesis in the gaseous phase are divided into physical vapor deposition PVD [7], which involves the condensation in an inert gas [8], and chemical vapor deposition CVD [9–12]. In this method, volatile compounds are heated to form vapor, mixed and transported using a carrier gas to a substrate, on which surface crystallization of the product takes place. In these methods, due to the determination of the appropriate conditions, the pair (usually oversaturated) becomes thermodynamically unstable, and therefore the formation of the condensed phase by homogeneous nucleation of solid phase occurs. The process is carried out with a high degree of vapor supersaturation resulting in an increased density of nucleation during the growth of nucleating agents. The key issue to achieve nanometric product is a fast extinguishing system, completed within a reasonable time, which is done by removing the source of supersaturation or extending the kinetics (cooling circuit), with the result that the particles do not grow [3].

Among these methods, various techniques for depositing material from a gaseous phase are highlighted. For example, it may be a magnetron sputtering [13, 14], plasma spraying [15–20], which consists in creating a plasma of high temperature, to 2000 K, and then delivered thereto, by a carrier gas, the material in the form of a powder that will be deposited. Hass et al. [15] concluded that a ceramic coating obtained this way may exhibit less consistency with the substrate and be characterized by higher porosity and non-uniform distribution of the pore volume fraction and shape. This fact significantly affects the thermal conductivity of the material. For another technique, which is the electron beam – physical vapor deposition (EB-PVD), the elements are rotated in the vapor plume, so that condensation of vapor molecules generally occurs unevenly. The result is a layer consisting of a porous, strongly

textured columnar structures of different lengths. Such a layer structure favors its flexibility and increases the thermal conductivity. Another example of physical vapor deposition is an electron beam – directed vapor deposition (EB-DVD) in which the material is not rotated. This results in a layer with a columnar structure, but with similar porosity, shape and length [15]. There is also a technique, which uses stellarator, specifically heliotron, known as the solar physical vapor deposition (SPVD) method [21].

Phase synthesis methods require specific gas pressure and temperature. Therefore, it is necessary to use specific and expensive equipment as well as facilities, among which plasma burners and magnetron stand out [13].

Methods for chemical synthesis from the gas phase can also be divided into several techniques. It should be noted that these methods differ mainly in the way of introducing the precursor and in the type of thermal energy source. In this group, chemical synthesis in the gaseous phase with the use of organometallic precursors [11, 22], liquid-injection chemical vapor deposition [23] and vapor phase hydrolysis [24] can be distinguished. And in the case, where the precursor is not introduced in the form of a vapor, but in the form of tiny droplets generated by a nebulizer, e.g. ultrasonic, spray pyrolysis techniques [25, 26] are highlighted, as well as others, e.g. aerosol assisted chemical vapor deposition (AACVD) [27].

Methods of synthesis and deposition from the gaseous phase result in the acquisition of a material with high purity. They allow to obtain the product, both in powder form and in the form of a layer applied to the substrate. The biggest limitation in the application of these methods is the need to use expensive and complex equipment.

3. Examples of the synthesis of nanometric zirconia in the liquid phase

Among liquid-phase synthesis methods, inter alia co-precipitation [28–31], hydrothermal methods [32–42], sol-gel method [43–48] and microemulsion method [49–56] are distinguished.

3.1. Synthesis of nano-zirconia by co-precipitation

Co-precipitation methods can be classified as conventional methods for the preparation of nanocrystals. Depending on the reaction conditions and substrates, the method allows to obtain particles with different diameters. Crystalline, non-agglomerated, nano-sized Al_2O_3-ZrO_2 prepared by Zhou and coworkers [57] can be cited as examples, which had a particle diameter of about 7 nm, followed by tetragonal zirconia polycrystals stabilized by yttrium (3 mol%.) with a diameter of 21.3 nm obtained by Hsu et al. [58] as well as lead catalysts based on CdO_2-ZrO_2 (in a single phase) with a diameter of 5.2 to 28.9 nm [59] and the Y_2O_3-ZrO_2 powder with a crystal diameter of 53–109 μm obtained by Aruna et al. [60].

The greatest advantage of the precipitation method is the possibility to obtain the product in amounts consisting of grams [28], and the ability to use cheaper and easier to

acquire equipment and nano-particle precursors [58, 61]. The disadvantage of this method is the inability to control the size of the resulting nanoparticles [28]. In order to reduce the agglomeration of nanoparticles, Wang et al. [61] proposed the synthesis of nano-ZrO_2 based on the direct precipitation using ethanol instead of water. The authors [61] used water, ethanol and water-ethanol mixtures at various ratios by volume, $ZrOCl_2$ i $NH_3 \cdot H_2O$, except that the surfactant was applied in the form of poly(ethylene glycol) 800 – PEG 800. Depending on the solvent used, products with different grain size and dispersion were obtained. The authors [61] have shown that the use of ethanol-limited agglomeration of particles, and no addition of PEG 800, additionally increase this effect [61].

3.2. Synthesis of nano-zirconia by hydrothermal methods

Hydrothermal methods are a group of methods among which hydrothermal decomposition, hydrothermal crystallization, homogeneous hydrothermal precipitation, microwave hydrothermal synthesis and hydrothermal synthesis, assisted by ultrasound and also hydrothermal synthesis under supercritical conditions [62] as well as hydrothermal oxidation [35] are highlighted. Hydrothermal methods relate to processes at hydrothermal conditions, which depend on temperature, pressure, pH, redox potential and environmental activity of the ingredients, which depends on the dielectric constant and the dissociation of water [35, 62]. It should be emphasized that, already in 1980, using the hydrothermal method, nanocrystalline powders based on ZrO_2 were obtained.

The hydrothermal method can be based on crystallization in an aqueous solution under increased pressure and at a temperature below the critical point of water. This process is conducted in an autoclave, where the substance concentration gradient is determined. The aterial is arranged in the lower part of the autoclave. This material is heat-solubilized, then travels up through the autoclave wherein at a lower temperature, due to the saturation of the solution, the crystallization of crystals occurs. Hydrothermal synthesis refers to synthesis of a compound in the hydrothermal solution under the influence of temperatures above 100°C and at a pressure above 1 atmosphere. For the synthesis of zirconium oxide (IV), the essence of the method is the production of ceramic sols, by chemical reaction in an aqueous or organic-aqueous solution, while applying heat and pressure, in the presence of an alkali or acid, which have a pseudo catalytic effect on the reaction [37]. Hydrothermal synthesis allows to obtain the three crystalline forms of zirconium oxide (IV): monoclinic, tetragonal [35] and the cubic [37], at the same time, with a lower rate of the monoclinic form. Depending on the reactants and conditions of the synthesis, it is possible to obtain particles with different diameters. The impact of metal precursors is significant, which has been demonstrated by the Caillot and co-workers [63] and derived by their surface characteristics and ICP analysis results compiled (based on [63]) and reported graphically in Figure 1. Based on the graph, it can be concluded that the addition of cerium, lanthanum and titanium resulted in an increase in the surface area of the zirconia products. However, it should be emphasized that this increase was not identical with the increasing proportion of each metal oxide in the final product.

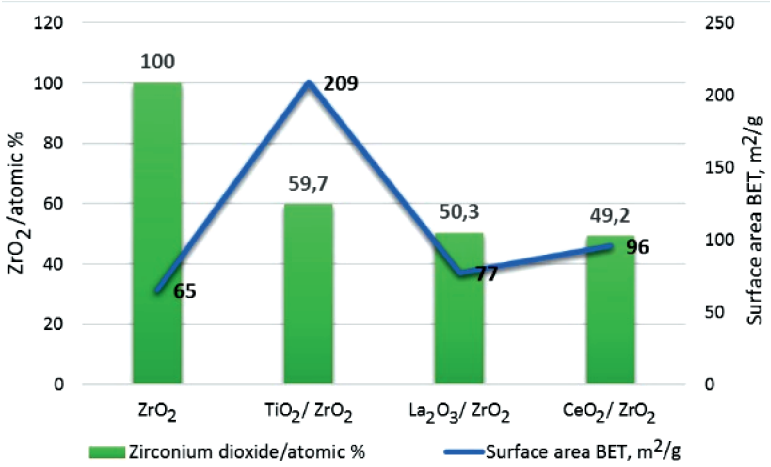


Fig. 1. Effect of metal precursor addition on the surface area of zirconia, on the basis [63]

Hydrothermal methods have both advantages and disadvantages. Among the advantages of these methods, we must distinguish the ability to conduct the process at low temperatures (max. 374°C), no need for a thermal treatment at high temperatures, high purity and quality of the products obtained, a small dispersion particle size, the ability to control the shape of the grains and nucleating agents. The increase in nucleation may be controlled by varying the concentration, increasing or decreasing the hydrolysis temperature and also increasing the duration of hydrolysis as well as by the introduction of surfactants during dehydration and changes in pH [62]. In contrast, the disadvantages of the hydrothermal methods include a high degree of complexity of the apparatus and its cost, as well as an inability to direct the observation of the process.

The microwave hydrothermal method, also called the solvothermal method, is a type of hydrothermal methods, which use microwave energy as a heat source. As a result, the reaction mixture does not have direct contact with the heating element and the time of whole process is considerably reduced. With the increase in pressure, the particles diameter increases and the monoclinic phase content in the product decreases. The result of the process is obtaining a material of high purity whose grain diameter is in the range of 5–50 nm [41, 42].

3.3. Synthesis of nano-zirconia by the sol-gel method

Another example of methods for the synthesis of nanometric zirconia is the sol-gel method. This type of chemical synthesis based on the specific chemical reactions taking place in solution, and in a further step on a series of transformations starting solution – first in the form of sol, then gel, ending with the conversion of the gel into a solid amorphous or crystalline structure [43, 44]. Therefore, this method is used to produce glass [45, 64, 65]. The conversion of the sol into gel is based on a dehydrated colloidal solution of material hydroxide particles, thereby producing a gelatinous substance, called the gel. These reactions are often carried

out with the use of alkoxides of various elements as starting materials [43]. In this case, the formation of a sol – gel system based on the hydrolysis and condensation reactions. The first reaction taking place in solution comprises of the hydrolysis (breakup) of alkoxy bond in the molecule Me-OR and the creation of hydroxyl bonds Me-OH. This reaction precedes in a mutual solvent for the alkoxide and water, with the participation of a suitable catalyst. This is followed by a condensation reaction between hydroxyl and alkoxy groups or two hydroxyl groups, resulting in the formation of bond type $(RO)_{n-1}Me\text{-}O\text{-}Me(OR)_{n-1}$, where *n* is the number of alkoxy groups in the molecule of alkoxide. The hydrolysis and condensation reactions occurring in parallel result in an increase in the number of Me-O bonds, which leads to initiation of the gelation process. The course of the described reaction depends on the temperature, pH, water and solvent content, polarity of the solvent, and type and concentration of the used catalyst. The catalysts, which are used in the sol-gel method, may be acidic or alkaline. In the first case, inorganic or organic acids are used, and in the second, it is mostly ammonium hydroxide. The increase in the number of Me-O bonds is the cause of the partial polymerization of alkoxides, which subsequently agglomerate. This results in an increase in the viscosity of the system, which, during prolonged drying, is converted to a gel by evaporation of the solvent [43]. In further steps, calcination is carried out.

The advantages of the sol-gel process include simplicity of use and low equipment cost, high purity raw materials and a high degree of homogeneity of the microstructure, the relatively low process temperature in comparison with synthesis methods from the gas phase like PVD (Physical Vapour Deposition) and CVD (Chemical Vapour Deposition), no toxic waste or its low participation, as well as the ability to receive color layers [43, 44]. In contrast, the disadvantages of this method include the high cost of chemical reagents, as well as what is one of the key issues – stability of technological parameters [43].

The sol-gel method is often used to receive various types of ceramic coatings. For the preparation of nanocrystalline sulfated zirconium oxide (IV), the process can be carried out in two stages. In the first step, hydrolysis and condensation take place, and in the second, sulfation with sulfuric acid or ammonium sulfate occurs. Whereas in a one-step method, the mentioned chemical reactions take place simultaneously [46]. The one-step synthesis method proposed by Mishra and colleagues [46] allowed to obtain a product with a particle diameter of 11–16 nm, specific surface area of 101–118 m^2/g and a diameter and pore size that were respectively 0.152–0.190 cm^3/g and 58–62 Å. These parameters (except for the diameter and pore size) are not very different from those, which characterize the product obtained by the two-step method. Particle diameter, surface area, pore diameter and pore size of the sulfated zirconium oxide (IV) were, respectively, 11 nm, 101 m^2/g, 0.087 cm^3/g and 37 Å [46]. All the obtained products were characterized by a tetragonal structure. Akkarin et al. [47], using the sol-gel method, have received nanocrystalline sulfated zirconium oxide (IV) on mesoporous silica used as a carrier. Zirconium acetylacetonate and tetraethyl orthosilicate were used as reactants. The sulfation reaction was carried out using *in situ* H_2SO_4. The authors [47] have also studied the effect of the amount of sulfuric acid added, as the S/Zr mole fraction, on structural properties. It has been shown that, with the mole fraction of 0–0.45, two dominant pore sizes, i.e. 3.5 nm and 10 nm, are observed. In contrast, when the S/Zr mole fraction is more

than 0.6, the pores with a diameter below 4.0 nm become negligible. It has also been observed that the increase of the S/Zr mole fraction in the range of 0.3 to 0.6 increased the gelling time, which resulted in a decrease of the pore volume [47]. De la Rosa et al. [48] received zirconium oxide (IV) doped with 0.5 wt%. La, Mn and Fe based on the sol-gel method. They analyzed the influence of the type of metallic precursor on the crystalline phase. The authors used zirconium alkoxide and precursors in the form of a metal acetate as substrates. The resulting products were characterized by a surface area in the range of 1.5 to 5.0 m^2/g and a pore size from 5 to approx. 60 nm. It has been shown that the zirconia doping with lanthanum and manganese promotes the formation of the monoclinic phase, and the dopant lanthanum-iron results in the formation of the tetragonal phase of ZrO_2. The resulting products were used in the catalytic combustion of trichlorethylene, so their thermal characteristics were carried out. The authors found an approximately 20% loss of mass when the gelation temperature was 80–200°C. This weight loss may be associated with the evaporation of physically absorbed water and ethanol from the solids. The second weight loss was observed at 200–380°C, which was explained as being the result of the alkoxy group oxidation [48]. The effect of the addition of the precursor metal on the zirconium dioxide crystal structure was also studied by Miyoshi et al. [66]. The authors demonstrated that even a small addition of yttrium (1 mol%) stabilizes the tetragonal phase. The effect of increasing the concentration of yttrium is the creation of equilibrium shares of monoclinic and tetragonal phases. The authors observed this effect in $Zr_{0.96}Y_{0.04}O_{1.98}$. This share phase change with increasing heat treatment temperature to the disadvantage of the monoclinic phase – in 1273 K the amount of this phase is negligible. For $Zr_{0.84}Y_{0.16}O_{1.92}$ both before and after heat treatment, the dominant phase in the crystal was a cubic phase [66].

The sol-gel method is an effective method for obtaining thin layers of ceramic. This has been proven in many publications [44, 45, 67, 68], but also by a combination of hydrothermal methods leading to the acquisition of nanocrystalline powders [32, 36].

3.4. Synthesis of nano-zirconia by microemulsion method

The microemulsion method has already been successfully applied for the preparation of nanoparticles by Boutonnet and co-workers [69] in 1982. The synthesis products by this technique have a uniform size distribution and dispersion. The essence of the method involves mixing together two microemulsions, where there are various reagents in the drops. Then, as a result of mixing, the emulsion liquid droplets collide and exchange reagents. The step of nucleation and crystal growth of product occurs within the droplets and is limited by their diameter (average diameter = 50 nm). In further steps, the reactions of the precipitation and calcination take place. A microemulsion is thermodynamically stable, isotropic and transparent dispersion, which consists of two mutually immiscible liquids – water and oil [51]. The third component of the emulsion is an emulsifier, which is introduced in the form of one or two surfactants with the assumption that the quantity does not exceed 10–15 mass%. of the entire system weight [56]. Into microemulsion, the addition of a co-surfactant is introduced. It prevents the so-called phenomenon of "Ostwald ripening". It occurs when the

monomer particles leak from small droplets of the dispersed phase into large particles through the continuous phase [56]. There are two main types of microemulsion – oil in water microemulsion (o/w) and water-in-oil (w/o) [55]. Most of the examples of zirconia nanoparticle synthesis using the microemulsion technique takes place in a water-in-oil system [53, 54, 70], but there are also known synthesis methods in an oil-in-water system, which are gaining in importance due to the possibility of the use of anionic surfactants [50, 52]. In the o/w microemulsion metal precursors and the precipitating agent are dissolved in the continuous phase, while in the w/o microemulsion, they are inside the droplets of the dispersed phase. Therefore, in the technique based on o/w microemulsion, the synthesis process is different. In a first step, a microemulsion is formed, and then metal cations, which will be adsorbed at the oil-water interface, are added. This is caused by Coulomb forces – the attraction between the metal cation and surfactant anion. The fact that the ions are strongly solvated in polar solvents is also not without significance. As a result, the cation is positioned at the oil-water interface, because it is stabilized by both water and surfactants molecules. After adding the flocculation agent to the system, the equilibrium is destroyed and consequently, at the interface, plurality of product particles are formed [52]. In the w/o microemulsions, the size and shape of the products depend largely on the used surfactant [37, 41, 42, 70], which is shown in Table 1.

Table 1. Examples of the zirconium dioxide synthesis using w/o microemulsion

Microemulsion components			Particle size (diameter) [nm]	Characteristics	Lit.
Oil phase	Water phase	Surfactant			
Heptane and cyclohexane	solution of zirconium salt	Span 80 + Arlacel 83 + isopropanol	$2 \cdot 10^3$–$8 \cdot 10^3$	spherical shape	71,72
Xylene	solution of zirconium and yttrium nitrate	Tween 80	$0{,}3 \cdot 10^3$–$1{,}0 \cdot 10^3$	fusiform and spherical shape	73
Cyclohexane	solution of zirconium salt	Triton X –100/ pentanol	4–20	fusiform shape and monoclinic	70
Cyclohexane	solution of zirconium salt	Triton X – 100/ hexyl alcohol	30–40	spherical shape	70
n-Octan	$ZrOCl_2$ solution	Span 80/Triton X-100/n-hexyl alcohol	7.2–23.7	spherical shape	53

The microemulsion method is a versatile technique for the synthesis of nanoparticles, which allows to control properties, such as size, geometry, morphology, homogeneity and surface [50]; therefore, it is regarded as one of the best synthesis methods. In addition, its application does not require the use of expensive and complicated laboratory equipment.

4. Conclusions

Nano-sized zirconia may be obtained from all phases of matter, ie. from the gas and the liquid phase as well as the solid. Methods for obtaining it are chosen based on the chemical synthesis or physical processes. This article discusses some methods for the synthesis of nano-zirconium dioxide from the liquid phase, focusing on the co-precipitation method, hydrothermal methods, sol-gel method and the microemulsion method. All of these methods have advantages and disadvantages. A great advantage of most of them is the lack of the need to use expensive and complicated laboratory equipment, the only exception being the hydrothermal method. Most liquid methods do not require high processing temperatures, unlike the gas phase synthesis methods. Products can be obtained both in the form of powders as well as layers deposited on the substrate. The big advantage is the ability to conduct modification of the final structure of the material, e.g. by applying suitable metal precursors and by carrying out the synthesis under appropriate conditions and in the presence of certain chemicals, e.g. surfactants. Thus, there are several ways to modify the methods of nanometric zirconia synthesis. Based on the cited literature, certain trends of the methods discussed can be seen as well. Hydrothermal methods are supported by the action of microwave energy and ultrasound. It also seems that, in the near future, interest in polymers, substances limiting the agglomeration of particles in aqueous solutions, will increase. Previously quoted example – PEG 800, which did not just cause a limit of the size of zirconium oxide particles, can provide a stimulus to explore this issue. The use, instead of water, volatile organic solvents, having a lower polarity than water, is also a direction of development of nano compound synthesis methods. We cannot determine which method of synthesis is the best because, when considering this issue, the following should be taken into account: access to equipment and chemical reagents, any financial outlay and predispositions as well as skills of people involved in the synthesis. The choice of synthesis methods can also be determined by the type of final material, e.g. powder or layered material.

References

[1] Pulit J., Banach M., Kowalski Z., *Chemical reduction as the main method for obtaining nanosilver*, Journal of Computational Theoretical Nanoscience 10, 2, 2013, 276–284.

[2] Marzec A., Pulit J., Kwaśny J., Banach M., *Nanometale– wybrane technologie wytwarzania*, Technical Translations vol. 1-Ch/2012, 95–107.

[3] Swihart M. T., *Vapor-phase synthesis of nanoparticles*, Current Opinion in Colloid and Interface Science 8, 2003, 127–133.

[4] Reguła T., Darłak P., Tchórz A., Lech-Grega M., *Próba wytworzenia kompozytu na osnowie Cu_xAl_y zbrojonego cząsteczkami Al_2O_3 przy pomocy procesu mechanosyntezy*, Prace Instytutu Odlewnictwa 1, 2010, 29–35.

[5] Goharshadi E. K., Hadadian M., *Effect of calcination temperature on structural, vibrational, optical, and rheological properties of zirconia nanoparticles*, Ceramics International 38 (3), 2012, 1771–1777.

[6] Xie Z., Ma J., Xu Q., Huang Y., Cheng Y.-B., *Effects of dispersants and soluble counter-ions on aqueous dispersability of nano-sized zirconia powder*, Ceramics International 30, 2004, 219–224.
[7] Khare J., Srivastava H., Singh C.H.P., Joshi M.P., Kukreja L.M., *Vapor phase synthesis of hexagonal shaped single crystal yttria stabilized zirconia nanoparticles using CO_2 laser*, Ceramics International 39, 2013, 1103–1109.
[8] Simchi A., Ahmadi R., Seyed Reihani S.M., Mahdavi A., *Kinetics and mechanisms of nanoparticle formation and growth in vapor phase condensation process*, Materials and Design 28, 2007, 850–856.
[9] Vasilyeva E.S., Tolochko O.V., Kim B.K., Lee D.W., Kim D.S., *Synthesis of tungsten disulphide nanoparticles by the chemical vapor condensation method*, Microelectronics Journal 40, 2009, 687–691.
[10] Gavillet J., Belmonte T., Hertz D., Michel H., *Low temperature zirconia thin film synthesis by a chemical vapour deposition process involving $ZrCl_4$ and O_2–H_2–Ar microwave post-discharges. Comparison with a conventional CVD hydrolysis process*, Thin Solid Films 301, 1997, 35–44.
[11] Srdic V.V., Winterer M., Miehe G., Hahn H., *Different zirconia-aluminia nanopowders by modifications of chemical vapour synthesis*, Nanostructured Materials 12, 1999, 95–100.
[12] Choi H.-S., Ryu C.-H., Hwang G.-J., *Obtention of ZrO_2–SiO_2 hydrogen permselective membrane by chemical vapor deposition method*, Chemical Engineering Journal 232, 2013, 302–309.
[13] Jiang J., Shen W., Hertz J.L., *Fabrication of epitaxial zirconia and ceria thin films with arbitrary dopant and host atom composition*, Thin Solid Films 522, 2012, 66–70.
[14] Yeh T.-H., Lin R.-D., Cherng J.-S., *Significantly enhanced ionic conductivity of yttria-stabilized zirconia polycrystalline nano-film by thermal annealing*, Thin Solid Films 544, 2013, 148–151.
[15] Hass D.D., Zhao H., Dobbins T., Allen A.J., Slifka A.J., Wadley H.N.G., *Multi-scale pore morphology in directed vapor deposited yttria-stabilized zirconia coatings*, Materials Science and Engineering A 527, 2010, 6270–6282.
[16] Li H., Khor K.A., Kumar R., Cheang P., *Characterization of hydroxyapatite/nano-zirconia composite coatings deposited by high velocity oxy-fuel (HVOF) spray process*, Surface and Coatings Technology 182, 2004, 227–236.
[17] Joulia A., Bolelli G., Gualtieri E., Lusvarghi L., Valeri S., Vardelle M., Rossignol S., Vardelle A., *Comparing the deposition mechanisms in suspension plasma spray (SPS) and solution precursor plasma spray (SPPS) deposition of yttria-stabilised zirconia (YSZ)*, Journal of the European Ceramic Society 34, 2014, 3925–3940.
[18] Dong H., Yang G.-J., Cai H.-N., Li C.-X., Li C.-J., *Propagation feature of cracks in plasma-sprayed YSZ coatings under gradient thermal cycling*, Ceramics International 41, 2015, 3481–3489.
[19] Dong H., Yang G.-J., Cai H.-N., Ding H., Li C.-X., Li C.-J., *The influence of temperature gradient across YSZ on thermal cyclic lifetime of plasma-sprayed thermal barrier coatings*, Ceramics International 41, 2015, 11046–11056.

[20] Gao L., Wei L., Guo H., Gong S., Xu H., *Deposition mechanisms of yttria-stabilized zirconia coatings during plasma spray physical vapor deposition,* Ceramics International 42, 2016, 5530–5536.
[21] Smits K., Grigorjeva L., Millers D., Kundzins K., Ignatans R., Grabis J., Monty C., *Luminescence properties of zirconia nanocrystals prepared by solar physical vapor deposition,* Optical Materials 37, 2014, 251–256.
[22] Bernard O., Huntz A.M., Andrieux M., Seiler W., Ji V., Poissonnet S., *Synthesis, structure, microstructure and mechanical characteristics of MOCVD deposited zirconia films,* Applied Surface Science 253, 2007, 4626–4640.
[23] Hemmer E., Kumakiri I. et al., *Nanostructured ZrO_2 membranes prepared by liquid-injection chemical vapor deposition,* Microporous and Mesoporous Materials 163, 2012, 229–236.
[24] Shi G., Yu F., Wang Y., Li R., *Synthesis of growth-controlled ZrO_2 nanocrystals via vapor phase hydrolysis,* Ceramics International 40, 2014, 13083–13088.
[25] Djurado E., Dessemond L., Roux C., *Phase stability of nanostructured tetragonal zirconia polycrystals versus temperature and water vapor,* Solid State Ionics 136–137, 2000, 1249–1254.
[26] Liu S., Jiang K., Zhang H., Liu Y., Zhang L., Su B., Liu Y., *Nano-nano composite powders of lanthanum–gadolinium zirconate and gadolinia-stabilized zirconia prepared by spray pyrolysis,* Surface & Coatings Technology 232, 2013, 419–424.
[27] Amézaga-Madrid P., Hurtado-Macías A., Antúnez-Flores W., Estrada-Ortiz F., Pizá-Ruiz P., Miki-Yoshida M., *Synthesis, microstructural, optical and mechanical properties of yttria stabilized zirconia thin films,* Journal of Alloys and Compounds 536S, 2012, S412–S417.
[28] Tao K., Dou H., Sun K., *Interfacial coprecipitation to prepare magnetite nanoparticles: Concentration and temperature dependence,* Colloids and Surfaces A: Physicochem. Eng. Aspects 320, 2008, 115–122.
[29] Chen Q., Rondinone A.J., Chakoumakos B.C., Zhang Z.J., *Synthesis of superparamagnetic $MgFe_2O_4$ nanoparticles by coprecipitation,* Journal of Magnetism and Magnetic Materials 194, 1999, 1–7.
[30] Song J.E., Lee D.K., Kim H.W., Kim Y.I., Kang Y.S., *Preparation and characterization of monodispersed indium–tin oxide nanoparticles,* Colloids and Surfaces A: Physicochem. Eng. Aspects 257–258, 2005, 539–542.
[31] Benavente R., Salvador M.D., Alcázar M.C., Moreno R., *Dense nanostructured zirconia compacts obtained by colloidal filtration of binary mixtures,* Ceramics International 38, 2012, 2111–2117.
[32] Chang Q., Zhou J., Wang Y., Meng G., *Formation mechanism of zirconia nano-particles containing pores prepared via sol–gel-hydrothermal method,* Advanced Powder Technology 21, 2010, 425–430.
[33] Chintaparty R., Palagiri B., Nagireddy R.R., Immareddy V.R., Madhuri W., *Effect of pH on structural, optical and dielectric properties of nano-zirconium oxide prepared by hydrothermal method,* Materials Letters 161, 2015, 770–773.

[34] Kumar R. V., Ghoshal A.K., Pugazhenthi G., *Fabrication of zirconia composite membrane by in-situ hydrothermal technique and its application in separation of methyl orange*, Ecotoxicology and Environmental Safety 121, 2015, 73–79.
[35] Yoshimura M., Sōmiya S., *Hydrothermal synthesis of crystallized nano-particles of rare earth-doped zirconia and hafnia*, Materials Chemistry and Physics 61, 1999, 1–8.
[36] Chang Q., Zhou J., Wang Y., Meng G., *Preparation and characterization of unique zirconia crystals within pores via a sol–gel-hydrothermal method*, Advanced Powder Technology 20, 2009, 371–374.
[37] Behbahani A., Rowshanzamir S., Esmaeilifar A., *Hydrothermal synthesis of zirconia nanoparticles from commercial zirconia*, Procedia Engineering 42, 2012, 908–917.
[38] Huang H.-L., Cao G.Z., Shen I.Y., *Hydrothermal synthesis of lead zirconate titanate (PZT or $Pb(Zr_{0.52}Ti_{0.48})O_3$) nano-particles using controlled ramping and cooling rates*, Sensors and Actuators A 214, 2014, 111–119.
[39] Ji X., Liu C. et al., *Lauric acid template synthesis of thermally stable lamellar crystalline zirconia via a reflux-hydrothermal route*, Materials Letters 122, 2014, 309–311.
[40] Chintaparty C. R., *Influence of calcination temperature on structural, optical, dielectric properties of nano zirconium oxide*, Optik 127, 2016, 4889–4893.
[41] Ao H., Liu X., Zhang H., Zhou J., Huang X., Feng Z., Xu H., *Preparation of scandia stabilized zirconia powder using microwave-hydrothermal method*, Journal of Rare Earths 33, 7, 2015, 746–751.
[42] Li C., Li K., Li H., Zhang Y., Ouyang H., Liu L., Sun C., *Effect of reaction temperature on crystallization of nanocrystalline zirconia synthesized by microwave-hydrothermal process*, Journal of Alloys and Compounds 561, 2013, 23–27.
[43] Porębska K., *Powłoki hydrofobowe na baize SiO_2 wytwarzane metodą zol-żel*, Budownictwo i Architektura 12(4), 2013, 257–267.
[44] Walczak M., *Charakterystyka powłok ceramicznych SiO_2 i SiO_2–TiO_2 otrzymywanych metodą zol-żel*, Postępy Nauki i Techniki 9, 2011, 80–90.
[45] Persson C., Unosson E., Ajaxon I., Engstrand J., Engqvist H., Xia W., *Nano grain sized zirconia–silica glass ceramics for dental applications*, Journal of the European Ceramic Society 32, 2012, 4105–4110.
[46] Mishra M.K., Tyagi B., Jasra R.V., *Synthesis and characterization of nano-crystalline sulfated zirconia by sol–gel method*, Journal of Molecular Catalysis A: Chemical 223, 2004, 61–65.
[47] Akkari R., Ghorbel A., Essayem N., Figueras F., *Synthesis and characterization of mesoporous silica-supported nano-crystalline sulfated zirconia catalysts prepared by a sol–gel process: Effect of the S/Zr molar ratio*, Applied Catalysis A: General 328, 2007, 43–51.
[48] De la Rosa J.R., Hernandez A., Rojas F., Ledezma J.J., *Sol–gel synthesis and characterization of novel La, Mn and Fe doped zirconia: Catalytic combustion activity of trichloroethylene*, Colloids and Surfaces A: Physicochem. Eng. Aspects 315, 2008, 147–155.
[49] López-Quintela M.A., Tojo C., Blanco M.C., García Rio L., Leis J.R., *Microemulsion dynamics and reactions in microemulsions*, Current Opinion in Colloid & Interface Science 9, 2004, 264–278.

[50] Malik M.A., Wani M.Y., Hashim M.A., *Microemulsion method: A novel route to synthesize organic and inorganic nanomaterials*, Arabian Journal of Chemistry 5, 2012, 397–417.
[51] Margulis-Goshen K., Magdassi S., *Organic nanoparticles from microemulsions: Formation and applications*, Current Opinion in Colloid & Interface Science 17, 2012, 290–296.
[52] Sanchez-Dominguez M., Pemartin K., Boutonnet M., *Preparation of inorganic nanoparticles in oil-in-water microemulsions: A soft and versatile approach*, Current Opinion in Colloid & Interface Science 17, 2012, 297–305.
[53] Duan G.-R., Yang X.-J., Huang G.-H., Lu L.-D., Wang X., *Water/span80/Triton X-100/n-hexyl alcohol/n-octane microemulsion system and the study of its application for preparing nanosized zirconia*, Materials Letters 60, 2006, 1582–1587.
[54] Tai C. Y., Hsiao B.-Y., Chiu H.-Y., *Preparation of spherical hydrous-zirconia nanoparticles by low temperature hydrolysis in a reverse microemulsion*, Colloids and Surfaces A: Physicochem. Eng. Aspects 237, 2004, 105–111.
[55] López-Quintela M.A., Rivas J., Blanco M.C., Tojo C., *Synthesis of nanoparticles in microemulsions*, [in:] L. M. Liz-Marzán, P. V. Kamat (Eds.), *Nanoscale Materials*, Springer US, 2003, 135–155.
[56] Witek E., Kochanowski A., Pazdro M., Bortel E., *Mikroemulsje jako źródło nanolateksów i nanoreaktorów*, Polimery 51, 2006, 507–516.
[57] Zhou M., Xu L. et al., *Investigation on the preparation and properties of monodispersed Al_2O_3–ZrO_2 nanopowder via Co-precipitation method*, Journal of Alloys and Compounds 678, 2016, 337–342.
[58] Hsu Y.-W., Yang K.-H., Chang K.-M., Yeh S.-W., Wang M.-C., *Synthesis and crystallization behavior of 3 mol% yttria stabilized tetragonal zirconia polycrystals (3Y-TZP) nanosized powders prepared using a simple co-precipitation process*, Journal of Alloys and Compounds 509, 2011, 6864–6870.
[59] Lan L., Chen S., Cao Y., Zhao M., Gong M., Chen Y., *Preparation of ceria–zirconia by modified coprecipitation method and its supported Pd-only three-way catalyst*, Journal of Colloid and Interface Science 450, 2015, 404–416.
[60] Aruna S.T., Arul Paligan B., Balaji N., Praveen Kumar V., *Properties of plasma sprayed yttria stabilized zirconia thermal barrier coating prepared from co-precipitation synthesized powder*, Ceramics International 40, 2014, 11157–11162.
[61] Wang S., Li X., Zhai Y., Wang K., *Preparation of homodispersed nano zirconia*, Powder Technology 168, 2006, 53–58.
[62] Dudnik E.V., *Modern methods for hydrothermal synthesis of ZrO_2-based nanocrystalline powders*, Powder Metallurgy and Metal Ceramics 48, 3–4, 2009, 238–248.
[63] Caillot T., Salama Z., Chanut N., Cadete Santos Aires F.J., Bennici S., Auroux A., *Hydrothermal synthesis and characterization of zirconia based catalysts*, Journal of Solid State Chemistry 203, 2013, 79–85.
[64] Montazerian M. et al., *Bioactivity and cell proliferation in radiopaque gel-derived CaO-P_2O_5-SiO_2-ZrO_2 glass and glass–ceramic powders*, Materials Science and Engineering C 55, 2015, 436–447.

[65] Montazerian M. et al., *Sol–gel synthesis, structure, sintering and properties of bioactive and inert nano-apatite–zirconia glass–ceramics*, Ceramics International 41, 2015, 11024–11045.
[66] Miyoshi S., Akao Y., Kuwata N., et al., *Water uptake and conduction property of nano-grained yttria-doped zirconia fabricated by ultra-high pressure compaction at room temperature*, Solid State Ionics 207, 2012, 21–28.
[67] You H.C., Chang C.-M., et al., *Facile preparation of sol–gel-derived ultrathin and high-dielectric zirconia films for capacitor devices*, Applied Surface Science 258, 2012, 10084–10088.
[68] Díaz-Parralejo A., Macías-García A., Sánchez-González J., Díaz-Díez M.Á., Cuerda--Correa E.M., *Influence of the experimental parameters on the synthesis process of yttria-doped zirconia sol-gel films*, Surface & Coatings Technology 204, 2010, 2257–2261.
[69] Boutonnet M, Kizzling J, Stenius P, *The preparation of monodisperse colloidal metal particles from microemulsions*, Colloids and Surfaces 5, 1982, 209–225.
[70] Ma T., Huang Y., Yang J., He J., Zhao L., *Preparation of spherical zirconia powder in microemulsion system and its densification behavior*, Materials and Design 25, 2004, 515–519.
[71] Lee M.H., Tai C.Y., Lu C.J., *Synthesis of spherical zirconia by precipitation between two water/oil emulsions*, Journal of the European Ceramic Society 19, 1999, 2593–2603.
[72] Tai C.Y., Lee M.H., Wu Y.C., *Control of zirconia particle size by using two-emulsion precipitation technique*, Chemical Engineering Science 56, 2001, 2389–2398.
[73] Qiu H.B., Gao L., Qiao H.C., Guo J.K., Yan D.S., *Nano-crystalline zirconia powder processing through innovative wet-chemical methods*, Nanostructured Materials 6, 1995, 373–376.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.17.022.6215

Dawid Łątka (dawid.latka@pk.edu.pl)
Institute of Building Materials and Structures, Faculty of Civil Engineering, Cracow University of Technology

Stress state laboratory verification in masonry structures according to the *flat jack* method

Laboratoryjna weryfikacji stanu naprężenia w konstrukcji murowej według metody *flat jack*

Abstract

In the article, the value of compressive stress defined according to the *flat jack* method with a theoretical value of the stress was compared. The test was performed in a laboratory using part of the 38 cm thick masonry wall according to the procedure described in the ASTM C1196 – 14a Standard. Significant correspondence between the obtained results confirms that this diagnostic method is useful in Polish conditions as well as it allows one to estimate approximately how accurate it is.

Keywords: semi-destructive tests, flat jack, masonry structure, stress state

Streszczenie

W artykule dokonano porównania wartości naprężenia ściskającego określonego według metody *flat jack* z wartością teoretyczną tego naprężenia. Test wykonano w laboratorium na fragmencie ściany murowej o grubości 38 cm według procedury opisanej w normie ASTM C1196 – 14a. Uzyskana duża zgodność wyników potwierdza przydatność metody diagnostycznej w warunkach polskich, jak również pozwala na orientacyjne oszacowanie jej dokładności.

Keywords: semi-destructive tests, flat jack, masonry structure, stress state

1. Standard adopted as basis for tests according to the flat jack method (FJ)

As part of a typical diagnostic procedure of historic masonry structures, tests according to the flat jack method cannot be missing. This method is the basis for evaluating the existing masonry structures in the United States as well as in the countries of Western and Southern Europe [1]. Unfortunately, in Poland, this method is known mainly from foreign literature and several domestic items (for example [2, 3]) and so far, it has been applied only by the author of the article to diagnose the existing buildings (for example, a historic building of the Municipal Theatre in Gliwice, Poland).

The simplest test using pressure flat jack, which is determining a level of compressive stress, is carried out in three steps. First of all, the length of the measurement base on the surface tested is read, and then, the joint located between these bases is removed; in order to do this, the most indicated solution is to use a circular saw. Thereby, the slot formed is smooth enough. As step two, mortar, which is removed, is replaced by a thin (3–8 mm) flat jack made of two smooth leaktight metal sheets joined together with tubes to fill the flat jack with a liquid medium, to vent it and to bleed it (pressure reduction). To make the flat jack thickness better match the width of the joint, a compensating insert is used and the pressure flat jack is preliminarily pumped so that the pressure value reaches half of the expected target value. The last step is to increase the pressure inside the flat jack using an external hydraulic pump and, at the same time, the length of the measurement bases is monitored. The test ends when each of the bases reaches the length at least equal to the initial value.

The procedure described above is founded by the American ASTM C1196 Standard [4] and by the European RILEM MDT.D.4. Instruction [5]. However, verification of the stress state does not exhaust all the flat jacks possibilities; they are also used to determine the stress – strain relationship (Young's modulus) according to ASTM C1197 [6] or RILEM MDT.D.5. [7] or the masonry mortar joint shear strength index according to ASTM C1531 [8].

2. Laboratory verification of the forced stress state

2.1. The purpose of the tests

In practice, to make the diagnosis, several shapes of pressure flat jacks are used, depending on the country where the tests are carried out, on the dimensions of masonry components and the values tested. In European studies, rectangular and oval-rectangular flat jacks are most commonly used. The purpose of this article is to assess how reliable this method is by comparing the value of the forced compressive stress to the value of this stress when using the most popular flat jack, i.e. the oval-rectangular. The methodology described in the introduction, i.e. methodology being in full compliance with the current ASTM C1196 Standard [4], was used in the testing. The main advantage is conducting tests using the 1:1 scale member, which, in contrast to the only existing domestic tests [9], will allow the author to obtain results similar to those that can in fact be expected.

2.2. Test bench description and equipment used

The subject of the tests was a wall sized 168 × 38 × 150 cm made of solid bricks (250 × 120 × 60 mm) bonded with lime-cement mortar. In order to maintain a reproduction of the actual conditions as well as possible, the masonry work was carried out by qualified masons using spacing strips to obtain a constant thickness of joint equal to 10–12 mm. To distribute the load more evenly on the whole width of the wall, the area tested was located below the half of its height. In addition, on the side tested as well as on the opposite side, 5 measurement bases were positioned to assess whether or not the wall was unevenly deformed. To record displacements in the area tested, 28 measurement bases in 7 columns (A to G) with 5 cm spacing were mounted on it (5, 10, 15 and 30 cm long). The bed joint designed for tests was placed right through the middle of all measurement bases. This condition is presented in Fig. 1 in detail.

A kit that was complete and suitable for a high-pressure flat jack, consisting of five components shown in Fig. 2, was used for testing. The pressure flat jack was made of two metal sheets 0.8 mm thick each. To make sure the readouts are highly stable, all screw connections were sealed with Teflon tape wound up to the conical threads.

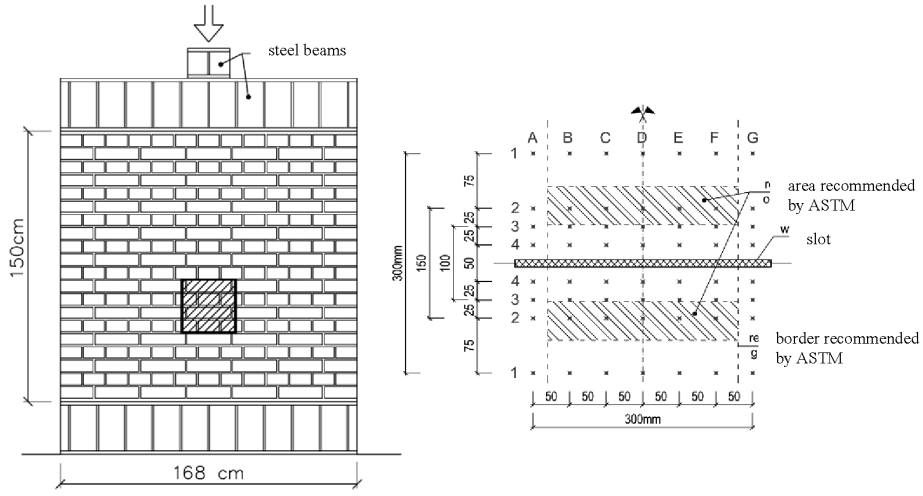


Fig. 1. Brickwall tested with marked measurement area (on the left) and picture of measurement bases made in this area (on the right)

The slot was made using a specialised peripheral-driven saw (Fig. 3) in dry cutting mode, which allowed to eliminate the wall moisture disturbances. The tests were conducted in the following sequence:

1. Measurement No. 1a of control bases located peripherally.
2. Wall load with known value of force – 980 kN.
3. Measurement No. 2a of control bases located peripherally in order to evaluate disturbances (e.g. possible eccentricities).

4. Measurement No. 1b of bases in the area of direct impact of the flat jack.
5. Tempering cutting – slot creation – Fig. 3.
6. Measurement No. 2b of bases in the area of direct impact of the flat jack.
7. Flat jack installation and preliminary pumping (in order to improve adhesion).
8. Measurement are taken No. 3b of bases in the area of direct impact of the flat jack by a predefined value of oil pressure until all the bases return to the length as in the 1b measurement – achieving the value of compensating pressure.

Fig. 2. Flat jack kit used to evaluate stress level – from the left: hydraulic hand pump, manometer, flexible hose, stopcock, oval-rectangular flat jack

Fig. 3. High precision slot cutting using the RING type circular saw and designer stand (on the left) and view of the wall during the flat jack test (on the right)

2.3. Determining the oil compensation pressure – p_{teor}

Readouts in vertical sections A and G proved to be impossible due to the collision of measuring devices with tubes transporting oil to/from the flat jack (in addition, they were located beyond the area recommended by [4] that has been shaded with lines in Fig. 1).

The other results recorded during the test were divided into 4 groups according to the length of the measurement base and are presented in sequence in graphs 1 to 4. The

relative level of base length return to the input value shown in graphs shows for which value of oil pressure (the so-called theoretical pressure, p_{teor}) the distance between the measuring points returns the value before sawing (ordinate 100% – compensation pressure). Due to the fact that the bases' return speed to their input length was differentiated, the interval of the compensation pressure value obtained was marked with vertical arrows. The detailed values of compensation pressure are shown in Table 1. For each of the bases and between them, the results were highly convergent – the standard deviation and the variation coefficient for (5; 10; 15; 30) cm bases amounted to, respectively: (0.10; 0.07; 0.08; 0.05) MPa and (4.5; 3.4; 3.7; 2.4)%.

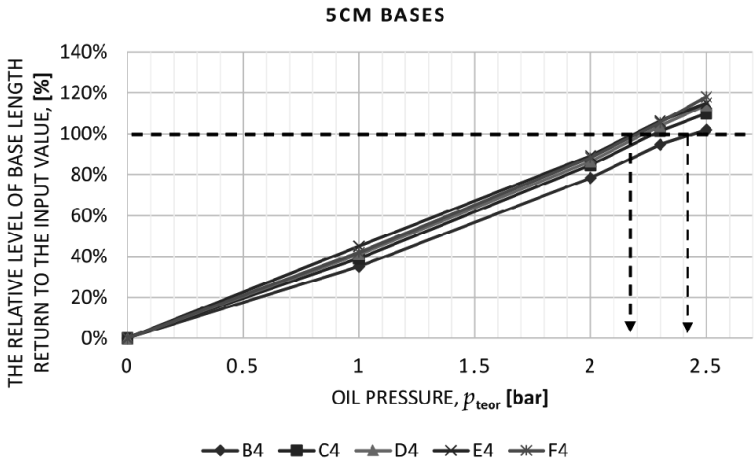


Fig. 4. 5 cm bases returning to the initial length in the theoretical pressure function

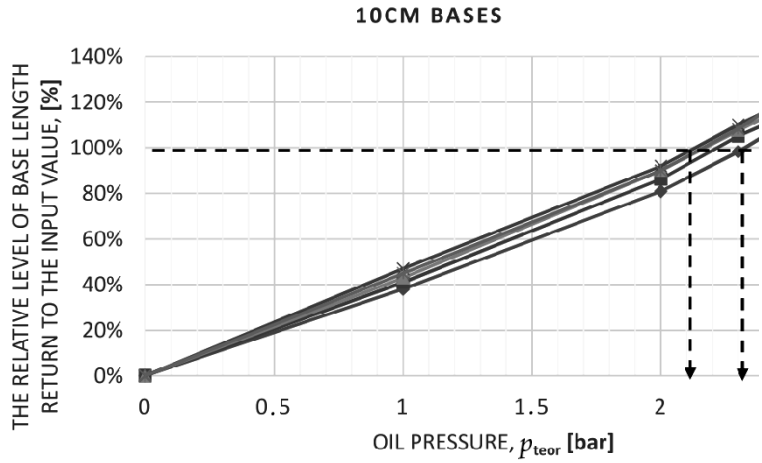


Fig. 5. 2. 10 cm bases returning to the initial length in the theoretical pressure function

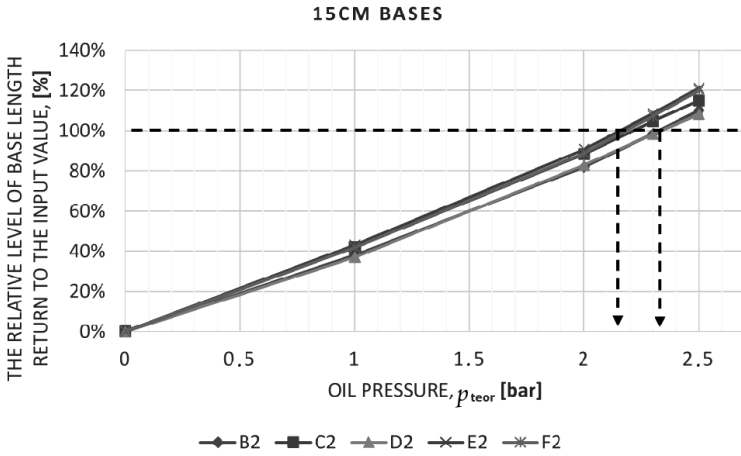


Fig. 6. 15 cm bases returning to the initial length in the theoretical pressure function

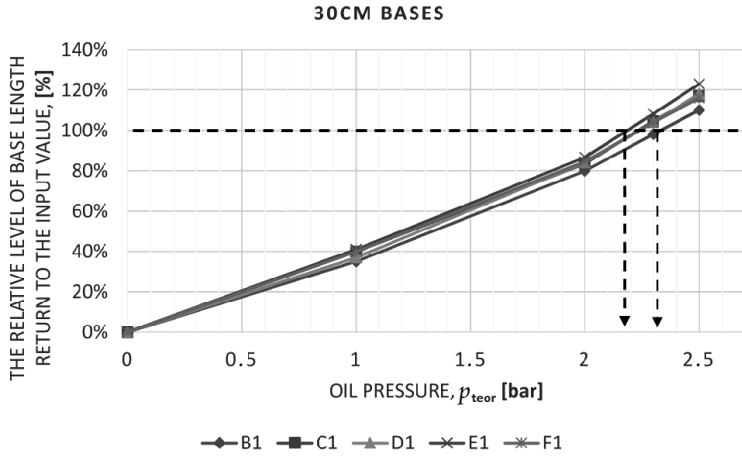


Fig. 7. 30 cm bases returning to the initial length in the theoretical pressure function

Table 1. The detailed values of the compensating pressure

Base	5 cm					10 cm				
Symbol	B4	C4	D4	E4	F4	B3	C3	D3	E3	F3
Comp. Pressure [MPa]	**2.44**	2.27	2.23	2.19	2.20	**2.32**	2.22	2.17	2.13	2.16
Base	**15 cm**					**30 cm**				
Symbol	B2	C2	D2	E2	F2	B1	C1	D1	E1	F1
Comp. Pressure [MPa]	**2.33**	2.21	2.33	2.16	2.18	**2.33**	2.23	2.24	2.18	2.24

2.4. Determining the correction coefficients – K_m, K_a

Unfortunately, the value of compensation pressure of oil cannot be directly identified with the value of the compression stress in the tested wall. This is due to the flat jack's own rigidity as well as due to the fact that it does not totally fill in the slot that was created after sawing (the surface of the flat jack is smaller than the cut-out). This is also why two correction coefficients should be entered, respectively: K_m, taking into account the reduction in flat jack's clamping to the wall and K_a, taking into account the flat jack's surface and slot's surface. Both coefficients take values from the $(0;1]$ interval.

The K_m coefficient should be determined before the first flat jack test is made. In order to do this, a laboratory test should be performed by placing the flat jack between the plates of a testing machine, and then by a trial flat jack load with oil pressure until the maximum value is reached. What is obtained as a result is a relationship between the oil pressure in the flat jack and the resistance posed by the machine's plates when preventing flat jack from being deformed. The procedure should be repeated 3 times and, as a result, an average value should be taken. A detailed description of the procedure is described in [4].

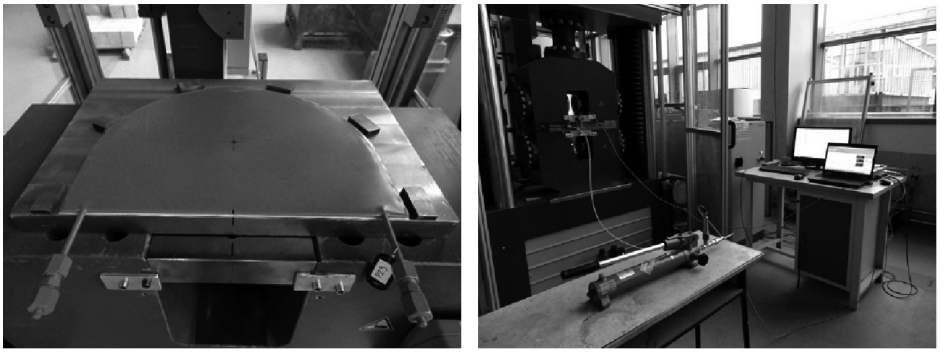


Fig. 8. Work station for determining the K_m coefficient

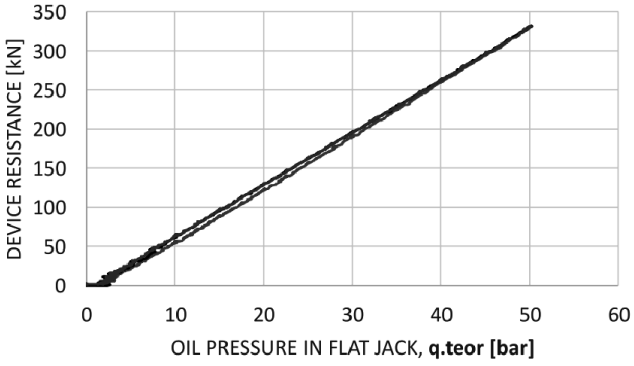


Fig. 9. The relationship between the device resistance and the oil pressure in the flat jack

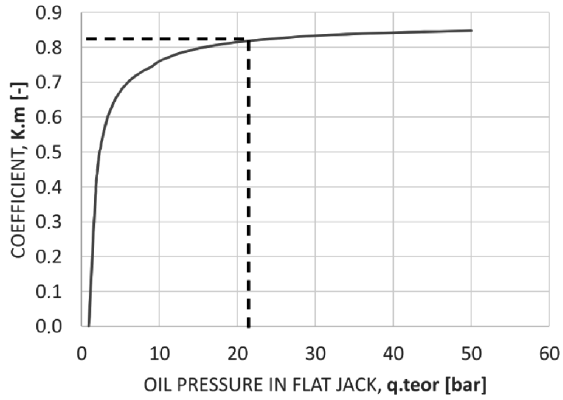


Fig. 10. The function of K_m coefficient

By analyzing the relationship between the oil pressure and the machine's resistance (graph 5), 3 specific intervals of oil pressure can be identified:

< 2 bars – flat jack's rigidity makes it impossible to transfer the load

[2;10] bars – measurement possible but poses a problem

> 10 bars – the most accurate measurement

Assuming that the resulting value of the test is the average value of the compensation pressure for the 15cm long base, the sought K_m (2.24 bar) coefficient = 0.82 (graph 6).

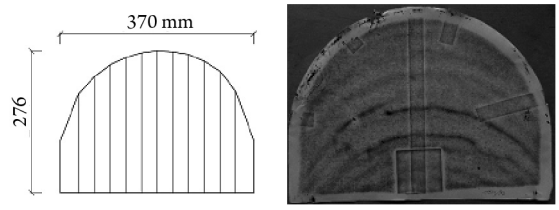


Fig. 11. The shape of the measured slot (on the left) and the flat jack active area (on the right)

To determine the value of K_a coefficient, the slot surface before testing was measured (864 cm^2) (Fig. 4) and the flat jack surface before testing was measured to (779 cm^2). The coefficient, which is the quotient of these values, was equal to 0.90.

In order to verify whether the flat jack adheres or not to the inner surface of the slot (whether or not the load is evenly distributed), carbon paper was placed. On the carbon paper, the flat jack's active area during loading was recorded (Fig. 4).

2.5. Comparison of the results obtained – evaluation of the method's reliability

When analysing the data contained in Table 2, almost no difference was observed in the average compensation pressure depending on the length of the measurement base. This means that, if a sufficient number of bases planned may not be applied according to standards adopted, other bases can be used.

Table 2. Verification of test results obtained

Base length	K_m coef.	K_a coef.	Avg. compensation pressure	Stress calculated	Average stress in the wall	Relative difference
[cm]	[–]	[–]	[MPa]	[MPa]	[MPa]	[%]
5	0.82	0.90	2.27	1.67	1.54	8.4
10			2.20	1.62		5.2
15			**2.24**	**1.65**		**7.1**
30			2.24	1.65		7.1

Ultimately, the pressure value evaluated according to the flat jack method was 1.65 MPa, while the average stress in the wall was 1.54 MPa. The difference is small, i.e. 7%, since deformation recorded on opposite sides of the wall is not homogeneous. The ratio of these deformations was 1.18, which would confirm that stresses on the tested side of the wall were higher.

3. Final conclusions

In terms of the methodology used in the article, the flat jack method is simple and effective way of conducting tests, which can be successfully used in diagnosing real objects in domestic conditions. However, you should bear in mind that, in order to make sure that the results are sufficiently reliable, professional equipment and extensive experience is necessary. The difference between the stress in theory and the stress measured was 7%. According to [2, 5, 10], this value does not normally exceed 15–20%, and we managed to demonstrate it. Previous domestic testing [9] produced small differences of about 1% as a result of using a very large surface flat jack compared to the size of the tested components, which is impossible in the case of on-site testing.

References

[1] Binda L., Saisi A., *Application of NDTs to the diagnosis of Historic Structures*, 7th International Symposium on Nondestructive Testing in Civil Engineering, July 2009, Nantes, France.

[2] Matysek P., *Identyfikacja wytrzymałości na ściskanie i odkształcalności murów ceglanych w obiektach istniejących*, Wydawnictwo PK, Kraków 2014.

[3] Stawiski B., *Konstrukcje murowe. Naprawy i wzmocnienia*, POLCEN, Warszawa 2014.

[4] ASTM C1196-14a, *Standard test method for in situ compressive stress within solid unit masonry estimated using flatjack measurements*, ASTM International, West Conshohocken, PA, 2014.

[5] RILEM Recommendation MDT.D.4. *In situ stress tests based on the flat-jack,* Materials and Structures, Vol 37, 2004, p.491–496.
[6] ASTM C1197-14a: *Standard test method for in situ measurement of masonry deformability properties using the flatjack method,* ASTM International, West Conshohocken, PA, 2014.
[7] RILEM Recommendation MDT.D.5, *In situ stress-strain behavior test based on the flat-jack,* Materials and Structures, Vol. 37, 2004, p. 497–501.
[8] ASTM C1531-15: *Standard test methods for in situ measurement of masonry mortar joint shear strength index,* ASTM International, West C., PA, 2015.
[9] Szwarowicz A., *Metoda badań wytrzymałościowych konstrukcji murowych in situ za pomocą poduszek ciśnieniowych*, ITB quarterly 4/2002, 2002, p.93–109.
[10] Noland J.L., Atkinson R.H., Schuller M.P., *Masonry evaluation using the flatjack method,* konferencja pt. Nondestructive evaluation of civil structures and materials, Colorado 1990.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
ENVIRONMENTAL ENGINEERING
DOI: 10.4467/2353737XCT.17.023.6216

Michael Poehler
Paul Uwe Thamsen
Chair of Fluid System Dynamics, Department of Fluid Dynamics and Technical Acoustics, Technische Universität Berlin

Design of a decentralised intelligent network for five wet pit pumping stations

Projektowanie zdecentralizowanej inteligentnej sieci pięciu przepompowni mokrych

Abstract

This paper presents a design for a decentralised intelligent network of five wet pit pumping stations in a rural area. This enables the network to automatically respond to inflow conditions divergent to the base flow and use the storage capacity of the system to prevent sewer overflows. First results for the base flow and a simulation of the network in EPANET are presented.

Keywords: base flow of wastewater, intelligent network, wet pit

Streszczenie

W artykule przedstawiono projekt zdecentralizowanej inteligentnej sieci pięciu przepompowni mokrych położonych w obszarze wiejskim. Sieć umożliwia automatyczną reakcję w warunkach przypływu przekraczających wartość przyjętego natężenia przepływu ścieków, a jej pojemność magazynowa pozwala zapobiec przelewaniu się ścieków. Przedstawiono także pierwsze wyniki dotyczące natężenia przepływu ścieków oraz model sieci w EPANET.

Słowa kluczowe: natężenie przepływu ścieków, inteligentna sieć, przepompownia mokra

1. Introduction

Wastewater infrastructure for rural areas is characterized by extensive pipe length and, due to small populations, small volumes of sewage. To run a wastewater network economically, the wastewater utilities have a big area to operate. This makes the maintenance and monitoring of the system very personnel-intensive.

However, existing wastewater infrastructures are increasingly facing growing demands, especially in rural areas. While large amounts of runoff during heavy rain events cause existing wastewater systems to exceed their limits, the systems are underused in dry weather. This leads to unequal system requirements.

By using remote control technology for hydraulic monitoring, control and regulation of rainwater tanks and the use of dynamic controls to exploit existing retention volumes in sewer systems, an increase in operational reliability can be achieved.

Fig. 1. Scheme of the wet pit pumping station network Grassau, displaying all five wet pit pumping stations and their relative positions to each other

The purpose of this research project was to prevent sewer overflow events of a wet pit pumping station (see Fig. 1, Grassau) located in the middle of a small village in a rural area. Therefore a decentralized, remote controlled network based upon the current operating condition of each wet pit of the system was installed.

2. Methods

In order to optimize the network and increase the operational reliability, an exact knowledge of the system is required. Also for the simulation in EPANET the hydraulic parameters of the system are needed. Besides, an inventory is also needed to identify unused retention volumes of the network.

2.1. Determination of characteristic curves

The characteristic curves of the pumps installed are later used to calculate the baseflow of wastewater in the system. They are also required to determine the operating point and to assess the performance or wear of the pumps since they are partially older than ten years.

To determine the characteristic curve, a mobile test stand was designed (see Fig. 2) and all ten pumps of the system were measured.

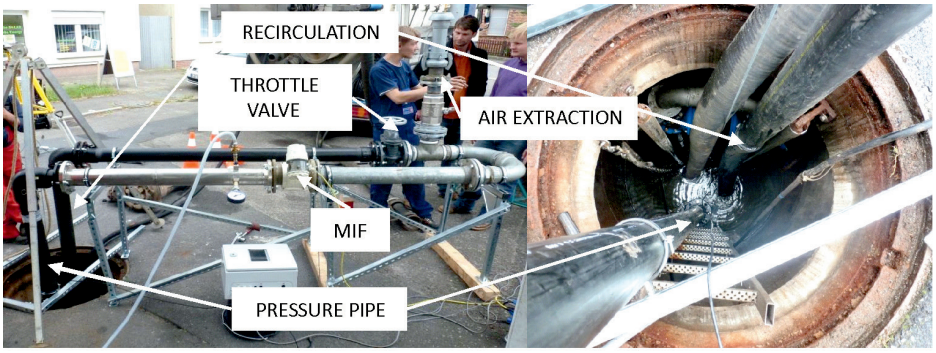


Fig. 2. Mobile test rig for determination of characteristic curve, in the case shown at wet pit pumping station Jeßnigk

The test rig allows to measurements of the pressure, the flow and the electric power consumption of the pump tested on the basis of ISO EN DIN 9906 [1]. Due to air pockets in the piping, a connection for a suction vehicle is needed to extract the air from the measuring section. The pressure sensor is installed directly at the water level of the pit at the discharge line of the pump.

2.2. Simulation of the network with EPANET

To simulate different scenarios of inflow and stormwater infiltration and the response of the system, the network was modelled in EPANET (see Fig. 3).

The network overview clearly shows the composition of the network. The data for the sizes of the wet pits, piping diameters and coordinates for inflow and outflow of every pit were provided by the wastewater network operator. After measuring the characteristic curves of the pumps,

they are implemented in EPANET using a multi-point-curve. Since EPANET cannot calculate open surface flows, every pipe is considered filled, which reduces the system capacity in relation to the real system. Before every pumping station, a knot is inserted which is defined as a source. In this way the dryweather inflow and the stormwater inflow into the wet pit pumping station is simulated. For the dryweather wastewater emergence, wastewater hydrographs are used (see Fig. 4). The stormwater inflow is simulated by a control, based upon probability of heavy rain events, rain intensity, runoff coefficient and the size of the catchment area [2].

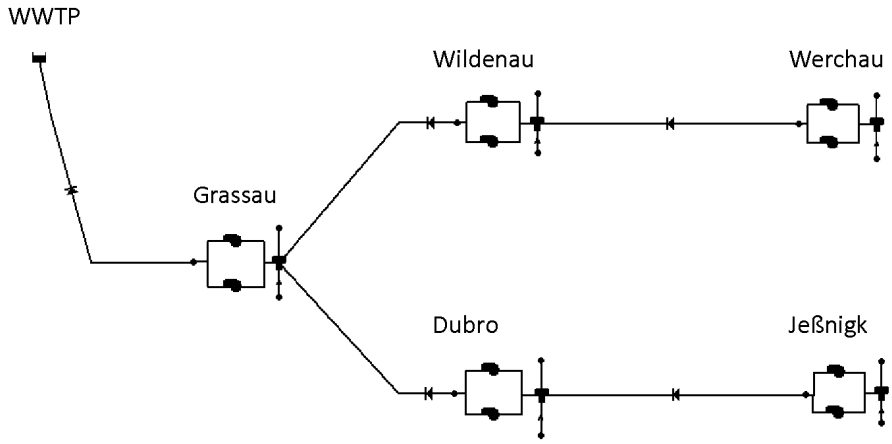


Fig. 3. Network model of the wastewater network

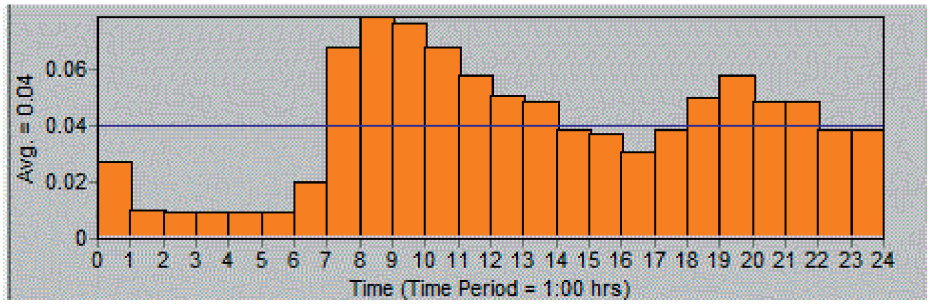


Fig. 4. Hydrograph for wastewater emergence at dryweather

2.3. Design of the decentralized network

For continuous online monitoring, all wet pits are equipped with pressure sensors, power consumption meters and a programmable logic controller (PLC). All signals from the pressure sensors, level meters, power consumption meters, on/off times of the pumps are saved onto a secure digital memory card (SD card) and sent via Email by a GPRS-

module. Since every PLC is equipped with a GPRS-module, data or information can be sent from every wet pit pumping station to the others. The operational reliability is increased by installing a redundancy in the control of the wet pits. If one PLC fails, the wet pit goes back to normal operation triggered by the installed level sensor.

2.4. Development of switching scenarios in dependence of load and flow

Normally, operation of a wet pit pumping station is controlled by a level sensor in the wet pit pumping station. Figure 5 shows the construction of a wet pit pumping station. In normal operation, the level rises up to a certain point. When the start-up level is reached the pump starts operating until the wastewater level falls below the end of operation level. The levels for switching the pumps on/off can be implemented in EPANET.

Normally, the start-up level is below the inflow pipe, but to increase the volume of the system, the limits of the start-up level are shifted up to 10 cm below the highest point of the pressure pipe. This enhances the storage capacity of the whole system (see Table 1). By this means, the overall capacity is nearly eight times bigger.

Table 1. Capacity of the wet pit pumping stations new vs. old control

Wet pit	Old control	New control	
	Volume in [m³]	Volume in [m³]	Factor
Grassau	0.94	12.88	13.7
Wildenau	1.26	12.88	10.2
Dubro	1.26	6.53	5.1
Werchau	1.26	9.42	7.4
Jeßnigk	1.88	11.0	5.8
Sum	6.6	52.71	7.98

The main goal of every control of the network is to prevent an overflow of any wet pit pumping station of the network. In the analysed network the wet pit in Grassau is the critical point, since all pumping stations feed this wet pit. In addition, it is also the smallest one.

During normal dry weather base flow, which was determined during long term monitoring of the system, the old levels for putting the pumps in and off operation apply. When there is a higher inflow into one or more wet pits, the levels of pump operation are extended to the ones mentioned above. To prevent an overflow, the wet pits communicate among themselves and exchange information about their levels. The basis of the control is that the upstream wet pit stops pumping into the wet pit downstream which has a strong rise in the level. Also the difference between the levels in the two lines of wet pits feeding Grassau is considered. The line with the least amount of wastewater according to the levels in the wet pits stops pumping to Grassau.

Every pumping station which is not pumping retains the wastewater to the extended level of 10 cm below the highest point of the pressure pipe. When this level is reached, the pumps start operating to prevent an overflow.

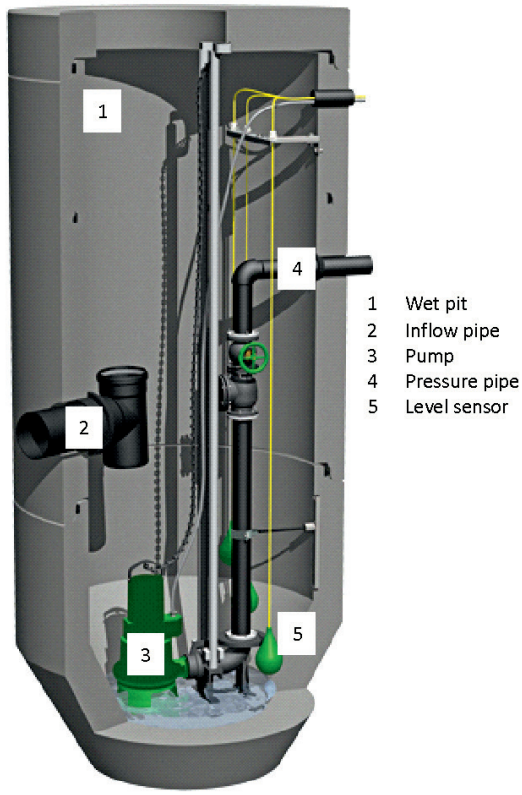


Fig. 5. Construction of a wet pit pumping station with submerged wastewater pumps

3. Results

In the following section the results for the developed switching scenarios, the dry weather base flow and the benefit of the control in dependence of load and flow are shown.

The first step was to determine the characteristic curves of all pumps in order to be able to draw a conclusion about the flow by looking at the pressure delivered by the pump. For redundancy, the monitoring of the electric power consumption, respectively the current draw, provides the same information, see Figure 6.

With this information it is possible to calculate the dry weather base flow of all wet pit pumping stations of the system by multiplying the operational time of every pump with their respective flow due to the determined characteristic curve and the continuous monitoring of the pressure signal. Table 2 shows the dry weather base flow of every wet pit of the system.

Table 2. Dry weather base flow

Wet pit	Inhabitants	Volume of wastewater in [m³/day]
Jeßnigk	286	34.56
Dubro	280	16.1
Werchau	144	13.95
Wildenau	182	7.84
Grassau	255	61.48

With these data, the network is simulated in EPANET and the inflow is distributed according to the wastewater hydrograph. Figure 7 shows the pumping intervals of the wet pits for the dry weather base flow during 24 hours.

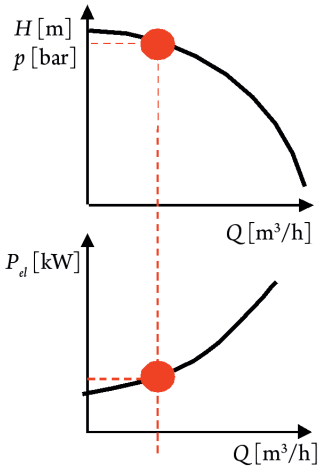


Fig. 6. Flow in dependency of pressure and electric power consumption

Since Grassau is the end of the line wet pit, the time of pump operation is the highest of all wet pits of the system. EPANET considers the pipes to be fully filled all the time, hence the level in the downstream wet pit increases strongly when the upstream wet pit goes into operation. In reality the pipes are not fully filled, so the peaks with the big gradient do not occur in reality.

Also the different levels of operation of every pumping station can be seen in Figure 7. For example the end of operation level for Werchau is a level of wastewater of 50 cm above the bottom of the wet pit whereas the start-up level is around 80 cm above the bottom.

To investigate the behaviour of the system for an additional inflow of stormwater, several different inflow scenarios have been investigated. Figure 8 shows the simulation of a heavy rain event with an estimated additional inflow of 5 m³/h into every wet pit, starting at 17h o'clock and ending at 19h o'clock.

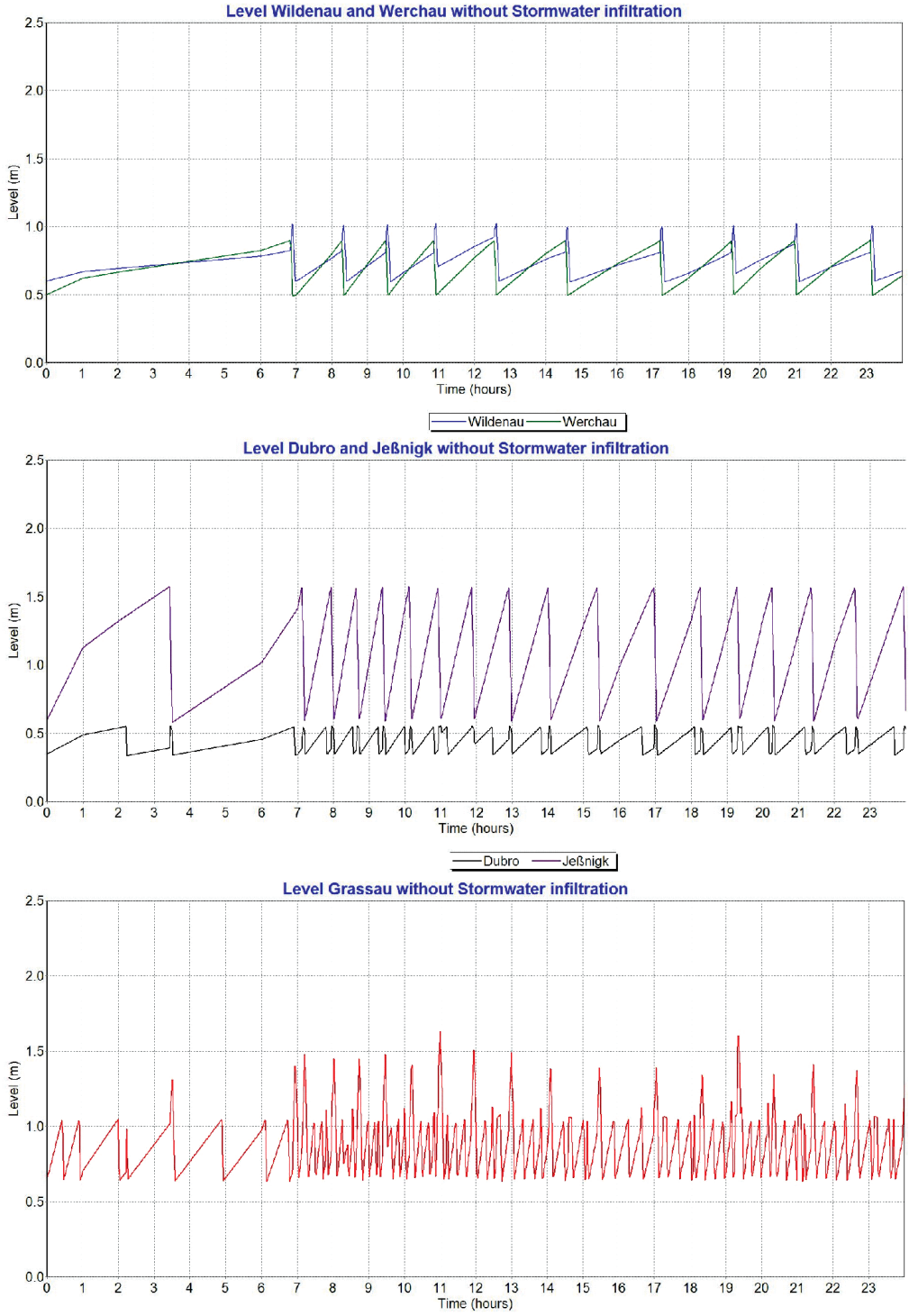


Fig. 7. Pumping intervals of the wet pits throughout one day at dry weather base flow

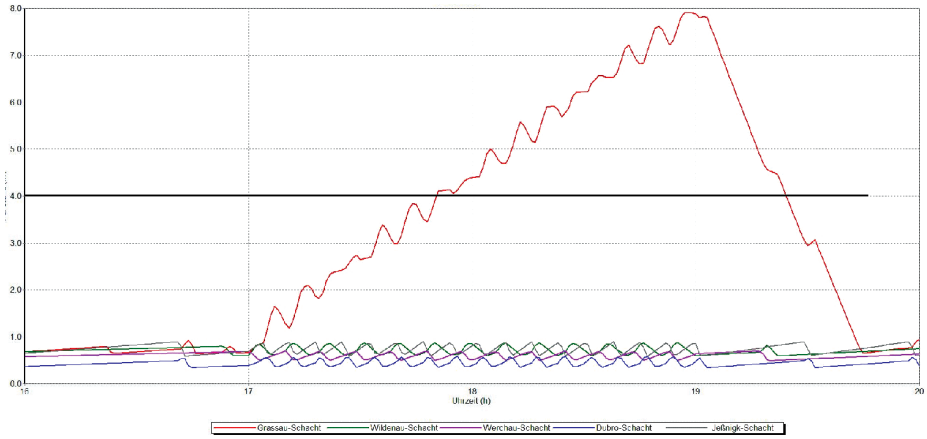


Fig. 8. Pumping intervals of the wet pits throughout one day with stormwater infiltration of 5 m^3/h for two hours into every wet pit

The ground level at Grassau is 4.0 m above the bottom of the wet pit. In the simulated case, the wet pit at Grassau would overflow due to the ongoing pumping of the other wet pits. In contrast to Grassau, all the other wet pits would stay in their normal level limits.

Applying the new control and the extended levels of pump operation, the system uses the existing retention volumes and an overflow at Grassau is prevented (see Fig. 9). The wet pit pumping station at Wildenau would show the highest level throughout the time of stormwater infiltration. As mentioned above, the levels are rising in every wet pit and consecutively, the wet pits are restraining their wastewater to support the downstream pit. The first pit which is retaining the wastewater is Werchau at around 17.20 h. The wastewater

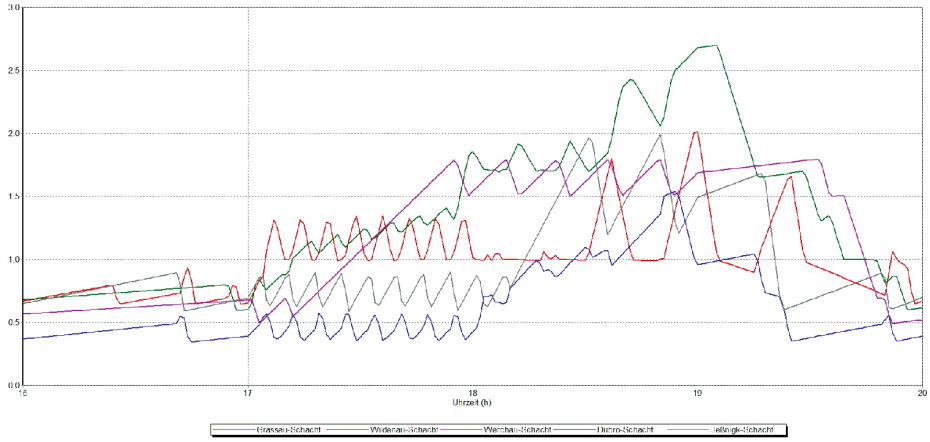


Fig. 9. Pumping intervals of the wet pits throughout one day with stormwater infiltration of 5 m^3/h for two hours into every wet pit using the developed control

is accumulated up to 10 cm below the pressurised pipe. When this level is reached, the pumps go back into operation. Theoretically this would go on until all wet pits have reached their critical level and then all the pumps would go into operation. Since Wildenau is the wet pit with the most retention volume, it is the last one to go back into operation.

For the simulated event of an infiltration of 5 m^3/h and using the developed control, the level of no wet pit pumping station in the system would be higher than the height of the pressurised pipe. The overflow of Grassau would be prevented.

References

[1] DIN EN ISO 9906: *Rotodynamic pumps – Hydraulic performance acceptance tests*, International Organisation for Standardisation, Geneva, Switzerland 2012

[2] Gujer W., *Siedlungswasserwirtschaft*, Springer-Verlag Berlin Heidelberg, 2007.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
Mechanics
DOI: 10.4467/2353737XCT.17.024.6217

Zygmunt Domagała
Krzysztof Kędzia (krzysztof.kedzia@pwr.edu.pl)
Department of Operation and Maintenance of Logistics, Transportation and Hydraulic Systems, Mechanical Faculty, Wroclaw University of Science and Technology

Analysis, modelling and verification of the phenomena occurring in a hydraulic prop during dynamic load

Analiza, modelowanie i weryfikacja zjawisk zachodzących w dwuteleskopowym stojaku hydraulicznym w warunkach obciążeń dynamicznych

Abstract

The paper describes the working conditions of hydraulic prop in particular issues related to their dynamics. On this basis, mathematical model was developed for given simplifying assumptions. The paper includes description of simulation model as well as an experimental verification performed in HSW Stalowa Wola.

Keywords: hydraulic prop, simulation model, experimental verification

Streszczenie

W artykule opisano warunki pracy podpory hydraulicznej, w szczególności zagadnienia związane z ich dynamiką. Na tej podstawie opracowano model matematyczny dla podanych założeń upraszczających. Praca zawiera opis modelu symulacyjnego, jak również weryfikację doświadczalną wykonaną w HSW Stalowa Wola.

Słowa kluczowe: podpora hydrauliczna, badania symulacyjne, weryfikacja eksperymentalna

1. Introduction

Because of the lack of substantial reserves of such energy carriers as crude oil or natural gas within the borders of Poland, hard coal is the primary energy source in this country. Hard coal is mined underground using longwall systems, this contributes to movements of rock mass and therefore poses a direct danger to human life. Consequently, designers of mining machinery must constantly seek solutions that ensure the required levels of productivity whilst guaranteeing the maximum safety of miners. Longwall coal bed extraction systems are commonly used to protect tunnels and work spaces that have been newly created by mining. Thanks to the development of such systems, coal can be mined from ever deeper coal beds where the rock mass is more active [3, 5].

Currently, the danger of rockburst is present in about half of all active longwalls, generating a high demand for additional protective measures capable of taking the dynamic loads produced by rockbursts. Since such loads are often heavier than the static loads, they can damage or destroy the roof support.

It is difficult to prevent rock mass tremors because the manner in which they occur is not sufficiently understood and because of the short duration and their unpredictable nature. Since the opportunities to measure rockbursts in the mine heading are limited, the rockbursting phenomenon is evaluated on the basis of its effects. Investigations in mine conditions pose many difficulties and involve high costs connected with the use of specialist equipment; therefore, instead of carrying out investigations in mine conditions, roof support components are tested on a special test stand. Such testing is highly complex in order to reproduce the actual load present in the rock mass. Experimental studies are supplemented with low-cost (not requiring specialist equipment) analytical studies. However, the results of the latter may contain errors stemming from the adopted model and from the simplified assumptions [2].

Under the action of the rock mass force, the commonly used double telescopic roof supports undergo damage mainly due to the cracking of the second-stage cylinders. This observation provided the basis for analytical and experimental studies of a double-telescopic hydraulic prop equipped with release valves in its two stages. A schematic of such a solution is shown in Fig. 1.

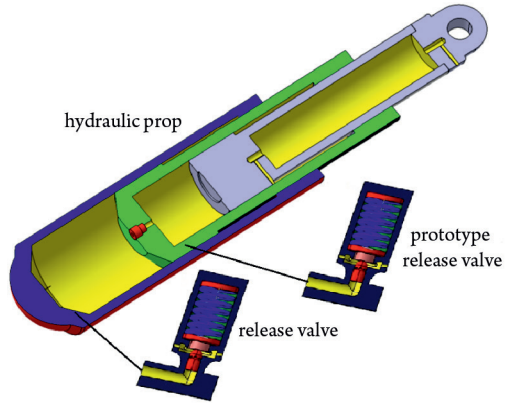


Fig. 1. Simplified schematic of hydraulic prop – concept (source: own work)

A hydraulic press (shown in Fig. 2) was used in the experimental studies, during which, the tested support was loaded for a specified period of time.

Fig. 2. Tests of actuators in a hydraulic press [2]

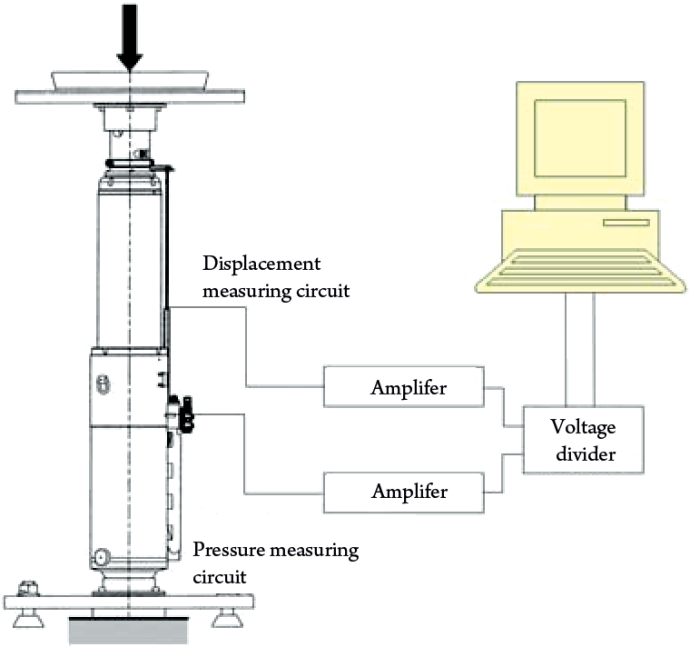


Fig. 3. Schematic of measuring system used for verification measurements of telescopic actuators [2]

Thus far, the phenomena arising in the rock mass have been difficult or virtually impossible to predict, thus, they cannot be effectively prevented. Additionally, the short duration of the tremor is unhelpful in determining and recording the place where the latter occurred. The tests of the telescopic actuators in the hydraulic press merely showed that the actuators performed well under the given load and load duration.

2. Mathematical model

A mathematical model was built assuming the discrete distribution of mass and elasticity and taking into account the constraints stemming from the MATLAB-Simulink software used for the computations. The hydraulic system of a powered roof support consists of many components such as actuators, flexible pipes, a pump, a tank and a valve block. The number of system components was considerably reduced in order to simplify the mathematical model and the simulation model. A schematic of the modelled hydraulic system is shown in Fig. 1. The symbols describing the hydraulic prop in Fig. 4 were used to formulate mathematical equations.

2.1. Equations of forces acting on hydraulic prop

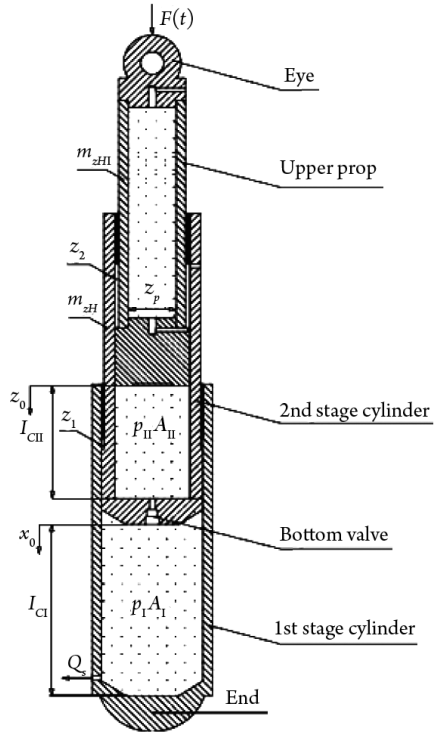


Fig. 4. Hydraulic prop (source: own work)

The equations of the forces acting on the particular stages of the hydraulic prop are as follows:

$$F(t) - F_{sb\,\mathrm{II}} - F_{st\,\mathrm{II}} - F_{sh\,\mathrm{II}} = 0 \tag{1}$$

$$P_{\mathrm{II}}(A_{\mathrm{II}} - A_d) - F_{sb\,\mathrm{I}} - F_{st\,\mathrm{I}} - F_{sh\,\mathrm{I}} = 0 \tag{2}$$

where:
F_{sb} – the prop's force of inertia, including mass m,
F_{st} – the viscous friction force,
F_{sh} – the force exerted by pressure onto the piston surface,
$F(t)$ – the rock mass impact force.

Equations (1) and (2) are considered for the following initial conditions:

$$x = x_{gr},\ \frac{dx}{dt} = 0,\ \frac{d^2x}{dt^2} = 0,\ z = 0,\ \frac{dz}{dt} = 0,\ \frac{d^2z}{dt^2} = 0$$

and for the following boundary conditions:

$$x = x_{gr},\ \frac{dx}{dt} = 0,\ \frac{d^2x}{dt^2} = 0,\ z = gr,\ \frac{dz}{dt} = 0,\ \frac{d^2z}{dt^2} = 0$$

2.2. Equation of forces acting on head of release valve I

Figure 5 schematically shows the prop's 1st stage release valve. The symbols describing the valve in the figure below were used to formulate the mathematical equations.

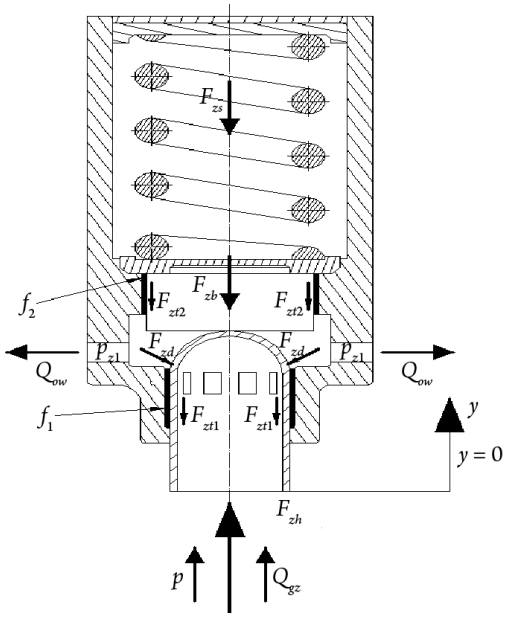


Fig. 5. Schematic of valve (source: own work)

The equation of the forces acting on the head of release valve I is as follows:

$$F_{zh} - F_{zb} - F_{zt} - F_{zs} - F_{zd} = 0 \tag{3}$$

where:
F_{zh} – the force exerted by pressure onto the valve head;
F_{zb} – the force of inertia;
F_{zt} – the force of viscous friction;
F_{zs} – the force produced by spring deflection;
F_{zd} – the hydrodynamic force.

The above equations are considered for the following initial conditions:

$$y = 0,\ \frac{dy}{dt} = 0,\ \frac{d^2y}{dt^2} = 0$$

and for the following boundary conditions:

$$y = y_{gr},\ \frac{dy}{dt} = 0,\ \frac{d^2y}{dt^2} = 0$$

2.3. Equation of rate of flow through the valve under-head chamber

$$Q_{gz} - Q_{cp} - Q_{gp} - Q_w = 0 \tag{4}$$

where:
Q_{gz} – the rate of the flow through the valve head,
Q_{cp} – the flow rate resulting from the compressibility of the working liquid in the under-head chamber,
Q_{gp} – the flow rate resulting from the displacement of the valve head,
Q_w – the rate of the flow through the valve under-head chamber.

The equation of forces and the equation of flow rates for the 2nd stage release valve are the same as for the 1st stage release valve.

2.4. Equation of rates of flow in the modelled system

Figure 6 shows a schematic of the modelled hydraulic system, including the adopted symbols.

For the above schematic, the following equations of flow were derived:

$$Q_{s1} - Q_{gz1} - Q_{c1} - Q_{g1} = 0 \tag{6}$$

$$Q_{s2} - Q_{gz2} - Q_{c2} - Q_{g2} = 0 \tag{7}$$

where:
Q_s – the liquid flow rate resulting from the movement of the upper prop,
Q_c – the rate of flow of the liquid resulting from its compressibility under the prop's piston,

Q_g – the liquid flow rate resulting from the displacement of the valve head,
Q_{gz} – the rate of the flow through the valve head.

The equation of forces and the equation of flow rates for the 2nd stage release valve are the same as for the 1st stage release valve.

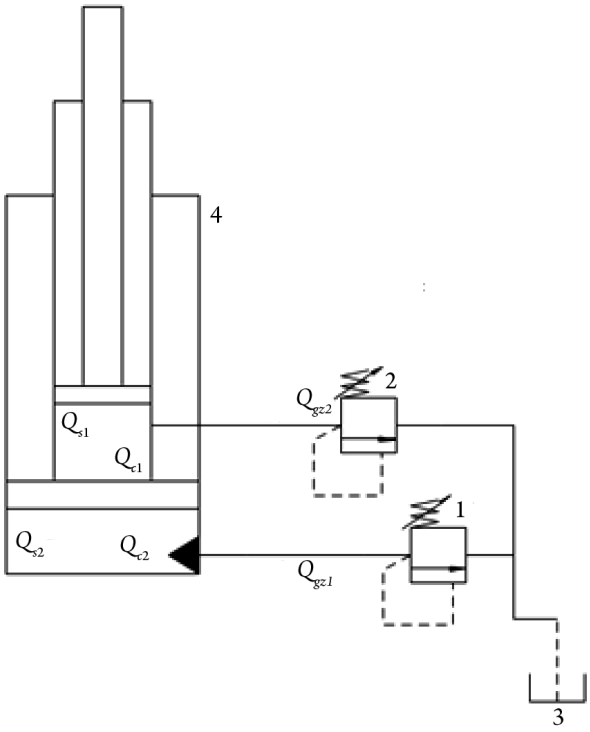


Fig. 6. Schematic of modelled hydraulic system, including adopted symbols (source: own work)

2.5. Mathematical model of hydraulic press

Equations of forces and flow rates characterising the hydraulic press are presented below. Figure 7 shows a schematic of the modelled press.

The equation of the forces acting on the hydraulic press is as follows:

$$F_{ph} - F_{pb} - F_{pt} - F_{sh2} = 0 \tag{8}$$

where:
F_{ph} – the force exerted by pressure onto the press piston,
F_{pb} – the force of inertia of the hydraulic press,
F_{pt} – the force of viscous friction,
F_{sh2} – the force of support in the 2nd stage of the actuator.

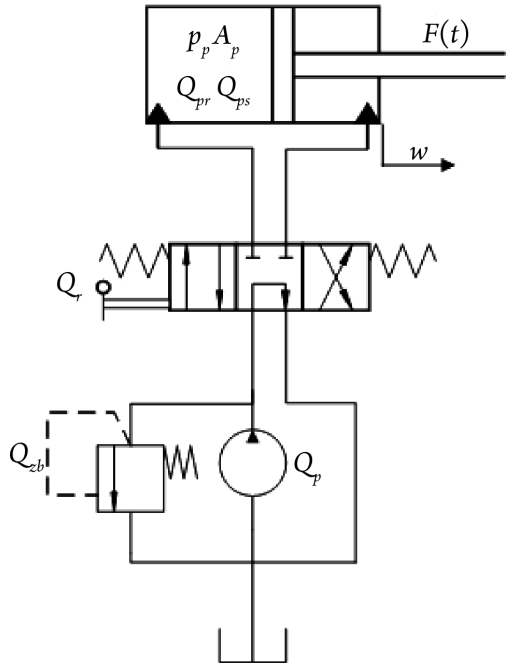


Fig. 7. Schematic of modelled hydraulic press (source: own work)

2.6. Equations of flow rates in the hydraulic press

The equations describing the balance of flow rates in the hydraulic press are as follows:

$$Q_p - Q_{pp} - Q_r - Q_{zb} = 0 \tag{9}$$

$$Q_r - Q_{ps} - Q_{pr} = 0 \tag{10}$$

where:

Q_p – the rate of delivery of the pump,
Q_{pp} – the rate of flow of the liquid resulting from its compressibility in the conduit between the pump and the distributor,
Q_r – the rate of flow of the liquid through the distributor,
Q_{zb} – the flow rate resulting from the flow of the liquid through the safety valve,
Q_{ps} – the rate of flow of the liquid resulting from its compressibility under the press piston,
Q_{pr} – the flow rate resulting from the displacement of the hydraulic press piston.

The MathWorks Matlab software environment was used to conduct simulation studies. The software enables one to perform numerical computations, test algorithms, modelling, run simulations and analysis, and visualisations of data.

3. Verification of mathematical model

Verification was performed by comparing the experimental graphs of pressure in the 2nd stage cylinder and of 2nd stage upper prop displacement with the respective graphs obtained from simulations.

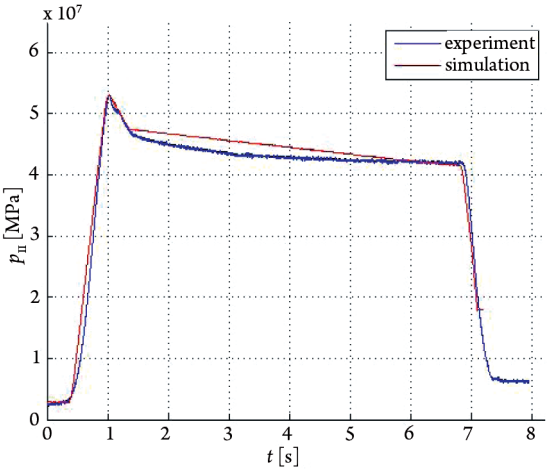


Fig. 8. Graphs of pressure in 2nd stage of hydraulic prop – test (41), modified model (source: own work)

Figure 8 shows exemplary experimental and simulated pressure curves for the hydraulic press. By comparing the curves, one can evaluate the validity of both the mathematical model and the simulation model of the investigated hydraulic system. The figure indicates that there are no grounds for rejecting the mathematical model. However, the experiment did not reproduce the actual conditions of the hydraulic work of the pit prop; therefore, in the simulation model, the loading of the prop by the hydraulic press was replaced by a force loading a single prop in accordance with the relation given by Stoiński [6].

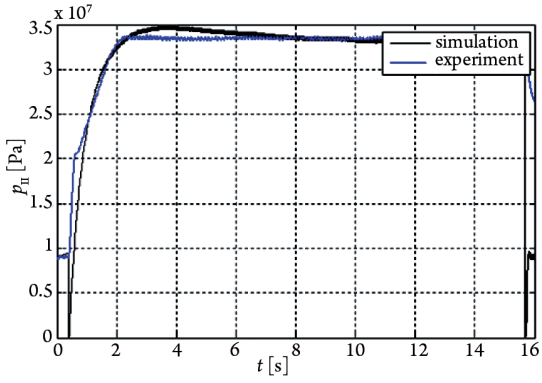


Fig. 9. Graphs of pressure in chamber under upper prop – test (25), modified model (source: own work)

4. Mathematical model of rock mass load

If the effect of the dynamic impact of the rock mass on the roof support is assumed to be a consequence of the instantaneous increase in load brought about by a rockburst, and if it is assumed that the particular sections of the roof support are uniformly loaded, the force loading a single prop can be expressed by the relation [6]:

$$f(t) = F_w + F_d[1 + k_d \cdot e^{-\delta t} \sin(\omega t - \varphi)] \tag{11}$$

where:

F_w – the initial supporting power of the prop, [N],
F_d – the force additionally loading the prop, [N],

$k_d = 0.102\omega v_0$ – a factor of safety $\left[\frac{\mathrm{m}}{\mathrm{s}^2}\right]$,

$\omega = 3.13\sqrt{\frac{k_s}{\eta_{tz} F_r}}$ – the angular frequency of the vibrating system, [s^{-1}],

δ – the damping coefficient of the vibrating system, [s^{-1}].

5. Simulation studies

The aim of the simulation studies was to minimise the pressures generated by an excitation in the individual sections of the hydraulic prop. The excitation represents an actual rockburst. The graph of the load generated by an exemplary rockburst is shown in the figure below.

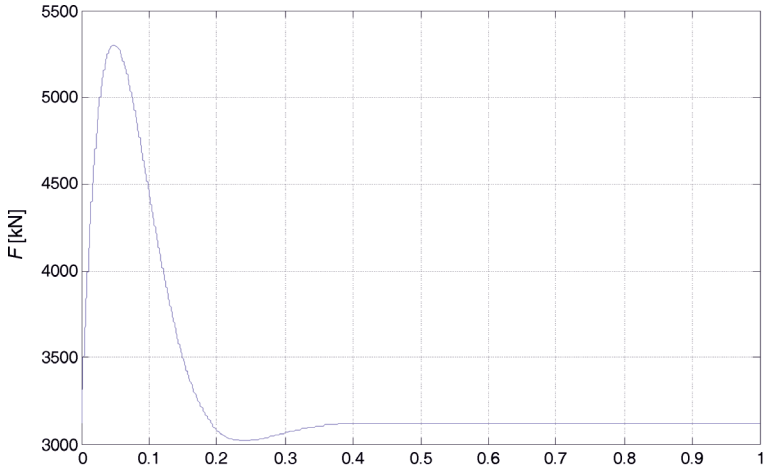


Fig. 10. Graph of load generated by rockburst (source: own work)

The distinctive feature of the considered rockburst is the fact that the maximum load values are reached in 0.05 s and they are nearly twice as high as the stable force value reached

in 0.4 s. This means that tests of hydraulic props performed through the use of a hydraulic press do not reproduce the conditions prevailing during the rapid collapse of the rock mass in a rockburst. The very heavy instantaneous loading of the pit prop results in an increase in pressure increment dp/dt whereby the pressure in the cylinders of the hydraulic actuators rapidly rises. In such conditions, the pressure exceeds the allowable pressure (Table 1) and the hydraulic prop cracks.

The considered hydraulic prop includes release (safety) valves which should open and release some of the emulsion in order to reduce the pressure in the actuators; however, the rate of pressure rise is so fast and the time in which the pressure reaches its maximum is so short that the release valve cannot open, as a result, the prop fails.

The question that thus arises is whether or not the release valve would open under the same load if the rate of pressure rise in the hydraulic system of the pit prop could be reduced.

An analysis of the problem showed that this is possible and can be achieved, for example, by:

- changing the initial tension of the spring of the release valves;
- changing the stiffness of the spring;
- increasing the volume of the hydraulic system.

Studies were carried out on this problem using a simulation model which had been successfully validated on a real object in HSW (Stalowa Wola Steel Mill).

First, the effect of the initial tension of the spring in the 2nd stage cylinder valve was studied. The results of this study are presented in Fig. 11.

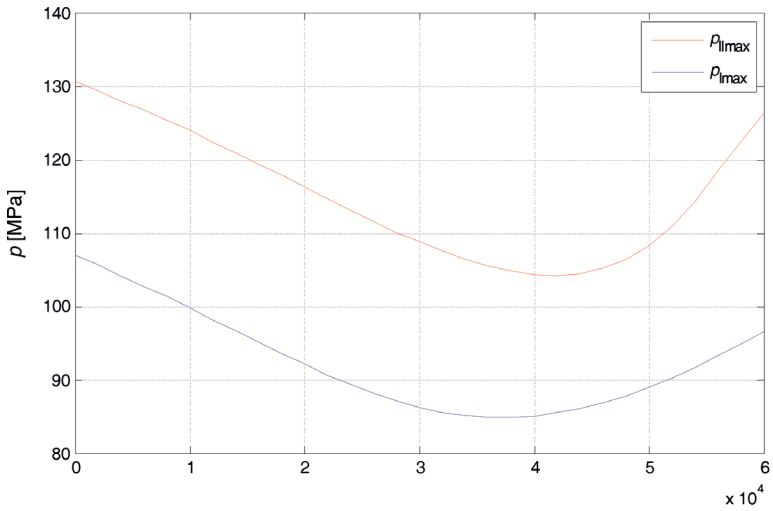


Fig. 11. Effect of initial spring tension c_{02} on maximum pressures in particular prop stages (source: own work)

Figure 11 shows that there exist such initial spring tensions at which pressures in the individual sections of the pit prop reach minimal values. The minimal pressures are reached at $c_{02} = 42\,000$ N for the actuator's second stage and at $c_{02} = 37\,000$ N for its first stage; however, the pressures reached are higher than those allowable for this actuator structure.

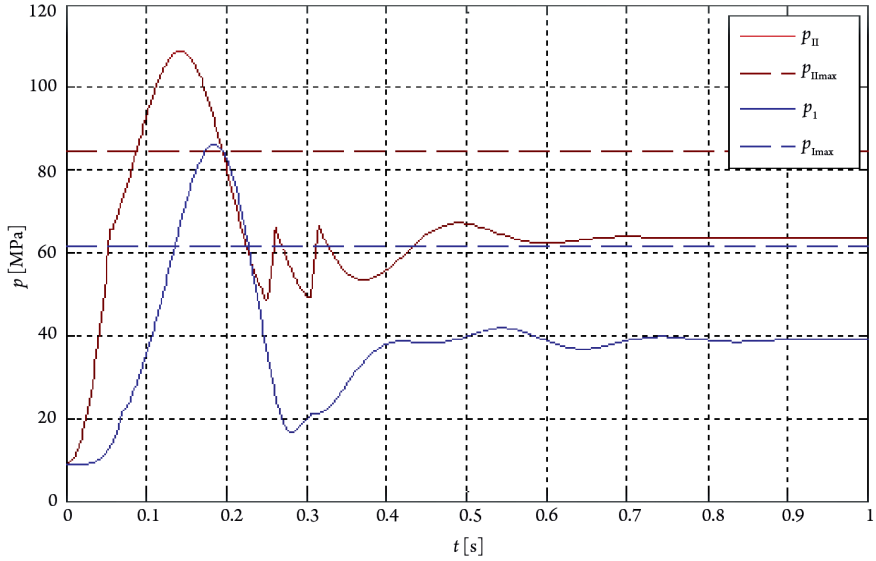


Fig. 12. Pressure curves for c_{02} = 42 000 N (source: own work) (P_{allow})

Figure 12 shows pressure curves for spring tension c_{02} = 42 000 N. It is clearly apparent that the pressures in both prop stages exceed the allowable values.

Another factor which can affect the performance of the system is the stiffness of the release valves' springs.

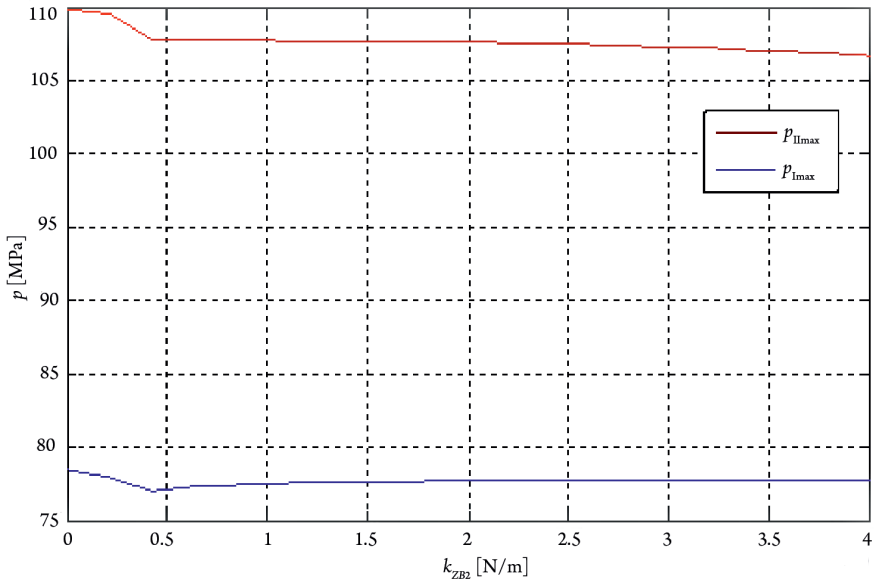


Fig. 13. Effect of initial spring stiffness k_{ZS2} on maximum pressures in individual prop sections (source: own work)

Figure 13 shows that the stiffness of the 2nd stage valve spring has no so strong effect on the maximum pressures as it had at the initial tension of this spring. The lowest pressures in the 2nd stage cylinder occurred at maximum adopted stiffness $k_{zs2} = 4 \cdot 10^6$ N/m.

The simulation model was then used to determine the effect of the initial spring tension for the first stage valve. The results are presented in Fig. 14.

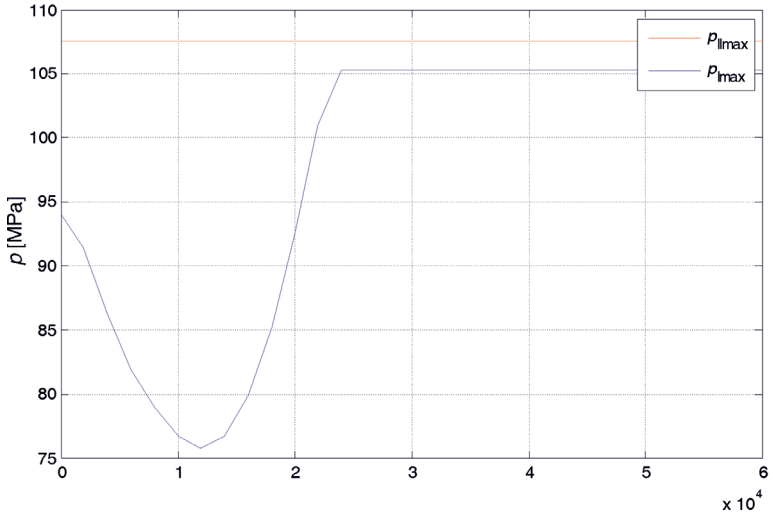


Fig. 14. Effect of initial spring tension c_{01} on maximum pressures in particular prop stages (source: own work)

It is immediately apparent that the initial tension of the spring in the 1st stage valve has no effect whatsoever on the maximum pressures in the second stage of the cylinder.

In the case of the cylinder's first stage, the best parameters were obtained for c_{01} = 12 000 N; however, the pressures are still higher than those allowable for the 1st stage cylinder.

Finally, the effect of the hydraulic system volume on pressure curves and on the maximum pressures was examined. A change in system volume undoubtedly has an effect on the rate of pressure rise (dp/dt); however, the drawback of this solution is that it entails an increase in the overall dimensions of the system. The most effective solution is to use hydraulic accumulators. Figure 15 shows the effect of the capacity of a hydro-pneumatic accumulator on the maximum pressures in the individual sections of the prop.

Figure 15 indicates that different ranges of accumulator volume have different effects on the performance of the system. The horizontal broken lines mark the allowable pressure values. The red lines and the blue lines represent the maximum pressures as a function of hydraulic accumulator volume for the 2nd stage and the 1st stage of the telescopic prop, respectively.

The figure shows that for a hydraulic accumulator volume of 100 litres. the pressures in the hydraulic prop do not exceed the allowable values. This problem is better illustrated in Fig. 16 where the pressure curves for the hydraulic support with and without a hydraulic accumulator are compared.

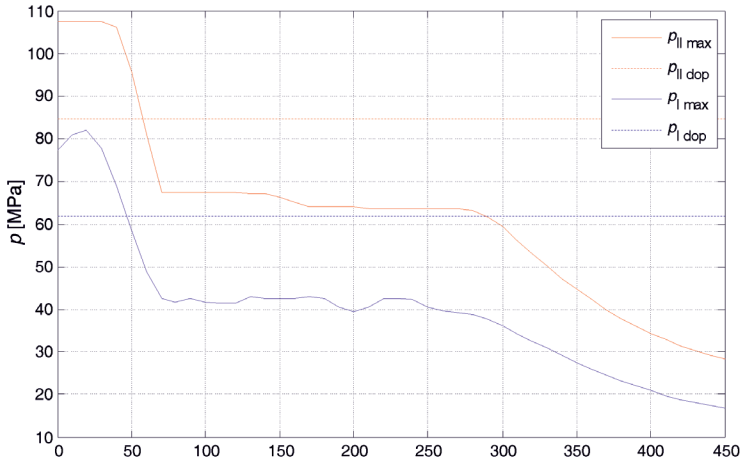


Fig. 15. Effect of hydro-pneumatic accumulator capacity on maximum pressures in particular prop stages [4]

The graphs indicate that the accumulator with the assumed volume was able to offset the whole rapid pressure rise and therefore reduce the rate of pressure increment dp/dt.

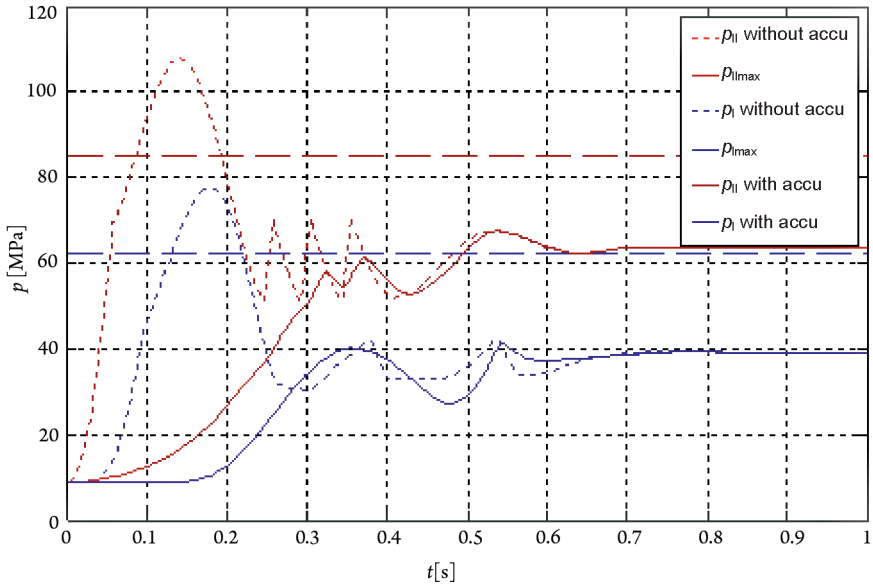


Fig. 16. Pressure curves for V_{0a} = 100 *l* (source: own work)

6. Conclusions

The simulation studies show that a change in the valve spring parameters had no significant effect on the maximum pressure values, whereas the use of a hydraulic accumulator could significantly offset the maximum pressures in the individual sections of the hydraulic prop and consequently protect the cylinders of the sections from destruction.

The use of hydraulic accumulators in roof support props can solve the problem of the hydraulic blockage of the release valves. This phenomenon is a serious problem due to the fact that as the supporting power of hydraulic props increases, so do their diameters and consequently, the volume of working liquid under the piston. Currently, the valves are not able to release the proper amount of working liquid in a sufficiently short time to ensure the required collapse of the prop. A hydraulic accumulator is suitable for this purpose since it takes over the working liquid excess and prevents the prop from being destroyed.

References

[1] Domagała Z., *Modelowanie i symulacja zjawisk zachodzących w zmechanizowanej obudowie ścianowej*, Maszyny Górnicze 4/2009.

[2] Domagała Z, Marianowski J., *Modelowanie i weryfikacja zjawisk dynamicznych zachodzących w teleskopowych stojakach hydraulicznych, Badanie, konstrukcja, wytwarzanie i eksploatacja układów hydraulicznych*, Gliwice 2013.

[3] Jaszczuk M., *Ścianowe systemy mechanizacyjne*, Śląsk, Wydawnictwo Naukowe, Katowice 2007.

[4] Kamiński G., *Analiza zjawisk dynamicznych zachodzących w dwuteleskopowym stojaku hydraulicznym*, Wrocław 2015.

[5] Praca zbiorowa, *Nowoczesne, niezawodne i bezpieczne systemy mechanizacyjne dla górnictwa*, Centrum Mechanizacji Górnictwa KOMAG, Gliwice 2008.

[6] Praca zbiorowa do redakcją A. Kidybańskiego, *Stateczność górotworu i obudowy przy łącznym obciążeniu statycznym i dynamicznym*, Główny Instytut Górnictwa, Katowice 2009.

[7] Stoiński K., *Obudowy górnicze w warunkach zagrożenia wstrząsami górotworu*, Główny Instytut Górnictwa, Katowice 2000.

[8] Stryczek S., *Układy hydrauliczne Tom I. Elementy i Tom II Układy*, Wydawnictwo Naukowo-Techniczne, Warszawa 2005.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
MECHANICS
DOI: 10.4467/2353737XCT.17.026.6219

Karolina Głogowska (k.glogowska@pollub.pl)
Janusz W. Sikora
Department of Polymer Processing, Faculty of Mechanical Engineering, Lublin University of Technology

The production and properties of micro-extruded low-density polyethylene microtubes

Otrzymywanie i właściwości mikrowytłoczyn z polietylenu małej gęstości

Abstract

Micro-extrusion is the process of obtaining microtubes with either simple or complex cross-sections which are used for the high-precision transport and protection of various of media. Through this process, it is possible to create tubing with interior diameters and wall thicknesses as small as 0.1 mm and 0.05 mm respectively. Micro-extruded tubing such as micro catheters and peripheral intravenous cannulas are employed in microfluidics, paediatrics and micro-analytics. This article presents the results of granulometric tests of micropellets produced by micro-extrusion with cold pelletizing and results of tensile strength tests of microrods of different diameters. The trials were performed using a standard twin-screw extruder and low-density polyethylene.

Keywords: polymer micro-extrusion, low-density polyethylene, micropellets

Streszczenie

Mikrowytłaczanie to proces wytwarzania mikrorurek prostych lub złożonych w przekroju poprzecznym, używanych do wysoce precyzyjnego przesyłu oraz ochrony różnych mediów. Dzięki temu procesowi możliwe jest uzyskanie rurek o średnicy wewnętrznej i grubości ścianek tak małych jak 0,1 mm i 0,5 mm. Rurki wytłoczone w ten sposób, np. mikrocewniki i kaniule dożylne, stosowane są w technikach mikroprzepływowych, pediatrii i mikroanalizie. Niniejszy artykuł przedstawia wyniki testów granulometrycznych mikrogranulek wytworzonych poprzez mikrowytłaczanie z granulowaniem na zimno, jak również testów wytrzymałości na rozciąganie mikrożyłek o różnych średnicach. Próby zostały przeprowadzone przy użyciu standardowej wytłaczarki dwuślimakowej i polietylenu małej gęstości.

Słowa klucze: mikrowytłaczanie polimerów, polietylen małej gęstości, mikrogranulki

1. Introduction

The constant progress of science leads to changes in the developmental of technology. One manifestation of these changes is the introduction of new and improved machinery, equipment, tools and technologies into the production process that can be used to manufacture higher quality products with new, hitherto unknown, functionalities. The polymer processing technology, in which one of the trends is to miniaturise both equipment and products, has not been left unaffected by this development. Examples of miniaturisation include processing procedures, such as micro-injection moulding and micro-extrusion performed using micro-injection moulding machines and micro-extruders [1–4].

The process of micro-extrusion can be used to form microrods or microtubes with either simple or complex cross-sectional shapes, and outer diameters and wall thickness as small as 0.2 mm and 0.05 mm, respectively [5, 6]. These products are mainly applied in medicine, in fields such as vascular surgery, and optics. Poly(vinyl chloride), poly(vinyl acetate) and silicone are used to produce various types of catheters with diameters as small as 2 mm [7, 8]; PVC and polyurethane (PUR) are applied in the production of tubing with even smaller diameters – 0.5 mm [9].

In coronary artery bypass graft surgeries, surgeons use artificial blood vessels called by-passes with diameters of as small as 3 mm, extruded from silicones or PUR [10]. Also common are different kinds of small-diameter (e.g. 0.9 mm) capillaries, used for transport, protection and dispensing of various media [11]. Sutures with diameters starting from 0.1 mm [12] are obtained from many kinds of plastics, such as PP, PVA, PE, PA, and PUR, by using spinnerets [13]. The process of extrusion pelletising is applied in the production of micropellets with diameters of less than 1 mm [14] which are readily used in the increasingly popular process of rotational moulding [15] or in the production of modifying concentrates, so-called masterbatches (colour masterbatches, lubricant masterbatches, etc.) [16]. Micropellets can be widely used for the mass-production of nanofiber through bubble electrospinning or bubbfil spinning [17].

The aim of the experiments was to produce microrods using a twin-screw extruder, and to process sections of them into micropellets with the aim of establishing the functional relationship between the diameter of the microrods and their tensile strength, as well as determining the basic granulometric properties of the micropellets, including their natural repose angle and bulk density.

2. Experimental

2.1. Material

The micro-extrudate and the micropellets obtained from it were produced from a low-density polyethylene (LDPE) sold under the trade name Malen E, symbol FABS 23-D022, manufactured by Basell Orlen Polyolefins based in Płock (Poland). The main properties of this

polymer are presented in Table 1. This polymer is widely used in the processing industry, mainly in the manufacturing of various types of films (e.g., packaging films and shrink films), as it is highly elastic and weldable.

Table 1. The basic properties of the test polymer, according to the manufacturer's data

Property	Unit	Value
Standard density at 23°C	[kg/m^3]	915–920
Melting point	[°C]	105–118
Degradation temperature	[°C]	300
Degree of crystallinity	[%]	40–50
Melt flow rate (MFR) (190°C/2.16 kg)	[g/10 min]	1.95
Shore hardness	[°Sh D]	48
Vicat softening temperature A50, (50°C/h 10 N)	[°C]	91

2.2. Test stand

The tests were carried out using an EHP 2 × 20 Sline Zamak Mercator twin-screw extruder, intended for use in the extrusion of composite materials based on powder- and fibre-reinforced thermoplastics (Fig. 1). The co-rotating twin screw extruder, which was a component of a technological line, had a system of two volumetric feeders. The extruder was equipped with cylindrical, segmented screws with a length-to-diameter ratio of 40, and a cross-type head with two circular cross-section dies. The technical specification of the extruder is presented in Table 2. The temperatures of the feed opening zone and each zone of the barrel were regulated with electric heaters and a water jacket cooling system. The plasticizing system of the extruder was more abrasion- and corrosion-resistant due to the use of inserts that were hot-pressed using hot isostatic pressing (HIP) technology.

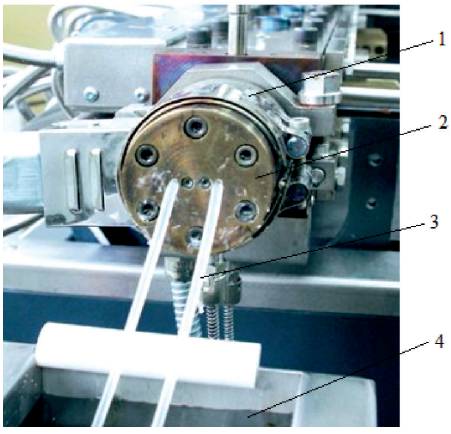


Fig. 1. A cross-type head with two 3 mm-diameter dies: 1 – extrusion head, 2 – heating component, 3 – extrudate, 4 – fragment of cooling bath

Table 2. Technical specification of the ZAMAK Mercator EHP 2×20 Sline extruder

Technical details	Unit	Value
Number of screws	[pcs.]	2
Screw diameter	[mm]	20
Working length of screws	[mm]	40D
Screw type	–	segmented
Barrel type	–	segmented, divided
Maximum rotational speed of screws	[rpm]	800
Number of barrel heating zones	[pcs.]	9
Maximum working pressure	[MPa]	20
Maximum working temperature	[°C]	400

Upstream of the cooling bath there was a G-16/32 Zamak Mercator pellet mill with a motor power of 1.66 kW, fitted out with an adjustable-speed puller and a milling cutter with 18 cutting edges, with an outer diameter of 125.25 mm, used for cutting the microrods into single micropellets. Microrod take-off speed and the rotational speed of the cutter were adjusted by a servo drive with a power of 1.6 kW via a control system. Microrods were passed onto a feed roller and then positioned by a pressure roller that cooperated with a drive roller (Fig. 2) – together, these formed a system of pull-up rollers. This system guided the feeding of the polymer in the direction of the rotating milling cutter and allowed for the simultaneous cutting of up to 4 microrods.

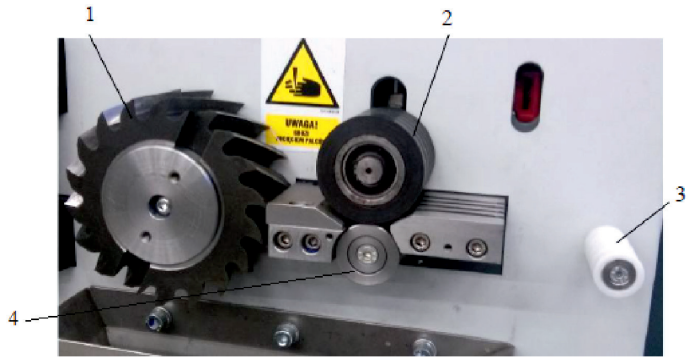


Fig. 2. The pelletizer: 1 – milling cutter, 2 – pressure roller, 3 – feed roller, 4 – drive roller

Tensile tests of the microrods were carried out using a Z010 Zwick/Roell tensile tester (Ulm, Germany). The Z010 machine operates with a maximum tensile load of 10 kN and a tensile speed of up to 2000 mm/min. The bulk pellet density and the natural repose angle of the micropellets were measured using a hopper with a closing slide, which was mounted and raised to a specific height on a measuring frame.

2.3. Test factors

A list of the most important test factors characterising the studied process was designed for the purposes of tensile and granulometric tests. The following test factors were considered:

Direct output factors:

- microrod diameter, *d*,
- diameter d and length of micropellets, *l*,
- microrod breaking force, *N*.

Indirect output factors:

- bulk density β, kg/m^3,
- angle of natural repose of pellets, deg,
- breaking stress σ, MPa.

Variable factors:

- microrod take-off speed in the pelletiser V – 6.1, 9.4, 14.6, 20.2, 27.1 m/min,
- rotational speed of the milling cutter – 80, 145, 195, 210, 335, 360, 600, 705 rpm.

Constant factors:

- processed material (LDPE),
- diameter of the extrusion head die 2 mm,
- others geometric elements of the extrusion head die and the plasticising unit,
- temperature in the individual zones of the plasticising system of the extruder (Table 3),

Table 3. The temperatures in the individual heating zones

Extruder component	Zone no.	Pre-set temperature [°C]
Feed opening zone	1	46.5
Plasticizing barrel	1	125
	2	130
	3	140
	4	150
	5	160
	6	160
	7	160
	8	170
	9	170
Connector	1	170
Head	1	170
	2	170

- polymer pressure in the extrusion head, 0.9 MPa;
- rotational speed of screws in the plasticising unit, 20 rpm.

Interfering factors:

- voltage, 219–241 V;
- relative air humidity, 55–65%;
- ambient temperature, 20–24°C. On the basis of the course of the micro-extrusion process, it was assumed that the impact of the interfering factors was small and could be neglected since those factors did not affect the results.

2.4. Method

To produce microrods of different diameters, the microrod take-off speeds in the pelletiser was manipulated. Microrod take-off speed was set on the front touch control panel of the pelletiser; the corresponding microrod diameters are shown in Table 4.

The microrods obtained in this way were measured, described and prepared for strength measuring. They were cut with scissors to the desired lengths (150 to 165 mm), so that they could be placed in the grips of the Zwick/RoellZ010 strength testing machine. Tensile tests of the microrods were performed according to the ISO 527-1:2012 standard at a test speed of 100 mm/min. Bulk density tests were carried out according to the ISO 60:1998 standard. Natural repose tests were prepared in accordance with the guidelines of ISO 6186:2001.

Table 4. The relationship between microrod diameter and take-off speed

Take-off speed [m/min]	Diameter [mm]
6.1	1.45
9.4	1.20
14.6	0.95
20.2	0.80
27.1	0.65

To obtain micropellets of different sizes, we used three microrod take-off speeds in the pelletiser and eight different speeds of the milling cutter (Table 5). The rotational speeds of the milling cutter were different each time as we wanted to obtain micropellets of the same length. Microrods with diameters of 0.8 mm and 0.65 mm were not pelletised, due to the milling cutter's inability to pelletize such small rods.

The smallest micropellet diameter obtained was 0.95 mm. As a final outcome, three different micropellet lengths and three different micropellet diameters were obtained giving a total of nine different combinations. Differences in the appearance of the micropellets are shown in photographs in Figures 3 to 5. Both microrods for strength tests and micropellets lengths were obtained at constant extruder setting parameters. The micropellets were dried at 50°C for approx. 24 hours.

Table 5. The relationship between the setting parameters of the Zamak Mercator G-16/32 pelletiser and the geometric dimensions of the pellets

Specimen No.	Pellet size		Take-off speed [m/min]	Rotational speed of milling cutter [rpm]
	Diameter [mm]	Length [mm]		
1	0.95	1.0	14.6	705
2		2.4		335
3		3.8		195
4	1.20	1.0	9.4	600
5		2.4		210
6		3.8		145
7	1.45	1.0	6.1	360
8		2.4		145
9		3.8		80

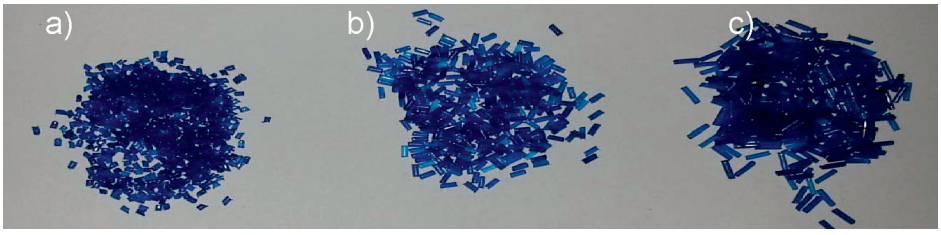


Fig. 3. Pellets with diameter d = 0.95 mm and length l: a) 1 mm, b) 2.4 mm, c) 3.8 mm

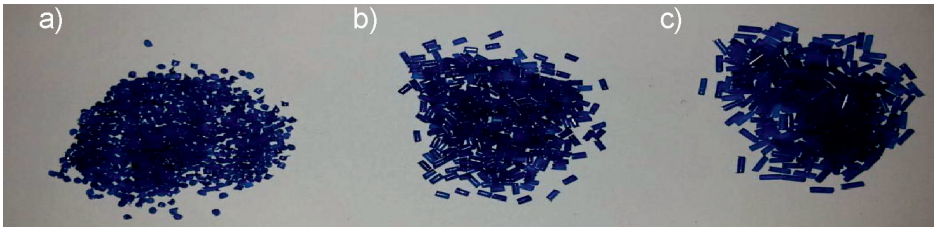


Fig. 4. Pellets with diameter d = 1.20 mm and length l: a) 1 mm, b) 2.4 mm, c) 3.8 mm

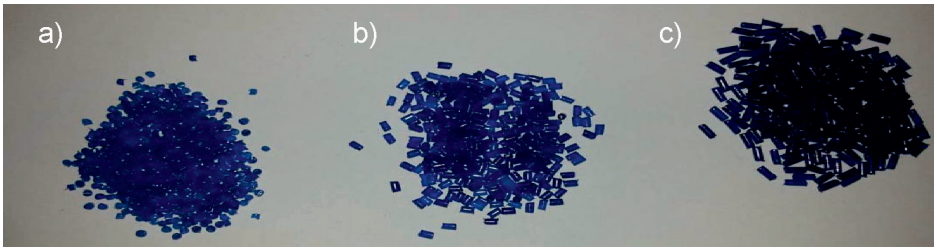


Fig. 5. Pellets with diameter d = 1.45 mm and length l: a) 1 mm, b) 2.4 mm, c) 3.8 mm

3. Results and discussion

Based on the results obtained in the experiments, the following graphs were plotted: tensile stress vs. strain for each specimen tested; breaking stress vs. microrod diameter; bulk density and natural repose angle vs. micropellet length at different pellet diameters. Regression equations were formulated where possible. Mathematical models made in this way describe in detail the connections between the studied indirect and direct factors in the accepted range of value changes.

3.1. Determination of strength parameters

The values of tensile stress as a function of strain for all tested microrods are shown in graphic form in Figure 6. The graph shows two tensile strength curves for microrods of each given diameter. Maximum tensile stress was equal to tensile stress at break in all cases. The curves of tensile stress as a function of strain did not differ significantly and the values of the test factors were comparable. The graph of tensile stress as a function of strain shows that out of all the microrods tested, those with a diameter of 0.65 mm showed the lowest strain at 200%, while the microrods with the largest diameter had approx. four times higher strain values. When analysing the results of the studies, it can be clearly concluded that the diameter of microrods has a significant influence on the strength properties of polymer products. This effect is probably caused by higher orientation of polymer chains in objects with a small diameter – a higher microrod take-off speed results in a higher orientation in a take-off direction and consequently increases ultimate tensile strength.

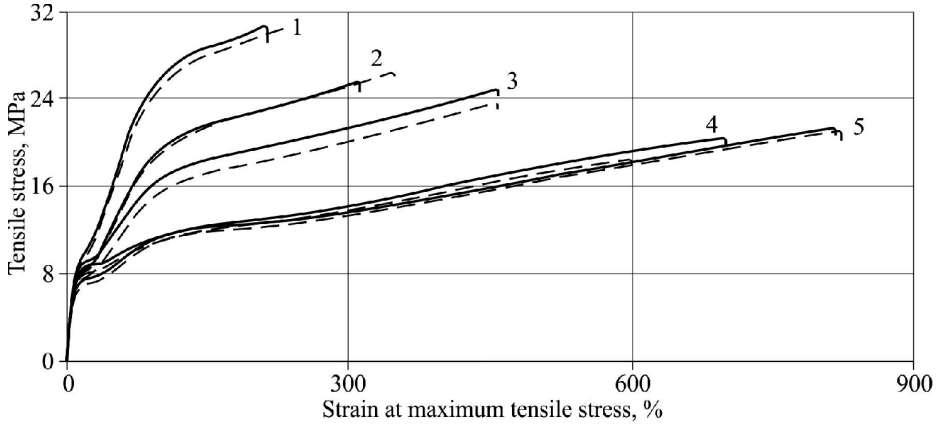


Fig. 6. Graph of stress σ vs. strain ε of test microrods with diameters of: 1–0.65 mm; 2–0.80 mm; 3– 0.95 mm; 4–1.20 mm; and 5–1.45 mm

Figure 7 shows a graph of stress needed to break a microrod as a function of microrod diameter. It was observed that along with changing microrod diameter, tensile stress at

break, and thus maximum tensile stress, increased. The maximum value of tensile stress at break was 30.60 MPa for the diameter of 0.65 mm, and the minimum value was 19.50MPa for the diameter of 1.20 mm, which represents a 36.27% decrease. This can be caused by a better packing of chains in the space and a large rapprochement of macroparticles, which results in stronger bonds among the chains and then gives improved mechanical properties of microrods of a smaller diameter.

On the basis of the study results, a computer-aided regression equation was determined – this describes the relationship between the tensile stress at break and microrod diameter using the approximation with the method of the smallest squares. The coefficients in the regression equation and the coefficient of correlation *r* were also determined by means of a computer programme.

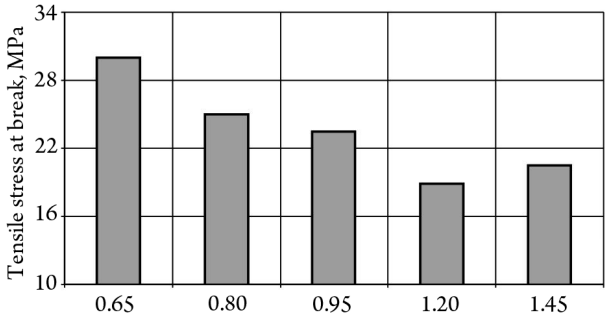


Fig. 7. Relationship between tensile stress at break and microrod diameter

In this case, the regression equation in the form of a second degree polynomial is as follows:

$$\sigma = 61.921d^2 - 65.829d + 25.589$$

The coefficient of correlation for this mathematical model is 0.983, this proves a very strong correlation between the tensile stress at break and the microrod diameter.

3.2. Granulometric tests

As before, regression equations describing the relationships between the bulk density and micropellet length (Fig. 8) and between the natural repose angle and micropellet length (Fig. 9) were determined as were the respective coefficients of correlations. The equations are as follows:

a) for micropellets diameter 0.95 mm:
   $\rho = 477.143\, l^2 + 43.571l - 10.714$; with $r = 0.999$;
   $\beta = 19.975\, l^2 - 1.391l + 0{,}115$; with $r = 0.999$;
b) for diameter 1.20 mm:
   $\rho = 488.469\, l^2 + 40.204l - 8.673$; with $r = 0.999$;
   $\beta = 1.081\, l^2 + 0.094l + 1.754$; with $r = 0.999$;

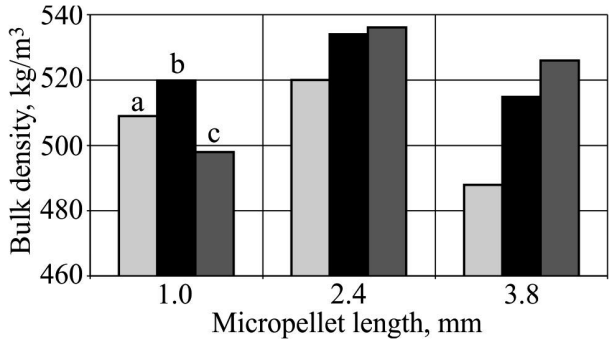


Fig. 8. The relationship between bulk density and micropellet length for different micropellet diameters: a) d = 0.95 mm, b) d = 1.20 mm, c) d = 1.45 mm

c) for diameter 1.45 mm:
$\rho = 438.918\, l^2 + 72.092l - 13.010$; with $r = 0.998$;
$\beta = 18.435\, l^2 - 0.773l + 0{,}038$; with $r = 0.999$.

Based on these results, it was found that LDPE micropellets with a diameter of 0.95 mm and a length of 3.8 mm had the lowest bulk density of 487.9 kg/m^3. The largest bulk density measured was 536.5 kg/m^3 for micropellets with a diameter of 1.45 mm and length of 2.4 mm.

The highest density values were recorded for micropellets with an average pellet length of 2.4 mm, and the lowest for the shortest micropellets. The difference between the highest and lowest bulk density of the test micropellets was approx. 10%. The bulk density of polymers depends upon, among other factors, the size and shape of pellets and the method of pellet placement in the hopper. The pellet placement is influenced by many factors: the method of introducing the pellets; height of their introduction; shape and dimensions of the hopper; geometrical features of the pellets [18].

The largest natural repose angle was measured for micropellets with a length of 1 mm and a diameter of 0.95 mm (Fig. 9). The differences between the smallest and the largest values of the natural repose angle were slight (approx. 2.6 degrees). Along with the decreasing micropellet length and diameter, the natural angle of repose increased. It can be assumed that this was due to the increase in frictional resistance of the external micropellets.

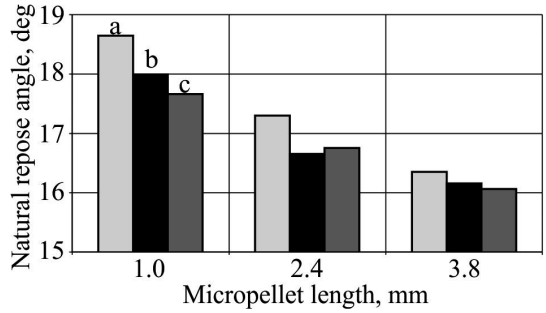


Fig. 9. The relationship between natural repose angle and micropellet length for different micropellet diameters: a) d = 0.95 mm, b) d = 1.20 mm, c) d = 1.45 mm

The value of the natural angle of repose is also affected by the shape and size of grains, as well as by the mutual mobility of the particles, which depends on the forces of adhesion between the individual particles, as well as the previously mentioned frictional resistance arising from the mutual movement of particles [19].

4. Conclusions

The experiments and the analysis of the strength properties of the microrods and the granulometric properties of the low-density polyethylene micropellets lead to the following conclusions:

- microrods with the smallest diameter showed the highest tensile strength and a correspondingly high stress at break, whilst at the same time showing the lowest strain – this indicates that the degree of longitudinal tension has a significant effect on strength properties, in particular, tensile stress;
- the dimensions of the micropellets, their lengths and diameters affect their bulk density and their natural repose angle. This means that it is necessary to choose one of the following options: either determin the geometric features of micropellets so that they match the existing design of the feed section of the plasticizing unit (in particular in an extruder), or adjust the design of the feed section of the plasticizing unit to the size of the micropellets that are to be processed. In this way, the process of plasticization will be optimized;
- the largest natural angle of repose was recorded for the smallest micropellets and vice versa. The values of the natural angle of repose can be used in designing the proper dimensions of the hopper of the plasticising unit to ensure the desired polymer flow.

References

[1] Jin G.B., Wang M.J., Zhao D.Y., Tian H.Q., Jin Y.F., *Design and experiments of extrusion die for polypropylene five-lumen micro tube*, Journal of Materials Processing Technology, 214, 1, 2014, 50–59.

[2] Wang H., Fu Z.H., Yao D.G., *Research of micro profile extrusion*, Advanced Materials Research, 87, 1, 2009, 125–129.

[3] Sahli M., Millot C., Roques-Carmes C., Khan Malek C., Barriere T., Gelin J.C., *Investigation of polymer inserts as prototyping tooling for micro injection moulding*, International Journal of Advanced Manufacturing Technology, 47, 1–4, 2010, 111–123.

[4] Bellantone V., Surace R., Trotta G., Fassi I., *Replication capability of micro injection moulding process for polymeric parts manufacturing*, International Journal of Advanced Manufacturing Technology, 67, 5–8, 2013, 1407–1421.

[5] Wang B.X., Wang W., Guo L.H., *The design of an extrusion head used to produce of micro-interventional catheter*/Electromachining & Mould, 4, 2005, 44–46.

[6] Jin G.B., Zhao D.Y., Wang M.J., Jin Y.F., Tian H.Q., Zhang J., *Study on design and experiments of extrusion die for polypropylene single-lumen micro tubes*, Microsystem Technologies, 21, 11, 2015, 2495–2503.
[7] Zhu C.W., Wu D.M., *Simulation of extrudate swell for multi-lumen precise medical catheter*, Plastics, 37, 6, 2008, 101–105.
[8] Tian H., Zhao D., Wang M., Jin G., Jin Y., *Study on extrudate swell of polypropylene in double-lumen micro profile extrusion*, Journal of Materials Processing Technology, 225, 2015, 357–368.
[9] Siracusano S., Ciciliato S., Ollandini G., Visalli F., *Catheters and Infections*, Clinical Management of Complicated Urinary Tract Infection, Intech, 2011, 83–98.
[10] Hayashi K., Takamizawa K., Saito T., Kira K. et al., *Elastic properties and strength of a novel small-diameter, compliant polyurethane vascular graft*, Journal of Biomedical Materials research Part B: Applied Biomaterials, 23, 1989, 229–244.
[11] Extrand C.W., Moon S.I., *Experimental measurement of forces and energies associated with capillary rise in a vertical tube*, Journal of Colloid and Interface Science, 407, 2013, 488–492.
[12] Murawski M., *The Onset and Progress of Microsurgical Technique*, Advances in Clinical and Experimental Medicine, 16, 5, 2007, 701–704.
[13] Callahan T.L., Lear W., Kruzic J.J., Maughan C.B., *Mechanical properties of commercially available nylon sutures in the United States*, Journal of Biomedical Materials Research Part B: Applied Biomaterials, 2016, DOI: 10.1002/jbm.b33600
[14] Thompson M.R., Xi C., Takacs E., Tate M., Vlachopoulos J., *Experiments and Flow Analysis of a Micropelletizing Die*, Journal of Polymer Engineering and Science, 44, 7, 2004, 1391–1402.
[15] Ramkumar P.L., Kulkarni D.M., Waigaonkar S.D., *Effect of Cycle Time on Mechanical Properties of LLDPE during rotational moulding process*, Polymers Research Journal, 8, 1, 2014, 19–33.
[16] Markarian J., *Pelletizing: choosing an appropriate method,* Plastics, Additives and Compounding, 6, 4, 2004, 22–26.
[17] He C.H., Li X.W., Liu P., Li Y., *Bubbfil spinning for fabrication of PVA nanofibers*, Thermal Science, 19, 2, 2015, 743–746.
[18] Sikora J.W., Samujło B., Stasiek A., *The extrusion of plasticized poly(vinyl chloride) in an extruder with a modified feed zone,* Part 1. Extrusion process, Journal of Polymer Engineering, 35, 2, 2015, 191–198.
[19] Dąbrowski A., *Adsorption – from theory to practice*, Advances in Colloid and Interface Science, 93, 2001, 135–149.

TECHNICAL TRANSACTIONS 2/2017
CZASOPISMO TECHNICZNE 2/2017
MECHANICS
DOI: 10.4467/2353737XCT.17.025.6218

Marek S. Kozień (kozien@mech.pk.edu.pl)
Dariusz Smolarski
Institute of Applied Mechanics, Faculty of Mechanical Engineering, Cracow University of Technology

Comparison of results of stress cycle counting by the direct spectral method and the Fu-Cebon method for bi-modal type stress processes

Porównanie rezultatów zliczania cykli naprężeń metodą spektralną bezpośrednią i metodą Fu-Cebona dla przebiegów o charakterze dwumodalnym

Abstract

This article compares the results of a high cycle fatigue analysis obtained through the application of the two methods dedicated to analysis of response of the structure of the bi-modal type – the direct spectral method and the Fu-Cebon one. the compared parameter is the lifetime for an assumed material S–N curve and stress spectrum defined in an article by Fu & Cebon.

Keywords: fatigue analysis, stress cycle counting, spectral method, bi-modal process

Streszczenie

W artykule dokonano porównania rezultatów wysokocyklowych analiz zmęczeniowych wykonanych przy zastosowaniu dwu metod dedykowanych dla przypadków odpowiedzi struktury o charakterze dwumodalnym: metody spektralnej bezpośredniej i metody Fu-Cebona. Porównywaną wielkością był czas życia konstrukcji przy założonej postaci krzywej S–N materiału i zmienności naprężeń o charakterystyce wykorzystywanej w artykule Fu i Cebona.

Słowa kluczowe: analiza zmęczeniowa, zliczanie cykli naprężeń, metoda spektralna, przebiegi dwumodalne

1. Introduction

In some cases the spectrum of vibrations of engineering structures have the bi-modal type. The example is the suspension system analysed by T.-T.Fu & D. Cebon [5].

Structural vibrations are the reason of changing deformation of material in time. the process is described by variable in time strain tensor components. The variation of strains accounts for material damage in a relatively large number of cycles for the elastic type of deformations. Changing in time strain tensor components are correlate with variable in time stress tensor components. In the paper, the high-cycle fatigue analysis based on the stress formulation is considered.

In the case of nonstationary or multimodal vibrations, the problem of cycle identification and cycle counting is crucial when conducting fatigue analysis. In literature, the following attempts at solving the problem are discussed: direct method (harmonic process, quasi-harmonic process) [4]; bi-modal dedicated methods [1, 2, 5, 14]; stress cycle counting method (e.g. the 'rain-flow' method) [4, 6, 11]; spectral methods (using the power spectral density function) [12, 13]; kurtosis analysis [15]. The most commonly used method is the 'rain-flow' method.

In their previous papers, the authors proposed the original method of stress cycle counting dedicated for the bi-modal type spectrum of stress – this is known as the spectral direct method [7–10]. T.-T. Fu & D. Cebon proposed in their article [5] a method of fatigue analysis that was also dedicated to the bi-modal process. The method proposed by T.-T.Fu & D. Cebon is based on using of the power spectra density description of the process and leads to determination of the probability density of the stress ranges. The method proposed by M. Kozień & D. Smolarski is a type of the time-domain method and is an alternative method of conducting fatigue analysis in cases of bi-modal stress history.

The aim of this paper is to compare results obtained by application of the two alternative but not being against methods: the spectral direct one and the Fu-Cebon one. The compared parameter is the lifetime of element for assumed material S–N curve and stress spectrum defined in article of T.-T.Fu & D. Cebon [5].

2. Basis of the spectral method

The determined bi-modal stress history can be theoretically defined with the following analytical form:

$$\sigma(t) = A_1 \sin(\omega_1 t + \varphi_1) + A_2 \sin(\omega_2 t + \varphi_2) \tag{1}$$

where:
A_1, A_2 – stress amplitudes of the harmonic components,
ω_1, ω_2 – angular frequencies of the harmonic components,
φ_1, φ_2 – phases of the harmonic components.

With such a formulation, the component frequencies f_1 and f_2 and corresponding periods T_1 and T_2 can be obtained using equations (2) and (3).

$$f_1 = \frac{\omega_1}{2\pi}, \quad f_2 = \frac{\omega_2}{2\pi} \tag{2}$$

$$T_1 = \frac{1}{f_1} = \frac{2\pi}{\omega_1}, \quad T_2 = \frac{1}{f_2} = \frac{2\pi}{\omega_2} \tag{3}$$

The basic assumptions and manner of application of the direct spectral method for bi-modal waveforms can be described as follows:

- For a given time-domain stress waveform, a spectrum is obtained. Let us assume that from this spectrum, the two frequencies f_1 and f_2 ($f_1 < f_2$) can be distinguished. They are characterized by periods T_1 and T_2, and amplitudes A_1 and A_2, respectively. The phase shift is not required and is therefore omitted in this method ($\varphi_1 = \varphi_2 = 0$). However, the proposed algorithm allows for the phase shift to be included in the calculation ($j_1 \neq 0$ or $j_2 \neq 0$) if necessary. The waveforms are analysed under the assumption that at the initial time, both of the harmonic components are of maximal (positive) value, which means that the stress value at time $t = 0$ is equal to $A_1 + A_2$.
- Period T_1 is the base period for the stress signal.
- Based on the values of periods T_1 and T_2, the so-called block of stress is determined. The block length (time range) T_B depends on the ratio T_1/T_2. It is the smallest integer number of period T_1, for which the ratio T_B/T_2 is an integer. In practical applications, this condition is nearly satisfied hence, assuming the value of T_B, is an arbitrary decision. The value of T_B depends on the precision of the determination of T_1 and T_2, usually by identification of frequencies f_1 and f_2.
- The primary stress cycle, which is the only cycle present within the block, has a stress amplitude of $A_1 + A_2$ and, if not stated otherwise (e.g. constant value present in FTT, static assembly stress or thermal stress), the effective mean stress is zero. This assumption is the basis for calculating the equivalent completely reversed uniaxial stress, e.g. Morrow's type (4) [4].
- The amplitudes of secondary stress cycles vary depending on the value of A_2 and the leading waveform of frequency f_1. Some of the identified cycles are not taken into account if they do not have a full stress-cycle form. For the high difference between frequencies f_1 and f_2, the amplitudes of the secondary cycles are approximately equal to A_2. The acquired data forms the basis of the calculation of the equivalent completely reversed uniaxial stress (4) [4].
- The obtained data which describes the identified stress cycles for a given waveform forms the basis of the fatigue analysis using the chosen stress cumulation hypothesis, e.g. Palmgreen-Miner's (5) [4].

In the analysis, the following parameters are used: effective mean stress (Sines stress) [4]; effective stress amplitude (according to the von Mises equivalent stress) [4]; equivalent completely reversed uniaxial stress (Morrow stress) (3) [4].

$$\sigma_{EQV} = \begin{cases} \dfrac{\sigma_a}{1-\dfrac{\sigma_m}{\sigma_u}} & \text{for } \sigma_m > 0 \\ \sigma_a & \text{for } \sigma_m \le 0 \end{cases} \tag{4}$$

$$B \cdot \sum_{i=1}^{N} \frac{N_i}{\left(N_f\right)_i} = 1 \tag{5}$$

$$T = B \cdot T_B \tag{6}$$

where:
σ_m – effective mean stress,
σ_a – effective stress amplitude,
σ_{EQV} – equivalent completely reversed uniaxial stress,
σ_u – ultimate stress,
N – total number of cycles identified in a block,
N_i – number of cycles with amplitude σ_i identified in a block,
$(N_f)_i$ – number of cycles to damage for stress with amplitude σ_i (S–N curve);
B – number of blocks;
T_B – time length of block;
T – estimated lifetime.

3. Estimation of values of amplitudes based on PSD function

The PSD function is calculated on the basis of a given signal $x(t)$ defined in the time domain. The two-sided PSD function $S_x(t)$ is defined as a Fourier transform of the autocorrelation function (7) and (8).

$$R_x(\tau) = \int_{-\infty}^{+\infty} x(t)\, x(t+\tau)\, dt \tag{7}$$

$$S_x(f) = \int_{-\infty}^{+\infty} R_x(\tau) e^{-i2\pi f \tau} d\tau \tag{8}$$

Because $S_x(t)$ is an even function of frequency (9), the commonly used the single-sided (or one-sided) power spectral density $G_x(f)$ (10) is defined and used.

$$S_x(-f) = S_x(f) \tag{9}$$

$$G_x(f) = \begin{cases} 0 \text{ for } f < 0 \\ S_x(f) \text{ for } f = 0 \\ 2S_x(f) \text{ for } f > 0 \end{cases} \tag{10}$$

Sometimes, the following problem arises: what are the averaged values of the signal when only the PSD function is known? The root mean square x_{RMS} value for a specific frequency range $f \in [f_1, f_2]$ can be calculated from its single-sided PSD function in the form (11) [3]. The amplitude of the signal, which is understood as the peak value x_{PEAK} of the harmonic signal, can be found based on formula (12).

$$x_{RMS} = \sqrt{\int_{f_1}^{f_2} G_x(f)\,df} \tag{11}$$

$$x_{PEAK} = \sqrt{2}\, x_{RMS} \tag{12}$$

4. Comparison of life time determined by the direct spectral and the Fu-Cebon methods

4.1. Analysed case

Let us consider after T.-T.Fu & D. Cebon's [5] fatigue analysis of the trailing arm of the car suspension system – this form is shown in Fig. 1.

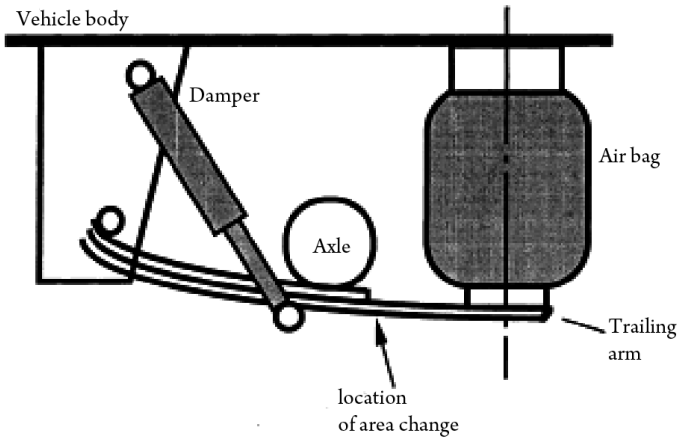


Fig. 1. A car suspension system [5]

The dominant stress component has the spectrum of the bi-modal type in steady-state case of its vibration, therefore the analysis takes a uniaxial form from the point of view of stress. The stress spectrum is defined as the single-sided power spectral density function shown in Fig. 2.

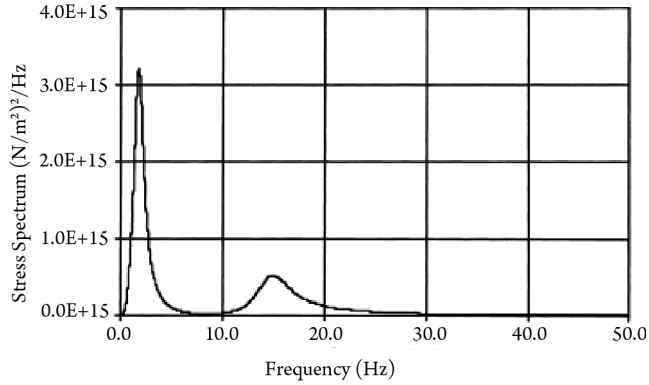


Fig. 2. One-sided power spectral density function of stress [5]

Based on the form of the one-sided power spectral density function of stress, it is possible to determine the probability function of the stress amplitudes using different methods presented in literature, e.g. pure bi-modal, Sakai, rain-flow, Rayleigh, Dirlik, Fu-Cebon [5]. With the application of the Fu-Cebon [5] method, the probability function of peak stress (stress amplitudes) can be visualised in the form shown in Fig. 3.

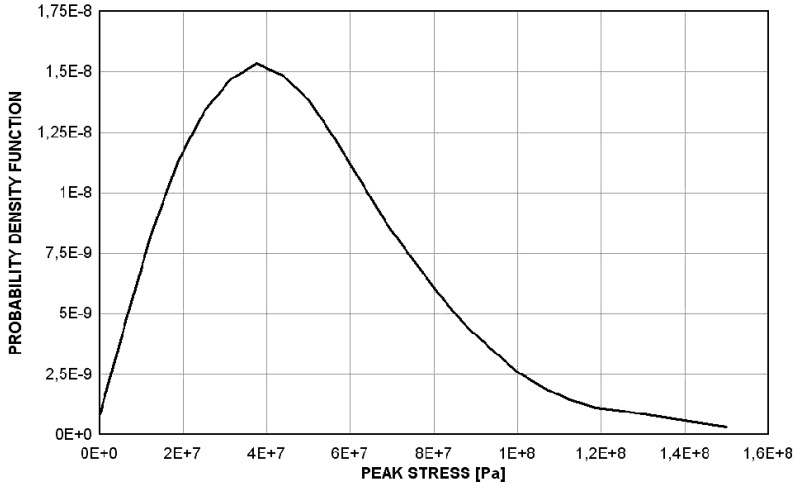


Fig. 3. Probability density function of peak stress [5]

Comparison was done for material the fatigue properties of which are described in the Wohler S–N curve in the form (13), where $m = 8$, and fatigue limit for $2.0 \cdot 10^6$ cycles is

100 MPa, hence $a = 2.0 \cdot 10^{22}$, where N_f is number of cycles for high-cycle fatigue damage and σ is stress amplitude for the completely reversed stress cycle. the ultimate stress σ_u is 400 MPa.

$$N_f = A\,\sigma^{-m} \tag{13}$$

4.2. Estimation of lifetime based on the Fu-Cebon method

The bi-modal idealised stress history considered by T.-T.Fu & D. Cebon [5] has the form (1) with assumption that $\varphi_1 = \varphi_2 = 0$. Moreover, it is assumed that the components are reasonably widely spaced in frequency, i.e. $\omega_1/\omega_2 > 4$. By applying the rain-flow counting method, the two distinct amplitude cycles are identified. the first with an amplitude $\sigma_1 = A_1 + A_2$ and $n_1 = \omega_1 T/2\pi$ cycles in T seconds and the second with an amplitude $\sigma_2 = A_2$ and $n_2 - n_1 = (\omega_1 - \omega_1)T/2\pi$ cycles. According to the Palmgren-Miner hypothesis, the total damage D is equal (14).

$$D = \frac{n_1}{N_1} + \frac{n_2 - n_1}{N_2} \tag{14}$$

where:
N_1 and N_2 – the cycles to failure at stresses σ_1 and σ_2, which follow the material S–N curve.

It is then assumed that the two components of equation (1) are narrow processes and their distributions of peaks $P_1(\sigma)$ and $P_2(\sigma)$ follow Rayleigh distributions. After suitable manipulations, the fatigue life estimated by this approach is [5]:

$$T = \frac{2\pi C}{\omega_1 \int_0^\infty \frac{P_1(\sigma)}{\sigma^{1/b}} d\sigma + (\omega_1 - \omega_2) \int_0^\infty \frac{P_2(\sigma)}{\sigma^{1/b}} d\sigma} \tag{15}$$

where:
$C = a$ and $1/b = -m$ for the material S–N curve equation (13).

Estimation of lifetime T_{FC} is done based on probability density function of peak stress determined with application of the Fu-Cebon formula (Fig. 3). The value of lifetime for a given material (the S–N curve) can be estimated from the Dirlik formula (16), where M^+ is the expected number of peaks in unit of time and $p(\sigma)$ is the probability density functions of the stress ranges [12].

$$T_{FC} = \frac{1}{M^+ \int_0^{+\infty} \frac{p(\Delta\sigma)}{N_f(\Delta\sigma)} d(\Delta\sigma)} \tag{16}$$

For the analysed case the obtained lifetime is $T_{FC} = 3.27 \cdot 10^9$ cycles.

4.3. Estimation of lifetime based on the direct spectral method

The power spectra density function of stresses (Fig. 2), show two components in frequency domain: f_1 = 2 Hz and f_2 = 15 Hz. The periods corresponding with this frequencies are respectively T_1 = 0.5 sec (the basic period), T_2 = 0.0(6) sec (the secondary period). Therefore, the ratio is T_1/T_2 = 7.5 – this means that one stress block has total length (time duration) equal to $T_B = 2T_1$ = 1s, T_2 = 1 sec. The values of amplitudes A_1 and A_2 were identified based on the stress PSD function (Fig. 2) with the application of formulas (11) and (12). The values of integral existing in (11), were calculated as areas of triangles with length 0.3125 Hz and heights respectively: $3.22 \cdot 10^{15}$ Pa^2/Hz for f_1 = 2 Hz and $1.26 \cdot 10^{14}$ Pa^2/Hz for f_2 = 15 Hz. This means that the very narrow frequency range connected with discretisation of the PSD function in the frequency domain was taken into account. the values of the identified stress amplitudes are approximately equal to: A_1 = 32 MPa, A_2 = 6 MPa.

Stress cycle identification for a given block depends on the assumed values of the phase angles φ_1 and φ_2. For uniaxial stress, there is no reason for different values of angles; however, due to the method of cycle identification in the spectral direct method, the values of quantitative parameters describing the secondary cycles depend upon the chosen value of phase angles $\varphi_1 = \varphi_2$. The two cases were analysed: $\varphi_1 = \varphi_2 = 0$ and $\varphi_1 = \varphi_2 = \pi/2$. Reconstructed stress function in time domain, which were basis for cycle identification, is shown in Fig. 4 for one block. Hence, the amplitudes of the main cycle (32 MPa) and the secondary cycles (6 MPa) and then application of the direct spectral method algorithm leads to identification the given in Table 1 and Table 2 groups of cycles for one block with length 1 sec.

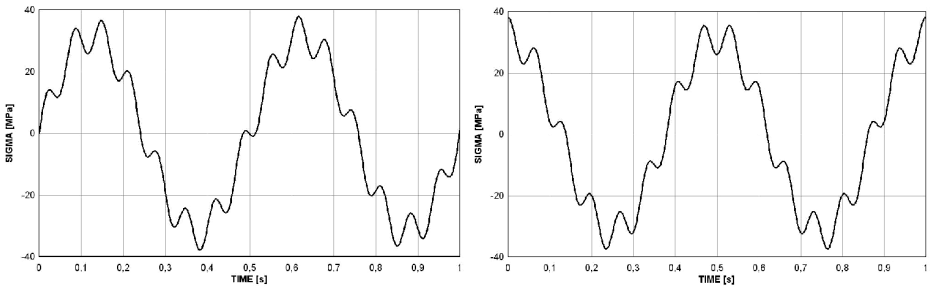


Fig. 4. Reconstructed variation of stress as functions of time: for $\varphi_1 = \varphi_2 = 0$ (left), for $\varphi_1 = \varphi_2 = \pi/2$ (right)

Even though there are differences in the time form of reconstructed stress functions, the number of identified stress cycles, especially secondary cycles, were the same in all cases (the one main cycle and the seventeen secondary cycles). After application of the Palmgren-Miner rule of stress cycle cumulation for a given material, the estimated lifetime was $T_{DS} = 4.6 \cdot 10^9$ cycles for both cases. This result is not far from that which was previously obtained after application of the Fu-Cebon method – $T_{FC} = 3.27 \cdot 10^9$ cycles.

Table 1. Identified cycles of one block for the case $\varphi_1 = \varphi_2 = 0$

NO.	σ_{MAX} [MPa]	σ_{MIN} [MPa]	$\overline{\sigma}_m$ [MPa]	$\overline{\sigma}_a$ [MPa]	σ_{EQV} [MPa]	NUMBER	TYPE
1	38.0	–38.0	0	38.0	38.0	1	main
2	14.1	11.7	12.9	1.2	1.2	1	secondary
3	34.0	25.8	29.9	4.1	4.2	1	secondary
4	36.6	25.8	31.2	5.4	5.5	1	secondary
5	20.3	17.1	18.7	1.6	1.6	1	secondary
6	–5.7	–7.5	0	0.9	0.9	1	secondary
7	–24.3	–30.4	0	3.1	3.1	1	secondary
8	–24.3	–37.8	0	6.8	6.8	1	secondary
9	–21.3	–25.8	0	2.3	2.3	1	secondary
10	0.9	–0.9	0	0.9	0.9	1	secondary
11	25.8	21.3	23.6	2.3	2.3	1	secondary
12	37.8	24.3	31.1	6.8	6.9	1	secondary
13	30.4	24.3	27.4	3.1	3.1	1	secondary
14	7,5	5.7	6.7	1.0	1.0	1	secondary
15	–17.0	–20.3	0	1.7	1.7	1	secondary
16	–25.8	–36.5	0	5.4	5.4	1	secondary
17	–25.8	–34.1	0	4.2	4.2	1	secondary
18	–11.8	–14.1	0	1.2	1.2	1	secondary

Table 2. Identified cycles of one block for the case $\varphi_1 = \varphi_2 = \pi/2$

NO.	σ_{MAX} [MPa]	σ_{MIN} [MPa]	$\overline{\sigma}_m$ [MPa]	$\overline{\sigma}_a$ [MPa]	σ_{EQV} [MPa]	NUMBER	TYPE
1	38.0	–38.0	0	38.0	38.0	1	main
2	38.0	23.0	30.5	7.5	7.6	1	secondary
3	28.2	22.9	25.6	2.7	2.7	2	secondary
4	4.2	2.4	3.3	0.9	0.9	2	secondary
5	–19.3	–23.0	0	1.9	1.9	2	secondary
6	–25.3	–37.3	0	6.0	6.0	2	secondary
7	–25.3	–32.4	0	3.6	3.6	2	secondary
8	–8.8	–10.9	0	1.1	1.1	2	secondary
9	17.2	14.5	15.9	1.4	1.4	2	secondary
10	35.4	26.0	30.7	4.7	4.8	2	secondary

4.4. Detailed conclusions of comparison

The analysis makes it possible to formulate the following detailed conclusions:

- Application of the spectral direct method gives lower errors if the spectral characteristics of the stress variation in time are given in form of the amplitude-frequency characteristics then the power spectra density function, because there is no need to transform from frequency-domain to time-domain. This is observation in the case of the spectral method, but not in comparison with the Fu-Cebon method.
- When leakage exists for spectral functions of stress (amplitude-frequency function or PSD one) around modal frequencies – this is often observed for realistic cases – the final results of cycle identification by the spectra direct method strongly depend on the chosen limits of integration in formula (8). For the discussed analysis, the integrations were performed for a very narrow frequency range. If the range is wider, the analysis would by much conservative.
- If the spectral characteristics of the stress variation in time are given only in the form of amplitude-frequency function or the power spectra density function, the reconstruction of the function is not unique due to a lack of information about values of the phase angles φ_1 and φ_2. For detailed reconstruction, the values of the angles must be known; however, if there is no information about these values, the method can be applied. If there is no other reasons, usually it is assumed that $\varphi_1 = \varphi_2 = 0$. For a uniaxial stress case, the condition $\varphi_1 = \varphi_2$ is always valid.

5. Conclusions

The fundamental conclusion resulting from the analysis is that the lifetime estimation of the analysed case is almost the same when identified using each method. This can be interpreted as verification of the possibility of using the spectral direct method for the fatigue analysis of the bi-modal process.

Moreover, the analyses make it possible to formulate the following conclusions*:*

- The bi-modal process exists in practical engineering applications.
- The methods based on the integration of the probability density of the stress ranges may not take the sum of modes into account.
- The results of analyses used spectral direct method applied to bi-modal type processes defined by power spectral density function strongly depend on the arbitrary chosen range of integration of the PSD for identified frequencies.
- The Fu-Cebon method based on the analysis of the PSD function can be applied in a natural way for the results experimentally obtained when the characteristics take into account random description. The spectral direct method can be applied in a natural way for the results of the dynamical response of a structure obtained from computer simulation based on the assumption of deterministic analysis.

References

[1] Benasciutti D., Tovo R., *On fatigue damage assessment in bimodal random process*, International Journal of Fatigue 29, 2007, 232–244.

[2] Braccesi C., Cianetti F., Lori G., Pioli D., *Fatigue behaviour analysis of mechanical component subject to random bimodal stress process: frequency domain approach*, International Journal of Fatigue 27, 2005, 335–345.

[3] Brandt A., *Noise and vibration analysis. Signal analysis and experimental procedures*, Wiley, Chichester 2012.

[4] Dowling N.E., *Mechanical behaviour of materials. Engineering methods for deformations, fracture and fatigue*, Prentice-Hall International Editors Inc., Englewood Cliffs, 1993.

[5] Fu T.-T., Cebon D., *Predicting fatigue for bi-modal stress spectral density*, International Journal of Fatigue 22, 2000, 11–21.

[6] Kocańda S., Szala J., *Podstawy obliczeń zmęczeniowych*, PWN, Warszawa 1997.

[7] Kozień M.S., Smolarski D., *Analytical simulation of application of FFT based spectral method of fatigue cycle counting for multiaxial stress on example of pulse excited beam*, Engineering Mechanics 19(5), 2012, 1–7.

[8] Kozień M.S., Smolarski D., *Formulation of the spectral direct method for cycle counting of bimodal multiaxial stress*, Book of Abstracts, 39th Solid Mechanics Conference SolMech, Kowalewski Z.L., Ranachowski Z., Widłaszewski J. (Eds.), Warszawa–Zakopane 2014, 305–306.

[9] Kozień M.S., Smolarski D., *Formulation of a direct spectral method for counting of cycles for bi-modal stress history*, Solid State Phenomena, Vol. 224, 2015, 69–74, DOI: 10.4028/www.scientific.net/SSP.224.69.

[10] Kozień M.S., Szybiński B., *Method of estimation of life time for vibrating engineering structures with irregular time history response*, Proceedings of the 5th International Conference on Very High Cycle Fatigue, Berlin 2012, 539–544.

[11] Ligaj B., *An analysis of the influence of cycle counting methods on fatigue life calculations of steel*, Scientific Problems of Machines Operation and Maintenance, Vol. 4 (168), 2011, 25–43.

[12] Niesłony A., *Wyznaczanie warstwic uszkodzeń zmęczeniowych metodą spektralną*, Studia i Monografie z. 233, Politechnika Opolska, Opole 2008.

[13] Niesłony A., Macha E., *Spectral method in multiaxial random fatigue*, Springer, Berlin–Heidelberg 2007.

[14] Sakai S., Okamura H., *On the distribution of rainflow range for Gaussian random process with bimodal PSD*, JSME International Journal, Serie A 38, 1995, 440–445.

[15] Van Baren J., Van Baren P., *the fatigue damage spectrum and kurtosis control*, Vibration Research Corporation, Jenison MI, December 2009.