NJD-iScience

№175

The Norwegian Journal of Development of the International Science

NJD-iScience

№175/2026

Norwegian Journal of development of the International Science

ISSN 3453-9875

It was established in November 2016 with support from the Norwegian Academy of Science.

DESCRIPTION

The Norwegian Journal of Development of the International Science, established in 2016 with support from the Norwegian Academy of Science, offers a platform for global researchers, academics, and students. Publishing 24 times annually, it covers current science topics. Submitting your article provides exposure to a diverse audience, enhancing your career. Our commitment to research and SEO ensures timely visibility. Submit to the Norwegian Journal of Development of the International Science today and join esteemed authors.

1000 copies
Norwegian Journal of development of the International Science
Iduns gate 4A, 0178, Oslo, Norway
email: publish@nor-ijournal.com
site: https://nor-ijournal.com

CONTENT

ARTS

THE STYLISTIC FEATURES OF ARAM KHACHATRYAN'S MUSICAL LANGUAGE

Papanyan N.
Vanadzor state university named after H. Tumanyan,
Faculty of Social Sciences,
Department of Art and Sports, Lecturer, Armenia

СТИЛИСТИЧЕСКИЕ ОСОБЕННОСТИ МУЗЫКАЛЬНОГО ЯЗЫКА АРАМА ХАЧАТУРЯНА

Папанян Н.А.
Преподаватель, факультет общественных наук, кафедра искусства и спорта Ванадзорский государственный университет имени О. Туманяна, Армения
https://doi.org/10.5281/zenodo.18840562

Abstract

The article examines the stylistic features of A. Khachaturian's music. It highlights the originality, richness, and intricate contrast of his rhythmic texture, as well as the specific traits of his polyphony, manifested not only in melodic writing but also in meter and rhythm, voice leading, and the interpretation of various genres. The study presents the characteristics of dance rhythm both in ballet music and in concert and symphonic works. It also explores the modal-intonational, metro-rhythmic, and timbral connections with Armenian folk-professional music, and reveals creative links with Russian and Western music.

Аннотация

В статье рассматриваются стилистические особенности музыки А.Хачатуряна. Оригинальность, богатство и прихотливая контрастность ритмической фактуры, особенности полифонии, проявляющиеся не только в мелодическом изложении, но и в метроритме, голосоведении, в трактовке различных жанров. Представлены особенности танцевальной ритмики как в балетной музыке, так и в концертных и симфонических произведениях. Рассмотрены ладоинтонационные, метроритмические, тембровые связи с армянской народно-профессиональной музыкой. Выявлены творческие связи с русской и зарубежной музыкой.

Keywords: rhythm, meter, polyphony, dance character, orchestration.
Ключевые слова: ритм, метр, полифония, танцевальность, оркестровка.

ВВЕДЕНИЕ. Каждая эпоха по-своему воспринимает гениальных художников. В наше время богатейшее творческое наследие Арама Хачатуряна покоряет жизненной энергией, страстным утверждением счастья, любви и права личности на свободу. С его творчеством в музыку вошёл новый мир идей и образов.

С самых первых шагов композитора его музыка привлекала слушателя красочностью, роскошью звучаний и темпераментом, неукротимой динамикой ритма. Не случайно Борис Асафьев назвал его «Рубенсом нашей музыки», подчёркивая живописность и богатство его художественного мышления.

Неиссякаемым источником творческого вдохновения Хачатуряна было богатейшее народное искусство — городская музыкальная традиция, ашугско-гусанское искусство, а также прихотливая и живая ритмика танцевальных жанров. Эти элементы в его произведениях получили новое художественное осмысление и стали важной частью его самобытного композиторского языка.

Армения — одна из древнейших культурных стран мира — исторически находясь на границе Востока и Запада, восприняла культуры различных миров и, в свою очередь, оказала влияние на их развитие. Проблему «Восток–Запад» Хачатурян поднял на новый, принципиально иной уровень, создав оригинальную модель художественного синтеза, в которой органично сочетаются национальное мышление и общечеловеческие ценности.

Несмотря на многочисленные книги, статьи и исследования, посвящённые Араму Хачатуряну, ряд эстетических и теоретических вопросов, связанных с его творчеством, до сих пор не получил всестороннего научного освещения. Настоящая статья представляет собой попытку охарактеризовать одну из важнейших сторон его творчества — стилистику ритма и ритмической драматургии, раскрывая её структурную и идейную роль.

МАТЕРИАЛЫ И МЕТОДЫ. В мировой музыке XX века чётко обозначились качественно новые подходы к арсеналу выразительных средств. Самостоятельную и первостепенную роль приобрели тембр, ритм, гармония, которые нередко развивались независимо от мелодии и фактуры.

В музыке Хачатуряна, насыщенной стремительной, напевной и пластичной мелодикой, своеобразным ключом к её пониманию стал ритм. Однако оригинальность, богатство и прихотливая контрастность ритмической фактуры для композитора не были самоцелью — они глубоко укоренены в истоках армянского народного искусства.

С древнейших времён в армянской народной музыке существовали три принципа развития ритма: танцевальный, песенный и речитативно-импровизационный. Творчески усвоив стихию ритма национального мышления, Хачатурян раскрывает процесс человеческого мышления через эмоции и чувства.[5] Для него танец становится «эмоцией мысли и мыслью чувства». В его музыке огромной эмоциональной силы нет «полумыслей» и «получувств», что во многом обусловлено чётко организованной ритмической стихией. Обращаясь к характерным народным ритмическим образам, Хачатурян насыщает их несвойственной им эмоциональностью и динамикой. Здесь ярко проявляется его исключительная способность к преобразованию народного материала, способствующая выражению идеи динамического развития.

Истоки богатства ритмической стихии его музыки коренятся не только в восточных (народных), но и в европейских традициях. Он творчески перенял традиции и принципы Л.Бетховена, западноевропейских композиторов-романтиков, П. Чайковского, С. Рахманинова, С. Прокофьева и других.[5]

Одной из характерных особенностей музыкального мышления композитора является полиритмия. В целом полифонические и полиметрические приёмы стали своеобразным почерком многих композиторов XX века — С. Прокофьева, Д. Шостаковича, Б. Бартока, К. Орфа и других. У Хачатуряна же полиритмия стала результатом обогащения ритмических структур и проявлением предельного напряжения ритмического чувства. Благодаря полиритмии даже вспомогательные голоса, лишённые яркой мелодической выразительности, приобретают самостоятельное значение.

Не ограничиваясь несколькими ритмически самостоятельными слоями, А. Хачатурян, при наличии общего метра для всех голосов, посредством смещения акцентов и наслоения аккордов на слабых долях создаёт впечатление полиметрии. Истоки подобной метроритмики уходят в исполнительские традиции народных инструментальных ансамблей — сазандаров. Импровизация ударного инструмента дафа во время исполнения представляет собой цепь бесконечных ритмических сочетаний, образующих с основной мелодией разнообразные полиритмические и полиметрические комбинации.

В произведениях композитора полиметрические соотношения возникают между темой и контрапунктом, двумя различными темами, мелодией и её аккордовым сопровождением. Ритмическая стихия в его музыке развивается в двух направлениях — танцевально-жанровом и песенно-импровизационном. Тем самым композитор остаётся верен принципам ритмического развития, характерным для музыкальной традиции народов Закавказья, особенно армянской ашугско-гусанской культуры.

В танцевально-жанровой сфере отчётливо противопоставлены два полюса образно-эмоционального содержания: страстно-мужественный, энергичный и живой, а так же - изящный и капризно-грациозный. Эти сферы связаны с укоренённой в национальной традиции ритмической стихией мужских и женских танцев.

Ритм тесно переплетается с особенностями оркестровки, поскольку тембровые факторы изложения существенно различаются в танцевальных, эпико-драматических и пейзажно-красочных эпизодах. Особенно примечательна роль ударных инструментов (как шумовых, так и с определённой высотой звучания). В партитурах балетов и симфонических произведений наряду с колоколами, ксилофоном, треугольником композитор использует и редко применяемые инструменты — флексатон (II часть Фортепианного концерта), тубафон («Танец египетских девушек», «Танец фей» из балета «Спартак»), деревянные кубики («Танец римской куртизанки», «Танец юных фракийцев с мечами»), а также кавказские народные инструменты — даф («Танец стариков», «Танец мужчин» в балете «Гаянэ») и дхол (доол) («Танец хлопка», «Танец старика и ковровщиц» в балете «Гаянэ»). Многие из этих инструментов выполняют не только ритмико-динамическую и колористическую, но и мелодическую, даже тематическую функцию. Так, во Второй симфонии колоколам поручена роль лейттемы — «тема тревоги, тема набата». В этой поистине «народной» симфонии композитор противопоставляет двойные образные сферы — набат и скорбный плач во вступлении, энергичный танец и протест-плач /1-я часть побочная партия/, траурное шествие и материнская скорбь /3-я часть, тема народной песни «Ворскан Ахпер» («Охотник-брат») и средневековая секвенция Dies irae/. Острота этих контрастов подчеркивается изобретательностью метроритма: размеренное шествие и импровизационное изложение. В этом отношении А. Хачатурян продолжает оперные и симфонические традиции русских классиков М. Глинки, П. Чайковского, С. Рахманинова, С. Прокофьева.

Другим проявлением концертности А. Хачатуряна являются его концерт-рапсодии, написанные в 1960-е годы, в которых получают дальнейшее развитие новаторские принципы «восточного» симфонизма. Здесь особенно ярко проявляются контрасты полярных ритмообразований, блестящее мастерство вариационного развития и тяготение к эпико-народным жанровым приёмам. Этот триптих рапсодий стал своеобразной энциклопедией технических трудностей современного инструментального искусства, где композитор с особой изобретательностью выступает в создании ломаных ритмических акцентов.[5]

Особого внимания заслуживает танцевальная ритмика А. Хачатуряна, наиболее ярко раскрытая как в балетах, так и в оркестровых и фортепианных произведениях. Одним из самых выдающихся балетов нашего времени является «Гаянэ» — благодаря сквозному развитию центрального лирического образа, методам его характеристики и трактовке танцевальности. Балет щедро насыщен разнообразными формами армянской народной танцевальной музыки, богатыми и прихотливыми ритмами, полиметрией. Вместе с тем они органично сочетаются и переплетаются с мелодической песенностью, перерастая в своеобразную ритмоинтонационную кантилену. Их предпосылки коренились в художественно отточенных движениях народных танцев, в изящных девичьих плясках с «поющими» движениями рук и плавной гибкостью женского корпуса. Танцы балета чётко подразделяются на две жанровые сферы. К страстно-мужественным, энергичным танцам относятся всемирно известный «Танец с саблями» (своего рода «визитная карточка» композитора), «Танец горцев», «Танец молодых курдов», «Лезгинка», в которых пунктирная ритмика становится исходной основой мелодики и гармонии, насыщенной секундовыми «гроздьями». В лирико-грациозных танцах, обращаясь к подлинным народным напевам — «Пшаты цар» («Дерево лора») (в «Танце сборщиц хлопка»), «Калосы пркен» («Танец ковровщиц»), — А. Хачатурян смело раскрывает и по-своему обновляет их ритмоинтонационные и стилистические приёмы. Интересно также обращение композитора к своеобразной мелодике и ритмике танцев других народов, например «Лезгинка», украинский «Гопак», плавная грациозная «Узундара» и др. Жизнерадостная и живописная танцевальность ярко выражена не только в балетных номерах («Вариации Нунэ», «Танец розовых девушек», «Танец Айше» и др.), но и в симфонических произведениях (главная партия финала Скрипичного концерта, побочные темы I и II частей Второй симфонии и др.). В балете «Спартак» ритмическая стихия приобретает особую драматическую напряжённость. Здесь рождаются новые для классического балета возможности движения — динамичные, подчёркнуто-мужественные, выражающие протест и гнев. Ритм становится основой развития лейтмотивов и тематических трансформаций.[2]

Свой творческий путь композитор начинал с фортепианных произведений — «Танец», «Вальс-каприс», «Поэма», Токката», где сочетаются изощренная мелодика и импульсивная танцевальная ритмика. Уже здесь проявились характерные черты его стиля — острая ритмическая пульсация, плотная аккордовая фактура с резкими martellato и песенно-импровизационное многоголосие. На основе народных танцевальных ритмов он создаёт смелые полиритмические сочетания.

Примером отточенности стиля стал Фортепианный концерт, в котором сочетаются возвышенно-патетическое и лирико-песенное начала. Здесь зарождается токкатность — одна из излюбленных форм его музыкального изложения. Ярким проявлением двух аспектов одухотворенной образности — ашугской импровизационности и импульсивной ритмической танцевальности — стали Скрипичный концерт, некоторые пьесы и песни.

Иной сферой творчества композитора является маршевость. А. Хачатурян был одним из первых советских композиторов, систематически создававших марши для духового оркестра. Его «Походные марши» отличаются богатой мелодикой и своеобразной ритмикой. «Зангезурский марш» (из кинофильма «Зангезур») понсути представляет собой своеобразную песню-шествие с широкой мелодией и оригинальным ритмическим рисунком. Если эту тему мысленно пропеть в медленном темпе, то она предстанет как мечтательное танцевальное адажио.[5]

Одной из характерных особенностей музыки А. Хачатуряна является полифония. Из полифонических форм он предпочитает те, которые не сковывают его свободное мышление. Первым творческим опытом стали его «Семь фуг», в которых на основе армянского музыкального материала органично сочетаются восточный фольклор и достижения западноевропейской и русской классической музыки. В то же время он выступает как новатор, нарушающий традиции. Об этом свидетельствуют слова самого композитора: «Для меня этот вечный конфликт осложнялся также стремлением соединить “прокрустово ложе” классической формы с требованиями своеобразного музыкального языка Востока». Не случайно впоследствии вместо традиционных прелюдии, токкаты и фантазии в «микроциклы» он вводит речитатив. Опыт, приобретённый в области полифонического стиля, стал плодотворной основой для дальнейшего формирования и развития полифонического мышления композитора. Полифония как средство противопоставления и развития образов входит в различные жанровые сферы — симфонию, концерт, балет. В разработочных разделах он широко обращается к имитации, создавая яркие фугато, либо активно использует остинатное изложение, прибегая к древней форме пассакалии. [4]

ЗАКЛЮЧЕНИЕ. Таким образом, в непрерывно обновляющемся звуковом потоке музыки А. Хачатуряна огромную роль играет стилистика использования как ритма, так и полифонического мышления. Он объединяет отдельные элементы в целостную систему, динамизирует развитие и определяет напряжённый пульс произведений. В отличие от многих композиторов, обращавшихся к восточным истокам, ритмическое мышление А. Хачатуряна поражает не сложностью метра, а блестящим искусством варьирования: смещением акцентов, дроблением ритмических фигур, изобретательным использованием синкоп. Яркий пример — переосмысление традиционной формулы вальса, насыщенной элементами восточного танца («Вальс» из «Маскарада», танец Айши из «Гаянэ»).

А. Хачатурян, опираясь на народные и национальные истоки, создал собственную традицию в армянской музыке — так же, как Мартирос Сарьян в живописи и Александр Таманян в архитектуре. Хачатуряновсский стиль сформировался под непосредственным воздействием традиций Комитаса, кто по праву считается основоположником армянской классической музыки. Творческая, научная и музыковедческая деятельность Комитаса и А. Хачатуряна сопоставима со вкладом Месропа Маштоца в истории формирования армянской науки и письменности.

Список литературы:

1. Мартынов И. «Арам Хачатурян», изд-во «Музыка»,Ленинград, 1978
2. Тигранов Г. «Балеты А. Хачатуряна», изд-во «Музыка», 1974
3. Тигранов Г. «Арам Ильич Хачатурян», М.Музыка, 1987
4. Чеботарян Г. «Полифония в творчестве Арама Хачатуряна», изд-во «Айастан», Ереван, 1969
5. Մ.Հարությունյան, Ա.Բարսամյան «Հայ երաժշտության պատմություն» /հնագույն շրջանից մինչև մեր օրերը/, Երևան 1996

EARTH SCIENCES

ASSESSMENT OF THE NATURAL REGULATİON LEVEL OF RİVERS AND THEİR RUNOFF DURİNG İNTENSİVE İRRIGATİON PERİOD

Hashimov E.
Doctor of National Security and Military Sciences, professor,
Technical University / National Defense University,
Baku, Republic of Azerbaijan
https://orcid.org/0000-0001-8783-1277
Teymurov M.
PhD in Geographical Sciences,
Head Researcher, Associate Professor,
Ministry of Science and Education, Institute of Geography,
Baku, Republic of Azerbaijan
https://orcid.org/0009-0006-8823-4670
İsayeva R.
Researcher in Geographical Sciences,
Ministry of Science and Education, Institute of Geography,
Baku, Republic of Azerbaijan
https://orcid.org/0009-0003-5240-1498
https://doi.org/10.5281/zenodo.18840564

Abstract

The article is devoted to the study of the natural regulation of rivers and the runoff during the intensive irrigation period. Since the Republic of Azerbaijan is located in an arid region, there is an acute shortage of water, and river waters constitute the main part of its water resources. However, since most of the rivers are drying up, their transit and natural regulation levels are unfavorable, the possibilities of using their water are weakened. The natural regulation of rivers refers to the favorable distribution of their runoff throughout the year. The main part of the river runoff occurs in the spring months, when the demand for water is low. During the irrigation period (summer months), when the demand for water is highest, the level drops. As a result, a serious disproportion arises between the volume of river water runoff and the possibilities of its use. The southern slope of the Greater Caucasus, which we have chosen as the study area, is the richest region in Azerbaijan with river waters. The study assessed the annual distribution of the runoff of the major rivers of this region, especially their shares in the spring and summer months. The natural regulation state of the rivers, the volume of spring runoff, and the possibilities of their accumulation for using irrigation in the summer months were assessed.

Keywords: water resources, river runoff regulation, annual runoff distribution, intensive irrigation period, spring and summer runoff.

Introduction

The Republic of Azerbaijan is among the countries with limited freshwater resources. Water resources in Azerbaijan are very small compared to other countries of the South Caucasus and account for approximately 15 percent of the water resources in the entire region. 67-70% of river water resources in Azerbaijan are formed outside our country's borders. As a result of the impact of climate change and excessive water withdrawal for use in neighboring countries, the water resources decreased by 26.4% to 26.2 billion cubic meters during 1980-2023. In addition, during this period, surface water resources entering the country's territory through transboundary rivers also decreased by 35.9% to 13.2 billion cubic meters, and local river flow decreased by 23.3% to 4.6 billion cubic meters [1].

According to the data of the State Statistical Committee of the Republic of Azerbaijan, the volume of renewable surface fresh water resources currently amounts to 26.2 billion cubic meters. Of this, 8.4 km^3 is formed from groundwater, 4.6 km3 from local river runoff and 13.2 km^3 from transboundary surface waters. 17.8 km^3 (67.9 %) of the renewable fresh water resources are formed from surface fresh water resources. If in 2000 the surface fresh water resources were 23.4 km^3, in 2023 they decreased by 24.1% to 17.8 [2].

The results of water resource projections are also not encouraging. According to recent research by the World Resources Institute (WRI), Azerbaijan ranks 59th out of 164 countries that will face water scarcity by 2050 and is classified as a country with a medium-high (20-40% water stress) risk category [3].

According to Azerbaijan's National Report to the UN Framework Convention on Climate Change, the volume of water resources (including transboundary water resources) is projected to decrease by 5-10% by 2040, 10-15% by 2070, and 10-15 and 15-25%, respectively, depending on the level of global temperature increase.

One of the most important aspects that complicates the use of water resources is the significant increase in demand for water against the background of their decline. According to the State Statistical Committee, while the volume of fresh water withdrawn from natural sources increased by 27.9% in 2001-2023, the volume of water consumed increased by 52.4%, including for irrigation and agricultural supply, almost doubled. On the other hand, the uneven distribution of river water resources throughout the year and across the territory significantly complicates their use and cannot meet the water demand that is increasing year by year.

Annual water consumption in Azerbaijan is 11-12 km^3. At the same time, serious losses occur during water distribution due to outdated infrastructure. As a result of the constantly increasing water demand of the economy and the population, an average water deficit of 4.5-7.0 km^3 occurs in Azerbaijan per year. This situation requires regulating the runoff of rivers and constructing artificial reservoirs [4-5].

Currently, in the republic, more than 60 reservoirs, each with a volume of more than 1 million m^3, can store up to 21 billion m^3 of water. However, the depletion of river water resources also has its negative effects in this area. Thus, the decrease in river water consumption reduces the volume of water entering the reservoirs, weakening their use. For example, between 2016 and 2020, the average annual flow of water entering the Mingachevir reservoir, the largest in Azerbaijan alone, decreased from 344 m^3 per second to 192 m^3, a decrease of 44.2%.

Methods and materials

The primary research materials were the observation data on climate and runoff of the Ministry of Natural Resources and Ecology of the Republic of Azerbaijan. The collection of important factors involved in the flow formation, such as elevation, slope, landscape and other components, was obtained through satellite images. From the climatic parameters, atmospheric precipitation and air temperature were preferred, and from the runoff quantities, river water discharge data was preferred. Both annual and monthly data of the mentioned hydrometeorological indicators were used for the study area. The restoration of observation series was carried out using both traditional and modern methods. The most important of the traditional methods were the Analogy and Interpolation methods [6-7].

Here, the main focus was on establishing relationship graphs between the average elevation of the basin and hydrometeorological parameters. The results of the relationships established based on the data of the observation stations were extrapolated for river basins without data. The homogeneity of the observation series and the reliability of the results obtained were checked using the Fisher and Student criteria [8-10].

The results obtained during the research process were also refined using modern scientific methods. The results obtained using modern Interpolation and Counter-Approach technologies were evaluated separately for each river basin using the ArcGIS program [11-12].

Results and discussions

The southern slope rivers of the Greater Caucasus were taken as the study area (Figure 1). This region is the richest in Azerbaijan in terms of river water resources. The right tributaries of the Kura River, the largest river in Azerbaijan, flow from here. The largest rivers of the southern slope of the Greater Caucasus are the Eyrichay, Pirsaat, Turyan, Goychay, Alijan, Girdiman, and Agsu rivers. The rivers are mainly fed by groundwater, rain, and snow. Most of the rivers are abundant in their upper reaches, but they also have a high flood risk. The river basins of the study area can be divided into two parts according to their physical and geographical characteristics and feeding regime:

1) Basins between the Mazym and Dashagıl rivers. They are distinguished by their higher water discharge and the predominance of the groundwater share in the feeding regime.

2) Basins between the Dashagıl and Sumgayit rivers. The water discharge of the rivers is low, and the share of rainwater in the feeding regime is greater.

In order to obtain meteorological indicators (precipitation and temperature) for various hydrometric stations and river basins of the southern slope of the Greater Caucasus, actual measurement data from existing meteorological stations were used.

It was shown that there is a uniform and highly correlated close relationship (R=0.96) between the average multi-year air temperature and the altitudes of the stations for the entire study area (Figure 1).

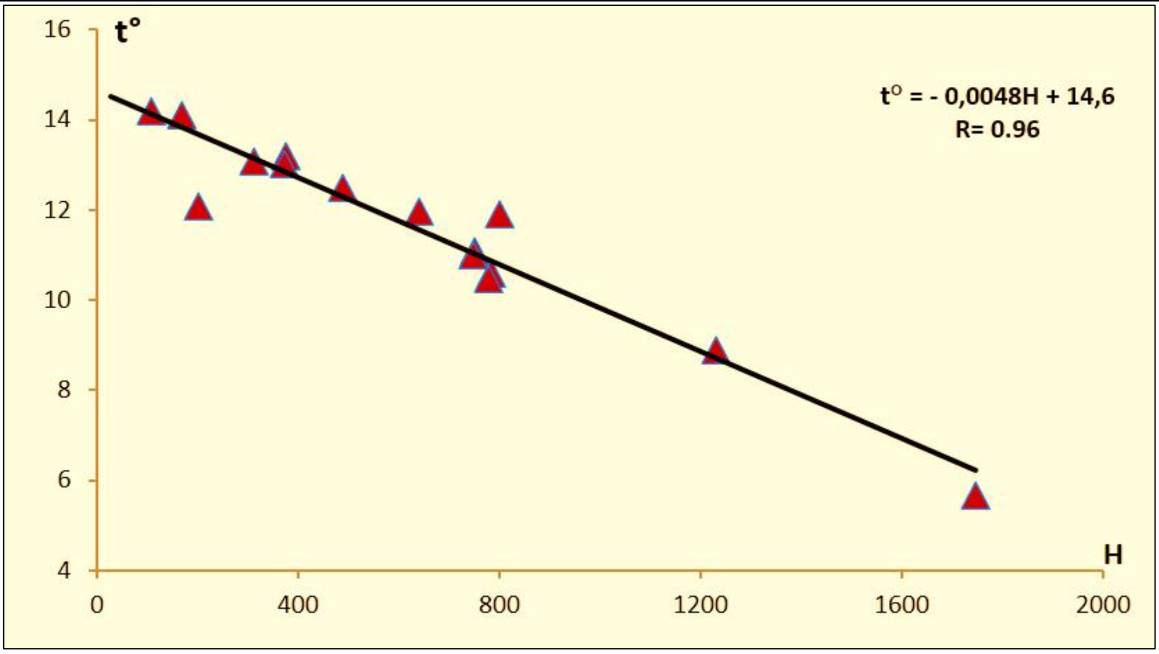


Figure 1. Relationship graph between mean multi-annual air temperature (t°) and mean altitude (H) on the southern slope of the Greater Caucasus

Unlike temperature, the relationship between atmospheric precipitation and altitude was established by dividing it into two parts due to its poor coverage for the entire study area. This division also covers two hydrologically distinct regions of the southern slope of the Greater Caucasus: within the Mazym-Dashagıl and Dashagıl-Sumgayit rivers basins (figure 2).

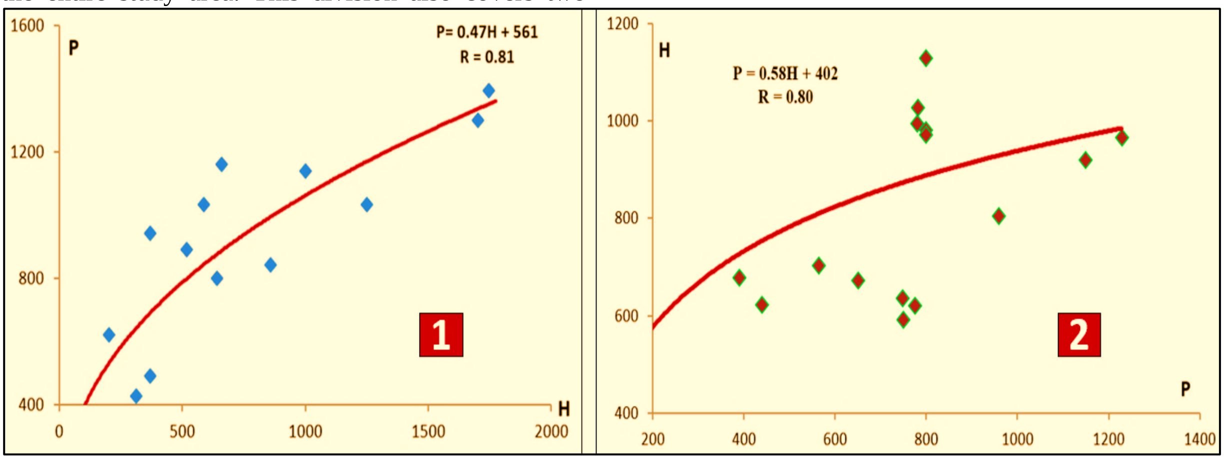


Figure 2. Relationship graphs between atmospheric precipitation (P) and mean elevation (H):
1– Between the Mazym and Dashagıl rivers. 2– Between the Dashagıl and Sumgayit rivers.

By selecting and interpolating the superior features in GIS comparison and calculation programs, average multi-year climate and flow data for spring and summer months were obtained for each river basin separately (table 1).

Table 1

Average multi-year spring and summer hydrometeorological data for river basins

Rivers	Temperature, Cº		Precipitation, mm		Runoff, mm	
	Spring	Summer	Spring	Summer	Spring	Summer
Mazym	9.4	21.2	227	219	120	119
Balaken	9.1	20.8	261	242	175	160
Katekh	8.6	20.3	276	301	169	189
Tala	9.5	21.3	247	252	137	138
Mukhakh	8.2	19.8	238	262	152	181
Kapychay	9.5	21.3	172	158	67.0	64.3
Kurmukh	8.3	19.9	260	288	166	211
Kashkachay	7.1	26.5	192	173	102	89.3
Shin	6.8	18.8	201	186	120	108
Kish	7.9	19.4	228	196	145	118
Dashagıl	8.0	19.5	237	261	167	181
Kungut	8.3	19.9	252	205	130	103
Zeyzit	7.3	18.7	258	210	158	136
Eyrichay	8.6	20.3	214	198	98.5	94.4
Alijan	10.9	23.1	188	128	48.1	25.7
Turyan	9.7	21.6	262	157	73.6	74.0
Goychay	11.1	23.3	243	129	71.2	59.2
Girdiman	7.7	19.2	275	218	141	108
Agsu	10.5	22.5	226	107	46.6	26.1
Pirsaat	9.4	21.2	127	70.5	41.0	27.5
Jeyrankechmez	12.1	24.4	96.1	29.6	7.33	2.71
Sumqayıt	9.8	21.8	104	44.9	15.4	7.32

Natural regulation of river runoff is a state of more even distribution of runoff throughout the year (months and seasons). In this case, the differences in river runoff between different months and seasons are very small. Most of the rivers of Azerbaijan have a high-water regime in the spring and autumn seasons, during which 60-70% (sometimes up to 80%) of the runoff passes. At the same time, a shortage of flow is observed in the rivers in the summer and winter seasons, when the combined share of the runoff of these two seasons in the annual runoff rate can rarely reach 50%. This situation indicates that the natural regulation of the runoff in the rivers of Azerbaijan is very low. Since most of the territory of Azerbaijan is located in a dry subtropical climate, sharp differences in the intra-annual runoff regime of the rivers also limit their use and serious problems arise when using water. The main way out of the situation is artificial regulation of the river runoff. Artificial regulation of rivers is the preservation of their resources during high water use periods (especially in the summer months) by maintaining them in various water levels until the stage of low water and high use.

Currently, water resources in Azerbaijan are mostly used for agricultural purposes in the summer (intensive irrigation period) [13]. However, the driest period of rivers is precisely the summer season. In the remaining seasons, there is no big problem in this regard. Despite the fact that the winter season is also dry, it is the season when river water is used the least. Therefore, there is no need to regulate the runoff of the preceding high-water autumn season with artificial interventions. Before the summer season, the regulation of the spring season flow is inevitable. In our research, we have mainly given extensive space to the study of the runoff of the spring and summer seasons.

Previously, a sufficient amount of research has been carried out on the study of the annual distribution of the rivers runoff of Republic of Azerbaijan [14-15].

However, our research was conducted in a completely different direction and purpose. In previous studies, the priority was mainly limited to providing quantitative indicators of the annual distribution of runoff based on hydrometeorological station data. We did not limit ourselves to the distribution ofrunoff by months and seasons. Research was conducted on the runoff regime of rivers, changes in the seasonal regime against the background of climate change, the level of natural regulation, the possibilities of spring runoff for conservation in watersheds, the assessment of spring runoff reserves for the purpose of intensive irrigation periods, and other new directions.

The most serious differences in the intra-annual distribution of the runoff are manifested in the ratio of spring and summer flows. In most rivers of the southern slope of the Greater Caucasus, spring is a high-water season, and summer is a low-water season. In such a case, the highest level of regulation (“very high”) was considered to be the summer runoff (Q_{sm}) higher than the spring runoff (Q_{sp}), i.e. $Q_{sm}/Q_{sp} > 1$. The values of this ratio between 0.90-1.00 were also accepted as a sufficient and “high” level of regulation. The situation where the summer runoff was less than half of the spring runoff ($Q_{sm}/Q_{sp} < 0.5$) was considered to be a very weak (“very low”) regulation. In the interval $Q_{sm}/Q_{sp} = 0.50\text{-}0.75$, i.e., if the summer runoff was from half to 3/4 of the spring runoff, this was considered weak (“low”). Finally, a “average” regulation within the range of $Q_{sm}/Q_{sp} = 0.75\text{-}0.90$ was adopted. The runoff of the rivers of the southern slope of the Greater Caucasus was mapped according to the level of natural regulation (Figure 3).

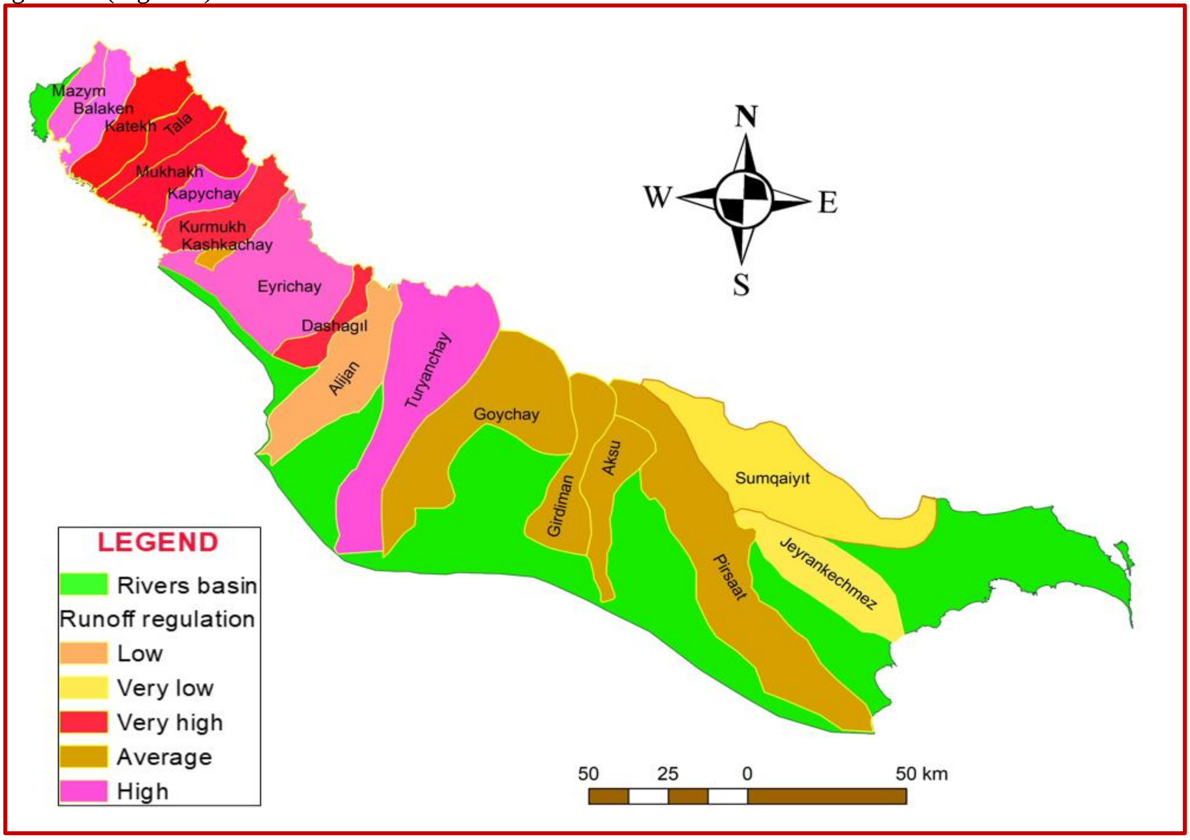


Figure 2. Natural regulation level of the rivers of the southern slope of the Greater Caucasus

The assessment of the natural regulation level of the rivers of the study area based on the ratio of summer and spring runoffs is given in the table below.

Table 2

Assessment of the natural regulation level based on the ratio of summer (Q_{sm}) and spring (Q_{sp}) runoffs

Rivers	Runoff layer, mm			Q_{sm}/Q_{sp}	Regulation level
	Perennial	Summer (Q_{sm})	Spring (Q_{sp})		
Mazym	381	119	120	0.992	High
Balaken	532	160	175	0.914	High
Katekh	548	189	169	1.118	Very high
Tala	429	138	137	1.007	Very high
Mukhakh	511	181	152	1.191	Very high
Kapychay	210	64.3	67.0	0.959	High
Kurmukh	561	201	166	1.211	Very high
Kashkachay	308	89.3	102	0.875	Average
Shin	366	108	120	0.900	High
Kish	426	118	145	0.814	Average
Dashagıl	577	181	167	1.084	Very high
Kungut	379	103	130	0.792	Average
Zeyzit	473	136	158	0.861	Average
Eyrichay	366	94.4	98.5	0.958	High
Alijan	136	25.7	48.1	0.534	Aşağı
Turyan	251	69.2	73.6	0.940	High
Goychay	213	59.2	71.2	0.831	Average
Girdiman	357	108	141	0.766	Average
Agsu	116	26.1	46.6	0.560	Average
Pirsaat	93.8	27.5	41.0	0.671	Average
Jeyrankechmez	15.4	2.71	7.33	0.369	Very low
Sumqayıt	32.5	7.32	15.4	0.475	Very low

As can be seen from Table 2, the natural regulation of rivers flowing from the western part of the study area (between the Mazym and Dashagil rivers) is better. The level of natural regulation in most rivers falls into the "high" and "very high" categories. The main reason for this is the high share of underground flow in the feeding of rivers. Due to the gradual and even distribution of the runoff in rivers fed by underground, they have better regulation, with high water discharge continuing until the summer months. Another advantage of rivers located between the Mazym and Dashagıl basins is that the average multi-year runoff share itself is also high. High annual runoff and partially good regulation allow for more rational use of the waters of the rivers located here.

The level of runoff regulation in rivers located in the eastern part of the study area (between the Dashagıl and Sumgayıt rivers) is low. This is explained by the high share of rainwater in the feeding of rivers. Both the low average annual runoff volume and the fact that the bulk of the runoff occurs during the spring rains have led to poor runoff regulation in most of the rivers located in the eastern part of this study area. Therefore, acute water shortages are observed during the period of intensive irrigation.

References:

1. Decree of the President of the Republic of Azerbaijan on approval of the "National Strategy for the Efficient Use of Water Resources". October 10, 2024.

2. The state statistical committee of the Republic of Azerbaijan. Water statistics in the Republic of Azerbaijan. 2014-2025.

3. World Resources Institute (WRI). Water resources reports.

4. Malik Garayev, İsmayil İsmayil, Elshan Hashimov. Wind, sun, and hydroenergy: a look ınto Azerbaıjan's green energy future. Collection of scientific papers ΛΌΓΟΣ. Seoul, South Korea. 2025. 8(1): 125-133.

5. Teymurov Movlud, Elshan G. Hashimov, İslam İslamov, Ramil Akhundov. Integrated assessment of water resources in Azerbaijan from quantitative and qualitative perspectives. Conference: Technologies, tools and strategies for implementing scientific research. 2026. Pages 109-120.

6. Tao Tao, Bernard Chocat. Uncertainty analysis of ınterpolation methods in rainfall spatial distribution (a case of small catchment in Lyon). Journal of Water Resource and Protection. 2009. Volume 1. No.2.

7. Vladimir Galakhov, Svetlana Samoilova, Olga Lovtskaya, Elena Mardasova. Basin-analogy for calculating spring flood volumes at insufficiency of meteorological data (Tuba river as a case study). Bulletin of the Tomsk Polytechnic University Geo Assets Engineering. 2024. 335(2):120-132.

8. İmanov F.A. Statistical methods in hydrometeorology. Baku. 2011. 272 pages.

9. Khalilov S.H., Safarov S.H. Monthly and annual norms of air temperature and atmospheric precipitation in the Republic of Azerbaijan (1691-1990), Baku, 2001, 229p.

10. Safarov S.H., Huseynov G.M., Safarov A.S. Non-stationary statistical model of air temperature variation with altitude for the Greater Caucasus and Nakhchivan Autonomous Republic. Baku, 2017, pp. 346-353.

11. Rza Makhmudov, Movlud Teymurov. Importance of using space information and GİS technologies in the process of selecting analogue terrains in hydrological research. Intercontinental Geoinformation Days. 2023. Volume 6. Pages 110-112.

12. Movlud A.Teymurov. (2022). The value of initial abstraction and estimation of run-off forming rainfall rate at a specific geographical location. Journal Science, Education and Innovation in the context of modern problems. Volume 5 (4), pp. 123-136.

13. Elnara Seyidova, Turkana Aliyeva, Latif Zeynalli. The impact of climate factors on agricultural development: ARDL bounds test approach. Agricultural and Resource Economics: International Scientific E-Journal. 2025. Volume 11 (4): 127-152.

14. M. Vinogradova, Z Altukhova, R Kashkay. Agrıcultural Sciences. 2017/11, Volume 11, Pages 11. Norwegian Journal of development of the International Science.

15. Кашкай Р.М., Теймуров М.А. Применение рационального метода при определении доли подземной составляющей водного баланса. Известия Наук о Земле НАНА. 2016. № 3-4, стр.83-89.

ECONOMIC SCIENCES

COMPARATİVE ANALYSİS OF WAGE INCOME TAXATİON İN TÜRKİYE AND AZERBAİJAN

Jafarova R.
https://orcid.org/0009-0008-3233-1222
https://doi.org/10.5281/zenodo.18840566

Abstract

This article provides a comprehensive comparative analysis of wage income taxation in the Republic of Azerbaijan and the Republic of Türkiye, focusing on the application of personal income tax. The objective of this study is to analyze the various characteristics and mechanisms involved in the taxation of wage income, with a particular emphasis on the methods used to calculate and apply tax rates in both countries. By comparing the tax systems in Azerbaijan and Türkiye, the article highlights the structural similarities and differences in wage tax frameworks, examining how these systems affect both individuals and the state. Additionally, the study explores the positive and negative aspects of the wage tax rates, considering their impact on social equity and economic efficiency. The research concludes with recommendations aimed at optimizing the existing tax structures, enhancing fairness, and reducing the tax burden for low and middle-income earners. The comparative methodology employed enables a deeper understanding of how tax policies can shape wage distribution, public finance, and overall economic sustainability in both countries.

Keywords: wage, Azerbaijan, Türkiye, income tax, tax rates, social equity, fiscal policy

Introduction

Tax is a compulsory and non-reciprocal monetary payment collected by the state or municipalities from taxpayers under legal coercion in order to finance public expenditures. Every individual inevitably encounters both indirect and direct taxes at some stage of life. One of the components of income subject to personal income tax among earnings and revenues is wages. A wage is defined as money, benefits in kind, and other advantages provided in return for labor.

When establishing tax systems, countries prioritize fairness and balance in tax distribution. However, wage income is not taxed at the same rate across all countries, and in certain cases, this may lead to inequality and an excessive tax burden.

1. Wage Income in Türkiye

In the Republic of Türkiye, wages constitute one of the elements of income subject to personal income tax. According to the Income Tax Law, wages are defined as monetary payments, benefits in kind (provision of goods as remuneration), and other advantages that can be expressed in monetary terms (such as housing or vehicle allocation), provided to employees who work under an employer and are attached to a specific workplace in return for services rendered.

For the purposes of implementing the Income Tax Law, the following payments are also considered wages:

- Pensions, disability, widow's and orphan's benefits paid by pension funds established by law and by the funds specified in Provisional Article 20 of Social Insurance Law No. 506, to the extent that they exceed the highest payment made to the highest-ranking civil servant.
- Monetary and in-kind payments and benefits provided in return for past or future services.
- Monetary and in-kind payments and benefits granted to members of the Grand National Assembly of Türkiye, Provincial General Assemblies, Municipal Councils, and members of permanent or temporary commissions established by special laws or administrative decisions, as well as to similar persons by virtue of their positions.
- Monetary and in-kind payments and benefits provided to chairpersons and members of Boards of Directors and Supervisory Boards, as well as to liquidators, by virtue of their positions.
- Payments and benefits provided to experts, official mediators, appraisers, sports referees, and members of competition juries.
- Transfer fees or similar payments and benefits made to athletes.
- Compensation and various payments made after the termination of an employment contract under mutual termination agreements or settlement agreements, including severance payments, job loss compensation, end-of-employment compensation, and employment security compensation.

1.1. Tax Rates

Tax rates are applied as indicated in the tariff schedule.

Income Tax Tariff for the 2025 calendar year (Article 103 of the Income Tax Law).

Up to **TRY 158,000**	15%
For income exceeding **TRY 158,000** up to **TRY 330,000**: TRY 23,700 for the first TRY 158,000, plus the applicable rate on the excess.	20%
For income exceeding **TRY 330,000** up to **TRY 800,000**: TRY 58,100 for the first TRY 330,000 (**For wage income:** TRY 58,100 for the first TRY 330,000 up to TRY 1,200,000), plus the applicable rate on the excess.	27%
For income exceeding **TRY 800,000** up to **TRY 4,300,000**: TRY 185,000 for the first TRY 800,000 (**For wage income:** TRY 293,000 for the first TRY 1,200,000), plus the applicable rate on the excess.	35%
For income exceeding **TRY 4,300,000**: TRY 1,410,000 for the first TRY 4,300,000 (**For wage income:** TRY 1,378,000 for the first TRY 4,300,000), plus the applicable rate on the excess.	%40

In 2006, a new system was introduced in Türkiye, and the number of tax brackets was reduced from five to four. The tax rates ranged from 15% to 35%.

1.2. Elements of Wage Income

In order for income to be taxed as wage income, it must possess certain characteristics:

1. **Subordination to an employer.**

One of the primary elements for income to qualify as wage income is subordination to an employer. An employer is a person who employs an employee (service provider) under their orders and instructions.

2. **Attachment to a workplace.**

Another essential element is being affiliated with a specific workplace, defined as the place where activities are carried out.

3. **Payment in return for services rendered.**

The final element is that payment must be made in consideration of services rendered. If no service is provided, the payment cannot be classified as a wage.

1.3. Gross and Net Wages

Wages are applied in two forms: gross and net wages.

Gross wage refers to the total income before any taxes or deductions are made. For example, if an employer pays TRY 7,000 annually, the gross salary amounts to TRY 7,000.

Calculation of Gross Wage

The calculation of gross wages depends on how the employee is compensated. For salaried employees, the gross wage is calculated by dividing the annual salary by the number of payment periods within a year.

Net wage refers to the amount paid after taxes and statutory deductions are made. When calculating the net wage, deductions such as income tax, stamp tax, employee social security contributions, and employee unemployment insurance contributions are subtracted from the gross wage.

1.4. Technical Characteristics of the Progressive Tariff

Progressive tax tariffs possess certain technical characteristics, such as bracket width differences, rate differences, and progression ratios.

The **bracket width difference** indicates the difference between the highest and lowest tax bases. The **rate difference** refers to the difference between the highest and lowest tax rates. For example, if the highest tax rate on wages is 40% and the lowest is 15%, the rate difference equals 25 percentage points.

Thus, since Türkiye applies a progressive and graduated system, it is appropriate to analyze the technical characteristics of progressive taxation within this framework.

2. Wage Income in Azerbaijan

This article briefly addresses the fundamental aspects of payroll calculation in Azerbaijan. A significant portion of the population in Azerbaijan earns their livelihood through wages.

Payroll is generally defined as the list of employees and the payments that the employer is required to make to them due to the employment relationship. Currently, payroll primarily refers to the payment of salaries to employees, as well as the withholding and remittance of income tax and mandatory social security contributions derived from wages and other employment-related income.

According to the Tax Code, the income tax rates applied to individual wages are determined as follows:

- In the oil and gas sector and the public sector, income up to AZN 2,500 per month is taxed at 14% of the taxable monthly amount; income exceeding AZN 2,500 is taxed at 25%.
- For individuals working outside the oil and gas sector and in the private sector, effective from 1 January 2019, monthly income up to AZN 8,000 is taxed at 0%, while the portion exceeding AZN 8,000 is taxed at 14%.

In addition to the 25% social insurance contribution, unemployment insurance contributions must also be paid at the rate of 0.5% by the employee and 0.5% by the employer from its own funds.

As of 1 January 2020, mandatory health insurance contributions are paid as follows: in the private oil and gas sector, 2% (1% employee and 1% employer) on salaries up to AZN 8,000, and 0.5% on the amount exceeding this threshold. In the public sector and the oil and gas sector, the rate is 2%.

Wage-related taxes and contributions in Azerbaijan can be summarized as follows:

- **Income tax:**
  - 0% (private non-oil and gas sector up to AZN 8,000)
  - 14% (private non-oil and gas sector exceeding AZN 8,000)
  - 14–25% (public service and oil and gas sector)
- **Mandatory social insurance contribution:** 25%

- **Mandatory unemployment insurance:** 1%
- **Mandatory health insurance:** 2% (1% employee, 1% employer for private sector income up to AZN 8,000; 0.5% for additional income), and 2% in the public or oil and gas sector.

Mandatory health insurance payments are made throughout the year in accordance with the established schedule.

Under Azerbaijan's wage taxation system, in addition to income tax (wage tax), a mandatory social insurance contribution of 25% is payable. For wages up to AZN 200, 3% of the social insurance contribution is deducted from the employee's salary and 22% is paid by the employer. If the monthly wage exceeds AZN 200, 10% is deducted from the employee and 15% is paid additionally by the employer.

Unlike income tax, foreign employees working under production-sharing agreements in the oil and gas sector are exempt from social insurance contributions. In other words, such foreign employees are subject only to income tax and are not required to make social contributions.

Update: As of 1 January 2025, the minimum wage in Azerbaijan is AZN 400.

"Gross" salary refers to the amount paid before tax deductions. "Net" salary refers to the amount paid to the employee after tax deductions. (Articles 102.1-1, 102.2, 102.3, 102.4, and 102.5 of the Tax Code.)

4. Comparison

4.1. Comparison in Terms of Taxation

In Türkiye, a progressive and graduated taxation system is applied. Tax rates increase from 15% up to 40%. At the same time, this structure may result in a relatively heavy tax burden for taxpayers.

In Azerbaijan, the system appears comparatively simpler, with income divided into two main sectors. As stated above, tax rates in the private sector range from 0% to 14%, while in the public and oil and gas sectors they range from 14% to 25%. Unlike Türkiye's multi-bracket progressive system, Azerbaijan applies a simpler form of progressive taxation.

4.2. Comparison of Wage Income Tax Rates

Country	**Country**	**Country**
Türkiye	15% – 40%	Progressive, graduated
Azerbaijan	Public and oil–gas sector: 14% and 25%; Private sector: 0% and 14%	Simplified (quasi-progressive)

4.3. Comparison in Terms of Exemptions

In Türkiye, a minimum wage exemption is applied, along with various additional exemptions and allowances.

In Azerbaijan, compared to Türkiye, a broader range of exemptions and tax privileges is implemented.

4.4. Comparison in Terms of Deductions from Gross Wages

In Türkiye, deductions such as unemployment insurance contributions and social security premiums are applied to gross wages.

In Azerbaijan, social insurance contributions are calculated based on income levels. Furthermore, the distribution of contributions varies between the employer and the employee. Employees working in the oil and gas sector are exempt from certain types of such deductions.

5. Conclusion

This study examined and compared the taxation of wage income in Türkiye and Azerbaijan. The findings reveal structural differences between the two countries' wage taxation systems.

In general, it has been determined that Türkiye applies a progressive and graduated taxation system with a relatively heavier tax burden. In contrast, Azerbaijan operates a comparatively simpler system with a lighter overall tax burden. Structurally, Türkiye demonstrates a "strong progressive rate" system, whereas Azerbaijan exhibits a "weak progressive rate" structure.

In Türkiye, simplifying the tax system, ensuring a more balanced distribution of tax rates, and expanding exemptions and allowances could help alleviate the tax burden and contribute positively to overall welfare. In Azerbaijan, although the system is relatively simple, ensuring tax equity may present challenges; therefore, regulatory improvements in this regard would be particularly important.

Below is the academic English translation of the **References** section. Original titles are translated into English, while preserving citation structure and academic formatting:

References:

1. Aghayev, O. I. *Personal Income Tax in Azerbaijan and Global Practices.*
2. Revenue Administration Presidency (2017). *Tax Guide for Individuals Earning Wage Income.* Publication of the Department of Taxpayer Services, No. 246, pp. 1–56.
3. Bay, H., & İnneci, A. (2017). Taxation of Wage Income in Türkiye in Terms of Tax Justice. *International Journal of Public Finance*, 2(2), 196–219.
4. Ergür, O. K., & Köstekçi, A. (2021). Tax Justice in the Taxation of Wage Income in Türkiye: The Case of Athletes' Taxation. *Journal of Human and Social Sciences Research*, 10(3), 2936–2951. https://doi.org/10.15869/itobiad.907247
5. Ertekin, Ş. (2021). Evaluation of the Fiscal Burden on Wages in Türkiye in Terms of the Principle of Justice. *Cumhuriyet University Journal of Economics and Administrative Sciences*, 22(1), 281–308.
6. İsa, Ş. H. S., & Şen, H. (2016). *Tax Theory and Policy.*
7. Kahraman, A. (2003). Wages in General and the Accrual of Wages.
8. Kılıç, Ö. (2018). Tax Justice in Türkiye. *Akademik Bakış: International Refereed Journal of Social Sciences*, (66), 396–411.
9. Öztürk, İ., & Ozansoy, A. (2011). Analysis of the Tax Burden on Wage Income. *Maliye Journal*, 161(2), 197–214.

10. Pashayeva, D. (2019). Directions for improving the tax administration and tax control under modern conditions in the Republic of Azerbaijan. Izvestiya. Journal of Varna University of Economics, 63(4), 344-359.

11. Şafiyeva, V. (2018). The Relationship Between Income Tax Applied in Azerbaijan and Tax Revenues, GDP, and Budget Revenues: Effects on Business Decisions. *TURAN-SAM*, 10(40), 781–787.

MATHEMATICAL SCIENCES

A COURSE OF ALGEBRA (PART II)

Frasser C.
Ph.D. in Engineering Sciences, Odessa National Polytechnic University, Ukraine
https://doi.org/10.5281/zenodo.18840568

Abstract

The exploration of algebraic operations finds a natural impetus in the challenge of solving polynomial equations of degree $n \geq 5$. Galois theory provides the conceptual framework for this inquiry, revealing that certain quintic equations with rational coefficients admit no general solution in radicals. This fundamental result, both accessible in statement and deep in implication, justifies a systematic development of the algebraic tools required for its understanding. Part II develops the language and techniques – particularly those concerning polynomials and matrices – necessary to formulate and approach such problems rigorously. It is worthy to mention that this final part is presented from a history-of-science perspective and explores modern abstract algebra with clarity and simplicity, while retaining an advanced outlook. Readers are encouraged to review Part I carefully before proceeding.

Keywords: Groups, Rings, Fields, Polynomials, Roots of Polynomials, Matrices.

POLYNOMIALS

The concepts of polynomials and their operations, as taught in post-secondary education, require precise definition and rigorous justification. For example, it is commonly stated that since $x = 1$ is a root of the equation $x^3 - x^2 + x - 1 = 0$, the polynomial on the left-hand side can be divided by the binomial $x - 1$ without remainder. However, if x is interpreted strictly as an unknown satisfying the equation, this situation effectively represents a disguised form of division by zero, rendering the meaning of the operation ambiguous – despite the fact that the formal procedure of polynomial division proceeds without difficulty. Furthermore, in certain branches of mathematics, such as combinatorial analysis, one frequently encounters "polynomials" of the form

$$a_0 + a_1 x + a_2 x^2 + \cdots$$

with infinitely many nonzero coefficients. These objects, known as *formal power series*, are considered independently of the numerical values of x for which the partial sums

$$S_n(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

may converge as $n \to \infty$. Unlike analysis, where convergence is central, algebra focuses on the algebraic rules governing the formation of coefficients under polynomial operations. By introducing polynomials in a rigorous manner – including those with infinitely many terms – we can define addition and multiplication so that these operations align consistently with the rules of the standard "level of post-secondary education."

Let an arbitrary ring $K = (M, +, \cdot)$ be given. Consider the set M^∞, whose elements are all infinite sequences of elements from M. If $a, b \in M^\infty$, where

$$a = (a_0, a_1, a_2, \dots),\ b = (b_0, b_1, b_2, \dots),$$

then we define $a = b$ if and only if $a_i = b_i$ for every $i = 0, 1, 2, \dots$. Addition and multiplication in M^∞ are defined as follows:

$$a + b = (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots),$$
$$a \cdot b = (a_0 b_0,\ a_0 b_1 + a_1 b_0, a_0 b_2 + a_1 b_1 + a_2 b_0, \dots);$$

In general, the *j*-th term of the product $a \cdot b$ has the form

$$a_0 b_j + a_1 b_{j-1} + \cdots + a_j b_0 = \sum_{k=0}^{j} a_k b_{j-k} = \sum_{\substack{k+i=j \\ k,i \geq 0}} a_k b_i.$$

(In the last expression, it is tacitly assumed that the indices k and i take only those values satisfying the condition $k + i = j$ for a given j, where both k and i are nonnegative integers). Finally, note that it is convenient to refer to the *j*-th term of a sequence as what would ordinarily be the $(j + 1)$-th term in standard counting since here we begin indexing from zero rather than from one.

We now prove that the *system* $K^\infty = (M^\infty, +, \cdot)$ *is a ring*.

I. $(M^\infty, +)$ is an *abelian group*.

If $a, b \in M^\infty$, then the *j*-th term of $a + b$ is $(a + b)_j = a_j + b_j$ $(j = 0, 1, 2, \dots)$. Thus, all the axioms of an abelian group for sequences follow directly from the fact that $(M, +)$ is an abelian group since each coordinate operation is inherited termwise.

- The zero element of $(M^\infty, +)$ is the sequence $(0, 0, 0, \dots)$, where 0 is the additive identity of K.
- The additive inverse of $a = (a_0, a_1, a_2, \dots)$ is $-a = (-a_0, -a_1, -a_2, \dots)$.

Therefore, $(M^\infty, +)$ is an abelian group.

II. $(M^\infty, \cdot)$ is a *groupoid*.

If $a, b \in M^{\infty}$, then the j-th term of the product sequence ab is given by

$$(ab)_j = \sum_{k+i=j} a_k b_i,$$

which is an element of M because $K = (M, +, \cdot)$ is a ring. Hence $ab \in M^{\infty}$. Therefore, multiplication is closed in M^{∞}, and $(M^{\infty}, \cdot)$ is a groupoid.

III. *Distributive properties*.

Let $a = (a_0, a_1, a_2, \dots)$, $b = (b_0, b_1, b_2, \dots)$, $c = (c_0, c_1, c_2, \dots) \in M^{\infty}$. The j-th term of the sequence $(a+b)c$ is

$$\sum_{k+i=j} (a_k + b_k) c_i = \sum_{k+i=j} (a_k c_i + b_k c_i) = \sum_{k+i=j} a_k c_i + \sum_{k+i=j} b_k c_i .$$

This shows that the j-th term of $(a+b)c$ is equal to the sum of the corresponding j-th terms of ac and bc. Thus, $(a+b)c = ac + bc$. The proof of the other distributive property, $c(a+b) == ca + cb$, is analogous.

Theorem 15. *If a ring K has any of the properties*:

1. *commutativity,*
2. *associativity,*
3. *existence of a unit,*
4. *absence of zero divisors,*

then the ring of sequences K^{∞} also has the same properties.

Proof. 1. *Commutativity*.

Suppose K is commutative, i.e., $xy = yx$ for all $x, y \in M$. For any $j \geq 0$, the j-th term of the product sequence ab is

$$\sum_{k+i=j} a_k b_i = \sum_{i+k=j} b_i a_k,$$

which is exactly the j-th term of the sequence ba. Hence $ab = ba$ in K^{∞}.

2. *Associativity*.

Suppose K is associative. For any $j \geq 0$, the j-th term of the sequence $a(bc)$ is

$$\sum_{k+i=j} a_k \left(\sum_{p+q=i} b_p c_q \right) = \sum_{k+i=j} \sum_{p+q=i} a_k (b_p c_q) = \sum_{k+p+q=j} a_k (b_p c_q).$$

Similarly, the j-th term of the sequence $(ab)c$ is

$$\sum_{i+q=j} \left(\sum_{k+p=i} a_k b_p \right) c_q = \sum_{i+q=j} \sum_{k+p=i} (a_k b_p) c_q = \sum_{k+p+q=j} (a_k b_p) c_q.$$

Since multiplication in K is associative, we have $a_k (b_p c_q) = (a_k b_p) c_q$ for all k, p, q and therefore $a(bc) = (ab)c$ in K^{∞}.

3. *Existence of a unit*.

If 1 is the multiplicative identity in K, then for any sequence $(a_0, a_1, a_2, \dots) \in K^{\infty}$, we have

$$(1, 0, 0, \dots) \cdot (a_0, a_1, a_2, \dots) = (a_0, a_1, a_2, \dots) \cdot (1, 0, 0, \dots) = (a_0, a_1, a_2, \dots).$$

Hence the element $(1, 0, 0, \dots)$ serves as the multiplicative identity of the ring K^{∞}.

4. *Absence of zero divisors*.

Assume K has no zero divisors, and consider two arbitrary nonzero elements in K^{∞}:

$$a = (0, 0, \dots, 0, a_m, a_{m+1}, \dots) \text{ and } b = (0, 0, \dots, 0, b_n, b_{n+1}, \dots),$$

where $a_m \neq 0$ and $b_n \neq 0$ $(m, n \geq 0)$ are, respectively, the first nonzero terms of a and b. Then the $(m+n)$-th term of the product ab is

$$a_0 b_{m+n} + a_1 b_{m+n-1} + \dots + a_{m-1} b_{n+1} + a_m b_n + a_{m+1} b_{n-1} + \dots + a_{m+n} b_0.$$

Since $a_i = 0$ for $i < m$ and $b_k = 0$ for $k < n$, all terms vanish except $a_m b_n$. Thus,

$$(ab)_{m+n} = a_m b_n.$$

Because $a_m \neq 0$, $b_n \neq 0$, and K has no zero divisors, it follows that $a_m b_n \neq 0$. Therefore, $ab \neq (0, 0, 0, \dots)$, which shows that K^{∞} also has no zero divisors. •

Corollary 2. *If K is an integral ring, then so is K^{∞}.* •

However, if K is a field, then the ring K^{∞} (which is obviously nontrivial) is not a field. Indeed, the nonzero element $(0, 1, 0, 0, \dots)$ is not invertible: if there were some $b = (b_0, b_1, b_2, \dots)$ such that

$$(0, 1, 0, 0, \dots) \cdot b = (1, 0, 0, \dots),$$

it would follow that $0 \cdot b_0 = 1$, i.e., $0 = 1$, a contradiction.

Now let K be an associative ring with unity. Then so is K^{∞}. It is straightforward to show that the set of elements of the form $(a_0, 0, 0, \dots)$ constitutes a subring of K^{∞}, which is isomorphic to K via the natural isomorphism

$$f: K \longrightarrow K^{\infty}, f(a_0) = (a_0, 0, 0, \dots).$$

Accordingly, such an element can be denoted simply by a_0. In particular, $0 = (0, 0, 0, \dots)$ and $1 == (1, 0, 0, \dots)$ are, respectively, the zero and unity of the ring K^{∞}.

For the element $(0, 1, 0, 0, \dots)$, we introduce the notation x. Then $(0, 0, 1, 0, \dots) = x \cdot x = x^2$, the element $(0, 0, 0, 1, 0 \dots) = x \cdot x^2 = x^2 \cdot x$ can be denoted by x^3, and so on. Assuming, by definition, that $x^0 = 1$, it is straightforward to prove that $x^m \cdot x^n = x^{m+n}$ for all integers $m, n \geq 0$. The element $(0, \dots, 0, a_m, 0, \dots)$ can be written in the form

$$(a_m, 0, 0, \dots) \cdot (0, \dots, 0, 1, 0, \dots) = a_m x^m,$$

or equivalently, $(0, \dots, 0, 1, 0, \dots) \cdot (a_m, 0, 0, \dots) = x^m a_m$. Hence, *every nonnegative integer power x^m commutes with all elements of the ring K^∞*, even though K^∞ itself need not be commutative.

Now consider an element $a = (a_0, a_1, a_2, \dots, a_m, 0, 0, \dots) \in M^\infty$, where $a_{m+1} = a_{m+2} = \cdots == 0$. It can be uniquely represented as

$$a = (a_0, 0, 0, \dots, 0, 0, \dots) + (0, a_1, 0, \dots, 0, 0, \dots) + (0, 0, a_2, \dots, 0, 0, \dots) + \cdots + (0, 0, 0, \dots, a_m, 0, \dots) =$$
$$= (a_0, 0, 0, \dots) + (a_1, 0, 0, \dots) \cdot (0, 1, 0, 0, \dots) + (a_2, 0, 0, \dots) \cdot (0, 0, 1, \dots) + \cdots +$$
$$+ (a_m, 0, 0, \dots) \cdot (0, 0, 0, \dots, 1, 0, \dots) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m.$$

That is, a is represented using the usual notation of a polynomial in terms of x. (One must prove the uniqueness of this representation.) Thus, every sequence in K^∞ with only finitely many nonzero terms corresponds to a polynomial, or more precisely – as is customary in higher algebra–to a *polynomial in a formal variable x over the ring K*. The word "formal" emphasizes that x is not just an element of K (or a "variable ranging over K"); rather, it is a well-defined, fixed element of the ring K^∞. The phrase "over the ring K" indicates that all coefficients lie in K. The subset of K^∞ consisting of such polynomials is denoted by $K[x]$.

Theorem 16. *The set $K[x]$ forms a subring of K^∞.*

Proof. (*Closure under addition*) Let

$$a = (a_0, a_1, \dots, a_m, 0, 0, \dots),\ b = (b_0, b_1, \dots, b_n, 0, 0, \dots) \in K[x].$$

Then their sum is $a + b = \left(a_0 + b_0, a_1 + b_1, \dots, a_{\max\{m,n\}} + b_{\max\{m,n\}}, 0, 0, \dots\right)$. Since all terms beyond index $\max\{m, n\}$ vanish, we conclude that $a + b \in K[x]$. Thus $(K[x], +)$ is a groupoid.

(*Closure under multiplication*) Consider the product ab. Since a has degree at most m and b has degree at most n, the degree of ab is at most $m + n$. Hence all terms beyond index $\max\{m, n\}$ vanish, and therefore $ab \in K[x]$. Thus $(K[x], \cdot)$is a groupoid.

(*Additive inverses*) For

$$a = (a_0, a_1, \dots, a_m, 0, 0, \dots) \in K[x],$$

the additive inverse is $-a = (-a_0, -a_1, \dots, -a_m, 0, 0, \dots) \in K[x]$.

Therefore, $K[x]$ is closed under addition, multiplication, and contains additive inverses. Since it is already a subset of the ring K^∞, we conclude that $K[x]$ is a subring of K^∞. •

Remark 17. As we already know, these three properties are sufficient for $K[x]$ to be a subring of K^∞. On this basis, $K[x]$ is called the *ring of polynomials in x over the ring K*. Elements of $K[x]$ will be written in polynomial form rather than as sequences.

Indeed, the coefficients of the powers $x^0 = 1, x^1 = x, x^2, \dots, x^m$ in the polynomial

$$A(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m$$

correspond to the first $m + 1$ terms of the sequence

$$(a_0, a_1, \dots, a_m, 0, 0, \dots),$$

which represents the same element (but in different notation) of the subring $K[x]$. Since two sequences in K^∞ are equal if and only if their corresponding terms are equal, it follows that *two polynomials are equal – that is, they represent the same element of $K[x]$ – if and only if their coefficients at each power of the formal variable coincide*.

Polynomials from $K[x]$ will be denoted by $A(x), B(x), C(x), \dots$, or simply $A, B, C, \dots$ in capital letters, so as not to confuse them with elements of the ring K.

If $A = A(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \in K[x]$, $a_m \neq 0, m \geq 0$, then the number m is called the *degree* of the polynomial A, denoted $\deg A$. In particular:

- If $\deg A = 0$, then $A = a_0$ contains no variable x, and thus corresponds to a nonzero element of K.
- For $a_0 = 0$, the zero polynomial $A = 0$ has no defined degree; nevertheless, it is convenient to assign it a negative degree (without specifying which).

Theorem 17. *Let A, $B \in K[x]$. Then*:

1. $\deg(A + B) \leq \max\ \{\deg A, \deg B\}$;
2. $\deg(A \cdot B) \leq \deg A + \deg B$.

Moreover, if the leading coefficients $a_m \neq 0$ and $b_n \neq 0$ of the polynomials A and B are not zero divisors in K, then equality holds in both cases. •

Remark 18. The first statement is straightforward. Note, however, that the degree of the sum of two polynomials may be strictly less than the degree of each individual polynomial, even when $\deg A = \deg B$. This occurs precisely when the leading coefficients (i.e., the coefficients corresponding to the highest power of x) of A and B are additive inverses in the ring K. To prove the second statement, let

$$A = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m, \text{ where } a_m \neq 0,\ m = \deg A,$$
$$B = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n, \text{ where } b_n \neq 0,\ n = \deg B.$$

then the highest possible power of x in the product $A \cdot B$ is $m + n$ and the leading term is $a_m b_n x^{m+n}$. This term is not zero whenever a_m and b_n are not zero divisors in K.

The inequalities (1) and (2) of Theorem 17 also remain valid when $A = 0$ or $B = 0$, due to the convention that $\deg(0) < 0$.

DIVISION WITH REMAINDER IN $K[x]$. We continue to assume that the ring K is associative with unity 1; hence, the same is true for $K[x]$. From now on, it will be more convenient to write polynomials in $K[x]$ in *descending* order of powers of x, rather than ascending order.

Theorem 18. *Let*

$$A = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0,\ B = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$$

be two polynomials in $K[x]$, where $B \neq 0$ and the leading coefficient b_n has an inverse element b_n^{-1} in the ring K. Then:

1. *There exist unique polynomials $Q_{left}, R_{left} \in K[x]$ such that*
$$A = Q_{left} \cdot B + R_{left},\ \deg R_{left} < \deg B.$$
2. *There exist unique polynomials $Q_{right}, R_{right} \in K[x]$ such that*
$$A = B \cdot Q_{right} + R_{right},\ \deg R_{right} < \deg B.$$

Proof. The polynomials Q_{left} and R_{left} are called the *left quotient* and the *left remainder* when dividing A (*the dividend*) by B (*the divisor*). Similarly, Q_{right} and R_{right} are called the *right quotient* and the *right remainder*. We establish their existence and uniqueness separately.

(*Left division*) The existence of such polynomials is established by following the usual scheme of polynomial long division, but with careful attention to the difference between left and right division in the case when the coefficient ring K is noncommutative. We begin with the left division.

$$\begin{array}{l|l} \ \ a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 & b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \\ \cline{2-2} - & \\ \ \ a_m x^m + a_m b_n^{-1} b_{n-1} x^{m-1} + \cdots & a_m b_n^{-1} x^{m-n} + (a_{m-1} - a_m b_n^{-1} b_{n-1}) b_n^{-1} x^{m-n-1} + \cdots \\ \cline{1-1} \qquad (a_{m-1} - a_m b_n^{-1} b_{n-1}) x^{m-1} + \cdots & \end{array}$$

If $\deg A < \deg B$, then we immediately set $Q_{left} = 0$, $R_{left} = A$. If $\deg A \geq \deg B$, then we begin by choosing the leading term of the left quotient Q_{left} as $a_m b_n^{-1} x^{m-n}$, so that when multiplied on the left by B, the leading term matches that of A:

$$a_m b_n^{-1} x^{m-n} \cdot (b_n x^n + \cdots) = a_m b_n^{-1} x^{m-n} \cdot b_n x^n + \cdots = a_m b_n^{-1} (x^{m-n} b_n) x^n + \cdots$$
$$= a_m b_n^{-1} (b_n x^{m-n}) x^n + \cdots = a_m (b_n^{-1} b_n)(x^{m-n}) x^n + \cdots = a_m \cdot 1 \cdot x^{m-n+n} = a_m x^m + \cdots$$

Subtracting this from A, we obtain

$$A - (a_m x^m + a_m b_n^{-1} b_{n-1} x^{m-1} + \cdots) = (a_{m-1} - a_m b_n^{-1} b_{n-1}) x^{m-1} + \cdots,$$

which is a polynomial of degree strictly less than m (equal to $m - 1$ if $a_{m-1} - a_m b_n^{-1} b_{n-1} \neq 0$, otherwise even smaller).

If the degree of this remainder is already less than $\deg B$, then we take it as R_{left} and set

$$Q_{left} = a_m b_n^{-1} x^{m-n}.$$

If instead the degree of the remainder is still at least $\deg B$, we proceed to determine the next term of the quotient. This second term (of degree less than $m - n$) is chosen so that, after multiplying the sum of the two first quotient terms by B, the resulting product agrees with A in both the leading term $a_m x^m$ and the next term $a_{m-1} x^{m-1}$.

This process is then continued inductively. At each step, the degree of the current remainder strictly decreases, ensuring that after finitely many steps we obtain a remainder of degree less than $\deg B$. We take this polynomial as R_{left}, while Q_{left} is defined as the sum of all the monomials chosen in the successive steps. Thus,

$$A = Q_{left} \cdot B + R_{left},\ \deg R_{left} < \deg B.$$

This description also covers the case of exact division, where $R_{left} = 0$. In this case we note that $\deg(0) = -\infty < \deg B$, and since $b_n \neq 0$ (due to its invertibility), we have $\deg B \geq 0$.

(*Right division*) The existence of Q_{right} and R_{right} is proved in a similar way (they do not necessarily coincide with Q_{left} and R_{left}), except that the quotient terms are chosen to act on the right of B. Equivalently, we select the quotient coefficients as indicated and multiply them on the left of B.

$$\begin{array}{l|l} \ \ a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 & b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0 \\ \cline{2-2} - & \\ \ \ a_m x^m + b_{n-1} b_n^{-1} a_m x^{m-1} + \cdots & b_n^{-1} a_m x^{m-n} + b_n^{-1}(a_{m-1} - b_{n-1} b_n^{-1} a_m) x^{m-n-1} + \cdots \\ \cline{1-1} \qquad (a_{m-1} - b_{n-1} b_n^{-1} a_m) x^{m-1} + \cdots & \end{array}$$

The uniqueness of the quotient and the remainder in both cases is proved in a completely similar way and we now consider the left division (omitting the index “left” for simplicity). Let two representations be obtained as follows:

$$A' = Q \cdot B + R = \tilde{Q} \cdot B + \tilde{R},\ \deg R < \deg B,\ \deg \tilde{R} < \deg B;$$

then $(Q-\tilde{Q})\cdot B=\tilde{R}-R$. For $Q-\tilde{Q}\neq 0$, we would have $\deg(Q-\tilde{Q})\geq 0$ and since the polynomial B has leading coefficient b_n, which is not a zero divisor in the ring K because it is invertible, $\deg\left((Q-\tilde{Q})\cdot B\right)=\deg(Q-\tilde{Q})+\deg B\geq \deg B$ while $\deg(\tilde{R}-R)\leq$
$\leq \max\{\deg\tilde{R},\deg(-R)\}<\deg B$, which is not possible. Therefore, $\tilde{Q}=Q$. It follows that $\tilde{R}-R=0\cdot B=0$, i.e., $\tilde{R}=R$. This completes the proof. •

If the ring K (and therefore $K[x]$) is commutative, then $Q_{left}=Q_{right}$ and $R_{left}=R_{right}$. In this case, the quotient and the remainder are simply denoted by Q and R. The following is a restatement of Theorem 18.

Theorem 18. (*Restated*) *For any polynomials $A,B\in K[x]$, where the leading coefficient b_n of B is invertible, there exist unique polynomials $Q,R\in K[x]$ such that*

$$A=QB+R=BQ+R,\ \deg R<\deg B.$$

If $R=0$, then B is called a *divisor* of A. In this case, we also say that B *divides* A or that A *is divisible by* B.

DIVISIBILITY OF POLYNOMIALS OVER A FIELD

Let’s study in detail the divisibility properties of polynomials under the assumption that the base ring K is a field P. Since every field is an integral ring, the ring $P[x]$ is also integral. For any $A,B\in P[x]$, the following properties hold:

$$\deg(A\pm B)\leq \max\{\deg A,\deg B\},$$
$$\deg(A\cdot B)=\deg A+\deg B.$$

Lemma 1. *Let $A,B,C,A_1,A_2,\dots,A_k,C_1,C_2,\dots,C_k\in P[x]$. Then*:

1. *If A is divisible by B, and B is divisible by C, then A is divisible by C.*
2. *If $A_1,A_2,\dots,A_k$ are each divisible by B, and $C_1,C_2,\dots,C_k$ are arbitrary, then*

$$A_1C_1+A_2C_2+\cdots+A_kC_k$$

is divisible by B. In particular, $A_1\pm A_2$ and A_1C_1 are divisible by B.

Proof.

1. Divisibility of A by B means that there exists $Q_1\in P[x]$ such that $A=BQ_1$. Divisibility of B by C means that there exists $Q_2\in P[x]$ such that $B=CQ_2$. Then $A=BQ_1=CQ_2\cdot Q_1=C\cdot Q_2Q_1$. Setting $Q_3=Q_2Q_1\in P[x]$, we get $A=CQ_3$. Hence, A is divisible by C.

2. If $A_1,A_2,\dots,A_k$ are divisible by B, then there exist $Q_1,Q_2,\dots,Q_k\in P[x]$ such that $A_1=BQ_1$, $A_2=BQ_2$,…, $A_k=BQ_k$. Then

$$A_1C_1+A_2C_2+\cdots+A_kC_k=BQ_1C_1+\ BQ_2C_2+\cdots+BQ_kC_k=BQ,$$

where $Q=Q_1C_1+Q_2C_2+\cdots+Q_kC_k\in P[x]$. Therefore, $A_1C_1+\ A_2C_2+\cdots+A_kC_k$ is divisible by B. •

Lemma 2. *Let $A,B\in P[x]$. The divisibility (or indivisibility) of A by B is preserved when both polynomials are multiplied by any nonzero elements of the field P.*

Proof. Suppose A is divisible by B. Then there exists $Q\in P[x]$ such that $A=BQ$. Let $c,d\in P\setminus\{0\}$. Then $cA=c(BQ)=dB\cdot cd^{-1}Q$, and since $d^{-1}\in P$, it follows that $cd^{-1}Q\in P[x]$. Thus, cA is divisible by dB. Conversely, suppose cA is divisible by dB. Then, there exists $\tilde{Q}\in P[x]$ such that $cA=\ dB\cdot\tilde{Q}$. Multiplying both sides by c^{-1} (which exists since $c\neq 0$) gives $A=B\cdot c^{-1}d\tilde{Q}$, and $c^{-1}d\tilde{Q}\in P[x]$. Hence, A is divisible by B. •

Corollary 3.

1. *Every $A\in P[x]$ is divisible by every $B\in P[x]$ with degree* $\deg B=0$; *that is, by every nonzero element of the field P. In other words, dividing a polynomial A by a nonzero constant rescales the coefficients of A and the quotient remains a polynomial in $P[x]$.*
2. *If* $\deg A=0$ *and A is divisible by B, then* $\deg B=0$. •

Remark 19. Unfortunately, pages 67 and 68 of the original Russian Version of this algebra course are missing. However, the subsequent exposition appears to be largely unaffected by this omission. We apologize for any inconvenience this may cause.

Remark 20. It is our view that the reconstructed material corresponding to the missing pages 67 and 68 can be presented in the following general form:

Theorem. (*Division Algorithm for Polynomials*) *Let $A,B\in P[x]$ with $B\neq 0$. Then there exist unique $Q,R\in P[x]$ such that*

$$A=BQ+R,\ \deg R<\deg B$$

Definition. A polynomial $D\in P[x]$ is called the *greatest common divisor* of $A,B\in P[x]$, denoted by $\gcd(A,B)$, if

1. both A and B are divisible by D;
2. D is divisible by any other common divisor of A and B.

By convention, the gcd is chosen *monic* (leading coefficient 1).

Lemma 3. *If D and D_1 are both* $\gcd(A,B)$, *then $D=cD_1$ for some $c\in P\setminus\{0\}$. In particular, the monic* gcd *is unique*.

Comment. Existence of gcds follows from the Euclidean algorithm for polynomials: repeated multiplication of the *Division Algorithm* yields the last nonzero remainder, which is the gcd of the two polynomials. •

Theorem-Lemma 1. *For any two polynomials $A,B\in P[x]$, with at least one of them nonzero, there exist $U,V\in P[x]$ such that $UA+VB=$* $\gcd(A,B)$. •

Remark 21. This result follows directly from the Euclidean algorithm: we express R_k in terms of R_{k-1} and R_{k-2}, then express R_{k-1} in terms of R_{k-2} and R_{k-3}, and so forth, until R_k is represented solely in terms of A and B. Subsequently, we multiply the polynomials that are factors of A and B by r^{-1}, where r is the leading coefficient of the last nonzero remainder R_k. This normalization converts R_k into a *monic polynomial* – a polynomial whose leading coefficient is 1. To avoid unnecessary complexity, we will demonstrate this process with a numerical example.

Example 24. Let $A = 64$, $B = 28$. Then

$$
\begin{array}{rrrr}
 & & 64 & \underline{\big|\,28\,} \\
 & & \underline{56} & 2 \\
 & 28 & \underline{\big|\,8\,} & \\
 & \underline{24} & 3 & \\
8 & \underline{\big|\,4\,} & & \\
\underline{8} & 2 & & \\
0 & & &
\end{array}
\qquad
\begin{array}{l}
64 = 28(2) + 8, \quad 8 = 64 - 2(28), \\
28 = 8(3) + 4, \qquad 4 = 28 - 3(8) \\
8 = 4(2); \\
4 = 28 - 3(8) = 28 - 3[64 - 2(28)] \\
\quad = 28 - 3(64) + 6(28) = -3(64) + 7(28); \\
\gcd(64, 28) = 4; \\
\text{here } U = -3, V = 7.
\end{array}
$$

It is not difficult to show in general that such a representation of $\gcd(A, B)$ is not unique (the proof is left to the reader).

Polynomials $A, B \in P[x]$ are called *coprime* if $\gcd(A, B) = 1$.

Corollary 4. (*from Theorem-Lemma 1*) *A and B are coprime if and only if there exist $U$and B such that $UA + VB = 1$.*

Proof. Statement "⇒" is a special case of Theorem-Lemma 1. Let's prove "⇐". Let there be such U and B. Then, by Lemma 1, polynomial 1 is divisible by any $cd(A, B)$ and by Corollary 3, the degree of it is 0, in particular, $\deg \gcd(A, B) = 0$. Hence, $\gcd(A, B) = 1$ due to the fact that this polynomial is monic. •

Theorem 19. *If* $\gcd(A, C) = 1$, *then* $\gcd(A, BC) = \gcd(A, B)$. ($A, B, C \in P[x]$; $A \neq 0$, $C \neq 0$.)

Proof. By Corollary 4, there exist $U_1, V_1 \in P[x]$ such that

$$U_1 A + V_1 C = 1,$$

and, by Theorem-Lemma 1 itself, there exist $U_2, V_2 \in P[x]$ such that

$$U_2 A + V_2 B = \mathrm{gdc}(A, B);$$

multiplying these two equalities, we get

$$U \cdot A + V \cdot BC = \gcd(A, B),$$

where $U = U_1 U_2 A + U_1 V_2 B + V_1 U_2 C \in P[x], V = V_1 V_2 \in P[x]$. From here, by Lemma 1, it follows that $\gcd(A, B)$ is divisible by any $cd(A, B)$, and therefore also by $\gcd(A, BC)$. Conversely, A and BC are divisible by any $cd(A, B)$, in particular, by $\gcd(A, B)$, therefore $\gcd(A, BC)$ is also divisible by $\gcd(A, B)$. Since both greatest common divisors are monic polynomials, it follows from Lemma 3 (which is missing in the original text but has been reconstructed in Remark 20) that $\gcd(A, BC) = \gcd(A, B)$. •

Corollary 5. *If* $\gcd(A, B) = \gcd(A, C) = \gcd(A, D) = \cdots = 1$, *then* $\gcd(A, BCD \dots) = 1$; *in other words, if a polynomial A is coprime with each of the polynomials $B, C, D, \dots$, then it is coprime with their product*. •

Theorem 20. *If A is divisible by B and C, and* $\gcd(B, C) = 1$, *then A is divisible by BC*. ($A, B, C \in P[x]$; $B \neq 0$, $C \neq 0$.)

Proof. Under the given conditions, there exists $Q \in P[x]$ such that $A = BQ$. Moreover, by Corollary 4, there exist $U, V \in P[x]$ such that $UB + VC = 1$. Multiplying the last equality by Q, we obtain

$$UBQ + VCQ = Q.$$

Since polynomial $BQ = A$ is divisible by C under the given conditions, Lemma 1 implies that Q is divisible by C; that is, there exists $D \in P[x]$ such that $Q = CD$. Hence,

$$A = BQ = B \cdot CD = BC \cdot D,$$

i.e., A es divisible by BC. •

REDUCIBILITY OF POLYNOMIALS

A polynomial $A \in P[x]$, $A \neq 0$ (i.e., $\deg A \geq 0$) is called *reducible over a field P* (or *in the ring $P[x]$*) if there exist polynomials $B, C \in P[x]$ such that $A = BC$, and

$0 < \deg B < \deg A$, $0 < \deg C < \deg A$. (3)

In other words, A is reducible if it can be expressed as the product of two polynomials over the same field P with positive degrees. Since $\deg A = \deg B + \deg C$, each of the inequalities in (3) implies the other. It follows directly from this definition that *polynomials of degree* 0 *and degree* 1 *are irreducible*.

If a polynomial $A \in P[x]$ (that is, over the field P) is irreducible over P and $\hat{P}$ is an extension of P (i.e., $P \subset \hat{P}$ and P is a subfield of $\hat{P}$), then $A \in \hat{P}[x]$ (a polynomial over a given field is also a polynomial over any of its extensions). However, over the extended field $\hat{P}$ the same polynomial may become reducible. For example, let $P = \mathbb{Q}$ and $\hat{P} = \mathbb{Q}(\sqrt{2})$. The polynomial $A = x^2 - 2$ belonging to $\mathbb{Q}[x]$ (and therefore to $\mathbb{Q}(\sqrt{2})[x]$) is irreducible over P, but reducible over $\hat{P}$. Indeed, if we try to factor $x^2 - 2$ as $(x + a)(x + b)$ with $a, b \in \mathbb{Q}$, and writing the equality in the form $x^2 + 0 \cdot x + (-2) = x^2 + (a + b) \cdot x + ab$, we will have $a + b = 0$ and $ab = -2$, which makes $a^2 + ab = 0$ and $ab = -2$. This implies $a^2 = 2$, which has no solution in $\mathbb{Q}$. However, over $\mathbb{Q}(\sqrt{2})$, the factorization $x^2 - 2 = (x + \sqrt{2})(x - \sqrt{2})$ is valid. Another example is $x^2 + 1$. Over $\mathbb{Q}$, $\mathbb{Q}(\sqrt{2})$ and $\mathbb{R}$ this polynomial is irreducible, but over $\mathbb{C}$ it factors as $x^2 + 1 = (x + i)(x - i)$. (You can verify the irreducibility of this polynomial over $\mathbb{R}$ yourself).

Remark 22. If polynomials $A, B \in P[x]$ have a common divisor *cd* over some extension of the field P, it is not necessary that *cd* $\in P[x]$. However, the greatest common divisor $\gcd(A, B)$ always belongs to $P[x]$. This is because, when computing the gcd using the Euclidean algorithm, only operations on the coefficients that remain within the field P (addition, subtraction, multiplication, and division) are performed. For example, polynomials

$$A = x^5 - x^3 + x^2 - 2x - 2,\ B = 2x^3 + x^2 - 4x - 2 \in \mathbb{Q}[x]$$

have a common divisor $cd(A, B) = x - \sqrt{2} \notin \mathbb{Q}[x]$, but their greatest common divisor satisfies $\gcd(A, B) = x^2 - 2 \in \mathbb{Q}[x]$. It follows directly from this remark

Lemma 4. *If monic polynomials $A, B \in P[x]$ are irreducible over the field P, then either $A = B$ or* $\gcd(A, B) = 1$. •

Remark 23. Let's also recall that, due to the absence of zero divisors in the ring $P[x]$, if $AC = BC$ with $C \neq 0$, then $A = B$; that is, both sides of the equality can be divided by a common factor of nonzero degree. This is precisely Theorem 19 on rings applied to $P[x]$.

Theorem 21. *Any polynomial $S \in P[x]$ can be expressed in the form $S = aA_1A_2 \dots A_k$, where $a \in P$ is the leading coefficient of S, and each $A_i \in P[x]$* for *$i = 1, 2, \dots, k$ is monic and irreducible over the field P. Note that the factors $A_1, A_2, \dots, A_k$ may include repetitions. For $S \neq 0$, this factorization is unique up to the order of the factors.*

Proof. For $S = 0$, we may, for instance, set $a = 0$ and choose any collection of monic irreducible polynomials $A_1, A_2, \dots$ in any quantity. In this case, uniqueness of the factorization clearly does not hold.

Let $S \neq 0$. If S is already monic, we set $a = 1$; otherwise, we write S in the form $S = a \cdot a^{-1}S$, where a is the leading coefficient of S, and $a^{-1}S$ is monic (the element $a \in P$ is invertible since $a \neq 0$). If this polynomial is irreducible over P, then, setting $a^{-1}S = A_1$, gives the desired factorization $S = aA_1$ (with $k = 1$). In the case where the polynomial $a^{-1}S$ is reducible, we factor it into a product of two polynomials of lower degree over the same field P. If at least one of these factors is itself reducible, we factor it further and continue this process iteratively. These factorizations can always be performed so that all factors are monic. Since each factorization produces polynomials of strictly smaller degree, and polynomials of degree 1 are irreducible, the process must terminate after finitely many steps, yielding the expression

$S = aA_1A_2 \dots A_k$, (4)

where $a \in P$ and polynomials $A_1, A_2, \dots, A_k$ are monic and irreducible over P, and

$$\deg A_1 + \deg A_2 + \dots + \deg A_k = \deg S,$$

$k \geq 0$ ($k = 0$ if $\deg S = 0$, i.e., $S = a$). We will prove the uniqueness of expression (4).

Assume that for some polynomial $S \in P[x]$, there exist two different factorizations satisfying the theorem conditions:

$$S = aA_1A_2 \dots A_k = bB_1B_2 \dots B_l,$$

where $a \neq 0$ and $b \neq 0$ since $S \neq 0$. In fact, $b = a$, because both a and b are the leading coefficients of S, and all remaining factors in this factorization are monic (assuming $k > 0$ and $l > 0$). Multiplying both sides of the equality by a^{-1}, we obtain

$A_1A_2 \dots A_k = B_1B_2 \dots B_l$. (5)

The Polynomial A_1 cannot be coprime with each of the polynomials $B_1, B_2, \dots, B_l$; if it were, then by Corollary 5, it would be coprime with their product. However, the product on the right-hand side of (5) is divisible by A_1 (since $\deg A_1 > 0$) because both sides of equality (5) represent the same polynomial. Hence, there exists some B_j ($1 \leq j \leq l$) such that $\gcd(A_1, B_j) \neq 1$. By Lemma 4, since A_1 and B_j are monic and irreducible, it follows that $A_1 = B_j$. Dividing both sides of (5) by A_1 on the left and by B_j on the right (see Remark 23), we obtain

$$A_2 \dots A_k = B_1B_2 \dots B_{j-1}B_{j+1} \dots B_l.$$

Applying the same reasoning to the new equality, we identify a factor on the right-hand side equal to A_2, cancel it from both sides, and continue in this manner until one side becomes 1, meaning all factors have been exhausted. Since both sides are monic polynomials of the same degree, the other side must be also 1. Consequently, $l = k$, and the polynomials $A_1, A_2, \dots, A_k$ coincide with $B_1, B_2, \dots, B_k$, though not necessarily in the same order. The statement now follows. •

By grouping identical irreducible factors in (4) (renumbering them accordingly and letting k denote the number of distinct factor groups rather than the total number of factors), we can write the polynomial $S \in P[x]$ in the form

$$S = aA_1^{r_1} \cdot A_2^{r_2} \cdot \dots \cdot A_k^{r_k},$$

where S is an arbitrary polynomial over the field P, $a \in P$, and $A_1, A_2, \dots, A_k$ are monic, irreducible polynomials over P, pairwise coprime (and thus distinct). In this case,

$$r_1 \deg A_1 + r_2 \deg A_2 + \dots + r_k \deg A_k = \deg S.$$

The natural number r_j is called the *multiplicity* of the factor A_j in the product S ($j = 1, 2, \dots, k$ with $k \geq 0$). The representation is unique up to a permutation of the factors $A_j^{r_j}$ among themselves.

ASSIGNMENT FOR INDEPENDENT WORK.

(*based on A.G. Kurosh*, *Course of Higher Algebra*).

1. Give a purely algebraic definition of the derivative of a polynomial. Derive the rules for differentiating the sum and the product of polynomials, and explain why results from mathematical analysis cannot be directly applied here.

2. Let P be a field of characteristic $p = 0$, and let

$$S = aA_1^{r_1}A_2^{r_2} \dots A_k^{r_k}$$

be the factorization of $S \in P[x]$ into irreducible factors. Prove that

$$\frac{S}{\gcd(S,S')} = aA_1A_2 \dots A_k,$$

that is, dividing S by the greatest common divisor of S and its derivative S' yields a polynomial containing the same irreducible factors as S, but each to the first power.

3. Investigate the role of the condition $p = 0$.

VALUES AND ROOTS OF POLYNOMIALS

We already know that in the expression

$$A(x) = a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$$

for the polynomial $A = A(x) \in P[x]$, the formal variable x is defined as the element $(0, 1, 0, \dots)$ of the ring of polynomials over the field P, and not as "a blank space to be filled with an arbitrary element of P". Nevertheless, the operation of *substituting x with an element $c \in P$ in the polynomial $A(x)$* can be rigorously defined by stipulating that the result of such a substitution is the element

$$A(c) = a_nc^n + a_{n-1}c^{n-1} + \dots + a_1c + a_0 \in P$$

obtained from c by performing the same operations (in the same order) by which $A(x)$ is obtained from x. This resulting element $A(c)$ is called *the value of the polynomial $A(x)$ at $x = c$*. If A and B are the same polynomial (written as $A = B$ or $A(x) = B(x)$), then clearly $A(c) = B(c)$ for any $c \in P$. However, the converse does not hold for every field P. For example, over the field $\mathbb{Z}_2$ (of characteristic $p = 2$) consisting of two elements $\{0, 1\}$–where $1 + 1 = 0$ – the polynomials $A(x) = x + 1$, $B(x) = x^2 + 1$ are distinct yet satisfy $A(0) = B(0) = 1$ and $A(1) = B(1) = 0$. We will not generalize this example here, but note the underlying reason for the phenomenon: the field $\mathbb{Z}_2$ contains so few elements that the variety of distinct polynomials over it is much greater than the number of possible function values.

An element c is called a *root* of a polynomial $A(x)$ (or of the equation $A(x) = 0$) if $A(c) = 0$. The element c need not to belong to the field P over which the polynomial A is defined, it may lie in some extension $\hat{P}$ of P. The definition of $A(c)$ remains valid in this case since $A(x) \in P[x]$ implies $A(x) \in \hat{P}[x]$. For example, the polynomial $x^2 - 2$ over the field of rational numbers has no roots in $\mathbb{Q}$, but has roots in its extension $\mathbb{Q}(\sqrt{2})$ as well as in larger (nonminimal) extensions containing it, such as $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, $\mathbb{R}$ and others.

Theorem 22. (*Little Bézout's Theorem, also known as the Polynomial Remainder Theorem*) *The remainder upon dividing a polynomial $A(x) \in P[x]$ by a monic first-degree binomial $x - c$ (which reduces to a monomial when $c = 0$), where $c \in P$, is equal to the value $A(c)$. In particular, $A(x)$ is divisible by $x - c$ if and only if c is a root of $A(x)$.*

Proof. When dividing $A(x)$ by $x - c$ with remainder, we have

$$A(x) = (x - c)Q(x) + R(x),$$

where $\deg R(x) < \deg(x - c)$. Moreover, both the remainder $R(x)$ and the quotient $Q(x)$ are uniquely determined by $A(x)$ and c. Since $\deg(x - c) = 1$, it follows that $\deg R(x) \le 0$; that is, $R(x)$ is constant and therefore an element of the field P. Thus, the equality

$$A(x) = (x - c)Q(x) + R \quad (6)$$

becomes an equality of elements of P when x is replaced by any element of P. In particular, substituting $x = c$ yields $A(c) = (c - c)Q(c) + R = R$. •

Thanks to Theorem 22, we can determine whether a polynomial $A(x)$ is divisible by a binomial $x - c$ without explicitly performing the division. Moreover, there exists an efficient procedure to compute both the remainder R and the coefficients of the quotient

$$Q(x) = q_{n-1}x^{n-1} + q_{n-2}x^{n-2} + \dots + q_1x + q_0,$$

requiring far fewer multiplications than directly substituting c into $A(x)$. This method is traditionally attributed to William Horner (18th-19th centuries), although historical evidence indicates that it was already known in China well before the Common Era.

Since the equality of two polynomials implies that the coefficients of corresponding powers of the formal variable x are identical, expanding the brackets on the right-hand side of the fully written equality (6)

$$a_nx^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 = (x - c)(q_{n-1}x^{n-1} + q_{n-2}x^{n-2} + \dots + q_1x + q_0) + R$$

and by equating the corresponding coefficients, we obtain

$$\begin{array}{ccc}
a_n = q_{n-1}, & & q_{n-1} = a_n, \\
a_{n-1} = q_{n-2} - cq_{n-1}, & & q_{n-2} = cq_{n-1} + a_{n-1}, \\
a_{n-2} = q_{n-3} - cq_{n-2}, & & q_{n-3} = cq_{n-2} + a_{n-2}, \\
\dots\dots\dots & \text{where} & \dots\dots\dots \\
a_2 = q_1 - cq_2, & & q_1 = cq_2 + a_2, \\
a_1 = q_0 - cq_1, & & q_0 = cq_1 + a_1, \\
a_0 = R - cq_0, & & R = cq_0 + a_0.
\end{array}$$

The sequential computation of q_{n-1}, q_{n-2},…, q_1, q_0, R can be conveniently represented using *Horner's scheme* (Table 4).

Table 4

	a_n	a_{n-1}	...	a_{k+1}	a_k	...	a_2	a_1	a_0
c	q_{n-1}	q_{n-2}	...	q_k	q_{k-1}	...	q_1	q_0	R

Example 25. Divide $A(x) = 2x^7 + 15x^6 + 240x^4 - 27x^3 + 10x^2 + 94x + 36$ by $x + 9$ and determine whether $A(-9) = 0$.

In this case, $c = -9$. After entering the coefficients of the polynomial $A(x)$ in the top row of Table 5, fill in the cells of the bottom row sequentially from left to right.

Table 5

	2	15	0	240	-27	10	94	36
-9	2	-3	27	-3	0	10	4	0

Remark 24. The coefficients of the last row of Table 5 (computed using Horner's scheme) can equivalently be obtained by applying *synthetic division* to the coefficients of $A(x)$ in the usual way:

2	15	0	240	-27	10	94	36	-9
	-18	27	-243	27	0	-90	-36	
2	-3	27	-3	0	10	4	0	

Thus, $Q(x) = 2x^6 - 3x^5 + 27x^4 - 3x^3 + 10x + 4$, $R = 0$, i.e., $A(x)$ is divisible by $x + 9$, and therefore $A(-9) = 0$. If we attempt to verify directly whether -9 is a root of $A(x)$, the calculation would require $8 + 7 + 5 + 4 + 3 + 1 = 28$ multiplications. In contrast, using Horner's scheme the multiplication operation is performed only 7 times (including one multiplication by zero).

From Theorem 22 it follows that *the zero element* ($c = 0$) *of the field P is a root of* $A(x) \in P[x]$ *if and only if* $a_0 = 0$. In this case, Horner's scheme becomes trivial: to compute the quotient $A(x)/x$, it suffices to reduce each power of x by 1 in $A(x)$.

TASK (optional): Describe and justify two variants – "right" and "left" – of Horner's schemes for polynomials over a noncommutative (but associative) ring with unity.

ABSOLUTE REDUCIBILITY OF POLYNOMIALS OVER THE FIELD $\mathbb{C}$. By the Fundamental Theorem of Algebra, any polynomial $A(x)$ of positive degree with complex coefficients has at least one complex root $c_1 \in \mathbb{C}$ *(The theorem does not apply to a polynomial of degree* 0 *or negative since in that case* $A(x) = 0$ *is not an equation).* By Theorem 22, $A(x)$ is divisible by the binomial $x - c_1$; that is, there exists a polynomial $Q_1(x) \in \mathbb{C}[x]$ such that $A(x) = (x - c_1)Q_1(x)$. If $\deg Q_1 > 0$, then $Q_1(x)$ also has a root $c_2 \in \mathbb{C}$ (possibly equal to c_1), so that $Q_1(x) = (x - c_2)Q_2(x)$ and hence $A(x) = (x - c_1)(x - c_2)Q_2(x)$. Continuing this process, we obtain a sequence of polynomials $A, Q_1, Q_2, ...$ each having degree one less than its predecessor. After n steps, we arrive at a constant $Q_n = a_n \in \mathbb{C}$, which is the leading coefficient of the original polynomial $A(x)$ *(why?)*. Thus, we obtain the factorization – unique up to permutation of the linear factors:

$A(x) = a_n(x - c_1)(x - c_2) ... (x - c_n)$. (7)

Since polynomials of degree zero or one are irreducible over any field, a decomposition of the form (7) represents an *extremely fine level of factorization*. For this reason, we say that polynomials over $\mathbb{C}$ are *absolutely reducible* in this field (or in the ring $\mathbb{C}[x]$).

If $A(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, then by expanding the right-hand side of equality (7) and equating the coefficients of corresponding powers of x, we obtain Vieta's formulas

$$a_n \cdot \sum_{j_1=1}^{n} c_{j_1} = -a_{n-1},$$

$$a_n \cdot \sum_{1 \le j_1 < j_2 \le n}^{n} c_{j_1} c_{j_2} = a_{n-2},$$

$$a_n \cdot \sum_{1 \le j_1 < j_2 < j_3 \le n}^{n} c_{j_1} c_{j_2} c_{j_3} = -a_{n-3},$$

$$\dots\dots\dots\dots\dots\dots$$

$$a_n \cdot \sum_{1 \le j_1 < j_2 < \ldots < j_n \le n}^{n} c_{j_1} c_{j_2} \ldots c_{j_{n-1}} = (-1)^{n-1} a_1,$$

$$a_n \cdot c_1 c_2 \ldots c_n = (-1)^n a_n,$$

Vieta's formulas are usually introduced in school only for the case of real coefficients and roots (often restricted to $n = 2$).

Let $c_1, c_2, \dots, c_n$ be the distinct roots of the polynomial A. Then the factorization of $A(x)$ can be written as

$$A(x) = a_n(x - c_1)^{r_1}(x - c_2)^{r_2} \dots (x - c_n)^{r_k},$$

where r_j denotes the multiplicity of the root c_j for $j = 1, 2, \dots, k$. In this case, it is clear that

$$r_1 + r_2 + \cdots r_k = n = \deg A.$$

REDUCIBILITY OF POLYNOMIALS OVER THE FIELD $\mathbb{R}$. In the ring $\mathbb{R}[x]$, there exist polynomials of degree greater than 1 that are irreducible over $\mathbb{R}$; for example, $x^2 + 1$. However, it turns out that no such irreducible polynomials of degree greater than 2 can exist.

Lemma 5. If $A(x) \in \mathbb{R}[x]$, $c \in \mathbb{C}$ *and* $A(c) = 0$, *then* $A(\bar{c}) = 0$. *In other words, if a complex number* $c = a + bi$ *is a root of a polynomial with real coefficients, then its conjugate* $\bar{c} = a - bi$, *with* $a, b \in \mathbb{R}$, *is also a root.*

Proof. Let

$$A(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_n, a_{n-1}, \dots, a_1, a_0 \in \mathbb{R}$. Suppose that

$$A(c) = a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0 = 0,$$

for some $c \in \mathbb{C}$. Using the basic properties of complex conjugation, namely

$$\overline{c_1 c_2} = \bar{c}_1 \bar{c}_2,\ \overline{c_1 + c_2} = \bar{c}_1 + \bar{c}_2,$$

and the fact that $\bar{a} = a$ for all $a \in \mathbb{R}$ (in particular, $\bar{0} = 0$), we obtain

$$\overline{A(c)} = \overline{a_n c^n + a_{n-1}c^{n-1} + \cdots + a_1 c + a_0} = a_n(\bar{c})^n + a_{n-1}(\bar{c})^{n-1} + \cdots + a_1\bar{c} + a_0 = A(\bar{c}).$$

Since $A(c) = 0$, it follows that $\overline{A(c)} = 0$. Thus, $A(\bar{c}) = 0$, which proves the claim. •

Theorem 23. *If* $A(x) \in \mathbb{R}[x]$ *and* $\deg A(x) > 2$, *then the polynomial* $A(x)$ *is reducible over* $\mathbb{R}$.

Proof. Since $\mathbb{R} \subset \mathbb{C}$, the polynomial $A(x)$, with real coefficients, also belongs to $\mathbb{C}[x]$. By the Fundamental Theorem of Algebra, $A(x)$ has at least one root $c = a + bi \in \mathbb{C}$. If $c \in \mathbb{R}$ (i.e., $b = 0$), then $x - c \in \mathbb{R}[x]$ divides $A(x)$. Hence, $A(x)$ is reducible over $\mathbb{R}$. If $c \notin \mathbb{R}$ (i.e., $b \neq 0$), then, by Lemma 5, $\bar{c} = a - bi$ is also a root. Thus, $x - c$ and $x - \bar{c}$ both divide $A(x)$. Since they are distinct irreducible monic polynomials in $\mathbb{C}[x]$, they are coprime by Lemma 4. Therefore, their product $(x - c)(x - \bar{c}) = x^2 - (c + \bar{c})x + c\bar{c}$ divides $A(x)$. Note that $c + \bar{c} = 2a \in \mathbb{R}$ and $c\bar{c} = a^2 + b^2 \in \mathbb{R}$, so this quadratic factor lies in $\mathbb{R}[x]$. Since $\deg A(x) > 2$, the existence of a quadratic divisor in $\mathbb{R}[x]$ implies that $A(x)$ is reducible over $\mathbb{R}$. •

Corollary 6. *Every polynomial of odd degree with real coefficients has at least one real root.*

Proof. Each non-real complex root of a polynomial with real coefficients occurs in a conjugate pair (by Lemma 5). Hence, the linear factors $(x - c)$ and $(x - \bar{c})$ can be combined into quadratic factors with real coefficients. Since the polynomial has odd degree, after pairing all non-real roots there must remain at least one unpaired linear factor. This factor has necessarily real coefficients since together with the quadratic factors it reproduces the original polynomial. Therefore, the unpaired linear factor corresponds to a real root. •

Remark 25. Corollary 6 remains valid in the presence of multiple roots. Indeed, if $c = a + bi$ with $b \neq 0$ is a root of multiplicity r, then $\bar{c} = a - bi$ is also a root of multiplicity r (prove this).

Remark 26. The existence of a real root for a polynomial of odd degree over $\mathbb{R}$ can be proved analytically. The function $y = A(x)$ is defined and continuous on the entire real line $(-\infty, +\infty)$. As $|x| \to +\infty$, the leading coefficient of $A(x)$ grows faster in absolute value than the sum of the remaining terms. Hence, there exists a sufficiently large positive number L such that the sign of $A(L)$ coincides with that of the leading coefficient. Since $\deg A$ is odd, the signs of $A(L)$ and $A(-L)$ are opposite. By the *Bolzano Theorem*, a continuous function on $[-L, L]$ that takes opposite signs at the endpoints, must vanish at least once in that interval. Therefore, $A(x)$ has at least one real root.

Example 26. This example illustrates Theorem 23. At first glance, the polynomial $x^4 + 4$ may appear irreducible over $\mathbb{R}$. However, Theorem 23 implies that this is not the case. One possible approach is to determine all fourth roots of -4 and then combine the conjugate pairs of linear factors (we leave this as an exercise). A simpler method is to use the following algebraic manipulation:

$$x^4 + 4 = (x^4 + 4x^2 + 4) - 4x^2 = (x^2 + 2)^2 - (2x)^2 = [(x^2 + 2) - 2x][(x^2 + 2) + 2x] =$$
$$= (x^2 - 2x + 2)(x^2 + 2x + 2).$$

Thus, $x^4 + 4$ factors into a product of two irreducible quadratics over $\mathbb{R}$.

Theorem 24. (*Factorization of Polynomials over* $\mathbb{R}$) *Any polynomial* $A(x) \in \mathbb{R}[x]$ *can be uniquely factored into irreducible polynomials over* $\mathbb{R}$ *in the following form*:

$$A(x) = a_n(x - c_1)^{r_1} \dots (x - c_k)^{r_k}(x^2 + p_1 x + q_1)^{s_1} \dots (x^2 + p_l x + q_l)^{s_l},$$

where

- $a_n, c_1, \dots, c_k, p_1, q_1, \dots, p_l, q_l \in \mathbb{R}$,
- $r_1, \dots, r_k, s_1, \dots, s_l \in \mathbb{N}$,
- *the degree condition holds*:

$$r_1 + \cdots + r_k + 2s_1 + \cdots + 2s_l = n = \deg A$$

- *and for each quadratic factor, the discriminant is negative*

$p_i^2 - 4q_i < 0$ *for all* $i = 1, \dots, l$. •

Remark 27. The discriminants must be negative because otherwise a quadratic trinomial would factor further into real linear factors, contradicting irreducibility over $\mathbb{R}[x]$.

POLYNOMIALS OVER THE FIELD $\mathbb{Q}$. For polynomials with rational coefficients, the problem of reducibility over the field $\mathbb{Q}$ is more subtle than in the cases of $\mathbb{C}$ and $\mathbb{R}$. Unlike those fields, there is no general universal criterion for irreducibility over $\mathbb{Q}$. One useful tool is the *Eisenstein criterion*, studied in standard algebra textbooks (e.g., D.G. Kurosh, *Course of Higher Algebra*). Despite its name, it is not a full criterion, but rather a *sufficient condition for irreducibility*. For example, it can be applied to show that the polynomial $x^n + x^{n-1} + \cdots + x + 1$ is irreducible over $\mathbb{Q}$ for any $n \geq 0$. This already demonstrates the existence of irreducible polynomials of arbitrarily large degree in $\mathbb{Q}[x]$. Thus, in practice, the main question shifts from general reducibility to the more concrete problem of finding all rational roots of a given polynomial.

It is convenient to regard the coefficients of the polynomial $A \in \mathbb{Q}[x]$ as integers. This does not affect the generality of the problem of finding roots because the roots clearly remain unchanged if the entire polynomial is multiplied by any nonzero number, in particular, by the least common denominator of all fractional coefficients. Thus, we may assume that

$$A(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}(x)$$

is a given polynomial with integer coefficients (i.e., defined over the ring $\mathbb{Z}$, and hence over the field $\mathbb{Q}$). We first consider integer roots, noting that the root 0 (together with its multiplicity) can be identified immediately.

Theorem 25. *If $x_0 \neq 0$ is an integer root of a polynomial $A \in \mathbb{Z}(x)$, then x_0 must be a divisor of the constant term a_0 of A.*

Proof. From

$$A(x_0) = a_n {x_0}^n + a_{n-1}{x_0}^{n-1} + \cdots + a_1 x_0 + a_0 = 0,$$

it follows that

$$a_0 = x_0 \cdot (-a_n {x_0}^{n-1} - a_{n-1}{x_0}^{n-2} - \cdots - a_1).$$

Thus, x_0 divides a_0 ($x_0 \mid a_0$), as claimed. •

By Theorem 25, all integer roots of a polynomial can be determined (or their absence established).

Example 27. Consider the polynomial

$$A(x) = 2x^7 - 2x^6 - 9x^5 - 7x^4 - 4x^3 - 4x^2 - 5x - 3.$$

The divisors of the constant term are ± 1 and ± 3. Hence, any integer root must belong to this set. To test these candidates, it is convenient to use the *Repeated Horner's Scheme* (also referred to as the *Multi-stage* or *Multi-story Method*). In this scheme, the steps ("stories") that produce a nonzero remainder R are omitted (we shade them in practice; see Table 6). Each root, since it may occur with multiplicity, requires at least two successive stories.

Table 6

	2	-2	-9	-7	-4	-4	-5	-3	
1	2	0	-9	-16	-20	-24	-29	-32	1 is not a root.
-1	2	-4	-5	-2	-2	-2	-3	0	
-1	2	-6	1	-3	1	-3	0		
-1	2	-8	9	-12	13	-16			-1 is a root of multiplicity 2.
3	2	0	1	0	1	0			
3	2	6	19	57	172				3 is a root of multiplicity 1.
-3	2	-6	19	-57	172				-3 is not a root.

Thus, the polynomial A has two integer roots: $x = -1$ (with multiplicity 2), and $x = 3$ (with multiplicity 1). The quotient obtained by dividing $A(x)$ by $(x+1)^2(x-3)$ is the fourth-degree polynomial $2x^4 + x^2 + 1$, whose coefficients appear in the last unshaded row of Table 6. This polynomial does not have any real roots since its values are strictly positive for all $x \in \mathbb{R}$.

In Example 27, only four numbers were subject to verification. However, if the constant term has a large number of divisors, some of them can be eliminated in advance, even before applying Horner's scheme, using the following considerations:

1. The direct computation of $A(1)$ and $A(-1)$ does not involve any multiplication and is computationally simple. Since both ± 1 divide any constant term $a_0 \neq 0$, they must always be checked.
2. Any other divisor $x_0 \neq \pm 1$ cannot be a root of A if at least one of the numbers $A(1)/(x_0 - 1)$ or $A(-1)/(x_0 + 1)$ is not an integer. Indeed, for any integer root $x_0 \neq \pm 1$ of $A(x) = (x - x_0)Q(x)$, $Q(x) \in \mathbb{Z}(x)$, it follows that $A(1)/(1 - x_0) = Q(1)$, $A(-1)/(1 + x_0) = Q(-1)$ must both be integers.

Example 28. Consider the polynomial

$$A(x) = 3x^6 - x^5 + 9x^4 + 3x^3 - 56x^2 + 54x - 12.$$

Since $A(1) = 0$, we conclude that $x = 1$ is a root. Using Horner's scheme (or direct polynomial division), we establish that its multiplicity is equal to two. Indeed, dividing $A(x)$ by $(x-1)^2$ gives the polynomial

$$Q_1(x) = 3x^4 + 5x^3 + 16x^2 + 30x - 12,$$

for which, clearly, $x = 1$ is not longer a root. In this case, $Q_1(1) = 42$, $Q_1(-1) = -28$. Hence, -1 is not a root either (otherwise we would divide $Q_1(1)$ by $x + 1$ as many times as the multiplicity of that root). Now, the divisors of the constant term -12 to be tested are: $\pm 2, \pm 3, \pm 4, \pm 6, \pm 12$.

Candidates $x_0 = -3, -4, 6, 12, -12$ are rejected since for them the numbers $42/(x_0 - 1)$ are not integers. Similarly, for $x_0 = 2, 4, -6$, the numbers $-28/(x_0 + 1)$ are not integers. (Candidates already excluded need not be tested again). Thus, only $x_0 = -2$ and $x_0 = 3$ remain, which must be tested using Horner's scheme (Table 7.)

Table 7

	3	5	16	30	-12	
-2	3	-1	18	-6	0	
-2	3	-7	32	-70		-2 is a root of multiplicity 1.
3	3	8	42	120		3 is not a root.

Thus, $A(x)$ has the following integer roots:

$x_1 = 1$ (of multiplicity 2), $x_2 = -2$ (of multiplicity 1).

After dividing $A(x)$ by $(x-1)^2(x+2)$, we get the polynomial

$$Q_2(x) = 3x^3 - x^2 + 18x - 6,$$

which has no integer roots.

To find all rational roots of a polynomial, we use the following result:

Theorem 26. *If $A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \in \mathbb{Z}(x)$ is a monic polynomial with integer coefficients, then all its rational roots are necessarily integers.*

Proof. Assume the contrary: Suppose $A(p/q) = 0$, where $p \neq 0, q \neq 0, \pm 1$ are integers, and the fraction p/q is irreducible, i.e., $\gcd(p,q) = 1$. Then

$$(p/q)^n + a_{n-1}(p/q)^{n-1} + \cdots + a_1(p/q) + a_0 = 0.$$

Multiplying through by q^{n-1} and rearranging gives

$$\frac{p^n}{q} = -(a_{n-1}p^{n-1} + a_{n-2}p^{n-2}q + \cdots + a_1pq^{n-2} + a_0q^{n-1}).$$

The right-hand side is an integer, so the left-hand side must also be an integer. But since $\gcd(p,q) = 1$, it follows that $\gcd(p^n, q) = 1$. Thus, $\frac{p^n}{q}$ is already in lowest terms, which forces $|q| = 1$. This contradicts the assumption that $q \neq \pm 1$. Therefore, every rational root of a monic polynomial with integer coefficients is necessarily an integer. •

By Theorem 26, all rational roots of the polynomial

$$A(x) = a_nx^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0 \in \mathbb{Z}(x),$$

where $a_n \neq 0$, can be determined as follows. Multiply the polynomial by a_n^{n-1} (which does not change the set of roots), and introduce a new formal variable $y = a_nx$. Then the polynomial $a_n^{n-1}A(x)$ can be rewritten as a monic polynomial in y:

$$B(y) = y^n + a_{n-1}y^{n-1} + a_{n-2}a_ny^{n-2} + \cdots + a_1a_n^{n-2}y + a_0a_n^{n-1}.$$

Note that, according with the general definition of the ring K^∞ of sequences and its polynomial subrings, this substitution of the formal variable amounts to expressing the polynomial $(a_0, a_1, \dots, a_n)$ in terms of the powers of the element $y = (0, a_n, 0, \dots, 0)$, rather than in terms of $x = (0, 1, 0, \dots, 0)$. The ring $K[y]$ of polynomials in the variable y is isomorphic to the ring $K[x]$; the natural isomorphism should be written explicitly.

And so, each rational root x_0 of $A(x)$ corresponds uniquely to a rational root $y_0 = a_nx_0$ of $B(y)$. By Theorem 26, this root y_0 must be an integer. Consequently, all rational roots of the original polynomial $A(x)$ can be obtained by finding all integer roots of $B(y)$ and dividing them by a_n.

Thus, in Example 28, the polynomial

$$Q_2(x) = 3x^3 - x^2 + 18x - 6$$

was considered. This polynomial has no integer roots; however, since it is not monic, fractional roots may exist. Multiplying Q_2 by 3^2 and making the substitution $y = 3x$, we obtain the monic polynomial

$$B(y) = y^3 - y^2 + 54y - 54,$$

which, in this case, factors immediately as

$$B(y) = (y-1)(y^2 + 54).$$

This factorization spares us from testing the numerous divisors of 54. The quadratic factor $y^2 + 54$ is strictly positive for all $y \in \mathbb{R}$, while the linear factor gives the simple root $y_0 = 1$. From this, we deduce the rational root $x_0 = 1/3$ also of multiplicity one (this should be verified rigorously) of the original polynomial

$$A(x) = 3x^6 - x^5 + 9x^4 + 3x^3 - 56x^2 + 54x - 12.$$

Thus, besides the rational roots $x = \frac{1}{3}, 1$, and -2 (the latter two found previously), the polynomial $A(x)$ has no other rational roots.

EMBEDDING A RING INTO A FIELD

As is well known, in the ring $\mathbb{Z}$ of integers, division is not always feasible: not for any $m, n \in \mathbb{Z}$ with $n \neq 0$, there exists $a \in \mathbb{Z}$ such that $m = an$. The need to divide numbers into equal parts – arising from both practical and theoretical considerations – led to the introduction of fractions. This ultimately resulted in the extension of the ring $\mathbb{Z}$ into the field $\mathbb{Q}$ of rational numbers. Let's now generalize and rigorously justify this historical process.

We have already studied field extensions; here, however, we must extend not a field but a ring, in such a way that the new structure strictly contains the original ring (or an isomorphic copy of it) as a subring, while itself forming a field. Since every field is also a ring, the notion of a *subring* is meaningful (similarly, a *subfield* of a ring is simply a subring that is, in fact, a field). The construction of a field containing a given ring – or an isomorphic image of it – as a subring is called an *embedding of the ring into a field*.

Now, since every field is associative, commutative and free of zero divisors, it follows that each of its subrings also satisfies these properties. Thus, these are necessary conditions for a ring to be embeddable in a field. Remarkably, these conditions are not only necessary but also sufficient: whenever a ring satisfies them, there exists a unique (up to isomorphism) minimal field into which the ring can be embedded.

Theorem 27. (*Field of Fractions and its Uniqueness*) *Any associative and commutative ring K without zero divisors* (*called integral ring*) *admits an embedding into a field, and all minimal fields containing K as a subring are isomorphic to one another*.

Proof. First, consider the trivial ring $(\{0\},+,\cdot)$. This ring is a subring of the field $\mathbb{Z}_2$, which is minimal since every field contains at least two distinct elements, 0 and 1. Moreover, a field consisting solely of 0 and 1 is unique up to isomorphism (to be proved separately).

Now let $K=(M,+,\cdot)$ be a nontrivial associative and commutative ring without zero divisors. In such a ring, the product of finitely many elements is well defined (associativity), and any equality may be reduced by cancelling a common nonzero factor (absence of zero divisors).

(*Construction of the formal fractions*) Let's define $M_0^2\subseteq M^2$ so that
$$M_0^2=\{(x,v)\in M\times M: v\neq 0\}.$$
Since K is nontrivial, we have $M_0^2\neq\emptyset$. For convenience, we denote an ordered pair (x,v) by the formal fraction x/v. On M_0^2, introduce the relation $(x,v)\sim(y,w)\Leftrightarrow xw=yv$. This is an equivalence relation:

- *Reflexive*: $xv=xv$.
- *Symmetric*: If $xw=yv$, then $yv=xw$.
- *Transitive*: If $xw=yv$ and $yu=zw$, then multiplying gives $xwu=yvu=zwv$, hence $xu=zv$.

The set of equivalence classes under this relation will be denoted by $\widehat{M}=M_0^2\,/\sim$. Each class is represented by a formal fraction x/v, corresponding to the equivalence class of (x,v). We now define addition and multiplication on $\widehat{M}$ so as to reproduce the familiar arithmetic operations with fractions.

(*Arithmetic of fractions*) Let $\hat{x},\hat{y}\in\widehat{M}$ and choose representatives $x/u, y/v$ from these classes. Define
$$\hat{x}+\hat{y}=[(xv+uy)/uv],\ \hat{x}\cdot\hat{y}=[xy/uv],$$
where brackets denote the equivalence class in $\widehat{M}$. Since $u\neq 0, v\neq 0$ and K has no zero divisors, we have $u\cdot v\neq 0$. Hence, both operations are well defined on M_0^2. The independence of representatives holds: if $x'/u'=x/u$ and $y'/v'=y/v$, then $x'u=u'x$ and $y'v=v'y$. It follows that
$$(x'v'+u'y')uv=x'v'uv+u'y'uv=x'u\cdot v'v+u'u\cdot y'v=u'x\cdot v'v+u'u\cdot v'y=u'v'(xv+uy),$$
i.e., $(x'v'+u'y')/u'v'=(xv+uy)/uv$. The analogous verification for multiplication is immediate. Thus, the binary operations defined above are consistent, and $(\widehat{M},+)$, $(\widehat{M},\cdot)$ are well defined groupoids.

Both groupoids are associative and commutative. Indeed, let $\hat{x},\hat{y},\hat{z}\in\widehat{M}$ and let $x/u, y/v, z/w$ be representatives of these classes. Since $(xv+uy)/uv\in\hat{x}+\hat{y}$, we obtain
$$[(xv+uy)w+uv\cdot z]/(uv\cdot w)=[(xv+uy)/uv]+z/w\in(\hat{x}+\hat{y})+\hat{z}.$$
On the other hand, since $(yw+vz)/vw\in\hat{y}+\hat{z}$, it follows that
$$[x\cdot vw+u(yw+vz)]/(u\cdot vw)=x/u+[(yw+vz)/vw]\ \in\hat{x}+(\hat{y}+\hat{z}).$$
But the left-hand side of these two fractions are equal (in fact, in this case they even coincide); hence they represent the same class. Therefore, $(\hat{x}+\hat{y})+\hat{z}=\hat{x}+(\hat{y}+\hat{z})$. Similarly, the relations
$$\hat{x}+\hat{y}=\hat{y}+\hat{x},\ (\hat{x}\hat{y})\hat{z}=\hat{x}(\hat{y}\hat{z}),\ \hat{x}\hat{y}=\hat{y}\hat{x}$$
are verified without difficulty.

In the groupoid $(\widehat{M},+)$, the neutral element 0 is the class of fractions of the form $0/t$ $(t\neq 0)$, and the opposite for the class $\hat{x}$, which contains the fraction x/u, is the class $(-\hat{x})$, which contains $(-x)/u$. Thus, $(\widehat{M},+)$ is an abelian group. The neutral element $\hat{1}$ of the groupoid $(\widehat{M},\cdot)$ is the class of fractions of the form t/t $(t\neq 0)$, and the inverse of the class $\hat{x}\neq 0$, containing the fraction x/u $(x,u\neq 0)$, is the class $\hat{x}^{-1}$, which contains u/x. Everything stated in this paragraph, including distributivity $\hat{x}(\hat{y}+\hat{z})=\hat{x}\hat{y}+\hat{x}\hat{z}$, can be verified directly.

Thus, for the algebraic system denoted by $\widehat{P}=(\widehat{M},+,\cdot)$, all conditions of the field definition have been satisfied (albeit in a different order). Hence, $\widehat{P}$ is a field.

(*Embedding of K into $\widehat{P}$*) Recall that a *homomorphism* is a mapping between two algebraic structures of the same type that *preserves their operations*. A *bijective homomorphism* is called an *isomorphism* since in this case the inverse mapping is also a homomorphism. Now, let's select a subset $\widehat{M}_K\subset\widehat{M}$ consisting of classes of fractions of the form xt/t $(t\neq 0)$, and show that the system $\widehat{K}=(\widehat{M}_K+,\cdot)$, with the same (induced) operations of addition and multiplication as in the field $\widehat{P}$, is a subring isomorphic to the original ring K. To establish this, it suffices to verify that $(\widehat{M}_K,+)$ and $(\widehat{M}_K,\cdot)$ are groupoids, and to construct a bijection $f: M\to\widehat{M}_K$ such that f is simultaneously an isomorphism from $(\widehat{M},+)$ onto $(\widehat{M}_K,+)$ and from $(\widehat{M},\cdot)$ onto $(\widehat{M}_K,\cdot)$.

If $\hat{x},\hat{y}\in\widehat{M}$, $xt/t\in\hat{x}$ and $yt/t\in\hat{y}$, then the class $\hat{x}+\hat{y}$ contains the fraction
$$(xt\cdot t'+t\cdot yt')/(t\cdot t')=(x+y)\cdot tt'/tt',$$
and therefore, also belongs to $\widehat{M}_K$. Similarly, the class $\hat{x}\cdot\hat{y}$ contains the fraction
$$(xt\cdot yt')/(t\cdot t')=xy\cdot tt'/tt',$$

whence $\hat{x} \cdot \hat{y} \in \widehat{M}_K$. The class $-\hat{x}$ contains the fraction $(-x)t/t$ and consequently $-\hat{x} \in \widehat{M}_K$. We now define $f(x) = \hat{x}$ for any $x \in M$, where x is the class of fractions of the form xt/t. We will show that the mapping $f: M \to \widehat{M}_K$ is an isomorphism of the rings K and $\widehat{K}$.

Injectivity: From $\hat{x} = \hat{y}$, i.e., $xt/t = yt'/t'$ for any fractions $xt/t \in \hat{x}$ and $yt'/t' \in \hat{y}$, it follows that $xtt' = ytt'$. By cancelling $tt' \neq 0$, we obtain $x = y$. Hence, f is injective.

Surjectivity: Every element of $\widehat{M}_K$ is of the form xt/t for some $x \in M$.

Homomorphism: From the definitions of addition and multiplication in $\widehat{M}$, one checks that

$$f(x+y) = f(x) + f(y) \text{ and } f(xy) = f(x) \cdot f(y).$$

But these properties were already established when proving that $(\widehat{M}_K, +)$ and $(\widehat{M}_K, \cdot)$ form groupoids. Therefore, f is an isomorphism of rings, and $K \cong \widehat{K} \subseteq \widehat{P}$.

Thus, the ring $K = (M, +, \cdot)$ is isomorphic to the subring $\widehat{K} = (\widehat{M}_K +, \cdot)$ of the field $\widehat{P} = (\widehat{M}, +, \cdot)$.

(*Uniqueness*) Let P be the minimal field containing K as a subring. For any $x, u \in M$ with $u \neq 0$, in P there exists an element xu^{-1} (not necessarily in K); we assign to it the corresponding class, namely, the element of the field $\widehat{P}$ representing the fraction x/u. This correspondence between certain elements of P and all of $\widehat{P}$ is well defined since the equality $xu^{-1} = yv^{-1}$ is equivalent to $xv = yu$, that is, the equality of the classes containing fractions x/u and y/v. Moreover, the sum and product in P correspond to the sum and product of their images in $\widehat{P}$ since $xu^{-1} + yv^{-1} = (xv + uy)(uv)^{-1}$ and $xu^{-1} \cdot yv^{-1} = (xy)(uv)^{-1}$. Consequently, $\widehat{P}$ is isomorphic to some subfield of P, and by minimality, P itself must be of this form. Thus, all minimal fields containing the given ring K as a subring are isomorphic to the field $\widehat{P}$, and hence to each other.

(*Conclusion*) Every integral ring K embeds into a field $\widehat{P}$, called its field of fractions. Any two minimal fields containing K as a subring are isomorphic to $\widehat{P}$. •

If K_0 is an arbitrary associative and commutative ring without zero divisors (i.e., an integral ring), then the polynomial ring $K_0[x]$ has the same properties. The minimal field containing $K_0[x]$ as a subring is called the *field of rational functions over K_0*.

Questions.

1. What is the field of rational functions over a trivial ring? Over the finite field $\mathbb{Z}_2$?
2. Can the field of rational functions over a nontrivial ring K_0 (an integral ring) be obtained by first embedding K_0 into a field P_0 and then constructing the field of rational functions over P_0?

We now assume that for the initial ring K_0 we have taken a field P. It is natural to view the field of rational functions over P as its simple transcendental extension, denoted by $P(x)$, since the element x does not satisfy any algebraic equation over P (why?). The elements of $P(x)$ are fractions of the form $A(x)/B(x)$, $A(x), B(x) \in P[x]$, $B(x) \neq 0$, considered up to cancelation of common factors in the numerator and denominator. Note the distinction between $P(x)$ and $P[x]$. The established expression "rational function" for such fractions is somewhat misleading, because two equal fractions (i.e., the same element of the field $P(x)$) do not necessarily coincide as functions of the variable $x \in P$. For example,

$$x(x-1)/[(x^3+1)(x-1)] = x(x+2) \,/\, [(x^3+1)(x+2)]$$

in the sense of equality of fractions over the field $\mathbb{R}$ since

$$x(x-1) \cdot (x^3+1)(x+2) = (x^3+1)(x-1) \cdot x(x+2).$$

However, when consider as functions, they differ: at $x = 1$, the left-hand side is undefined, while the right-hand side takes the value ½; at $x = -2$ the left-hand side equals $2/7$, while the right-hand side is undefined. Such discrepancies occur only at finitely many values of x. After complete reduction of fractions, the resulting functions coincide wherever they are defined (or fail to exist at the same points). Hence, in mathematical analysis, the distinction between "rational function" and "fraction" is not particularly problematic. To simplify notation, we shall omit the explicit "(x)" in the symbols (in capital letters) of polynomials whenever the same variable is understood, but not a specific value.

Let $B \in P[x]$ be a polynomial of positive degree, irreducible over the field P. A fraction of the form A/B^r, $A \in P[x]$, $\deg A < \deg B$, $r \in \mathbb{N}$, is called a *simple fraction with respect to B*. By convention, we also include the zero element of the field $P(x)$ among such fractions, which may be regarded as the special case $\deg A < 0$.

Theorem 28. (*Partial Fractions over a Field*) *Let $C/S \in P(x)$, where $S = aB_1^{r_1}B_2^{r_2} \dots B_k^{r_k}$ is the factorization of the polynomial S into irreducible monic polynomials $B_1, B_2, \dots, B_k$ with exponents $r_1, r_2, \dots, r_k$ and $a \in P$. Then the fraction C/S can be expressed, in one and only one way, as a sum of a polynomial and simple fractions with respect to $B_1, B_2, \dots, B_k$, whose denominators are among*

$$B_1, B_1^2, \dots, B_1^{r_1}, B_2, B_2^2, \dots, B_2^{r_2}, \dots, B_k, B_k^2, \dots, B_k^{r_k}.$$

Proof.

(*Reduction of the scalar factor*) Because $a \neq 0$ (otherwise $S = 0$, which is impossible), we may assume without loss of generality that $a = 1$. Indeed, if $a \neq 1$, we replace the numerator C by $a^{-1}C$.

(*Existence*) We first prove the existence of the desired representation, beginning with the special case $k = 1$, i.e., when the denominator of the given fraction has the form $S = B^r$.

Dividing C by B, we obtain $C = BQ_{r-1} + A_r$, $\deg A_r < \deg B$, where the notation of the quotient Q_{r-1} and remainder A_r will soon prove useful. Hence,

$$C/S = C/B^r = Q_{r-1}/B^{r-1} + A_r/B^r,$$

where the second term is a simple fraction (possibly zero). Dividing Q_{r-1} further by B, we obtain $Q_{r-1} = BQ_{r-2} + A_{r-1}$, $\deg A_{r-1} < \deg B$, whence

$$Q_{r-1}/B^{r-1} = Q_{r-2}/B^{r-2} + A_{r-1}/B^{r-1}.$$

Therefore,

$$C/S = Q_{r-2}/B^{r-2} + A_{r-1}/B^{r-1} + A_r/B^r,$$

where the last two terms are simple fractions. Continuing this process, we arrive at

$$C/S = Q_1/B + A_2/B^2 + \cdots + A_{r-1}/B^{r-1} + A_r/B^r.$$

Finally, dividing Q_1 by B, we obtain

$$C/S = Q + A_1/B + A_2/B^2 + \cdots + A_r/B^r,$$

where $Q, A_1, A_2, \ldots, A_r \in P[x]$, and all fractions on the right-hand side are simple. This proves existence for $k = 1$.

(*Induction on k*) Let $k > 1$. Recall that our convention regarding the omission of the multiplication symbol sometimes makes it possible to write products without brackets. In the expansion

$$S = B_1^{r_1} \cdot B_2^{r_2} \ldots B_k^{r_k}$$

the two factors $B_1^{r_1}$and $B_2^{r_2} \ldots B_k^{r_k}$ are relatively prime. The reason is that the B_i are distinct irreducible, $\gcd\left(B_1, B_2^{r_2} \ldots B_k^{r_k}\right) = 1$, hence also $\gcd\left(B_1^{r_1}, B_2^{r_2} \ldots B_k^{r_k}\right) = 1$. By Corollary 4, there exist polynomials $U, V \in P[x]$ such that

$$U \cdot B_1^{r_1} + V \cdot B_2^{r_2} \ldots B_k^{r_k} = 1.$$

It follows that $C = CV \cdot B_2^{r_2} \ldots B_k^{r_k} + CU \cdot B_1^{r_1}$, and hence

$$C/S = CV/B_1^{r_1} + CU/\left(B_2^{r_2} \ldots B_k^{r_k}\right).$$

The first term, as we have already established, admits a representation as the sum of a polynomial and simple fractions with respect to B_1. Proceeding with the second term in the same way as for the case $k = 2$, and analogously treating the original fraction C/S when $k > 2$, after k steps – and collecting all polynomials of type Q – we arrive at the desired representation (the general form of which will be introduced a little later). This completes existence by induction.

(*Uniqueness*) Suppose that

$$C/S = Q + \cdots + A/B_i^r + \cdots = \tilde{Q} + \cdots + \tilde{A}/B_i^r + \cdots$$

are two different representations, where $A \neq \tilde{A}$ $(1 \leq i \leq k, 1 \leq r \leq r_i)$. Thus, there exist two simple fractions with the same denominator B_i^r, but different numerators. Such fractions must necessarily occur (why?). Among all such pairs, let's choose one with the largest exponent r. Multiplying both sides of

$$\left(A - \tilde{A}\right)/B_i^r = \tilde{Q} - Q + \cdots$$

by

$$B_1^{r_1} \ldots B_{i-1}^{r_{i-1}} B_i^{r-1} B_{i+1}^{r_{i+1}} \ldots B_k^{r_k},$$

we obtain an equality in which the left-hand side still contains simple fractions, while the right-hand side becomes a nonzero polynomial – a contradiction (why?).

Hence, uniqueness holds, and the theorem is proven. •

Remark 28. To represent the original fraction as the sum of a polynomial and simple fractions in general form, note that the numerators of the terms associated with different denominators B_i are distinct. Therefore, their notation requires two indices: A_{ij} denotes the numerator corresponding to the denominator B_i^j for $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, r_i$. The representation, whose existence and uniqueness are guaranteed by Theorem 28, takes the general form

$$C/S = Q + A_{11}/B_1 + A_{12}/B_1^2 + \cdots + A_{1r_1}/B_1^{r_1} +$$
$$+ A_{21}/B_2 + A_{22}/B_2^2 + \cdots + A_{2r_2}/B_2^{r_2} + \quad (8)$$
$$\cdots\cdots\cdots\cdots$$
$$+ A_{k1}/B_k + A_{k2}/B_k^2 + \cdots + A_{kr_k}/B_k^{r_k},$$

the indices are read digit by digit: e.g., A_{11} is read as "A one one" (not "A eleven"), and A_{1r_1} as "A one r_1". By convention, the indices are not separated by a comma (sometimes one writes $A_{(1,r_1)}$, but not $A_{1,1}$). In (8), $Q, A_{ij} \in P[x]$, $\deg A_{ij} < \deg B_i$ for all $i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, r_i$. If $A_{ij} = 0$, then the corresponding fraction represents the zero element of the field $P(x)$ and that term disappears from (8) when its numerator vanishes.

In the particular case $P = \mathbb{R}$, considering the reducibility of polynomials over the real field, we obtain the following result.

Corollary 7. *Every fraction $C/S \in P(x)$, where*

$$S = a(x - c_1)^{r_1} \ldots (x - c_k)^{r_k}(x^2 + p_1x + q_1)^{s_1} \ldots (x^2 + p_lx + q_l)^{s_l},$$

with

$a, c_1, \ldots, c_k, p_1, \ldots, q_l \in \mathbb{R}$, $r_1, \ldots, r_k, s_1, \ldots, s_l \in \mathbb{N}$, $p_i^2 - 4q_i < 0$ *for all* $i = 1, \ldots, l$,

is uniquely representable as a sum of a polynomial over $\mathbb{R}$ *and simple fractions*:

$$C/S = Q + \alpha_{11}/(x - c_1) + \alpha_{12}/(x - c_1)^2 + \cdots + \alpha_{1r_1}/(x - c_1)^{r_1} +$$
$$+ \cdots + \alpha_{k1}/(x - c_k) + \alpha_{k2}/(x - c_k)^2 + \cdots + \alpha_{kr_k}/(x - c_k)^{r_k} +$$
$$+ (\beta_{11}x + \gamma_{11})/(x^2 + p_1x + q_1) + (\beta_{12}x + \gamma_{12})/(x^2 + p_1x + q_1)^2 +$$
$$+ \cdots + \left(\beta_{1s_1}x + \gamma_{1s_1}\right)/(x^2 + p_1x + q_1)^{s_1} + \cdots +$$

$$+ (\beta_{l1}x + \gamma_{l1})/(x^2 + p_l x + q_l) + (\beta_{l2}x + \gamma_{l2})/(x^2 + p_l x + q_l)^2 +$$
$$+ \dots + (\beta_{ls_l}x + \gamma_{ls_l})/(x^2 + p_l x + q_l)^{s_l},$$

where all coefficients
$\alpha_{11}, \dots, \alpha_{kr_k}, \beta_{11}, \dots, \beta_{ls_l}, \gamma_{11}, \dots, \gamma_{ls_l} \in \mathbb{R}$. •

In practice, it is convenient to find such a representation not by the general scheme of Theorem 28, but rather by the *method of undetermined coefficients*. To avoid cumbersome formulas, we illustrate the problem with an example.

Example 29. (*Partial fraction decomposition*) Consider

$$C = 8x^7 + 4x^6 + 2x^5 + 11x^4 + 15x^3 + 9x^2 + 10x + 11,$$

$S = 8x^6 + 4x^5 - 2x^4 - 5x^3 - 2x^2 + 4x - 1.$

(*Division of C by S*)

$$\frac{C}{S} = x + \frac{4x^5+16x^4+17x^3+5x^2+11x+11}{8x^6+4x^5+2x^4-5x^3-2x^2+4x-1}.$$

(*Factorization*)

$$= x + \frac{(x+1)(4x^4+12x^3+5x^2+11)}{(x+1)(2x-1)^3(x^2+x+1)}.$$

(*Cancellation and monic denominator*)

$$= x + \frac{\frac{1}{2}x^4+\frac{3}{2}x^3+\frac{5}{8}x^2+\frac{11}{8}}{\left(x-\frac{1}{2}\right)^3(x^2+x+1)}.$$

Here we not only reduced the fraction by cancelling $(x + 1)$, but also made the denominator monic by multiplying the numerator by $1/8$. In the resulting fraction we have

$$k = l = 1,\ c_k = c_1 = \frac{1}{2},\ r_k = r_1 = 3,\ p_l = p_1 = q_l = q_1 = 1, s_l = s_1 = 1.$$

We now write the fraction as a sum of simple fractions with undetermined coefficients (not yet defined explicitly):

$$\frac{\frac{1}{2}x^4+\frac{3}{2}x^3+\frac{5}{8}x^2+\frac{11}{8}}{\left(x-\frac{1}{2}\right)^3(x^2+x+1)} = \frac{\alpha}{\left(x-\frac{1}{2}\right)} + \frac{\beta}{\left(x-\frac{1}{2}\right)^2} + \frac{\gamma}{\left(x-\frac{1}{2}\right)^3} + \frac{\delta x+\varepsilon}{(x^2+x+1)}.$$

(*Reduction to a common denominator*) Multiplying both sides by $\left(x-\frac{1}{2}\right)^3(x^2+x+1)$, we obtain the polynomial identity

$$\alpha\left(x-\frac{1}{2}\right)^2(x^2+x+1) + \beta\left(x-\frac{1}{2}\right)(x^2+x+1) + \gamma(x^2+x+1) + (\delta x+\varepsilon)\left(x-\frac{1}{2}\right)^3$$
$$= \frac{1}{2}x^4 + \frac{3}{2}x^3 + \frac{5}{8}x^2 + \frac{11}{8}$$

(*Expansion and coefficient comparison*) Expanding the left-hand side, grouping by powers of x, and equating coefficients gives the system

$$\alpha + \delta = 1/2,$$
$$\beta - \frac{3}{2}\delta + \varepsilon = 3/2,$$
$$\frac{1}{4}\alpha + \frac{1}{2}\beta + \gamma + \frac{3}{4}\delta - \frac{3}{2}\varepsilon = 5/8,$$
$$-\frac{3}{4}\alpha + \frac{1}{2}\beta + \gamma - \frac{1}{8}\varepsilon = 0,$$
$$\frac{1}{4}\alpha - \frac{1}{2}\beta + \gamma - \frac{1}{8}\varepsilon = 11/8,$$

(*Solving the system*)

$$\alpha = \frac{3}{2}, \beta = 0, \gamma = 1, \delta = -1, \varepsilon = 0.$$

(*Final representation*)

$$\frac{C}{S} = x + \frac{\frac{3}{2}}{x-\frac{1}{2}} + \frac{1}{\left(x-\frac{1}{2}\right)^3} - \frac{x}{x^2+x+1} = x + \frac{3}{2x-1} + \frac{8}{(2x-1)^3} - \frac{x}{x^2+x+1}.$$

Remarks 29. In general, the system of equations for the unknowns $\alpha_{11}, \dots, \beta_{ls_l}, \gamma_{ls_l}$ will always have a unique solution. This follows automatically from the existence and uniqueness of the partial fraction decomposition (Corollary 7). (As an exercise, one may verify that the number of equations coincides with the number of unknowns, though this fact alone does not guarantee solvability or uniqueness; the general theory of such systems is studied in linear algebra courses.)

A separate justification for the process described in Corollary 7 is unnecessary, as it constitutes an immediate consequence of the proof of Theorem 28 itself. Only one subtle point (which, apparently, not everyone noticed) requires clarification. Specifically, in equality (8) we immediately took the quotient of dividing C by S as Q. Then, in the course of the proof, this polynomial gradually developed from many terms. One might ask: could this process introduce any additional terms beyond the original quotient?

In fact, from the uniqueness of the representation $C = SQ + R$, $\deg R < \deg S$, it follows that the representation $C/S = Q + R/S$ is also unique, where the fraction R/S is *proper* (the degree of the numerator is less than that of the denominator). Since every simple fraction is proper, the only gap in the proof is to confirm that the *sum of* *proper fractions remains a proper fraction*. We recommend proving this general statement independently, and, as a simpler exercise, verifying that the product of proper fractions is also a proper fraction.

Question. Will the ratio of two proper fractions itself be a proper fraction?

Finally, in the context of mathematical analysis – where Theorem 28 and Corollary 7 are applied primarily to the integration of rational functions – other methods for decomposing a proper fraction into a sum of simple fractions are also studied.

MATRICES

Consider a nine-story building with five entrances. Let a_{ij} denote the total number of residents in those apartments located on the i-th floor whose doors open to the stairwell of the j-th entrance. The distribution of residents can then be represented by the matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{91} & a_{92} & a_{93} & a_{94} & a_{95} \end{bmatrix},$$

which has nine rows and five columns. This matrix describes the complete distribution of residents in the building according to stairwell. The entries a_{ij} need not be numerical; they may, for example, represent sets (such as lists of residents).

(*General definition*) A $(p \times q)$-matrix over a set M is a function

$$A: \{1, 2, \dots, p\} \times \{1, 2, \dots, q\} \longrightarrow M,$$

which assigns to each ordered pair (i, j) with $i \in \{1, 2, \dots, p\}$, $j \in \{1, 2, \dots, q\}$ an element $a_{ij} \in M$. The notation for a matrix is written as

$$A = (a_{ij}) = (a_{ij})_{p\times q} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix},$$

where the indices p and q, denoting the numbers of rows and columns, may be omitted when the context makes them clear. Two $(p \times q)$-matrices $A = (a_{ij})_{p\times q}$ and $B = (b_{ij})_{p\times q}$ are said to be *equal*, written $A = B$, if and only if $a_{ij} = b_{ij}$ for all $i \in \{1, 2, \dots, p\}$ and $j \in \{1, 2, \dots, q\}$. It is not meaningful to speak of equality or inequality between matrices of different sizes (even if they share the same p or q).

Let $K = (M, +, \cdot)$ be a ring. The sum of two $(p \times q)$-matrices with entries in K (that is, in M) is defined *componentwise*, meaning entry by entry. If $A = (a_{ij})_{p\times q}$, $B = (b_{ij})_{p\times q}$, then

$$A + B = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ a_{21} & a_{22} & \cdots & a_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pq} \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1q}+b_{1q} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2q}+b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}+b_{p1} & a_{p2}+b_{p2} & \cdots & a_{pq}+b_{pq} \end{bmatrix} = (a_{ij} + b_{ij})_{p\times q}.$$

With respect to addition, the set of all $(p \times q)$-matrices (with fixed p and q) over K forms an abelian group. The neutral element is the zero $(p \times q)$-matrix, denoted by $(0_{ij})_{p\times q}$, whose entries are all equal to the zero element of the ring K. For any matrix $A = (a_{ij})$, the additive inverse is given by $-A = (-a_{ij})$. Defining the product of matrices is a more subtle matter, but it is also considerably more interesting.

The *componentwise* multiplication of matrices, $(a_{ij})_{p\times q} \cdot (b_{ij})_{p\times q} = \left((a_{ij}) \cdot (b_{ij})\right)_{p\times q}$ gives rise to a ring of $(p \times q)$-matrices that inherits the essential properties of the underlying ring: commutativity, associativity, the existence of a unity, the absence of zero divisors, and the invertibility of nonzero elements. This componentwise product, introduced by J. Hadamard, is well known in combinatorial analysis. However, the corresponding matrix ring, considered as an algebraic system, is of particular interest not because it involves new objects such as matrices, but because its structure can be more naturally explained through the already well-studied framework of polynomials.

Let's suppose that, for infinite sequences of elements of a ring K, we define multiplication in the same componentwise fashion as addition:

$$(a_0, a_1, a_2, \dots) \cdot (b_0, b_1, b_2, \dots) = (a_0 b_0, a_1 b_1, a_2 b_2, \dots).$$

Clearly, this again produces a ring that is structurally as "good" as the original ring K (a fact that can be checked and verify). However, in this construction the situation is quite limited: if an element d is given by an expression such as $ab + c$, then the i-th term of the resulting sequence satisfies $d_i = a_i b_i + c_i$, $i = 0, 1, 2, \dots$ so that each term depends only on the corresponding terms of the original sequences. No interaction occurs between different indices. In this sense, the resulting ring of sequences is nothing more than a "mechanical mixture" of infinitely many independent copies of the ring K.

By contrast, consider the alternative definition of multiplication:

$$(a_0, a_1, a_2, \dots) \cdot (b_0, b_1, b_2, \dots) = (a_0 b_0,\ a_0 b_1 + a_1 b_0,\ a_0 b_2 + a_1 b_1 + a_2 b_0, \dots).$$

Here, the i-th term of the product (for $i > 0$) is influenced not only by the i-th terms of the factors but also by earlier terms. This construction leads precisely to the polynomial ring $K[x]$, which represents a qualitatively new algebraic system. Such a fruitful definition of multiplication is not merely a clever invention, but rather a reflection of deeper quantitative principles of the real world – principles that find natural mathematical expression in the operations on polynomials. As other examples – such as the field $\mathbb{Q}(\sqrt{2})$, the field of complex numbers, and the subsequent historical development of algebra – demonstrate, such "linking of different components" proves fruitful precisely when it is incorporated into the definition of multiplication rather than addition. Naturally, however, statements of this kind should not be interpreted categorically since doing so would only restrict scientific creativity.

In order to obtain a "linked" multiplication of matrices rather than an entry-by-entry product, we consider the following example.

Example 30. Suppose there are 3 products from which 5 different dishes are prepared by mixing them in various proportions. Let a_{ij} denote the number of grams of the *j*-th product required to prepare one gram of the *i*-th dish ($i = 1, 2, 3, 4, 5$; $j = 1, 2, 3$). Furthermore, each gram of the *j*-th product contains b_{j1} grams of proteins, b_{j2} grams of water, b_{j3} grams of fat and b_{j4} grams of carbohydrates (these are referred to as the 1st, 2nd, 3rd, and 4th components of the product, respectively). The task is to determine the component composition of each dish.

To find the amount c_{ik} (in grams) of the *k*-th component in the *i*-th dish ($k = 1, 2, 3, 4$), one simply sums the contributions of the three products. For instance, the fat content ($k = 3$) of the second dish ($i = 2$) is obtained by adding up the quantities contributed by the first, second, and third products: $c_{23} = a_{21} b_{13} + a_{22} b_{23} + a_{23} b_{33}$. In general, the entries of the desired matrix

$$C = (c_{ik})_{5\times 4}$$

are obtained from the given matrices

$$A = (a_{ij})_{5\times 3},\ B = (b_{jk})_{3\times 4},$$

according to the rule

$$c_{ik} = a_{i1} b_{1k} + a_{i2} b_{2k} + a_{i3} b_{3k},\ i = 1, \dots, 5,\ k = 1, \dots, 4. \quad (9)$$

Example 31. Here is a more abstract instance. Let the variables x_1, x_2, x_3, x_4, x_5 be expressed through the variables y_1, y_2, y_3 by the equalities

$$\begin{aligned} x_1 &= a_{11} y_1 + a_{12} y_2 + a_{13} y_3, \\ x_2 &= a_{21} y_1 + a_{22} y_2 + a_{23} y_3, \\ x_3 &= a_{31} y_1 + a_{32} y_2 + a_{33} y_3, \\ x_4 &= a_{41} y_1 + a_{42} y_2 + a_{43} y_3, \\ x_5 &= a_{51} y_1 + a_{52} y_2 + a_{53} y_3, \end{aligned} \quad (10)$$

where a_{ij} are constants (for example, real numbers). The variables y_1, y_2, y_3, in turn, are expressed through z_1, z_2, z_3, z_4 as follows:

$$\begin{aligned} y_1 &= b_{11} z_1 + b_{12} z_2 + b_{13} z_3 + b_{14} z_4, \\ y_2 &= b_{21} z_1 + b_{22} z_2 + b_{23} z_3 + b_{24} z_4, \\ y_3 &= b_{31} z_1 + b_{32} z_2 + b_{33} z_3 + b_{34} z_4. \end{aligned} \quad (11)$$

It is required to express the x-variables directly in terms of the z-variables. The solution is obvious: replace all the y-variables in (10) with their expressions from (11), expand the brackets and group the z-variables. As a result, we obtain

$$\begin{aligned} x_1 &= c_{11} z_1 + c_{12} z_2 + c_{13} z_3 + c_{14} z_4, \\ x_2 &= c_{21} z_1 + c_{22} z_2 + c_{23} z_3 + c_{24} z_4, \\ &\vdots \qquad \vdots \\ x_5 &= c_{51} z_1 + c_{52} z_2 + c_{53} z_3 + c_{54} z_4, \end{aligned} \quad (12)$$

where the entries of the matrix $C = (c_{ik})$ are expressed in terms of the entries of the matrices $A = (a_{ij})$ and $B = (b_{jk})$ by the same equalities (9) of Example 30.

Of course, it was not the problem of rational nutrition, but the theory of *linear transformations* that served as the basis for the creation of the theory of matrices, in which precisely this method of obtaining the matrix C from A and B plays a key role.

Let $A = (a_{ij})_{p\times q}$ and $B = (b_{jk})_{q\times r}$ be matrices over a ring K. The sizes of A and B are chosen so that the number of columns of the first matrix equals the number of rows of the second. The product $A \cdot B$ is defined to be the matrix $C = (c_{ik})_{p\times r}$ with entries

$$\sum_{j=1}^{q} a_{ij} b_{jk}\ (i = 1, 2, \dots, p;\ k = 1, 2, \dots, r).$$

Whenever matrices are multiplied without explicitly indicating their sizes, it is always assumed that their dimensions are compatible as described above; otherwise, the matrix product is not defined. Alternatively, without introducing C explicitly, the product of a $(p \times q)$-matrix A by a $(q \times r)$-matrix B can be written as

$$A \cdot B = (a_{ij})_{p\times q} \cdot (b_{jk})_{q\times r} = \left(\sum_{j=1}^{q} a_{ij}b_{jk}\right)_{p\times r},$$

which is, as expected, a $(p \times r)$-matrix.

Matrix multiplication over a ring K is not commutative, even when the ring K itself is commutative.

First, if $A = (a_{ij})_{p\times q}$ and $B = (b_{ij})_{p\times q}$ with $p \neq q$, then equality $A \cdot B = B \cdot A$ does not even make sense. Indeed:

- The product $A \cdot B$ is only defined when the number of columns of A equals the number of rows of B.
- If A is $p \times q$ and B is $p \times q$ then $A \cdot B$ is undefined unless $q = p$.
- Even if $p = r$, the products

$$(a_{ij})_{p\times q} \cdot (b_{jk})_{q\times p} = (A \cdot B)_{p\times p} \text{ and } (b_{ij})_{q\times p} \cdot (a_{jk})_{p\times q} = (B \cdot A)_{q\times q}$$

are of different sizes ($p \times p$ vs. $q \times q$) and thus cannot be equal.

Second, even when $p = q = r$ so that both products exist and have the same size, they are *not necessarily equal*, as shown by the following example.

Example 32. Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

over the ring $\mathbb{Z}$ ($p = q = r = 2$). Then

$$AB = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, BA = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix},$$

which are not equal.

Therefore, matrix multiplication is not commutative. Moreover, while both distributive laws

$$A(B + C) = AB + AC \text{ and } (B + C\,)A = BA + CA$$

hold, they require proof. We shall prove only the first identity since the argument for the second is similar.

Claim 1. *Let $A = (a_{ij})_{p\times q}$, $B = (b_{jk})_{q\times r}$ and $C = (c_{jk})_{q\times r}$ be matrices over a ring K. Then*

$$A(B + C) = AB + AC$$

whenever the sizes make the corresponding addition and multiplication defined.

Proof. Matrix addition is defined only when the summands have the same dimensions. Thus,

$$B + C = (b_{jk} + c_{jk})_{q\times r}.$$

Then

$$A(B + C) = \left(\sum_{j=1}^{q} a_{ij}(b_{jk}+c_{jk})\right)_{p\times r} = \left(\sum_{j=1}^{q}(a_{ij}b_{jk} + a_{ij}c_{jk})\right)_{p\times r}$$

$$= \left(\sum_{j=1}^{q} a_{ij}b_{jk} + \sum_{j=1}^{q} a_{ij}c_{jk}\right)_{p\times r} = \left(\sum_{j=1}^{q} a_{ij}b_{jk}\right)_{p\times r} + \left(\sum_{j=1}^{q} a_{ij}c_{jk}\right)_{p\times r} = AB + AC.$$

This equality holds because the ring K satisfies the distributive law $a(b + c) = ab + ac$ •

Claim 2. *Let $A = (a_{ij})_{p\times q}$, $B = (b_{jk})_{q\times r}$, $C = (c_{kl})_{r\times s}$ be matrices over a ring K whose multiplication is associative. Then $A \cdot BC = AB \cdot C$ whenever the sizes make the products defined.*

Proof.

$$A \cdot BC = (a_{ij})_{p\times q} \cdot \left(\sum_{k=1}^{r} b_{jk}c_{kl}\right)_{q\times s} = \left(\sum_{j=1}^{q} a_{ij}\left(\sum_{k=1}^{r} b_{jk}c_{kl}\right)\right)_{p\times s}$$

$$= \left(\sum_{j=1}^{q}\sum_{k=1}^{r} a_{ij} \cdot b_{jk}c_{kl}\right)_{p\times s} = \left(\sum_{j=1}^{q}\sum_{k=1}^{r} a_{ij}b_{jk} \cdot c_{kl}\right)_{p\times s} = \left(\sum_{k=1}^{r}\sum_{j=1}^{p} a_{ij}b_{jk} \cdot c_{kl}\right)_{p\times s}$$

$$= \left(\sum_{k=1}^{r}\left(\sum_{j=1}^{p} a_{ij}b_{jk}\right)c_{kl}\right)_{p\times s} = \left(\sum_{j=1}^{p} a_{ij}b_{jk}\right)_{p\times r} \cdot (c_{kl})_{r\times s} = AB \cdot C$$

•

Remark 30. (*Entrywise explanation*) Fix $i \in \{1, \ldots, p\}$ and $l \in \{1, \ldots, s\}$. Then:

$(A \cdot BC)_{il} = \sum_{j=1}^{q} a_{ij}(BC)_{jl}$ (definition of product)

$= \sum_{j=1}^{q} a_{ij}\left(\sum_{k=1}^{r} b_{jk}c_{kl}\right)$ (definition of $(BC)_{jl}$)

$= \sum_{j=1}^{q}\sum_{k=1}^{r} a_{ij} \cdot b_{jk}c_{kl}$ (expand the double sum)

$= \sum_{j=1}^{q}\sum_{k=1}^{r} a_{ij}b_{jk} \cdot c_{kl}$ (associativity in K)

$= \sum_{k=1}^{r}\sum_{j=1}^{p} a_{ij}b_{jk} \cdot c_{kl}$ (finite sums can be reordered)

$= \sum_{k=1}^{r}\left(\sum_{j=1}^{p} a_{ij}b_{jk}\right)c_{kl}$ (group the inner sum)
$= \sum_{k=1}^{r}(AB)_{ik}\, c_{kl}$ (definition of $(AB)_{ik}$)
$= (AB\cdot C)_{il}$ (definition of product).

Since this holds for all i, l, we conclude $A\cdot BC = AB\cdot C$

(*Key properties used*)

- Associativity of multiplication in K: $a_{ij}\cdot b_{jk}c_{kl} = a_{ij}b_{jk}\cdot c_{kl}$.
- Commutativity and associativity of addition in K: finite sums can be regrouped and reordered.

Due to the associativity of matrix multiplication, the solution to the previous problem of variable transformations (see Example 31) can be written in a very compact form.

Namely, along with the coefficient matrices

$$A = (a_{ij})_{5\times 3},\ B = (b_{jk})_{3\times 4},\ C = (c_{kl})_{5\times 4},$$

we also introduce the variable column matrices:

$$X = (x_j)_{5\times 1} = \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\end{bmatrix},\ Y = (y_k)_{3\times 1} = \begin{bmatrix} y_1\\ y_2\\ y_3\end{bmatrix},\ Z = (z_l)_{4\times 1} = \begin{bmatrix} z_1\\ z_2\\ z_3\\ z_4\end{bmatrix}.$$

Then the system of equations (10) can be written equivalently as the single matrix equation

$$X = AY.\ (10A)$$

Similarly, the system (11) is equivalent to the matrix equation

$$Y = BZ.\ (11A)$$

Substituting the expression for Y from (11A) into (10A) and using the associativity of matrix multiplication, we obtain

$$X = A\cdot BZ = AB\cdot Z = CZ,$$

where $C = AB$.

But the equality $X = CZ$ is precisely the matrix form of the system (12).

Strictly speaking, the entries of these matrices do not belong to the original ring K over which the coefficient matrices A and B are defined, but rather to the polynomial ring

$$K[x_1, \dots, x_5, y_1, y_2, y_3, z_1, \dots, z_4]$$

in twelve variables over K. However, if we interpret these variables as taking values in K, then all of the above is valid without additional justification.

If the ring K has unity 1, then the identity $(p\times p)$-matrix

$$I_p = \begin{bmatrix} 1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}_{p\times p}$$

serves as the *left neutral element* with respect to multiplication for all $(p\times q)$-matrices, and the identity $(q\times q)$-matrix I_q serves as the *right neutral element*. Indeed, for any $(p\times q)$-matrix A over K, one has

$$I_p\cdot (A)_{p\times q} = (A)_{p\times q} \text{ and } (A)_{p\times q}\cdot I_q = (A)_{p\times q}.$$

However, when $p \neq q$ there is *no two-sided neutral element*, because the set of $(p\times q)$-matrices does not form a groupoid under matrix multiplication. This resolves the apparent contradiction with the remark on the uniqueness of the neutral element in a groupoid (Theorem 1).

Example 33. The following calculations illustrate that the product of two nonzero matrices over K may be the zero matrix even when the ring K itself has no zero divisors:

$$\begin{bmatrix} 0 & 1\\ 0 & 0\end{bmatrix}\cdot\begin{bmatrix} 1 & 0\\ 0 & 0\end{bmatrix} = \begin{bmatrix} 0 & 0\\ 0 & 0\end{bmatrix} = (0)_{2\times 2},\ \begin{bmatrix} 1 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix} 0 & 1 & 0\\ 1 & 1 & 0\\ 0 & -1 & 0\end{bmatrix} = \begin{bmatrix} 0 & 0 & 0\end{bmatrix} = (0)_{1\times 3}.$$

We shall present additional properties of matrix multiplication as needed.

From the discussion above, it is evident that the set $M(K)$ of all matrices over an arbitrary ring K (even over a field) does not form a satisfactory algebraic structure of ring type with respect to matrix addition and multiplication. The same holds for any subset consisting of $(p\times q)$-matrices with fixed p and q whenever $p \neq q$.

Only when $p = q$ does the set $M_p(K)$ of all *square* $(p\times p)$*-matrices* form a ring $(M_p(K), +, \cdot)$. For brevity, we shall continue to denote this ring simply by $M_p(K)$. The fact that $M_p(K)$ is indeed a ring – and that properties such as associativity and the existence of a multiplication identity (but not necessarily commutativity or the absence of zero divisors) carry over from the base ring K – can be established by collecting and organizing the relevant statements and examples presented earlier.

The ring $M_1(K)$ is clearly isomorphic to K; in fact, it essentially coincides with K if we ignore the notational distinction between the symbols $(a_{ij})_{1\times 1} = [a_{11}]$ and a_{11}. (You could say this, though it brings to mind O. Henry's remark: "She sat in pink dreams and in a kimono of the same color." And from which novel is this, by the way?)

Before examining the rings $M_p(K)$ in detail, let's first consider several important examples of transformations of $(p \times q)$-matrices achieved by multiplying a given matrix by suitable *transforming matrices*. Throughout, we shall assume that the ring K is associative and has a multiplicative identity 1.

Transformation I. (*Row permutation*, *also known as row swap*)

Suppose that the matrix $A = \left(a_{ij}\right)_{p\times q}$ $(p > 1)$ over a field K needs to be rearranged as follows: Place the first row of A in the i_1-th position, the second row in the i_2-th position ,…, and the p-th row in the i_p-th position. This arrangement can be achieved by multiplying A on the left by a suitable $(p \times p)$ permutation matrix

$$S\begin{bmatrix} 1 & 2 & \cdots & p \\ i_1 & i_2 & \cdots & i_p \end{bmatrix},$$

in which each column contains exactly one entry equals to 1 and $p - 1$ entries equal to 0; the 1 in the first column is located in the i_1-th row, the 1 in the second column is in the i_2-th row, and so on. For example:

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a & b & c & d \\ x & y & z & t \\ \alpha & \beta & \gamma & \delta \end{bmatrix} = \begin{bmatrix} \alpha & \beta & \gamma & \delta \\ a & b & c & d \\ x & y & z & t \end{bmatrix}.$$

where the left factor is the transforming matrix

$$S\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}.$$

Transformation II. (*Row replacement by row sum*)

Consider the following modification of a matrix A: the entries of the *i*-th row are replaced by the sum of themselves and λ times the corresponding entries of the *k*-th row, where $\lambda \in K$ is fixed and $i \neq k$. The *k*-th row itself remains unchanged. In other words, the transformation

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{iq} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kq} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix} \to \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} + \lambda a_{k1} & a_{i2} + \lambda a_{k2} & \cdots & a_{iq} + \lambda a_{kq} \\ & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kq} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pq} \end{bmatrix}$$

is equivalent to left-multiplication of A by the matrix $E_{ik}(\lambda) \in M_p(K)$, which is identical to the identity matrix I_p except that the zero element in the (i, k)-position is replaced by λ. Similarly, right-multiplication of A by the matrix $E_{kj}(\lambda) \in M_q(K)$ (with $k \neq j$) means that the entries of the j-th column are replaced by the sum of themselves and λ times the corresponding entries of the k-th column. For example:

$$\begin{bmatrix} a & b & c & d \\ x & y & z & t \\ \alpha & \beta & \gamma & \delta \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b-d & c & d \\ x & y-t & z & t \\ \alpha & \beta-\delta & \gamma & \delta \end{bmatrix},$$

where the second factor on the left-hand side is the transforming matrix $E_{42}(-1) \in M_4(K)$. However, if in some $(4 \times q)$-matrix it were necessary to replace the second row by its difference with the fourth row, then one would multiply on the left of A by $E_{24}(-1) \in M_4(K)$.

Transformation III. (*Row scaling*)

Multiplying the *i*-th row (i.e., each entry in that row) of a matrix A on the left by $\lambda \in K$ can be accomplished by multiplying the entire matrix A on the left by the matrix $E_{ii}(\lambda) \in M_p(K)$, which is obtained from the identity matrix I_p by replacing the *i*-th diagonal entry 1 with λ. For example:

$$E_{22}(3) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, E_{22}(3) \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 9 & 12 \\ 5 & 6 \end{bmatrix}.$$

Analogously, performing similar operations on the columns of A (instead of its rows) – where multiplication of entries by λ is now performed from the right – is equivalent to multiplying A on the right by the corresponding matrix $E_{jj}(\lambda) \in M_q(K)$, which is obtained from the identity matrix I_q by replacing the *j*-th diagonal entry 1 with λ. These column-transforming matrices differ from the row-transforming ones not only in size (when $p \neq q$) but also in that the roles of rows and columns are interchanged.

Remark 31. (*Elementary transformations*)

Transformations of types I-III applied to the rows or columns of a matrix are called *elementary transformations*.

- The corresponding elementary row transformation matrices are:

$$S\begin{bmatrix} 1 & 2 & \cdots & p \\ i_1 & i_2 & \cdots & i_p \end{bmatrix}, E_{ik}(\lambda) \in M_p(K), E_{ii}(\lambda) \in M_p(K),$$

where S represents a row permutation, $E_{ik}(\lambda)$ (with $i \neq k$) represents adding λ times row k to row i, and $E_{ii}(\lambda)$ represents multiplying row i by λ.

- The corresponding elementary column transformation matrices are:

$$S\begin{bmatrix} i_1 & i_2 & \cdots & i_p \\ 1 & 2 & \cdots & p \end{bmatrix}, E_{kj}(\lambda) \in M_q(K), E_{jj}(\lambda) \in M_q(K),$$

where S represents a column permutation, $E_{kj}(\lambda)$ (with $k \neq j$) represents adding λ times column j to column k, and $E_{jj}(\lambda)$ represents multiplying column j by λ.

These matrices are collectively called *elementary matrices*.

Let's illustrate, with an example, how a sequence of elementary transformations can be expressed as a product of their corresponding transformation matrices.

Example 34. Let the matrix

$$A = \begin{bmatrix} a & b & c & d \\ x & y & z & t \\ 1 & 2 & 3 & 4 \end{bmatrix}_{3\times 4} \quad \text{(over } \mathbb{Z}\text{)}$$

be subjected to the following elementary operations, in order:

1. Add the first row to the second row.
2. Swap the first and third columns.
3. Subtract twice the fourth column from the first column.
4. Multiply the second row by -1.

Only the first operation is performed directly on A; each subsequent operation is applied to the matrix obtained from the previous step.

(*Stepwise transformation*)

$$A \xrightarrow{\text{R2}\leftarrow\text{R2+R1}} \begin{bmatrix} a & b & c & d \\ x+a & y+b & z+c & t+d \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

$$\xrightarrow{\text{swap C1}\leftrightarrow\text{C3}} \begin{bmatrix} c & b & a & d \\ z+c & y+b & x+a & t+d \\ 3 & 2 & 1 & 4 \end{bmatrix}$$

$$\xrightarrow{\text{C1}\leftarrow\text{C1}-2\text{C4}} \begin{bmatrix} c-2d & b & a & d \\ z+c-2(t+d) & y+b & x+a & t+d \\ 3-2(4) & 2 & 1 & 4 \end{bmatrix}$$

$$\xrightarrow{\text{R2}\leftarrow-\text{R2}} \begin{bmatrix} c-2d & b & a & d \\ 2(t+d)-(z+c) & -(y+b) & -(x+a) & -(t+d) \\ -5 & 2 & 1 & 4 \end{bmatrix} = A'.$$

(*Matrix Factorization*)

Using elementary (row/column) transformation matrices and associativity of multiplication, the same process can be written as:

$$A \to E_{21}(1)\cdot A \to E_{21}(1)\,A\cdot S\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix} \to E_{21}(1)\,A\,S\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix}\cdot E_{41}(-2)$$

$$\to E_{22}(-1)\cdot E_{21}(1)\,A\,S\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix} E_{41}(-2)$$

$$= E_{22}(-1)E_{21}(1)\cdot A\cdot S\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix} E_{41}(-2) = UAV = A',$$

where

$$U = E_{22}(-1)E_{21}(1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\cdot\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$V = S\begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix} E_{41}(-2) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\cdot\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}.$$

(*Verification*)

$$UAV = UA\cdot V = \left(\begin{bmatrix} 1 & 0 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\cdot\begin{bmatrix} a & b & c & d \\ x & y & z & t \\ 1 & 2 & 3 & 4 \end{bmatrix}\right)V$$

$$= \begin{bmatrix} a & b & c & d \\ -a-x & -b-y & -c-z & -d-t \\ 1 & 2 & 3 & 4 \end{bmatrix}\cdot\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ -2 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} c + d(-2) & b & a & d \\ -(c+z) - (d+t)(-2) & -(b+y) & -(a+x) & -(d+t) \\ 3 + 4(-2) & 2 & 1 & 4 \end{bmatrix} = A'.$$

Transformation matrices U (for rows) and V (for columns) perform operations on matrix A that are more complex than elementary ones since they represent a sequence of several elementary operations. When constructing U, the order of the matrices in the product is reversed, while when forming B, the order is direct. It is important to note that during the process of altering a matrix directly, only the internal order of row operations and the internal order of column operations must be preserved. However, the two types of operations (rows and columns) may be interleaved in any way. This follows from the associativity of matrix multiplication (and operations in general – recall the multiplication of permutations). Thus, in Example 34, writing the transformation as $A \to UAV$ means that the row operations were applied first and the column operations second (with the order in each preserved). However, a different alternation of row and column operations had initially been specified. One can verify independently that if

$$U' = E_{12}(1) \cdot E_{22}(-1), V' = E_{41}(-2) \cdot S \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{bmatrix},$$

then $U'AV \neq UAV \neq UAV'$.

It is clear that, in the general case, if the row operations correspond in order to the transforming matrices $U_1, U_2, \dots, U_m$, and the column operations correspond to $V_1, V_2, \dots, V_n$, then the original matrix A will finally transform into UAV, where $U = U_m U_{m-1} \dots U_1$ and $V = V_1 V_2 \dots V_n$.

Let's now consider rings of square matrices of order $p \in \mathbb{N}$ over an associative ring K with identity 1. Each ring $M_p(K)$ is itself associative and has an identity element $I = I_p$ (the zero element is the matrix $0 = \left(0_{ij}\right)_p$). In what follows, for matrices in $M_p(K)$ with fixed p, we will omit the explicit mention of the dimension unless necessary. Our focus will be on the problem of invertibility of elements in the ring $M = M_p(K)$.

A matrix $A \in M$ is called *nonsingular* (or *nondegenerate*) if it is invertible; that is, if there exists a matrix A^{-1}, called the *inverse* of A, such that

$$A^{-1} \cdot A = A \cdot A^{-1} = I.$$

Conversely, a matrix is called *singular* (or *degenerate*) if it is noninvertible – that is, if not such inverse exists. Singular matrices include, in particular, the zero matrix.

According to the remark on the uniqueness of the inverse element in a monoid (Remark 3), if for a matrix $A = \left(a_{ij}\right)$ there exist matrices X and Y such that $AX = YA = I$, then it follows that $X = Y = A^{-1}$; that is, A is nonsingular. Directly finding the inverse matrix $A^{-1} = \left(x_{ij}\right)$ amounts to solving a system of p^2 linear equations in the p^2 unknowns $x_{11}, x_{12}, \dots, x_{pp}$. This is equivalent to a single matrix equation $\left(a_{ij}\right) \cdot \left(x_{ij}\right) = I$ or $\left(x_{ij}\right) \cdot \left(a_{ij}\right) = I$, with the same solution, even when the ring K is noncommutative. Thus, the nonsingularity of A is equivalent to the simultaneous solvability of the systems $\left(a_{ij}\right) \cdot \left(x_{ij}\right) = I$ and $\left(x_{ij}\right) \cdot \left(a_{ij}\right) = I$. If the existence of solutions in K for both systems is established, then in fact it suffices to solve only one of them. Conversely, if it is known that at least one of the systems is inconsistent, then A is a singular matrix.

Example 35. Over $K = \mathbb{Z}_6$ (the ring of residue classes modulo 6), consider the ring $M_2(\mathbb{Z}_6)$ of (2×2)-matrices. Although the elements 2 and 3 of $\mathbb{Z}_6$ are zero divisors (hence noninvertible), the matrix

$$\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \in M_2(\mathbb{Z}_6)$$

is invertible: indeed

$$A^2 = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 4+9 & 6+6 \\ 6+6 & 9+4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod 6,$$

so $A^{-1} = A$.

By contrast, the matrix

$$B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

is singular: if

$$BX = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I,$$

then comparing entries yields the system $2x_{11} = 1, 2x_{12} = 0, x_{21} = 0, x_{22} = 1$. The equation $2x_{11} = 1$ has no solution in $\mathbb{Z}_6$ because 2 is noninvertible, so no right inverse exists and B is singular.

As an exercise, investigate singularity/invertibility of the matrices

$$\begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}, \begin{bmatrix} 5 & 1 \\ 5 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 5 \\ 4 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 4 \\ 4 & 0 \end{bmatrix}$$

and find inverses for the nonsingular ones.

Example 36. Let A be a $(p \times p)$-matrix over a commutative ring K that is block/upper-triangular in the sense that all entries below the main diagonal are zero (so the matrix has diagonal entries $a_{11}, a_{22}, \dots, a_{pp}$). Fix k with $1 \leq k \leq p$:

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1,k-1} & a_{1k} & a_{1,k+1} & \cdots & a_{1p} \\ 0 & a_{22} & a_{23} & \cdots & a_{2,k-1} & a_{2k} & a_{2,k+1} & \cdots & a_{2p} \\ 0 & 0 & a_{33} & \cdots & a_{3,k-1} & a_{3k} & a_{3,k+1} & \cdots & a_{3p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a_{k-1,k-1} & a_{k-1,k} & a_{k-1,k+1} & \cdots & a_{k-1,p} \\ 0 & 0 & 0 & \cdots & 0 & a_{kk} & a_{k,k+1} & \cdots & a_{kp} \\ 0 & 0 & 0 & \cdots & 0 & 0 & a_{k+1,k+1} & \cdots & a_{k+1,p} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & a_{p,k+1} & \cdots & a_{pp} \end{bmatrix}.$$

If the diagonal entry a_{kk} is *noninvertible in* K, then A is singular. Assume by contradiction that there exists $X=(x_{ij})$ with $XA=I$.

From the p^2 equations obtained by equating the entries of the matrix

$$XA=\begin{bmatrix} x_{11}a_{11} & x_{11}a_{12}+x_{12}a_{22} & \cdots & x_{11}a_{1k}+x_{12}a_{2k}+\cdots+x_{1k}a_{kk} \\ x_{21}a_{11} & x_{21}a_{12}+x_{22}a_{22} & \cdots & x_{21}a_{1k}+x_{22}a_{2k}+\cdots+x_{2k}a_{kk} \\ \cdots & \cdots & \cdots & \cdots \\ x_{k1}a_{11} & x_{k1}a_{12}+x_{k2}a_{22} & \cdots & x_{k1}a_{1k}+x_{k2}a_{2k}+\cdots+x_{kk}a_{kk} \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

componentwise to those of the matrix I, we obtain

$x_{11}a_{11}=1,$

$x_{21}a_{11}=0,\ x_{21}a_{12}+x_{22}a_{22}=1,$

$x_{31}a_{11}=0,\ x_{31}a_{12}+x_{32}a_{22}=0,\ x_{31}a_{13}+x_{32}a_{23}+x_{33}a_{33}=1,$

……

$x_{k1}a_{11}=0,\ x_{k1}a_{12}+x_{k2}a_{22}=0,\ x_{k1}a_{1,k-1}+\cdots+x_{k,k-1}a_{k-1,k-1}=0,$

$$x_{k1}a_{1k}+\cdots+x_{k,k-1}a_{k-1,k}+x_{kk}a_{kk}=1.$$

The first equation implies a_{11} is invertible so the remaining equalities in the first column force $x_{21}=x_{31}=\cdots=x_{k1}=0$.

Now compare entries in the second column using the already-known vanishing of the x_{i1}:

$$x_{21}a_{12}+x_{22}a_{22}=1,\ x_{32}a_{22}=0,\ \ldots,\ x_{k2}a_{22}=0,$$

and we deduce $x_{22}a_{22}=1$. Hence a_{22} is invertible and $x_{32}=\cdots=x_{k2}=0$. Continuing this way, we find successively that $a_{33},\ldots,a_{k-1,k-1}$ are invertible and many x_{ij} vanish, until finally the equation coming from the (k,k)-entry reads

$$x_{kk}a_{kk}=1,$$

which contradicts the assumed noninvertibility of a_{kk}. Thus, no such X exists and A is singular.

Further important examples of inversion are the transformation matrices already known to us.

Example 37. For any permutation $i_1,i_2,\ldots,i_p$ of the numbers $1,2,\ldots,p$, the matrix

$$S\begin{bmatrix} 1 & 2 & \cdots & p \\ i_1 & i_2 & \cdots & i_p \end{bmatrix}$$

is nonsingular. Its inverse can be verified directly and is given by

$$S^{-1}\begin{bmatrix} i_1 & i_2 & \cdots & i_p \\ 1 & 2 & \cdots & p \end{bmatrix},$$

where the unit of the 1st column lies in the first row, the unit of the 2nd column lies in the second row, and so forth. Indeed, it is clear that to restore the original matrix A from the product SA, one must move the i_1-th row back to the first position, the i_2-th row to the second, and so on. Let's make Example 37 concrete with $p=3$.

Take the numbers 1,2,3. Suppose we choose the permutation $(i_1,i_2,i_3)=(2,3,1)$.

(*Construct S*)

$$S\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}.$$

This matrix encodes:

- the 1st row of A goes to position 2,
- the 2nd row goes to position 3,
- the 3rd row goes to position 1.

So, multiplying SA permutes the rows of A.

(*Compute the inverse* S^{-1})

$$S^{-1}\begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix}.$$

This says:

- move row 2 back to row 1,
- move row 3 back to row 2,
- move row 1 back to row 3.

(*Test with a sample matrix*) Let

$$A=\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}.$$

Then

$$SA = \begin{bmatrix} e & f \\ a & b \\ c & d \end{bmatrix}.$$

Here the rows are permuted: row 3 is now 1st, row 1 is 2nd, row 2 is 3rd. Now, apply S^{-1}:

$$S^{-1}(SA) = A.$$

So, the permutation is perfectly undone.

(*Key points used*)

- S permutes the rows of matrix A according with the chosen permutation.
- S^{-1} restores the original order.

Example 38. For any $\lambda \in K$, the inverse matrix of $E_{ik}(\lambda)$ is $E_{ik}(-\lambda)$. For instance, when $p = 4, i = 2$, and $k = 3$, we have

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \lambda & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

This is natural: if the matrix A' is obtained from A by adding the k-th row, multiplied by λ, to the i-th row, then to recover A from A', one must add the k-th row, multiplied by $-\lambda$, to the i-th row.

Example 39. The matrix $E_{ii}(\lambda)$ is nonsingular if and only if λ is an invertible element of the ring K (proof left to the reader). In that case, the inverse is given by $E_{ii}(\lambda^{-1})$.

Analogous statements hold for matrices which, when multiplied on the right, transform the columns of a given matrix A.

Thus, in principle, the problem of inverting matrices $A \in M_p(K)$ is solved. However, the explicit computation of A^{-1} (or the determination of whether A is singular) is, in general, rather labor-intensive. The process can be simplified by assuming that every nonzero element of K is invertible; that is, by assuming K is a skew field. To proceed, we require the matrices of elementary transformations together with Example 36 (which serves as a lemma). In that lemma, the condition that an element a_{kk} is noninvertible is replaced with the simpler condition $a_{kk} = 0$. In practice, K is usually even a field, but commutativity of multiplication is never required.

Let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}, A \in M_p(K),$$

where K is a skew field, $p \in \mathbb{N}$. We describe a process which, if it terminates at some stage, shows immediately that the matrix A is singular, and which, if carried out to the end, either yields the inverse A^{-1}, or shows that it does not exist.

(*First stage*) If $a_{11} = a_{21} = \cdots = a_{p1} = 0$, the *process terminates at the first stage*, and A is singular. Otherwise, choose i such that $a_{i1} \neq 0$. Multiply the i-th row on the left by a_{i1}^{-1} and move it to the first row (the other rows may be rearranged for convenience). The corresponding transformation matrices are $E_{ii}(a_{i1}^{-1})$ and $S\begin{bmatrix} i & \cdots \\ 1 & \cdots \end{bmatrix}$, so that

$$A' = \begin{bmatrix} 1 & a'_{12} & \cdots & a'_{1p} \\ a'_{21} & a'_{22} & \cdots & a'_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{p1} & a'_{p2} & \cdots & a'_{pp} \end{bmatrix} = S\begin{bmatrix} i & \cdots \\ 1 & \cdots \end{bmatrix} E_{ii}(a_{i1}^{-1})\, A.$$

If $a'_{21} = \cdots = a'_{p1} = 0$, the first stage is complete. Otherwise, for each $r \geq 2$ with $a'_{r1} \neq 0$, add to the r-th row the first row multiplied on the left by $-a'_{r1}$. The corresponding matrices are $E_{r1}(-a'_{r1})$. At the end of the first stage, the matrix has the form

$$A'' = \begin{bmatrix} 1 & a''_{12} & \cdots & a''_{1p} \\ 0 & a''_{22} & \cdots & a''_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a''_{p2} & \cdots & a''_{pp} \end{bmatrix}.$$

(*Second stage*) If $a''_{22} = \cdots = a''_{p2} = 0$, the *process terminates at the second stage*, and A is singular. Otherwise, as in the first stage, a leading 1 is produced in position $(2, 2)$, the entries are cleared below it, and finally a''_{12} is eliminated (if nonzero) by adding to the first row the second row multiplied on the left by $-a''_{12}$ (This step is skipped if $a''_{12} = 0$).

At the end of the second stage (if the process has not terminated earlier), the matrix has the form

$$A''' = \begin{bmatrix} 1 & 0 & a'''_{13} & \cdots & a'''_{1p} \\ 0 & 1 & a'''_{23} & \cdots & a'''_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a'''_{p3} & \cdots & a'''_{pp} \end{bmatrix}$$

(*General stage*) Proceeding similarly, at each stage k either:

1. we obtain a matrix $\tilde{A}$ whose all entries in column k below the pivot position vanish, in which case the process terminates and A is singular; or

2. a pivot element is produced in position (k, k), zeros are cleared below it, and the column entries above it are eliminated.

(*Termination*) Two cases are possible:

1. If all p stages of the process are completed, the result is the identity matrix, and the product of the elementary matrices used gives A^{-1}.

2. If the process terminates earlier, the resulting matrix has a zero column in the reduced block, implying that A is singular.

Remarks 32. We have described a row-reduction procedure (*Gaussian elimination* with left multiplications) that either produces a factorization from which A^{-1} can be computed, or else terminates earlier, proving that A is singular.

Notation: all scalar inverses and scalar multiplications of rows are performed on the left (since K may be noncommutative). Elementary row operations are represented by left multiplication with elementary matrices.

First case. $I = U_m U_{m-1} \dots U_1 A$, where $U_1, U_2, \dots, U_m$ are the matrices corresponding to the elementary row transformations performed at all p stages, in order: first the operations of the initial stage (in sequence), then those of the second stage, and so on. Thus, $U = U_m U_{m-1} \dots U_1$ is the only candidate for the role of A^{-1} since it satisfies the relation $UA = I$. However, it is still necessary to verify directly that $AU = I$. This verification is essential, because $UA = I$ only guarantees that U is a *left inverse* of A, but not necessarily the true inverse. Avoiding this check is inadvisable: even if (under certain additional assumptions on K and A) the verification were theoretically unnecessary, in practice it remains crucial since the likelihood of an error in at least one transformation is fairly high.

Second case. If the process terminates at the k-th stage ($1 \le k \le p$), the matrix $\tilde{A}$ is singular by Example 36. We now show that the original matrix A must also be singular.

Indeed, $\tilde{A} = U_m U_{m-1} \dots U_1 A$, where $U_1, U_2, \dots, U_m$ are the matrices corresponding to the elementary transformations that convert A into $\tilde{A}$. Each U_i is nonsingular. If A were nonsingular, then by the theorem on the invertibility of a product in a monoid, the matrix $\tilde{A}$ would also be nonsingular – a contradiction.

Example 40. (*Matrix Inversion Example*) The procedure for inverting the matrix A, although already described and justified, has not yet been presented in a concise form. Before doing so, let's consider an illustrative example.

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix}, (K = \mathbb{Q}, K = \mathbb{R} \text{ or } K = \mathbb{C}; p = 3).$$

We perform a sequence of elementary row transformations. Above each arrow, we place the symbol of the corresponding transformation matrix. The reduction of A to the identity matrix is carried out as follows:

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \xrightarrow{E_{33}(1/2)} \begin{bmatrix} 0 & 1 & 1 \\ 3 & 0 & 1 \\ 1 & 1/2 & 0 \end{bmatrix} \xrightarrow{S\left[\begin{smallmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \end{smallmatrix}\right]} \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 1 \\ 3 & 0 & 1 \end{bmatrix} \xrightarrow{E_{31}(-3)} \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 1 \\ 0 & -3/2 & 1 \end{bmatrix}$$

$$\xrightarrow{E_{32}(3/2)} \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 5/2 \end{bmatrix} \xrightarrow{E_{12}(-1/2)} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 5/2 \end{bmatrix} \xrightarrow{E_{33}(2/5)} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E_{23}(-1)} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{E_{13}(1/2)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I.$$

Since each elementary transformation is simultaneously applied to the identity matrix I, the inverse A^{-1} is obtained as the product

$$U = E_{13}(1/2) E_{23}(-1) E_{33}(2/5) E_{12}(-1/2) E_{32}(3/2) E_{31}(-3) S \begin{bmatrix} 3 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix} E_{32}(1/2).$$

Carrying out the matrix multiplications step by step:

$$U = \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2/5 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1/2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3/2 & 1 \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} = \begin{bmatrix} -1/5 & 1/5 & 1/5 \\ 2/5 & -2/5 & 3/5 \\ 3/5 & 2/5 & -3/5 \end{bmatrix}.$$

As an exercise, one may explicitly verify the product of all factors (instead of computing them sequentially, as above), which replaces the need for seven smaller examples.

(*Verification*)

$$AU=\begin{bmatrix}0&1&1\\3&0&1\\2&1&0\end{bmatrix}\cdot\begin{bmatrix}-1/5&1/5&1/5\\2/5&-2/5&3/5\\3/5&2/5&-3/5\end{bmatrix}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}.$$

Thus, $U=A^{-1}$ since both $AU=UA=I$.

Example 41. (*Matrix inversion using the method of elementary transformations*) Having understood the procedure described earlier, it is now easy to follow an elegant implementation (due to D.N. Krylov), which avoids the need to explicitly write out the transformation matrices.

Given a matrix $A=(a_{ij})\in M_p(K)$, we append the identity matrix $I=I_p$ to its right, thereby forming a $(p\times 2p)$-matrix:

$$\left[\begin{array}{cccc|cccc}a_{11}&a_{12}&\cdots&a_{1p}&1&0&\cdots&0\\a_{21}&a_{22}&\cdots&a_{2p}&0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots&\vdots&\vdots&\ddots&\vdots\\a_{p1}&a_{p2}&\cdots&a_{pp}&0&0&\cdots&1\end{array}\right]. \tag{13}$$

Applying a sequence of elementary row transformations to this augmented matrix, we transform its left half into the identity matrix. Simultaneously, the right hand becomes a matrix U satisfying $UA=I$ (why?). To complete the argument, one verifies directly that $AU=I$, and hence $U=A^{-1}$.

This technique is called *inversion by the method of elementary transformations* and is conventionally denoted by

$$[A|I]\xrightarrow{E_1,E_2,\dots,E_k}[I|A^{-1}].$$

(*Illustrative Computation*) Consider the matrix A of Example 40.

$$\left[\begin{array}{ccc|ccc}0&1&1&1&0&0\\3&0&1&0&1&0\\2&1&0&0&0&1\end{array}\right]\longrightarrow\left[\begin{array}{ccc|ccc}0&1&1&1&0&0\\3&0&1&0&1&0\\1&1/2&0&0&0&1/2\end{array}\right]\longrightarrow\left[\begin{array}{ccc|ccc}1&1/2&0&0&0&1/2\\0&1&1&1&0&0\\3&0&1&0&1&0\end{array}\right]\longrightarrow$$

$$\longrightarrow\left[\begin{array}{ccc|ccc}1&1/2&0&0&0&1/2\\0&1&1&1&0&0\\0&-3/2&1&0&1&-3/2\end{array}\right]\longrightarrow\left[\begin{array}{ccc|ccc}1&1/2&0&0&0&1/2\\0&1&1&1&0&0\\0&0&5/2&3/2&1&-3/2\end{array}\right]\longrightarrow$$

$$\longrightarrow\left[\begin{array}{ccc|ccc}1&0&-1/2&-1/2&0&1/2\\0&1&1&1&0&0\\0&0&5/2&3/2&1&-3/2\end{array}\right]\longrightarrow\left[\begin{array}{ccc|ccc}1&0&-1/2&-1/2&0&1/2\\0&1&1&1&0&0\\0&0&1&3/5&2/5&-3/5\end{array}\right]\longrightarrow$$

$$\longrightarrow\left[\begin{array}{ccc|ccc}1&0&-1/2&-1/2&0&1/2\\0&1&0&2/5&-2/5&3/5\\0&0&1&3/5&2/5&-3/5\end{array}\right]\longrightarrow\left[\begin{array}{ccc|ccc}1&0&0&-1/5&1/5&1/5\\0&1&0&2/5&-2/5&3/5\\0&0&1&3/5&2/5&-3/5\end{array}\right].$$

Thus, the right-hand side is precisely A^{-1}, and the relation $AU=I$ has already been verified.

(*Another approach*) Instead of working with rows, one can form a $(2p\times p)$-matrix

$$\begin{bmatrix}A\\I_p\end{bmatrix}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1p}\\a_{21}&a_{22}&\cdots&a_{2p}\\\vdots&\vdots&\ddots&\vdots\\a_{p1}&a_{p2}&\cdots&a_{pp}\\1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{bmatrix}$$

and now apply elementary column transformations instead of row transformations. The goal is to transform the top block into I_p; simultaneously, the bottom block will become a matrix V such that $AV=I$. It then remains to check directly that also $VA=I$. We suggest carrying this out independently for the same (3×3)-matrix A from the previous example, and verifying that the resulting matrix V coincides with the earlier inverse U. Elementary transformations with columns can also be performed on a matrix written in augmented form as in (13), provided that every operation on the columns of the left block is simultaneously applied to the corresponding columns of the right block.

Question. The idea that arises is to transform both the rows and columns of matrix (13), with the goal of reducing the left half to the identity matrix. For example,

$$\left[\begin{array}{ccc|ccc} 0 & 1 & 1 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{swap columns: } S\left[\begin{smallmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{smallmatrix}\right]} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \end{array}\right]$$

$$\xrightarrow{C_2 \leftarrow C_2 - C_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & -1 & 0 \\ 1 & -1 & 2 & 0 & 1 & 0 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 - R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & -1 & 0 \\ 0 & -1 & 2 & 0 & 1 & -1 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 + R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & -1 & 0 \\ 0 & 0 & 5 & 1 & 0 & -1 \end{array}\right] \xrightarrow{R_3 \leftarrow \frac{1}{5} R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1/5 & 0 & -1/5 \end{array}\right]$$

$$\xrightarrow{R_2 \leftarrow R_2 - 3R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2/5 & -1 & 3/5 \\ 0 & 0 & 1 & 1/5 & 0 & -1/5 \end{array}\right].$$

Thus, the candidate for the inverse is

$$W = \begin{bmatrix} 0 & 0 & 1 \\ 2/5 & -1 & 3/5 \\ 1/5 & 0 & -1/5 \end{bmatrix}.$$

However, direct verification shows that $W \neq A^{-1}$. *What went wrong with this approach*?

One of the main applications of matrix inversion is the solution of systems of linear equations

$$\left.\begin{matrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = b_2 \\ \cdots\cdots\cdots\cdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pp}x_p = b_2 \end{matrix}\right\},$$

where the coefficients a_{ij}, the unknowns x_j, and the constant terms b_j belong to a certain field P. Defining

$$A = (a_{ij})_{p\times p} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{bmatrix}, X = (x_j)_{p\times 1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}, B = (b_j)_{p\times 1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix},$$

we can rewrite the system compactly as

$$AX = B. \quad (14)$$

A *solution* is not only the process of determining the values of the unknowns $x_1, x_2, \dots, x_p$, but also the ordered tuple of values $\check{x}_1, \check{x}_2, \dots, \check{x}_p \in P$. Substituting these values into the system gives true equalities in P. Equivalently, a solution (or *root*) of (14) is a vector $\check{X} = (\check{x}_j)_{p\times 1}$ for which $A\check{X} = B$.

Claim 3. *If A is a nonsingular matrix, then equation* (14) *has a unique solution $\check{X} = A^{-1}B$.* •

Indeed, assuming first that some solution $\check{X}$ exists, we obtain $\check{X} = A^{-1}B$ from the equality $A\check{X} = B$ by multiplying by A^{-1} on the left, and then by substituting into (14) we verify that this unique candidate is indeed a solution (compare with Theorem 4 on the unique solvability of equations in a group).

Corollary 8. *If* $b_1 = b_2 = \cdots = b_p = 0$ (in this case, the system of equations is called *homogeneous*) *and the matrix A is nonsingular, then the only solution to the system is the trivial* (*zero*) *solution* $\check{x}_1 = \check{x}_2 = \cdots = \check{x}_p = 0$. •

Remark 33. As follows from the above, for a homogeneous system

$$\left.\begin{matrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p = 0 \\ \cdots\cdots\cdots\cdots \\ a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pp}x_p = 0 \end{matrix}\right\}$$

to admit nontrivial solutions (that is, when at least one $\check{x}_j \neq 0$), it is necessary that the coefficient matrix $A = (a_{ij})_{p\times p}$ be singular. In fact, this condition is also sufficient: a homogeneous system of p equations in p unknowns with a singular matrix necessarily possesses nontrivial solutions. The proof, however, is not entirely straightforward and will be given later in the general theory of linear systems (semester 2).

Exercise. Apply the matrix approach to the systems of equations from Exercises $4 - 4''$ of the introductory lecture (see Part I of this course), and also to the corresponding homogeneous systems obtained by replacing all constants on the right-hand sides with zeros.

MATERIAL FOR INDEPENDENT STUDY.

(*based on G.E. Shilov*, *Mathematical Analysis*, *Part I*: *Finite-Dimensional Linear Spaces*, *Moscow*, *Nauka*, 1969)

- Matrix multiplication (over an arbitrary ring), including block partitioning.
- Matrix transposition; the transposition theorem for the product of matrices over a commutative ring.

These topics are neither difficult nor lengthy. Their distribution near the end of the examination, and their inclusion in the exam syllabus, will serve as a useful test of whether students have acquired at least minimal skills in the independent study of mathematical literature during the semester.

CONCLUSION

Only in this part of the course – intended for the entire first semester – does our teaching methodology differ significantly from the traditional one. Beginning in the second semester, one may, for example, adopt the following textbook as a basis:

N. V. Efimov, E. R. Rozendorn, *Linear Algebra and Multidimensional Geometry*, *Moscow*, *Nauka*, 1970, supplemented with the theory of determinants from either G. E. Shilov or A. G. Kurosh.

This consideration motivated us to publish the initial part of this algebra course separately, without waiting for the completion of the subsequent volumes.

CONCLUDING NOTE

(Carlos E. Frasser)

Professor Zykov's approach to algebra as a scientific discipline emphasizes the independent development of fundamental concepts and the gradual transition to more general results. I would add that this methodology remains pedagogically valuable today: it enables students not only to master computational techniques but also to develop the capacity to reason abstractly about groups, rings, fields, polynomials, and matrix theory.

In this spirit, the exercises and independent study materials serve as a crucial bridge between formal theory and personal discovery. My experience confirms that such a balance between guidance and autonomy is essential for fostering genuine mathematical maturity in students.

References

1. Frasser C.E. *A Course of Algebra* (*Part I*). 2024. ArXiv:2411.11873.

SINGLE-MODE VORTEX AS AN ELEMENTARY VORTEX GEOMETRIC STRUCTURE OF AN IDEAL FLUID

Milyute E.
Milyus A.
IRG "LITAVEM-3"
Vilnius, Lithuania
https://doi.org/10.5281/zenodo.18840570

Abstract

A single-mode vortex is considered as an elementary solution to the equations of an ideal incompressible fluid in Arnold's geometric formulation [1]. It is shown to be defined by a single harmonic mode and to represent a divergence-free velocity field with helical streamline geometry. In the Lie algebra of volume-preserving diffeomorphisms, a single-mode vortex corresponds to an Abelian element whose commutator with itself is zero, ensuring the absence of nonlinear self-generation of modes and the stationarity of the solution [2]. It is established that such vortices are geodesic trajectories on the group of diffeomorphisms and serve as fundamental building blocks of more complex vortex structures.

Keywords: vortical dynamics, topological methods in hydrodynamics, single-mode vortex, kern, algebra of Lie, diffeomorphism, helical geometry.

1. General Definition. In the theory of an ideal incompressible fluid, particularly in Arnold's geometric formulation [1], elementary vortex structures called single-mode vortices play a fundamental role. A single-mode vortex represents a velocity field corresponding to a single harmonic mode in the Fourier expansion. In its most general form, such a field can be described using a scalar stream function.

$$\psi(x) = A\,Cos(k\cdot x), \quad (1)$$

where $x \in \mathbb{R}^n$ is the spatial coordinate, $k \in \mathbb{R}^n$ is the fixed wave vector, A is the amplitude, and $k \cdot x$ denotes the scalar product. The corresponding velocity field is defined through the oblique gradient operator

$$\upsilon = \nabla^{\perp}\psi \text{ , (2)}$$

Which guarantees the fulfillment of the incompressibility condition $\nabla \cdot \upsilon$. In the two-dimensional case, this means

$$\upsilon(x) = \left(\frac{\partial\psi}{\partial y}, -\frac{\partial\psi}{\partial x}\right) = Ak^{\perp}\,Sin(k\cdot x), \quad (3)$$

where $k^{\perp} = (-k_2, k_1)$ is the vector orthogonal to k.

Geometrically, a single-mode vortex is characterized by the fact that all of its dynamic properties are determined by a single spatial harmonic. Lines of constant phase $k \cdot x = Const$ form a regular family of surfaces along which the fluid moves. In cylindrical coordinates, these lines take the form of helical curves, and the corresponding vortex structure realizes helical geometry. In this sense, a single-mode vortex represents the simplest helical structure possessing complete spatial regularity.

From the perspective of Arnold's geometric mechanics [1], a single-mode vortex has a particularly important algebraic property. The space of all admissible velocity fields of an ideal fluid can be viewed as an infinite-dimensional Lie algebra of divergence-free vector fields. In this space, the elementary harmonics $e_k = e^{i(k\cdot x)}$ form a natural basis, and the corresponding vortex fields can be expressed in terms of these modes. The Lie commutator of two modes is defined by the relation

$$[e_k, e_\ell] = (k \wedge \ell)e_{k+\ell}\,, \quad (4)$$

where $(k \wedge \ell)$ is the oriented area of the parallelogram constructed on the vectors k and ℓ. In particular, for a single-mode vortex the following holds:

$$[e_k, e_k] = 0\,. \quad (5)$$

This means that a single-mode vortex does not interact with itself and does not generate new modes. Dynamically, this corresponds to the velocity field maintaining its structure over time and not undergoing nonlinear deformation. Thus, a single-mode vortex is a stationary solution to the Euler equations and represents a geodesic line in the group of volume-preserving diffeomorphisms.

The vortex field corresponding to a single-mode structure also has a simple form. Vorticity is defined as the velocity curl:

$$\omega = \nabla \times \upsilon\,. \quad (6)$$

Substituting the expression for velocity shows that the vorticity is also proportional to the same harmonic:

$$\omega \propto Cos\,(k\cdot x). \quad (6^*)$$

This means that the vortex is completely described by a single spatial mode and does not contain more complex spatial scales.

Geometrically, a single-mode vortex corresponds to a minimal helical surface, in particular a helicoid. Such a surface can be parameterized as

$$X(r,\theta) = (r\,Cos\,\theta, r\,Sin\,\theta, a\theta), \quad (7)$$

where the parameter a is determined by the components of the wave vector. The radial coordinate remains constant along the streamlines, and the fluid moves strictly along helical trajectories. This reflects the fundamental regularity of the single-mode structure [2-4].

From a physical perspective, a single-mode vortex is an elementary vortex configuration that contains no internal interaction between different spatial scales. Increasingly complex vortex structures can be obtained as a superposition of several single-mode vortices. It is the interaction of different modes that leads to radial transport, geometric deformation, and the formation of more complex surfaces such as catenoids and spherical structures. However, the single-mode vortex remains the fundamental building block of vortex dynamics, defining the simplest stable state of vortex motion.

Thus, a single-mode vortex represents an elementary divergent-free velocity field corresponding to a single harmonic mode, possessing helical geometry and being a stationary solution to the equations of ideal hydrodynamics. Its fundamental role is that it serves as the basic element from whose interactions increasingly complex vortex structures emerge.

2. Single-mode vortex in the context of the Lie algebra of diffeomorphisms. The formalization of a single-mode vortex naturally arises within the framework of the geometric theory of ideal fluids, where the space of all admissible motions is considered as a group of volume-preserving diffeomorphisms of the flow domain. Let M be a smooth oriented manifold with a fixed volume shape. Then the group of all smooth volume-preserving diffeomorphisms $\mathrm{Diff}_\mu(M)$, forms an infinite-dimensional Lie group. The corresponding Lie algebra $\mathfrak{diff}_\mu(M)$ consists of all divergence-free vector fields on M, that is, those velocity fields $v(x)$ for which the incompressibility condition holds.

$$\nabla \cdot v = 0. \tag{8}$$

The algebraic structure of this Lie algebra is given by the commutator of vector fields, defined as

$$[u,v] = (u \cdot \nabla)\, v - (v \cdot \nabla) u \,. \tag{9}$$

In this formulation, the Euler equation for an ideal fluid acquires a particularly simple geometric meaning: it describes geodesic motion on the group $\mathrm{Diff}_\mu(M)$ with respect to a right-invariant metric induced by kinetic energy. The fluid velocity vector at each instant is an element of the Lie algebra $\mathfrak{diff}_\mu(M)$, and its evolution is determined by the intrinsic geometry of this group. Thus, the dynamics of vortices can be completely described in terms of the Lie algebra of divergence-free diffeomorphisms.

In the case of a periodic domain, for example, a two-dimensional torus T^2, the natural basis of the Lie algebra is the harmonic modes of the form

$$e_k(x) = e^{i(k\cdot x)} \,, \tag{10}$$

where k is an integer wave vector. These modes form a complete basis in which any divergence-free field can be expanded in a Fourier series. The commutator of two such modes has the form (4), where the quantity

$$k \wedge \ell = k_1\ell_2 - k_2\ell_1 \tag{11}$$

represents the oriented area of the parallelogram constructed on the vectors k and ℓ. This relation is fundamental, since it shows that the interaction of two vortex modes leads to the emergence of a new mode corresponding to the sum of the wave vectors.

A single-mode vortex corresponds to the choice of a single basis element of the Lie algebra. Let the velocity field be defined by a single mode e_k. Then the corresponding vector field has the form (3) and belongs to the Lie algebra $\mathfrak{diff}_\mu(M)$. Since the commutator of this mode with itself is zero (5), a single-mode vortex is an Abelian element of the Lie algebra [2]. This means that the corresponding velocity field does not generate new modes during its own evolution and retains its structure over time. Geometrically, this corresponds to the fact that a single-mode vortex defines a one-dimensional Lie subalgebra, which generates a one-parameter subgroup of diffeomorphisms. This subgroup describes the helical motion of fluid particles.

In cylindrical coordinates, such motion takes the form of a helical trajectory:

$$z = a\theta \,, \tag{12}$$

where the parameter a is determined by the components of the wave vector. Thus, a single-mode vortex realizes the simplest geometric structure, corresponding to a one-parameter subgroup of the group of volume-preserving diffeomorphisms. Its stability is a direct consequence of the fact that it corresponds to a geodesic line on the group $\mathrm{Diff}_\mu(M)$ [2].

If the velocity field contains more than one mode, the corresponding element of the Lie algebra ceases to be Abelian [2-4]. The commutators of the different modes become non-zero, and a nonlinear interaction arises, leading to the generation of new modes. This corresponds to going beyond the one-dimensional subalgebra and transitioning to a more complex geometry of motion. In this sense, a single-mode vortex is an elementary geometric object in the Lie algebra of diffeomorphisms, serving as a fundamental building block for more complex vortex structures.

Thus, the Lie algebra of volume-preserving diffeomorphisms provides a natural and rigorous mathematical framework for describing vortices. A single-mode vortex is an elementary element of this algebra, corresponding to one harmonic mode and one geodesic trajectory in the diffeomorphism group. Its geometric simplicity and stability make it a fundamental object, the interactions of which give rise to increasingly complex vortex and geometric structures [2-4].

References:

1. V.I.Arnold, B.A. Khesin. (1998) Topological Methods in Hydrodynamics, Springer
2. E.Milyute (2022) Conservation of differential inclusion in the conformal transformation of a spherical vortex onto itself, *LXIII Conference of the LMS, dedicated to the 100th anniversary of the Lithuanian University, June 16-17, KTU, VDU, Kaunas*
3. E.Milyute (2023) Investigation of Conformal Mappings in Solving Problems of Vortex Dynamics.
4. E.Milyute, A.Milyus (2024) Radius-vector of the Vector Potential and its Connection with the Kern of an Ellipsoidal Vortex, *NJD-iScience*, **134**, PP. 30-37

MEDICAL SCIENCES

IMMUNOHISTOCHEMICAL EVALUATION OF ENDOMETRIAL REACTIVITY IN HYPERPLASTIC PROCESSES

Asanbekova S.
The applicant of the department of Human anatomy named after Y.A. Rakhimov of SEI Avicenna Tajik State Medical University
Kurbonov S.
Doctor of medical sciences, professor of the department of Human anatomy named after. Y.A. Rakhimov of SEI Avicenna Tajik State Medical University
https://doi.org/10.5281/zenodo.18840572

Abstract

In endometrial hyperplasia (EHP), increased proliferative and reactive processes have been detected in the glands and stroma, accompanied by significant changes in the expression of the immunohistochemical markers Ki-67 and CD34. Quantitative analysis revealed that the nuclear expression level of Ki-67 reaches 58%. Such high proliferative index values confirm the active nature of cell division underlying the pathogenesis of EHP during the reproductive period. Analysis of the angiogenesis marker CD34 revealed the presence of intense vascular neoplasms, which is also characteristic of endometrial hyperplasia.

Keywords: endometrium, immunohistochemistry, hyperplastic process, Ki67 and CD34 markers.

Endometrial hyperplastic processes (EHP) including hyperplasia and polyps are histopathological changes in the endometrium resulting from the proliferation of the uterine mucosa, leading to thickening and an increase in its volume. These changes remain the most common gynecological pathology in women of reproductive age. Over the past decades, many researchers have noted an increase in the incidence of EHP, which averages 5-6% of all gynecological morbidity [3, 2, 7, 6].

Particular attention is paid to the study of the molecular biological marker of proliferation, the Ki-67 protein, which is recognized as the most informative and convenient tool for assessing tissue proliferative activity [1, 5]. The study of cell proliferation in various tissue remains one of the fundamental tasks of modern morphology. Thanks to significant advances in molecular biology, morphologists have gained the ability to perform immunohistochemical labeling of cells in various phases of the cell cycle [4].

To evaluate immunohistochemical (IHC) changes in the endometrium in hyperplastic processes using the proliferation marker Ki-67 and the angiogenesis marker CD 34.

Prospective studies of the endometrial function (diagnostic curettage of the uterine cavity) were conducted at two departments of Obstetrics and Gynecology №1 and № 2 of SEI of Avicenna Tajik State Medical University, at Republican Maternity Hospitals № 1 and № 2. The study included 11 reproductive-age patients with hyperplastic pathological processes and a confirmed absence of malignant neoplasms.

For IHC studies, histological sections 2-3 μm thick were stained with special antibodies of Ki-67 and CD-45 markers of the uterine mucosa using the automated system of Ventana Benchmark (Roche, Switzerland). The expression level of markers was assessed quantitatively (as a percentage) and classified as follows: weak expression – up to 20% of positively stained cells; moderate expression – from 20% to 60% of cells; pronounced (strong) expression – more than 60% of cells.

The use of IHC in studying of the endometrium with hyperplastic changes demonstrated a direct correlation between tissue structural changes and its biological potential. The study assessed the expression level of the IHC Ki-67 marker and the proliferative index in the endometrial tissues. It was found that Ki-67 expression is closely linked to the nature of endometrial proliferative disorders. The identified structural changes, such as glandular hyperplasia, were also accompanied by a pronounced positive reaction to the Ki-67 marker. The proliferative index in the Ki-67 was 58%, which corresponds to a high level of cell division and confirms the active pathogenetic nature of the process (Figures 1 and 2).

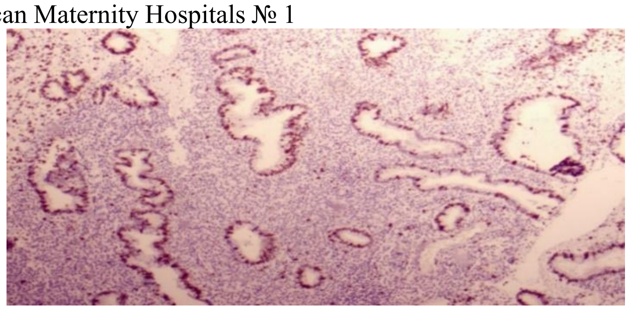


Figure 1. – The degree of expression of the IHC marker Ki-67 in the endometrial glands in HPP. Microphoto. Staining: DAB-hematoxylin. Magnification 50.

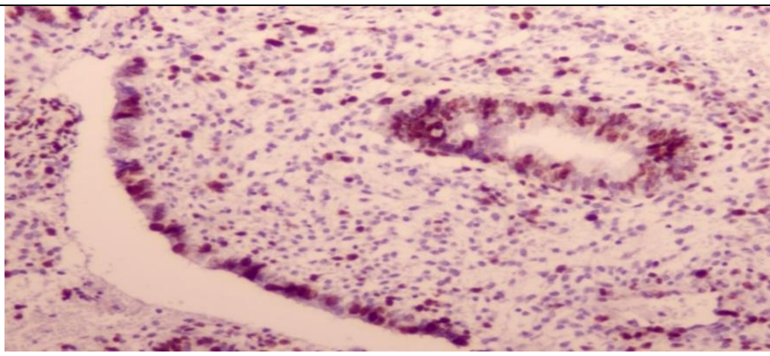


Figure 2. – The degree of expression of the IHC marker Ki-67 in the endometrial glands in HPP. Microphoto. Staining: DAB-hematoxylin. Magnification 100.

The study results showed that the Ki-67 proliferative index was intense in both the glandular epithelium and the endometrial stroma. Morphological analysis revealed that the Ki-67 marker localizes in cell nuclei with varying intensity as distributed between the karyoplasm and nucleolus, corresponding to different phases of the cell cycle. The variability of subcellular localization of Ki-67 confirms its ability to label cells at different stages of proliferation and mitotic activity.

Thus, in the case of EHP, the express level of the Ki-67 marker in the endometrium is 58%. Such high proliferative index values confirm the active nature of cell division, which underlies the pathogenesis of EHP during the reproductive period.

The functional significance of CD 34 and methodological aspects in EHP as a marker of the vasculature, which is a glycoprotein expressed on the surfaces of endothelial cells, making it a reliable marker for the identification of blood vessels.In endometrial studies, CD34 analysis allows for the assessment of the degree of tissue vascularization and the intensity of microangiogenesis in endometrial endothelial cell death. When interpreting staining results, a specific brown precipitate formed during the reaction with DAB indicated positive CD34 expression in the vascular endothelium. The analysis took into account both the signal localization and the staining intensity, reflecting the density of the vascular network in the samples (Figure 3).

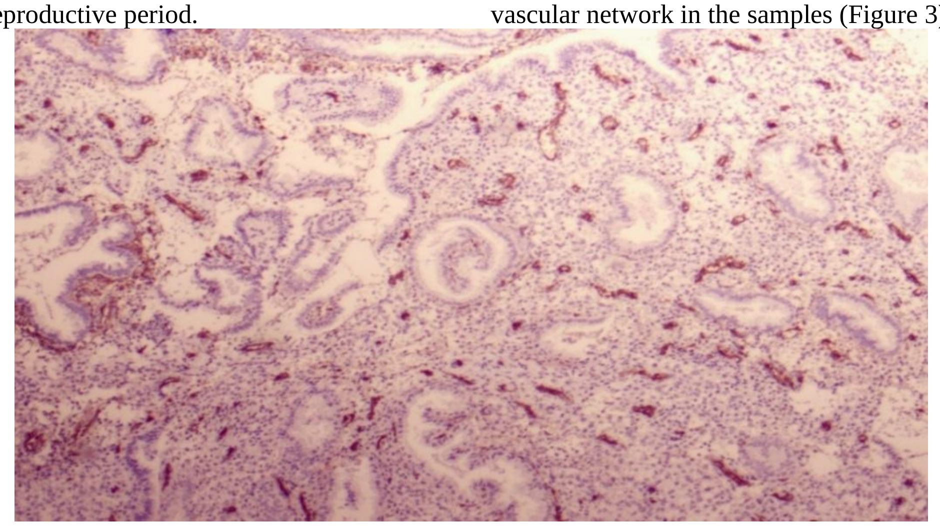


Fig. 3. – Immunohistological examination of CD 34-positive brown blood vessels in the endometrium. The appearance of small spiral arteries is noted. Magnification 50.

The obtained data demonstrate adequate expression of CD34 in endometrial vessels, with clear identification of both small capillaries and larger vascular trunks. Dilation lumen and high vascular density indicate enhanced angiogenesis. Specific brown staining (accumulation of the reaction product with DAB) confirms positive expression of CD34 in vascular endothelial cells.

Microscopic specimens reveal a dense network of exchange vessels and components of the microcirculatory bed. At the level of the terminal vessels of the endometrium, their dilation, tortuosity, and the formation of numerous anastomoses are noted, indicating increased stromal vascularization in this pathology (Figure 3). The study results are consistent with literature data [1,5].

Thus, various variants of endometrial hyperplasia process have diverse immunohistochemical characteristics. High expression of the angiogenesis marker CD34 confirms the intense nature of neovascularization in endometrial hyperplastia.

References:

1.Safarova, S.M. Apoptosis and angiogenesis in diagnostics and tactical selection of treatment of hyperplastic processes [Text] / S.M. Safarova, M.H. Kurbonova, D.B. Sattoroka // Materials of the VII Congress of Obstetricians and Gynecologists of the Republic of Tajikistan. - Dushanbe. - 2022. - p. 283-287.

2.Sviridova, N.I., Tkachenko L.V., Yakhontova M.A., Gritsenko I.A., Maksimov S.N., Puryaseva K.A. Hyperplastic processes of the endometrium: modern approaches to diagnosis and treatment. Obstetrics, Gynecology and Reproduction.-2024; 18(1): 83-95. Https://doi.org/10.17749/2313-7347/ob.gyn.rep.2023.464

3. Khushvakhtova, E.H. Diagnostic value of various methods of examination of patients of reproductive age with hyperplastic endo-myometrial processes [Text] /, E.H. Khushvakhtova, M.H. Kurbonova, F.M. Akilova // Materials of the VII Congress of Obstetricians and Gynecologists of the Republic of Tajikistan.-Dushanbe.-2022.-p.289-293.

4. Chibisova, G.M., Khabarov, S.V. Kompleksnoe opredelenie onkomarkerov SA125, NE4 i indeksa ROMA kak factor prognoza razvitya raka yaichnikov [Complex definition of oncoprotein CA125, HE4 and ROMA index as a factor in the prognosis of ovarian cancer]. Journal of New Medical Technologies.2018;25(3):15-20/DOI. 10.24411/1609-2163-2018-16158.

5. Khaskhachykh, D.A., Morphological and immunohistochemical characteristics of endometrial layers in women of reproductive age with endometrial hyperplasia without atypia. Scientific digest of association of obstetricians and gynecologists of Ukraine. DOI:10.35278/2664-0707. 2(53).2023.298047.

6. Kossaï, M. Penault-Llorca, F. Role of Hormones in Common Benign Uterine Lesions: Endometrial Polyps, Leiomyomas, and Adenomyosis // Advances in Experimental Medicine and Biology. – 2020.- № 61.-P. 37-58.

7. Venturella, R. Subcutaneous progesterone for endometrial polyps in premenopausal women: a preliminary retrospective analysis [Text] / R. Venturella, G. Miele, K. Cefali, [et al.] // J Minim Invasive Gynecol 2019; 26(1): 143–147.

IMMUNOHISTOCHEMICAL ANALYSIS OF FALLOPIAN TUBE REACTIVITY IN TUBAL ECTOPIC PREGNANCY

Qodirova F.
The applicant of the department of Human Anatomy named after Ya.A. Rakhimov of SEI "Avicenna Tajik State Medical University"
Kurbonov S.
Doctor of medical sciences, of the department of Human Anatomy named after Ya.A. Rakhimov of SEI "Avicenna Tajik State Medical University"
https://doi.org/10.5281/zenodo.18840574

Abstract

An immunohistochemical study assessing the expression of the proliferative marker Ki-67 revealed a wide range of proliferative changes in the fallopian tubes of patients with tubal ectopic pregnancy, reflecting pronounced reactive processes in response to inflammatory damage of the organ. An increase in the level of proliferative activity with an increase in the number of immunopositive cells was established.

Keywords: fallopian tubes, tubal ectopic pregnancy, immunohistochemistry, Ki-67 marker.

An ectopic pregnancy (EP) is the abnormal implantation of a fertilized egg outside the uterine cavity. This condition remains a life-threatening complication of early gestational pregnancy and is considered one of the most severe gynecological pathologies. One of the leading complications of EP is intra-abdominal (intraperitoneal) bleeding, which continues to significantly contribute to maternal mortality in developed countries. [2,7-8].

Recent global studies indicate that the incidence of ectopic pregnancy ranges from 0.5% to 6% [6]. A global analysis of the incidence of EP revealed a 10.2% decrease between 1990 and 2019. These epidemiological trends can be explained by the effective implementation of strategies to prevent reproductive system infections [9]. Much attention is paid to the study of the molecular biological marker characterizing proliferation – the Ki-67 protein, which is the most informative and adequate diagnostic indicator [1,5]. The study of cell proliferation in various tissues is one of the fundamental tasks of morphology. Due to significant advances in molecular biology, morphologists now have the opportunity to immunohistochemically label cells at different phases of the cell cycle [4].

To evaluate the characteristics of the proliferative activity of cellular elements of the fallopian tube in women of reproductive age during tubal pregnancy by immunohistochemical analysis of the expression of the Ki-67 marker.11 reproductive-age women (18-45 years) with tubal ectopic pregnancies were examined. All patients underwent surgery at the clinical sites of the departments of Obstetrics and Gynecology №1 and №2 of SEI Avicenna Tajik State Medical University. Serial histological sections of 3 µm thickness were prepared for immunohistochemical (IHC) analysis. The study protocol included standard preanalytical preparation steps: deparaffinization, rehydration, and high-temperature antigen retrieval. In order to assess the proliferative activity of tissues, incubation with polyclonal antibodies to the Ki-67 marker was carried out. The Ki-67-immunopositive cell index was determined as the ratio of the number of positively stained nuclei to the total number of cells (as a percentage per 100 counted cells) [3]. The laboratory stage of the work was carried out on the basis of the biotechnology laboratory of the Republican Oncology research center of the Ministry of health and social protection of population of the Republic of Tajikistan (RORC MHSPP RT).

Results of the study and their discussion. During an immunohistochemical study of fallopian tube fragments during tubal pregnancy, the expression level of the Ki-67 marker and the proliferation index in various fallopian tube structures were assessed.

The results showed that proliferative activity (Ki-67 index) was pronounced in both the epithelium and stroma of the fallopian tube (FT). Morphological analysis revealed that the Ki-67 marker localizes in cell nuclei with varying intensity and is distributed between the karyoplasm and nucleolus, corresponding to different phases of the cell cycle. Proliferative activity in the stroma during tubal pregnancy (TP) was 45-50%, while in women in the control group, the proliferation levels remained consistently low, did not exceed 10-12%. A study of the expression of the Ki-67 marker in the fallopian tube during tubal pregnancy revealed differences in the level of cell proliferation. This indicator indicates cell cycle activation, which is characteristic of reactive changes and inflammatory processes accompanying EP (Figures 1 and 2).

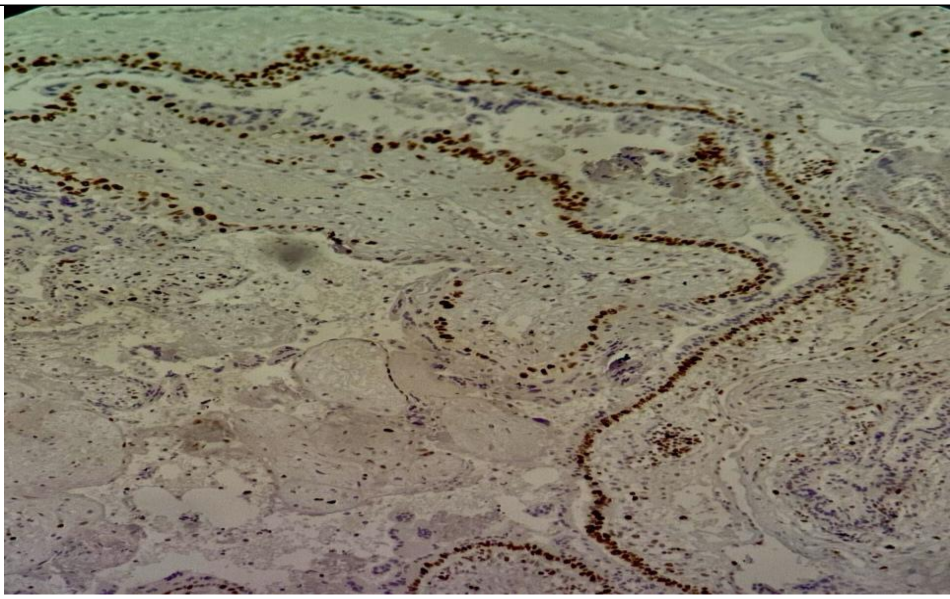


Figure 1. – The level of Ki-67 expression and proliferative activity of the fallopian tube mucosa. In the embryo implantation zone. Staining: DAB-hematoxylin. Magnification 80.

Figure 2. – The level of Ki-67 expression and proliferative activity of the fallopian tube mucosa in the area of fertilized egg implantation. Staining: DAB-hematoxylin. Magnification 80.

Thus, the different subcellular localization of Ki-67 confirms its ability to label cells at different stages of proliferation and mitotic activity.

Referenecs:

1. Sulaimanova, R.T. Expression of the Ki-67 marker during prenatal administration of estrogens to the ovaries of the offspring. Bulletin of new medical technologies. Electronic publication. - 2022, №3. - P. 73-77.
2. Titova, G.P. Structural changes in the fallopian tubes in patients with ectopic pregnancy [Text] / G.P. Titova, M.M. Damirov, I.V. Anchabadze [et al.] // N.V. Sklifosovsky Journal Emergency Medical Care. 2020; 9 (4): 598-605. https://doi/org/10.23934/2223-9022-2020-9-4-598-605.
3. Sharafutdinova, L.F. Morphofunctional changes in the nervous, immune and reproductive systems under the influence of nanosized titanium dioxide in the form of rutile: author's abstract of diss… d.b.s. Kazan, 2019. 43 p.
4. Chibisova G.M., Khabarov S.V. Kompleksnoe definition oncomarkers SA125, NE4 and indexa ROMA kak factor prognoza development cancer yaichnikov [Complex definition of oncoprotein CA125, HE4 and ROMA index as a factor in the prognosis of ovarian cancer]. Journal of New Medical Technologies.2018;25(3):15-20/DOI. 10.24411/1609-2163-2018-16158.
5. Khaskhachikh, D.A., Potapov O.V., Poslavsk O.V. Factors of resistance to protests therapy in endometriosis hyperplasia in women. Morphologia. - 2023, №17(1). P. 56-62.
6. Mullany, K., Minneci M., Monjazeb R.C., Coiado O. Overview of ectopic pregnancy diagnosis, management, and innovation. *Women's Health (Lond)*. 2023 Jan-Dec. 19:17455057231160349. [QxMD MEDLINE Link] .

7. Temirova, D.O. Ectopic pregnancy is a disease requiring emergency assistance | Text] / D.0. Temirova // Modern Science and Research. - 2025. - 4(2), - P. 920-928. Retrieved from https://inlibrary.uz/index. Php/science-research/article/view/67580.

8. Tulandi, T. Patient education: Ectopic (tubal) pregnancy (Beyond the Basics) [Text] /T. Tulandi // www.uptodate.com © 2025 5, 12:15 Up To Date. Inc. and/or its affiliates. All Rights Reserved.

9. Sargin, M.A. Is interstitial pregnancy clinically different from corneal pregnancy? A case reports. [Text] / M.A. Sargin, N. Tug, S. Ayas [at al.]. // Clin Diagnosis Res. (2015) 9(4): QD05-6. doi : 10.7 860/ JCDR/ 2015/ 12198.5836.

UDC 616.24-006.6:615.849.1

RESULTS OF INTRAARTERIAL CHEMOTHERAPY INTO BRONCHIAL ARTERIES IN LOCALLY ADVANCED NON-SMALL CELL LUNG CANCER

Issametov D.
Oncologist-chemotherapist of the highest qualification category, interventional surgeon, master of health organization, head of the department of chemotherapy and endovascular oncology of the State Center of Healthcare in Shymkent, chief freelance chemotherapist at the Healthcare Institution of Shymkent, Republic of Kazakhstan
Abdrakhmanov R.
PhD, Oncologist-chemotherapist of the highest qualification category,Head of the Chemotherapy Department, KazIOR,Almaty, Republic of Kazakhstan
Datkayeva F.
Oncologist-chemotherapist of the highest qualification category, doctor of the department of chemotherapy and endovascular oncology of the State Center of Healthcare in Shymkent, Republic of Kazakhstan
Amangeldy A.
Oncologist-chemotherapist of the department of chemotherapy and endovascular oncology of the State Center of Healthcare in Shymkent, Republic of Kazakhstan
Ahmedov B.
Oncologist-chemotherapist of the department of chemotherapy and endovascular oncology of the State Center of Healthcare in Shymkent, Republic of Kazakhstan
Abdimanapova A.
Oncologist-chemotherapist of the department of chemotherapy and endovascular oncology of the State Center of Healthcare in Shymkent, Republic of Kazakhstan
https://doi.org/10.5281/zenodo.18840576

Abstract

Background: Lung cancer remains a pressing issue in modern clinical oncology worldwide. Approximately 60-70% of newly diagnosed lung cancers are inoperable, necessitating the search for alternative treatment methods, including drug therapy. However, the issue of sensitivity to existing treatment regimens remains a significant aspect and is a key factor determining survival and local control.

Keywords: primary lung cancer, locally advanced inoperable lung cancer, transarterial chemotherapy, targeted therapy

Abstract. Lung cancer, despite modern trends in the development of medical oncology and science, remains one of the key tasks of clinical medicine.

On the one hand, this is due to the fact that the overwhelming number of patients at the time of initial diagnosis have an unresectable/inoperable stage of the disease, on the other hand, the latest predictive indicators in the world still indicate a high incidence of morbidity and mortality in terms of the structure of cancer incidence.

In this regard, there is an urgent need to find alternative ways to treat the malignant process, including drug therapy. Drug therapy for lung cancer in the world today is represented by an extensive list of optional cytostatics, targeted agents, and immunoages.

In the Republic of Kazakhstan, these data are statistically different from those of the Russian Federation in terms of the structure of cancer incidence per capita and amount to 3,615 new cases for 2021 and 3,925 identified new cases of lung cancer for 2022. [2,4].

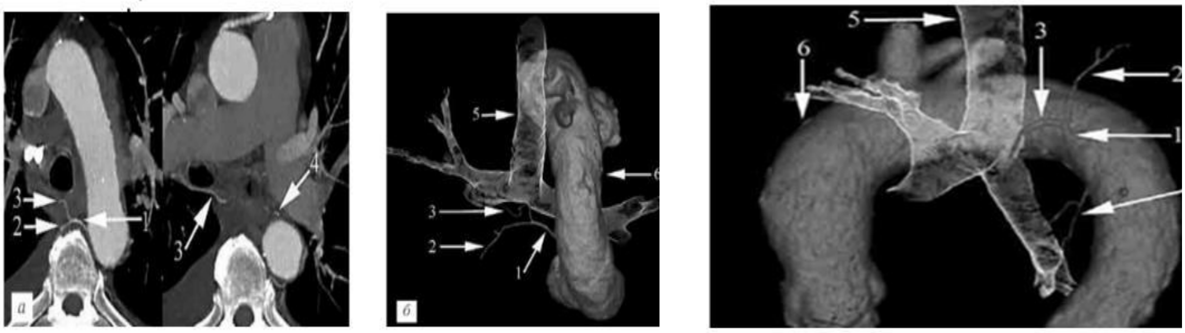
Almost 40-50% of cases of verified lung cancer do not carry a detectable mutational load; accordingly, the targeted use of genetically modified therapy options with an inherent selective effect is not applicable in these cases, but is limited only by a combination of existing chemotherapy regimens, which also have an expected and very limited therapeutic effect [3,5 ,6,7].

In this regard, in a number of cases, we widely use the technique of loco-regional intra-arterial chemotherapy in the branch of the bronchial artery on the affected side of the right or left lung with an increase in the bioavailability of cytostatics with its local effect and minimal toxic effect in unresectable lung cancer.

The purpose of the study was to evaluate the possibility of using intra-arterial treatment methods for unresectable locally advanced non-small cell lung cancer.

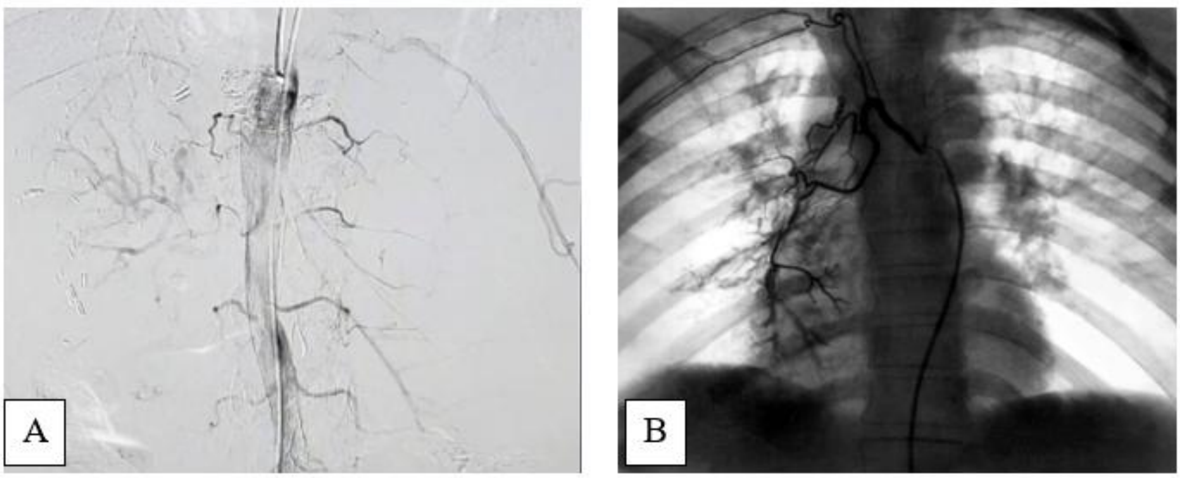
Materials and methods.

The technique of intra-arterial chemoinfusion is reduced to puncture and puncture followed by catheterization of the femoral artery using the Seldinger technique in the projection of the right/left inguinal fold. Selective catheterization of the bronchial artery is carried out by aortography of the thoracic aorta with identification of the branches of the bronchial artery (bronchial arteries arise at the level of ThV, ThVI, the so-called orthotopic option). In any variant of origin, the pulmonary arteries enter the parenchyma in the area of their roots and run parallel to the bronchi, located in their outer connective tissue layer. Cauldwell et al. [8] described four classic variants of the origin of the bronchial arteries from the aorta: (a) the first type (40.6% of cases), when two bronchial arteries arise from the aorta on the left, and one on the right from the intercostal-bronchial trunk (IRB); (b) the second type (21% of cases), when there is one bronchial artery on the left and one MRBS on the right; (c) the third type (20% of cases), when two bronchial arteries arise from the left and two from the right, while one of the arteries on the right is presented as MRBS;(d) the fourth type (9.7% of cases), in which one artery is identified on the left and two on the right, while one of the arteries on the right is presented as MRPS. The left bronchial artery usually arises from the anterolateral surface of the descending thoracic aorta and follows the direction of the root of the left lung. On the right, the intercostal-bronchial trunk usually arises from the right posteromedial surface of the descending thoracic aorta and follows in a cephalad direction until it divides into the intercostal (or intercostal) artery and the right bronchial artery. The latter then turns in a caudal direction and follows along the right main bronchus. The duration of the procedure is 20-30 minutes. The procedure time from the beginning of the procedure when working in “4 hands” (2 angiosurgeons) is 25-35 minutes. (depends on the anatomy of the vessels). As the catheter is installed, the infusion of cytostatics continues for 60 minutes each, followed by compression of the femoral artery area until complete hemostasis (10-15 minutes) or suturing the puncture hole with an Angiosil device measuring 6-7 F.

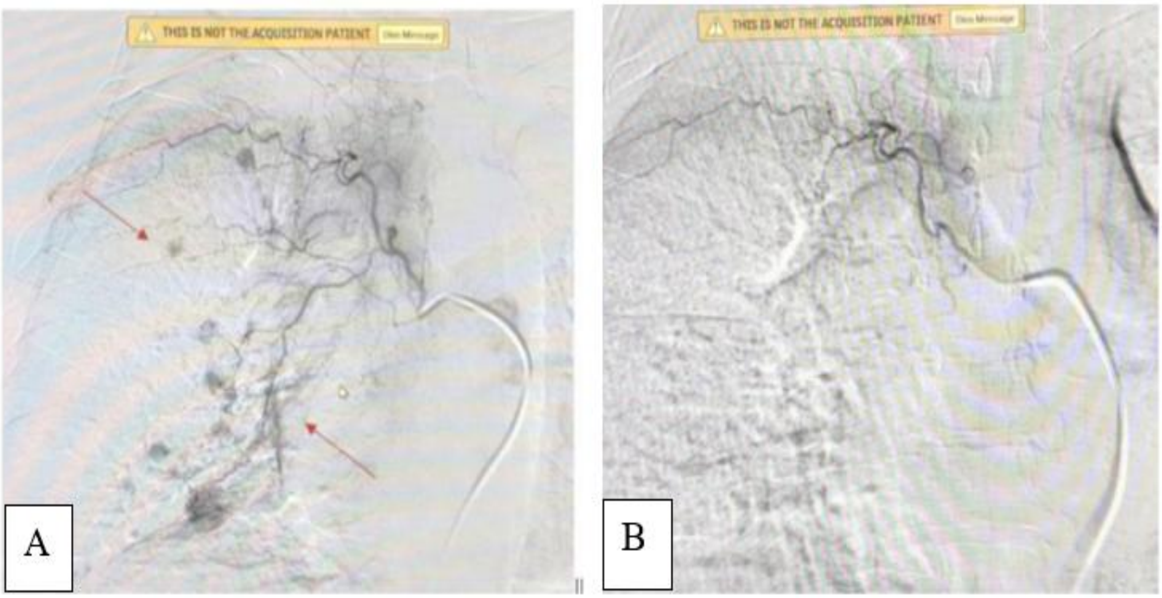


Pic.1 *A typical variant of the origin of the bronchial arteries from the aorta (type 1 according to Cauld well et al., type 2 according to Yoshiaki et al.). One right bronchial artery arises from the intercostal-bronchial trunk, two left bronchial arteries arise from the descending part of the thoracic aorta through a common trunk: and fragments of computed tomograms in MIR mode into three-dimensional reconstructions of bronchial arteries. 1 intercostal bronchial trunk, 2 intercostal artery, 3 right bronchial artery, 4 two left bronchial arteries immediately below the origin of the aorta, 5 trachea and bronchi, 6 larynx.*

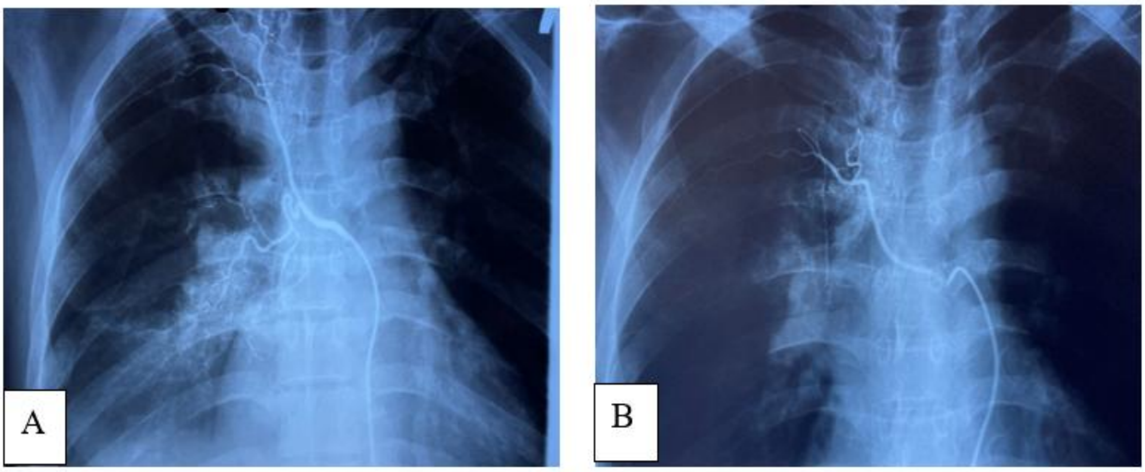
All of 52 patients aged 52 to 60 years were treated in the period from 2022 to 2024. In all patients, stage IIIa-IIIb of the process was established (T3N0M0 – T2-3bN1M0). All patients at the time of initial or repeated treatment (relapse) had a histologically verified variant of non-small cell lung cancer. Morphologically, lung adenocarcinoma was detected in 18 patients (28,1%), squamous cell carcinoma was detected in 46 patients (71,9%). In all cases, radiological confirmation of the local extent of the process was noted through CT/MRI/PET/CT examinations. All patients received intra-arterial chemotherapy in the branch of the bronchial artery according to the “TR” regimen (Docetaxel 75 mg/m2 + Cisplatin 75 mg/m2) in a standard therapeutic dose. The interval between courses of intra-arterial chemotherapy was carried out once every 21 days from the date of administration of chemotherapy. Among the side effects, one can note the inherent cytotoxicity of the drugs with manifestations of anemia, leukemia and thrombocytopenia, which were corrected through the use of L-CSF (leukocyte colony-stimulating factor), glucocorticoids and targeted symptomatic replacement therapy.



Pic. 1 *Exploratory angiography at the stage of bronchial artery catheterization (A). Angiography of the right brochialis artery, anatomy variations (B).*



Pic. 2 *Angiography a. bronchialis on the right.*
Central cancer of the right lung with intraorgan metastatic foci.
Selective catheterization (A). Condition after 3 courses of HIBA (B)



Pic. 3 *Arteriography a. bronchialis on the right.*
Central cancer of the right lung with intraorgan metastatic foci.
Selective catheterization (A). Condition after 3 courses of HIBA (B)

Results. Analysis of the immediate results of treatment showed the effectiveness of the technique in 52 patients (81.25%). Thus, out of 64 patients, 34 primary patients (53.1%) with a clinical diagnosis of lung cancer after 3 courses of transarterial chemoinfusion with the “Docetaxel+Cisplatin” regimen, according to the control examination (FBS, CT), partial regression of the process was noted, 2 (3,1%) patients noted progression of the process due to a violation of the intercourse interval due to the duration of recovery of the blood picture after a course of chemotherapy. In 28 patients, stabilization of the process with positive dynamics was noted, which made it possible to transfer the unresectable stage of the disease to a resectable tumor process. All 52 patients, within an interval of up to 3 weeks after the third course of transarterial chemoinfusion, based on the results of control examinations, underwent radical surgical treatment. Thus, in 28 (43.7%) patients with partial regression of the process, forehead surgery and bilobectomy were performed, in the remaining 22 (34.3%) patients an extended combined pneumonectomy was performed.

Discussion.The effectiveness of the transarterial administration of cytostatics for primary non-small cell lung cancer was demonstrated by clinically significant tumor regression after 1 course of chemoinfusion in 12 patients (18.75%), stabilization of the process in 20 patients, progression of the process in 1 patient. Immediate postoperative complications were observed in 3 patients (hematoma of the puncture area of the femoral triangle according to Seldinger). The median follow-up in all patients in the postoperative period exceeded 18 months. According to various authors, it is quite variable [3,4,6,9].

Among the complications inherent in chemotherapy, we also noted grade 2-3 emetogenic reactions. All these phenomena were manageable with accompanying therapy and did not require blood transfusion additions.

Conclusions. The use of minimally invasive interventional regional methods for treating primary non-small cell lung cancer demonstrates a high overall response rate, a statistically significant median survival as part of complex therapy, and an improvement in the quality of life of patients.

References:

1 A.D. Kaprin, V.V. Starinsky, A.O. Shakhzodova. Malignant neoplasms in Russia in 2021. (morbidity and mortality). Moscow 2022. Page 14.

2 Kaidarova D.R., Oncological care in the Republic of Kazakhstan for 2023. "Kazakhstan Pharmaceutical Bulletin". 2024 Page1.

3 Issametov D., Maulenov Z.O., Ismailov S.T., Assabayeva R.I «East Europian science journal».,part 1. 43-47 p. doi:10.31618/ESSA.2782-1994.2025.1.109

4 Isametov D.,Gantsev Sh.,Maulenov Zh.,Arybzhanov D.,Tursumetov D., Materiales de la coтayrencia cientifica internacional extranjera «Tendencias en el dessarollo cientifico y desafios mundiales», «Resultados de laquimionfusion en la arteria bronquial en pacientes con cancer de pulmon de celulas no pequenas localmente avanzado» Junio 2024 /doi 10.37539/240620.2024.63.48.001, ISBN 978-5-00213-288-1

5 Gantsev Sh.Kh., Gantsev K.Sh., Kaidarova D.R., Kzyrgalin Sh.R., Kobyakov G.L., Kudaibergenova I.O., Lipatov O.N., Lyadov V.K. , Polatova D.Sh., Rakhmatullina I.R., Safin Sh.M., Semiglazov V.V., Tryakin A.A., Ulrich E.A., Umezawa K. – M., "Oncology" 2023. – 704 p.

6 Alaswad M. «Locally advanced non-small cell lung cancer: current issues and recent trends». Rep Pract Oncol Radiother. 2023 Jun 26;28(2):286-303. doi: 10.5603/RPOR.a2023.0019.

7 Hendriks JMH, Billiet C, De Bondt C, Van Schil PE. «Multimodality treatment including surgery related to the type of involvement in locally advanced non-small cell lung cancer. Cancers» (Basel). Cancers (Basel). 2022 Mar 25;14(7):1656 doi: 10.3390/cancers14071656. 2022 Mar 25;14(7):1656. doi: 10.3390/cancers14071656.

8 Chen X, Tong X, Qiu Q, Sun F, Yin Y, Gong G, Xing L, Sun X.«Radiomics nomogram for predicting locoregional failure in locally advanced non-small cell lung cancer treated with definitive chemoradiotherapy». Acad Radiol. 2022 Feb;29 Suppl 2:S53-S61. doi: 10.1016/j.acra.2020.11.018.

9 J.Thorac Dis. Chidi AP, Antonoff MB «Hybrid video-assisted thoracoscopic sleeve lobectomy: an elegant minimally invasive technique for locally advanced non-small cell lung cancer». 2023 Apr 28;15(4):1533-1535. doi: 10.21037/jtd-23-77.

10 Cauldwell E. W. The bronchial arteries: an anatomic study of 105 human cadavers / E. W. Cauldwell, R. G. Siekert, R. E. Lininger et al. // Surg. Gynecol. Obstet.— 1948.— Vol. 86.— P. 395–412.

PEDAGOGICAL SCIENCES

BLENDED LEARNING ONE OF THE TOOLS TO INCREASE INVOLVEMENT OF STUDENT AND CREATE SOFT SKILLS

Polumeeva I.
Saint-Petersburg University of Management Technologies and Economics
Saint-Petersburg
https://doi.org/10.5281/zenodo.18840578

Abstract

The article by Polumeeva I.N. considers the issue of increasing student involvement in the educational process through blended learning, the impact of blended learning on the development of soft skills, indicates student satisfaction with blended learning, the advantages and disadvantages of blended learning from the perspective of students are noticed, draws attention to the models of blended learning that can improve the effectiveness of subject learning.

Keywords: blended learning, traditional learning, soft skills, ECTS system, feedback, educational portal.

The challenges of today's rapidly changing society, where skills become obsolete every 3-5 years, increase the requirements for graduates in terms of their academic achievements and competencies. The use of blended learning technologies can contribute to improving the system of professional and personal competencies of graduates. [1, p. 240] Every graduate requires both hard skills (professional skills) and soft skills (non-professional skills) that help them in their work, communication, decision-making, and adaptation to new conditions. Blended learning promotes the formation of soft skills, as this approach combines traditional learning methods with distance and online methods, which allows you to develop creativity, critical thinking, communication, self–discipline, and time management.

The combination of traditional teaching methods with digital technologies provides an opportunity to increase students' involvement in the educational process. A survey on the satisfaction of the blended learning was conducted among first-year students (who were studied in a blended learning) at the end of the first semester. A total of 75 students participated in the survey. As shown in Figure 1, 92% of the respondents expressed their satisfaction with the blended format of education, while 5% were dissatisfied and 3% found it difficult to answer.

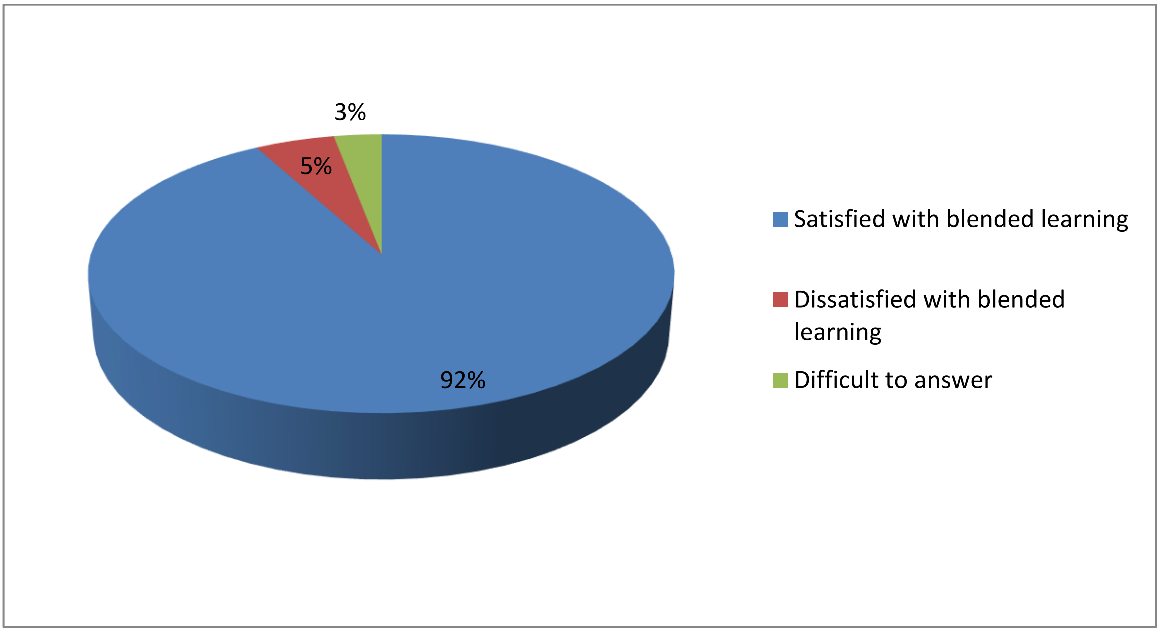


Figure 1. Satisfaction with blended learning among first-year students.

The same group of recipients was asked to fill out a questionnaire, where they could select several points, with the choice of the pros and cons of blended learning. It should be noted that the pros and cons of blended learning were compiled, according to the first-year students in the last academic year. The purpose of the questionnaire was to find out why the students showed a high level of satisfaction with this learning format. As can be seen from Figure 2, the most frequently selected point is the ability to monitor progress – 50 appeals and control deadlines 35 appeals. And this is not surprising in the conditions of the ECTS system, when it is important for a student to be aware of his/her rating not only at the end of the semester, as a result of work on the subject, but also during the semester, in order to be able to control the mastery of the subject. As we can see here, it is important for students to be informed during their studies, to be independent and self-disciplined in following deadlines.

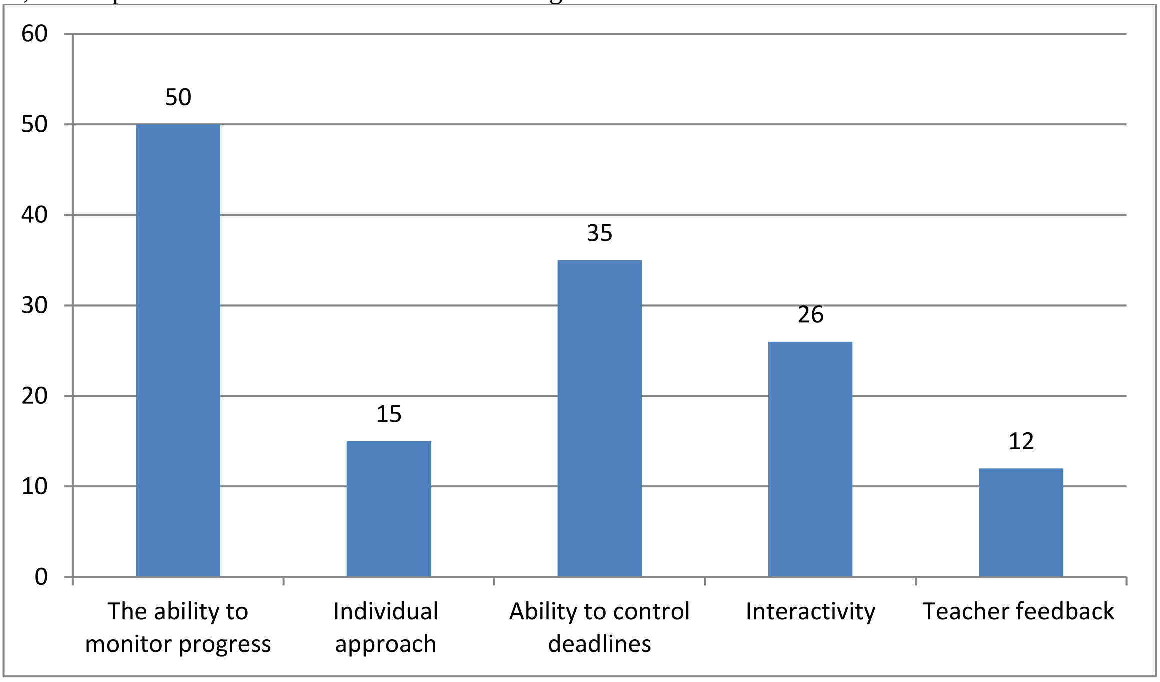


Figure 2: Benefits of blended learning.

It should be noted that an electronic course on the discipline "Foreign Language" was created for students studying in a mixed format on the university's educational portal, where topics were specified and study materials were provided. The portal also features an electronic attendance log for classroom sessions. The course includes homework assignments, such as completing tests and exercises, preparing presentations on specified topics, and following a minimum number of slides and a presentation plan provided by the teacher. Students are responsible for preparing their own information and questions for the audience. In the classroom the teacher explained the material and answered the questions that arose during self-study, conducted online testing, followed by analysis, and so on. Here we see the use of two mixed learning models - the Face-to-Face driver model and the Flipped classroom model. One more model should be mentioned, which received a wide response from students and was used in 100% of cases – this is the Flex Model. If a student misses classes for a valid reason (such as illness), they may lose points, and if the absences are prolonged, they may fall behind the rest of the class in terms of learning the subject. To prevent this, such students have the opportunity to complete assignments and tests alongside the rest of the class in the course on the educational portal, and if they have any questions, they can contact the teacher by sending a message on the educational portal. The presence of a electronic course on the educational portal contributes to improving the effectiveness of the discipline, as all stakeholders have the opportunity to view the material at any time, monitor the students' ratings, and check the attendance of classroom sessions. [2] The blended format allows students being not in the classroom to receive feedback from the teacher. This was also indicated by the recipients in the questionnaire, with 12 responses (Figure 2). A mixed format is also an individual approach that allows students to learn at their own pace, for example, some students may need one hour to prepare a presentation, while others may need one day. In the questionnaire, the individual approach was selected 15 times. In the online communication format, the teacher adjusts the level and intensity of the workload for individual students, which allows the entire group to participate in the organization of time and level of learning, and increases student motivation [3, p. 601-602].

The questionnaire also asked the recipients to select the disadvantages of blended learning. As can be seen in Figure 3, the most common issue was low internet speed, with 15 responses, followed by additional time and increased computer usage, with 9 and 8 responses, respectively. Additionally, 5 respondents indicated that they found it inconvenient to work on a phone screen, and only 3 respondents expressed dissatisfaction with electronic textbooks.

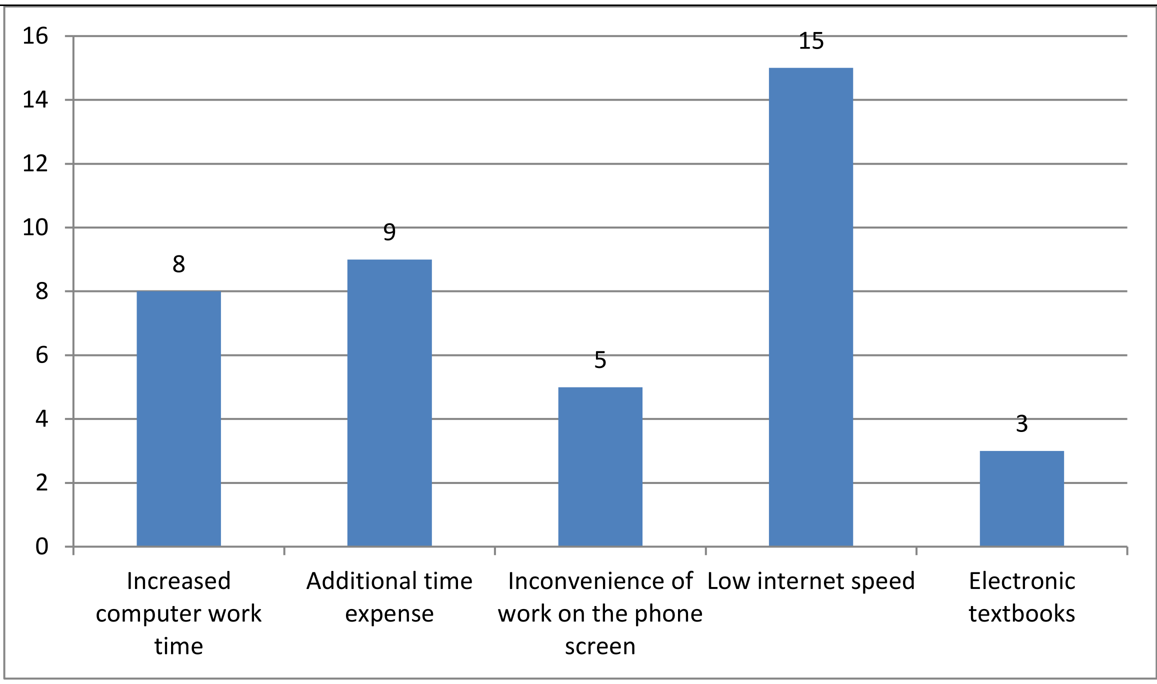


Figure 3. Student-perceived disadvantages of blended learning.

Conclusion

With the use of blended learning models, students' motivation to study the subject increases, the educational trajectory and assessment of results are clear, and there is a personal-oriented aspect and increased degree of responsibility of the participants in the educational process.[4] Thus, students have to use also soft skills to study the subject.

References:

1. Chupanov A.Kh., Gasanova D.I. The specifics of using educational technologies in the implementation of elements of distance and blended learning in higher education programs // MNKO. 2024. No. 3 (106). p. 240-242.

2. Krechetnikov K.G. Features of the organization of blended learning// Modern problems of science and education. 2019 No.4 URL: https://science-education.ru/ru/article/view?id=29019 (date of access: 25.02.2026)

3. Ostapuk S.I., Larina T.V. Integration of elements of the blended learning educational concept in teaching a business foreign language at the Business Institute of BSU // Collection of articles of the International scientific and practical conference. Minsk. 2022 p. 601-603

4. Yanchenko I.V. Blended Learning at a University: From Theory to Practice// Modern Problems of Science and Education. 2016 No. 5 URL: https://science-education.ru/ru/article/view?id=25417 (date of access: 26.02.2026).

ISSUES OF THE METHOLOGY OF MUSICAL HESRING DEVELOPMENT

Papanyan N.
Vanadzor state university named after H.Tumanyan,
Faculty of Social Sciences,
Department of Art and Sports, Lecturer, Armenia

ԵՐԱԺՇՏԱԿԱՆ ԼՍՈՂՈՒԹՅԱՆ ԿԼԱՍՏԵՐԱՅԻՆ ԶԱՐԳԱՑՄԱՆ ՄԵԹՈԴԻԿԱՅԻ ՀԱՐՑԵՐ

Նելլի Պապանյան
Դասախոս, հասարակական գիտությունների ֆակուլտետ,
արվեստի և սպորտի ամբիոն, Վանաձորի
Հ.Թումանյանի անվան պետական համալսարան, Հայաստան
https://doi.org/10.5281/zenodo.18840580

Abstract

- The article analyzes the goal-oriented process of simultaneous complex development of all component-aspects of musical consciousness in the "Solfeggio" course of the Children's Music School - Intonational hearing education and diverse methodical means and forms included in its development are presented

- The methodological problem of auditory assimilation and meaning of the patterns of coordination of the tonal structure of the 1st-3rd grades of the Children's Music School is emphasized

- The hypothetical "house" model used from the musical arsenal of the "Solfeggio" subject

- The problem of mastering and understanding the music of the 20th century and the rooting of new methodological and pedagogical principles of musical ear development

-Detailed and in-depth study of the national tonality system within the scope of the "Solfeggio" subject of the Children's Music School

- The development of the sense of metre rhythm, new approaches to the evaluation system

- The development of stylistic taste, intonation and the development of metrical hearing and memory as a comprehensive process.

Սեղմագիր. Հոդվածում դիտարկված են մանկական երաժշտական դպրոցում «Սոլֆեջո» առարկայի դասավանդման մեթոդական խնդիրներ՝ հիմնված ազգային լադաինտոնացիոն հենքի վրա և կապված

- ինտոնացիոն ընկալման,
- մետրառիթմի զգացողության,
- երաժշտական լսողության և հիշողության զարգացման
- գեղագիտական ճաշակի և գեղարվեստական մտածելակերպի ձևավորման հետ:

Ինտոնացիոն լսողության զարգացումը բաժանված է 2 կլայստերի՝

- ինստրուկտիվ-նախապատրաստական
- գործնական-նպատակային

Կարևորվում է երաժշտական ընկալումների զարգացման համակարգի ձևավորումը թե՝ դասական, թե՝ ժամանակակից երաժշտության յուրացման ոլորտում:

Keywords: intonation, musical ear, memory, tonality, metre, interval, chord, cluster.
Հանգույցային բառեր. ինտոնացիա, երաժշտական լսողություն, հիշողություն, լադատոնայնություն, մետրառիթմ, ինտերվալ, ակկորդ, կլաստեր

Ներածություն.Մանկական երաժշտական դպրոցում «Սոլֆեջո»-ն առաջին սիստեմատիկ դասընթացն է երաժշտատեսական առարկաների շարքում, որը աշակերտներին տալիս է կայուն հիմք հետագա երաժշտական զարգացման համար: Սոլֆեջոն գործնական դասընթաց է, որի խնդիրն է սովորողների մոտ ձևավորել գործնական հմտություններ՝ երգեցողություն նոտաներով, թերթից ընթերցում, երաժշտական թելադրություն, լսողական վերլուծություն: Առարկայի կարևոր նպատակներից է համակարգված գիտելիքների և հմտությունների ձեռքբերումը երաժշտության արտահայտչական համակարգի ընկալման և գործնական կիրառման համար: Մանկավարժի հիմնական խնդիրն է սովորողների երաժշտական ընդունակությունների համակողմանի և ներդաշնակ զարգացումն ու դաստիարակումը, գեղեցիկի և հնչյունային աշխարհի նպատակաուղղված հույզային ընկալումը:

Այս հոդվածի նպատակն է օգնել երիտասարդ մանկավարժներին սոլֆեջոյի դասերին մոտենալ ստեղծագործաբար՝ կիրառելով նորագույն մեթոդներ և ուսուցողական տարբեր հնարներ, որոնք դեռևս այդքան էլ լուսաբանված չեն երաժշտամեթոդական գրականության մեջ:

Նյութը և մեթոդները.Ծփումը գեղեցիկի հետ զուգակցվում է մարդու հոգևոր ներաշխարհի, իդեալների և համամարդկային արժեքների իմաստավորման հետ: Այդ պատճառով երաժշտական դասերն իրենց նշանակությամբ համընդհանուր գեղագիտական զարգացման կարևորագույն մասն են կազմում: Երաժշտական արվեստի էությունը և ստեղծագործությունների գեղարվեստական իմաստավորումը, արժևորումը տեղի է ունենում երաժշտական կրթության բոլոր օղակներում և բոլոր դասերի ժամանակ: «Սոլֆեջո» առարկայի դասընթացը նպատակաուղղված է երաժշտական լսողության և ավելի լայն իմաստով երաժշտական գիտակցության բոլոր բաղադրիչ-կողմերի համընթաց կոմպլեքսային զարգացման խնդիրն իրագործելուն: Այդ բոլոր կողմերի համակցումը և փոխլրացումը, որպես զուգահեռ տարվող մանկավարժական մեթոդական աշխատանք, հիմնարար գերխնդիրն է և դասավանդման պրոցեսը: Այդ համակցված խնդիրների հանգույցի մեջ կարելի է առանձնացնել ինտոնացիոն լսողության, մետրառիթմի զգացողության, երաժշտական հիշողության, լսածի վերլուծության, ճաշակի, գեղարվեստական-գեղագիտական կողմերը զարգացնելու մեթոդական հնարները: Անկասկած, ինտոնացիոն լսողության դաստիարակությունը ամենից առաջնայինն է իր մեջ ներառված բոլոր մեթոդական միջոցների և ձևերի ամենաբազմազան կիրառմամբ:

Քանի որ ինտոնացիոն լսողության դաստիարակությունը բաժանվում է երկու ինքնուրույն պլանի (առաջինը կարելի է բնութագրել որպես ինստրուկտիվ-նախապատրաստական, որի ենթամասերն են կազմում ինտոնացիոն վարժանքները, կոնկրետացված թեստերը, երկրորդը կարելի բնութագրել որպես պրակտիկ-նպատակային, որի ենթամասերն են թերթից ընթերցումը, երաժշտական թելադրության ունկնդրումը և գրառումը), ապա սկսենք նրանց մեթոդաբանության պարզաբզանումը վոկալ-ինտոնացիոն կարողությունների զարգացման խնդիրներից: ՄԵԴ-ի «Սոլֆեջո» առարկայի հիմնական մեթոդական խնդիրն է I-ից մինչև III -րդ դասարանը լադատոնայնական կառույցի համակարգման օրինաչափությունների լսողական յուրացումը և իմաստավորումը: Այն պետք է գիտակցվի և ընկալվի որպես հիմնականում դիատոնիկ մաժոր և մինոր լադերի կառույց-բլոկ (կարկաս): Նպատակահարմար է դա բացատրել որպես լադի կմախքը (արմատուրա-կարկասը) կազմավորող սյուների (երեք աստիճանների) փոխհարաբերակցում (դրանք են ներքևի տոնիկան որպես I աստիճան, V դոմինանտային աստիճանը, որը բաժանում է լադը 2 տետրախորդների և վերևի տոնիկան, VIII կամ I աստիճանը որպես օկտավային ամբողջականացում): Իսկ I և V աստիճանների միջև հանդես եկող III աստիճանը դիտարկվում է որպես լադի տեսակը տարբերակող, զանազանող աստիճան: Լադի անկայուն աստիճանների հարաբերակցությունը տոնիկայի հետ իմաստավորված կապի բացատրելուց հետո, կարելի է բացատրել նաև V աստիճանից ցած գտնվող անկյուն աստիճանների (II-ը, IV-ը) վարընթաց լուծումները, և V աստիճանից վեր գտնվող անկյուն աստիճանների (VI-ը, VII-ը) վերընթաց լուծումները: Սակայն բացատրություններով սահմանափակվելն այնքան էլ նպաստավոր չէ ցածր դասարաններում, ուստի խաղերգային տարրերի կիրառումը գրեթե անհրաժեշտ է I-ից մինչև III դասարանների աշակերտների հետ մեթոդական աշխատանքի պրակտիկայում: [1]

20-րդ դարի բոլոր առաջատար դպրոցների «Սոլֆեջո» առարկայի խաղերգային զինանոցից ես նախընտրում եմ ենթադրյալ կամ հիպոտետիկ տնակի մոդելը: Այդ մոդելի մեջ լադի I աստիճանը տնակի հիմքն է, իսկ III աստիճանը շքամուտքն է (որը նաև մաժոր և մինոր լադերը տարբերակող աստիճանն է): Մյուս բոլոր աստիճանները տան տարբեր հարկերն են: Քանի որ ուսուցիչը սովորեցնում է աշակերտներին լավ և արագ կողմնորոշվել այդ տնակի մեջ, ապա մաժոր և մինոր լադերի կառուցվածքին համապատասխան հարևան աստիճանների ամբողջատոնային և կիսատոնային հարաբերակցումը լավ է նպաաստում աստիճանների տոնային հարաբերությունները լսողությամբ յուրացնելու համար:

Այդ մոդելի մեջ լավ է իմաստավորվում նաև աստիճանների վերընթաղ և վարընթաց կարգը հիմքից մինչև տանիքը և հակառակը: Լադի VIII աստիճանը իբրև տանիքը (I աստիճանի օկտավային կրկնությունը վերևում) լավ է ամբողջացնում լադը և որպես տանիք կարծես միահյուսում-շաղկապում է տեսողական-ենթիմաստային և լսողական տպավորությունները լադի ամբողջականության վերաբերյալ: Մեթոդական իմաստով լադի կայուն և անկայուն աստիճանների գիտակցման և լսողությամբ յուրացնելու ունակությունները «տեղ հասցնելու» ավանդական միտումը յուրաքանչյուր կայուն աստիճանի շուրջը «պտտվող» անկյուն աստիճանների լուծման խորը, համակողմանի ըմբռնումն է: Չցանկանալով կրկնել ավանդական մեթոդաբանության մեջ վաղուց ի վեր արմատավորված աշխատանքի ձևերը և սկզբունքները, ես ուզում եմ մատնանշել հետևյալը, որ լադատոնայնության ըմբռնումը ժամանակակից երաժշտության դիրքերից այսօր պարտադրում է որոշ իմաստով մոտենալ այն դասվանդելու խնդիրներին: Ելնելով այդ դիրքերից ժամանակակից լադատոնայնության ըմբռնելու լսողությանն դաստիարակության խնդիրները որոշ իմաստով այլ են: Իմ կարծիքով, այսօր այլևս հնարավոր չէ բացառել 20-րդ դարի երաժշտության յուրացումը, իմաստավորումը և դրա հետ կապված լսողության զարգացման նոր մեթոդական – մանկավարժական սկզբունքների արմատավորման պրոբլեմը: Այդ հիմնախնդրի ամենադժվար կողմը կապվում է ժամանակակից երաժշտության մեջ գերիշխող այսպես կոչված դիսոնանտային տոնայնության յուրացումը ապահովող նոր մեթոդիկայի ստեղծագործական մշակման հետ: Հրատապ եմ համարում այն հարցը, թե ինչպես, օրինակ, սեկունդա ինտերվալը կարող է յուրացվել միանգամից երկու հարևան աստիճանների միաժամանակյա հնչմամբ: Այսպես դասատուն դաշնամուրի վրա նվագում է լադի I աստիճանը, իսկ աշակերտը պետք է սովորի նրա հետ միասին երգել լադի II աստիճանը: Դա լավ գիտակցելու համար առարկան դասավանդող դասատուն, դիմելով XX-XXI դարերի երաժշտությունից քաղված օրինակներին, պետք է մատչելի և պատկերավոր ձևով բացատրի իր սաներին, որ սուր կամ դիսոնանտային հարաբերակցումը հարևան աստիճաննեի միջև մի դեպքում կարող է հանդես գալ իբրև անկայունության պարպում կամ լուծում (երբ անկայուն աստիճանը լուծվում է հարևան կայուն աստիճանի մեջ), մյուս դեպքում անկայուն սեկունդաների հանդես գալը կարող է կինել որպես ժամանակակից երաժշտության մեջ հաճախ հանդես եկող մեղեդային գծի շերտավորված հաստացում (կլաստեր): Ես կարծում եմ, որ շատ ճկուն ձևով պետք է կարողանալ զուգակցել դասական երաժշտության լադահարմոնիկ ֆունկցիոնալ ակորդիկայի մասնակցությամբ տարբեր (ծրագրով նախատեսված) տոնայնությունների գամմաների երգեցողությունը, տարբեր աստիճանների դասական լուծումը, անսովոր կլաստերային նվագակցությամբ երգելու հետ: [2]

Մյուս անկայուն աստիճանի՝ սեպտիմայի, յուրացումը կարելի է շատ մատչելի ձևով բացատրել նրանով, որ տոնիկական ակորդը նորագույն երաժշտության մեջ ոչ բոլոր դեպքերում լադի կայուն աստիճանների միաժամանակյա համակցումն է: Տոնիկական ակորդի մեջ սեպտիմայի առկայությունը օրինաչափ է ժամանակակից երաժշտության համար, ուստի I և VII աստիճանների միաժամանակյա «վերադրումը» նույնպես հարկավոր է շուտ մշակել արդեն ցածր դասարանների աշակերտների համար:

Այսպիսով, կլաստերների ընկալման և նրանց վերարտադրման տեխնոլոգիան 1-ից 3-րդ դասարանի աշակերտների համար պետք է մեթոդական իմաստով լավ մշակել: Ես սովորաբար աշխատում եմ, որքան հնարավոր է, շատ դիմել խաղերգային տարրերի գործնական կիրառմանը: Դրա պատճառով մեր փոխաբերական իմաստով տոնայնական մոդել-տնակը կարող է նաև կուբիկ-ռուբիկի նման ծալվել, նոր կառուցվածքային տեսք ստանալ, խոյանալ դեպի վեր կամ հորիզոնական ձևով «փռվել» հատակին (այս դեպքում, իհարկե, տնակի բոլոր հարկաբաժինները տեղադրվում են իրար կողքի, նույն հարթության վրա և դա պատկերավոր իմաստով հենց նմանվում է կլաստերին): Բացի դրանից, այդ տնակի մեջ կարող ենք նրա բոլոր հարկերը բարձրանալ և իջնել կամ ոտքով, կամ վերելակով և կամ միացնել I հարկում տեղադրված մոնիտորը և տեսնել բոլոր մյուս հարկերի հեռուստապատկերը: Երբ մենք ոտքով ենք բարձրանում կամ իջնում, դա աստիճանական շարժումն է, վերելակի միջոցով մենք արագընթաց ձևով կարող ենք թռչել տարբեր, հեռու գտնվող աստիճանների վրայով: Մոնիտորային խցիկը դա պահակի առանձնասենյակն է, որտեղ կարելի է լինել միանգամից բոլոր հարկերում:

Մյուս փոխաբերական իմաստով ես կլաստեր համեմատում եմ լավ մարզված կրկեսային ակրոբատի հետ, որն իր «շալակն է առել» մյուս բոլոր աստիճանները և ազատ քայլում է արենայի վրա: Անկասկած, ՄԵԴ-ի աշակերտների մոտ ժամանակակից լսողությունը կազմավորելու համար հարկավոր է որքան հնարավոր է շատ նվագել XX և XXI երաժշտությունը: Ես առանձնացրել եմ Ս. Պրոկոֆևի «Ալբոմ պատանեկության համար» դաշնամուրային պիեսների ժողովածուն: Սակայն դրա հետ միասին նվագում եմ պիեսներ Ա. Խաչատրյանի «Մանկական ալբոմ» (1-ին և 2-րդ տետր) շարքից, Բ.Բարտոկի «Միկրոկոսմոս» ժողովածուների շարքից, Կ.Օրֆի «Schurlwerk» շարքից, Պ. Հինդեմիտի, Դ. Միյոյի, Հ. Էյսլերի, Ի.Ստրավինսկու և այլ ժամանակակից կոմպոզիտորների գործերը: Հայ երաժշտության մեջ ես նախընտրում եմ Ստ. Նաղդյանի մանկական պիեսները, սակայն հարկ եմ համարում ցուցադրել նաև Ա. Բաբաջանյանի «6 պատկեր» դաշնամուրային շարքի պիեսները, Ա.Հարությունյանի, Է.Բաղդասարյանի, Տ.Մանսուրյանի և ավելի երիտասարդ սերնդի կոմպոզիտորների երկերը: Բոլոր դեպքերում չափազանց կարևոր եմ համարում լադա-ինտոնացիոն վարժությունները երգել և՛ solo, և՛ միաժամանակ սովորել երգել ու նվագել: Դիսոնանտայնության շեմը հաղթահարելու համար լավ կլիներ, երբ ուսուցիչը նվագելով տոնիկան առանձին կամ տոնիկական կվինտան սովորեցներ աշակերտներին երգել անկայուն աստիճանները: IV և V դասարանների աշակերտների համար անհրաժեշտ եմ համարում այն վարժանքների հաճախակի կիրառումը, երբ որևէ տոնայնության տոնիկական կվինտայի սահմանում բազմակի անգամ կրկնվող վերընթաց և վարընթաց գամմայակերպ աստիճանական շարժումը զուդակցվում է տարբեր աստիաններից սկսվող նույնանման հաջորդականության երգեցողությամբ: Այսպիսով, միաժամանակ նվագելը ու երգելը լավ է ներազդում ժամանակակից լադազգացողության զարգացման վրա: Գուցե V դասարանի աշակերտների մոտ կարող ստացվել նաև լադատոնայնական պրոեկցիան, այսինքն, տվյալ տոնայնության գաման կարողանալ երկձայն կամ եռաձայն երգել՝ օգտագործելով այդ տոնայնության աստիճանների գամմա-պրոեկցիան սկսած միաժամանակ I, IV, V աստիճաններից: Բոլոր դեպքերում աշակերտների դիսոնանտայնության զգացողությունը հարկավոր է կարողանալ կոփել: Նույնը կարելի է փորձել նաև V և VI դասարանների աշակերտների հետ, երբ ցած հնչող տոնայննությունները (G-dur g-moll, As-dur, A-dur, a-moll, B-dur) միաժամանակ երգվեն (կամ երգվեն և նվագվեն) սկսած I, IV, VII աստիճաններից: Նույնանման ձևով կարելի է երգել սեկվենցիոն բնույթի ինտոնացիոն վարժանքներ՝ օգտագործելով կլաստերային «վերադրումները»: [3]

Միգուցե ես չափից ավելի եմ «շահագործում» ժամանակակից լսողության համար այդչափ անհրաժեշտ կլաստերային ինտոնացիոն փորձի դաստիարակությունը, սակայն, համոզված եմ, որ դրանից ավելի էֆեկտիվ մեթոդական աշխատանքի սկզբունք և միջոցների ամբողջականացված համակարգ պարզապես չի կարող գոյություն ունենալ: Մինչև VII դասարանը աշակերտները պետք է լավ իմանան, թե ինչ բան է սոնորիկան և գծայնությունը: VI և հատկապես VII դասարանի աշակերտները պատք է պատկերացում ունենան ընդլայնված տոնայնության մասին: Դա նշանակում է, որ աստիճանների քանակը չի կարող լինել միայն 7-ը:

Այս բոլորի կապակցությամբ ցանկանում եմ հատկապես շեշտել այն հանգամանքը, որ դասական լադազգացողության և ժամանակակից լադազգացողության զուգահեռ դաստիարակությունը պետք է իրագործել չափազանց ճկուն, փոխլրացնելով մեկը մյուսով: Առավելապես անհրաժեշտ է, որ ՄԵԴ-ը ավարտած աշակերտը ստանա լադատոնայնության զարգացման «ուրվագիծ-համայնապատկերը», այսինքն, ամենաընդհանրական ձևով հասկանա, թե ինչպես է զարգացել լադատոնայնությունը:

Հաջորդ հարցը, որ ես կուզենայի լուսաբանել, կապված է թերթից երգելու (ընթերցելու) մշակույթի ճիշտ կազմավորման և զարգացման հետ: Այդ հարցում, ըստ իս, պետք է հետևել ընթերցվող համարների գեղարվեստական արտահայտչականության և բովանդակությունից եկող բնույթի և տրամադրությունների արտահայտման կարողությունների աստիճանական զարգացմանը: Այդ իմաստով սովորած համարը պետք է փորձել կատարել արտահայտելով դինամիկայի, ֆրազավորման յուրահատուկ առանձնահատկությունները: Այստեղ խնդիրն այն է, որ երաժշտական նյութի գեղարվեստական ընկալումը պետք է զարգացնել և խորացնել ոչ միայն մասնագիտական դասերի ժամանակ, այլ նաև «Սոլֆեջո» առարկայի դասերի ընթացքում: Մյուս կողմից անհրաժեշտ է զարգացնել աշակերտների մոտ նյութի մտապահման կարողությունները, տալ նրանց որոշակի ռացիոնալ ցուցումներ, թե ինչպես նրանք կարող են մտապահել սովորվող համարը:

Ինչպե՞ս է հարկավոր առանձնացնել հիմնականը երկրորդականից, ինչի՞ վրա է պետք սկզբնապես բևեռել ուշադրությունը, տրամաբանորեն վերլուծելով ուսումնասիրվող համարը: Դասատուի միջոցով աշակերտը սովորում է անգիր անելու ժամանակ հաշվի առնել կառուցվածքը, կրկնվող հատվածները, սեկվենցիաները, անկայունության վերաճը, բնորոշ ինտերվալների լուծումները և այլն: Մետրառիթմական պարբերականության և կառուցվածքային քառակուսայնության մասին ստացված գիտելիքները նույնպես շատ են նպաստում լավ և արագ մտապահելու խնդիրը հեշտացնելու համար: Այդ հարցի հետ սերտորեն առնչվում է երաժշտական մտածողության զարգացմանը մեծապես նպաստող ունկնդրվող նյութի տրամաբանական վերլուծությունը: Խաղերգային, թեստային-հանելուկային բնույթի ամենաբազմազան առաջադրանքների կիրառումը շատ կարևոր է: Սակայն այսօր ավելի առաջնահերթ խնդիր է երաժշտական նյութի ամբողջական զարգացման օրինաչափությունների և սկզբունքների իմաստավորումը: Անհրաժեշտ է մշտապես ուղղորդել աշակերտի մտայնական ընկալման մեջ ամրագրելու սկզբնավորում-զարգացում-ամփոփում (ընդհանրացում) եռափուլ ամբողջական զարգացման գաղափարը: Դարձյալ խաղերգային-փոխաբերական իմաստով դա հիմք-սյուն-(խոյակ)-տանիք ամբողջական և ամբողջացնող տրամաբանության իմաստավորումն է: Դա ավելի հիմնարար ձևով աշակերտի գիտակցության մեջ ամրացնելու համար ուսուցիչը պետք է կանոնավոր ձևով բոլոր սովորած և թերթից ընթերցվող համարներում առանձնացնի էքսպոզիցիան կամ թեման անվանելով այն initio: Աշակերտը նույնպես կընկալի ուսուցչից զանազանել այն, ինչ սկզբնապես է շարադրվում նրանից, ինչը հետագայում դրա հիմքի վրա է զարգանում (կամ շարունակվում):[1]

Հաջորդ հարցը, որ ես կուզենայի առաջ քաշել, դա ազգային լադահամակարգի մանրամասն և խորը ուսումնասիրությունն է մանկական երաժշտական դպրոցի ՄԵԴ/ «Սոլֆեջո» առարկայի սահմաններում: Այդ կապակցությամբ պետք է ասել, որ ազգային երաժշտության լադահամակարգը /այսպես կոչված ութձայնը/ պետք է հիմք դառնա, ուսումնասիրվի և պարզաբանվի մանկական երաժշտական դպրոցի «Սոլֆեջո» առարկայի նոր ծրագրում: Դեռևս պատմական միջնադարի ժամանակաշրջանում ազգային երաժշտական միտքը մշակել և ստուգաբանել է ազգային երաժշտության լադերի կառույցները և նրանց անվանումները: Ինչպես հայտնի է, այդ լադերը կոչվում են ձայն, կողմ և դարձվածք /օրինակ Ա-ձայն, Ա-կողմ, Ա-ձայնի դարձված/: Մենք պարտավոր ենք անվանել մեր ազգային լադերն այնպես, ինչպես դրանք անվանել է ժամանակին Մեծն Կոմիտասը: Արդյոք «Սոլֆեջո» առարկայի դասավանդման փորձը կարող է հետագայում ազատագրվել բոլոր օտարամուտ շղարշներից և մեծապես կնպաստի երաժշտության մեջ մեր ազգային նեցուկի պահպանմանը: Գուցե ուրույն ազգայինի կորսման մտավախությունը մեզ կստիպի առանցքային նշանակություն տալ «Սոլֆեջո» առարկայի դասավանդման սկզբունքներին, փորձելով շաղկապել հին ավանդականը այսօր արմատավորված նորի և նորագույնի հետ:

Նախապես ներողություն եմ խնդրում այն բանի համար, որ չեմ կարող մեջբերումներ կատարել որևէ հեղինակի գիտամեթոդական աշխատությունից: Պատճառն այն է, որ ես չեմ հանդիպել ինձ հետաքրքրող հարցերի պատասխանները /օրինակ, լռության կլաստերային զարգացման մեթոդիկայի վերաբերյալ չեմ կարդացել որևէ հեղինակի մոտ/:

Մետրառիթմի զգացողության զարգացումը իր մեթոդական խնդիրների իրագործման իմաստով պահանջում է այսօր վերանայել, վերակառուցել պլանային աշխատանքի բնույթն ու սկզբունքները: Չափազանց գործածական և տարածված էր երաժշտական ստեղծագործությունների ժամանակի ընթացքը համեմատել փոխաբերական իմաստով զարկերակի պարբերական բաբախումների հետ: Սակայն այդ համեմատությունը ավելի շուտ հարաբերական է, բայց ոչ բաձարցակ: Ավելի ճիշտ է երաժշտությունը համեմատել տարբեր և որոշակի բարձրություն ունեցող ձայնաշարվածքի ազատ պլաստիկ պարի հետ: Մյուս իմաստով երաժշտությունը կարելի է համեմատել բարձունքում ազատ ճախրող թռչունի հետ: Գուցե դա է պատճառը, որ այսօր չափազանց կարևոր է դպրոցի աշակերտների մոտ սերմանել երաժշտական ժամանակի ռելյատիվ ընկալումը: Այդ կարելի է արմատավորել այն բանից հետո, երբ թույլ և ուժեղ մասերի հավասարաչափ զարկերակը գիտակցվի իբրև շարժման ընդհանրական սխեմա:

Ուստի պետք է ստույգ ձևով բացատրել, որ տակտերի ուժեղ մասերը չունեն նույնանման ծանրակշռությային շեշտվածություն: Ի՞նչ է դրանից եզրահանգվում: Նախ աշակերտը սովորում է և ընտելանում կոնկրետ ընկալել այն, որ ուժեղ մասերի տարբեր աստիճանի շեշտվածությունը ենթարկվում է ոչ թե զարկերակային մեխանիկայի օրենքներին, այլ երաժշտական շարժման որոշակի ենթիմաստներից բխող էներգետիկայի օրենքներին:

Իմ առաջարկը կայանում է նրանում, որ մետրառիթմիկ կառույցների տարբերակումը շատ կարևոր է տարբեր մետրերի ճկուն անցումները նրբորեն զգալու և վերարտադրելու համար: Օրինակ, ազատ, իմպրովիզացիոն առաջադրանքներից մեկը կարող է լինել այն, որ 2/4 մետրով ընթացող ինտոնացիոն կառույցը հարկավոր է տեղադրել 3/4 մետրի մեջ, կիրառելով 1/4 տևողություն ունեցող պաուզաները: Որպես տնային առաջադրանք այդ հանձնարարականը պետք է կատարվի մի քանի տարբերակներով: Ստեղծագործական իմպրովիզացիայի մեկ այլ օրինակ կարող է լինել այն հանձնարարությունը, երբ որևէ վարժության նոտագրված օրինակը տրվում է աշակերտին այնպես, որ բոլոր տակտերում չի գրառված ռիթմական շարժումը: Երաժշտական շարժման բնույթի ճիշտ կռահումը և լավ զգացողությունը պետք է երևա մի քանի իմաստներով: Նախ սեկվենցիոն բնույթի շարժում ունեցող տակտերում աշակերտը պետք է կռահի և պահպանի ելակետային ռիթմապատկերը: Կադանսային եզրափակման նշանակություն ունեցող տակտերում /երբ գրված են քիչ քանակությամբ հնչյուններ/ աշակերտը պետք է տեղադրի խոշոր տևողությունները: Չմոռանանք նշել, որ ռիթմավորում չունեցող տակտերում գրառված է միայն հնչյունների բարձրությունը: Իմ կարծիքով 4-րդ և 5-րդ դասարանների աշակերտներին արդեն կարելի է հանձնարարել այդպիսի առաջադրանքների կատարումը:

Մի քանի խոսք ուզում եմ ասել աշակերտի ունակությունների և կատարած աշխատանքի գնահատման սկզբունքի մասին: Իմ կարծիքով ավելի նպատակահարմար է կիրառել 20 բալյան գնահատման սանդղակը: Բալային չափանիշների կատեգորիաները տրվում են ըստ 4 հինգբալյան սահմանումների: Յուրաքանչյուր հինգբալյան ենթասանդղակը ներառում է վատ, բավարար, լավ, շատ լավ, գերազանց չափանիշները: Իսկ բալային նոմինացիաները չորսն են՝

1. երաժշտականություն
2. երաժշտական հիշողություն
3. երաժշտական մտածողություն
4. անալիտիկ վերլուծություն

Եզրակացություն. Այսպիսով, ավարտելով տվյալ գիտամեթոդական ուսումնասիրությունը շեշտադրում եմ այն փաստը, որ ոճական ճաշակի ձևավորումը, ինտոնացիոն և մետրառիթմիկ լսողության և հիշողության զարգացումը համապարփակ գործընթաց է, որը պետք է համակցված կիրառվի, ոչ միայն «Սոլֆեջո» առարկայի դասերին, այլև երգչախմբի, անհատական պարապմունքների /դաշնամուրի, ջութակի, քանոնի, երգեցողության և այլ մասնագիտական ուղղվածության դասարաններում/ ժամանակ:Այդ դեպքում դպրոցը կպատրաստի գիտակ և գրագետ շրջանավարտներ:

Գրականության ցանկ:

1. Алексеев Э.”Проблемы формирования лада”, Москва – 1976

2. Бочкарев О. “Некоторые виды диатонических ладоконструкций в современной музыке” /сборник “Музыка и современность”, выпуск 7/, Москва – 1971

3. Рети А. “Ладотональность в современной музыке”, Москва – 1986

METHODOLOGY FOR ANALYZING CHORAL WORKS

Papanyan N.
Vanadzor state university named after H. Tumanyan,
Faculty of Social Sciences,
Department of Art and Sports, Lecturer, Armenia

МЕТОДОЛОГИЯ АНАЛИЗА ХОРОВОГО ПРОИЗВЕДЕНИЯ

Папанян Н.А.
Преподаватель, факультет общественных наук, кафедра искусства и спорта Ванадзорский государственный университет имени О. Туманяна, Армения
https://doi.org/10.5281/zenodo.18840582

Abstract

This article discusses methods of analytic of a choral work. The aim of this work is to prepare students for independent chorister activity, which requires a comprehensive knowledge of choral music.

Maximum approximation to a correct interpretation and choice of choir master's techniques to overcome vocal choral difficulties. The student studies choral work in class under the control of a teacher and alone, at home this working in this way should be done systematically and according to a plan. It implies a trough study of the work, its comprehensive analysis and implementation in conducting. If musical analysis is focused on which musical and expressive and for native means are used to create a certain content, than the performing analysis must answer the question of how to deliver it to a listener, how to perform the work. The choral work is built according to a certain scheme analysis of a musical-theoretical means, general executive analysis (temp, agogy, dynamic tones, phrase, means of achievement and distribution of a clumsiest, analysis of a conductor of expression, analysis of technical difficulties (ways of consideration of vocal, intonation, rhythmic, ensemble, difficulties with diction and other. Musical- theoretical analysis includes the breakdown of genre features, musical form, master control, harmonic language, texture (the way the letters are), elements of polyphony of a choral score.

Аннотация

В статье рассматриваются методы анализа хорового произведения. Целью такой работы является подготовка студентов к самостоятельной хормейстерской деятельности, для чего необходим навык всестороннего изучения хоровой партитуры. Максимальное приближение к верной исполнительской трактовке и выбор хормейстерских приемов по преодолению вокально-хоровых трудностей.

Изучением хорового произведения студент занимается в классе под руководством педагога и в домашних условиях самостоятельно. Работа в этом направлении должна проводиться системно и по определенному плану. Она подразумевает тщательное изучение произведения, его всесторонний анализ и воплощение в дирижировании. Если в музыкальном анализе основной упор делается на то, при помощи каких музыкально-выразительных и формообразующих средств создан тот или иной художественный образ, передается то или иное содержание, то исполнительский анализ должен ответить на вопрос, как донести его до слушателя, как следует исполнять сочинение. Анализ хорового произведения строится по определенной схеме: анализ музыкально-теоретических средств, исполнительский анализ (темп, агогика, динамические оттенки, фразировка, средства достижения и распределения кульминаций и т.д.), анализ дирижерских средств выразительности, анализ технических трудностей (рассмотрение вокальных, интонационных, ритмических, ансамблевых, дикционных и других трудностей). Музыкально-теоретический анализ включает разбор жанровых особенностей, музыкальной формы, ладотональные особенности, гармонический язык, фактура (склад письма), элементы полифонии хоровой партитуры.

Keywords: conductor, academic folk choirs, style of presentation, tempo, harmony, voice leading, agogy.

Ключевые слова: дирижер, народные , академические хоры, стиль изложения, темп, агогика, голосоведение, гармония

Введение. Музыкально-исполнительское искусство всегда находилось в центре внимания как музыковедов, так и педагогов. В этой сфере особое место занимает дирижирование как одна из самых сложных и в то же время условных областей исполнительского искусства. Данный вид исполнительства не имеет своего «собственного инструмента», с которым можно вести ежедневную работу: носителем его искусства является вокальный или инструментальный коллектив. Чтобы взаимодействовать со своим «инструментом», дирижёр должен выработать собственный «язык жестов», который и становится основой его исполнительской техники.

В разные времена к дирижированию нередко обращались блестящие музыканты других специальностей, однако результат не всегда был удовлетворительным. Известный русский дирижёр Евгений Светланов утверждал, что быть блестящим музыкантом ещё не значит быть блестящим дирижёром. А Сергей Рахманинов считал, что «дирижирование — это индивидуальный дар, талант, который не может быть приобретён».

Слово diriger — французского происхождения и означает «направлять», «руководить». Следовательно, дирижёр как исполнитель руководит группой музыкантов — оркестром, вокальным или инструментальным ансамблем, оперной труппой и т.д. Греческое слово χορός («хор»), означающее «толпа», в армянском языке получило значение «певческий коллектив» (երգչախումբ) — группа певцов-исполнителей, где каждая партия, состоящая как минимум из трёх человек, исполняет одну и ту же мелодию. Руководитель хора называется дирижёром или хормейстером, а руководитель церковного хора — регентом.

По исполнительской манере хоры подразделяются на академические и народные. Исторический и практический опыт, а также новейшие научные исследования подтверждают, что коллективное пение оказывает большое влияние не только на эмоционально-эстетическое состояние ребёнка, но и на его интеллектуальное развитие. Хор — это уникальный «живой» музыкальный инструмент, основой звучания которого является ансамбль вокальных унисонов.

Материалы и методы. Истоки хорового искусства восходят к Древнему Египту, Вавилону, Древней Греции. Именно в Египте возникла «хейрономия» — своеобразная система управления певцами с помощью движений рук, пальцев, мимики и головы. Хоровое искусство носило культовый характер.

В Древней Греции формирование хорового искусства было связано с театром, особенно с трагедией. Основное действие разворачивалось на специальной площадке — скене (прообраз современной сцены), а хор располагался впереди, на орхестре. Хор выступал либо от имени автора, либо как действующее лицо. В античный период Аристоксен говорил, что медицина очищает тело, а музыка — душу. Хоровое пение было обязательным, а слово «необразованный» толковалось как «неспособный петь в хоре».

Хоровое искусство прошло огромный путь развития, достигло высоких достижений и стало учебной дисциплиной в колледжах, консерваториях и музыкально-педагогических вузах. Важным требованием в подготовке специалиста-хормейстера является самостоятельный анализ произведения и составление письменной аннотации.

Такая работа должна иметь следующую структуру:

1. Введение

• Историко-стилевой анализ произведения

• Сведения об авторах (композиторе и поэте)

• Литературный текст

2. Основная часть

• Музыковедческий анализ

• Вокально-хоровой анализ

• Вокально-хоровые трудности и пути их преодоления

• Исполнительский анализ

3. Заключение

• Значение изучаемого произведения с точки зрения современного искусства

• Список литературы

В историко-стилевом анализе представляются конкретные общественно-исторические условия, в которых протекала жизнь и творчество композитора и поэта, формировались их эстетические взгляды и основные творческие принципы.

Стиль в искусстве определяется как исторически сложившаяся общность образной системы, средств художественной выразительности и приёмов, обусловленная единством идейно-исторического содержания. Понятие «художественный стиль» используется в двух значениях:

1. Совокупность отличительных признаков произведений искусства, архитектуры, музыки, поэзии, литературы, направления или эпохи.
2. Метод, приём, манера работы.

В истории музыкальной науки сложилась целостная система понятий и теорий, разделённых на этапы: Античность, Средневековье, Возрождение, барокко, классицизм, романтизм. В XIV–XV веках термин «стиль» применялся исключительно в изобразительном искусстве. В музыке он стал использоваться в эпоху Возрождения, когда сформировались нидерландская (Ж. Депре), римская (Дж. Палестрина), венецианская (А. Габриели) школы с господствующей вокальной полифонией. В XVIII веке Ж.-Ж. Руссо дал более точное определение стиля как отличительного характера композиции и исполнения.

В эпоху романтизма выделяются три этапа:

- Ранний — 10–20-е годы XIX века сочетание классицизма и романтизма, с классической логикой, но наполненный новым содержанием (К. М. Вебер, Ф. Шуберт, Н. Паганини, Ф. Мендельсон).
- Средний — 30–40-е годы XIX века период расцвета, характеризующийся преобладанием программности, яркой эмоциональностью, развитием виртуозности и синтезом искусств (Г. Берлиоз, Р. Шуман, Ф. Лист, Ф. Шопен, Дж. Верди и др.).
- Поздний — вторая половина 19-го – начало 20-го века, характеризуется психологической глубиной, эмоциональной экспрессией, формированием новых стилевых тенденций (Р. Вагнер, А. Брукнер, Г. Малер и др.). [2]

В XX веке возникло множество новых направлений и стилевых течений: импрессионизм, экспрессионизм, неоклассицизм, атональная музыка, крайние проявления модернизма и авангардизма.

- Знание исторически сложившихся исполнительских традиций и стиля композитора позволяет выстроить методы исследования
- Ознакомление с произведениями одного композитора разных жанров (хоровых, фортепианных, симфонических, камерных, вокальных).
- Изучение особенностей его стиля, музыкального языка, нотного текста, фактуры и авторских указаний.

Обязательно указывается полное название хорового произведения, его жанр (народная песня, обработка народной песни, хоровой романс, поэма, часть хорового цикла, сюиты, кантаты, оратории или фрагмент из оперы) и тип исполнения (a cappella или с сопровождением). Важно рассмотреть роль и значение данного произведения в творчестве композитора, его предпочтительные поэтические источники и индивидуальное отношение к литературному тексту — изменён он или сохранён точно.

- Если анализируемый хор является частью крупного произведения, необходимо охарактеризовать весь цикл и его основную идею.
- Если анализируемое произведение представляет собой обработку народной песни, следует охарактеризовать первоисточник и представить изменения, внесённые в процессе обработки.
- Если существуют произведения других композиторов на тот же текст, можно провести сравнительный анализ с точки зрения интерпретации.

В анализе **поэтического текста** раскрываются особенности рифмы и интонации, форма изложения, количество строф, четверостиший, строк и т.д. Каждый композитор по-своему интерпретирует содержание поэтического текста, подчёркивая наиболее близкие ему мысли, что проявляется в особенностях мелодического и хорового изложения. Известны многочисленные случаи, когда прекрасная музыка создавалась на основе не слишком ценного текста, и наоборот. Анализ поэтического текста должен способствовать всестороннему раскрытию композиторского замысла. [2]

В **музыковедческом анализе** представляются жанровые особенности произведения:

- Стиль изложения — гомофонно-гармонический, полифонический, смешанный
- Тональный план, приёмы гармонизации• Голосоведение
- Метр, размер, ритмический рисунок, темп, агогика

Указанные параметры музыкальной формы являются средствами раскрытия художественного образа произведения.

В **тональном плане** обязательно отмечаются:

- Тональная устойчивость или неустойчивость
- Отклонения — к доминантовым тональностям как средство усиления динамического напряжения или к субдоминантовой сфере как проявление более мягкого напряжения
- Переменный лад
- Разновидности мажора и минора
- Лады, характерные для национальных традиций

Гармонический анализ целесообразно начинать с общей характеристики произведения, поскольку колористическая сторона гармонии является важнейшим выразительным средством. Среди особенностей гармонического языка следует подчеркнуть:

- Простоту изложения, прозрачность или насыщенность фактуры
- Классическую, романтическую или современную усложнённую гармонию
- Типы и роль каденций

В разделе **голосоведения** выделяются:

- Мелодическое развитие основного голоса, тип изложения — поступенный, скачкообразный, плавный, прерывистый и др.
- Гармонически-аккордовое
- Мелодико-гармоническое (мелодия и сопровождающие голоса имеют одинаковый ритмический рисунок)
- Гомофонно-гармоническое (мелодия обладает ритмической самостоятельностью, остальные голоса создают гармонический фон)
- Полифоническое изложение — контрастное, подголосочное, имитационное (канон, фуга, фугато и др.)
- Комбинированное, когда в одном произведении объединены различные виды хорового изложения
- При анализе **технических трудностей** рассматриваются следующие важные моменты:
- Определение диапазона и тесситуры каждой партии

Вид вертикального и горизонтального напряжения

- Анализ технических и исполнительских трудностей каждой партии и произведения в целом. [1]

Заключение. В результате анализа различных средств музыкальной выразительности определяются кульминации произведения и его отдельных разделов, выявляются спокойные и напряжённые участки с точки зрения динамики музыкальной формы. На этом этапе рассматриваются все средства, способствующие достижению кульминации — тональные, тембровые, ритмические.

Каждый раздел аналитической работы должен быть подкреплён примерами. Чрезвычайно важен анализ исполнительских приёмов, включающий вокальные, интонационные, ритмические, ансамблевые, дикционные и другие трудности, связанные со спецификой хора как «инструмента». Эту работу следует начинать со всестороннего изучения партитуры, чтобы стало ясно содержание, настроение и характер произведения.

Цель данной методической статьи — помочь студентам правильно выстроить свою аналитическую работу. Для обучающихся по специальности «хоровое дирижирование» обязательна письменная аннотация произведений различных стилей и жанров.

Список литературы:

1. Андрианова Л. Методика преподавания хорового дирижирования-М.1969г.
2. Колесникова Н. Основы теории и методики обучения дирижированию, Учебно-методическое пособие. Владимир 2012
3. Медынь Я. Методика преподавания дирижерско-хоровых дисциплин. М. Музыка 1978
4. Ольхов К. Некоторые вопросы теории дирижирования и методики начального этапа обучения. В кн. Хоровое искусство, вып. 3 -Л Музыка 1977

PHYSICAL SCIENCES

ON THE ROLE OF THE "LEMNISCATIC CORE" IN LAME ELLIPSOIDAL COORDINATES

Milyute E.
IRG "LITAVEM-3", Vilnius, Lithuania
https://doi.org/10.5281/zenodo.18840584

Abstract
The functions introduced by Gabriel Lamé in connection with the distribution of heat in a homogeneous ellipsoid [1-2], which have since been given Lamé's name, have become the subject of a large number of studies. The works of Liouville and Heine [3-4], independent but simultaneous, are the first in history and among the most important. Then come the excellent studies of Charles Hermite, continued by various authors on the integration of the differential equation satisfied by these functions. Closer to them are two separate groups of significant works: F. Klein [5-6] and his students generalize the Lamé functions and consider them in a completely new way.

Keywords: ellipsoidal vortex, lemniscatic kernel, Lame functions of the first and second kinds, conformal mapping, Lame ellipsoidal coordinates, lemniscatic strip, confocal (homofocal) surfaces, potential theory.

1. History of the issue

The functions introduced by Gabriel Lamé in connection with the distribution of heat in a homogeneous ellipsoid [1-2], which have since been named after Lamé, have become the subject of a large number of studies. The works of Liouville and Heine [4, 3], independent but simultaneous, are the first in history and among the most important. Then come the excellent studies of Charles Hermite, continued by various authors on the integration of the differential equation satisfied by these functions. Closer to them are two separate groups of significant works: F. Klein [5-6] and his students generalize Lamé's functions and examine them in a completely new way.

Following the remarkable introduction of a short treatise by Carl Neumann, a follower of Heine in potential theory, "On the Expansion of a Function with Imaginary Argument" (1862) [7], German mathematicians in the late 19th century developed an interest in Lamé functions for solving conformal problems. In his work "Development of Functions of a Complex Variable from Lamé Functions for Given Spherical Functions" (1882) [8], F. Lindemann laid the foundations for a simple solution of "conformal" problems by means of a transition from Lamé functions to spherical functions, without resorting to complex theoretical calculations. His student O.T. Volk (1925) proved the validity of using the Lame function (21) for an imaginary quantity for a conformal mapping [9,10] (in the formula of Lindemann and Volk, $\boldsymbol{c}$ was replaced by $\boldsymbol{a}$; since at that time due attention was not paid to the precise interpretation of quantities and their belonging to the generally accepted quantities in the calculus of an ellipse) and the research was continued by Volk's student O.E. Stanaitis (1937, 1939), who studied the integral properties of the Lame functions [11].

Based on the above, we can study in detail the mechanism of circulating vortex flows on layered (on different layers) surfaces of an ellipsoidal vortex using Lame ellipsoidal coordinates.

Therefore, there is very little information left in modern mathematics about such special Lamé functions; we can further very briefly dwell on some proofs of well-known Lamé functions, which can be found in this article.

2. Spherical Laplace function (3R).

We have a differential equation:

$$\frac{1}{Sin\,\theta}\frac{d^2f}{d\varphi^2}+\frac{1}{Sin\theta}\frac{d}{d\theta}\left(Sin\theta\frac{df}{d\theta}\right)+n(n+1)f=0, \quad (1)$$

where n is a positive integer. Solutions regular on the surface of a sphere are called spherical Laplace functions [8]. For each value of n, as is known, there exist $2n+1$ functions $f_{mn}(m=1,2,3,\ldots,2n+1)$, which are expressed as follows:

$P_{nm}(Cos\theta)\ Cos\ m\varphi,\ P_{nm}(Cos\theta)\ Sin\ m\varphi$, (2)

where $m=0,1,2,\ldots,n$, and $P_{nm}(Cos\theta)$ are Legendre functions of the n-th order. Next, we assume that in the region $-\pi\le\varphi\le\pi,\ 0\le\theta\le\pi$ these functions are orthogonal and generally unrelated.

Introducing elliptic coordinates [11]:

$$\begin{cases} x=Cos\theta=\frac{\mu\nu}{bc} \\ y=Sin\theta\ Cos\varphi=\frac{\sqrt{\mu^2-b^2}\sqrt{b^2-\nu^2}}{b\sqrt{c^2-b^2}} \\ z=Sin\ \theta Sin\varphi=\frac{\sqrt{c^2-\mu^2}\sqrt{c^2-\nu^2}}{c\sqrt{c^2-b^2}} \end{cases} \quad (3)$$

Then equation (1) changes to another form:

$$\frac{d^2f}{du^2}+\frac{d^2f}{dv^2}+n(n+1)(\mu^2-\nu^2)f=0, \quad (4)$$

Where the values of u and υ denote:

$$u=\int_b^\mu\frac{d\mu}{\sqrt{\mu^2-b^2}\sqrt{c^2-\mu^2}},\ \upsilon=\int_b^\mu\frac{d\nu}{\sqrt{b^2-\nu^2}\sqrt{c^2-\nu^2}} \quad (5)$$

Here μ and ν vary within the interval:

$$b\le\mu\le c,\ -b\le\nu\le b,\ 0<b<c$$

3. The "lemniscatic kernel" of the Lamé function of the first kind and its harmonic analysis

When searching for solutions to the differential equation (1), we will use a special Lamé formula that connects two Lame functions $E_1(\mu)$ and $E_2(\nu)$ of two arguments μ and ν :

$$f=E_1(\mu)E_2(\nu) \quad (6)$$

And we find the Lamé functions $E_1(\mu)$ and $E_2(\nu)$ of the quantities $z,\sqrt{z^2-b^2},\sqrt{z^2-c^2}$ satisfy the same differential equation $E_{ns}(z)$ – Lamé functions of the first kind of degree $n\ at\ z=\mu=\nu$

$$(z^2-b^2)(z^2-c^2)\frac{d^2E_{nm}}{dz^2}+z(2z^2-b^2-c^2)\frac{dE_{nm}}{dz}+$$
$$+\left((b^2+c^2)\,\upsilon_s^n-\,n(n+1)z^2\right)E_{nm}=0\ ,\ (7)$$

where υ_s^n is the root of some algebraic equation:

$$\upsilon_s^n\sim\alpha_1 n^2+\beta_1 n+\gamma_1=\alpha_1(n+\ 0.5)^2+$$
$$+(\beta_1-\alpha_1)+\gamma-0.25\ \alpha_1,\ (8)$$

where $0\le\alpha_1\le\ 1, b, c$ are real numbers. From here we have that for odd $n\implies n\to\sigma=n/2$, for even $n\implies n\to\sigma=(n-1)/2$.

Thus, the function (7) has $s=\sigma+1$ functions,

$$K_{ns}(z)=a_0z^n+a_1z^{n-2}+\cdots\ (9)$$

As a solution, where the coefficients $a_1, a_2, \ldots$ are determined by integration using the infinite series method. To determine the coefficients of the class $K_{ns}(z)$ of functions E_{ns}, we first find them from the system:

$$\begin{cases} 2(2n-1)a_1=p(K-n^2)a_0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \\ 2m(2n+1-2m)a_m=p\left[K-\left((n+2-2m)\right)^2\right]a_{m-1}+ \\ +q(n+4-2m)(n+3-2m)a_{m-2}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{\sigma+1}=0=\left[K-\left((n-2\sigma)\right)^2\right]a_0+ \\ q(n+4-2\sigma)(n+1-2\sigma)a_{\sigma-1}\ ,\ (10) \end{cases}$$

Where the notations p, q, K mean:

$$p=b^2+c^2, q=b^2c^2,\ K=\upsilon_s^n\ (11)$$

Eliminating the coefficients $a_1, a_2, \ldots, a_\sigma$ from the system of equations (10), we obtain an algebraic equation of degree $\sigma+1$, which defines K. It has been proven that this algebraic equation has $s+1$ distinct real roots. From this and therefore follows that there exist $\sigma+1$ functions $K(z)$, that are not linearly related. Therefore, each individual case will always have its own formula, which may not be applicable to another case. Lamé functions possess this unique property, so they cannot be precisely defined for use as a general formula for other cases.

Then, upon substitution, we have

$$t=\int_c^z\frac{\sqrt{(b^2+c^2)\alpha_1-z^2}}{(z^2-b^2)(z^2-c^2)}dz\ ,y=\frac{u}{\sqrt[4]{(b^2+c^2)\alpha_1-z^2}}\ (12)$$

then equation (7), which will have the form

$$\frac{d^2u}{dt^2}+\left[\left(n+\frac{1}{2}\right)^2+\frac{z^2+4B(\beta_1+\gamma_1)}{4A}+\frac{z^4-b^2c^2}{2A^2}-\right.$$
$$\left.-\frac{5z^2(z^2-b^2)(z^2-c^2)}{4A^3}\right]u=0,\ (13)$$

where $B=b^2+c^2\ , A=\left((b^2+c^2)\alpha_1-z^2\right)$.

So we have a final solution

$$e^{it}=\zeta\ eps\left(\int_c^z\frac{z-\sqrt{z^2-\alpha_1B}}{\sqrt{(z^2-b^2)(z^2-c^2)}}dz\right)=\zeta\ eps\ \mathfrak{S}\ (14)$$

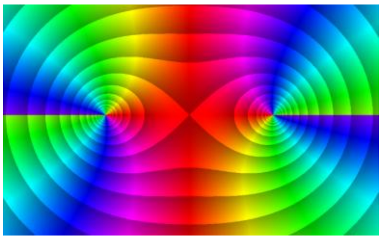
from here we obtain the "*lemniscatic kernel*" function for the Lamé integral:

$$\zeta=\frac{\sqrt{z^2-b^2}-\sqrt{z^2-c^2}}{\sqrt{c^2-b^2}}=eps\left(-\int\ \frac{zdz}{(z^2-b^2)(z^2-c^2)}\right)\ (15)$$

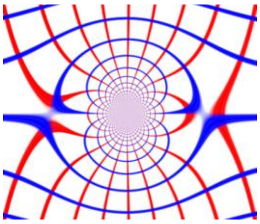
In the conformal mapping, function (15) for $|\zeta|=\rho\ \ (0\ \le\rho\ \le1$ or $1\ \le\rho\le\ \infty)$ yields a system of curves that are very similar in system to the *sofocal curves* of the lemniscate and the sofocal curves of the Cassini ovals [8-10, 18]. Formula (15) gives a system of annular rings connected through the center, as shown in Fig.1.1 and 1.2.

From Fig.1.1a, it is clear that at the conformal mapping of function (15) there are two cuts here, and the semi-hyperbolas pass from one quadrant through the cuts to the other quadrant strictly according to the inversion law. Here we see a *focal point* C_1, which is connected through the cuts to *another focal point* C_2.

We can denote this visual imaginary connection by a special function that will connect both the conformal relation and the spherical one through the center of coordinates.



a



b

Fig.1.1 a) Conformal mapping of a function (26) [1,18]; b) The $\mathfrak{S}$ function is surrounded by a blue circle in the center. This function is described in more detail in the paper [18].

$$f(x,y,z)=E(\mu^2)E(\nu^2),\qquad(16)$$

where

$$\Delta_2F=0\ (17)$$

And since function (15) has two discontinuities, we will assume that in the Riemann variation we are dealing with a two-sheeted Riemann surface, and this surface is doubly connected, as shown in Fig. 1.2 from the work [18].

From this it is clear that the curves $\left|e^{i\zeta}\right| = \rho$ cannot intersect with the curves $|\zeta| = \rho$, since the function $\mathfrak{S}$ has a non-zero real part (Fig. 1.2) and it affects other isogonal lines that fall within the region of the kernel of the function $\mathfrak{S}$.

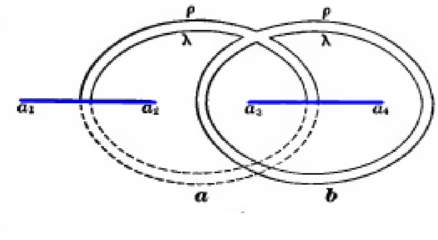


Fig.1.2 Doubly connected Riemann surface of function (15) according to Lindemann [18].

4. Lamé functions of the second kind.

We obtain, as usual, by Euler's method, the second solution of the Lamé equation: this function, known as the Lamé function of the second kind and important in potential theory, is defined by the expression

$$\boldsymbol{E^n(\rho) = (2n + 1)R^n(\rho)} \int_{-\infty}^{\rho} \frac{\boldsymbol{\rho d\rho}}{\boldsymbol{[R^n(\rho)]^2\sqrt{(\rho^2-a^2)(\rho^2-b^2)(\rho^2-c^2)}}} \quad (18)$$

or with a variable *u*

$$\boldsymbol{E^n(z) = (2n+1)R^n(z)\int_0^u \frac{du}{[R^n(z)]^2}} \quad (19)$$

When $z \to \mathbf{0}$ or $\boldsymbol{\rho} \to \infty$, principal value $\boldsymbol{E^n} = \boldsymbol{\rho^{-n-1}}$, and the value $\boldsymbol{R^n} = \boldsymbol{\rho^n}$.

The introduction and study of this second solution are due to Liouville and Heine [3-4]. Various authors have also dealt with it; F. Lindemann [8] developed the quantity $(\boldsymbol{z_1} - \boldsymbol{z})^{-\mathbf{1}}$, where $\boldsymbol{z}$ and $\boldsymbol{z_1}$ are any two real quantities successively occurring in products of the type $\boldsymbol{R^n(z)E^n(z_1)}$; he also studies other expansions of the same kind in series of Lamé functions of the first or second kind: the regions of convergence of these various series are, in general, bounded by curves of the fourth order.

The fundamental role played in the problem of equilibrium of a rotating fluid by equations involving the product $\boldsymbol{R^nE^n}$ has prompted authors studying the problem to seek a form for the function $\boldsymbol{E}$ that is characteristic of a numerical solution. Lyapunov [12-14] presents expressions involving symmetric functions of the roots of the functions $\boldsymbol{R}$, or more precisely, the polynomial $\boldsymbol{f(\rho^2)}$ adjacent to these functions: these formulas are still quite complex.

But if we take into account the *principle of orthogonality* and the *principle of chirality*, then the formula of orthogonal identity (formula (20) from work [18]) is solved using the law of mapping of an ellipsoidal vortex (see Chapter 6 from work [18]) or using the law of symmetry of the lemniscatic strip.

$$\boldsymbol{\frac{X+iY}{r-Z} = \frac{r+Z}{X-iY}} = \mathbb{Z} = \boldsymbol{Const} \quad (20)$$

Thanks to this law of symmetry and the mapping of the lemniscatic strip, the problem is solved more easily

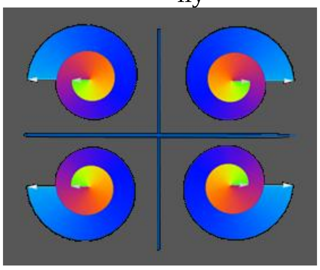


Fig. 1.3 Mirror and inversion symmetry with respect to the OX and OY axes of the simply connected Riemann model (helical spiral)

Riemann's two-sheet geometry (see Fig. 2.3 from work [18] and Fig.1.3). With such an inversion mapping, the flow direction along the lemniscatic strip is preserved; this is even more clearly seen in conformal geometry.

This phenomenon must often be taken into account when describing the distribution of "discharged" matter in the interior of an ellipsoidal closed vortex. Here, in this case, we can easily transition from spherical (elliptical) trigonometry to hyperbolic geometry by equating the spherical radius to the imaginary unit [15-16, 18].

Each function $\boldsymbol{E(z)}$ can also be used as an entire function $\boldsymbol{E(\zeta)}$, where from the system of equations (22-28) we can find the function for the "*lemniscatic core*" for ellipsoidal vortex:

$$\boldsymbol{\zeta = \frac{\sqrt{z^2-b^2}-\sqrt{z^2-c^2}}{\sqrt{c^2-b^2}}} \quad (21)$$

5. The relationship of Lamé homofocal surfaces with elliptical coordinates of field potentials in the 3R system and their characteristics.

Heine himself noted that it was very difficult to solve spatial physical problems of heat and potential in three dimensions without a suitable coordinate system. It was only after Lamé introduced special constructions that simplified the approach to studying three known geometric figures simultaneously: the ellipsoid, the hyperboloid of one sheet, and the hyperboloid of two sheets. Before Lamé, scientists used polar coordinates on a spherical (ellipsoidal) surface, which at the time was computationally cumbersome.

Let us rewrite the system of these figures "according to Lame" in the words of Heine [3]:

Let there be a system of surfaces subject to the conditions

$\boldsymbol{mx^2+ny^2+lz^2} = 1$, (22)

where $\boldsymbol{m}=\frac{1}{\lambda^2}$, $\boldsymbol{n}=\frac{1}{\lambda^2-b^2}$, $\boldsymbol{l}=\frac{1}{\lambda^2-c^2}$

and $\boldsymbol{\lambda > c}$, $\boldsymbol{c > \lambda > b, b > \lambda > 0}$

than

$\frac{x^2}{\rho^2}+\frac{y^2}{\rho^2-b^2}+\frac{z^2}{\rho^2-c^2}=\mathbf{1}$ Ellipsoid

$\frac{x^2}{\mu^2}+\frac{y^2}{\mu^2-b^2}-\frac{z^2}{c^2-\mu^2}=\mathbf{1}$ One-sheet (23) hyperboloid

$\frac{x^2}{\nu^2}-\frac{y^2}{\rho^2-\nu^2}-\frac{z^2}{c^2-\nu^2}=\mathbf{1}$ Two-sheeted hyperboloid

Such surfaces are usually called *confocal* in German works, and Lame called them *homofocal* (Lame, p. 156) [2].

From system (23) we find the values for *x, y, z* for elliptical coordinates:

$$\boldsymbol{x}=\frac{\boldsymbol{\rho\mu\nu}}{\boldsymbol{bc}}$$

$$\boldsymbol{y}=\frac{\sqrt{\rho^2-b^2}\sqrt{\mu^2-b^2}\sqrt{b^2-\nu^2}}{b\sqrt{c^2-b^2}} \quad (24)$$

$$\boldsymbol{z}=\frac{\sqrt{\rho^2-c^2}\sqrt{c^2-\mu^2}\sqrt{c^2-\nu^2}}{c\sqrt{c^2-b^2}}$$

From here we find expressions for the quantities:

$\boldsymbol{\xi = x-\sqrt{x^2-1}}$ (25)

$\boldsymbol{\eta}=\frac{\sqrt{\mu^2-b^2}-i\sqrt{c^2-\mu^2}}{\sqrt{c^2-b^2}}$ (26)

And additional formulas for finding the relationship between two independent coordinate systems:

$$\begin{cases} 2\sqrt{\mu^2-b^2}=\sqrt{c^2-b^2}\,(\eta^{-1}+\eta) \\ 2\sqrt{\mu^2-c^2}=\sqrt{c^2-b^2}\,(\eta^{-1}-\eta) \end{cases} \quad (27)$$

$$\begin{cases} 2\sqrt{\mu^2-b^2}=\sqrt{c^2-b^2}\,\boldsymbol{Cos\ \varphi} \\ 2\sqrt{c^2-\mu^2}=\sqrt{c^2-b^2}\,\boldsymbol{Sin\ \varphi} \end{cases}$$

from the system of equations (27) we find

$\boldsymbol{(27) \Rightarrow \eta = Cos\ \varphi - i\ Sin\ \varphi}$ (28)

The most striking thing is that Lamé anticipated the ease of calculating such complex three-dimensional structures. At the time Lamé published his work on isogonal lines and surfaces, few scientists paid attention to the importance of Lamé's work, but I note that Kelvin took these works with him to England in order to understand the mathematical formulas that he later applied in developing the stability of the rotation of the ellipsoid and in order to be able to apply these conical coordinates. His two-volume treatise on philosophy (on physics) [17], jointly with Tait, had a great influence on the direction of theoretical physics at that time in Europe and in Russia as a whole. Chebyshev himself repeatedly spoke about the problem of the stability of the ellipsoid (a problem raised by Kelvin) [12]. In my case, the ellipsoidal vortex, due to its unique core structure (Fig. 5.2 from [18]), allows its emerging internal vortex domains (flows) to rotate according to the geropoloid (herpoloid) law. These vortex flows in the closed system of the ellipsoidal vortex are characterized by a very high level of self-organization and a clear hierarchy in motion. Here, the vortex motion (not to be confused with the rotational type of solenoidal motion) has many analogs in nature of manifestations with semi-lemniscate motions, which are clearly observed in some cases when equations (29) and (30) are satisfied.

$\boldsymbol{\xi}=\frac{\sqrt{x^2-b^2}}{\sqrt{a^2-b^2}}$, $\boldsymbol{\eta}=\frac{\sqrt{a^2-b^2}}{\sqrt{a^2-b^2}}$ (29)

$$\boldsymbol{\omega = u + iv} = \frac{\sqrt{x^2-b^2}-\sqrt{x^2-c^2}}{\sqrt{c^2-b^2}} \quad (30)$$

For $\boldsymbol{b = 2}$ and $\boldsymbol{c = 1}$, clear signs of a doubly connected three-field (six-field) lemniscate are observed [18]

6. Conclusions

Based on the above, we can conclude that the theory of Lamé functions forms a comprehensive mathematical foundation for studying problems of potential, heat conduction, and spatial conformal transformations in ellipsoidal coordinates. Historical development, beginning with the work of Gabriel Lamé and continued by Joseph Liouville and Eduard Heine, then Neumann, Lindemann, and his successor, Otto Volk, has shown that these functions possess unique analytical properties and require an individual approach in each specific case.

The relationship between the spherical Laplace functions and Lamé's ellipsoidal coordinates allows us to move from the standard spherical description to a system of second-order homofocal surfaces - the ellipsoid and hyperboloids. This representation significantly simplifies the solution of spatial problems and reveals the internal structure of solutions through the separation of variables and the construction of functions - integrals of the first and second kind. The introduction of a "*lemniscatic kernel*" and analysis of the corresponding conformal mappings demonstrate that solutions to the Lamé equation are naturally interpreted in the geometry of *lemniscatic strips* [18], significantly simplifying the study of the mathematical model of a two-sheeted Riemann surface. This emphasizes the importance of the principles of orthogonality, symmetry, and analytic continuation in the study of these functions.

Geometric interpretation through *homofocal surfaces* and *ellipsoidal coordinates* makes it possible to apply the obtained results to the modeling of complex domain-like closed vortex structures, where ordered and hierarchical dynamics of internal flows are observed.

Thus, Lamé functions act as a link between analytical methods, geometric constructions, and physical vortex models, providing a universal framework for describing complex spatial closed vortex processes.

References:

1. E. Miliūtė. (2023) Development of functions of a complex variable on the Lamé functions in conformal geometry. *64-oji Lietuvos matematikų draugijos konferencija*, Vilnius, Lietuva

2. G.Lamé. (1837) *J. Liouville* **2**, 247; (1839) **4**, 126; (1869) Leçons sur les Coordonnées Curvilignes. Paris, Mallet-Bachelier

3. E. Heine. (1864) Handbuch der Kugelfunktionen, erste Auflage.

4. J. Liouville. Lettres sur diverses questions

d'analyse et de physique mathématique concernant l'ellipsoïde, adressées à M.P.H.Blanchet, *Journ. de Math*. T. **XI**, P. 261; *Liouville Journ*. T. **IV**

5. F. Klein (1922) Gesammelte mathematische Abhand-lungen. Band 2, Berlin

6. F. Klein (1881) Über Lamésche Funktionen. *Math.Ann.* Bd. **19** :237-246

7. C. Newmann. (1862) Ueber die Entwicklung einen Funktion mit imaginären Argument nach den Kugelfunktionen, erster und zweiter Art, Halle

8. F. Lindemann. (1882) Entwieklung der Funktionen einer complexen Variabeln nach Laméschen Funktionen und nach Zugeordenten der Kugelfunktionen, *Mathematische Annalen*, vol. **19**.

9. O.T.Volk. (1925) Ueber die Entwicklung von Funktionen zweier komplexen Veränderlichen nach Lamèschen Funktionen, *Mathematische Zeitschrift*, Band **23**, PP. 224-237

10. O.T.Volk. (1925) Die Abbildung $\zeta = \left(\sqrt{z^2 - b^2} - \sqrt{z^2 - c^2}\right)/\sqrt{c^2 - b^2}$, *MGF Darbai*, T.**2**, PP: 34-50

11. O.E.Stanaitis. (1937) Iš Lamè funkcijų teorijos. *MGB Darbai*, T **XI**, I sąsiuvinis, PP. 2-24

12. А.Л.Цыкало. (1988) Александр Михайлович Ляпунов (1857-1918), М.: Наука

13. А.М.Ляпунов. (1885) Теория потенциала

14. А.М.Ляпунов. (1954) Собрание сочинений. т. 1, Изд-во Акад. наук ССР, Москва

15. N.J. Lobatschefskij. (1902) Pangeometrie. 1856 (übersetzt und herausgegeben von H.Liebmann), Leipzig

16. H. Liebmann. (1923) Nichteuklidische Geometrie, Walter de Gruyter & Co., Berlin

17. W.*Thomson, P.G.Tait. (1867) Treatise on Natural Philosophy. Oxford.* 2nd edition, 1883

18. E. Milyute (2023) Investigation of Conformal Mappings in Solving Problems of Vortex Dynamics.

NARRATIVE OF UNIVERSE ON THE BASE OF CONCEPT OF SCIENTISTS AND IN CONTEST OF SCIENCE

Dr Hossain K.
Vice Chancellor, Bangladesh Maritime University
Dhaka, Bangladesh
https://doi.org/10.5281/zenodo.18840586

Abstract

As we know the universe is 13.8 billion years old and that emerging from a hot, dense singularity in the Big Bang and has been expanding ever since. The majority of physics theories about the universe's formation and development are based on General Relativity (GR). Humanity's viewpoint has gradually changed from one of a deterministic, predictable cosmoses to one of probabilistic outcomes and inherent uncertainty due to the successive development of physics theories. This change is mostly caused by the transition from classical to contemporary physics, which has a significant impact on hypotheses regarding the universe's formation and future predictions. The generally accepted Big Bang hypothesis resulted from the integration of relativity with astronomical evidence, including as Edwin Hubble's discovery of an expanding cosmos and the detection of cosmic microwave background radiation. The cosmic microwave background's small quantum fluctuations from the early cosmos may be studied to find "fingerprints of creation" that eventually developed into the large-scale structures we see today. The Big Bang, Cosmic Inflation, and Quantum Foam are all included in the universe's genesis framework. However, things like the exact form and behavior of dark energy, which is speeding up the universe's expansion, make the future mostly uncertain. The Big Freeze, Big Bounce, Big Crunch, and Big Rip ideas of the universe's end do not fully account for the replies of the future. The future is therefore still uncertain and mostly unpredictable. This voyage through uncertainty reveals physics as an ever-evolving subject that has transitioned from the pursuit of flawless prediction to an embrace of probability and the dynamic, open-ended character of cosmic research. Based on contemporary concept of physics and ideas of scientists, this research investigates the universe's genesis and future. With the chronological development of physics theories, an attempt has been made to educate and analyses the universe's formation and consequences as well in a story.

Keyords: Big bang, black hole, GR, QR, JWST, LQG, CCC, CMB, Theory of Everything

1. Introduction

The scientific narrative of the universe is a dynamic, constantly evolving story rooted in evidence-based, testable hypotheses (cosmology), as opposed to mythical or static narratives. Casement of scientific discovery is unparalleled, whether one is seeing a novel phenomenon or unravelling a long-standing riddle. Physicists have continuously investigated the cosmos to discover because physics offers the basic rules that control all matter and energy, from the smallest subatomic particles to the greatest cosmic structures, it is crucial to comprehend the future of life and the universe. The principles of physics govern how life itself functions. One of humanity's most noble pursuits is the quest to comprehend the cosmos and our environment. The most basic science, physics, is essential to these endeavors. Scientists like Einstein, Max Planck, Niels Bohr, Newton, and Archimedes have led challenges and achieved ground-breaking discoveries throughout history. These groundbreaking discoveries and cutting-edge hypotheses are what give physics its excitement.[1] The exert scientific mysteries, using their creativity and logic to unearth basic and unexpected hypotheses about nature.[2] Improving our comprehension of the cosmos is one of physics' most important accomplishments. Scientists have investigated the beginnings of the universe, the formation of stars and galaxies, and the fundamental characteristics of space and time by using physical principles.[3] Regardless of their level of notoriety, all physicists play a vital role in advancing our knowledge by presenting and answering these fundamental questions about creation, life, the cosmos, and the future. Physics has roots in ancient civilizations and is different from current understanding, even though it was formerly limited to astronomy.

The stars in the universe were well understood by the ancient Greeks, Egyptians, Mesopotamians, and new world cultures like the Maya and Aztecs.[4] Each of them studied the sky to observe and forecast the path of the sun, moon, and stars in order to track the seasons and gain a basic grasp of the universe.[5] Their mathematics was occasionally amazing and accurate, despite the fact that they lacked the tools and understanding of the sky to fully comprehend these concepts. Every major civilization practiced some form of sun worship or devotion, and some of these practices were strikingly similar despite their lack of contact with contemporary technology. [6] In a few of instances, their math's was

[1] https://physics.wvu.edu/undergraduate-students/major-in-physics/why-study-physics, accessed on 13 Jan 2026

[2] https://www.linkedin.com/pulse/importance-physics-man-society-dr-muhammad-attique-khan-shahid, accessed on 13 Jan 2026

[3] https://cehsweb.health.umt.edu/sites/default/files/webform/steer/transcripts/Importance-of-Physics.pdf#:~:text=, accessed on 13 Jan 2026

[4] http://blog.nmai.si.edu/main/2017/08/indian-beliefs-eclipse.html, accessed on 13 Jan 2026

[5] http://galileoandeinstein.physics.virginia.edu/lectures/thales.html, accessed on 13 Jan 2026

[6] https://explorable.com/ancient-physics, accessed on 13 Jan 2026

amazing and accurate, but they lacked the tools and the knowledge necessary to fully understand these concepts. Surprisingly, some of the customs of sun worship and devotion were quite similar across all major civilizations, even in the absence of contemporary technology.[7] Particularly with theories like general relativity and quantum mechanics, modern physics has questioned traditional ideas of reality and determinism, raising the possibility that the universe is considerably more complicated and bizarre than previously thought.[8] The attempt to combine quantum mechanics, which governs particle behavior at the smallest scales, with general relativity, which explains gravity and the universe's large-scale structure, is one of physics' biggest unsolved problems.[9]

One of the most fascinating areas of physics is still the search for a Theory of Everything (ToE). If successful, it would provide a cohesive framework for comprehending all physical phenomenon, from particle interactions to the behavior of galaxies. This research examines the relationship between Einstein's theory of relativity and quantum mechanics to investigate the nature of time, reality, and the future.[10] Relativity implies a deterministic universe by introducing the idea of a four-dimensional space-time continuum and the "block universe," where all events—past, present, and future exist simultaneously. However, the probabilistic behavior of particles and the notion that results only become certain upon measurement in quantum mechanics cast doubt on this determinism. Different interpretations of this quantum indeterminacy exist: the Many-Worlds interpretation restores determinism by proposing branching realities for every potential result, whereas the Copenhagen interpretation proposes a fundamentally random reality. The discussion is further enhanced by other ideas like decoherence, entanglement, and the shortcomings of rival theories like pilot-wave mechanics. Therefore, the history of physics is fascinating and fascinating, and the future of human civilization is yet unknown. Nearly a century ago, it was established that the universe is expanding. However, it wasn't until the 1990s, with the advent of powerful telescopes, particularly orbital ones, and the era of precise cosmology, that scientists realized exactly how this happened.[11]

Nonetheless, the most widely accepted account of human civilization and the modern world is the Big Bang theory. The universe began with the Big Bang, which happened around 13.8 billion years ago. From a very heated and dense condition, it began to expand and grow. Fundamental particles emerged as the cosmos cooled, giving rise to hydrogen and helium. Gravitational forces brought these gases together into stars and galaxies over millions of years, sparking nuclear fusion. This process produced heavier elements, which are briefly described in the cosmic calendar and eventually helped build planets and life. The idea that sentient life originated from something random may be a fundamental flaw in the big bang theory. Similarly, Einstein's well-known formula, $E=mc^2$, may be called into doubt since, provided it is matter or has mass, trash may likewise be used in place of radioactive material to create atom bombs. Some people also wonder how the big bang came to be in the first place. And everything that is or will ever be came from nothing. The study also examines how our perception of the "present" is complicated by awareness and the relativity of simultaneity, posing profound queries about whether the future is fixed or still developing. Spacetime is not smooth at Planck scales (very tiny), but rather a chaotic, fluctuating "foam" where virtual particles appear and disappear, implying that everything has a quantum origin. According to the Big Bang theory, the universe began to expand from a very hot, dense singularity around 13.8 billion years ago, as shown by CMB, element abundances, and expansion. Stretching quantum fluctuations into cosmic seeds for structure and resolving Big Bang paradoxes (such as flatness/horizon difficulties) is a fast, exponential growth phase that occurs shortly after the Planck period (about 10^{-32} to 10^{-31} seconds). Consequently, three dominant theories of physics that are compatible emerge. First, String Theory suggests that basic particles are vibrating strings, which might unite gravity with other forces, need more dimensions, and provide methods for characterizing the quantum gravity regime close to the Big Bang/Singularity.[12]

The anti-de Sitter/conformal field theory connection (AdS/CFT correspondence), which links string theory to another kind of physical theory known as a quantum field theory, was discovered by theorists in late 1997.[13] The fact that the complete theory lacks a proper definition in every situation is one of the difficulties with string theory.[14] Second, the singularity problem is resolved by Loop Quantum Gravity (LQG), which quantizes spacetime into discrete "loops," implying that the Big Bang was a "Big Bounce" from an earlier contracting universe rather than a singularity.[15] By arguing that spacetime is quantized—that is, composed

[7] http://www.desy.de/user/projects/Physics/Relativity/SR/light_mass.html, accessed on 17 Jan 2026

[8] http://hyperphysics.phy-astr.gsu.edu/hbase/Forces/qed.html, accessed on 17 Jan 2026

[9] https://cehsweb.health.umt.edu/sites/default/files/webform/steer/transcripts/Importance-of-Physics.pdf, accessed on 17 Jan 2026

[10] Hossain K A, (2026), Evaluation of Creation and Future of Universe on the Basis of Contemporary Theories of Physics, Journal of Liberal Arts and Humanities (JLAH) Issue: Vol. 7; No. 1; January 2026 (pp. 1-93) ISSN 2690-070X (Print) 2690-0718, Doi: 10.48150/jlah.v7no1. 2026.a1, accessed on 30 Jan 2026

[11] https://scitechdaily.com/astrophysicists-developed-a-new-theory-to-explain-dark-energy/, accessed on 03 Nov 2025

[12] Duff, Michael (1998). "The theory formerly known as strings". Scientific American. 278 (2): 64–9. Bibcode:1998SciAm.278b..64D. doi:10.1038/scientificamerican0298-64, accessed on 30 Jan 2026

[13] Maldacena, Juan (2005). "The Illusion of Gravity" (PDF). Scientific American. 293 (5): 56–63. Bibcode:2005SciAm.293e..56M. doi:10.1038/scientificamerican1105-56. PMID 16318027, accessed on 17 Jan 2026

[14] Smolin, Lee (2006). The Trouble with Physics: The Rise of String Theory, the Fall of a Science, and What Comes Next. New York: Houghton Mifflin Co. ISBN 978-0-618-55105-7, accessed on 17 Jan 2026

[15] Rovelli, Carlo (August 2008). "Loop Quantum Gravity" (PDF). Living Reviews in Relativity. 11 (1)

of discrete units—rather than continuous, LQG theory integrates quantum mechanics with general relativity.[16] Instead of being indefinitely split, it suggests that time and space are composed of fundamental, discrete units, such as pixels on a screen.[17] Conformal Cyclic Cosmology (CCC) comes in third. This hypothesis, proposed by Nobel laureate scientist Roger Penrose, holds that the cosmos continuously cycles through "aeons" in which a conformal change that effectively "squashes" spacetime connects the Big Bang of one aeon. This concept suggests that matter (photons principally) from a prior aeon develops the seeds for the new one, perhaps leaving "Hawking points" (circular patterns) in the Cosmic Microwave Background (CMB) as evidence, avoiding a traditional Big Bang singularity and the necessity for beginning conditions.[18] These ideas fit into the inflationary Big Bang story by attempting to connect General Relativity (GR) on a large scale with Quantum Mechanics (QM) on a tiny scale to explain the early moments of the universe.[19]

The enigmas surrounding the universe's beginnings have long piqued human interest and provoked thought throughout history. Humanity has attempted to solve the many mysteries of the universe that surrounds us, whether via the transmission of stories, space exploration, or the development of contemporary scientific theories. From at least Democritus and classical Greek philosophy to Newton and Bentley's discussions and the 21st century, the universe and its origin has been a philosophical topic that has led to advances in philosophy of science and physics. It has also inspired highly sophisticated technology, clever engineers, and a wide range of devices. Since Einstein offered both his well-known General Relativity theory and empirical and Gedanken experiments (thought experiments and conceptual scenarios, frequently in physics, used to explore theories, principles, or hypotheses by imagining outcomes, especially when real experiments are impossible or impractical, famously used by Einstein to develop relativity by envisioning riding light beams or falling freely, illuminating concepts like gravity, light speed, and quantum mechanics), pointing towards QM properties, the Standard Model and main theories have been deemed stuck or even gone astray due to incompatibilities between GR and QM, particularly with regard to the unification of the universe and its origin, such as galaxy formation and the CCP, Dark Matter and Dark Energy. Based on data-driven discovery and symbolic-regression, a state-of-the-art framework was proposed to address these problems that were made worse by JWST 2020–2025 results, the CCP, and naturally the other concerns, such as the unification between QM and GR. The findings led to three compatible leading hypotheses and a modest multiverse approach link with entropic formation of a unique, unstipulated or time-wise universe within a quantum foam, Big Bang, and inflation-based primaeval framework. String Theory, Loop Quantum Gravity, and Conformal Cyclic Cosmology are three complementary theories that investigate this and provide routes to a Theory of Everything (ToE). In order to explain and evaluate the sequential evolution of concepts and theories of physics in order to forecast and correlate the birth of the universe and its ultimate future, this analytical study has based on current concept of scientific theories and ideas of scientists.

2. Methodology and Literature Review

2.1. Basic Scientific Concept. Many cultures throughout the world created elaborate myths to explain how the cosmos came into being long before scientific investigation took the stage. These symbolic and metaphorical tales gave communities a framework for understanding the intricacies of their environment and their own importance. According to Egyptian theology, the cosmos originated from the primordial waters of Nun, the creator God, and these waters created an abyss with infinite possibilities. Atum, the sun god, was credited with using his creative abilities to give the universe shape and order. Similarly, ancient Greek cosmogeny describes how the cosmos was created by the primordial God Chaos, from which the gods Gaia (Earth), Uranus (Sky), and other primordial divinities emerged. These ancient myths provide cultural tales that reflect social values, beliefs, dreams, and ideas, but they don't provide an intelligible explanation for any of the natural happenings in our universe.

Through contemporary cosmology, the 20th century saw a shift in our comprehension of the universe's origins. Edwin Hubble's revolutionary finding of an expanding universe in the 1920s served as the foundation for the Big Bang Theory, which has fundamentally changed how we think about the evolution of the cosmos. According to the Center for Astrophysics, the Big Bang Theory suggests that the universe started from a dense, highly hot single place around 13.8 billion years ago. The "singularity," as it is called, is the point at which space, time, and matter as we know them now began.[20] Subatomic particles combined to become atoms as space cooled and expanded over time, and these atoms then developed into far-off galaxies, stars, and

5, doi:10.12942/lrr-2008-
5. PMC 5256093. PMID 28179822, accessed on 17 Jan 2026

[16] Muxin, H. (2011). "Cosmological constant in loop quantum gravity vertex amplitude". Physical Review D. 84 (6) 064010. arXiv:1105.2212, doi:10.1103/PhysRevD.84.064010, accessed on 17 Jan 2026

[17] Smolin, Lee (2006). "The Case for Background Independence". In Rickles, D.; French, S.; Saatsi, J. T. (eds.). The Structural Foundations of Quantum Gravity. Clarendon Press. pp. 196ff. arXiv:hep-th/0507235. Bibcode:2005hep.th....7235S. ISBN 978-0-19-926969-3, accessed on 17 Jan 2026

[18] Roger Penrose (2006). "Before the Big Bang: An Outrageous New Perspective and its Implications for Particle Physics" (PDF). Proceedings of the EPAC 2006, Edinburgh, Scotland: 2759–2762. Bibcode:2006epac.conf.2759R, accessed on 04 Nov 2025

[19] DeAbreu, A.; et al. (2015). "Searching for concentric low variance circles in the cosmic microwave background". Journal of Cosmology and Astroparticle Physics. 2015 (12): 031, doi:10.1088/1475-7516/2015/12/031, accessed on 17 Jan 2026

[20] Wollack, Edward J. (10 December 2010). "Cosmology: The Study of the Universe". Universe 101: Big Bang Theory. NASA. Archived from the original on 14 May 2011, accessed on 04 Nov 2025

planets.[21] In the end, it influenced the current structure of the cosmos including our own solar system. The Big Bang Theory has become the dominant model in cosmology due to its capacity to explain a wide range of data, including the distribution of galaxies and cosmic microwave background (CMB) radiation. However, the Big Bang Theory has its own drawbacks, just like any other theory, and has spurred continuous debates and investigations into other hypotheses of cosmic development.[22]

An astronomer by the name of Georges Lemaître had a brilliant notion in 1927. According to him, the cosmos began as a single point a very long time ago. He claimed that the cosmos may continue to grow and extend until it reached its current size. Edwin Hubble, an astronomer, discovered that other galaxies were eclipsing us only two years later. And that's not all. Galaxies that were further away were traveling more quickly than those that were nearer to us. This implied that, as Lemaître believed, the cosmos is continually expanding. Everything had been closed together a long time ago if things were drifting apart. The cosmos started out as a mixture of light and energy and hot, tiny particles. It was very different from what we see now. Everything cooled as it grew larger and occupied more space. The little particles clustered. They created atoms. These atoms then formed a group. Stars and galaxies are the result of atoms coming together over a very long period of time.[23] Larger atoms and clusters of atoms were produced by the earliest stars. More stars were born as a result. Galaxies were clumping together and colliding at the same moment. Asteroids, comets, planets, and black holes originated as a result of the birth and death of new stars. They refer to it as the "Big Bang." The cosmos is 13,800,000,000 years old, or 13.8 billion years old, according to current knowledge. That's a lengthy period.[24]

The Big Bang hypothesis is compelling and shows that, according to popular belief, the cosmos did not originate from a "primordial atom explosion," but rather from an extraordinarily hot and dense starting condition. Even However, the scientific community's widespread acceptance of it does not ensure that it has the last word about the universe's distant beginnings. Science has regarded this hypothesis as a legitimate paradigm for more than a century. However, a number of historical gaps remain, and as the idea spreads across academic and educational settings worldwide, fresh questions are raised. Its widespread acceptance does not prove that it is infallible or that the cosmos began precisely as it was suggested. A scientific hypothesis should also be backed up by the actual data that supports it. New ideas, like the "dead universe" notion covered in this article, may now be supported by the same data that originally supported the Big Bang. Unquestionably, the Big Bang hypothesis is a widely recognized model, yet the cosmological model of the dead universe theory could end up being the most likely. Technological developments and mathematical computations in the field of quantum computing may provide a clearer proof of this new theory's validity than the efforts of astrophysicists seeking accuracy to support the theory alone.

However, Alexander Friedmann and Georges Lemaître had already established the theoretical foundation that would challenge conventional wisdom about the universe before Edwin Hubble solidified his place in the study of the cosmos. A Russian mathematician named Friedmann was the first to use the theory of relativity's equations to forecast an expanding universe in 1922. At first, this hypothesis was viewed with suspicion. Simultaneously, in 1927, the Belgian astronomer and priest Georges Lemaître separately put out a similar concept that contained the idea of a "primeval atom" and the theoretical forerunner of the Big Bang. Edwin Hubble became a transformational figure in the environment that these creative minds had constructed. He devoted his professional life to the study of galaxies' redshifts, a phenomenon he himself brought to light via careful studies. Hubble presented his findings in 1929, supporting and extending Friedmann and Lemaître's hypotheses by demonstrating a direct correlation between a galaxy's apparent brightness and redshift.

The idea of the universe's expansion was elevated from a mathematical abstraction to an experimentally provable fact by this finding, which is now known as Hubble's Law. In addition to supporting the work of his forebears, Hubble's contribution marked the beginning of a new era in cosmology, when the notion of a dynamic and expanding cosmos became a fundamental component of the contemporary understanding of space and time. Hawking hypothesized in 1988 that Hubble's measurements suggested the cosmos was infinitely dense and infinitely tiny at a point known as the Big Bang. All scientific rules and, hence, any ability to forecast the future would be ineffective under such circumstances. Events that occurred before now might not have an impact on what is happening now. Since there would be no observable ramifications, their presence might be disregarded. According to Hawking (1988), it is possible to argue that time started at the Big Bang since there would have been no definition for prior periods.[25]

According to a scientific theory known as the "heat death of the universe" (also known as the "Big Chill" or "Big Freeze"), the universe will eventually reach a point where it has no more thermodynamic free energy and, having reached maximum entropy, will no

[21] Stone, M. (1977). "Semiclassical methods for unstable states". Phys. Lett. B. 67 (2): 186–188. Bibcode:1977PhLB...67.186S. doi:10.1016/0370-2693(77)90099-5, accessed on 03 Nov 2025

[22] Caldwell, Robert R.; Kamionkowski, Marc; Weinberg, Nevin N. (2003). "Phantom Energy and Cosmic Doomsday". Physical Review Letters. 91 (7) 071301, doi:10.1103/PhysRevLett.91.071301, accessed on 03 Nov 2025

[23] https://spaceplace.nasa.gov/dark-matter/en/, accessed on 17 Jan 2026

[24] https://spaceplace.nasa.gov/big-bang/en/#:~:text=, accessed on 17 Jan 2026

[25] Hawking, S. (1988) A Brief History of Time. Bantam Books, accessed on 17 Jan 2026

longer be able to support any more thermodynamic processes.[26],[27] Heat death simply means that temperature variations or other processes may no longer be used to do work; it does not indicate any specific absolute temperature. This is the point at which the cosmos enters thermodynamic equilibrium in the terminology of science. The universe will continue to expand indefinitely if its curvature is hyperbolic or flat, or if dark energy is a positive cosmological constant. After a long time, a heat death is predicted, in which the cosmos cools to near equilibrium at a very low temperature.[28] Discussion of the application of the first two principles of thermodynamics to universal processes is where the concept of the universe's heat death originates. Based on previous studies on the dynamical theory of heat, Lord Kelvin presented the idea in 1851 that "heat is not a substance, but a dynamical form of mechanical effect, we perceive that there must be an equivalency between mechanical work and heat, as between cause and effect."[29] It should be anticipated, nonetheless, that an isolated system broken up into subsystems would not always reach thermodynamic equilibrium and will instead stay in a non-equilibrium steady state. The second rule of thermodynamics is not violated since entropy will be transferred from one subsystem to another but its generation will be zero.[30],[31]

According to the "dead universe" idea, there once was an ancient universe that "died" and produced primordial black holes. This universe may have been more than a trillion times bigger than the cosmos as it is now. According to this view, the construction of our universe was not initiated by the first black holes after the collision. Rather, it suggests that only the finest particles of a dead cosmos make up the visible universe. In contrast to the Big Bang hypothesis, it proposes a novel model for the expansion process and a revolutionary cosmological model that provides a more thorough account of the universe's beginnings. Unlike the theory of universal expansion, which suggests an aggressive expansion that results in the breakdown of matter and the end of the universe, this theory really explains a simple separation of galaxies.

Astrophysicists cannot afford to be unfair about the universe-size views put out in this study, nor can they be trivial about the possibilities this technique offers with reference to the ancient universe's moribund state. This particular treatise retains the right to reject the notion of the cosmos as an information processor and multiverse theories, should the criticisms be sarcastic. According to this idea, our visible universe is influenced by the rules of a dead array of star bodies and other elements. Within the context of the "dead universe" idea, two theories are proposed. First, the word "dead" is reinterpreted to refer to a cosmos whose primary feature from the beginning is the inherent lack of light, going beyond the conventional idea of star extinction. According to this idea, light is a cosmic abnormality that results from supermassive bodies fusing and colliding in the vastness of a primordially black cosmos. Furthermore, according to this idea, the world in which we live was not created by fusions or black holes. According to the first theory, the fundamental elements of this primordial cosmos include things like supermassive black holes, dark energy, and dark matter. Interestingly, under some conditions, light appears—possibly as a result of intricate gravitational interactions—acting as a trigger for the shift to a lighted universe similar to what we see today. According to the second theory, the death that destroyed all galaxies and put out the light of a once-bright universe was caused by an ancestral universe that was far bigger than the universe as we know it now. Important proof of cosmic processes ending in the universe's present observable condition may be found in this previous universe.

The observable universe is thought to have been in a state of death trillions of years prior to the Big Bang. In the framework of known physics, the appearance of light, spurred by high activity in this dead world, would be comparable to the oddity of black holes. As a result, the cosmos and galaxies lived in the absence of light, immersed in the gloom of the dead universe yet kept organized by the rules of physics. According to this viewpoint, the cosmos has always existed, although initially in a condition of darkness, rather than just emerging. This idea describes a lengthy time after the Big Bang, but before the first stars formed, when the universe was engulfed in darkness. It draws comparisons to the "cosmic dark ages" postulated by the Big Bang theory. Galaxies and the cosmos therefore existed in the absence of light, immersed in the darkness of the dead universe, but maintained in a condition structured by the rules of physics. According to this viewpoint, the cosmos has always been in a condition of death, having previously existed in complete darkness, rather than having just emerged. This idea compares the Big Bang theory's "cosmic dark ages," which describe a protracted time following the Big Bang but before the first

[26] https://map.gsfc.nasa.gov/universe/uni_fate.html, accessed on 17 Jan 2026

[27] Dyer, Alan (2007-07-24). Insiders: Space. Simon & Schuster Books for Young Readers. pp. 40–41. ISBN 978-1-4169-3860-6, accessed on 17 Jan 2026

[28] Plait, Philip (2008). Death from the Skies!. Viking Adult (published 16 October 2008). p. 259. ISBN 978-0-670-01997-7, accessed on 03 Nov 2025

[29] homson, Sir William. (1851). "On the Dynamical Theory of Heat, with numerical results deduced from Mr Joule's equivalent of a Thermal Unit, and M. Regnault's Observations on Steam" Excerpts. [§§1–14 & §§99–100], Transactions of the Royal Society of Edinburgh, March 1851, and Philosophical Magazine IV, 1852. This version from Mathematical and Physical Papers, vol. i, art. XLVIII, pp. 174, accessed on 03 Nov 2025

[30] Lemishko, Sergey S.; Lemishko, Alexander S. (2017). "Cu2+/Cu+ Redox Battery Utilizing Low-Potential External Heat for Recharge". The Journal of Physical Chemistry C. 121 (6) (published 30 January 2017): 3234–3240. doi:10.1021/acs.jpcc.6b12317, accessed on 17 Jan 2026

[31] Lemishko, Sergey S.; Lemishko, Alexander S. (2020). "Non-equilibrium steady state in closed system with reversible reactions: Mechanism, kinetics and its possible application for energy conversion". Results in Chemistry. 2 100031 (published 8 February 2020). doi:10.1016/j.rechem.2020.100031, accessed on 17 Jan 2026

stars formed, during which the universe was dominated by darkness. The light as a cosmic anomaly: what about it? A second hypothesis may be taken into consideration, which goes beyond the traditional perspective that upholds the natural paradigm of the cosmos in its present condition.

This second viewpoint offers a different explanation for celestial phenomena, such as star activity, and is based on the idea of the "Dead Universe." It's difficult to avoid thinking about the oddity of the Sun's and other stars' deceiving frenetic activity when we gaze at them. It may seem impossible to reconcile the idea of a dead and lifeless cosmos with the powerful light and radiation emissions from these celestial bodies. But if we accept the Dead Universe idea, we may see this activity as an aberration that led to the creation of the universe as we know it. There are two theories that dominate the Dead Universe idea. According to the first hypothesis, light would be seen as an extraterrestrial presence in the otherwise inert universe, which is in its natural state of death. The second hypothesis presents a more ambitious idea: a universe that is billions of times bigger than our own and has progressively descended into an eternal state of death. The whole nature of existence was changed in this vast galaxy of light and regular stars, resulting in a world where light and stellar occurrences were every-day.[32] We are forced to reevaluate our comprehension of cosmic occurrences by exploring these theories. They transform from simple anomalies into mysterious hints that allude to the deep complexities of the universe's creation and its death.

2.2 Dead Universe Theory. According to the Dead Universe idea, there would be no brilliant stars, solar flares, or any radiant energy in the universe, and it would be completely inactive. According to this paradigm, starlight and the energetic events that accompany it would be seen as anomalous disruptions in a cosmos that would otherwise continue to exist in an eternally peaceful state. We would view stellar events such as coronal mass ejections and solar flares as transient departures from the inert equilibrium that defines the Dead Universe. In this ostensibly unchanging cosmos, these expressions of tremendous energy activity would be seen as oddities that resulted from extraordinary circumstances or catastrophic occurrences. As a result, adopting the Dead Universe idea forces us to reconsider how we interpret starlight and other celestial phenomena. They become indicators of a fundamental rupture that gave origin to the universe as we know it, rather than being seen as natural features of the cosmos. This alternate viewpoint calls into question our preconceived notions and encourages us to investigate fresh approaches to comprehending the nature and genesis of the universe.

➢ Anomaly of Light: A key component of the physics of the world as we know it is light, a basic kind of electromagnetic energy. Invoking complexity is not all that is required to suggest that light is an aberration in this theory; rather, it provides solutions to some of the most important concerns in classical physics. This method offers a fresh perspective on the nature of the cosmos in addition to reinterpreting well-known physical ideas. Thermonuclear events in the cores of stars are the main source of light in this complicated process.

➢ Nuclear Fusion: Nuclear fusion is the main process by which stars produce light. Nuclear fusion is the process by which hydrogen atoms combine to generate helium in the stellar core, particularly in stars like the Sun. A little portion of the atoms' mass is transformed into energy during this fusion in accordance with Einstein's well-known equation, E = mc2. Heat and light are the products of this energy's release.

➢ Pressure of Radiation and Gravitational Pressure: Within a star, nuclear fusion generates massive quantities of energy in the form of radiation and high-energy particles. This radiation exerts an outward push in all directions. At the same time, the massive mass of the star creates a powerful gravitational attraction that attempts to compress it into the core. Hydrostatic equilibrium is established when these two forces—gravity inward and radiation pressure outward—balance each other.

➢ Fusion Cycle: The proton-proton cycle, which releases photons (light particles) when four hydrogen nuclei unite to produce a helium nucleus, is the main fusion mechanism in the sun and other stars of comparable size.

➢ Gravitational Pressure: Only in stars can nuclear fusion take place because of the tremendous gravitational pressure in their cores, which pushes the hydrogen nuclei closer together until they can overcome their electrical repulsion and fuse.

➢ Hydrostatic equilibrium: The force of gravity that is attempting to squeeze the star is balanced by the outward pressure created by the light produced by nuclear fusion. The star remains stable and in its present state because of this hydrostatic equilibrium.

➢ Pressure of radiation and gravitational pressure: Nuclear fusion produces enormous amounts of energy in the form of high-energy particles and radiation within a star. There is an outward pressure from this radiation in every direction. Concurrently, the star's enormous mass creates a strong gravitational pull that tries to compress it in the direction of the core. When these two forces—gravity inward and radiation pressure outward—balance one another, hydrostatic equilibrium is reached.

➢ Stellar stability: The star stabilizes when hydrostatic equilibrium is reached. The stellar structure will alter in response to any perturbation that results in an imbalance between radiation pressure and gravity. For instance, gravity will start to crush the star if radiation pressure drops, raising the core's warmth and pressure. In order to reestablish balance, this may cause the rate of nuclear fusion to accelerate. However, if radiation pressure is too high, it might overcome gravity and cause the star to expand, which could lead to an explosion or the ejection of stellar material.

➢ Stellar lifecycle: Understanding hydrostatic equilibrium is essential to comprehending the star's life

[32] Hossain K A, (2025), Emotional and Artificial Intelligence: Narrative of Human and Machine Relation, Transactions on Engineering and Computing Sciences - Vol. 13, No. 05 Publication Date: September 07, 2025 DOI:10.14738/tmlai.1305.19324, accessed on 07 Feb 2026

cycle. Stars maintain this balance for the most of their lifespan, staying steady and producing energy via nuclear fusion. But once nuclear fuel is used up, radiation pressure drops and gravity takes over. This may lead to a variety of outcomes, including transformation into a red giant, explosion, or even a black hole, depending on the star's mass.

- Dynamic equilibrium: It is crucial to remember that hydrostatic equilibrium is a dynamic balance rather than a static condition. Energy generation, stellar material migration, and other physical interactions cause conditions within a star to change continuously. To guarantee that these changes take place in a balanced and regulated way and maintain the star's relative stability across its lifetime, hydrostatic equilibrium is necessary. the idea that light did not exist at the beginning of the cosmos and was generated later. It is indisputable that darkness came before light, regardless of whether one supports the creationist theory—which holds that the cosmos was covered in darkness before God declared, "Let there be light"—or the scientific understanding of these early occurrences.
- Primitive Elements: Despite being well-established ideas in contemporary cosmology, black holes, dark matter, and dark energy are often seen as emerging phenomena rather than necessarily as the universe's original constituents. The dead universe idea, however, offers a tenable account for their genesis, identifying them as essential components of an earlier inactive universe. This hypothesis offers one of the first logical attempts to explain these mysterious phenomena, despite the fact that dark matter and dark energy are topics of great study and controversy, with their origins still up for question.
- Expansion of Cosmic Understanding: These concepts test our cosmic imagination and provide a rich environment for scholarly debates and imaginative fiction. These theoretical ideas aim to deepen our understanding of the universe's potential states and the underlying processes that shape its development and eventual finality, even if they are still far from the present consensus in science. Therefore, these claims permit speculative investigation based on alternative theories and hypotheses while acknowledging the limits of accepted scientific knowledge.

According to the Dead Universe hypothesis, the idea of star birth is flipped to universal death, and the universe as we know it is the remnant of a vastness that has long since vanished. According to this theory, black holes are the last resting place of a cosmos that has run out of energy rather than the sources of creation. These primordial black holes are the last gravitational lighthouses of a vanished universe, not generating singularities. Despite their seeming freshness, the galaxies and stars we see are really the remnants of a long-extinguished cosmic fire. The mysterious components of our cosmos, dark matter and dark energy, might be seen as a faint echo of this last catastrophic catastrophe. The ideas of the "Big Freeze," "Big Rip," "Big Crunch," and "Big Slurp" provide dramatic scenarios based on the ongoing expansion, contraction, or phase transitions of space-time, among other theories explaining the ultimate destiny of the universe. On the other hand, the "Dead Universe" idea offers a more tranquil and essentially different conclusion for the universe.

There is no explosive or catastrophic event that signifies the end of the universe, according to the Dead Universe idea. The cosmos just reverts to its normal state, which is devoid of the oddities that define the visible universe, in its place. According to this idea, light is an aberration that needs nuclear fusions and other energy processes to emerge rather than emerging on its own. According to this interpretation, the universe's primordial and perpetually black condition is momentarily disturbed by the presence of light. According to this viewpoint, if there were living things in this early cosmos, they would see light as odd, fascinating, or an invader in the eternal state of darkness. According to the hypothesis, the cosmos ends with a quiet and unavoidable return to darkness rather than a dramatic explosion, as is the case with star death. According to science, life as we know it may have originated from this aberration, from a sequence of mishaps and strong mergers in the early universe that would eventually pull the universe back to its original state. According to the hypothesis, the universe would eventually reabsorb everything back into its original form, devoid of light and life, where prevailing gravitational forces will enable this return to the primitive and dark state, just as light and life evolved from remarkable occurrences.

The Big Bang and other cosmological theories' common idea of continuous expansion is challenged by this hypothesis. It foretells a slow descent into a quiet equilibrium, when the universe gradually disappears into the shadowy backdrop of its primordial existence, eternally trapped in a cycle of fleeting brightness and ultimate darkness, rather than a violent or catastrophic end. Galaxies are moving apart at a pace proportionate to their distance from one another, as shown by the expansion of the cosmos, a phenomenon that has been seen and explained by Hubble's Law. This fundamental aspect of contemporary cosmology is consistent with the "Dead Universe" idea; however, the ramifications of this expansion are interpreted quite differently. According to the theory of the Dead Universe, the observed expansion is not the consequence of an initial impulse from an explosion, as in the Big Bang, but rather a straightforward separation of galaxies brought about by gravity and other as-yet-undiscovered laws arising from the Dead Universe's very nature. It is believed that this movement is an expression of the inherent and lingering characteristics of a universe that is no longer functioning in the conventional mind.

In other words, the idea of the Dead Universe aims to explain why we perceive things, while Hubble's Law explains what we see. It implies that the dynamics of the visible world are now governed by the unidentified rules of the Dead world, which might be remnants or echoes of a past cosmic reality. Galaxies may continue to separate long after the initial energy of the Big Bang should have dissipated due to factors that vary from the well-known classical gravity. The expansion would thus represent a slow return to the basic and quiescent condition of the Dead world, a final state of rest after the end of anomalies like light and the intricate structures that define our current world, rather than a sign of ongoing development or birth. As a result, the Dead

Universe idea provides an interesting alternative to accepted cosmological theories and sheds light on the ultimate destiny of the universe. Newton's law of universal gravitation, which states that mass-containing bodies attract one another gravitationally and that this force is directly proportional to the masses of the bodies and inversely proportional to the square of the distance between their centers, appears to be applied in this description. Gravity would be a key factor in determining how the observable world moves, according to the idea of the Dead world we are investigating. Taking this view, we can hypothesize that our visible universe is being pulled back to a more uniform and peaceful state by the gravitational pull of a Dead Universe that contains bodies of immeasurable mass. This would suggest that the gravity of such cosmic superstructures may ultimately oppose or even defeat the forces causing the universe's expansion, such as dark energy.

This idea differs from previous cosmological models and is incompatible with cyclic models of creation and rebirth, the Big Bang, and multiverse theories. In sharp contrast to the multiverse theory, it promotes the notion of a perpetually dying cosmos that produces smaller galaxies and sometimes some bigger ones as it dies. According to the many hypotheses, the universe will end dramatically in the future—possibly 200 billion years from now. The dead universe hypothesis, on the other hand, holds that this catastrophe has already happened and that the universe is just continuing on its path to complete extinction. These theories may be challenging because of how they approach the study of space and time. This new hypothesis is more in line with Penrose's theories (Penrose, 2004).[33] "Let us start by exploring the arena within which all the phenomena of the physical universe appear to take place: spacetime," said Penrose. For the most of the remainder of this book, Penrose's explanation of this idea is crucial. Spacetime was not Einstein's original concept, and he was not especially excited about it when he first heard of it, despite what seems to be the popular belief and his excellent use of it in his formulation of the general theory of relativity. First, we need to know why "spacetime" exists. What is wrong with seeing space and time as distinct concepts instead of trying to combine these apparently disparate ideas into a single one? Furthermore, when we reflect on the great earlier relativistic discoveries of Galileo and Newton, we see that they, too, may have benefited greatly from the spacetime viewpoint.

2.3 Additional Significant Concept. According to the theory of the Dead Universe, there is a vast and singular universe in which the process of death continues to have an impact on the observable universe. These particles can be thought of as the last remnants of the once-vigorous dead universe's energy, which continue to create galaxies as memories of death. But despite the formation of new galaxies, the universe is contracting rather than expanding, as suggested by the Big Bang and other accepted hypotheses. One analogy for our observable universe is a big balloon that is progressively losing its stellar vibrancy. The observable cosmos today only makes up a tiny portion of the once-active universe—roughly 0.001%—and is essentially nothing in comparison to its former splendor. Furthermore, the Dead Universe idea suggests that galaxies gradually shrink in size as a result of this fatal decay process. It is improbable that a galaxy bigger than the biggest ones we are aware of will appear in front of our sight. The Big Bang theory and the data that the James Webb Telescope will evaluate in the future stand in sharp contrast to this. On the edge of time, galaxies will undergo a period of complete and silent death. The Big Rip, Big Crunch, Big Freeze, and Big Bounce hypotheses of the universe's conclusion stand in stark contrast to this viewpoint.

- Big Rip: This theory suggests a catastrophic expansion that tears everything apart, from galaxies to atoms.
- Big Crunch: Proposes a scenario where the universe’s expansion reverses, leading to a collapse into a singularity.
- Big Freeze: Predicts a universe that expands forever, becoming increasingly cold and dark as stars burn out.
- Big Bounce: Suggests that the universe goes through infinite cycles of expansion and contraction.

However, the Dead Universe concept does not foreshadow a violent ending. Instead, it implies that the gradual and silent drifting away of galaxies is caused by the gravitational remnants of the dead universe. This delayed separation lacks the violent dynamics suggested by other hypotheses. The concepts of cosmic inflation and conventional expansion are replaced with the idea of a silent, steady detachment of galaxies, influenced by the decaying forces of a much larger, now mostly dead universe. Because it does not have the catastrophic elements that other models indicate, this smooth separation highlights a more tranquil finish. A gradual winding down of cosmic activity is also implied by the fading universe effect, leading to galaxies that becoming smaller and less active until complete cosmic silence is reached. The gravitational and creative remnants of the dead universe are thought to influence the formation of new galaxies inside our observable universe.

The unique aspect of the Dead cosmos hypothesis is that it rejects the conventional expansionist and inflationary theories in favor of a concept in which the fading leftovers of a much bigger cosmos control cosmic events and galaxy formation. As opposed to the violent or cyclic finishes suggested by other cosmological theories, this results in a slow, silent spread of matter and energy. Thus, this hypothesis offers a distinctive viewpoint on cosmic development, highlighting a gradual, tranquil fall and the birth of galaxies as a reminder of a once-living, now mostly destroyed world. This hypothesis proposes a cosmos that gradually disappears rather than terminating in a cataclysmic catastrophe, providing a fresh framework that contrasts with the more dramatic cosmological theories. To further this

[33] Penrose, R. (2004) The Road to Reality: A Complete Guide to the Laws of the Universe. AbeBooks, accessed on 07 Nov 2025

theory, it may additionally include the idea of a membrane (brane). As a kind of cosmic memory, the dead cosmos produces new galaxies as it decays. These galaxies are not random; rather, they are impacted by gravitational forces and the residual architecture of the collapsing universe. This mechanism promotes the ongoing production of galaxies inside a single universe rather than the creation of numerous universes.

Our world is neither growing, inflating, or going through cycles of rebirth, according to the Dead world idea. Rather, it has been in a state of deterioration for perhaps trillions of years and is progressively declining. According to this scenario, galaxies are constantly developing and then vanishing as a result of the dead universe's leftovers, indicating that the cosmos is gradually approaching complete death. Our universe is gradually winding down, heading toward a quiet and total end of cosmic activity, rather than expanding or renewing. It's possible that the whole visible universe is contained inside a unique brane floating in a higher-dimensional region. A portion of the dead cosmos that is approaching death is symbolized by this brane. Nonetheless, new galaxies may emerge inside this brane as a result of the creative leftovers of the dead universe. These galaxies arise when the leftovers of the dead cosmos continue to impact the formation of smaller and smaller galaxies. These new galaxies are formed in part by cosmic events like as explosions that occur during this fading phase (Randall, 2005).[34] "The universe may be governed by hidden dimensions, beyond the familiar three of space and one of time, opening up new avenues for exploration and understanding," said Randall.

Therefore, it is possible to hypothesize that our dead world is contained inside a brane that floats in a higher-level realm. Our world enters a condition of death inside this region, existing as a membrane separate from other possible realities. The physical principles and remains of the dead universe may have an impact on this brane, causing galaxies to continue to form. The interactions and residual energy of the dead universe's structures produce these galaxies, which are a component of the cosmic memory of the dead universe. According to this idea, the physics of the dead universe impacting the visible world is what causes the gravitational anomalies and the curvature of time and space that are seen in our universe. The structure and behavior of our visible universe are significantly shaped by the fading relics of the Dead Universe, especially supermassive bodies and dark energy. It would be fair to cite a number of established physical concepts and findings in order to support Dead Universe hypothesis:

- The foundation for comprehending gravity on a broad scale is Newton's Law of Universal Gravitation.
- The concept of gravity as the curvature of space-time was revised by Einstein's General Relativity. The "Dead Universe" might be used to examine the function of singularities and event horizons in black holes.
- The Hubble Space Telescope's observations and Hubble's Laws provide factual proof of the universe's expansion, which may be explained by the allure of a vast "Dead Universe."
- Quantum Cosmology: By examining the effects of quantum mechanics on cosmic scales, this field may provide light on how matter and energy in our universe may be affected by a "Dead Universe."
- Investigations into Dark Energy and Dark Matter: Research on these enigmatic elements of the cosmos may help us comprehend the forces that interact or compete with gravity in the "Dead Universe."
- The following notable scientists might be cited: Stephen Hawking: For his research on singularities and black hole characteristics.
- Alongside Hawking, Roger Penrose explored hypotheses on the nature of space-time.
- Three astronomers, Saul Perlmutter, Brian P. Schmidt, and Adam G. Riess, received the Nobel Prize for their discovery on the acceleration of the universe's expansion.
- Kip Thorne was a theoretical physicist who made important advances to our knowledge of gravity and gravitational waves.

2.4 Difficulties of Exicting Theories. When scientists look at the constancy and homogeneity of cosmic microwave background radiation, the horizon issue becomes clear. The thermal radiation from the creation of our universe, about 380,000 years ago, is known as CMB radiation. Soon after the Big Bang, when visible light was first able to travel freely and unhindered, this radiation took place.[35] The furthest regions of space were previously in thermal equilibrium, according to the apparent uniformity of CMB, which represents temperature fluctuations at a scale of one part in 100,000. Put another way, the temperature of the universe's farthest regions used to be the same, indicating that heat was dispersed equally in all directions. These areas are separated by great distances, nevertheless. Given the antiquity of the universe and the speed of light, which is around 186,000 miles per second, it should be physically impossible for these places to have ever been near enough to directly interact and equilibrate since the Big Bang. In other words, the horizon issue poses the intriguing question of how the far-flung regions of the cosmos may possibly have such identical temperatures and properties.[36]

The form and general curvature of the cosmos, however, are the subject of the flatness issue. Omega (Ω), a measure of curvature, describes how mass and

[34] Randall, L. (2005) Warped Passages: Unraveling the Mysteries of the Universe's Hidden Dimensions. HarperCollins Publishers, accessed on 17 Jan 2026

[35] Frampton, P. H. (1977). "Consequences of Vacuum Instability in Quantum Field Theory". Physical Review D. 15 (10): 2922–28. Bibcode:1977PhRvD.15.2922F. doi:10.1103/PhysRevD.15.2922, accessed on 07 Nov 2025

[36] Wang, Yun; Kratochvil, Jan Michael; Linde, Andrei; Shmakova, Marina (2004). "Current observational constraints on cosmic doomsday". Journal of Cosmology and Astro-Particle Physics. 2004 (12): 006. arXiv:astro-ph/0409264. Bibcode:2004JCAP...12.006W. doi:10.1088/1475-7516/2004/12/006, accessed on 17 Jan 2026

energy affect the form of the cosmos in accordance with Einstein's Theory of Relativity. A universe with "Ω = 1" is flat, meaning it has no curvature and satisfies the critical density condition, which states that the expansion rate of the universe should ultimately slow down and get closer to zero without ever achieving it. It indicates that the universe's expansion is always slowing down. According to the original Big Bang Theory, the universe should have been extremely near to critical density (Ω = 1/flat in form) just after the Big Bang. However, if the universe expanded, even a little departure from the critical density would become more pronounced over time, creating a highly curved cosmos that is either "open" ($\Omega < 1$) or "closed" ($\Omega > 1$). However, the cosmos as we now see it via scientific tools is flat. The issue then becomes: How is it even possible?

Modern cosmologists have proposed a number of hypotheses to better explain the characteristics and phenomena of the cosmos in order to address these types of issues.[37] The cosmic inflation idea, initially put out by physicist Alan Guth in the 1980s, is among the most sobering and scientifically supported hypotheses. A fraction of a second after the Big Bang, an exponential expansion occurred, in accordance with Guth's cosmic inflation hypothesis. The observable structure and composition of the cosmos that we see today were made possible by this inflationary phase. There is visible scientific data that supports Guth's idea. Additionally, it addresses a number of long-standing cosmological puzzles, like as the flatness and horizon problems. According to cosmic inflation theory, the universe expanded exponentially in the first fraction of a second after the Big Bang, which addresses the horizon issue. The cosmos was expanded beyond its observable horizon during this inflationary phase, which allowed far-off areas to reach thermal equilibrium and make causal touch. According to this hypothesis, the expansion made it possible for the far-flung regions of the cosmos to interact and affect one another, bringing them all to the same temperature. Stated differently, the current universe might have expanded faster than the speed of light during this early inflationary phase based on the physics explained by the cosmic inflation hypothesis.[38]

The issues of time and distance impeding thermal equilibrium would have been resolved. The scale factor of the universe, which establishes the relative sizes of spatial dimensions (the size of space itself), increased during the time of fast and massive expansion, according to cosmic inflation theory, which addresses the flatness issue. Therefore, during this inflationary phase, any little departures from a flat geometry in the early cosmos would have been significantly stretched out and weakened. Put differently, the universe would have been more evenly flat as a result of the fast expansion, which would have smoothed out these discrepancies. The energy density associated with the inflation field overtook other types of energy, like as radiation and matter, during the expansion of the universe. The geometry of the whole cosmos would have been leveled by this supremacy, approaching a flat shape. Therefore, inflationary cosmology since the 1980s offers convincing answers to these types of issues about the universe's formation. It establishes the groundwork for contemporary cosmological ideas and transforms our understanding of early dynamics. Researchers believe that quantum fluctuations in the space-time fabric, which are predicted by quantum physics, are what caused this inflation.[39] Tiny variations at these quantum levels are thought to have been amplified during inflation, introducing anomalies and differences that ultimately gave rise to the earliest galaxies, galaxy clusters, and macro-level cosmic structures.

The idea that our universe may just be one of many in a vast "multiverse" is being explored by scientists due to developments in cosmology. According to this hypothesis, there may be an endless number of universes, each with unique physical rules, constants, and traits. The multiverse theory offers an intriguing explanation for some of the most perplexing characteristics of the world, even if it is still theoretical and outside the scope of current empirical investigation. For instance, a multiverse scenario where each area has distinct features might justify the exact tuning of constants and parameters in our world to sustain life. In such scenario, life as we know it would not be supported by the design of our own universe, but would instead be the result of chance and happenstance.[40] Numerous other worlds in the multiverse may not be able to sustain this kind of life. After discussing the universe's early beginnings, you may be wondering, "How will it end?" Although scientists have some suggestions, there is no way to tell for sure. The Big Rip hypothesis, the Big Freeze theory, and the idea of rapid expansion all provide clues about the possible futures of the universe.

Researchers hypothesized that gravity will eventually halt the universe's expansion when all of the universe's stuff pulls on itself to rejoin once the Big Bang Theory for the universe's origin was solidly confirmed. They thought the expansion would finally be halted by gravity. After then, there would be a recoil, which would lead everything to gradually come back together—possibly all the way back to a single point. This notion was dubbed the Big Crunch by researchers. It even sparked the idea that the cosmos may undergo a recurring cycle of rebounds as it repeatedly expands and shrinks due to conflicting forces vying for supremacy. However, the pace of expansion of the cosmos is not slowing down, according to scientific observations.

[37] Kirshner, Robert P. (13 April 1999). "Supernovae, an accelerating universe and the cosmological constant". Proceedings of the National Academy of Sciences. 96 (8): 4224–4227. Bibcode:1999PNAS...96.4224K. doi:10.1073/pnas.96.8.4224, accessed on 07 Nov 2025

[38] Caldwell, Robert R.; Kamionkowski, Marc (2009). "The Physics of Cosmic Acceleration". Annu. Rev. Nucl. Part. Sci. 59 (1): 397–429. arXiv:0903.0866. Bibcode:2009ARNPS.59..397C. doi:10.1146/annurev-nucl-010709-151330, accessed on 17 Jan 2026

[39] Tegmark, Max (2014). Our Mathematical Universe: My Quest for the Ultimate Nature of Reality (1 ed.). Knopf. ISBN 978-0-307-59980-3, accessed on 17 Jan 2026

[40] Hawking, Stephen W. & Moss, I. G. (1982). "Supercooled phase transitions in the very early universe". Physics Letters B. 110 (1): 35–38. Bibcode:1982PhLB.110...35H. doi:10.1016/0370-2693(82)90946-7, accessed on 07 Nov 2025

Rather, it is really rising. This surprising discovery, which came from researching supernovae in the late 1990s, implies that a mysterious force known as dark energy is driving the expansion of the universe and resisting gravity on a cosmic scale. This fast expansion is driven by dark energy, which has important ramifications for the future of our universe. It implies that galaxies will keep separating at an accelerating rate.

The Big Rip Theory paints a striking and dramatic image of one possible outcome for our cosmos by following the universe's accelerated expansion to its ultimate end. According to this idea, the repelling force of dark energy intensifies with time and has the capacity to overcome all other forces, including the gravitational pull of stars, galaxies, and subatomic particles.[41] The Big Rip hypothesis predicts that galaxies will move apart as the universe expands at an ever-increasing rate, which is actually occurring now.[42] Dark energy may eventually overwhelm the gravitational forces that hold galaxies, stars, planets, and atoms together. Cosmic structures would be destroyed, matter would disintegrate into its constituent elements, and spacetime itself would be torn apart at the most fundamental level as a consequence of this cataclysmic catastrophe. In other words, everything in the cosmos would be "ripped" to bits by dark energy.

The Big Freeze Theory, often called the Heat Death Theory, predicts a slower and more muted end for the cosmos. The Big Freeze postulates that matter and energy will progressively thin out over eons while the cosmos continues to expand at an accelerating rate owing to dark energy. Existing stars will gradually burn out and new stars will cease to emerge as galaxies separate and the cosmos becomes colder and more arid. The cosmos will eventually attain a maximum entropy state, in which there is no possibility of matter interaction and all energy is evenly distributed.[43] The cosmos would turn into a chilly, black emptiness in this condition, which some physicists refer to as Heat Death. Life, light, and any distinguishable structure or activity would be absent.[44]

Cosmology still faces challenges, doubts, and unanswered questions despite advances in understanding the universe's beginnings and development. For instance, dark matter and dark energy together make up around 95% of the universe's total mass energy, yet current astrophysics and cosmology still don't fully understand these aspects of our world. We know very little about them, despite the fact that we can deduce their presence and even quantify them to some extent.[45] Researchers are also still perplexed by the singularity's enigmatic genesis, which is the apparent beginning point from whence the cosmos formed. In an effort to develop a cohesive description of the cosmos, modern scientific theories like loop quantum gravity and string theory have tried to combine Einstein's relativity with quantum mechanics. At best, this work is still unfinished.

One of humanity's riddles that has fascinated myths, philosophies, and scientific investigations is the origin of our universe. Our understanding of the origins of the universe has changed throughout time, moving from cosmological tales that portray primordial chaos to modern cosmological theories developed via complex mathematical analysis and computation. Our curiosity, inventiveness, and will to solve the puzzles of our own existence in the great universe are reflected in this development. The grandeur, complexity, and magnificence that characterize our ever-growing comprehension of the universe humble us as we dig further into cosmic dynamics via scientific investigation.[46] Every cosmological theory offers an intriguing view of the universe's creation and development, whether it is about the Big Bang, Cosmic Inflation, or the notion of a multiverse full of realities. Curiosity, awe, and a profound sense of interconnectedness with the universe and one another on Earth are all sparked by it. Therefore, we should keep trying to comprehend the cosmos and see where the truth takes us.[47]

According to the Dead Universe idea, our understanding of the universe is based on the remnants of an ancient reality whose splendor has long since vanished into the abyss. It's possible that the cosmos we live in is like a cosmic afterglow, a muted remnant of a once-bright and vast cosmic past. The black holes that fill our night sky are assumed to represent the remains of a past cosmic end, the cemeteries of galaxies and stars that have long ago died, rather than the initiators of genesis. Every star system and nebula that we can see with our telescopes may be an expression of cosmic memory, a faint echo from a cosmos that has passed its peak. According to this perspective, dark energy and dark matter are reinterpreted as the lingering characteristics of this ancient era, perhaps the last remnants of a once vibrant cosmic framework. The remnants of an earlier structure in a condition of stately and methodical breakdown are what give rise to the young galaxies that we see, rather than a vacuum. Our universe may have been partly

---

[41] Caldwell, Robert R. (2002). "A phantom menace? Cosmological consequences of a dark energy component with super-negative equation of state". Physics Letters B. 545 (1–2): 23–29. arXiv:astro-ph/9908168. Bibcode:2002PhLB.545...23C. doi:10.1016/S0370-2693(02)02589-3, accessed on 17 Jan 2026

[42] Wang, Yun; Kratochvil, Jan Michael; Linde, Andrei; Shmakova, Marina (2004). "Current observational constraints on cosmic doomsday". Journal of Cosmology and Astro-Particle Physics. 2004 (12): 006. arXiv:astro-ph/0409264. Bibcode:2004JCAP...12.006W. doi:10.1088/1475-7516/2004/12/006, accessed on 17 Jan 2026

[43] Frampton, P. H. (1976). "Vacuum Instability and Higgs Scalar Mass". Physical Review Letters. 37 (21): 1378–1380. Bibcode:1976PhRvL.37.1378F. doi:10.1103/PhysRevLett.37.1378, accessed on 07 Nov 2025

[44] https://science.nasa.gov/dark-matter/, accessed on 17 Jan 2026

[45] Glanz, James (1998). "Breakthrough of the year 1998. Astronomy: Cosmic Motion Revealed". Science. 282 (5397): 2156–2157, doi:10.1126/science.282.5397.2156a. S2CID 117807831, accessed on 17 Jan 2026

[46] Crane, Leah (29 June 2024). de Lange, Catherine (ed.). "How big is the universe, really?". New Scientist. p. 31,

[47] https://www.apu.apus.edu/area-of-study/math-and-science/resources/origin-of-the-universe/, accessed on 17 Jan 2026

sculpted by the debris left by its predecessor, just how stars are born from dense cosmic nurseries.

In their billions of years of existence, the brilliant stars and galaxies we see may very likely represent the last products of a previous universe. Nearing its conclusion, it still had the potential to produce new stars, suggesting that the end of one cosmic cycle and the beginning of another are inextricably linked, culminating in a final event estimated to occur in about 200 billion years. The unchangeable rules of conservation and transmutation that govern all of natural reality are followed by these occurrences that are visible in our current cosmos. These young galaxies might be seen as the last remnants or shining memories of a no longer-existing universe. The final whisper of a cosmos that once flourished in size and energy riches, they are pieces of a huge star legacy. As a result, we are living in the twilight of a magnificent cosmic history, seeing what may be called the "last dance" of matter and light from a vanished cosmos. What we see as our starry reality is just the remnant—a little but still vibrant part of an existence that is far larger than we can comprehend and that transcends both space and time. In essence, all that exists, everything we see, and everything we may learn is nothing more than the surviving remnants of time and space, the eternal mark of the dead cosmos. "Matter tells space how to curve, and space tells matter how to move." In "The Fabric of the Cosmos: Space, Time, and the Texture of Reality," Brian Greene states."[48]

The cosmos becomes less dense and complicated as this process progresses. Dense groupings of matter and energy have given way to increasingly large, empty areas interspersed with solitary islands of star activity. Thus, the James Webb Space Telescope's observations of newborn galaxies provide a window into this process of decline, illuminating the latter phases of a universe we are just starting to comprehend. In this image, the ancestral universe's demise was a protracted phenomenon rather than an instantaneous catastrophe, allowing new structures to gradually rise from its rubble. Black holes store the history of everything that previously been in their gravitational abysses, acting as the last defenders of the cosmic memory of the previous universe rather than the initiators of a new beginning. Black holes do have mass. Depending on the kind of black hole, its mass may be far larger than that of the Earth or the Sun. Supermassive black holes may have masses that are millions or billions of times that of the Sun, whereas stellar black holes typically have masses between a few and tens of times that of the Sun.

These objects are known as "black holes" because they are areas of space with such strong gravity that nothing can escape from them, not even light. This "trapping" characteristic is referred to as a "hole," even though it is not a hole in the conventional sense of an entrance or hollow. The term "black" is employed because a black hole is entirely dark and neither emits nor reflects light, making it seem "black" to any observer as light cannot escape from it. A phenomenon known as gravitational collapse may occur when certain stars, which are much more massive than the Sun, approach the end of their lifetimes. The pressure holding the star against gravity vanishes when they run out of nuclear fuel, and the star falls in on itself. A supernova may be produced by this collapse, and the surviving core may generate a stellar black hole, depending on the star's initial mass. This is an illustration of a black hole that comes from a "dead star." We become closer to the idea of a dead universe with dimensions larger than our visible universe in this manner.

The orbits of the planets in the solar system, including Earth, would initially stay intact, at least for a while, but the repercussions would be severe if the Sun abruptly ceased providing heat and light. This is due to the fact that the planets' orbit around the Sun is maintained by gravity rather than light. Light is a kind of energy that an object emits, and gravity is a result of its mass. The Sun's mass would still be there, so its gravity would still affect the planets even if it abruptly ceased generating light, which would indicate that nuclear processes are no longer occurring in its core. On the other hand, life on Earth and the planet's climate would suffer greatly in the absence of heat and light. The planets' orbits may change over time if the Sun underwent a process that drastically changed its mass, such as turning into a white dwarf. The trajectory of celestial bodies circling the Sun would be impacted by changes in the Sun's mass, which would also change its gravitational pull.

The idea of a universe that is made up of galaxies, supermassive black holes, dark matter, dark energy, and where the rules of physics are still in effect despite the absence of light emission is not out of the ordinary. Based on the presented theoretical reasoning, I can assert that such a world exists and that scientific knowledge may soon expose it. However, from a scientific perspective, the assertion that a basic fact, such a "dead universe," exists requires a considerable body of theoretical and empirical data that can be independently confirmed by testing and observations. Like many ideas and hypotheses that have come before it, this one is now in the realm of theoretical conjecture and should be treated cautiously until such proof is presented and confirmed by the scientific community's examination. Nearly all massive galaxies, including the Milky Way, have these black holes at their cores. Their masses are between millions and billions of times greater than the Sun's. Although their precise genesis is still unknown, it is thought that they expand over time by gathering materials and other black holes. Although they are an essential component of the dynamics and development of galaxies, they are not regarded as "dead galaxies."

Usually, a galaxy that has stopped producing stars is referred to as a "dead galaxy." Galaxies may "die" in terms of producing stars for a variety of reasons, including interactions with neighboring galaxies or the loss of gas, which is the fuel for star formation. These galaxies may contain supermassive black holes at their cores, but they do not turn into black holes. In contrast to the idea of multiverses proposed by numerous speculative

[48] Greene, B. (1999) The Fabric of the Cosmos: Space, Time, and the Texture of Reality. Vintage Books, accessed on 07 Nov 2025

theories, the "dead universe" hypothesis suggests that the aggregate deaths of celestial bodies converge in the construction of a single preceding world, which justifies and merits further investigation. These views often diverge from mathematical models and the scientific data that has been thoroughly examined and validated. The dead universe idea, on the other hand, provides a viewpoint that fits within the current knowledge of the world while also offering a possible foundation for further research since it is consistent with known findings and physics principles.

While the Dead Universe theory not only complies with the physical laws already in place but also offers substitute explanations that are easily verified through observation and experimentation, the consensus surrounding the Big Bang theory currently seems to be waning. Therefore, the "expansion of the universe" might be seen as a progressive separation from a previous universe, a remnant still impacting the present cosmos, driven by the rules of gravity rather than as a sign of dynamic development. One may think of this occurrence as the universe's last sigh as it progressively releases its energy. We are seeing a cooling and quiescent process in which space-time expands, matter and energy are gradually redistributed, and a condition of lasting tranquility is sought. A cosmos full of thoughtful calm and the genuine quiet that follows the bright intermission of the stars would come from this process, which tends to reduce and finally stop the development of new galaxies. The observable universe, its progeny, will expire just as its parent universe did.

2.5 Experimental Indication from the James Webb Space Telescope. The Big Bang hypothesis has long served as the foundation of cosmology, giving us a picture of a universe that was created from a singularity and expanded for around 13.5 billion years. But as more information emerges, it is becoming more and more obvious that this story has serious problems, opening the door for a different viewpoint—the "dead universe" idea that I put out. A completely new strategy is suggested by the Dead Universe idea. It suggests that the cosmos is a huge and potentially everlasting continuum where notions of origin and end are relativized, rather than being the product of an explosion. The findings of the James Webb provide hard data that casts doubt on the central idea of the Big Bang; therefore, this is more than simply a theoretical theory. The astonishing preservation of ancient galaxies that should exhibit interactions and mergers, as expected by the conventional model, points to a much more intricate and nonlinear cosmic past.

A major obstacle to modern cosmology is the observation of celestial objects that seem to be older than the universe's age as determined by the Big Bang concept. How can a universe that is now estimated to be 13.5 billion years old be reconciled with the presence of these mature structures? In an attempt to resolve this discrepancy, the Dead cosmos theory suggests that these galaxies are not only anomalies but are indicators of an ancient cosmos whose history goes beyond the temporal scale defined by the Big Bang. According to this hypothesis, traditional cosmological timelines could need to be updated in light of fresh data, potentially improving our comprehension of the universe's past and present. Furthermore, new data cast doubt on the Big Bang model's central tenet—the universe's alleged uniform expansion. The behavior of distant and old galaxies does not support fast and continuous expansion. A fundamental issue is brought up by this: what if the cosmos is not expanding evenly, or maybe not at all? According to a new idea put out by scientists, the universe may be in a more complicated, static, or inverted state than previously thought, one in which space and time are relative rather than absolute and interwoven in a manner that we are still learning about.

This is an opportunity to fundamentally reconsider our conception of the universe, not just a challenge to the prevailing story. Since the primordial nature of the Dead Universe was not light but rather the darkness of dark matter and supermassive bodies, the theory suggests a "timeless" universe or one that creates its own strange body, light. Beginning and end are human concepts rather than universal realities. From the standpoint of the dead universe, the fusing of black holes and the ensuing formation of stars can be seen as inexplicable occurrences for the creatures who inhabit this world. Imagine a society developing in the middle of unending darkness, where light is a concept that is abstract and almost legendary. The abrupt appearance of brilliant spots in the sky would be incomprehensible to them, an abnormality in an otherwise gloomy setting. It's possible that the UNO matter equation is similar to a window tint or solar control screen in that we see light while we are inside, but when we gaze from the outside in, we see no light at all and everything appears to be black and lightless. Therefore, because of the existence of a substance that I characterize as neutral, a world that is enmeshed in death inside a black fabric may show a reality of glorious brightness that we cannot perceive.

This is not just a challenge to the dominant narrative, but also a chance to radically reevaluate how we think about the cosmos. The idea proposes a "timeless" world or one that generates its own peculiar body, light, as the fundamental nature of the Dead world was not light but rather the blackness of dark matter and supermassive bodies. Rather of being universal truths, beginning and end are human notions. The merging of black holes and the subsequent creation of stars may be seen from the perspective of the dead universe as unexplainable events for the animals who live here. Imagine a culture growing up in the midst of endless darkness, where light is an almost mythical and abstract idea. They would not understand the sudden apparition of bright spots in the sky, an anomaly in a generally dreary scene. It's likely that the UNO matter equation functions similarly to a solar control screen or window tint in that, from the outside, everything seems black and lightless, but when we look in, we see light. Therefore, a world entangled in death within a black fabric may reveal a reality of glorious brightness that we are unable to perceive due to the existence of a substance that I describe as neutral.

Theories of cosmology that suggest different kinds of "barriers" or areas of transition in the cosmos. A black hole's event horizon, for instance, functions as a point of no return when gravitational pull is so powerful that not even light can escape, rendering it invisible from the outside. This is similar to looking at a dark window from the outside; you can't see through it, which gives the impression that there isn't any light or activity inside, yet there is really hidden treasure. Extending this to our concept, if there were a "UNO matter" that functioned as this type of cosmic hue, it might be a hypothetical substance or field that exists within the universe's structure and interacts with light and other forms of energy in a way that, when viewed from a particular angle, obscures the activity or underlying structure of the cosmos. Theoretically, such material may be in charge of the events we see, like as the impacts of dark matter, which affects galaxy motion but doesn't produce any observable light or radiation, making it "UNOl," or "invisible," to our existing observational techniques.

In cosmology, the term "domain wall" refers to a hypothesized structure that may serve as a partition between several vacuum state types or phases in the universe, analogous to the interface between two bubbles. Similar to our idea of "UNO matter" film, it's a theoretical idea that may help explain cosmic separation or transition regions. Though they may be helpful in explaining ideas, analogies are usually used sparingly and are always supported by sound reasoning and actual data in scientific papers. Furthermore, the basic principles of this dead cosmos may be called into question by the sheer absence of light as a primal constituent. Even though they live in a world where darkness is king, light may be seen as an encroachment or even as a metaphysical impossibility. These thoughts make us wonder whether, given our constrained viewpoint as cosmic watchers, we can indeed understand the whole of the universe. It's possible that the dead cosmos represents a spectrum of existence that is beyond our complete comprehension, and that what we take for granted are just a tiny portion of cosmic reality.

Perhaps the core of the dead world is in fact in opposition to the fundamental nature of light. Light, an object alien to its existence, was created by the mergers of black holes and supermassive entities, which constituted the universe's primordial essence. This cosmos will endure indefinitely, with light shining in contrast to its own endless darkness. This does not, however, imply that the substance of this dead world is our visible reality. Strange bodies like the galaxies we see are the result of particle unions and irregular behaviors that changed the universe's initial order. In this way, unless there is a creator entity for the dead universe, we are only encroaching on this reality by accident. Light is a peculiar aspect of the dead world's existence, if we may say so, because it will always exist with its own rules and nature. This peculiar universe contains light as a fundamental component of its essence and nature. In this way, the existence of humans and life as we know it is not debatable. According to the Abrahamic faiths, "no one can deny that the universe is more for darkness, chaos, and obscure mystery than for a reality of light."[49] This is an exciting time to investigate, ask questions, and perhaps learn more about the real nature of the universe. Since we are in the eternal time of the dead cosmos, the present will always be our time (Thorne, 1994). All of the old issues and many new ones would be raised by a single knowledge that integrates the quantum and classical worlds, which would answer some but leave the majority unsolved."[50]

The dark matter conundrum is a problem in physics. Such stuff is thought to be made up of hypothetical, indescribable particles called sterile neutrinos or compact, supermassive objects like primordial black holes. But the fundamental idea that is thought to explain dark matter resonates more in the context of the Dead Universe hypothesis than it does in the Big Bang framework. The idea that this process produced energy, like the mysterious cosmic mystery of dark energy, is supported by the presence of a previous and extinct world that was completely dark. This hypothesis holds that the rules of the previous universe are still in force and that dark energy is not the cause of cosmic expansion (Lee Smolin 2006). The Dead Universe Theory considers particle behavior, radio waves, and dark matter. However, for strictly scientific reasons, a creative agency does not disprove the Dead Universe idea. The goal of science is not to prove that the divine exists. It simply looks at natural things and uses empiricism to explain them. It also doesn't exist in order to reject the divine. Therefore, let's focus on what can be explained in the Theory of the Dead Universe and put aside what cannot be explained (Carroll, 2010). "Consider a world where any one of these figures was different. A world devoid of atoms, stars, planets, humans, and all other forms of life as we know them would exist. It would be a history-free universe. But the rules of physics as we know them would be perfectly compatible with such a reality. So why do we live in a world that is perfect for us?"[51]

2.6 Clarification for the Cold Spot in the Universe. A gigantic, under-dense area known as a super-void that CMB photons pass through and lose energy (integrated Sachs-Wolfe effect) is the best explanation for the "Cold Spot" in the cosmos, which is a sizable, abnormally chilly patch in the Cosmic Microwave Background (CMB), the afterglow of the Big Bang.[52] Although the dominant theory is a super-void, the spot's enormous size and depth defy conventional theories; some even speculate on unusual origins, such as the imprint of another universe, yet the majority of the data suggests a massive emptiness. The WMAP (Wilkinson Microwave Anisotropy Probe) satellite made the discovery in 2004, and data from the ESA's Planck mission subsequently verified it. This location defies the conventional explanation based on the Big Bang theory's prediction of the universe's homogeneity

[49] Thorne, K. (1994) Black Holes and Time Warps: Einstein's Outrageous Legacy. W. W. Norton & Company, accessed on 07 Nov 2025

[50] Smolin, L. (2006) The Life of the Cosmos. Oxford University Press, accessed on 17 Jan 2026

[51] Carroll, S. (2010) From Eternity to Here: The Quest for the Ultimate Theory of Time. Goodreads

[52] Wright, E.L. (2004). "Theoretical Overview of Cosmic Microwave Background Anisotropy". In W. L. Freedman (ed.). Measuring and Modeling the Universe. Carnegie Observatories Astrophysics Series. Cambridge University Press. p. 291, ISBN 978-0-521-75576-4, accessed on 07 Nov 2025

since it is around 70 microkelvins cooler than the average CMB.[53] According to this viewpoint, anomalies like as the Cold Spot are direct representations of the harsh circumstances and rules of the previous universe rather than just random fluctuations or the result of unidentified cosmic superstructures. The temperature fluctuations seen in the cosmic background radiation might be caused by gravitational pull or other lingering factors from this dead cosmos. Because of its state of cosmic death, temperature is a property of the dead universe rather than the visible cosmos. The idea that multiverses are colliding with this universe seems unlikely given its billion-year history; undoubtedly, other cold spots would have been encountered, but they do not exist because, for trillions of years, this cold spot served as the connection between this universe and the mother universe.

2.7 Constraint of Big Bang Theory. Anisotropies like the Cold Spot are difficult for the Big Bang theory to adequately explain, despite its effectiveness in describing the majority of the universe's known properties, such as cosmic expansion and the preponderance of light elements. The standard model predicts that cosmic inflation would cause temperature variations in the CMB to be comparatively consistent across vast scales. Because of its size and depth, the Cold Spot is difficult to incorporate into this model without the need for more intricate explanations, such as uncommon statistical fluctuations or massive, undiscovered cosmic superstructures. Thus, the Horizon Problem (uniform CMB temperature) is one of the main drawbacks of the Big Bang hypothesis, which explains a large portion of cosmology.[54], Flatness Problem (universe's geometry)[55], and Monopole Problem (lack of magnetic monopoles)[56], which are addressed by the Inflation theory, but this introduces its own issues like explaining the initial conditions and the nature of dark matter/energy.[57] Additionally, the theory has trouble explaining the matter-antimatter imbalance and the first singularity (Planck Epoch). According to the Inflation Theory, the cosmos expanded very quickly (exponentially) in its early stages. It was created in the 1980s to address a number of issues with the conventional Big Bang hypothesis, which holds that the universe expands over time somewhat slowly.[58] It is uncertain what specific particle physics process causes inflation. A number of inflation model predictions were verified by observation in 1992. For instance, temperature anisotropies seen by the COBE satellite reveal almost scale-invariant spectra, as anticipated by the inflationary paradigm, and WMAP data also provide compelling support for inflation.[59] However, some scientists dissent from this position.[60],[61]

Determining why the cosmos seems statistically homogenous and isotropic in line with the cosmological principle is known as the "horizon problem."[62] For instance, since the molecules in a gas canister are in thermal equilibrium, they are dispersed uniformly and isotopically. This is because the gas has had enough time to interact with one another to dissipate anisotropies and inhomogeneities.[63] Regions are too rapidly separated by gravitational expansion in the Big Bang scenario without inflation. Thus, there is not enough time for the early cosmos to reach equilibrium. Two widely distant parts of the visible universe cannot have equilibrated in a Big Bang with just the matter and radiation described in the Standard Model since they travel apart faster than the speed of light and have never had causal touch.[64] A cosmological fine-tuning issue with the Big Bang model of the universe is the flatness problem, often referred to as the oldness problem. The Universe is flat to within a few percent, according to observations of the cosmic microwave background.[65] Flatness rises as the universe expands. The early universe must therefore have been remarkably near to flat. The density of matter and energy in the universe influences the curvature of space-time in traditional cosmology, which is based on the Friedmann equations. A flat universe necessitates a very particular critical number.[66] It is found that the universe's present density is really near to this crucial number. Given that the overall density would rise quickly over cosmic time

[53] https://www.universetoday.com/articles/finally-an-explanation-for-the-cold-spot-in-the-cosmic-microwave-background, accessed on 07 Nov 2025
[54] Carrigan, Richard A.; Trower, W. Peter (1983). Magnetic Monopoles. doi:10.1007/978-1-4615-7370-8. ISBN 978-1-4615-7372-2, accessed on 17 Jan 2026
[55] Peacock, J. A. (1998). Cosmological Physics. Cambridge: Cambridge University Press. ISBN 978-0-521-42270-3,
[56] Hooper, Dan (October 6, 2009). Dark Cosmos: In Search of Our Universe's Missing Mass and Energy. Harper Collins. ISBN 978-0-06-197686-5, accessed on 17 Jan 2026
[57] https://ui.adsabs.harvard.edu/abs/1992BASI...20..157N/abstract, accessed on 17 Jan 2026
[58] Tyson, Neil deGrasse and Donald Goldsmith (2004), Origins: Fourteen Billion Years of Cosmic Evolution, W. W. Norton & Co., pp. 84–85, accessed on 07 Nov 2025
[59] Tsujikawa, Shinji (28 April 2003). "Introductory review of cosmic inflation". arXiv:hep-ph/0304257, accessed on 07 Nov 2025
[60] Guth, Alan H. (1997a). The Inflationary Universe: The quest for a new theory of cosmic origins. Basic Books. pp. 233–234. ISBN 978-0-201-32840-0, accessed on 17 Jan 2026
[61] Earman, John; Mosterín, Jesús (March 1999). "A Critical Look at Inflationary Cosmology". Philosophy of Science. 66 (1): 1–49. doi:10.1086/392675. JSTOR 188736. S2CID 120393154 , accessed on 07 Nov 2025
[62] Misner, Charles W.; Coley, A A; Ellis, G F R; Hancock, M (1968). "The isotropy of the universe". Astrophysical Journal. 151 (2): 431. Bibcode:1998CQGra. 15..331W. doi:10.1088/0264-9381/15/2/008, accessed on 07 Nov 2025
[63] Misner, Charles; Thorne, Kip S. & Wheeler, John Archibald (1973). Gravitation. San Francisco: W. H. Freeman. pp. 489–490, 525–526. ISBN 978-0-7167-0344-0, accessed on 17 Jan 2026
[64] Weinberg, Steven (1971). Gravitation and Cosmology. John Wiley. pp. 740, 815. ISBN 978-0-471-92567-5,
[65] https://map.gsfc.nasa.gov/universe/uni_matter.html, accessed on 07 Nov 2025
[66] Lightman, Alan P. (1 January 1993). Ancient Light: Our Changing View of the Universe. Harvard University Press. ISBN 978-0-674-03363-4, accessed on 17 Jan 2026

if it deviated from the crucial value,[67] The density of the early cosmos must have been even nearer the critical density, deviating from it by no more than one part in 1062. Cosmologists wonder how the original density was so precisely tailored to this "special" number.[68]

2.8 Disagreement with Cosmic Radiation Theory. The Cold Spot indicates anisotropies that could call for novel physics or modifications to existing cosmogonic theories, even if cosmic background radiation typically supports a uniform and homogeneous cosmos as anticipated by inflation. Given this, the Big Bang hypothesis is unable to directly explain this anomaly, which begs the issue of whether the model should be modified or expanded. An alternate explanation for the Cold Spot is provided by the Dead cosmos idea, which proposes that it is a "umbilical cord" from a previously collapsed cosmos. According to this idea, our visible world is only a residue of both the dead universe and a much bigger, older cosmos. Irregularities like the Cold Spot might be caused by gravity and effects from this former cosmos, which is currently only partly present. By adding the concepts of a multiverse or cosmic cycles of death and rebirth, this method not only provides an explanation for the anomaly but also advances our knowledge of the cosmos. An interesting explanation for the genesis of anomalies such as the Cold Spot in the cosmic microwave background radiation is offered by the Dead Universe idea. According to this view, the Cold Spot is a direct result of the temperature condition of a now-extinct predecessor universe rather than just a random fluctuation. Consider a massive, aged universe that gradually cools down to a large area with little thermal energy, similar to a cold chamber in space.

This comparison may be compared to opening a little door between a warmer space, like a kitchen, and an excessively cold one, like a freezer. The influence of the dead universe on the surrounding space, particularly at locations where the interaction is most strong, is comparable to the instantaneous heat exchange that takes place. In the context of the visible cosmos, this thermal exchange produces a region that is considerably colder, which we identify as the Cold Spot. Like the way cold air from a freezer mixes with warmer air in a kitchen, this equation aims to measure the direct impact of the dead universe's severe cold on the visible universe. This comparison, together with the equation that goes with it, offers a vivid and scientifically sound explanation of how the temperature of the cosmos as it exists now may be affected by an old and cold universe. Thorough observations and a thorough examination of the cosmic microwave background radiation's anisotropies, searching for particular patterns that would support this thermal interaction on a cosmic scale, are necessary for the validation of this idea.

According to the hypothesis of the previous universe's enormous gravitational magnitude, space-time may naturally distort; this phenomenon is referred to as a "gravitational well" in astrophysics and is what bends light, for instance. What astrophysics has failed to explain may be explained by the theory that the visible universe lies within the womb of a deceased mother universe that perished trillions of years ago, the same destiny as our world, which came from the womb of the previous mother. The cosmos's fabric may have been bent by the gravitational pull of the old universe, causing a "slippery" to advance across space without really moving. Despite being acknowledged as the explanation for the universe's beginning, the Big Bang hypothesis contains flaws, such as the inability to account for the universe's ongoing expansion. This alternate theory may be supported by studies linking particle accelerators that show micro-explosion-like occurrences. The presence and behavior of black holes, which provide indirect evidence of this process, could be interpreted as an expansion driven by the residual action of the gravity of a previous universe, if the observable universe originated from a "dead universe." The persistence of gravitational rules, which seem to have governed unchanged since the primordial state, might be evidence of the close relationship between the present universe and its potential antecedent in a more expansive setting.

The fact that expansions brought on by explosive events often include a degree of unpredictability in the motion of the associated particles is a relevant point of contention in the Big Bang hypothesis. But the universe's expansion points to a more regular and orderly development, perhaps driven by ideas that modern physics hasn't yet completely clarified. In reference to the description of the “Explosion” of the Big Bang itself, the phrase can be considered unsuitable if seen in the context of typical explosions. What physical processes would support such a concept if such an event did not fall within the conventional definition of an explosion? A source of energy that can support the Big Bang theory, which postulates the expansion of spacetime itself, is necessary. Furthermore, the Big Bang process refers to the expansion of spacetime itself rather than an explosion inside an existing place. The Big Bang model, which sometimes appears to waver in its explanations about the precise nature of the original event, leaves gaps that may be filled by the alternative explanation offered by the "Great Dead Universe" idea, which could provide a thorough account of cosmic expansion.

Furthermore, a traditional explosion's claimed chaotic and random beginning may appear incompatible with the universe's regularity and ordered structure. Explosive characteristics may be part of the universe's nature, according to scientific research, especially those based on quantum physics concepts. Therefore, the beginning circumstances and physical rules of this earlier universe may be the governing keys of the expansion we see today if the visible world is impacted by a previous cosmic heritage (Rees 2000). "The theory of everything is a bold attempt in theoretical physics to bring together the weak nuclear force, strong nuclear force, electromagnetism, and gravity—the four fundamental

[67] Peacock, J. A. (1998). Cosmological Physics. Cambridge: Cambridge University Press. ISBN 978-0-521-42270-3,

[68] Cox, Brian; Forshaw, J. R. (2017). Universal: a guide to the cosmos (1 ed.). Boston, MA: Da Capo Press. ISBN 978-0-306-82270-4. OCLC 973019447, accessed on 17 Jan 2026

forces of the universe." Sean Carroll said this in "From Eternity to Here: The Quest for the Ultimate Theory of Time."[69] “Each and every atom in your body originated from an exploding star. Furthermore, your left hand's atoms most likely originated from a different star than your right. You are all stardust, which is the most beautiful thing I've ever heard about physics. According to Lawrence M. Krauss, "A Universe from Nothing: Why There Is Something Rather than Nothing" (Krauss, 2012).[70] Every every atom in your body came from a star that exploded. Moreover, the atoms in your left hand most likely came from a different star than those in your right. The most exquisite thing I've ever heard about physics is that you are all stardust. As stated in "A Universe from Nothing: Why There Is Something Rather than Nothing" by Lawrence M. Krauss.

2.9 Contemporary Cosmology. Theoretical and experimental efforts interact dynamically in modern cosmology, which is always changing to overcome new obstacles. In order to reconcile theory with new empirical evidence at every stage, the field requires methodical reconstruction. The discovery of supernova fading marked a turning point in this continuing discussion,[71] a phenomenon that revealed the limitations of the Friedmann–Lemaitre–Robertson–Walker metric (here Friedmann metric). The cosmological constant was suggested to resolve this contradiction and bring the theoretical predictions into line with actual findings.[72] Galaxies are becoming further away from us, according to current surveys and astronomical studies. Understanding how structures arise and how galaxies evolve against the background of the late-time universe's accelerating expansion is a major question at the center of contemporary cosmological arguments.[73] Based on the cosmological principle, the Friedmann model has successfully explained the development of the universe in accordance with actual findings.[74],[75],[76] .[77],[78],[79] However, the mystery of dark energy and the force driving cosmic acceleration remains a tenacious challenge in contemporary physical cosmology .[80],[81] Concepts like the cosmological constant or dark energy-dominated scenarios are used in a variety of attempts to explain cosmic acceleration. Nonetheless, there are important questions raised by the cosmological constant.[82],[83],[84]Adding to these difficulties is the potential violation of the cosmological principle when homogeneity or isotropy falters in galaxy structure formation.[85],[86]

Surveys of three-dimensional redshift look further into space and uncover features without a transition to homogeneity.[87],[88] Now, questions arise regarding the

[69] Rees, M. (2000) Just Six Numbers: The Deep Forces That Shape the Universe. Hachette Book Group, accessed on 07 Nov 2025

[70] Krauss, L. M. (2012) A Universe from Nothing: Why There Is Something Rather than Nothing. Atria Books

[71] Yang, T.; Liu, T.; Huang, J.; Cheng, X.; Biesiada, M.; Wu, S. Simultaneous measurements on cosmic curvature and opacity using latest HII regions and H(z) observations. Eur. Phys. J. C 2024, 84, 3, accessed on 07 Nov 2025

[72] Riess, A.G.; Filippenko, A.V.; Challis, P.; Clocchiattia, A.; Diercks, A.; Garnavich, P.M.; Gilliland, R.L.; Hogan, C.J.; Saurabh, J.; Kirshner, R.P.; et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. Astron. J. 1998, 116, 1009–1038, accessed on 17 Jan 2026

[73] Hu, J.P.; Wang, Y.Y.; Hu, J.; Wang, F.Y. Testing the cosmological principle with the Pantheon + sample and the region-fitting method. Astron. Astrophys. 2024, 681, A88, accessed on 07 Nov 2025

[74] Li, J.; Yang, Y.; Yi, S.; Hu, J.; Wang, F.; Qu, Y. Constraints on the Cosmological Parameters with Three-Parameter Correlation of Gamma -Ray Bursts. Astrophys. J. 2023, 953, 58, accessed on 07 Nov 2025

[75] Perez, J.d.; Park, C.; Ratra, B. Current data are consistent with at spatial hypersurfaces in the λCDM cosmological model but favor more lensing than the model predicts. Phys. Rev. D 2023, 107, 063522, accessed on 07 Nov 2025

[76] Khadka, N.; Zajace, K.M.; Prince, R.; Panda, S.; Czerny, B.; Aldama, M.L.; Jaiswal, V.K.; Ratra, B. Quasar UV/X-ray relation luminosity distances are shorter than r reverberation-measured radius-luminosity relation luminosity distances. Mon. Not. R. Astron. Soc. 2023, 522, 1247–1264, accessed on 17 Jan 2026

[77] Li, J.; Yang, Y.; Yi, S.; Hu, J.; Wang, F.; Qu, Y. Constraints on the Cosmological Parameters with Three-Parameter Correlation of Gamma -Ray Bursts. Astrophys. J. 2023, 953, 58, accessed on 17 Jan 2026

[78] Perez, J.d.; Park, C.; Ratra, B. Current data are consistent with at spatial hypersurfaces in the λCDM cosmological model but favor more lensing than the model predicts. Phys. Rev. D 2023, 107, 063522, accessed on 07 Nov 2025

[79] Khadka, N.; Zajace, K.M.; Prince, R.; Panda, S.; Czerny, B.; Aldama, M.L.; Jaiswal, V.K.; Ratra, B. Quasar UV/X-ray relation luminosity distances are shorter than r reverberation-measured radius-luminosity relation luminosity distances. Mon. Not. R. Astron. Soc. 2023, 522, 1247–1264, accessed on 17 Jan 2026

[80] Mégier, E.A. Square–torsion gravity, dark matter halos and the baryonic Tully–Fisher relation. Eur. Phys. J. C 2020, 80, 1157, accessed on 17 Jan 2026

[81] Bernard, R.C.; Grandon, D.; Said, J.L.; Cardenas, V.H.; Bernardo, R.C. Parametric and nonparametric methods hint dark energy evolution. Phys. Dark Universe 2022, 36, 101017, accessed on 17 Jan 2026

[82] Del Popolo, A.; Le Delliou, M. Small Scale Problems of the LCDM Model: A Short Review. Galaxies 2017, 5, 17, accessed on 17 Jan 2026

[83] Gomes, L.G. Breaking the Cosmological Principle into pieces: A prelude to the intrinsically homogeneous and isotropic space times. arXiv 2024, arXiv:2401.01992, accessed on 17 Jan 2026

[84] Nesterov, A.I. Spacetime Foam and Solution of the Cosmological Constant Problem. arXiv 2024, arXiv:2401.04,

[85] Melia, F.; Shevchuk, A.S. The Rh = ct universe. Mon. Not. R. Astron. Soc. 2012, 419, 2579–2586, accessed on 11 Nov 2025

[86] Abdalla, E.; Abellán, G.F.; Aboubrahim, A.; Agnello, A.; Akarsu, Ö.; Akrami, Y.; Alestas, G.; Aloni, D.; Amendola, L.; Anchordoqui, L.A.; et al. Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associatedwith the cosmological tensions and anomalies. J. High Energy Astrophys. 2022, 34, 49–211, accessed on 19 Jan 2026

[87] Conn, A.R.; Lewis, G.F.; Ibata, R.A.; Parker, Q.A.; Zucker, D.B.; McConnachie, A.W.; Martin, N.F.; Valls-Gabaud, D.; Tanvir, N.; Irwin, M.J.; et al. The Three-Dimensional Structure of the M31 Satellite System; Strong Evidence for an Inhomogeneous Distribution of Satellites. Astrophys. J. 2013, 766, 120, accessed on 19 Jan 2026

[88] Doliva-Dolinsky, A.; Martin, N.F.; Yuan, Z.; Savino, A.; Weisz, D.R.; Ferguson, A.M.; Ibata, R.A.; Kim, S.Y.; Lewis, G.F.; McConnachie, A.W.; et al. The PAndAS View of the

steadfastness of the cosmological principle.[89] On the highest statistical scale accessible, the distribution of galaxies in recent observations (or light) and simulations of the distribution of dark matter (or matter) show notable inhomogeneity.[90] The search for the cosmological principle in the present visible light or matter distribution in the cosmos is complicated by the much more inhomogeneous matter distribution.[91] Recent model-independent examinations of a brilliant red galaxy sample utilizing the Sloan Digital Sky Survey's Sixteenth Data Release (SDSS-IV DR16) revealed an angular scale of cosmic homogeneity of 60–80 h−1 Mpc.[92] A homogeneity test for the matter distribution based on the Baryon Oscillation Spectroscopic Survey Data Release 12 CMASS galaxy sample has contested this conclusion.[93] Up to a radius of 300 h−1 Mpc, which is about the maximum statistically possible scale, it was discovered that the observed distribution of matter is statistically improbable to be a random arrangement.

The discovery of massive quasar groups (LQGs) intensifies the controversy by implying an underlying inhomogeneity that is incompatible with accepted cosmological theories.[94],[95] These discoveries highlight the need of a thorough cosmological reexamination.[96] To prevent misinterpretations of the sources under analysis, accurate testing of the standard model's prediction on the spatial distributions of luminous astronomical sources must be based on high-resolution cosmological simulations involving a sizable sample of isolated galaxies using reliable data-driven detectors.[97] Three-dimensional catalogs provide a complicated picture of inhomogeneous galaxy distributions, but two-dimensional predictions seem consistent with isotropy and homogeneity.[98] Attempts to adopt a single viewpoint are complicated by these conflicting results on the shift to homogeneity.[99],[100] These discoveries' disparate character calls into question the widely accepted notion of cosmic homogeneity and isotropy.[101] Understanding cosmic acceleration and the need for an extra dark energy component may be affected by the consequences.[102],[103]

Given the observable abnormalities of the standard model and its lack of physical reason, researchers think it essential to investigate alternate theories of dark energy or its modified forms to account for the universe's cosmic acceleration.[104],[105],[106],[107] The suggested model incorporates dark energy and modified gravity theories in which the scalar field takes the place of the

Andromeda Satellite System. IV. Global Properties. Astrophys. J. 2023, 952, 72, accessed on 19 Jan 2026

[89] Matos, T.; Urena-Lopez, L.A.; Lee, J. Short Review of the main achievements of the Scalar Field, Fuzzy, Ultralight, Wave, BEC Dark Matter model. Front. Astron. Space Sci. 2024, 11, 1347518, accessed on 19 Jan 2026

[90] Gu, Q.; Guo, Q.; Zhang, T.; Cautun, M.; Lacey, C.; Frenk, C.S.; Shao, S. The spatial distribution of satellites in galaxy clusters. Mon. Not. R. Astron. Soc. 2022, 514, 390–402, accessed on 19 Jan 2026

[91] Kim, Y.; Park, C.G.; Noh, H.; Hwang, J. CMASS galaxy sample and the ontological status of the cosmological principle. Astron. Astrophys. 2022, 660, A139, accessed on 19 Jan 2026

[92] Andrade, U.; Gonçalves, R.S.; Carvalho, G.C.; Bengaly, C.A.; Carvalho, J.C.; Alcaniz, J.U. The angular scale of homogeneity with SDSS-IV DR16 Luminous Red Galaxies. arXiv 2022, arXiv:2205.07819v2, accessed on 19 Jan 2026

[93] Marzo, G.; Labini, F.S.; Pietronero, L. Zipf's law for cosmic structures: How large are the greatest structures in the universe? Astron. Astrophys. 2021, 651, A114, accessed on 19 Jan 2026

[94] Clowes, R.G.; Harris, K.A.; Raghunathan, S.; Campusano, L.E.; Soechting, I.K.; Graham, M.J. A structure in the early Universe at z= 1.3 that exceeds the homogeneity scale of the R-W concordance cosmology. Mon. Not. R. Astron. Soc. 2013, 429, 2910–2916, accessed on 19 Jan 2026

[95] Morrison, S.; Pieri, M.; Som, D.; Pérez-Ràfols, I.; Morrison, S. Probing large-scale UV background inhomogeneity associated with quasars using metal absorption. Mon. Not. R. Astron. Soc. 2020, 506, 5750–5763, accessed on 13 Nov 2025

[96] Pandey, B.; Sarkar, S. Probing large scale homogeneity and periodicity in the LRG distribution using Shannon entropy. Mon. Not. R. Astron. Soc. 2016, 460, 1519–1528, accessed on 19 Jan 2026

[97] Cautun, M.; Bose, S.; Frenk, C.S.; Guo, Q.; Han, J.; Hellwing, W.A.; Sawala, T.; Wang, W. Planes of satellite galaxies: When exceptions are the rule. Mon. Not. R. Astron. Soc. 2015, 452, 3838–3852, accessed on 13 Nov 2025

[98] Labini, F.S. Inhomogeneities in the universe. Class. Quantum Gravity 2011, 28, 16, accessed on 13 Nov 2025

[99] Meszaros, A. An Oppositeness in the Cosmology: Distribution of the Gamma-Ray Bursts and the Cosmological Principle. Astron. Nachrichten 2019, 340, 564–569, accessed on 13 Nov 2025

[100] Milaković, D.; Lee, C.; Molaro, P.; Webb, J.K.; Milakovic, D. Methods for quasar absorption system measurements of the fine structure constant in the 2020s and beyond. arXiv 2023, arXiv:2212.02458, accessed on 13 Nov 2025

[101] Abdalla, E.; Abellán, G.F.; Aboubrahim, A.; Agnello, A.; Akarsu, Ö.; Akrami, Y.; Alestas, G.; Aloni, D.; Amendola, L.; Anchordoqui, L.A.; (2022), Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associatedwith the cosmological tensions and anomalies. J. High Energy Astrophys. 2022, 34, 49–211, accessed on 19 Jan 2026

[102] Ellis, G.F. Inhomogeneity effects in Cosmology. Class. Quantum Gravity 2011, 28, 164001, accessed on 13 Nov 2025

[103] Arun, K.; Gudennavar, S.B. Prasad. Effects of Dark Matter in Star Formation. arXiv 2019, arXiv:1902.01815, accessed on 19 Jan 2026

[104] Abaca, F.; Zamora, D.J. Multicomponent holographic dark energy model with generalized entropy. arXiv 2024, arXiv:2401.17324, accessed on 13 Nov 2025

[105] Tian, S. The relation between cosmological redshift and scale factor for photons. Astrophys. J. 2017, 846, 1538–4357, accessed on 19 Jan 2026

[106] Shahalam, M.; Ayoub, S.; Avlani, P.; Myrzakulov, R. Dynamical system analysis in descending dark energy model. arXiv 2024, arXiv:2402.01270, accessed on 13 Nov 2025

[107] Wang, B.; Abdalla, E.; Atrio-Barandela, F.; Pavón, D. Further understanding the interaction between dark energy and dark matter: Current status and future directions. arXiv 2024, arXiv:2402.00819v1, accessed on 13 Nov 2025

cosmological constant.[108],[109] The Friedmann model and the concordance model of cosmology in general are challenged by recent data such as the unexplained Hubble parameter tensions, large-scale anisotropies, and enormous disk galaxies at higher redshifts. For instance, the Hubble parameter derived from Type Ia supernovae and the redshift of their host galaxies is different from that derived from the cosmic microwave background (CMB) radiation.[110],[111] Other hypotheses suggest that the classic redshift model, as a distance–scale factor connection, may be imperfect, whereas the concordance model's incompleteness is one potential reason.[112],[113] Resolving these facts encourages changes to some of the fundamental principles that form the basis of cosmology.[114] Investigating current Hubble tensions may be made more credible by altering the conventional redshift relation.[115]

According to several other hypotheses, cosmic acceleration is an emergent phenomenon.[116] Cosmological redshift is the primary impact of cosmic evolution on photon propagation. Cosmological redshift is a theoretical function of the scale factor obtained from the Friedmann metric in the standard model. Instead of depending on its theoretical form, researchers are increasingly using data to rebuild this scale factor–redshift association.[117],[118] The unknown function of the measured redshift, which raises the equation's degree of freedom, is a disadvantage of remapping cosmological models. Before using a parametric method, this problem has been resolved by adding function parameterization using Taylor expansion. A cosmological model that modifies the conventional redshift relation to explain the universe's accelerating expansion is one example of related study.[119] Under the premise that the Friedmann model is insufficient, it has been shown that combining Friedmann equations with a version of redshift remapping may result in a self-consistent framework.[120],[121] The parametric,[122] non-parametric,[123] and modified standard redshift models,[124] are expected to address the cosmological constant problem.

All of these lofty goals, meanwhile, are dependent on a necessary prerequisite and a wealth of precise and comprehensive cosmological data. Notwithstanding the increasing amount of observable evidence, enduring constraints need a cautious evaluation of the correctness and completeness of the present cosmological models.[125],[126],[127] Equipped with a 3.6 gigapixel camera, the future Vera Rubin Observatory might conduct a groundbreaking ten-year expedition,[128] prepared to examine the whole night sky and explore the complexities

---

[108] Wamalwa, D.S.; Omolo, J.A. Generalized relativistic dynamics in a non-inertial reference frame. Indian J. Phys. 2010, 84, 1241–1255, accessed on 19 Jan 2026

[109] Mandal, S.; Pradhan, S.; Sahoo, P.K.; Harko, T.; Mandal, S. Cosmological observational constraints on the power law f(Q) type modified gravity theory. Eur. Phys. J. C 2023, 83, 1–18, accessed on 19 Jan 2026

[110] Wang, B.; Abdalla, E.; Atrio-Barandela, F.; Pavón, D. Further understanding the interaction between dark energy and dark matter: Current status and future directions. arXiv 2024, arXiv:2402.00819v1, accessed on 19 Jan 2026

[111] Riess, A.G.; Yuan, W.; Macri, L.M.; Scolnic, D.; Brout, D.; Casertano, S.; Jones, D.O.; Murakami, Y.; Anand, G.S.; Breuval, L.; et al. A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SHOES Team. Astrophys. J. 2022, 934, L7, accessed on 13 Nov 2025

[112] https://ui.adsabs.harvard.edu/abs/2005coex.conf.555S/abstract, accessed on 19 Jan 2026

[113] Chan, M.H. The cosmological ultra-low frequency radio background: A solution to the Hubble tension and the 21-cm excess trough. Eur. Phys. J. C 2023, 83, 509, accessed on 19 Jan 2026

[114] Aluri, P.K.; Cea, P.; Chingangbam, P.; Chu, M.; Clowes, R.G.; Hutsemékers, D.; Kochappan, J.P.; Lopez, A.M.; Liu, L.; Martens, N.M.; et al. Is the Observable Universe Consistent with the Cosmological Principle? arXiv 2023, arXiv:2207.05765v4, accessed on 19 Jan 2026

[115] Seshavatharam, U.V.S.; Lakshminarayana, S. A Rotating Model of a Light Speed Expanding Hubble-Hawking Universe. Phys. Sci. Forum 2023, 7, 43, accessed on 13 Nov 2025

[116] Bassett, B.A.; Fantaye, Y.; Hložek, R.; Sabiu, C.; Smith, M. A Tale of two redshift. arXiv 2015, arXiv:1312.2593v2, accessed on 13 Nov 2025

[117] Wojtak, R.; Prada, F. Testing the mapping between redshift and cosmic scale factor. Mon. Not. R. Astron. Soc. 2016, 458, 3331–3340, accessed on 19 Jan 2026

[118] Wojtak, R.; Prada, F. Redshift remapping and cosmic acceleration in dark-matter-dominated cosmological models. Mon. Not. R. Astron. Soc. 2017, 470, 4493–4511, accessed on 13 Nov 2025

[119] Tian, S. The relation between cosmological redshift and scale factor for photons. Astrophys. J. 2017, 846, 1538–4357, accessed on 19 Jan 2026

[120] Wojtak, R.; Prada, F. Redshift remapping and cosmic acceleration in dark-matter-dominated cosmological models. Mon. Not. R. Astron. Soc. 2017, 470, 4493–4511, accessed on 13 Nov 2025

[121] Green, S.; Wald, R. How well is our universe described by an FLRW model? Class. Quantum Gravity 2014, 31, 1–31, accessed on 19 Jan 2026

[122] Bassett, B.A.; Fantaye, Y.; Hložek, R.; Sabiu, C.; Smith, M. A Tale of two redshift. arXiv 2015, arXiv:1312.2593v2, accessed on 13 Nov 2025

[123] Wojtak, R.; Prada, F. Redshift remapping and cosmic acceleration in dark-matter-dominated cosmological models. Mon. Not. R. Astron. Soc. 2017, 470, 4493–4511, accessed on 19 Jan 2026

[124] Tian, S. The relation between cosmological redshift and scale factor for photons. Astrophys. J. 2017, 846, 1538–4357, accessed on 19 Jan 2026

[125] Riess, A.G.; Yuan, W.; Macri, L.M.; Scolnic, D.; Brout, D.; Casertano, S.; Jones, D.O.; Murakami, Y.; Anand, G.S.; Breuval, L.; et al. A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SHOES Team. Astrophys. J. 2022, 934, L7, accessed on 13 Nov 2025

[126] Pâris, I.; Petitjean, P.; Ross, N.P.; Myers, A.D.; Aubourg, É.; Streblyanska, A.; Bailey, S.; Armengaud, É.; Palanque-Delabrouille, N.; Yèche, C.; et al. The Sloan Digital Sky Survey Quasar Catalog: Twelfth data release. Astrophys. J. 2017, 597, A79,

[127] Di Valentino, E. Cosmological tensions: Hints for a new concordance model? In Proceedings of the The Sixteenth Marcel Grossmann Meeting, Online, 5–10 July 2021, accessed on 30 Jan 2026

[128] Riess, A.G.; Casertano, S.; Yuan, W.J.; Bowers, B.; Macri, L.; Zinn, J.C.; Scolnic, D. Cosmic Distances Calibrated to 1%

of the cosmos.[129] Again, in 2015 the parametric model proposed by Bassett et al.[130] introduces modifications to the traditional redshift paradigm, seeking to refine our understanding of cosmic dynamics. By adding parameters that account for changes in the redshift space, this model enables a more sophisticated interpretation of observational data. By adding certain characteristics, the model takes into account intricate features of cosmic processes and offers a more thorough and precise depiction of findings connected to redshift. However, the non-parametric model, developed by Wojtak and Prada in 2017, adopts a different strategy by eschewing predetermined parameters, enabling more freedom in simulating cosmic occurrences.[131] The non-parametric model allows for a more flexible and data-driven examination of redshift-related phenomena since it does not impose preset parameters as parametric models do. When the underlying dynamics are complicated and difficult to capture by preset parameters, this approach is invaluable.

It offers a more flexible instrument for analyzing data from observations. The local cosmos has a diverse range of developed star systems (Norris et al. 2014),[132] from globular clusters (Brodie & Strader 2006; Kruijssen 2014; Renzini et al. 2015)[133],[134],[135] to compact elliptical galaxies (Faber 1973)[136], ultrafaint dwarfs (e.g., Simon & Geha 2007)[137], and ultra-diffuse spheroids (van Dokkum et al. 2017)[138], each of which presumably has its own characteristic formation pathway. In addition to their ancient antiquity, several of these systems have tremendous star concentrations (Forbes and Bridges 2010)[139] suggest that at $z \gtrsim 1.5$, when the universe's gas concentrations were generally much greater, the bulk of their star formation took place. Getting a sensitive, high-resolution image of the distant cosmos is one potentially viable avenue to explore the development of these local systems. Fortunately, such discoveries may be made by combining the magnifying effect of gravitational lensing with the capability of long exposures with the Hubble Space Telescope, as recently done in the ambitious Hubble Frontier Fields (HFF) mission (Coe et al. 2015; Lotz et al. 2017)[140],[141]. Since many of the progenitors of local systems are probably required, these sensitive observations enable us to probe to extremely low luminosities. Numerous galaxies may be analyzed at extremely high spatial resolution because of the strong lensing magnifications from huge galaxy clusters, which stretch them by significant factors. As it has discussed in Bouwens et al. (2021b) [142], this stretching can reliably be estimated up to linear magnifications of ∼30× (or total magnification factors of ∼50×). It is also available in Bouwens et al. 2017a, c[143],[144], where similar though smaller limits were presented with the then-current models, and Meneghetti et al. 2017)[145]. Given the small inferred sizes of the fainter lensed sources identified by Kawamata et al. (2018)[146] and Bouwens et al. (2021b).[147] It is intriguing to consider these sources in relation to other tiny star-forming systems, such as star clusters or cluster complexes, as well as other stellar systems that they may develop into today. An initial look at such comparisons was already executed in an earlier unpublished study by our group (Bouwens et al. 2017b)[148] and also by Kikuchihara et al. (2020)[149]. One significant early conclusion from these investigations was that, in size/mass space, lensed z = 6–8 galaxies seem to fall between ultracompact dwarfs/globular clusters and compact elliptical galaxies, with masses and diameters that fall between 107 and 108 M⨀ and around 50–500 pc. Over the last ten years, significant advancements have been made in the discovery of galaxies that formed within a billion years of the Big Bang, at z > 6. It is fantastic, mostly because

Precision with Gaia EDR3 Parallaxes and Hubble Space Telescope Photometry of 75 Milky Way Cepheids Confirm Tension with ΛCDM. Astrophys. J. Lett. 2021, 908, L6, accessed on 13 Nov 2025

[129] Brough, S.; Collins, C.; Demarco, R.; Ferguson, H.C.; Galaz, G.; Holwerda, B.; Martinez-Lombilla, C.; Mihos, C.; Montes, M. The Vera Rubin Observatory Legacy Survey of Space and Time and the Low Surface Brightness Universe. arXiv 2020, arXiv:2001.11067v1, accessed on 30 Jan 2026

[130] Bassett, B.A.; Fantaye, Y.; Hložek, R.; Sabiu, C.; Smith, M. A Tale of two redshift. arXiv 2015, arXiv:1312.2593v2, accessed on 30 Jan 2026

[131] Wojtak, R.; Prada, F. Redshift remapping and cosmic acceleration in dark-matter-dominated cosmological models. Mon. Not. R. Astron. Soc. 2017, 470, 4493–4511, accessed on 13 Nov 2025

[132] Norris M., Kannappan S., Forbes D. A. et al. 2014 MNRAS arXiv:1406.6065v1, accessed on 13 Nov 2025

[133] Brodie J. P., Romanowsky A. J., Strader J. and Forbes D. A. 2011 AJ 142 199, accessed on 13 Nov 2025

[134] Kruijssen J. M. D. 2014 CQGra 31 244006, accessed on 30 Jan 2026

[135] Renzini A., D'Antona F., Cassisi S. et al. 2015 MNRAS 454 4197, accessed on 13 Nov 2025

[136] Faber S. M. 1973 ApJ 179 423, accessed on 30 Jan 2026

[137] Simon J. D. and Geha M. 2007 ApJ 670 313, accessed on 30 Jan 2026

[138] Dessauges-Zavadsky M., Schaerer D., Cava A., Mayer L. and Tamburello V. 2017 ApJL 836 L22, accessed on 13 Nov 2025

[139] Forbes D. A. and Bridges T. 2010 MNRAS 404 1203, accessed on 30 Jan 2026

[140] Coe D., Bradley L. and Zitrin A. 2015 ApJ 800 84, accessed on 30 Jan 2026

[141] Lotz J. M., Koekemoer A., Coe D. et al. 2017 ApJ 837 97, accessed on 13 Nov 2025

[142] Bouwens R. J., Oesch P. A., Stefanon M. et al. 2021b AJ 162 47, accessed on 30 Jan 2026

[143] Bouwens R. J., Illingworth G. D., Oesch P. A. et al. 2017a ApJ 843 41, accessed on 13 Nov 2025

[144] Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S. and Stefanon M. 2017c ApJ 843 129, accessed on 13 Nov 2025

[145] Meneghetti M., Natarajan P., Coe D. et al. 2017 MNRAS 472 3177, accessed on 30 Jan 2026

[146] Kawamata R., Ishigaki M., Shimasaku K. et al. 2018 ApJ 855 4, accessed on 30 Jan 2026

[147] Bouwens R. J., Oesch P. A., Stefanon M. et al. 2021b AJ 162 47, accessed on 30 Jan 2026

[148] Bouwens R. J., Illingworth G. D., Oesch P. A. et al. 2017b arXiv:1711.02090, accessed on 30 Jan 2026

[149] Kikuchihara S., Ouchi M., Ono Y. et al. 2020 ApJ 893 60, accessed on 30 Jan 2026

to confirmation spectra from huge ground-based observatories and photos from the Hubble Space Telescope. The James Webb Space Telescope (JWST) and ground-based Extremely Large Telescopes (ELTs) will push the "high redshift frontier" even further during the next years.[150] The Nancy Grace Roman Space Telescope, also known as the Roman Space Telescope, Roman, or RST, is an infrared space telescope being developed by NASA that is expected to launch into a Sun-Earth L2 orbit by May 2027.[151] The narrow field of view of these facilities, particularly JWST, and the Roman Space Telescope's (previously WFIRST) sensitivity confined to the near-infrared (near-IR, $\lambda < 2\mu m$)[152] and EUCLID wide-field imaging space missions, and which mean that a crucial piece of the jigsaw remains missing. That is a wide-field imaging survey,[153] Finding the rarest most massive and bright galaxies at the greatest redshifts—whose progenitors are probably the earliest galactic structures to form—requires research at near and mid-IR wavelengths (necessarily from space). For galaxies at $z > 10$, NIR spectroscopy at $\lambda > 2\mu m$ (corresponding to the rest-frame optical frame) is also required to get comprehensive information (metallicity, star mass).[154]

With the completion of the JWST mission, which is projected to uncover a multitude of faint galaxies at high redshift and address the function of these early galaxies in the reionization of the inter-galactic medium, the astrophysics landscape is anticipated to be very rich between 2035 and 2050. The most powerful radio telescope on the planet right now is the Millimeter/submillimeter Array (ALMA). An international intergovernmental radio telescope project called the Square Kilometre Array (SKA) is being constructed in South Africa (mid-frequency) and Australia (low-frequency). By then, SKA will have investigated the molecular emission and dust re-emission from several of these objects, and ALMA will be a well-developed facility.[155] It was low luminosity sources that caused the Universe to reionize.[156],[157],[158] Only in groups or proto clusters will these low brightness sources be detectable. Given that the early galaxies were quite faint, this is probably the case. Finding proto clusters between z~6 and z~15 would thus reveal the re-ionization history of the Universe.[159] At wavelengths less than 2 microns, the Rubin Observatory (formerly LSST) on Earth and EUCLID and the Roman Space Telescope in space will investigate rare and brilliant sources at high redshift as well as transients like distant supernovae.[160] After many decades, new facilities like Athena will enable AGN to reach previously unheard-of distances in the X-ray. However, a wide-field IR survey mission with spectroscopy and imaging operating beyond 2 microns has to be addressed in the future, since there is a significant gap in the parameter space that is yet unexplored.[161] This incomplete voyage into the enigmatic cosmos will continue. However, in order to inform the formation and origin of the universe based on the current development of the notion of physics, this analytical study effort will assess and explain the chronological evolution of concepts and theories of physics. The author will also make an attempt to research, describe, and evaluate the universe's genesis and future.

3. Progress of Modern Theory of Science

The foundation of all natural sciences is physics. Our knowledge of chemistry, biology, astronomy, and geology is based on its principles. Using universal rules to describe a wide range of events is one of physics' greatest achievements. From commonplace items like lightbulbs, cell phones, and can openers to biological systems like muscles, lungs, and brains; from artistic endeavors like paints, piccolos, and pirouettes to technological innovations like cameras, cars, and cathedrals; from natural disasters like hurricanes, tsunamis, and earthquakes to subatomic particles like quarks, DNA, and black holes, it aids in our understanding of how the world functions. By showing relationships between apparently unconnected occurrences, physics helps us make sense of the cosmos. It offers strong resources to stimulate innovation, broaden our horizons, and cultivate creativity. Physics serves as the basis for the majority of contemporary technology, including the equipment and techniques utilized in scientific, technical, and medical research and development. Technology based on physics is crucial to manufacturing.[162] Energy solutions (nuclear, solar power), medical imaging (MRI, X-ray, ultrasonogram, etc.), and quantum computing are all products of applied physics and are entwined with the future of human existence. Modelling the Earth's climate system requires physics because it

[150] Carroll, S. M., & Ostlie, D. A. (2017). An Introduction to Modern Astrophysics. Cambridge University Press, accessed on 17 Nov 2025

[151] https://www.nasa.gov/missions/roman-space-telescope/nasa-tool-gets-ready-to-image-faraway-planets/, accessed on 17 Nov 2025

[152] https://roman.gsfc.nasa.gov/, accessed on 30 Jan 2026

[153] https://www.nasa.gov/missions/euclid/new-images-from-euclid-mission-reveal-wide-view-of-the-dark-universe/, accessed on 17 Nov 2025

[154] Santos MG, Silva MB, Pritchard JR, Cen R, Cooray A (2011) Probing the first galaxies with the Square Kilometer Array. AAP 527: A93. https://doi.org/10.1051/0004-6361/201015695, accessed on 17 Nov 2025

[155] https://www.almaobservatory.org/en/home/, accessed on 30 Jan 2026

[156] David Tong. ""Lecture 2: The Hot Universe"". Lectures on Cosmology. University of Cambridge, accessed on 17 Nov 2025

[157] Wise, John H. (2019). "Cosmic reionisation". Contemporary Physics. 60 (2): 145-163, doi:10.1080/00107514.2019.1631548, accessed on 30 Jan 2026

[158] Furlanetto, Steven R.; Oh, S. Peng (July 2008). "The History and Morphology of Helium Reionization". The Astrophysical Journal. 681 (1): 1–17, doi:10.1086/588546. ISSN 0004-637X, accessed on 30 Jan 2026

[159] https://iopscience.iop.org/article/10.1088/0004-637X/750/2/137, accessed on 30 Jan 2026

[160] LIGO Scientific Collaboration, & Virgo Collaboration. (2016). Observation of Gravitational Waves from a BinaryBlack Hole Merger. Physical Review Letters, 116(6), 061102, accessed on 30 Jan 2026

[161] https://link.springer.com/article/10.1007/s10686-021-09703-1, accessed on 30 Jan 2026

[162] http://phystec.physics.cornell.edu/content/why-study-physics, accessed on 30 Jan 2026

enables scientists to forecast changes in atmospheric conditions, sea levels, and weather patterns—all of which are vital for the survival and adaption of life on Earth in the future.

The Copernican mathematician, astronomer, and physicist Galileo Galilei (1564–1642) was an Italian who made several contributions. He modified the telescope, carried out empirical research, and found new celestial phenomena. At the period, the primary disciplines of law, medicine, and theology—all of which were intimately related to philosophy—overshadowed his function as a mathematician at the university. Galileo discovered that bodies do not fall with weight-proportional velocities. Even though the legend of him throwing weights from the Leaning Tower of Pisa is probably untrue, he did prove that a projectile's trajectory is a parabola and discovered inertia, which foreshadowed Newton's laws of motion.[163] Among them is what is now called Galilean relativity, which was the first precise explanation of space and time's characteristics outside of three-dimensional geometry. Galileo is sometimes considered to be the "father of modern observational" approaches in astronomy,[164] the "father of modern physics", the "father of science",[165]and "the father of modern science". "Galileo, perhaps more than any other single person, was responsible for the birth of modern science," according to Stephen Hawking.[166] Galileo's support for heliocentrism caused controversy since religious orthodoxy maintained a geocentric or Tychonic understanding of the Solar system, which resulted in his prosecution by the Inquisition. His discovery of the Jovian moons, which he published in 1610, helped him get a position at the Medici court as a mathematician and philosopher.[167] In addition to drawing a sizable readership for his own writings, such as the Discourses and Mathematical Demonstrations Concerning Two New Sciences, he was expected to engage in discussions with Aristotelian thinkers. Following his detention, his writings, including the Dialogue Concerning the Two Chief World Systems, were published all over the world.[168,169] Galileo's emphasis on conducting experiments and creating mathematical explanations of motion made experimentation a cornerstone of natural science philosophy.[170]

The British mathematician and physicist Sir Isaac Newton (1642–1727), a member of the Royal Society of England, developed a single framework for explaining how the cosmos functions. In 1672, he was made a Fellow of the Royal Society of England. Later, from 1703 until his death in 1727, he served as President of the Society.[171] The link between objects and their motion is described by Newton's three laws of motion. Additionally, he developed the law of universal gravitation, which explains how planets and other heavenly objects move as well as how things fall on Earth.[172] In order to accomplish his goals, Newton created calculus, a new area of mathematics that Gottfried Leibniz also independently created. This branch turned into a crucial instrument for later developments in several branches of physics. In the *Philosophiae Naturalis Principia Mathematica*, also known as the *Mathematical Principles of Natural Philosophy*, Newton described his findings in full. Its publication in 1687 signaled the beginning of the modern age in astronomy and mechanics.[173] The Cartesian mechanical theory, which holds that all movements are caused by instantaneous forces from corpuscles, was contested by Newton. He replaced the notion that things follow preset natural courses with the knowledge that all future movements can be mathematically predicted using present motion, mass, and the forces at work by using his three laws of motion and the rule of universal gravitation.[174] Nevertheless, the observed celestial movements did not quite fit a Newtonian viewpoint, and Newton—who was also deeply interested in theology—thought that God's involvement was required to preserve the stability of the solar system.[175]

Continental thinkers disagreed with Newton's ideas, with the exception of his mathematical techniques, and deemed his lack of philosophical justifications for motion and gravity intolerable.[176] The Continental and British intellectual traditions began to

[163] Thornton, Stephen T.; Marion, Jerry B. (2004). Classical Dynamics of Particles and Systems (5th ed.). Brooke Cole. p. 49. ISBN 0-534-40896-6, accessed on 30 Jan 2026

[164] Singer, Charles (1941), A Short History of Science to the Nineteenth Century, Clarendon Press, page 217,

[165] Weidhorn, Manfred (2005), The Person of the Millennium: The Unique Impact of Galileo on World History, iUniverse, p. 155, ISBN 0-595-36877-8, accessed on 30 Jan 2026

[166] https://www.inventionandtech.com/content/galileo-and-birth-modern-science, accessed on 26 Nov 2025

[167] https://www.newyorker.com/magazine/2013/12/16/a-very-rare-book, accessed on 26 Nov 2025

[168] Drake, Stillman (1978), Galileo at Work: His Scientific Biography, Chicago, Illinois: University of Chicago Press, ISBN 0-226-16226-5, accessed on 30 Jan 2026

[169] Biagioli, Mario (1993), Galileo, Courtier: The Practice of Science in the Culture of Absolutism, Chicago, Illinois: University of Chicago Press, ISBN 0-226-04559-5, accessed on 26 Nov 2025

[170] Brading, Katherine (August 2019). "A note on rods and clocks in Newton's Principia". Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics. 67: 160–166, doi:10.1016/j.shpsb.2017.07.004, accessed on 30 Jan 2026

[171] Matthews, Michael R. (2000). Time for Science Education: How Teaching the History and Philosophy of Pendulum Motion Can Contribute to Science Literacy. New York: Springer Science+Business Media, LLC. p. 181. ISBN 978-0-306-45880-4, accessed on 30 Jan 2026

[172] Cheng, K. C.; Fujii, T. (1998). "Isaac Newton and Heat Transfer". Heat Transfer Engineering. 19 (4): 9–21. doi:10.1080/01457639808939932. ISSN 0145-7632, accessed on 30 Jan 2026

[173] Whiteside, D.T., ed. (1970). "The Mathematical principles underlying Newton's Principia Mathematica". Journal for the History of Astronomy. 1. Cambridge University Press. pp. 116–138, accessed on 30 Jan 2026

[174] Rynasiewicz, Robert A. (22 August 2011). "Newton's Views on Space, Time, and Motion". Stanford Encyclopedia of Philosophy. Stanford University, accessed on 30 Jan 2026

[175] Bloye, Nicole; Huggett, Stephen (2011). "Newton, the geometer" (PDF). Newsletter of the European Mathematical Society (82): 19–27. MR 2896438, accessed on 30 Jan 2026

[176] Goldstine, Herman H. (1980). A History of the Calculus of Variations from the 17th Through the 19th Century.

sharply diverge about 1700. Intense, continuous, and often personal disagreements between Newtonian and Leibnizist adherents about who should have precedence in the development of calculus's analytical methods—which each had separately invented—were the main source of this division.[177] Leibnizian calculus notation was widely used everywhere except Britain at first because the Cartesian and Leibnizian traditions predominated on the European continent. Although he made it clear in his works that a philosophical explanation of gravity was not necessary to deduce its effects in reality, Newton himself was uncomfortable with the lack of one.[178] Throughout the 18th century, continental natural philosophers increasingly adopted the Newtonians' strategy of eschewing ontological metaphysical explanations in favor of mathematically specified movements.[179],[180]

Newton developed a theory of color, which he described in *Opticks*, and constructed the first functional reflecting telescope. The fact that a prism divides white light into the several colors that are visible to the human eye spectrum served as the foundation for his hypothesis.[181] In 1690, Christiaan Huygens put up a rival wave theory to explain the behavior of light, in contrast to Newton's description of light as consisting of microscopic particles. However, the wave hypothesis did not acquire much momentum until the 19th century because of the great reputation of Newton and the general trust in mechanical philosophy. In addition, Newton proved the generalized binomial theorem, investigated power series, created an empirical rule of cooling, and developed a technique for calculating the roots of an equation function.[182] Simon Stevin's work on decimals served as inspiration for his studies on infinite series.[183] Most significantly, Newton showed that the same natural laws—which are neither hazardous nor unpredictable—control the movements of things on Earth and in space. Newton removed any residual skepticism about heliocentrism by demonstrating how his theory of gravity and Kepler's rules of planetary motion aligned.[184] Newton successfully established the foundation for contemporary mathematics and science by combining the concepts put forward throughout the Scientific Revolution.[185]

Springer New York. pp. 7–21. ISBN 978-1-4613-8106-8, accessed on 30 Jan 2026

177 Newton, I. (1999). The Principia, The Mathematical Principles of Natural Philosophy. Translated by Cohen, I.B.; Whitman, A. Los Angeles: University of California Press, accessed on 29 Nov 2025

178 Guicciardini, Niccolò (1999), Reading the Principia: The Debate on Newton's Methods for Natural Philosophy from 1687 to 1736, New York: Cambridge University Press, accessed on 03 Feb 2026

179 Hall, A. Rupert (1980), Philosophers at War: The Quarrel between Newton and Leibniz, New York: Cambridge University Press. accessed on 03 Feb 2026

180 Bertolini Meli, Domenico (1993), Equivalence and Priority: Newton versus Leibniz, New York: Oxford University Press, accessed on 29 Nov 2025

181 Wilson, Raymond N. (2013). "1.1 Period 1608–1672". Reflecting Telescope Optics I: Basic Design Theory and its Historical Development. Springer. pp. 1–10. ISBN 978-3-662-03227-5, accessed on 03 Feb 2026

182 Barth, Nils R. (2004). "Computing Cavalieri's Quadrature Formula by a Symmetry of the n-Cube". The American Mathematical Monthly. 111 (9): 811–813. doi:10.2307/4145193, accessed on 29 Nov 2025

183 Błaszczyk, Piotr; Katz, Mikhail; Sherry, David (2012), "Ten misconceptions from the history of analysis and their debunking", Foundations of Science, 18: 43–74, doi:10.1007/s10699-012-9285-8, accessed on 03 Feb 2026

184 "Orbits and Kepler's Laws". NASA Solar System Exploration. 26 June 2008. Archived from the original on 2022-12-13, accessed on 03 Feb 2026

185 Palomo, Miguel (2 January 2021). "New insight into the origins of the calculus war". Annals of Science. 78 (1): 22–40. doi:10.1080/00033790.2020.1794038, accessed on 03 Feb 2026

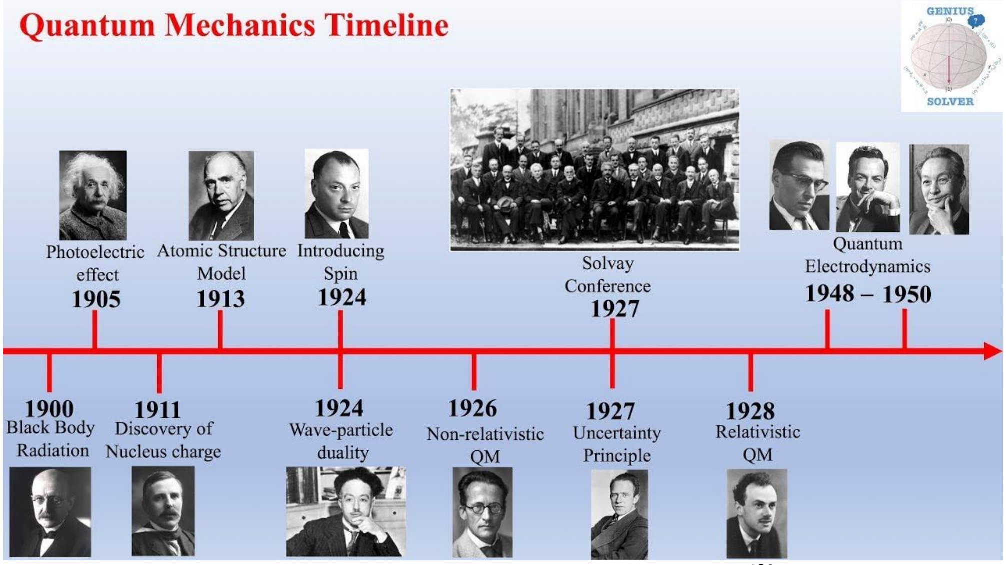


Figure 1: *A brief history of quantum mechanics*[186]

After relocating to Bern, Switzerland, in 1905, a 26-year-old German physicist named Albert Einstein proved how motion between the observer and the object of observation impacts time and space measurements.[187] Significant developments in quantum theory were also made by Einstein. E=mc2, his mass-energy equivalency formula derived from special relativity, has been named "the world's most famous equation."[188] His contributions to theoretical physics are among the best intellectual works ever created. Even though he did not develop the theory of relativity, Einstein recognized that there is an absolute limit to speed and that the speed of light in a vacuum is constant, or the same for all observers. This has less impact on a person's day-to-day activities since most objects travel at far slower speeds than light. However, the theory of relativity states that an object's length will decrease as measured by an observer on Earth as it approaches the speed of light because its connected clocks will operate more slowly. Einstein also developed the equation E=mc2, which asserts the equality of mass and energy.[189] Einstein argued that electromagnetic principles should be universally applicable, independent of reference frame, and that the speed of light is constant across all inertial reference frames. In 1921, he was awarded the Nobel Prize in Physics for his research on the photoelectric effect. Science was drastically altered by Einstein's groundbreaking theory of relativity. His theory of relativity is still his most significant contribution, despite the fact that he made several other significant contributions.

[186] https://www.youtube.com/watch?v=EXq7P2OzgbE, accessed on 02 Jan 2026
[187] Whittaker, E. (1 November 1955). "Albert Einstein. 1879–1955". Biographical Memoirs of Fellows of the Royal Society. 1: 37–67. doi:10.1098/rsbm.1955.0005, accessed on 03 Feb 2026
[188] Yang, Fujia; Hamilton, Joseph H. (2010). Modern Atomic and Nuclear Physics. World Scientific. p. 274. ISBN 978-981-4277-16-7, accessed on 03 Feb 2026
[189] Miller, Arthur I. (1981), Albert Einstein's special theory of relativity. Emergence (1905) and early interpretation (1905–1911), Reading: Addison–Wesley, ISBN 978-0-201-04679-3, accessed on 03 Feb 2026

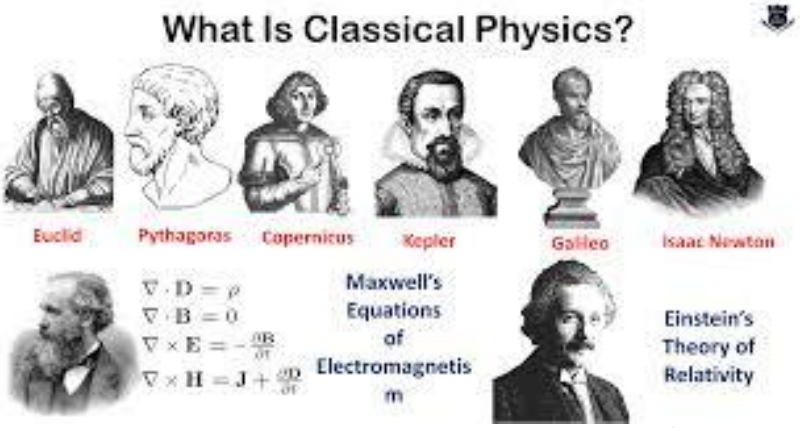


Figure 2: *Journey of classical to modern physics* [190]

Albert Einstein in 1905 develop "special theory of relativity" and that made the ether "superfluous" to physical theory and maintained that measurements of length and time changed according on the observer's motion with regard to the object being observed.[191] Furthermore, the equation suggested that energy and mass were equal quantities. In order to provide an accurate explanation for the dispersion of blackbody radiation, theoretical physicist Max Planck suggested a constant in 1900.[192] In another publication the same year, Einstein asserted that electromagnetic radiation was transmitted in definite amounts, or quanta. The peculiarities of the photoelectric effect were explained by this premise. [193] The special theory of relativity explains the connection between the concepts of space and time and physical phenomena. The idea, which had a big impact on both domains, was born out of the contradictions between electromagnetic and Newtonian physics. The first historical concern was whether it was pertinent to discuss motion in terms of the "ether" that carries electromagnetic waves and if it was feasible to detect such motion, as the Michelson–Morley experiment unsuccessfully sought to do. Both of these questions and the concept of ether were refuted by Einstein in his special theory of relativity. However, his basic formulation does not incorporate entire electromagnetic theory. "What is time?" is the first question it poses. "Absolute, true, and mathematical time, by itself and from its own nature, runs equably without connection to anything external," was Newton's straightforward and mathematical explanation in the Principia (1686).[194] This definition is basic to all classical physics.

When the bright Einstein challenged it, he found it to be deficient. Instead, two observers travelling in relation to each other will have separate time scales, and each "observer" must use their own time scale. This has a comparable effect on position measurements. Space and time become concepts that are fundamentally dependent on the observer. Each observer controls their own space-time framework or coordinate system. Since there is no absolute frame of reference, each witness of a given event provides a unique measurement that is yet equally legitimate (and reconcilable). According to Einstein's relativity postulate, "the basic rules of physics are same for two observers who have a constant relative velocity with regard to each other," which establishes what is absolute. The discovery of "Poincaré symmetry," a new symmetry law of nature that replaced "Galilean symmetry," by special relativity had a profound effect on physics. It started off as a revision of the electromagnetic theory.'[195] Another long-lasting impact on dynamics was special relativity.[196] It quickly became evident that relativistic dynamics distinguishes between rest mass, an invariant (observer-independent) feature of a particle or system of particles, and the energy and momentum of a system, despite its original credit for combining mass and energy. [197] The latter two are not invariant between observers, although they are preserved independently in every situation. The definition of "mass" in particle physics has changed over time; since the late 20th century, it almost always refers to rest mass. Therefore, the part of an object's or system's total mass that is unaffected by the system's overall motion is known as invariant mass, rest mass, intrinsic mass, proper mass, or simply mass in the context of bound systems. In particular, it is a characteristic of the

[190] https://www.instagram.com/p/DUGN-M0ifcd/, accessed on 03 Feb 2026

[191] Mermin, N. David (2009). It's about time: understanding Einstein's relativity (Sixth printing, and first paperback printing ed.). Princeton, NJ and Oxford: Princeton University Press. ISBN 978-0-691-14127-5, accessed on 03 Dec 2025

[192] Peres, Asher; Terno, Daniel R. (6 January 2004). "Quantum information and relativity theory". Reviews of Modern Physics. 76 (1): 93–123, doi:10.1103/RevModPhys.76.93, accessed on 03 Feb 2026

[193] Goody, R. M.; Yung, Y. L. (1989). Atmospheric Radiation: Theoretical Basis (2nd ed.). Oxford University Press. ISBN 978-0-19-510291-8, accessed on 03 Dec 2025

[194] https://www.britannica.com/topic/Principica, accessed on 03 Feb 2026

[195] https://www.feynmanlectures.caltech.edu/I_52.html, accessed on 03 Dec 2025

[196] Greenwood, D.T. (1997). Classical Dynamics. Dover books on mathematics. Dover Publications. ISBN 978-0-486-69690-4, accessed on 03 Dec 2025

[197] Flego, Silvana; Plastino, Angelo; Plastino, Angel Ricardo (2011-12-20). "Information Theory Consequences of the Scale-Invariance of Schröedinger's Equation". Entropy. 13 (12). MDPI AG: 2049–2058, doi:10.3390/e13122049, accessed on 03 Feb 2026

system's total momentum and energy that holds true for all frames of reference connected by Lorentz transformations.[198] The invariant mass is equal to the entire mass in the "rest frame" if the system has a center-of-momentum frame. In the end, the system's total or relativistic mass is more than the invariant mass, but the invariant mass doesn't change.[199]

After expanding this to include all states of motion, including non-uniform acceleration, Einstein created the general theory of relativity in 1916. The gravitational force at every point in space was explained by the curvature of space-time, a new concept that Einstein presented in this theory. The curvature of space-time replaced Newton's universal rule of gravity. Einstein claimed that the manifestation of gravitational force in the traditional sense is caused by the geometry of space. The curvature of space-time caused by the existence of a mass determines the space-time path that all freely moving objects take.[200] Additionally, experiments supported this theory's prediction that gravity would have an impact on light. This aspect of relativity explained the characteristics of the cosmic microwave background radiation, predicted black holes, and explained the phenomena of light bending around the sun. This discovery led to serious problems with the conventional Steady-State paradigm.[201] For his contributions to blackbody radiation, relativity, and the photoelectric effect, Einstein received the Nobel Prize in 1921. The ultimate acceptance of Niels Bohr's model of the atom, Einstein's theories of relativity, and the quantized nature of light transmission led to a thorough endeavor to rebuild physics on new fundamental principles, which created as many problems as they addressed.[202] Extending relativity to accelerating reference frames in the 1910s gave rise to the "general theory of relativity."[203] In order to conclude that space is curved and limited in size, Einstein proposed an equivalence between the inertial force of acceleration and the force of gravity. This led to the prediction of phenomena like gravitational lensing and the distortion of time in gravitational fields.[204] Therefore, anything that bends light from a distant source as it moves towards an observer—such as a point particle or a cluster of galaxies—is known as a gravitational lens. Albert Einstein's general theory of relativity describes the magnitude of gravitational lensing.[205]

A second theoretical issue was how to explain the distribution of electromagnetic radiation emitted by a black body, even though relativity resolved the conflict between electromagnetic phenomena as shown by Michelson and Morley. Experiments revealed that at shorter wavelengths, towards the ultraviolet end of the spectrum, the energy approached zero, but classical theory predicted it should become infinite. The new theory of quantum mechanics resolved this stark disparity, sometimes referred to as the ultraviolet disaster.[206] The theory of atoms and subatomic systems is known as quantum mechanics. The idea was conceived and developed over the first thirty years of the twentieth century. By arguing that energy is released in discrete packets known as "quanta" rather than continually, Max Planck (1858–1947) presented the fundamental concepts of quantum theory in 1900.[207] This groundbreaking research on black-body radiation established the groundwork for contemporary quantum mechanics and resolved a significant physics issue. For his discovery, he received the 1918 Nobel Prize in Physics. Planck's constant h, a basic constant that connects a quantum's energy to its frequency ($E=h\nu$), was first presented by his work.[208] When Louis de Broglie claimed that matter may be seen as acting like a wave in a manner similar to that of electromagnetic waves (wave–particle duality) and the Compton Effect demonstrated that light has velocity and can scatter off particles, the quantum theory gained acceptance.[209]

The photoelectric phenomenon was explained by Einstein in 1905 using the quantum theory, while the stability of Rutherford's atom and the light frequencies

[198] https://books.google.com.bd/books?id=Nv5GAyAdijoC&pg=PA1073&redir_esc=y#v=onepage&q&f=false,

[199] Roche, J (2005). "What is mass?" (PDF). European Journal of Physics. 26 (2): 225. Bibcode:2005EJPh...26.225R. doi:10.1088/0143-0807/26/2/002, accessed on 03 Feb 2026

[200] Galison, Peter (Winter 2000). "Einstein's Clocks: The Question of Time". Critical Inquiry. 26 (2): 355–389. doi:10.1086/448970. JSTOR 1344127. S2CID 1444844 66, accessed on 03 Feb 2026

[201] Fixsen, D. J. (2009). "The Temperature of the Cosmic Microwave Background". The Astrophysical Journal. 707 (2): 916–920, doi:10.1088/0004-637X/707/2/916, accessed on 03 Feb 2026

[202] Bohr, N. (December 1925). "Atomic Theory and Mechanics1". Nature. 116 (2927): 845–852. Bibcode:1925Natur. 116..845B. doi:10.1038/116845a0. ISSN 0028-0836, accessed on 03 Feb 2026

[203] Wald, Robert M. (1984), General Relativity, University of Chicago Press, ISBN 978-0-226-87033-5, accessed on 03 Dec 2025

[204] Drakeford, Jason; Corum, Jonathan; Overbye, Dennis (March 5, 2015). "Einstein's Telescope - video (02:32)". The New York Times, accessed on 03 Feb 2026

[205] Pekka Teerikorpi; Mauri Valtonen; K. Lehto; Harry Lehto; Gene Byrd; Arthur Chernin (2008). The Evolving Universe and the Origin of Life: The Search for Our Cosmic Roots (illustrated ed.). Springer Science & Business Media. p. 165. ISBN 978-0-387-09534-9, accessed on 03 Feb 2026

[206] Ehrenfest, P. (1911). "Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle?" [In which features of the light quantum hypothesis does thermal radiation play an essential role?]. Annalen der Physik. 341 (11): 91–118. Bibcode:1911AnP...341...91E. doi:10.1002/andp.19113 411106, accessed on 03 Dec 2025

[207] A. Whitaker. The New Quantum Age: From Bell's Theorem to Quantum Computation and Teleportation, Oxford University Press, 2011, ISBN 978-0-19-958913-5, accessed on 03 Feb 2026

[208] https://www.nist.gov/physics/explainers/planck-constant, accessed on 03 Dec 2025

[209] Wittke, J.P; Dicke, R.H (June 1, 1961) [1960]. "11". In Holladay, W.G. (ed.). Introduction to Quantum Mechanics (eBook). Vol. 16. Nashville, Tennessee: ADDISON WESLEY LONGMAN INC (published January 1, accessed on 03 Dec 20251978). p. 10. doi:10.1063/1.3057610. ISBN 978-0-201-01510-2, accessed on 03 Feb 2026

emitted by hydrogen gas were explained by Danish scientist Niels Bohr in 1913 using the same constant.[210] In the 1920s, full-scale quantum mechanics replaced the quantized theory of the atom. In 1925, Werner Heisenberg, Max Born, and Pascual Jordan developed new principles of "quantum" rather than "classical" mechanics in matrix form. These principles rejected the idea of causality and were based on the probabilistic link between discrete "states".[211] Heisenberg, Wolfgang Pauli, Paul Dirac, and Erwin Schrödinger developed quantum mechanics in great detail. In 1926, Schrödinger proposed an analogous theory based on waves.[212],[213] Once again, the "Copenhagen interpretation" and Heisenberg's 1927 "uncertainty principle," which states that it is impossible to measure position and momentum exactly and simultaneously [214] of quantum mechanics (named after Bohr's hometown) persisted in rejecting the idea of basic causality, despite the fact that detractors like Einstein would figuratively claim that "God does not play dice with the universe."[215] In order to study and explain atomic-level occurrences, the new quantum mechanics became an essential instrument. The research on photons conducted by Indian physicist Satyendra Nath Bose in the 1920 [216] Bose-Einstein statistics and the theory of the Bose-Einstein condensate are based on quantum mechanics.[217] Any particle in quantum mechanics may be either a boson (statistically Bose–Einstein) or a fermion (statistically Fermi–Dirac), according to the spin–statistics theorem.[218] All basic bosons, including the photon, which transmits electromagnetism, were eventually discovered to transfer forces. Particles "like electrons and nucleons" are called fermions, and they are often found in matter. Later, Fermi-Dirac statistics were used in many other fields, including as semiconductor design and astronomy.

Classical physics refers to the earlier sciences, while "modern physics" refers to the work based on quantum and relativity theories. This is because of the conceptual differences between the physics theories that were discussed in the 19th century and those that were most historically prominent in the first decades of the 20th century. The split, which was first used to describe "classical mechanics" in mechanics, ultimately came to describe quantum and relativistic phenomena.[219] Physicists like Max Planck and Hendrik Lorentz, well-known scientists who identified problems that conventional theories could not explain, were the first to push this characterization. This division was first proposed as a notion as a result of their participation in and contributions at the 1911 Solvay Conference.[220] The names of several physics textbooks reflect this split. For instance, Goldstein's Classical mechanics prelude shows why physics students might still benefit from studying the subject.[221] Arthur Beiser begins Concepts of contemporary Physics by defining contemporary physics as follows: Max Planck's discovery of the energy quantization component of blackbody radiation in 1900 marked the beginning of modern physics. Albert Einstein's equally groundbreaking theories of relativity and quantum theory of light came shortly after.[222] A few chapters on the shortcomings of classical physics open Kenneth Krane's Modern Physics, a book on quantum and relativity ideas.[223] The Classical Theories in book one and The Modern Theories in volume two of E.T. Whittaker's two-volume History of the Theories of Aether and Electricity (1900–1926) are subtitled.[224]

Paul Dirac's 1928 formulation of a relativistic quantum theory was the first of many quantum theories developed as the philosophically minded continued to dispute the underlying nature of the cosmos.[225] However, theoretical formulations that produced limitless

[210] Bohr, N. (1928). "The Quantum Postulate and the Recent Development of Atomic Theory". Nature. 121 (3050): 580–590. Bibcode:1928Natur. 121..580B. doi:10.1038/121580a0, accessed on 03 Dec 2025

[211] Herbert S. Green (1965). Matrix mechanics (P. Noordhoff Ltd, Groningen, Netherlands) ASIN: B0006BMIP8,

[212] Von Meyenn, Karl (1 February 2001). "Wolfgang Pauli". Physics Today. 54 (2): 43–48, doi:10.1063/1.1359709, accessed on 03 Feb 2026

[213] Pauli, Wolfgang (1980-01-01). General principles of quantum mechanics. Springer-Verlag. ISBN 9783540098423, accessed on 03 Feb 2026

[214] Buckley, Paul; Peat, F. David; Bohm; Dirac; Heisenberg; Pattee; Penrose; Prigogine; Rosen; Rosenfeld; Somorjai; Weizsäcker; Wheeler (1979). "Leon Rosenfeld". In Buckley, Paul; Peat, F. David (eds.). A Question of Physics: Conversations in Physics and Biology. University of Toronto Press. pp. 17–33. ISBN 978-1-4426-5166-1, accessed on 03 Dec 2025

[215] Kragh, Helge (1999), Quantum Generations: A History of Physics in the Twentieth Century, Princeton, New Jersey: Princeton University Press, accessed on 03 Dec 2025

[216] Wali, Kameshwar C (2009), Satyendra Nath Bose: his life and times (selected works with commentary), Singapore: World Scientific, ISBN 978-981-279-070-5, accessed on 03 Feb 2026

[217] Leggett, Anthony J. (2022). Quantum liquids: Bose condensation and Cooper pairing in condensed-matter systems (First published in paperback ed.). Oxford: Oxford University Press. ISBN 978-0-19-285694-4, accessed on 03 Dec 2025

[218] Dirac, Paul Adrien Maurice (1981-01-01). The Principles of Quantum Mechanics. Clarendon Press. p. 149. ISBN 9780198520115, accessed on 03 Feb 2026

[219] Kragh, Helge (2015). "The "new physics"". In Saler, Michael T. (ed.). The fin-de-siècle world. The Routledge worlds. London: Routledge. ISBN 978-0-415-67413-3, accessed on 03 Feb 2026

[220] Staley, Richard (December 2005). "On the Co-Creation of Classical and Modern Physics". Isis. 96 (4): 530–558. doi:10.1086/498592. ISSN 0021-1753, accessed on 03 Dec 2025

[221] Goldstein, Herbert (1980). Classical mechanics. Addison-Wesley series in physics (2 ed.). Reading, Mass: Addison-Wesley Pub. Co. ISBN 978-0-201-02918-5, accessed on 03 Dec 2025

[222] Beiser, Arthur (2003). Concepts of modern physics (International ed., 6. ed.). Boston: McGraw-Hill. ISBN 978-0-07-244848-1, accessed on 03 Feb 2026

[223] Krane, Kenneth S. (2020). Modern physics (4 ed.). Hoboken, New Jersey: John Wiley & Sons, Inc. ISBN 978-1-119-49548-2, accessed on 03 Dec 2025

[224] Whittaker, Edmund T. (1989). A history of the theories of aether & electricity. 2: The modern theories, 1900 - 1926 (Repr ed.). New York: Dover Publ. ISBN 978-0-486-26126-3, accessed on 03 Feb 2026

[225] Bhaumik, Mani L. (2022). "How Dirac's Seminal Contributions Pave the Way for Comprehending Nature's Deeper

energies hampered efforts to fully quantize electromagnetic theory into the 1930s. Only after World War II, when Julian Schwinger, Richard Feynman, and Sin-Itiro Tomonaga independently proposed the renormalization approach, which made it possible to build a robust quantum electrodynamics (QED), was this issue deemed sufficiently overcome.[226] As the concept of the quantization of fields through "exchange forces" governed by an exchange of transient "virtual" particles—which were permitted to exist in accordance with the laws governing the uncertainties inherent in the quantum world—grew in popularity, new theories of fundamental particles proliferated.[227] The uncertainty principle limits the existence of virtual particles, which are theoretical transitory particles that have some of the properties of conventional particles but may spontaneously originate from vacuum at short time and space ranges.[228] Notably, Hideki Yukawa postulated that a particle with a mass between that of the electron and the proton mediated a strong but short-range force that held the nucleus' positive charges together.[229] This particle, known as the "pion," was discovered in 1947 as one of several particles that were found after World War II. These particles were first discovered to be ionising radiation left by cosmic rays, but they are now increasingly being created in more advanced and potent particle accelerators.[230]

The development of the laser, which won the 1964 Nobel Prize in Physics, was one of the major breakthroughs of the era outside of particle physics.[231] Additionally, the study of superconductivity both theoretically and experimentally, particularly the development of a quantum theory of superconductivity by Vitaly Ginzburg and Lev Landau, which won the 1962 Nobel Prize in Physics, and its subsequent explanation via Cooper pairs (1972 Nobel Prize in Physics).[232] One of the first instances of quasiparticles was the Cooper pair. According to Einstein, a single theory can account for all of the basic interactions in nature.[233] Several efforts to "merge" multiple interactions resulted in unified field theories. A gauge theory, which is a generalization of the concept of symmetry, is one of many formulations of such theories.[234] At last, the unification of strong, weak, and electromagnetic interactions was achieved by the Standard Model.[235] Every effort to combine gravity with another force was unsuccessful. In her experiment, "Chien-Shiung Wu" shattered parity in weak interactions,[236] a series of discoveries were created thereafter.[237] New basic quantum theories were made possible by the interaction of these particles via decay and scattering. These new particles were given some order by Murray Gell-Mann and Yuval Ne'eman, who categorized them based on a number of characteristics, starting with what Gell-Mann called the "Eightfold Way."[238],[239] While its further development, the quark model,[240] at first seemed inadequate to describe strong nuclear forces,[241] The development of quantum chromodynamics in the 1970s finalised a set of basic and exchange particles, permitting the brief emergence of rival theories such the S-Matrix,[242] It made it possible to create a "standard model" based on

Designs". Quanta. 8 (1): 88–100. arXiv:2209.03937. doi:10.12743/quanta. v8i1.96, accessed on 03 Dec 2025

[226] Schweber, Silvan (1994), QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga, Princeton, New Jersey: Princeton University Press, accessed on 03 Feb 2026

[227] Griffiths, D.J. (2008). Introduction to Elementary Particles (2nd ed.). John Wiley & Sons. p. 65. ISBN 978-3-527-40601-2, accessed on 03 Feb 2026

[228] Peskin, M.E., Schroeder, D.V. (1995). An Introduction to Quantum Field Theory, Westview Press, ISBN 0-201-50397-2, p. 80, accessed on 03 Feb 2026

[229] Yukawa, Hideki (1982). Tabibito (人) = The Traveler. World Scientific. pp. 46–47 & 118, 121–123, 10, Foreword, 141 & 163. doi:10.1142/0014. ISBN 9971-950-10-3, accessed on 03 Dec 2025

[230] Galison, Peter (1997), Image and Logic: A Material Culture of Microphysics, Chicago, Illinois: University of Chicago Press, ISBN 0-226-27917-0, accessed on 03 Feb 2026

[231] Taylor N (2022). "The Invention of the Laser". Optical Tweezers. Methods Mol Biol. Vol. 2478. pp. 3–10. doi:10.1007/978-1-0716-2229-2_1. ISBN 978-1-0716-2228-5, accessed on 03 Dec 2025

[232] Fossheim, Kristian; Sudboe, Asle (2005). Superconductivity: Physics and Applications. John Wiley and Sons. p. 7. ISBN 978-0-470-02643-4, accessed on 03 Feb 2026

[233] Braibant, Sylvie; Giacomelli, Giorgio; Spurio, Maurizio (2011). Particles and Fundamental Interactions: An Introduction to Particle Physics (illustrated ed.). Springer Science & Business Media. p. 109. ISBN 9789400724631,

[234] O'Raifeartaigh, Lochlainn; Straumann, Norbert (2000-01-01). "Gauge theory: Historical origins and some modern developments". Reviews of Modern Physics. 72 (1): 1–23, doi:10.1103/RevModPhys.72.1, accessed on 03 Feb 2026

[235] R. Oerter (2006). The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics (Kindle ed.). Penguin Group. p. 2. ISBN 978-0-13-236678-6, accessed on 03 Dec 2025

[236] Oertelt, Nadja (June 2, 2017). "Meet Chien-Shiung Wu, the "Queen of Nuclear Research" and destroyer of natural laws". massivesci.com, accessed on 03 Feb 2026

[237] https://home.cern/science/physics/antimatter, accessed on 03 Dec 2025

[238] Johnson, George (May 24, 2019). "Murray Gell-Mann, Who Peered at Particles and Saw the Universe, Dies at 89". Obituaries. The New York Times. ISSN 0362-4331, accessed on 03 Feb 2026

[239] Lawrence Joffe (14 May 2006). "Obituary: Yuval Ne'eman | Science". The Guardian. London, accessed on 03 Feb 2026

[240] Gell-Mann, M. (4 January 1964). "A Schematic Model of Baryons and Mesons". Physics Letters. 8 (3): 214–215. Bibcode:1964PhL.... 8..214G. doi:10.1016/S0031-9163(64)92001-3, accessed on 03 Dec 2025

[241] Ragheb, Magdi. "Chapter 4 Nuclear Processes, The Strong Force" (PDF). University of Illinois. Archived from the original (PDF) on 2012-12-18, accessed on 03 Feb 2026

[242] D.J. Gross; F. Wilczek (1973). "Ultraviolet behavior of non-abelian gauge theories". Physical Review Letters. 30 (26): 1343–1346. Bibcode:1973PhRvL.30.1343G. doi:10.1103/PhysRevLett.30.1343, accessed on 03 Feb 2026

gauge invariance mathematics that effectively explained all forces save gravitation and is still widely used in its field.[243],[244]

The Standard Model, based on the 'Yang–Mills theory'[245] groups the 'electroweak interaction theory'[246] and quantum chromodynamics into the "gauge group" structure.[247] Abdus Salam, Steven Weinberg, and Sheldon Glashow are credited for formulating the standard model's unification of the electromagnetic and weak interactions.[248] Neutral weak currents were subsequently seen in experiments to validate the electroweak hypothesis,[249],[250] and distinguished by the 1979 Nobel Prize in Physics. Fundamental particle physics, especially the Big Bang idea put out as a result of Einstein's general theory of relativity, has shed light on the early universe's cosmology since the 1970s.[251],[252] But since the 1990s, astronomical discoveries have also presented new problems, such the need for fresh theories on galaxy stability, or "dark matter," and the apparent acceleration of the universe's expansion, or "dark energy."[253],[254] Though many theorists thought that supersymmetry and string theory would be a promising path forward, no theory that reconciles general relativity with the Standard Model has yet been discovered, despite the fact that accelerators have confirmed most aspects of the Standard Model by detecting expected particle interactions at various collision energies. The way these threads move across space and communicate with one another is explained by string theory.[255] A string behaves as a particle on distance scales greater than the string scale; its vibrating state determines its mass, charge, and other characteristics. According to string theory, the graviton, a quantum mechanical particle that conveys the gravitational pull, is one of the string's many vibrational states.[256] As a result, string theory is a quantum gravity theory.[257] However, since it started running in 2008, the Large Hadron Collider has not discovered any evidence in favour of supersymmetry or string theory.[258]

After Einstein's General Theory of Relativity was published in 1915, cosmology could be considered a serious study subject, but it wasn't until the "Golden age of general relativity" that it became widely accepted in science.[259] Edwin Hubble and Vesto Slipher discovered the expansion of the universe in the 1920s by measuring the redshifts of Doppler spectra from galactic nebulae, around ten years later, during what was known as the "Great Debate."[260],[261] The Big Bang hypothesis was developed by Georges Lemaître and George Gamow using Einstein's general relativity.[262] Fred Hoyle, Thomas Gold, Jayant Narlikar, and Hermann Bondi developed a competing hypothesis known as the steady

---

[243] O'Raifeartaigh, Lochlainn; Straumann, Norbert (2000-01-01). "Gauge theory: Historical origins and some modern developments". Reviews of Modern Physics. 72 (1): 1–23, doi:10.1103/RevModPhys.72.1. ISSN 0034-6861, accessed on 03 Feb 2026

[244] Ortín, T. (2015). Gravity and Strings. Cambridge University Press. ISBN 978-0-521-76813-9,

[245] https://ncatlab.org/nlab/show/theory+of+everything, accessed on 03 Feb 2026

[246] Salam, A.; Ward, J. C. (1959). "Weak and electromagnetic interactions". Nuovo Cimento. 11 (4): 568–577. Bibcode:1959NCim...11.568S. doi:10.1007/BF02726525, accessed on 03 Dec 2025

[247] J. Greensite (2011). An introduction to the confinement problem. Springer. ISBN 978-3-642-14381-6, accessed on 03 Dec 2025

[248] Hasert, F. J.; Faissner, H.; Krenz, W.; Von Krogh, J.; Lanske, D.; Morfin, J.; Schultze, K.; Weerts, H.; Bertrand-Coremans, G. H.; Lemonne, J.; Sacton, J. (1973-09-03). "Search for elastic muon-neutrino electron scattering". Physics Letters B. 46 (1): 121–124. Bibcode:1973PhLB...46.121H. doi:10.1016/0370-2693(73)90494-2, accessed on 03 Dec 2025

[249] Hasert, F. J.; Kabe, S.; Krenz, W.; Von Krogh, J.; Lanske, D.; Morfin, J.; Schultze, K.; Weerts, H.; Bertrand-Coremans, G. H.; Sacton, J.; Van Doninck, W. (1973-09-03). "Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment". Physics Letters B. 46 (1): 138–140. Bibcode:1973PhLB...46..138H. doi:10.1016/0370-2693(73)90499-1, accessed on 03 Dec 2025

[250] The discovery of the weak neutral currents, CERN courier, 2004-10-04, accessed on 03 Dec 2025

[251] Stookey, Lorena Laura (2004). "Ymir Motif". Thematic Guide to World Mythology. Bloomsbury Publishing USA. pp. 213–214. ISBN 978-0-313-03937-9, accessed on 03 Feb 2026

[252] NASA/WMAP Science Team (6 June 2011). "Cosmology: The Study of the Universe". Universe 101: Big Bang Theory. Washington, D.C.: NASA. Archived from the original on 29 June 2011, accessed on 03 Dec 2025

[253] Overbye, Dennis (20 February 2017). "Cosmos Controversy: The Universe Is Expanding, but How Fast?". The New York Times, accessed on 03 Feb 2026

[254] Scharping, Nathaniel (18 October 2017). "Gravitational Waves Show How Fast the Universe is Expanding". Astronom, accessed on 03 Feb 2026

[255] Becker, K.; Becker, M.; Schwarz, J. H. (2006). String Theory and M-Theory: A Modern Introduction. Cambridge University Press. ISBN 978-0-521-86069-7, accessed on 03 Feb 2026

[256] Green, Michael; Schwarz, John; Witten, Edward (2012). Superstring Theory. Vol. 1: Introduction. Cambridge University Press. ISBN 978-1-107-02911-8, accessed on 03 Dec 2025

[257] Kiefer, Claus (2012). Quantum gravity. International series of monographs on physics (3rd ed.). Oxford: Oxford University Press. pp. 1–4. ISBN 978-0-19-958520-5, accessed on 03 Dec 2025

[258] Woit, Peter (20 October 2013). "Last Links for a While". Not Even Wrong, accessed on 03 Dec 2025

[259] Crelinsten, Jeffrey. "Einstein's Jury: The Race to Test Relativity Archived 28 August 2014 at the Wayback Machine". Princeton University Press. 2006. Retrieved on 13 March 2007. ISBN 978-0-691-12310-3, accessed on 03 Dec 2025

[260] Hubble, Edwin (December 1926). "Extragalactic nebulae". Astrophysical Journal. 64 (64): 321–369. Bibcode:1926ApJ.... 64..321H. doi:10.1086/143018, accessed on 03 Dec 2025

[261] Giordano, Nicholas (2009). College Physics: Reasoning and Relationships. Cengage Learning. pp. 421–424. ISBN 978-0-534-42471-8, accessed on 03 Dec 2025

[262] Ryden, Barbara (17 November 2016). Introduction to Cosmology (2 ed.). Cambridge University Press. doi:10.1017/9781316651087. ISBN 978-1-107-15483-4, accessed on 02 Dec 2025

state theory.[263] In the 1960s, Robert Woodrow Wilson and Arno Allan Penzias confirmed the existence of cosmic microwave background radiation, which supported the big bang theory over the steady state theory.[264] Among others, George Smoot et al. (1989) used data from the Cosmic Background Explorer (CoBE) in their later work,[265] and these findings were improved by the Wilkinson Microwave Anisotropy Probe (WMAP) satellites.[266] Alan Guth proposed inflation theory in the 1980s, the same decade as the COBE data. Dark energy and dark matter issues have recently taken center stage on the cosmological agenda.[267]

On July 4, 2012, scientists at the Large Hadron Collider at CERN [268] declared the discovery of a new subatomic particle that closely resembles the Higgs boson. This particle might be crucial to explaining the existence of variety and life in the cosmos as well as the reason why elementary particles have mass.[269] It is now referred to by some physicists as a "Higgslike" particle. "It's something that may, in the end, be one of the biggest observations of any new phenomena in our field in the last 30 or 40 years, going way back to the discovery of quarks, for example," said Joe Incandela of the University of California, Santa Barbara.[270] "This is a big moment for particle physics and a crossroads – will this be the high water mark or will it be the first of many discoveries that point us towards solving the really big questions that we have posed," stated Michael Turner, a cosmologist at the University of Chicago and chairman of the physics center board.'[271] The concept of the Higgs field (also known as "cosmic molasses") was created in 1964 by six scientists, including Peter Higgs, who worked in three separate groups. Higgslike fields are crucial to string theory and universe theories, despite the fact that they have never been seen.[272] They may get infused with energy that exerts an antigravitational force under certain circumstances, according to the peculiar accounting of Einsteinian physics. These fields have been suggested as the origin of the universe's early inflation, a massive expansion, and as the source of the dark energy that seems to be speeding up the universe's expansion today.[273]

The basic physical theory that explains how matter and light behave is called quantum mechanics; its peculiar properties usually arise at and below the size of atoms.[274] A key component of the history of contemporary physics is the development of quantum mechanics. The creation of quantum concepts to describe specific phenomena, such as solar emission spectra, blackbody radiation, and the photoelectric effect, marks the beginning of the main chapters of this history. This period is known as the "old" or "older" quantum theories.[275] A number of experimental discoveries of atoms that didn't make intuitive sense within the framework of classical physics gave rise to the discipline of quantum physics in the late 1800s and early 1900s. The understanding that matter and energy may be seen as discrete packets, or quanta, having a minimum value attached to them was one of the fundamental discoveries.[276] For instance, energy is delivered by light with a defined frequency in quanta known as "photons." At this frequency, the energy of each photon will be the same and cannot be divided into smaller units.[277] Actually, the Latin origin of the term "quantum" means "how much." Understanding quantum principles changed how we

---

[263] Kragh, Helge (1999). Cosmology and Controversy: The Historical Development of Two Theories of the Universe. Princeton University Press. ISBN 978-0-691-02623-7, accessed on 02 Dec 2025

[264] Fixsen, D. J. (2009). "The Temperature of the Cosmic Microwave Background". The Astrophysical Journal. 707 (2): 916–920, doi:10.1088/0004-637X/707/2/916,

[265] Horgan, J. (1992). "Profile: George F. Smoot – COBE's Cosmic Cartographer". Scientific American. 267 (1): 34–41. doi:10.1038/scientificamerican0792-34, accessed on 02 Dec 2025

[266] Glenday, C., ed. (2010). Guinness World Records 2010: Thousands of new records in The Book of the Decade!. Bantam Books. p. 7. ISBN 978-0553593372, accessed on 02 Dec 2025

[267] da Silva, Wilson (March 2, 2015). "The physicist who inflated the Universe". Cosmos,

[268] oel Achenbach (March 2012). "The God Particle". National Geographic Magazine. Archived from the original on 25 February 2008, accessed on 02 Dec 2025

[269] Steinhardt, Paul J. (2011). "The inflation debate: Is the theory at the heart of modern cosmology deeply flawed?". Scientific American. Vol. 304, no. 4. pp. 18–25. Bibcode:2011SciAm.304d..36S. doi:10.1038/scientifica merican0411-36, accessed on 02 Dec 2025

[270] Ian Sample at Cern, Geneva (2012-07-04). "Higgs boson: it's unofficial! Cern scientists discover missing particle | Science | guardian.co.uk". London: Guardian, accessed on 03 Feb 2026

[271] Turner, Michael S. (1999). Gibson, Brad K.; Axelrod, Rim S.; Putman, Mary E. (eds.). The Third Stromlo Symposium, the galactic halo: proceedings of a symposium held at Australian Academy of Science, Canberra, Australia, 17-21 August, 1998. Vol. 165. San Francisco, Calif.: Astronomical Society of the Pacific. p. 431. ISBN 1-886733-86-4, accessed on 02 Dec 2025

[272] Griggs, Jessica (Summer 2008) The Missing Piece Archived 20 December 2016 at the Wayback Machine Edit the University of Edinburgh Alumni Magazine, p. 17, accessed on 02 Dec 2025

[273]https://www.nytimes.com/2012/07/05/science/cern-physicists-may-have-discovered-higgs-boson-particle.html?pagewanted=3&_r=1&ref=science, accessed on 02 Dec 2025

[274] Feynman, Richard; Leighton, Robert; Sands, Matthew (1964). The Feynman Lectures on Physics. Vol. 3. California Institute of Technology. Retrieved 19 December 2020. Reprinted, Addison-Wesley, 1989, ISBN 978-0-201-50064-6, accessed on 03 Feb 2026

[275] Whittaker, Edmund T. (1989). A history of the theories of aether & electricity. 2: The modern theories, 1900 - 1926 (Repr ed.). New York: Dover Publ. ISBN 978-0-486-26126-3, accessed on 02 Dec 2025

[276] Howard, Don (1990). "Nicht Sein Kann Was Nicht Sein Darf, or The Prehistory of EPR, 1909–1935: Einstein's Early Worries About The Quantum Mechanics of Composite Systems" (PDF). In Miller, A. I. (ed.). Sixty-Two Years of Uncertainty. New York: Plenum Press. pp. 61–111, accessed on 03 Feb 2026

[277] Tipler, Paul A.; Llewellyn, Ralph A. (1999), Modern Physics, vol. 3, W.H. Freeman & Co., p. 3, ISBN 978-1-57259-164-6, LCCN 98046099, accessed on 03 Feb 2026

thought about the atom, which is made up of an electron-surrounded nucleus.[278] According to early theories, electrons were particles that orbited the nucleus in a manner similar to that of satellites orbiting the Earth. Instead, electrons are understood by modern quantum physics to be dispersed inside orbitals, which are mathematical representations of the likelihood that electrons will reside in several locations within a particular range at any given moment.[279] When an electron gains or loses energy, it may move from one orbital to another, but it cannot be located in between orbitals. A few key ideas contributed to the development of quantum physics' foundations:

- Wave-particle duality: Matter and energy, like electrons and light, may have both particle-like and wave-like characteristics. In processes like interference and diffraction, light acts like a wave; but, in other situations, like the photoelectric effect, which includes discrete energy packets, light behaves like a particle.[280] According to Louis de Broglie, the concept also applies to matter, which has a wave character.[281],[282] This idea has existed since the beginning of quantum research. It explains the results of studies that demonstrated that matter and light had either particle or wave characteristics, depending on how they were measured. We now know that these many types of energy are neither particles nor waves. They are unique quantum entities that are difficult for us to imagine.
- Superposition: Imagine touching a pond's surface simultaneously at two distinct locations. From each point, waves would radiate outward and eventually overlap to create a more intricate design. This is a wave superposition. Similar to this, in quantum physics, particles with wavelike characteristics, like electrons and photons, may combine to form what is known as superposition. This phrase refers to an item that is simultaneously a composite of many different states. A ripple on a pond's surface, which is the result of two waves overlapping, is comparable to a superposed item. In mathematics, an equation with several solutions or outcomes may be used to describe an object in superposition.[283]
- Uncertainty principle: A key idea in quantum physics is the uncertainty principle, which asserts that there is a limit to how accurately certain pairs of physical attributes, such as location and momentum, may be known at the same time.[284] This implies that you can determine a particle's momentum less accurately the more exactly you know its location, and vice versa. This idea is not the product of faulty measuring instruments, but rather an inherent characteristic of nature.[285] It is a mathematical idea that shows a trade-off between opposing viewpoints. This indicates that two characteristics of an item, such its location and velocity, cannot be accurately determined simultaneously in physics.[286] For instance, we can only determine an electron's speed to a certain extent if we measure its location correctly.[287]

Entanglement: refers to a quantum phenomenon in which, independent of their distance from one another, two or more particles become connected and have entangled features.[288] Faster-than-light communication is not possible, yet measuring one entangled particle's property instantly affects the other's corresponding property. In a broader context, "entanglement" may also mean being fully immersed in a challenging or complicated circumstance.[289] This phenomenon happens when two or more items, even if they are extremely far away, are linked in a manner that allows them to be considered as a unified system. It is impossible to properly explain the state of one item in that system without knowing the state of the other.[290] Similarly, discovering details about one item often reveals something about another, and vice versa.[291]

It is somewhat honest for a guy on Earth to determine the location and speed of a moving object. We can instantly determine the location of items by measuring an automobile moving at 100 km/h or a person strolling

---

[278] Heisenberg, W. (1930), Physikalische Prinzipien der Quantentheorie (in German), Leipzig: Hirzel English translation The Physical Principles of Quantum Theory. Chicago: University of Chicago Press, 1930, accessed on 03 Feb 2026

[279] https://www.cantorsparadise.com/mathematical-uncertainty-principles-d45134b10e4d, accessed on 02 Dec 2025

[280] Millikan, R. (1916). "A Direct Photoelectric Determination of Planck's "h"". Physical Review. 7 (3): 355–388. Bibcode:1916PhRv....7..355M. doi:10.1103/PhysRev.7.355, accessed on 03 Feb 2026

[281] Leroy, Francis (2003). A Century of Nobel Prize Recipients: Chemistry, Physics, and Medicine (illustrated ed.). CRC Press. p. 141. ISBN 0-8247-0876-8. Extract of page 141, accessed on 03 Feb 2026

[282] M. J. Nye. (1997). "Aristocratic Culture and the Pursuit of Science: The De Broglies in Modern France". Isis. 88 (3) (Isis ed.): 397–421. doi:10.1086/383768, accessed on 03 Feb 2026

[283] https://scienceexchange.caltech.edu/topics/quantum-science-explained/quantum-superposition,

[284] https://www.britannica.com/science/uncertainty-principle, accessed on 02 Dec 2025

[285] Sen, D. (2014). "The Uncertainty relations in quantum mechanics" (PDF). Current Science. 107 (2): 203–218. Archived (PDF) from the original on 2019-09-24, accessed on 02 Dec 2025

[286] Landau, Lev Davidovich; Lifshitz, Evgeny Mikhailovich (1977). Quantum Mechanics: Non-Relativistic Theory. Vol. 3 (3rd ed.). Pergamon Press. ISBN 978-0-08-020940-1, accessed on 03 Feb 2026

[287] https://www.youtube.com/watch?v=MBnnXbOM5S4, accessed on 03 Feb 2026

[288] Lee, K. C.; Sprague, M. R.; Sussman, B. J.; Nunn, J.; et al. (2 December 2011). "Entangling macroscopic diamonds at room temperature". Science. 334 (6060): 1253–1256, doi:10.1126/science.1211914, accessed on 03 Feb 2026

[289] Horodecki, Ryszard; Horodecki, Pawel; Horodecki, Michal; Horodecki, Karol (2009). "Quantum entanglement". Reviews of Modern Physics. 81 (2): 865–942. arXiv:quant-ph/0702225. Bibcode:2009RvMP...81..865H. doi:10.1103/RevModPhys.81.865, accessed on 02 Dec 2025

[290] https://scienceexchange.caltech.edu/topics/quantum-science-explained/entanglement, accessed on 03 Feb 2026

[291] Holland, Connor M.; Lu, Yukai; Cheuk, Lawrence W. (8 December 2023). "On-demand entanglement of molecules in a reconfigurable optical tweezer array". Science. 382 (6675): 1143–1147, doi:10.1126/science.adf4272, accessed on 03 Feb 2026

at 9 km/h. However, a basic mathematical property known as the uncertainty principle makes these computations impossible in the quantum world of particles. According to the widely accepted Copenhagen interpretation, quantum systems live in a superposition of possibilities until they are measured. Since just one result is chosen at random at the time of measurement, the precise future is unclear.[292] Interestingly, German physicist and Nobel laureate Werner Heisenberg showed in 1927 that the inability to precisely know some combinations of properties simultaneously, such a particle's position and momentum, further limits overall prediction.[293] To put it another way, if a person were reduced to the size of an electron, we could only accurately ascertain its position or speed, not both at once. Although the Heisenberg uncertainty principle is well recognized in quantum physics, issues in pure mathematics and conventional physics also follow a similar concept. This idea basically applies to any object that has wave-like properties. Because of the nature of quantum mechanics theory, quantum objects are special because they naturally exhibit wave-like characteristics.[294] The area of study that studies matter and energy at the most fundamental level is called quantum physics. Its objective is to comprehend the characteristics and actions of the basic components of the natural world.[295]

Figure 4: What will future navigation and reality look like296

Consider a ripple in a pond to understand the fundamental idea of the uncertainty principle. You would follow a number of passing peaks and troughs to gauge its pace. Your understanding of the wave's speed will improve with the number of peaks and troughs you track, but your ability to pinpoint its precise location will diminish as it is scattered across those peaks and troughs. On the other hand, you would only need to watch a tiny portion of the wave in order to identify the location of a single peak, giving up information about its speed. The uncertainty principle essentially emphasizes a trade-off between two related qualities, such as location and speed. Another comparison is a rollercoaster: when the automobile reaches the top of a hill, a picture can reveal its location but not its speed; as it descends the hill, speed can be measured but its precise location is less definite. The idea shows how to balance two complementing factors, such location and speed.[297] Because subatomic particles may behave like waves, the basic law is applicable in the quantum domain. The uncertainty principle in quantum physics is often misunderstood to imply that our measurements are imprecise or uncertain.[298] In actuality, every event exhibiting wave-like behavior is fundamentally characterized by uncertainty.[299]

4. Investigation of Different Concept of Modern Science and Uncertainty

4.1 Einstein's Special Theory of Relativity and the Block Universe. Two events that seem simultaneously to one observer may not occur simultaneously to another observer travelling at a different speed, according to a fundamental tenet of special relativity. This is due to the fact that, regardless of an observer's motion, the speed of light remains constant. The theory

[292] Sen, D. (2014). "The Uncertainty relations in quantum mechanics" (PDF). Current Science. 107 (2): 203–218. Archived (PDF) from the original on 2019-09-24, accessed on 13 Oct 2025

[293] Werner Heisenberg (1989), Encounters with Einstein and Other Essays on People, Places and Particles, Princeton University Press, p. 53, accessed on 03 Feb 2026

[294] http://www.physics.udel.edu/~bnikolic/teaching/phys624/lectures/what_is_solid_state_physics.pdf, accessed on 13 Oct 2025

[295] https://scienceexchange.caltech.edu/topics/quantum-science-explained/quantum-physics, accessed on 03 Feb 2026

[296] https://www.facebook.com/groups/virtualrealitys/posts/1777999482292708/, accessed on 02 Feb 2026

[297] https://www.zmescience.com/feature-post/pieces/heisenberg-uncertainty-principle-095235/, accessed on 03 Feb 2026

[298] https://www.livescience.com/50776-thermodynamics.html, accessed on 03 Feb 2026

[299] https://scienceexchange.caltech.edu/topics/quantum-science-explained/uncertainty-principle, accessed on 13 Oct 2025

unifies space and time into a single, four-dimensional continuum known as spacetime in order to reconcile the perspectives of many observers. By fusing space and time into a single, four-dimensional entity known as space-time, Einstein's special theory of relativity completely changed our understanding of both concepts. The idea of the "block universe" resulted from this. Because of this, all earthly events—past, present, and future—exist concurrently and forever as components of a fixed continuum. According to this perspective, time exists all at once, just like space, rather than flowing.[300] The relativity of simultaneity, which means that various observers may disagree about what is occurring right now based on their relative velocities, is a crucial consequence. The philosophical concept of eternalization, which holds that every instant in time is equally real, is therefore supported by the collapse of the idea of a globally shared "present."[301] Determinism is implied by the block universe as the future seems to be regulated by the same unchangeable laws of physics as the past.[302] Determinism is implied by the block universe as the future seems to be regulated by the same unchangeable laws of physics as the past.[303] This idea, which implies a deterministic cosmos in which every occurrence is already "there" in spacetime, has significant philosophical ramifications.

4.2 Determinism Versus Indeterminism in Relativity. If a theory can only have one conceivable future conclusion given all beginning circumstances and the rules of physics, it is said to be deterministic.[304] Special relativity theory is completely deterministic as it deals with spacetime in an inertial (non-accelerating) context. By guaranteeing that all witnesses agree on the order of cause-and-effect occurrences, it supports the notion of causality. In addition to being locally deterministic—that is, given enough knowledge in a region of spacetime, its immediate future is uniquely determined—General Relativity (GR) defines gravity as the curvature of spacetime.[305] The fundamental equations of the theory are deterministic Einstein Field Equations. Accordingly, relativity proposes a deterministic cosmos in which the whole timeline—including the future—is fixed and preset. But this raises a question: can the future be really open if relativity reflects the large-scale cosmos accurately? We must go to quantum mechanics, which adds true indeterminacy to physical law, to overcome this conflict between determinism and an open future.[306] The theory predicts singularities, when the equations break down and predictability is lost, such those at the core of black holes or the Big Bang. It is generally accepted that a theory of quantum gravity must ultimately replace classical general relativity as it is inaccurate in these extreme circumstances. A boundary known as a Cauchy horizon may emerge in certain particular, but potentially unphysical, solutions to the equations (such as some charged black holes). A failure of determinism is shown by the possibility of various potential futures arising from the same basic circumstances beyond this horizon.[307] The spacetime must be "globally hyperbolic" in order to provide global determinism, in which a whole "now" slice of spacetime determines the entire future of the universe. However, this attribute is absent from many intriguing GR solutions, suggesting that determinism could not be a universal feature in all potential relativistic worlds.

4.3 Quantum Mechanics and the Indeterminacy of Reality. In contrast to classical determinism, quantum mechanics describes systems in terms of wave functions, which store probability for every potential event rather than specific values. Quantum systems exist in superposition—a composite of all conceivable states—prior to measurement.[308] Measurement yields a single result, but it is only possible to anticipate the outcome probabilistically rather than deterministically. Numerous experiments have confirmed this fundamental aspect of quantum physics.[309] Wave function collapse, the process by which the probabilistic wave function simplifies to a single conclusion, remains one of the most enigmatic aspects of quantum theory. Determining whether the future is fixed or open requires an understanding of whether this collapse is genuine and what causes it. By claiming that a system's attributes are not definite until measured and that results are dictated by probability distributions rather than strict causality, quantum mechanics introduces the indeterminacy of reality.[310] This idea, also referred to as the uncertainty principle or quantum indeterminacy, states that even with full knowledge of a quantum state (its wavefunction), one can only forecast the probability of various possibilities. This basic uncertainty casts doubt on classical determinism and suggests that nature is fundamentally probabilistic and subject to observation at the subatomic level.

---

300 https://iai.tv/articles/einstein-and-the-block-universe-auid-2065, accessed on 03 Feb 2026

301 Einstein, Albert; Podolsky, Boris; Rosen, Nathan (1935). "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?". Phys. Rev. 47 (10): 777–780, doi:10.1103/PhysRev.47.777, accessed on 03 Feb 2026

302 https://www.informationphilosopher.com/freedom/special_relativity.html, accessed on 02 Dec 2025

303 https://goldennumberratio.medium.com/the-block-universe-interpretation-of-einsteins-relativity-spacetime-is-wrong-224cd9a038a5, accessed on 03 Feb 2026

304 https://inters.org/determinism-indeterminism, accessed on 03 Feb 2026

305 https://www.quora.com/What-is-the-difference-between-indeterminism-and-determinism-Which-do-physicists-believe-is-true-of-our-world, accessed on 02 Dec 2025

306 https://medium.com/neo-cybernetics/determinism-and-indeterminism-b1e3a58b548f, accessed on 02 Dec 2025

307 James, William. “The Dilemma of Determinism.” In The Will to Believe and Other Essays in Popular Philosophy and Human Immortality. New York: Dover Publications, 1956, accessed on 03 Feb 2026

308 Jaeger, Gregg (September 2014). "What in the (quantum) world is macroscopic?". American Journal of Physics. 82 (9): 896–905. Bibcode:2014AmJPh.82..896J. doi:10.1119/1.4878358, accessed on 03 Feb 2026

309 Bialynicka-Birula, Zofia (2009), "Why photons cannot be sharply localized", Physical Review A, 79 (3): 7–8, doi:10.1103/PhysRevA.79.032112, accessed on 03 Feb 2026

310 https://thequran.love/2025/03/11/quantum-indeterminacy-how-quantum-mechanics-introduces-true-indeterminism-and-creates-free-will/#:~:text=, accessed on 03 Feb 2026

Probabilities take the role of classical certainty in quantum mechanics. For instance, the wavefunction simply gives a chance for when a radioactive atom could decay; the precise time is not known.[311] The measuring process is essential because it affects a particle's state by "collapsing" its wavefunction from a superposition of possibilities into a single, certain result. Intrinsic uncertainty is a characteristic of quantum systems and is not caused by measurement mistake or external influence. For example, the Kochen-Specker theorem shows that all measurable attributes of a quantum state cannot have sharp, definite values at the same time. Classical determinism, which maintains that all occurrences are foreseeable given sufficient knowledge about the system's starting state, is essentially challenged by challenge to determinism theory. According to quantum indeterminacy, reality is regulated by chance within the confines of those probabilities rather than being preset at its most basic level.[312] Although quantum indeterminacy is a fundamental idea, other interpretations try to explain its meaning. One such interpretation is that every quantum event splits the cosmos into several branches, each of which exists in a distinct reality.[313] Additionally, it implies that a kind of objective, spontaneous collapse takes place without the need for observation. So, the exact nature of quantum indeterminacy and its metaphysical implications (whether it represents a true indeterminacy in reality or is merely a limitation of our knowledge) remains a topic of ongoing discussion among physicists and philosophers.[314]

[311] https://fiveable.me/key-terms/principles-of-physics-iv/quantum-indeterminacy,

[312] https://pmc.ncbi.nlm.nih.gov/articles/PMC7517274/, accessed on 02 Dec 2025

[313] https://arxiv.org/abs/quant-ph/0012007, accessed on 03 Feb 2026

[314] https://www.researchgate.net/publication/232968685_Quantum_Mechanics_and_Metaphysical_Indeterminacy,

Figure 5: Quantum physics and prove parallel is it fact or fiction315

4.4 The Copenhagen Interpretation and Non-Determinism. Quantum mechanics (QM), not relativity, is the primary cause of indeterminism in contemporary physics. The traditional Copenhagen interpretation of QM holds that the results of quantum measurements are essentially random and independent of past occurrences. In order to give a more comprehensive understanding of determinism in the cosmos, theoretical physics is now working to unify these two pillars of contemporary research into a single theory of quantum gravity.[316] The philosophical ramifications of this scientific determinism on ideas like free will may be investigated.[317] One of the earliest and most extensively taught interpretations of quantum physics is the Copenhagen interpretation. According to this theory, a quantum system stays in superposition until a measurement is conducted, at which time the wave function collapses to a single, definitive result. This adds a basic element of randomness to the cosmos, which means that even with perfect knowledge of the present, it is impossible to forecast the future with confidence. But when mixed with relativity, this creates problems. Observers may disagree on whether the wave function of a certain event has already collapsed because they may perceive simultaneity differently. This disparity shows how difficult it is to reconcile quantum collapse with relativistic space-time and casts doubt on the notion of a universally accepted reality.[318]

[315] https://vocal.media/01/fact-or-fiction-could-quantum-physics-prove-parallel-universes, accessed on 02 Dec 2025

[316] Bishop, Robert. C. and Kronz, Frederick M. "Is Chaos Indeterministic?" In Language, Quantum, Music: Selected Contributed Papers of the Tenth International Congress of Logic, Methodology & Philosophy of Science, Florence, August 1995. Edited by Maria Luisa Dalla Chiara, Roberto Giuntini and Federico Laudisa. Boston: Kluwer Academic Publishers, 1999, accessed on 03 Feb 2026

[317] Smith, Leonard A. "Disentangling Uncertainty and Error: On the Predictability of Nonlinear Systems." In Nonlinear Dynamics and Statistics. Edited by A. I. Mees. Boston: Birkäuser, 2001, accessed on 03 Feb 2026

[318] Hossain K A, (2026), Quantum Foundations and Technological Futures: A Critical Analysis of Interpretative Frameworks and Socio-Economic Projections, Transactions on Engineering and Computing Sciences; Vol. 14 No. 01; 2026, pp 59-118; https://doi.org/10.14738/tmlai.1401.19883, ISSN: 2756-2638, accessed on 02 Feb 2026

Figure 6: Copenhagen vs Many Worlds Interpretation of Quantum Mechanics319

4.5 The Many-Worlds Interpretation and Determinism Restored. By suggesting that every potential result of a quantum event really takes place in distinct, parallel worlds, the Many-Worlds Interpretation (MWI) restores determinism. The global wavefunction, which develops deterministically in accordance with the Austrian physicist Schrödinger equation, instead of a single world collapsing into a single result,[320] just divides into many separate, non-interacting branches (worlds). The development of the multiverse as a whole is completely predictable and deterministic, yet from the viewpoint of an observer inside any one universe, the result seems random. A probabilistic "collapse" of the wave function, which is essential to the Copenhagen interpretation of quantum mechanics, is avoided by the MWI.[321] Once again, the wave function keeps evolving rather than collapsing. The wave function branches as a result of the observer being entangled with the system during a measurement or interaction. Each branch, however, stands for a distinct potential measurement result. In Schrödinger's cat, for instance, the world divides into two parts: one where the cat is alive and the other where it is dead. The Schrödinger equation governs the singular, deterministic process of the multiverse's whole development.[322] However, the result seems random from the viewpoint of a single observer inside one branch.[323] Because their own wave function is entangled with the measurement result, an observer can only ever perceive one of the possible outcomes.[324] Consequently, the MWI presents a bold but sophisticated alternative: it completely rejects the wave function's collapse. Instead, it suggests that all possible outcomes of a quantum event occur, each in its own parallel branch of the universe. The Schrödinger equation describes the deterministic evolution of the wave function, and every potential outcome is realised in some branch.[325] While all other versions of themselves experience different results, observers get entangled with these branches and encounter just one result. The cosmos is completely deterministic from a global standpoint, with all branches coexisting. Since the branching respects causal bounds like the speed of light, this interpretation is consistent with Einstein's relativity. However, since they only live in one branch, spectators see randomness from a subjective perspective.[326] Here, we have discovered that randomness is a result of our constrained viewpoint inside a particular branch rather than a basic aspect of reality.[327] Last but not least, this deterministic perspective results from acknowledging the universal wave function as physically

[319] https://www.youtube.com/watch?v=OjrEudqgZ1M, accessed on 02 Dec 2025
[320] Schrödinger, E. (1926). "An Undulatory Theory of the Mechanics of Atoms and Molecules" (PDF). Physical Review. 28 (6): 1049–70, doi:10.1103/PhysRev.28.1049. Archived from the original, accessed on 03 Feb 2026
[321] Everett, Hugh; Wheeler, J. A.; DeWitt, B. S.; Cooper, L. N.; Van Vechten, D.; Graham, N. (1973). DeWitt, Bryce; Graham, R. Neill (eds.). The Many-Worlds Interpretation of Quantum Mechanics. Princeton Series in Physics. Princeton, New Jersey: Princeton University Press. p. v. ISBN 0-691-08131-X, accessed on 03 Feb 2026
[322] Tegmark, Max (1998). "The Interpretation of Quantum Mechanics: Many Worlds or Many Words?". Fortschritte der Physik. 46 (6–8): 855–862. arXiv:quant-ph/9709032. Bibcode:1998ForPh. 46..855T. doi:10.1002/(SICI)1521-3978(199811)46:6/8<855::AID-PROP855>3.0.CO;2-Q, accessed on 03 Feb 2026
[323] Brown, Harvey R.; Christopher G. Timpson (2016). "Bell on Bell's Theorem: The Changing Face of Nonlocality". In Mary Bell; Shan Gao (eds.). Quantum Nonlocality and Reality: 50 years of Bell's theorem. Cambridge University Press. pp. 91–123. arXiv:1501.03521. doi:10.1017/CBO9781316219393.008, accessed on 02 Dec 2025
[324] Everett, Hugh (1957). "Relative State Formulation of Quantum Mechanics". Reviews of Modern Physics. 29 (3): 454–462, doi:10.1103/RevModPhys.29.454, accessed on 03 Feb 2026
[325] Schlosshauer, Maximilian; Kofler, Johannes; Zeilinger, Anton (August 2013). "A snapshot of foundational attitudes toward quantum mechanics". Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, (2013), 44 (3): 222–230, doi:10.1016/j.shpsb.2013.04.004, accessed on 02 Dec 2025
[326] Skyrms, Brian (1976). "Possible Worlds, Physics and Metaphysics". Philosophical Studies. 30 (5): 323–332. doi:10.1007/BF00357930. S2CID 170852547, accessed on 02 Dec 2025
[327] DeWitt, Bryce S. (1970). "Quantum mechanics and reality". Physics Today. 23 (9): 30–35. Bibcode:1970PhT....23i..30D. doi:10.1063/1.3022331. See also Ballentine, Leslie E.; Pearle, Philip; Walker, Evan Harris; Sachs, Mendel; Koga, Toyoki; Gerver, Joseph; DeWitt, Bryce (1971). "Quantum-mechanics debate". Physics Today. 24 (4): 36–44. Bibcode:1971PhT....24d..36.. doi:10.1063/1.3022676, accessed on 02 Dec 2025

real and letting it change in accordance with the basic quantum mechanical equations without the need for further, ad hoc restrictions like wave function collapse.[328]

4.6 Relativity, Decoherence, and the Propagation of Branching. Our comprehension of physical reality was completely transformed by the simultaneous discovery of quantum theory and Einstein's general theory of relativity. The concepts of spacetime and quantum superposition are perhaps the most unique aspects of each of these theories. Finding a coherent way to unify quantum theory with general relativity is the biggest difficulty facing current physicists.[329] Any theory that combines quantum mechanics and relativistic gravitation must unquestionably be able to characterize spacetime as having quantum-mechanical degrees of freedom, whose states exist in a complex Hilbert space and can be arranged in quantum superpositions of various configurations.[330], [331] Branching is not instantaneous across the universe under the Many-Worlds paradigm. Rather, it propagates at or below the speed of light in accordance with relativistic causality. Different realities are defined by entangled webs, which are made up of interrelated quantum events. Here, decoherence is essential because it explains how quantum branches effectively become separated from one another.[332] Each branch exhibits classical behavior as quantum interference between branches diminishes over time. A physical explanation for the change from quantum probability to classical experiences is provided by decoherence.[333] This results in an objective interpretation of the present as the particular entanglement network of which the observer is a member. However, because of the relativity of simultaneity, how each observer perceives "now" varies according on how they move through space-time.[334] This research enhances MWI's reputation as a solid framework for relativistic quantum theory by elucidating the interaction between branching, entanglement, and relativity.[335] It also draws attention to unsolved ontological issues: a thorough explanation of nonlocal branching requires a more thorough comprehension of the existence of entangled states in space and time. Future research will fill this gap.[336],[337]

[328] Byrne, Peter (2010). The Many Worlds of Hugh Everett III: Multiple Universes, Mutual Assured Destruction, and the Meltdown of a Nuclear Family. Oxford University Press. ISBN 978-0-19-955227-6, accessed on 02 Dec 2025

[329] C. Marletto and V. Vedral, gravitationally induced entangle-ment between two massive particles is sufficient evidence ofquantum effects in gravity, Phys. Rev. Lett. 119, 240402, (2017), accessed on 02 Dec 2025
(PDF) Relativity and decoherence of spacetime superpositions. Available from: https://www.researchgate.net/publication/368332890_Relativity_and_decoherence_of_spacetime_superpositions [accessed Nov 28 2025], accessed on 03 Feb 2026

[330] S. Bose, A. Mazumdar, G. W. Morley, H. Ulbricht, M. Toroˇs, M. Paternostro, A. A. Geraci, P. F. Barker, M. S. Kim, andG. Milburn, Spin entanglement witness for quantum gravity, Phys. Rev. Lett. 119, 240401 (2017), accessed on 02 Dec 2025

[331] D. L. Danielson, G. Satishchandran, and R. M. Wald, Black holes decohere quantum superpositions, arXiv preprintarXiv:2205.06279 (2022), accessed on 03 Feb 2026

[332] https://arxiv.org/html/2302.03259v2, accessed on 03 Feb 2026

[333] R. Gambini, J. Olmedo, and J. Pullin, Quantum black holes in loop quantum gravity, Classical and Quantum Gravity 31, 095009 (2014), accessed on 02 Dec 2025

[334] C. Kiefer, Conceptual problems in quantum gravity and quantum cosmology, ISRN Mathematical Physics 2013, 1–17 (2013), accessed on 02 Dec 2025

[335] Vaidman, L. (2021). Many-Worlds Interpretation of Quantum Mechanics. The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.). https://plato.stanford.edu/archives/fall2021/ entries/qm-manyworlds/, accessed on 03 Feb 2026

[336] Wallace, D. and C. G. Timpson (2010). Quantum Mechanics on Spacetime I: Spacetime State Realism. The British Journal for the Philosophy of Science 61, pp. 697-727, accessed on 02 Dec 2025

[337] Wallace, D. (2012). The Emergent Multiverse: Quantum Theory according to the Everett Interpretation. Oxford: Oxford University Press, accessed on 03 Feb 2026

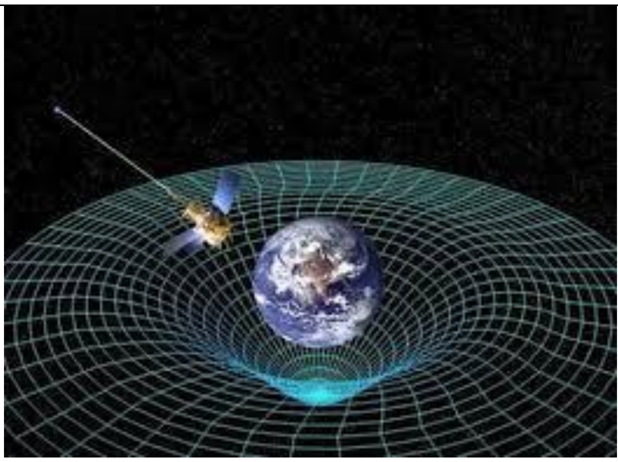


Figure 7: Relativity considered a theory rather than a law338

4.7 Pilot-Wave Theory and Its Challenges. Another interpretation of quantum physics that reinstates determinism without mentioning parallel universes is the de Broglie–Bohm pilot-wave theory.[339] As the first alternative formulation of quantum mechanics, de Broglie and Bohm's pilot-wave theory has a unique position in the history of modern physics.[340] It asserts that a guiding wave function that changes in accordance with quantum laws influences the particular paths that particles take. This gives particle behavior a precise, non-random explanation.[341] One of the difficulties with pilot-wave theory is that it is incompatible with special relativity, which necessitates a non-local, deterministic, and non-relativistic framework.[342] Its increased mathematical complexity, the ambiguous physical meaning of the "pilot wave," and the absence of reciprocal interaction between the particle and wave are further difficulties. Despite providing a deterministic substitute for the Copenhagen interpretation's unpredictability, these problems make it less well-liked than conventional interpretations. Nonetheless, there are substantial challenges in harmonizing the pilot-wave theory with special relativity.[343] The theory's non-local, instantaneous links between particles seem to go against the idea that nothing can move faster than the speed of light, which presents a significant obstacle.[344] A fundamental tenet of special relativity, Lorentz invariance, is violated by pilot-wave theory as it necessitates a preferred frame of reference.[345] According to the hypothesis, there is a true guiding wave, but it doesn't exist in the three-dimensional world we see; instead, it resides in an abstract, multi-dimensional realm. This makes it challenging to comprehend its physical nature, particularly because it cannot be identified as a separate item. It does not readily fit into the relativistic framework, which holds that no influence can move faster than light, since it contains non-local effects, where changes in one location may instantly impact distant particles. Its status as a comprehensive, contemporary physics theory is weakened by this incompatibility with relativity.[346] The hypothesis explains a one-way interaction in which the particle is guided by the pilot wave but has no effect on the wave. This goes against the basic idea of action-reaction in classical physics, which states that forces are always reciprocal. There are no new testable predictions made by pilot-wave theory that would set it apart from conventional quantum mechanics. The effectiveness of the theory is predicated on the idea of "quantum equilibrium," which postulates that the probabilistic character of quantum mechanics results from our incomplete understanding rather than from underlying randomness.

4.8 Consciousness and the Nature of the Present. The ability to analyses and compare mental experiences in a single instant, or "present time," is made possible by consciousness, a state of awareness that connects our perspective of the past, present, and future. Philosophically speaking, awareness is closely linked to the present; according to certain traditions,

[338] https://www.facebook.com/groups/physicsisfun109/posts/354094670602717/, accessed on 02 Dec 2025
[339] https://journals.aps.org/prd/abstract/10.1103/PhysRevD.54.2564, accessed on 02 Dec 2025
[340] https://www.researchgate.net/publication/391956708_The_trouble_with_pilot-wave_theory_a_critical_evaluation, accessed on 03 Feb 2026
[341] https://medium.com/swlh/on-pilot-wave-theories-and-general-issues-in-quantum-mechanics-2d2414320b15, accessed on 03 Feb 2026
[342] https://theconversation.com/why-did-louis-de-broglie-nobel-laureate-in-physics-abandon-his-own-pilot-wave-theory-244594, accessed on 03 Feb 2026
[343] https://www.aps.org/policy/reports/popa-reports/energy/climate.cfm, accessed on 03 Feb 2026
[344] https://link.springer.com/chapter/10.1007/978-3-319-55420-4_6, accessed on 03 Feb 2026
[345] uis Gonzalez-Mestres (1997-05-26). "Absence of Greisen-Zatsepin-Kuzmin Cutoff and Stability of Unstable Particles at Very High Energy, as a Consequence of Lorentz Symmetry Violation". Proceedings of the 25th International Cosmic Ray Conference (Held 30 July - 6 August). 6: 113, accessed on 03 Feb 2026
[346] A. C. de la Hamette, V. Kabel, E. Castro-Ruiz, and C. Brukner, falling through masses in superposition: quantum reference-frames for indefinite metrics (2021), accessed on 05 Dec 2025
(PDF) Relativity and decoherence of spacetime superpositions. Available from: https://www.researchgate.net/publication/368332890_Relativity_and_decoherence_of_spacetime_superpositions [accessed Nov 28 2025], accessed on 03 Feb 2026

consciousness is the capacity to see and value the "now" as the sole reality in which existence takes place. Others, such as philosopher Henri Bergson, see consciousness as an ongoing flow of "duration," a conflict between the present's immediacy and memory's expansive reach.[347] According to one idea, consciousness is thought to be the brain's greatest ability to connect space and time into global mental representations, presumably via quantum mechanical processes inside the brain. Many stress the significance of "conscious living," which is defined as being in the now and embracing the situation as it is rather than being preoccupied with regrets from the past or worries about the future. According to Eckhart Tolle, awareness is one with the present, and the self is essentially a sensation of presence or being that is the same as the present.[348] Consciousness is a constant "duration" that connects the accumulated potential of the past with the immediate present, according to Henri Bergson.[349] According to Matthew David Segall, awareness is a process that allows for freedom and choice by contrasting the present with what may be.[350] According to certain beliefs, physical reality comes after awareness, which is essential. According to a new theoretical concept, time, space, and matter originate from consciousness rather than substance.[351] Some theories make a distinction between the subjective, knower, or self that experiences the ideas and sensations that the brain processes and the processing itself.[352] Some believe that consciousness arises from matter or the brain, while others contend that awareness may really create matter.[353]

The subject of how quantum fields and quantum particles behave in a superposition of spacetimes seems to be outside the scope of theoretical and experimental studies in the absence of a comprehensive theory of quantum gravity.[354] The concept of a block universe raises an important philosophical and experiential query: what does it mean to perceive the present if all events have already occurred? The "now" we perceive is not an instantaneous moment but rather a prolonged period, according to consciousness, which results from neuronal processes developing over time.[355] These prolonged present runs counter to naïve notions of a clear temporal border between the past and the future, as well as presentism and the idea that only the present instant is genuine. Rather, a more nuanced, dispersed understanding of the present that is consistent with relativity's relative simultaneity is supported by our conscious awareness of time.[356] These cast doubt on oversimplified conceptions of time and pose more profound queries about the subjective perception of temporal flow.[357] A non-Turing test for machine consciousness based on a computer's capacity to generate philosophical conclusions was proposed by Victor Argonov in 2014.[358] In 2023, Nick Bostrom made the case that having a strong belief that large language models (LLMs) are not sentient would be unjustified. He also discussed whether aspects of consciousness theory are true and how they relate to machines.[359] According to him, awareness is a degree,[360] He maintained that, in principle, computers may be much more intelligent than people.[361]

Does the future already exist? That is the central question. It is still unsolved and mostly relies on one's preferred interpretation of quantum physics. According to the Copenhagen perspective, the future is essentially unpredictable and can only be influenced by observation. In a large multiverse, the future is decided for those who favor Many-Worlds, but each observer follows a branching, unpredictable route.[362] Physics is still struggling to reconcile the direction and flow of time, which explains why humans can recall the past but not the future.[363] In order to comprehend time, determinism, and the structure of reality, the movie concludes by emphasizing the significance of combining physical theory with philosophical understandings and the subjective realities of consciousness. Let's talk about the universe's transience, highlighting the fact that all matter—from stars and planets to galaxies—emerged from primordial cosmic beginnings and will eventually return to a fundamental, formless state. In a huge and mostly hostile universe, Earth is said to be a unique, life-supporting oasis. Earth, a planet in the solar system

[347] Robert C. Grogin, The Bergsonian Controversy in France, 1900–1914, Univ of Calgary Press (May 1988), ISBN 0919813305, accessed on 05 Dec 2025

[348] Rottgardt, Beate (27 May 2008). "Amerikas Guru stammt aus Lünen" [America's guru comes from Lünen]. Ruhr Nachrichten. Archived from the original on 15 January 2011, accessed on 05 Dec 2025

[349] Merquior, J. G. (1987). Foucault (Fontana Modern Masters series), University of California Press, p.11. ISBN 0-520-06062-8, accessed on 03 Feb 2026

[350] https://footnotes2plato.com/mds/, accessed on 03 Feb 2026

[351] https://footnotes2plato.substack.com/p/the-nature-of-consciousness-and-what, accessed on 05 Dec 2025

[352] https://www.organism.earth/library/document/out-of-your-mind-1, accessed on 03 Feb 2026

[353] P.A. Guertin (2019). "A novel concept introducing the idea of continuously changing levels of consciousness". Journal of Consciousness Exploration & Research. 10 (6): 406–412. Archived from the original on 2021-12-15, accessed on 03 Feb 2026

[354] Jaynes J (2000) [1976]. The Origin of Consciousness in the Breakdown of the Bicameral Mind. Houghton Mifflin. ISBN 0-618-05707-2, accessed on 05 Dec 2025

[355] https://www.frontiersin.org/articles/10.3389/fphy.2013.00006/full, accessed on 03 Feb 2026

[356] https://iopscience.iop.org/article/10.3847/0004-637X/830/2/83, accessed on 05 Dec 2025

[357] V. Kabel, A.-C. de la Hamette, E. Castro-Ruiz, and C. Brukner, Quantum conformal symmetries for spacetimes in superposi-tion (2022), accessed on 03 Feb 2026

[358] Victor Argonov (2014). "Experimental Methods for Unraveling the Mind-body Problem: The Phenomenal Judgment Approach". Journal of Mind and Behavior. 35: 51–70. Archived from the original on 2016-10-20, accessed on 03 Feb 2026

[359] Leith S (2022-07-07). "Nick Bostrom: How can we be certain a machine isn't conscious?". The Spectator,

[360] https://www.nytimes.com/2023/04/12/world/artificial-intelligence-nick-bostrom.html, accessed on 05 Dec 2025

[361] Shulman C, Bostrom N (August 2021). "Sharing the World with Digital Minds". Rethinking Moral Status: 306–326. doi:10.1093/oso/9780192894076.003.0018. ISBN 978-0-19-289407-6, accessed on 03 Feb 2026

[362] https://philsci-archive.pitt.edu/24763/7/branching%202025.pdf, accessed on 05 Dec 2025

[363] https://www.thp.uni-koeln.de/gravitation/zeh/Spacetime-Branching.pdf, accessed on 03 Feb 2026

and a member of the Milky Way galaxy, is a part of a cosmos that is always changing due to cycles of creation and destruction. This dynamic universe spans unfathomably vast periods, laying the groundwork for a future-focused story that examines how celestial planets, including Earth, will change and ultimately meet their demise.

4.9 The String-theory. A theoretical framework known as "string theory" substitutes one-dimensional objects known as "strings" for point-like particles. These strings vibrate in many ways, and each vibrational mode represents a distinct particle type, such an electron or a quark. Although the theory has not yet been scientifically confirmed, it seeks to unite the basic forces of nature, including gravity, and all forms of matter into a unified framework. String theory has the ability to include all matter and the four basic forces—gravity, electromagnetism, strong force, and weak force—into a single, coherent quantum mechanical framework, making it a contender for a "theory of everything." The theory's need for more dimensions than the three spatial and one temporal dimension of our universe—specifically, 10 or 11 dimensions in most formulations—is a crucial component. It is hypothesized that the additional dimensions are "compactified," or rolled up, and hence invisible at our size. String theory is still a theoretical idea despite its mathematical elegance.[364] As of yet, it has not produced testable predictions that have been confirmed by experimentation. The fact that the complete theory lacks a proper definition in every situation is one of the difficulties with string theory. Another problem is that the theory is believed to explain a vast array of potential universes, which has made attempts to create particle physics theories based on string theory more difficult.[365] Because of these problems, some members of the community have questioned the worth of further string theory unification research and criticized these methods of approaching physics.[366]

Figure 8: The frictional concept of string theory367

Stephen Wolfram, a mathematician, computer scientist, and physicist, recently shocked the scientific world by declaring that he had essentially worked out how the universe works. In the scientific world, Wolfram is well-known. He is the creator of Wolfram Alpha, the search engine that AskJeeves dreamed it were, Wolfram Language, a math-based programming language that powers the well-known Mathematica system, and the Wolfram Physics Project.[368] His contributions date back to his early years, when he wrote three books on physics before the age of 14.[369] His contributions date back to his early years, when he wrote three books on physics before the age of 14.[370] In order to do this, physicists and other scientists have developed ideas ranging from "we're all living in a computer simulation" (which raises the issue of what the universe the computer is in is comprised of) to numerous worlds (as in, more than one universe). Therefore, referring to a physics theory as "wacky" suggests a whole different degree of strangeness. "The fact that our universe is expanding was discovered almost a century ago, but scientists only realized how exactly this happens in the 90s of last century, when powerful telescopes (including orbital ones) appeared and the era of exact cosmology began," said Artyom Astashenok. Three to four billion years after the universe's creation, measurements and data analysis revealed that the universe is not only expanding, but expanding at an accelerated rate.[371] Additionally, he said, "And so the concept was created that the Universe is filled with some "dark energy," which has extraordinary qualities, rather than regular matter. It is called "Dark Energy" because no one understands

[364] Becker, Katrin; Becker, Melanie; Schwarz, John (2007). String theory and M-theory: A modern introduction. Cambridge University Press. ISBN 978-0-521-86069-7, accessed on 05 Dec 2025

[365] Zwiebach, Barton (2009). A First Course in String Theory. Cambridge University Press. ISBN 978-0-521-88032-9, accessed on 03 Feb 2026

[366] Merali, Zeeya (2011). "Collaborative physics: string theory finds a bench mate". Nature. 478 (7369): 302–304 [303]. Bibcode:2011Natur.
478..302M. doi:10.1038/478302a, accessed on 03 Feb 2026

[367] https://www.space.com/17594-string-theory.html, accessed on 05 Dec 2025

[368] Johnson, Bobbie (March 9, 2009). "British search engine 'could rival Google'". The Guardian. UK: Guardian News and Media, accessed on 03 Feb 2026

[369] https://writings.stephenwolfram.com/2020/04/finally-we-may-have-a-path-to-the-fundamental-theory-of-physics-and-its-beautiful/, accessed on 03 Feb 2026

[370] https://www.reddit.com/r/Physics/comments/16socnp/is_wolfram_physics_considered_a_legitimate/, accessed on 05 Dec 2025

[371] https://www.spacedaily.com/reports/Evidence_suggests_key_assumption_in_discovery_of_dark_energy_is_in_error_999.html, accessed on 05 Dec 2025

what it is or how it operates. And this energy makes up 70% of the universe.

The experts at the Immanuel Kant Baltic Federal University (IKBFU) offered their own notion about the "Dark Energy," which is the subject of several hypotheses. They clarified: "It has long been known that two metal plates put in a vacuum are attracted to one another. This phenomenon is known as the Casimir effect, after the Dutch scientist Hendrik Casimir. Since there is nothing in the hoover, it would seem that this cannot be. However, in accordance with quantum theory, particles are always appearing and disappearing there, and a very little attraction results from their contact with plates, which mark certain limits of space (which is crucial). Additionally, there is a theory that roughly the same thing occurs in space. On the contrary, only this causes more repulsion, which speeds up the universe's expansion. In other words, there is a manifestation of the universe's limits but ultimately no "Dark Energy." Naturally, this does not imply that it comes to an end, although complicated topology may occur. You may use the Earth as an example. Ultimately, it is finite yet also without bounds. The Earth and the Universe vary in that we deal with two-dimensional space in the former and three-dimensional space in the latter. According to Artem Astashenok, the published article expands on the concepts in Alexander Teplyakov's thesis and offers a mathematically sound model of the universe where additional repulsion takes place and the law of universal gravitation and the acceleration of the universe's expansion do not conflict.[372]

Recently, two physicists from IKBFU in Russia put forward a completely new theory of the universe. In essence, our whole universe is a component of an immensely huge quantum system spanning a "uncountable" multiverse, according to their study, which combines the absurd notion that we live in a computer simulation with the astounding "many worlds" hypothesis. When we think about quantum systems, such as Google's and IBM's quantum computers, we often picture a machine that uses qubits, which are subatomic particles, to carry out quantum computations. For now, these computers are helpful as a means of investigating the difference between classical and quantum reality, but they may eventually be able to do sophisticated computations that are now beyond the capabilities of classical computers.[373] The IKBFU academics who conducted the aforementioned study, Artyam and Valerian Yurov, argue that everything in the world—including the universe itself—should be thought of as a quantum object. This implies that we are already in "quantum reality" and do not need to examine qubits or subatomic particles. It's all quantum! In the opening of their work, Yurov and Yurov claim to have reversed the prevailing theories of theoretical physics. The article continues by providing a mathematical explanation of how the whole universe is a quantum item. In other words, it has quantum features that should include superposition, much like a small subatomic particle.[374] There must be something out there for our universe to interact with, even if it means using incredibly counterintuitive quantum mechanics to interact with itself in multiple states at once. This is because, theoretically, our universe should be able to exist in multiple places or states at once.

Expanding quantum mechanics to massive things, such as a single cell, has the drawback of making other theoretical quantum properties less logical. In this instance, "decoherence," or the process by which quantum objects "collapse" from many states into the physical state seen in our classical observations, doesn't seem to hold up at the cosmic scale.[375] The "many worlds" idea is then expanded to include numerous universes, or multiverses, by the researchers' calculations based on their assumptions. If the universe is a quantum item, it must interact with something, and that something is most likely other universes. This is the main notion here. The findings do not, however, explain why our world and everything in it would exist as something like to a single qubit in a massive quantum computer that concurrently spans numerous realities. If humans aren't the magical observers that measure the quantum realm and force it to "collapse" into classical reality, then we could just be gears in a machine; perhaps the cosmos is a qubit, and we are the qubits. Maybe humans are merely background noise that the universes overlook as they do their computations. Perhaps we are really living in a computer simulation. But rather than being the preferred NPCs of some sophisticated species, we are only mathematical components that support the operation of the operating system. Ultimately, one must question the viability of a unified theory of the universe that draws inspiration from both Nick Bostrom's simulation hypothesis and the arcade gaming age in comparison to M-theory, relativity, or other established solutions.[376] However, a rising tide lifts all ships, and Wolfram's present enthusiasm will undoubtedly produce some intriguing mathematical outcomes.

5. Ultimate Finding from Scientific Theories and Difficulties

5.1 Limitation of Theory of Science or Physics.

To yet, no physical hypothesis is thought to be exactly correct. Rather, physics has advanced via a sequence of "successive approximations" that enable more precise forecasts over an ever-expanding array of phenomena. Therefore, some physicists maintain that the sequence of approximations will never end in the "truth" and that it is incorrect to mix theoretical models with the actual nature of reality.[377] Einstein himself

[372] https://doi.org/10.1142/S0218271819501761, accessed on 03 Feb 2026

[373] https://thenextweb.com/news/ibm-algorithms-key-quantum-utility-first-european-data-centre, accessed on 05 Dec 2025

[374] https://thenextweb.com/news/spacex-inspires-iqm-quantum-computing, accessed on 05 Dec 2025

[375] https://thenextweb.com/news/study-our-universe-may-be-part-of-a-giant-quantum-computer,

[376] https://philosophybreak.com/articles/simulation-with-bostrom/, accessed on 05 Dec 2025

[377] Results, search (17 December 2006). The New Cosmic Onion: Quarks and the Nature of the Universe. CRC Press. ISBN 978-1-58488-798-0, accessed on 05 Dec 2025

sometimes voiced this opinion.[378] Whether a theory of everything is worthy of being referred to as the basic law of the universe is a topic of philosophical discussion within the physics community.[379] The hard reductionist viewpoint holds that the theory of everything is the basic law and that all other theories that apply to the cosmos follow from it. According to a different perspective, emergent laws—which control how complex systems behave—should be seen as equally basic.[380] The theory of natural selection and the second law of thermodynamics are two instances of emergent laws.[381] Emergence proponents contend that emergent laws are distinct from low-level, microscopic rules, particularly those that describe complex or living systems. According to this perspective, emergent rules are just as essential as a universal theory.

Weinberg notes that it is difficult to determine the exact trajectory of a real projectile in Earth's atmosphere.[382] Therefore, how can we be certain that our theory of projectile motion is sufficient? Newton's equations of motion and gravity, according to Weinberg, are "well enough" for basic situations, such as how planets move in empty space. We can be quite certain that these concepts will apply to more complicated situations since they have worked so effectively on simpler ones. For instance, general relativity is commonly regarded as a legitimate theory even if it contains equations without precise solutions since all of its equations with exact solutions have been confirmed via experimentation.[383] In a similar vein, a theory of everything must function for a variety of straightforward cases such that we can be very certain it will function for any scenario in physics. Combining quantum physics with general relativity often leads to challenges in developing a theory of everything since the equations of quantum mechanics start to break when the force of gravity is applied to them. The fundamental premise is that a quantum system loses its ability to exist as an autonomous entity when it interacts with a measurement device because their respective wave functions become entangled.[384] The uncertainty principle is one effect of the fundamental quantum formalism. In its most well-known version, this asserts that no quantum particle preparation may imply simultaneous, accurate predictions for a measurement of the particle's location and momentum.[385]

Once again, the zero-energy universe hypothesis postulates that there is precisely zero energy in the cosmos overall, with the negative energy of gravity precisely cancelling out the positive energy of matter.[386] Although the zero-energy universe hypothesis necessitates the existence of both a matter field with positive energy and a gravitational field with negative energy, some physicists, including Lawrence Krauss, Stephen Hawking, and Alexander Vilenkin, refer to this condition as "a universe from nothingness."[387] Popular sources address the theory in general. The strong CP problem, the absolute mass of neutrinos, matter–antimatter asymmetry, and the nature of dark matter and dark energy are only a few of the significant issues that still need to be resolved in the field of physics outside of the Standard Model.[388],[389] The Standard Model's mathematical foundation, which is still incompatible with general relativity, is another major issue. Under extreme circumstances, such as within known spacetime gravitational singularities like those at the Big Bang and at the centers of black holes beyond their event horizons, both theories collapse due to this contradiction.[390] On the other hand, Big Bang cosmology's chronology explains the universe's past and future. According to research released in 2015, the earliest phases of the universe's existence occurred 13.8 billion years ago, with a 68% confidence level uncertainty of around 21 million years.[391]

On a timescale of millions of trillions of years, it is hypothesized that black holes would appear to evap-

[378] The Theory of Everything: Are we getting closer, or is a final theory of matter and the universe impossible? Debate between John Ellis (physicist), Frank Close and Nicholas Maxwell, accessed on 05 Dec 2025

[379] https://www.bbc.co.uk/programmes/p004y24b, accessed on 03 Feb 2026

[380] O'Connor, Timothy; Wong, Hong Yu (February 28, 2012). "Emergent Properties". In Edward N. Zalta (ed.). The Stanford Encyclopedia of Philosophy (Spring 2012 Edition), accessed on 03 Feb 2026

[381] Darwin, Charles (1872). The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life (6th ed.). London: John Murray. OCLC 1185571, accessed on 05 Dec 2025

[382] Einstein, letter to Felix Klein, 1917. (On determinism and approximations.) Quoted in Pais (1982, Ch. 17),

[383] Landau, L. D.; Lifshitz, E. M. (1975), The Classical Theory of Fields, v. 2, Elsevier Science, Ltd., ISBN 978-0-08-018176-9, accessed on 05 Dec 2025

[384] Greenstein, George; Zajonc, Arthur (2006). "8 Measurement". The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics (2nd ed.). Jones and Bartlett. p. 215. ISBN 978-0-7637-2470-2. Archived from the original on 2023-01-02, accessed on 03 Feb 2026

[385] Landau, Lev D.; Lifschitz, Evgeny M. (1977). Quantum Mechanics: Non-Relativistic Theory. Vol. 3 (3rd ed.). Pergamon Press. ISBN 978-0-08-020940-1. OCLC 2284121, accessed on 05 Dec 2025

[386] "A Universe From Nothing lecture by Lawrence Krauss at AAI". YouTube. 2009. Archived from the original on 2021-12-15, accessed on 03 Feb 2026

[387] Isham, Christopher (14 July 1994). Quantum Cosmology and the Origin of the Universe (Speech). Lecture presented at the conference Cosmos and Creation. Cambridge University, accessed on 05 Dec 2025

[388] Hammond, Richard (1 May 2008). "The Unknown Universe: The Origin of the Universe, Quantum Gravity, Wormholes, and Other Things Science Still Can't Explain". Proceedings of the Royal Society of London, Series A. 456 (1999): 1685, accessed on 05 Dec 2025

[389] Womersley, J. (February 2005). "Beyond the Standard Model" (PDF). Symmetry Magazine. Archived from the original (PDF) on 17 October 2007, accessed on 03 Feb 2026

[390] Overbye, Dennis (11 September 2023). "Don't Expect a 'Theory of Everything' to Explain It All". The New York Times. Archived from the original on 11 September 2023, accessed on 03 Feb 2026

[391] Planck Collaboration (October 2016). "Planck 2015 results. XIII. Cosmological parameters". Astronomy & Astrophysics. 594: Article A13. arXiv:1502.01589. Bibcode:2016A&A...594A..13P. doi:10.1051/0004-6361/201525830, accessed on 03 Feb 2026

orate almost instantly, rare quantum tunnelling phenomena would appear to occur frequently, and quantum or other phenomena so improbable that they might occur only once in a trillion years might occur numerous times.[392] Once again, the hypothetical collection of all worlds is known as the multiverse.[393] All of space, time, matter, energy, information, and the physical laws and constants that characterize them are assumed to be included in these universes. The terms "parallel universes," "flat universes," "other universes," "alternate universes," "multiple universes," "plane universes," "parent and child universes," "many universes," or "many worlds" refer to the several realities that make up the multiverse. The multiverse is often thought of as a "patchwork quilt of separate universes all bound by the same laws of physics."[394] What will happen to the cosmos in the end is the question now. In the end, the cosmos will descend into near-absolute stillness and extreme darkness. Entropy will rule all processes, and all stars, galaxies, and structures will have either vanished or disintegrated. The universe's dark age, which might last forever, would be characterized by motionlessness, quiet, and eternity, broken only by brief, inconsequential quantum occurrences. But the last word is still elusive. There might be fresh beginnings at the nexus of general relativity and quantum physics, or the cosmos could stay inactive forever. In either scenario, the story concludes with humility and awe in the face of a huge, enigmatic, and incompletely understood universe.[395]

5.2 Possibility of the Alcubierre Warp Drive Ship. One potential faster-than-light propulsion method, the Alcubierszre warp drive, has long been limited by its need on unusual negative-energy materials.[396] It begs the question: would superluminal travel be technically possible in the near future? It could be theoretically feasible. One of the most fascinating possibilities in contemporary theoretical physics is the concept of faster-than-light (FTL) travel. The Alcubierre warp drive, which Miguel Alcubierre initially presented in 1994, is one of the suggested models[397] proposes a way for a spaceship to surpass the speed of light without going against special relativity. In order to accomplish this, spacetime is expanded behind the spaceship and contracted in front of it, enabling the ship to "surf" on a wave of spacetime. Despite its theoretical allure, the Alcubierre drive has a number of basic issues that have been thoroughly studied over the previous thirty years. These difficulties include the need for negative energy, the viability of dynamically modifying spacetime curvature, the possible consequences for time travel and closed time like curves (CTCs), and the massive energy requirements of maintaining a warp bubble. The need for a negative energy density to maintain the warp bubble is one of the main challenges in implementing an Alcubierre drive.[398] But it's thought that the quantity needed for a macroscopic warp drive is orders of magnitude more than what can be made in a lab right now. Lentz's recent work[399] and Bobrick & Martire[400] implies that, although these models are still theoretical, certain spacetime geometries could provide modified warp bubble solutions that reduce or eliminate the demand for exotic matter. Furthermore, new developments in semiclassical gravity and quantum field theory suggest that energy condition violations could not be as severe as originally believed, providing a potential avenue for further study.[401]

Another major challenge is the engineering of spacetime itself to achieve the anticipated contraction and expansion. Theoretical studies by Natáriopresent alternate warp drive metric formulations that do not need an explicit expansion area, implying that various geometries might result in more useful designs. Furthermore, new research on metric engineering in higher-dimensional theories and modified gravity models (such string theory and f(R) gravity) suggests potential ways to deform spacetime without the need for impractical energy densities.[402] Additionally, while experimental proof is still difficult, White et al.'s numerical simulations at NASA's Eagleworks laboratory have investigated possible methods of manipulating spacetime via local energy fluctuations.[403] The possible connection between warp drive physics and time travel is an important and often discussed topic. An operational Alcubierre drive might potentially allow backward time travel, which would result in causality

[392] Razavy, Mohsen (2003). Quantum Theory of Tunneling. World Scientific. pp. 4, 462. ISBN 978-981-256-488-7,

[393] https://www.newscientist.com/article/mg26234971-300-we-are-closer-than-ever-to-finally-proving-the-multiverse-exists/, accessed on 03 Feb 2026

[394] Swain, Frank (2017). The Universe Next Door: A Journey Through 55 Alternative Realities, Parallel Worlds and Possible Futures. London: New Scientist. p. 12. ISBN 978-1-4736-5867-7, accessed on 05 Dec 2025

[395] Hossain K A, (2025), Appraisal of Chronological Development of Physics with Creation of Universe and Uncertainty of Future, Transactions on Engineering and Computing Sciences - Vol. 13, No. 06 Publication Date: December 19, 2025 DOI:10.14738/tmlai.1306.19717, accessed on 05 Feb 2026

[396] Auci, M. (2021) Multi-Bubble Universe Model: A Quantum-Relativistic Gravitational Theory of Space-Time. Journal of Modern Physics, 12, 179-21, accessed on 05 Feb 2026

[397] Alcubierre, M. (1994) The Warp Drive: Hyper-Fast Travel within General Relativity. Classical and Quantum Gravity, 11, L73-L77, accessed on 03 Feb 2026

[398] Casimir, H.B.G. (1948) On the Attraction between Two Perfectly Conducting Plates. Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen, 51, 793-795, accessed on 05 Dec 2025

[399] Lentz, E.W. (2021) Breaking the Warp Barrier: Hyper-Fast Solitons in Einstein-Maxwell-Plasma Theory. Classical and Quantum Gravity, 38, Article ID: 075015, accessed on 05 Dec 2025

[400]Bobrick, A. and Martire, G. (2021) Introducing Physical Warp Drives. Classical and Quantum Gravity, 38, Article ID: 105009, accessed on 03 Feb 2026

[401] Visser, M. (1996) Lorentzian Wormholes: From Einstein to Hawking. AIP Press, accessed on 05 Dec 2025

[402] Capozziello, S. and Faraoni, V. (2011) Beyond Einstein Gravity: A Survey of Gravitational Theories for Cosmology and Astrophysics. Springer, accessed on 03 Feb 2026

[403] White, H.G., March, P., Lawrence, J., Vera, J., Sylvester, A., Brady, D. and Bailey, P. (2013) Measurement of the Warp Field Dynamics of a Lorentzian Wormhole. Journal of the British Interplanetary Society, 66, 242-247, accessed on 06 Dec 2025

paradoxes since any superluminal travel in general relativity entails the existence of closed time-like curves (CTCs).[404] After studying the stability of CTCs in warp drive spacetimes, Everett and Roman came to the conclusion that quantum backreaction effects could be able to avoid these paradoxes.[405] Furthermore, Barceló *et al*.[406] have investigated whether Hawking's chronology preservation methods and quantum inequalities may stop CTCs from forming in physically plausible situations. However, there is still disagreement regarding whether causality violations are an inevitable aspect of FTL travel or if there is a mechanism that may fix the problem. The Alcubierre drive's energy requirements continue to be a significant obstacle; preliminary calculations indicate that they would surpass the entire mass-energy of the visible universe.[407] The goal of recent studies has been to improve these computations so they are more physically sound. Pfenning and Ford[408] established limitations on warp bubble energy distributions, showing that energy needs may be greatly decreased by thinning the warp bubble wall. Furthermore, soliton-based warp bubble designs may enable subluminal preparation prior to superluminal speeds, thereby lowering energy requirements, according to recent research by Lentz.[409] In the meanwhile, studies on manipulating quantum vacuum energy, including those carried out by Davis and Puthoff,[410] suggests that the extraction of useful negative energy from the vacuum itself might be made possible by future developments in quantum field control.

In a recent study, Bi Qiao from the University of Wuhan Technology's Physics Department presents a unique method based on extended gauge transformations. He also creates a new formula that allows the electromagnetic tensor to be directly converted to the Weyl tensor. This method offers a workable alternative for the Alcubierre warp drive, in which precise electromagnetic field manipulation causes spacetime expansion (negative curvature) at the spacecraft's back and compression (positive curvature) at its front.[411] By doing this, it gets around Einstein's field equations' strict requirements for negative energy and exotic matter, providing a strong theoretical basis for the development of warp-drive propulsion.[412] He has shown and said that "a more practical path to realizing warp drive physics may be an electromagnetic field-induced Weyl curvature mechanism." He demonstrates that "a dramatic reduction in energy demands suggests that future technological advancements may be able to achieve faster-than-light travel." Additionally, his study offers crucial information on the controllability and safety of upcoming warp-driven spacecraft.[413] The suggested electromagnetic tensor to Weyl tensor modulation provides a potential path towards the practical implementation of warp-drive spacecraft, building upon the unification of the four basic interactions and generalized gauge transformations in primary fiber bundles. This method turns superluminal travel from a theoretical speculation into a possibly feasible engineering reality by removing the need on negative energy exotic matter and drastically lowering energy limitations. This discovery might enable mankind to transcend the boundaries of our solar system in the not-too-distant future, bringing about the interstellar period. [414]

6. Contemporary Concept of Universe Creation and it's Future

6.1 Preparation of Cosmic Calendar. An illustration of the universe's 13.8-billion-year history condensed into a single calendar year is the Cosmic Calendar. Carl Sagan popularized it in order to help people comprehend the length of cosmic history.[415] The Geologic Calendar is a comparable metaphor used to illustrate the geologic time scale and the evolution of life on Earth.[416] According to the Cosmic Calendar, human history represents a very small portion of cosmic time. In comparison to our species, the cosmos is unfathomably ancient. The majority of significant cosmic occurrences took place long before humanity arrived.[417] The

[404] Thorne, K.S. (1994) Black Holes & Time Warps: Einstein's Outrageous Legacy. W. W. Norton & Company,

[405] Everett, A.E. and Roman, T.A. (1997) Superluminal Travel and the Violation of the Weak Energy Condition. Physical Review D, 56, 2100-2108, accessed on 06 Dec 2025

[406] Barceló, C., Liberati, S. and Visser, M. (2002) Chronology Protection in Lorentzian Traversable Wormholes. Journal of High Energy Physics, 11, 42, accessed on 03 Feb 2026

[407] Klinkhamer, F.R. and Volovik, G.E. (2008) The Alcubierre Drive and the Energy Conditions. Physics Letters B, 673, 214-217, accessed on 06 Dec 2025

[408] Pfenning, M.J. and Ford, L.H. (1997) The Unphysical Nature of 'Warp Drive'. Classical and Quantum Gravity, 14, 1743-1751, accessed on 03 Feb 2026

[409] Davis, E. and Puthoff, H.E. (2000) Quantum Vacuum Energy Extraction and Applications to Space Propulsion. AIAA Journal of Propulsion and Power, 16, 763-768, accessed on 03 Feb 2026

[410] Davis, E. and Puthoff, H.E. (2000) Quantum Vacuum Energy Extraction and Applications to Space Propulsion. AIAA Journal of Propulsion and Power, 16, 763-768, accessed on 06 Dec 2025

[411] Bi, Q. (2023) Large Scale Fundamental Interactions in the Universe. Journal of Modern Physics, 14, 1703-1720,

[412] Auci, M. (2023) Quantum-Relativistic Properties of the Space-Time Bubbles and Their Evolution in a Multi-Bubble Universe. Journal of Modern Physics, 14, 1045-1066, accessed on 03 Feb 2026

[413] Qiao, B. (2024) Further Exploration of the Gauge Transformation across Fundamental Interactions. Journal of Modern Physics, 15, 2317-2334, accessed on 06 Dec 2025

[414] https://www.scirp.org/journal/paperinformation?paperid=141629, accessed on 03 Feb 2026

[415] Bell, Elizabeth A.; Boehnike, Patrick; Harrison, T. Mark; et al. (19 October 2015). "Potentially biogenic carbon preserved in a 4.1-billion-year-old zircon" (PDF). Proc. Natl. Acad. Sci. U.S.A. 112 (47): 14518–21. Bibcode:2015PNAS.11214518B. doi:10.1073/pnas.1517557112. ISSN 1091-6490, accessed on 03 Feb 2026

[416] Witt, Steven (December 2024). "Libraries at the intersection of history and the present". IFLA Journal. 50 (4): 691–695. doi:10.1177/03400352241299307. ISSN 0340-0352, accessed on 03 Feb 2026

[417] Borenstein, Seth (19 October 2015). "Hints of life on what was thought to be desolate early Earth". Excite. Yonkers, New York: Mindspark Interactive Network. Associated Press, accessed on 03 Feb 2026

Big Bang occurred at midnight on January 1st, according to this representation, and the present instant corresponds to the conclusion of December 31st, just before midnight.[418] There are 37.8 million years per cosmic day, 1.58 million years per cosmic hour, and 438 years every cosmic second at this size. If today is December 31 at 23:59:59 and the Big Bang occurred on January 1 at 00:00, then the remaining events are shown on a graph below in order.

COSMIC CALENDAR

January to March: Early Universe

Jan 1: Big Bang
Jan 10: First star form
Jan 13: First galaxies appear

April to August: Galaxy Growth

May: Milky Way begins to form
August: Our Milky Way stabilizes

September: Birth of the Solar System

Sep 1: Formation of the Sun
Sep 2: Formation of Earth
Sep 21: Earliest life (simple cells)

October — November: Evolution

Oct 9: Photosynthesis begins
Nov 12: First complex cells (eukaryotes)
Nov 15: First multicellular life
Nov 24: First animals (simple forms)

December: Life Explodes

Dec 5: Fish evolve
Dec 14: First land plants
Dec 17: Amphibians
Dec 18: Reptiles
Dec 26: Dinosaurs appear
Dec 30 (early morning): Dinosaurs go extinct
Dec 30 (afternoon): Mammals diversify

December 31 — Humans Arrive

Dec 31, 22:24: First humans (Homo sapiens)
Dec 31, 23:59:35: Agriculture starts
Dec 31, 23:59:50: Ancient civilizations (Egypt, Mesopotamia)
Dec 31, 23:59:59: All of recorded human history. → 1 second on the cosmic calendar.

Figure 9: Cosmic Calendar in short[419]

6.2 Earth and the Solar System's that Cloose to Near Future (2100–3000 AD). Significant astronomical and environmental changes are highlighted in projections into the 22nd and 23rd centuries. Sea level rise and ice cap melting are accelerated by global warming, which is brought on by increasing greenhouse gas levels. This forces coastal communities to relocate, mirroring the migrations of ancient humans. The Earth's magnetic field is progressively disrupted by solar phenomena including sunspot cycles, rising solar flares, and coronal mass ejections, which also pose a danger to contemporary technological systems like power grids, satellites, and communication infrastructure. By 2200 AD, increased volcanic activity releases aerosols that momentarily cool the planet, yet solar wind interactions with the Earth's magnetosphere still result in bright auroral displays. Around 2300 AD, polar melting inhibits ocean currents like the Gulf Stream, causing tropical regions to warm and portions of Europe to cool. This amplifies climatic disruption. Long-term orbital cycles known as Milankovitch cycles start to have a greater impact on climate as Earth approaches 2500 AD. These include variations in Earth's axial tilt, eccentricity (or orbital shape), and precession (wobble), which may cause both small ice ages and significant ecological disruptions. Human civilization faces significant problems as a result of these changes. Concurrently, cosmic ray levels are impacted by the solar system's motion across the galaxy, which may have an impact on Earth's temperature and atmospheric chemistry.

By 3000 AD, there may be protracted glacial periods, and gradual gravitational changes between planets would start to slightly change their orbits over far longer times.[420] In his book "The Future of Humanity: Terraforming Mars, Interstellar Travel, Immortality, and Our Destiny Beyond Earth," renowned theoretical physicist, futurist, and popular science author Michio Kaku explores how humanity can become a multiplanetary species to ensure survival. It covers everything from string theory and artificial intelligence to interstellar travel and colonization, offering an optimistic vision of space exploration.[421] Our existing civilization is classified as mid-level Type 0 by Michio Kaku. What does it signify? In 1964, Russian astronomer Nikolai Kardashev identified several civilization types and assigned them numbers ranging from 1 to 5. He made no mention of a Type 0. Another renowned astronomer, Carl Sagan, emphasized that by measuring our present condition inside a Type 0 parameter. From Amazonian tribal civilizations to our energy-hungry modern cities and nations, Sagan characterized human life as existing somewhere between Type 0.5 and 0.6. We were proficient in flying at 0.5. We arrived at the Moon at 0.6. As we are ready to establish a lasting presence on our closest celestial neighbor, we are now at 0.7. We will advance to 0.8, which should occur before 2100, after fusion energy is mastered and quantum computing becomes widely used. Kardashev's scale corresponded with the energy consumption of civilizations.[422] We are now at 17.5 Terawatt-hours, which is still a long way from Kardashev's first prediction of Type 1, which would increase our energy consumption until it is 100,000 times more than it is now. In the next 100 to 200 years, Kaku thinks that milestone will be reached and exceeded. This implies that humans will capture all of the solar energy that reaches Earth—roughly 100

[418] Blanchard, Therese Puyau (1995). "The Universe At Your Fingertips Activity: Cosmic Calendar". Astronomical Society of the Pacific. Archived from the original on 2007-12-16, accessed on 03 Feb 2026

[419] https://visav.phys.uvic.ca/~babul/AstroCourses/P303/BB-slide.htm, accessed on 03 Feb 2026

[420] https://jaxroam.medium.com/will-humanity-make-it-to-year-3000-on-earth-and-what-will-it-be-like-9940a308c948, accessed on 03 Feb 2026

[421] https://www.kirkusreviews.com/book-reviews/michio-kaku/the-future-of-humanity/, accessed on 06 Dec 2025

[422] https://www.nytimes.com/2018/05/25/books/review/the-future-of-humanity-michio-kaku.html, accessed on 06 Dec 2025

quadrillion Watt-hours. Our Type 1 society will be able to regulate volcanoes, earthquakes, and the weather.

Could we do this in a century or two? I don't believe it. But it seems plausible by the year 2500 CE. What more can we accomplish between that year and the next 500? Is it possible for humans to reach each Kardashev's Type 2 and use all of the Sun's energy, not only the sunlight that reaches our planet? Can we move Earth and the other planets in the Solar System using our Type 2 technologies? Can the extinction-level event link to the extinction of the dinosaurs no longer be considered a danger to human civilization if we reach Type 2? By 3000 CE, all of this appears improbable. When you consider how far we have come from 2500 BCE, the time when the Great Pyramids were built, to where we are today, however, then leaps of futuristic insight like this seem possible.[423] The Great Pyramid of Giza, built for Pharaoh Khufu around 2560 BCE (part of the 4th Dynasty Old Kingdom), was a staggering engineering feat using massive stone blocks, a feat achieved with copper tools and immense human labor, with no wheels or advanced machines It demonstrates extraordinary organizational and engineering expertise for that time period, given that building took around 20 to 30 years and involved transporting blocks weighing tons.[424] Consider our previous advancements. We used fire, water, and wind energy to power our preindustrial societies as early as 2500 BCE. Electricity and atomic energy have been introduced in the present.[425]

Human knowledge quadrupled every 1,500 years prior to the year 0 CE. It quadrupled every century by 1900 CE. Every 25 years by 1945. It now occurs once every twelve months, but it will soon occur once every twelve hours. By the year 3000 CE, which is 977 years from now, our society and the planet will hardly resemble what we know now. Both our world and ourselves will be re-engineered. We shall then go to Type 3, 4, 5, and beyond, mastering the Solar System and beyond. As we go through each of these phases, try to picture how human society will look at that point. What will limit us? Based on the speed of light, Einstein proposed a speed limit for the universe. Scientists are investigating the idea of Star Trek's warp drive, which bends spacetime to enable faster-than-light (FTL) travel without breaking relativity locally. NASA and others are looking into theoretical frameworks like the Alcubierre Drive, but there are still many obstacles to overcome, such as the requirement for a huge amount of negative energy, which makes it purely theoretical and far from being implemented in practice.[426] Scientists at NASA are working to make it happen now. Can we create the transporter from Star Trek and transmit ourselves around the galaxy if we can grasp quantum entanglement?[427] Furthermore, the fictitious wormholes that show up in science fiction films like Contact may serve as the equivalent of metropolitan subways, enabling humans to travel between star systems and come across other species that are going through similar phases of civilization to our own.[428]

6.3 Ecological and Solar Evolution Over Millennia (4000 AD–10,000 AD). While solar history is predicted to stay steady and continue its normal cycles within the existing knowledge of stellar physics, geological evolution will be dominated throughout the millennia between 4000 AD and 10,000 AD by continuous sea-level rise caused by the gradual melting of ice sheets.[429] Earth's topography is still being shaped by tectonic processes as millennia go by. While growing volcanic activity, particularly from super-volcanoes like Yellowstone, raises the possibility of extended climate cooling or "nuclear winter" conditions, mountain ranges like the Himalayas are rising.[430] The Earth gradually warms due to the Sun's increasing brightness, which also pushes the globe toward a wet greenhouse state and a critical tipping point when water starts to evaporate into the atmosphere, endangering biodiversity and ocean stability.[431] Reversals of the geomagnetic field finish their cycles, which might decrease Earth's magnetic shield and change how charged particles are distributed in the atmosphere. These modifications might have a domino impact on technology and the environment.[432] The subduction and spreading of the ocean bottom, as well as the assembly and disintegration of continents, are examples of plate tectonic processes that take place over millions of years. Significant large-scale tectonic movements are not expected between 4000 and 10,000 AD, but there will be continuous, slow alterations and sporadic volcanic eruptions (including potentially unanticipated super-volcano occurrences).[433] The Earth's climate system will continue to adapt to human influences. Over the millennia, it is anticipated that changes in rainfall patterns, a rise in the frequency of severe weather events, and changes in

[423] https://www.reddit.com/r/AskEngineers/comments/17rn2bd/how_could_the_great_pyramid_have_been_built_in_20/, accessed on 03 Feb 2026

[424] https://www.respectegypttours.com/blog/when-were-ancient-pyramids-built, accessed on 06 Dec 2025

[425] https://www.21stcentech.com/world-3000-beyond/, accessed on 06 Dec 2025

[426] Choi, M., Okyay, M.S., Dieguez, A.P., Ben, M.D., Ibrahim, K.Z. and Wong, B.M. (2024) QRCODE: Massively Parallelized Real-Time Time-Dependent Density Functional Theory for Periodic Systems. Computer Physics Communications, 305, Article ID: 109349, accessed on 06 Dec 2025

[427] https://gamerant.com/star-trek-warp-drive-technology-explained/, accessed on 06 Dec 2025

[428] https://www.youtube.com/watch?v=Z8axMaBL4uo, accessed on 06 Dec 2025

[429] Abreu, J.A., Beer, J., Steinhilber, F., Tobias, S.M. and Weiss, N.O., 2008, "For how long will the current grand maximum of solar activity persist?", Geophys. Res. Lett., 35, L20109, accessed on 03 Feb 2026

[430] https://cc.oulu.fi/~usoskin/personal/lrsp-2008-3Color.pdf, accessed on 03 Feb 2026

[431] Abreu, J.A., Beer, J., Ferriz-Mas, A., McCracken, K.G. and Steinhilber, F., 2012, "Is there a planetary influence on solar activity?", Astron. Astrophys., 548, A88, accessed on 03 Feb 2026

[432] Argento, D.C., Reedy, R.C. and Stone, J.O., 2013, "Modeling the earth's cosmic radiation", Nucl. Instrum. Methods B, 294, 464–469, accessed on 03 Feb 2026

[433] Donadini, F., Korte, M. and Constable, C., 2010, "Millennial Variations of the Geomagnetic Field: from Data Recovery to Field Reconstruction", Space Sci. Rev., 155, 219–246, accessed on 03 Feb 2026

ecosystem boundaries (such as desertification) would continue to alter landscapes.[434] The heliosphere, the Sun's protective bubble, contracts when the solar system passes through thicker interstellar gas clouds, letting more cosmic rays into the atmosphere.[435] On geological scales, this might change climate patterns and promote cloud formation via atmospheric ionization. These elevated radiation levels may accelerate biological evolution by raising the rates of mutations, some of which may be detrimental and others of which may be beneficial.[436]

6.4 Long-Term Geological and Cosmic Changes (10,000 AD–1 million AD). Although it is a relatively little era in Earth's whole geological history, the time between 10,000 and 1 million AD was sufficient for significant changes to the planet's ecology and surface. Earth's tectonic plates will continue to shift throughout the course of the next hundreds of thousands to millions of years, progressively forming a new supercontinent.[437] Ocean circulation, temperature zones, and ecosystems will undergo significant changes as a result of this event, which happens around every 300–500 million years. Continued tectonic plate movement will result in slow continental drift. Even if it moves slowly, this movement over a million years may change coastlines, create new mountain ranges, and cause changes in the sites of earthquakes and volcanoes.[438] Natural cycles of warming and cooling will occur in the Earth's climate system. It is expected that there will be many ice ages over this time, which will result in widespread glaciation, increasing and declining sea levels, and major changes in the patterns of erosion and sedimentation. Global ice volume fluctuations will significantly alter coasts around the globe. Areas that are now close to sea level may become dry ground during glacial maximums or submerged during warm seasons.[439] Large meteor strikes, enormous volcanic eruptions, significant earthquakes, and large-scale landslides are examples of uncommon occurrences whose likelihood grows over a million-year timeline, but their precise timing is uncertain. The Sun's activity and the planets' orbits will mostly not change from what they are now. For billions of years, the Sun will remain in its main sequence phase. The Milky Way galaxy's stars are always moving. As stars move out of our line of sight, the constellations that are visible from Earth will gradually change form over a million years, becoming somewhat different from the patterns we see now.[440] The solar and terrestrial magnetic fields will continue to control variations in the flow of cosmic rays that reach Earth, which over extended periods of time may have slight effects on the atmosphere's chemistry and temperature. While volcanic activity continues across tectonic boundaries, erosion continues to sculpt the surface. Continental drift may ultimately cause the Atlantic Ocean to shut, resulting in significant collisions between landmasses and the creation of tall mountain ranges. At the same time, the Sun will keep becoming brighter.[441]

Its enhanced radiation will gradually evaporate ocean water, transforming most of Earth into a desert environment over time, even while it is still in its main-sequence phase. In the larger scheme of things, the Milky Way galaxy is headed toward merging with the Andromeda galaxy; this process is already in progress

[434] Barnard, L., Lockwood, M., Hapgood, M.A., Owens, M.J., Davis, C.J. and Steinhilber, F., 2011, “Predicting space climate change”, Geophys. Res. Lett., 38, L16103, accessed on 03 Feb 2026

[435] Bezuglov, M.V., Malyshevsky, V.S., Malykhina, T.V., Torgovkin, A.V., Fomin, G.V. and Shramenko, B.I., 2012, “Photonuclear channel of ^{7}Be production in the Earth’s atmosphere”, Phys. Atom. Nucl., 75, 393–397, accessed on 03 Feb 2026

[436] Haigh, J.D., Winning, A.R., Toumi, R. and Harder, J.W., 2010, “An influence of solar spectral variations on radiative forcing of climate”, Nature, 467, 696–699, accessed on 03 Feb 2026

[437] Rescher, Nicholas (1998). Predicting the future: An introduction to the theory of forecasting. State University of New York Press. ISBN 978-0-7914-3553-3, accessed on 03 Feb 2026

[438] Stars in the Milky Way galaxy are in constant motion. Over a million years, the constellations visible from Earth will subtly change shape as stars drift from our perspective, making them slightly different from the patterns we see today, accessed on 03 Feb 2026

[439] Adams, Fred C.; Laughlin, Gregory (1 April 1997). "A dying universe: the long-term fate and evolution of astrophysical objects" (PDF). Reviews of Modern Physics. 69 (2): 337–372, doi:10.1103/RevModPhys.69.337. ISSN 0034-6861, accessed on 03 Feb 2026

[440] Blackman, J. W.; Beaulieu, J. P.; Bennett, D. P.; Danielski, C.; et al. (13 October 2021). "A Jovian analogue orbiting a white dwarf star". Nature. 598 (7880): 272–275, doi:10.1038/s41586-021-03869-6, accessed on 03 Feb 2026

[441] Ferreira, Becky (13 October 2021). "Astronomers Found a Planet That Survived Its Star's Death – The Jupiter-size planet orbits a type of star called a white dwarf, and hints at what our solar system could be like when the Sun burns out". The New York Times. Archived from the original on 28 December 2021, accessed on 03 Feb 2026

and should be completed in billions of years.[442] The center supermassive black holes will combine into a single, bigger entity as a result of this merger, which will also drastically restructure star systems and start fresh waves of star formation. Planetary systems will be rearranged due to gravitational perturbations brought on by these galactic movements.[443]

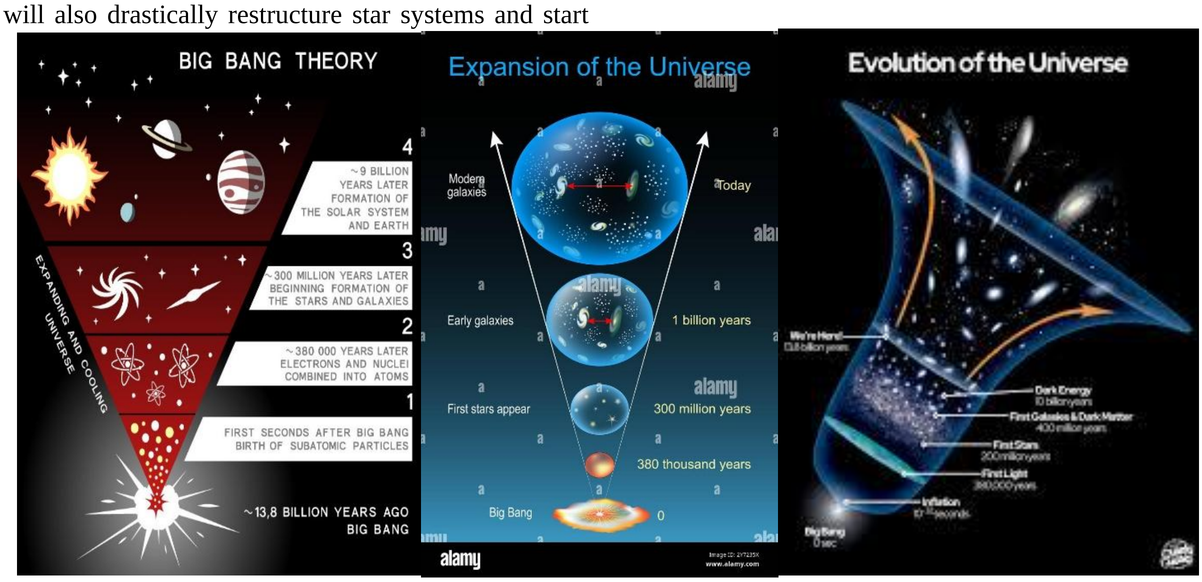


Figure 10: The Big Bang and evaluation of universe444,445,446

6.5 The Far Future of Earth and the Solar System (1 million–10 billion Years).

Mass extinctions might result from hits from asteroids the size of the Chicxulub impactor over this enormous duration.[447] But the Sun will undergo the biggest change. The Sun will run out of hydrogen fuel in around 5 billion years and turn into a red giant, swallowing all the inner planets, including Earth, in the process.[448] Earth will be burned and devoid of its atmosphere if it is not eaten, making it uninhabitable. A white dwarf—a dense, cooling star remnant—will be left behind when the Sun sheds its outer layers, creating a stunning but transient planetary nebula.[449] This white dwarf will continue to fade and lose heat over billions of years. As the solar system progressively turns into a cold, inert region, the gas giants and Kuiper Belt objects will freeze and their orbits may change.[450] The solar system will essentially be dead—silent, frozen, and drifting across the dark expanse of the galaxy—when the white dwarf cools into a black dwarf in around 10 billion years.[451]

6.6 The Degenerate Era and the Dark Universe (10 trillion–10^150 Years).

When all stars have run out of fuel, this period—known as the degenerate era—will start. The absence of hydrogen gas will prevent the formation of new stars.[452] Red dwarf stars are now the longest-lived, and all surviving white dwarfs will cool into black dwarfs. Eventually, it will burn out. Galaxies will spread over billions of years as gravitational forces propel stars and

[442] Komatsu, E.; Smith, K. M.; Dunkley, J.; Bennett, C. L.; et al. (2011). "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation". The Astrophysical Journal Supplement Series. 192 (2): 18. arXiv:1001.4731. Bibcode:2011ApJS.192...19W. doi:10.1088/0067-0049/192/2/18, accessed on 03 Feb 2026

[443] Landstreet, John D. (2003). Physical Processes in the Solar System: An introduction to the physics of asteroids, comets, moons and planets. Keenan & Darlington. p. 121. ISBN 978-0-9732051-0-7, accessed on 03 Feb 2026

[444] https://www.lode.de/blog/the-big-bang, accessed on 03 Feb 2026

[445]https://www.alamy.com/evolution-of-the-universe-universe-expansion-cosmic-timeline-and-evolution-of-stars-galaxy-and-universe-after-big-bang-general-relativity-illust-image623922262.html, accessed on 03 Feb 2026

[446] https://www.facebook.com/groups/physicsisfun109/posts/529983169680532/, accessed on 07 Nov 2025

[447] Haddok, Eitan (29 September 2008). "Birth of an Ocean: The Evolution of Ethiopia's Afar Depression". Scientific American. Archived from the original on 24 December 2013, accessed on 03 Feb 2026

[448] Berski, Filip; Dybczyński, Piotr A. (November 2016). "Gliese 710 will pass the Sun even closer: Close approach parameters recalculated based on the first Gaia data release". Astronomy & Astrophysics. 595: L10, doi:10.1051/0004-6361/201629835. ISSN 0004-6361, accessed on 03 Feb 2026

[449] Horner, J.; Evans, N. W.; Bailey, M. E. (2004). "Simulations of the Population of Centaurs I: The Bulk Statistics". Monthly Notices of the Royal Astronomical Society. 354 (3): 798–810, doi:10.1111/j.1365-2966.2004.08240.x, accessed on 03 Feb 2026

[450] Finkleman, David; Allen, Steve; Seago, John; Seaman, Rob; et al. (June 2011). "The Future of Time: UTC and the Leap Second". American Scientist. 99 (4): 312. arXiv:1106.3141. Bibcode:2011arXiv1106.3141F. doi:10.1511/2011.91.312, accessed on 07 Nov 2025

[451] Bilham, Roger (November 2000). "NOVA Online | Everest | Birth of the Himalaya". pbs.org. Archived from the original on 19 June 2021, accessed on 07 Nov 2025

[452] Loeb, Abraham; Batista, Rafael; Sloan, W. (2016). "Relative Likelihood for Life as a Function of Cosmic Time". Journal of Cosmology and Astroparticle Physics. 2016 (8): 040. arXiv:1606.08448. Bibcode:2016JCAP...08.040L. doi:10.1088/1475-7516/2016/08/040, accessed on 07 Nov 2025

debris into intergalactic space.[453] As they merge and devour surrounding matter, black holes emerge as the dominating cosmic formations. Even they eventually start to evaporate due to Hawking radiation, a gradual process that may take up to 1010010^ {100}10100 years or more. Cosmic microwave background radiation will eventually become undetectable and the cosmos will be a sparse, dark environment full of low-energy particles. Time will become irrelevant in any meaningful sense when particle interactions become so uncommon.[454] The universe will achieve total thermodynamic equilibrium, or maximum entropy.

6.7 Speculative Scenarios Beyond the Heat Death (Post-10^150 Years). According to speculative physics ideas, quantum fluctuations may sometimes result in transient occurrences at this last stage, such as the creation of new particles or even black holes.[455] Such events are very uncommon, however. Whole universes could be recreated by quantum tunneling or vacuum decay, according to certain ideas like everlasting inflation and Poincare recurrence. This might potentially begin cosmic development with a new Big Bang.[456] Boltzmann brains and self-aware creatures that momentarily exist as a result of random fluctuations are further possibilities.[457] The problems of time, entropy, and awareness in endless worlds are highlighted by these strange beings.[458] At the smallest scales, quantum foam might persist as a dynamic, flickering background and possibly the final trace of physical activity in an otherwise static universe.[459]

[453] Adams, F. C.; Bodenheimer, P.; Laughlin, G. (2005). "M dwarfs: planet formation and long-term evolution". Astronomische Nachrichten. 326 (10): 913–919. Bibcode:2005AN....326.913A. doi:10.1002/asna.200510440, accessed on 07 Nov 2025

[454] Tayler, Roger John (1993). Galaxies, Structure and Evolution (2nd ed.). Cambridge University Press. p. 92. ISBN 978-0-521-36710-3, accessed on 03 Feb 2026

[455] Douglas, M. (21 March 2003). "The statistics of string / M theory vacua". JHEP. 0305 (46): 046. arXiv:hep-th/0303194. Bibcode:2003JHEP...05.046D. doi:10.1088/1126-6708/2003/05/046, accessed on 07 Nov 2025

[456] https://ludus-vitalis.org/html/textos/30/30-03_kwiatkowska_szatzschneider.pdf, accessed on 06 Dec 2025

[457] https://ui.adsabs.harvard.edu/abs/2002nap.book10248F/, accessed on 03 Feb 2026

[458] Linde, Andrei (2007). "Sinks in the landscape, Boltzmann brains and the cosmological constant problem". Journal of Cosmology and Astro-particle Physics. 2007 (1): 022, doi:10.1088/1475-7516/2007/01/022. ISSN 1475-7516, accessed on 03 Feb 2026

[459] Andreassen, Anders; Frost, William; Schwartz, Matthew D. (12 March 2018). "Scale-invariant instantons and the complete lifetime of the standard model". Physical Review D. 97 (5) 056006, doi:10.1103/PhysRevD.97.056006, accessed on 07 Nov 2025

Figure 11: We know very less about dark matter and dark energy460

7. Scientific Theories about the Future of Universe and Contradiction

7.1 The Big Bang nucleosynthesis. The idea that explains how the elements formed in the early cosmos is called Big Bang nucleosynthesis. When the cosmos was about three minutes old and its temperature fell below the point at which nuclear fusion could take place, it came to an end. Only the lightest elements were created since Big Bang nucleosynthesis could only function for a short time. It mostly created deuterium, helium-4, and lithium from hydrogen ions or protons. Only tiny amounts of other elements were generated. In 1948, Robert Herman, Ralph Asher Alpher, and George Gamow devised the fundamental theory of nucleosynthesis.[461] Because the Big Bang nucleosynthesis hypothesis links the abundances of primordial light elements with the characteristics of the early universe, it was used for many years as a physics probe during the Big Bang.[462] It may be used to test the equivalency principle in particular,[463] to test neutrino physics and investigate dark matter.[464] According to certain cosmologists, Big Bang nucleosynthesis indicates the existence of a fourth "sterile" neutrino type.[465]

One of the biggest endeavors in cosmology is to comprehend the creation and development of the biggest and oldest structures, including as quasars, galaxies, clusters, and superclusters. Cosmologists investigate a concept of hierarchical structure creation in which the biggest items, such superclusters, are still assembling while smaller objects form first.[466] Surveying the visible galaxies is one approach to quantify the matter power spectrum and create a three-dimensional image of the universe's galaxies. The Sloan Digital Sky Survey and the 2dF Galaxy Redshift Survey take this tack.[467] Cosmologists employ simulations to examine the gravitational aggregation of matter in the cosmos as it clusters into filaments, superclusters, and voids. This

[460] https://www.discovery.com/science/we-have-liftoff--congratulations-to-nasa-and-spacex, accessed on 06 Dec 2025
[461] Peebles, Phillip James Edwin (April 2014). "Discovery of the hot Big Bang: What happened in 1948". The European Physical Journal H. 39 (2): 205–223, doi:10.1140/epjh/e2014-50002-y, accessed on 06 Dec 2025
[462] Burles, Scott; Nollett, Kenneth M.; Turner, Michael S. (May 2001). "Big Bang Nucleosynthesis Predictions for Precision Cosmology". The Astrophysical Journal. 552 (1): L1 – L5, doi:10.1086/320251, accessed on 03 Feb 2026
[463] Boucher, V.; Gérard, J.-M.; Vandergheynst, P.; Wiaux, Y. (November 2004). "Cosmic microwave background constraints on the strong equivalence principle". Physical Review D. 70 (10) 103528, doi:10.1103/PhysRevD.70.103528, accessed on 03 Feb 2026
[464] Cyburt, Richard H.; Fields, Brian D.; Olive, Keith A.; Yeh, Tsung-Han (January 2016). "Big bang nucleosynthesis: Present status". Reviews of Modern Physics. 88 (1) 015004, doi:10.1103/RevModPhys.88.015004, accessed on 06 Dec 2025
[465] Lucente, Michele; Abada, Asmaa; Arcadi, Giorgio; Domcke, Valerie (March 2018). "Leptogenesis, dark matter and neutrino masses". arXiv:1803.10826, accessed on 06 Dec 2025
[466] Heß, Steffen; Kitaura, Francisco-Shu; Gottlöber, Stefan (November 2013). "Simulating structure formation of the Local Universe". Monthly Notices of the Royal Astronomical Society. 435 (3): 2065–2076. arXiv:1304.6565. Bibcode:2013MNRAS.435.2065H. doi:10.1093/mnras/stt1428, accessed on 03 Feb 2026
[467] Cole, Shaun; Percival, Will J.; Peacock, John A.; Norberg, Peder; Baugh, Carlton M.; Frenk, Carlos S.; et al. (2005). "The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications". Monthly Notices of the Royal Astronomical Society. 362 (2): 505–534, doi:10.1111/j.1365-2966.2005. 09318.x, accessed on 03 Feb 2026

is another way to comprehend structure development.[468] Since there is a lot more-dark matter in the universe than visible, baryonic matter, the majority of simulations only include non-baryonic cold dark matter, which should be enough to comprehend the world on the biggest scales. More sophisticated models are beginning to investigate the origin of individual galaxies and include baryons. Cosmologists examine these simulations to see if they match the galaxy surveys and to interpret any differences.[469]

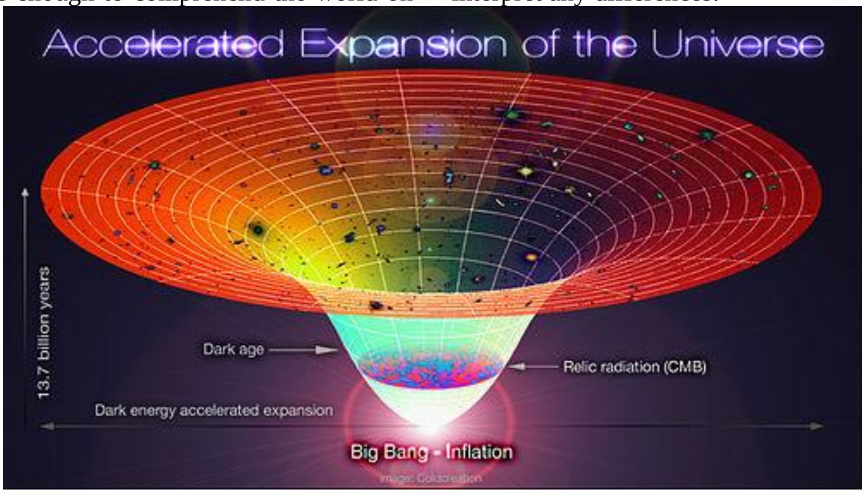


Figure 12: Lambda-CDM, accelerated expansion of the universe470

The cosmic microwave background, structure creation, Big Bang nucleosynthesis, and galaxy rotation curves all provide evidence that non-baryonic dark matter makes up around 23% of the universe's mass, whereas visible, baryonic matter makes up just 4%.[471] Because dark matter acts like a cool, non-radiative fluid that forms haloes surrounding galaxies, its gravitational effects are well known. Dark matter's particle physics nature is still entirely unknown, and it has never been found in a lab. A stable supersymmetric particle, a huge particle that interacts weakly, a big particle that interacts gravitationally, an axion, and a massive compact halo object are some of the possible options in the absence of observational constraints. An effect from brane cosmology or a change of gravity at tiny accelerations (MOND) are alternatives to the dark matter theory. Gravitational lensing may be explained by TeVeS, a kind of MOND. On the other hand, a hypothesis called Modified Newtonian Dynamics (MOND) suggests altering Newton's equations to take into consideration the observable characteristics of galaxies.[472] Modified gravity is the outcome of changing Newton's law of gravity, while modified inertia is the effect of changing Newton's second law.[473] Compared to the adjusted gravity version, the latter has not gotten as much attention. One of the most well-known ideas in this class, its main goal is to explain galaxy rotation curves without the need for dark matter. MOND was created by Israeli scientist Mordehai Milgrom in 1982 and debuted in 1983.[474]

In addition to the 23% dark matter and 4% baryons that make up the universe's energy density, there must be another component that accounts for 73% of the universe's energy density if it is flat. We refer to this as dark energy. It must not cluster in haloes like dark matter and baryons to avoid interfering with Big Bang nucleosynthesis and the cosmic microwave background. Strong observational evidence for dark energy exists since the universe's entire energy density is known because to limitations on its flatness, yet the quantity of clumping matter is precisely measured and far lower than this. In 1999, observations showed that the universe's expansion had started to progressively accelerate, supporting the theory for dark energy.[475] Dark energy is unknown except for its density and clustering characteristics. The cosmological constant (CC), which

[468] Percival, Will J.; et al. (2007). "The Shape of the Sloan Digital Sky Survey Data Release 5 Galaxy Power Spectrum". The Astrophysical Journal. 657 (2): 645–663. arXiv:astro-ph/0608636. Bibcode:2007ApJ...657.645P. doi:10.1086/510615, accessed on 03 Feb 2026

[469] Kuhlen, Michael; Vogelsberger, Mark; Angulo, Raul (November 2012). "Numerical simulations of the dark universe: State of the art and the next decade". Physics of the Dark Universe. 1 (1–2): 50–93, doi:10.1016/j.dark.2012.10.002, accessed on 06 Dec 2025

[470] Weinberg, Steven (2008). Cosmology. Oxford University Press. ISBN 978-0-19-852682-7, accessed on 06 Dec 2025

[471] Schramm, David N.; Turner, Michael S. (1998-01-01). "Big-bang nucleosynthesis enters the precision era". Reviews of Modern Physics. 70 (1): 303–318, doi:10.1103/RevModPhys.70.303, accessed on 03 Feb 2026

[472] Andrei, Cosmin; Ijjas, Anna; Steinhardt, Paul J. (12 April 2022). "Rapidly descending dark energy and the end of cosmic expansion". Proceedings of the National Academy of Sciences. 119 (15) e2200539119, doi:10.1073/pnas.2200539119. ISSN 0027-8424, accessed on 03 Feb 2026

[473] Jacob Bekenstein & M. Milgrom (1984). "Does the missing mass problem signal the breakdown of Newtonian gravity?". Astrophys. J. 286: 7–14, doi:10.1086/162570, accessed on 06 Dec 2025

[474] Milgrom, M. (1983). "A modification of the Newtonian dynamics as an alternative to the hidden mass hypothesis". Astrophysical Journal. 270: 365–370, doi:10.1086/161130, accessed on 06 Dec 2025

[475] Perlmutter, Saul; Turner, Michael S.; White, Martin (1999). "Constraining Dark Energy with Type Ia Supernovae and Large-Scale Structure". Physical Review Letters. 83 (4): 670–673, doi:10.1103/PhysRevLett.83.670, accessed on 06 Dec 2025

is 120 orders of magnitude greater than the measured value, is predicted by quantum field theory to be similar to dark energy.[476] Several string theorists, including Steven Weinberg, have cited the "weak anthropic principle."[477] In this case, physicists see a world with a tiny cosmological constant since a universe with a bigger cosmological constant would not allow for the existence of physicists or any life.[478] The weak anthropic principle is self-evident (there must be at least one universe with a cosmological constant (CC) that permits life to exist given the existence of living observers), but it does not attempt to explain the context of that universe, which is why many cosmologists find this explanation unsatisfactory.[479] The question of the universe's ultimate destiny is probably going to be resolved with a deeper comprehension of dark energy. The fast expansion caused by dark energy is inhibiting the formation of formations bigger than superclusters in the present cosmic epoch. It is uncertain whether the acceleration will ultimately reverse, result in a Big Freeze, or follow some other scenario, or if it will continue continuously, perhaps even rising till a Big Rip.[480]

[476] Adler, Ronald J.; Casey, Brendan; Jacob, Ovid C. (July 1995). "Vacuum catastrophe: An elementary exposition of the cosmological constant problem". American Journal of Physics. 63 (7): 620–626. Bibcode:1995AmJPh. 63..620A. doi:10.1119/1.17850, accessed on 03 Feb 2026

[477] Schombert, James. "Anthropic principle". Department of Physics at University of Oregon. Archived from the original on 2012-04-28, accessed on 03 Feb 2026

[478] Delia Schwartz-Perlov & Alexander Vilenkin (2006). "Probabilities in the Bousso-Polchinski multiverse". Journal of Cosmology and Astroparticle Physics. 0606 (6): 010, doi:10.1088/1475-7516/2006/06/010, accessed on 06 Dec 2025

[479] Siegfried, Tom (11 August 2006). "A 'Landscape' Too Far?". Science. 313 (5788): 750–753. doi:10.1126/science.313.5788.750, accessed on 30 Nov 2025

[480] Fernández-Jambrina, L. (September 2014). "Grand rip and grand bang/crunch cosmological singularities". Physical Review D. 90 (6) 064014. arXiv:1408.6997. Bibcode:2014PhRvD.90f4014F. doi:10.1103/PhysRevD.90.064014, accessed on 03 Feb 2026

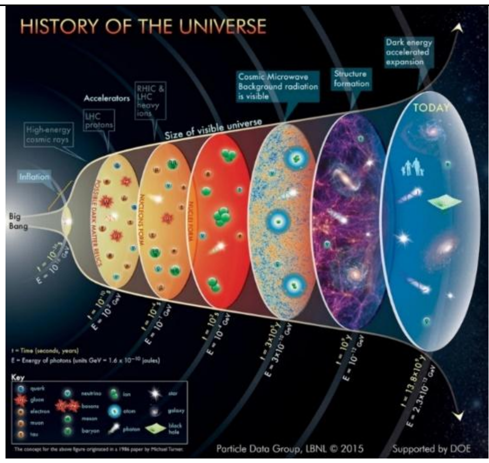


Figure 13: History of universe481

Physics is still unsure about the existence and fate of primordial black holes in our universe.[482] Cosmologists must find cosmic rays with energies higher than the GZK cutoff and determine if this indicates a high-energy breakdown of special relativity.[483] The Greisen–Zatsepin–Kuzmin limit, also known as the GZK cutoff or GZK limit, is a theoretical upper limit on the energy of cosmic ray protons that travel to our galaxy from other galaxies via the intergalactic medium. They must ascertain the concept of equivalency,[484] if Einstein's general theory of relativity is the best way to explain gravity,[485] and if the fundamental laws of physics governing the cosmos are universal.[486] Cosmologists must explain biophysical cosmology, a branch of physical cosmology that examines and comprehends life as both an intrinsic and a component of physical cosmology. It emphasizes that since life is a natural part of the cosmos, it occurs often.[487],[488] Different scientific theories have projected a number of conceivable futures, such as the possibility that the cosmos existed for a limited or infinite amount of time or to explain how and under what conditions it began.[489] The density of the cosmos may influence its destiny. The "Big Freeze" scenario below results from the majority of data to date, which is based on measurements of the mass density and the rate of expansion, favoring a universe that will continue to expand endlessly.[490] Alternative hypotheses are still conceivable, however, and findings are not definitive. By "weighing" the universe—for instance, by calculating the respective contributions of matter, radiation, dark matter, and dark energy to the critical density—one might choose between these difficult possibilities (discussed below). More specifically, opposing hypotheses are assessed in light of information on distant supernovas, galaxy clustering, and cosmic microwave background anisotropies.[491]

7.2 The Big Freeze or Heat Death. The conflict between the pull or push of gravity and the velocity of expansion determines how the cosmos evolves. The Hubble Constant measures the present rate of expansion, while the density and pressure of the universe's matter determine how strong gravity is.[492] The density

[481] https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcQugbskfiB8KnCrD8M85sPWiRpJ0S79eCcH9O-nJ0DpRQ&s, accessed on 03 Feb 2026

[482] Kovetz, Ely D. (2017). "Probing Primordial Black Hole Dark Matter with Gravitational Waves". Physical Review Letters. 119 (13)131301, doi:10.1103/PhysRevLett.119.131301, accessed on 30 Nov 2025

[483] Takeda, M.; et al. (10 August 1998). "Extension of the Cosmic-Ray Energy Spectrum beyond the Predicted Greisen-Zatsepin-Kuz'min Cutoff". Physical Review Letters. 81 (6): 1163–1166, doi:10.1103/PhysRevLett.81.1163, accessed on 30 Nov 2025

[484] Boucher, V.; Gérard, J.-M.; Vandergheynst, P.; Wiaux, Y. (November 2004). "Cosmic microwave background constraints on the strong equivalence principle". Physical Review D. 70 (10) 103528, doi:10.1103/PhysRevD.70.103528, accessed on 30 Nov 2025

[485] Turyshev, Slava G. (2008). "Experimental Tests of General Relativity". Annual Review of Nuclear and Particle Science. 58 (1): 207–248. arXiv:0806.1731. Bibcode:2008ARNPS. 58..207T. doi:10.1146/annurev.nucl.58.020807.111839, accessed on 04 Feb 2026

[486] Uzan, Jean-Philippe (March 2011). "Varying Constants, Gravitation and Cosmology". Living Reviews in Relativity. 14 (1) 2. arXiv:1009.5514. Bibcode:2011LRR....14....2U. doi:10.12942/lrr-2011-2, accessed on 04 Feb 2026

[487] Dick, Steven J. (2020). "The Biophysical Cosmology: The Place of Bioastronomy in the History of Science". Space, Time, and Aliens. Cham: Springer International Publishing. pp. 53–58. doi:10.1007/978-3-030-41614-0_4, accessed on 30 Nov 2025

[488] Mack, Eric (March 19, 2015). "There may be more Earth-like planets than grains of sand on all our beaches – New research contends that the Milky Way alone is flush with billions of potentially habitable planets – and that's just one sliver of the universe". CNET. Archived from the original on December 1, 2023, accessed on 04 Feb 2026

[489] Pogge, Richard W. (February 24, 2014). ‘Essay: The Folly of Giordano Bruno" by Thomas Digges in 1576 in his A Perfit Description of the Caelestial Orbes, in which Digges both presents and extends the Copernican system, suggesting that the Universe was infinite., accessed on 04 Feb 2026

[490] https://map.gsfc.nasa.gov/universe/uni_fate.html, accessed on 04 Feb 2026

[491] Hu, Wayne; Dodelson, Scott (September 2002). "Cosmic Microwave Background Anisotropies". Annual Review of Astronomy and Astrophysics. 40 (1): 171–216, doi:10.1146/annurev.astro.40.060401.093926,

[492] https://map.gsfc.nasa.gov/universe/uni_expansion.html, accessed on 04 Feb 2026

determines the destiny of the cosmos if the matter's pressure is low, as it is for the majority of matter types that we are aware of. The cosmos will expand indefinitely if its density is lower than the critical density, as shown by the green or blue arcs in the graph above.[493] Gravity may eventually limit the rate of expansion, but at densities below the critical density, the material's gravitational force is insufficient to halt or reverse the outward expansion. The cosmos will gradually chill as it expands until it can no longer support any life, which is why it is sometimes referred to as the Big Chill or Big Freeze.

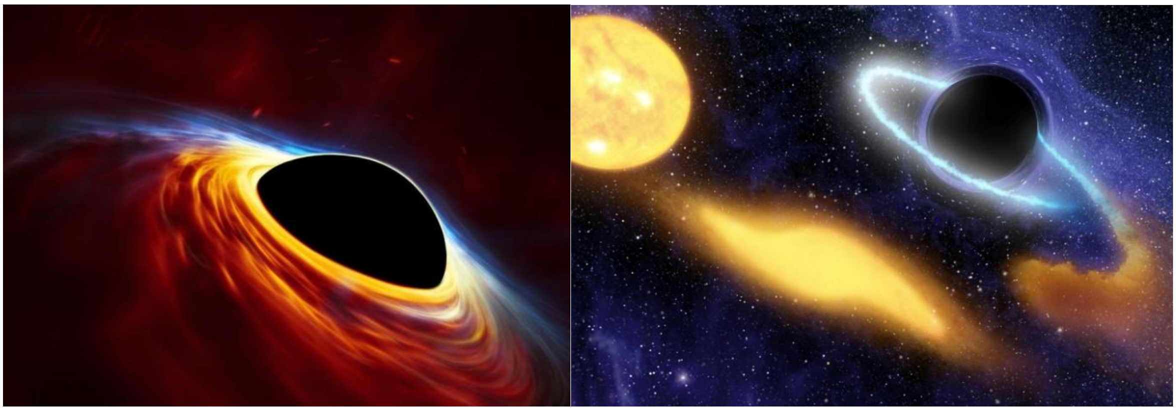


Figure 14: Black holes don't emit light and they might make dark matter shine494,495

The Big Freeze (or Big Chill), also referred to as the "heat death of the universe," is a scenario in which the universe's temperature asymptotically approaches absolute zero due to ongoing expansion.[496] According to this theory, the world ultimately achieves a state of maximum entropy, where everything is distributed equally and there are no energy gradients—all of which are necessary to support information processing, including life. This situation is now more likely to occur.[497] The gas supply required for star formation will ultimately run out, although stars are predicted to develop properly for 1012–1014 (1–100 trillion) years.[498] The cosmos will gradually and unavoidably get darker when current stars run out of fuel and stop shining. Black holes will eventually rule the cosmos, but since they release Hawking radiation, they will eventually vanish.[499] According to the Poincaré recurrence theorem, thermal fluctuations may cause a spontaneous entropy drop over an unlimited amount of time,[500],[501] and the fluctuation theorem.[502],[503] Any of the three spatial models can accommodate the heat death scenario, but it necessitates that the universe eventually reach a minimum temperature.[504] It could only take place in a flat or hyperbolic geometry in the absence of dark energy. It may also happen in a closed universe with a positive cosmological constant.

[493] https://map.gsfc.nasa.gov/universe/bb_concepts.html, accessed on 04 Feb 2026

[494] https://www.forbes.com/sites/startswithabang/2020/07/07/could-black-holes-be-the-dark-matter-our-universe-needs/, accessed on 10 Feb 2026

[495] https://www.space.com/29382-black-holes-make-dark-matter-shine.html, accessed on 10 Feb 2026

[496] Glanz, James (1998). "Breakthrough of the year 1998. Astronomy: Cosmic Motion Revealed". Science. 282 (5397): 2156–2157. Bibcode:1998Sci...282.2156G. doi:10.1126/science.282.5397.2156a, accessed on 30 Nov 2025

[497] Wang, Yun; Kratochvil, Jan Michael; Linde, Andrei; Shmakova, Marina (2004). "Current observational constraints on cosmic doomsday". Journal of Cosmology and Astro-Particle Physics. 2004 (12): 006, doi:10.1088/1475-7516/2004/12/006, accessed on 30 Nov 2025

[498] Stahler, S. W. & Palla, F. (2004). The Formation of Stars. Weinheim: Wiley-VCH. ISBN 3-527-40559-3, accessed on 30 Nov 2025

[499] Adams, Fred C.; Laughlin, Gregory (1997). "A dying universe: the long-term fate and evolution of astrophysical objects". Reviews of Modern Physics. 69 (2): 337–372, doi:10.1103/RevModPhys.69.337, accessed on 30 Nov 2025

[500] Tegmark, Max (May 2003). "Parallel Universes". Scientific American. 288 (5): 40–51. arXiv:astro-ph/0302131. Bibcode:2003SciAm.288e..40T. doi:10.1038/scientificamerican0503-40, accessed on 04 Feb 2026

[501] Werlang, T.; Ribeiro, G. A. P.; Rigolin, Gustavo (2013). "Interplay Between Quantum Phase Transitions and the Behavior of Quantum Correlations at Finite Temperatures". International Journal of Modern Physics B. 27: 1345032. arXiv:1205.1046. Bibcode:2013IJMPB.2745032W. doi:10.1142/S021797921345032X, accessed on 30 Nov 2025

[502] Xing, Xiu-San; Steinhardt, Paul J.; Turok, Neil (2007). "Spontaneous entropy decrease and its statistical formula". arXiv:0710.4624, accessed on 04 Feb 2026

[503] Linde, Andrei (2007). "Sinks in the landscape, Boltzmann brains and the cosmological constant problem". Journal of Cosmology and Astroparticle Physics. 2007 (1): 022, doi:10.1088/1475-7516/2007/01/022, accessed on 30 Nov 2025

[504] Yurov, A. V.; Astashenok, A. V.; González-Díaz, P. F. (2008). "Astronomical bounds on a future Big Freeze singularity". Gravitation and Cosmology. 14 (3): 205–212, doi:10.1134/S0202289308030018, accessed on 30 Nov 2025

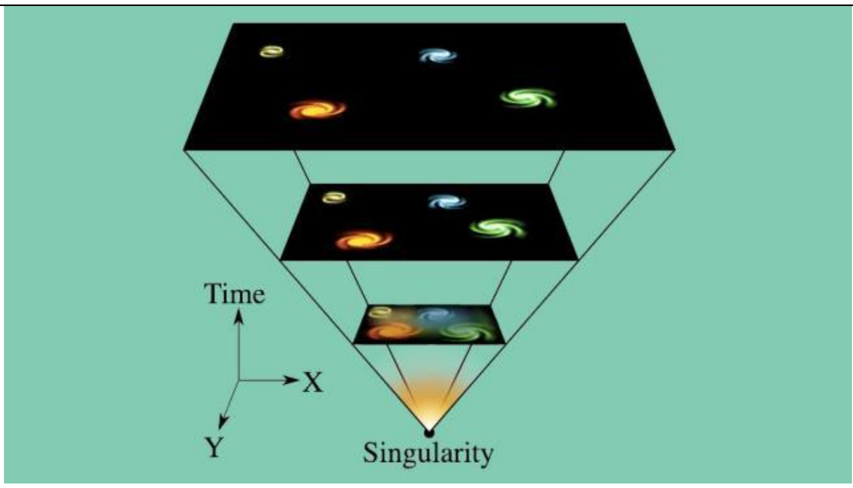


Figure 15: The Big Freeze theory suggests, the universe expansion will continue and have no end[505]

In "Time Without End: Physics and Biology in an Open Universe," published in 1979, scientist Freeman Dyson[506] suggested a scenario for the far future in which intelligent life may process an endless number of ideas to attain a type of immortality.[507] Initially, this idea—known as "Dyson's eternal intelligence"—was based on the open cosmos, a cosmological model that extends indefinitely. The universe would continue to cool as it expands, although more slowly, in the case of a zero cosmological constant, which would result in a flat or open universe devoid of dark energy.[508] According to Dyson's theory, sentient entities might store a limited quantity of energy and use it in ever decreasing amounts. These creatures would hibernate for extended periods of time after each thinking process, which allowed the cosmos to chill even further. It is theoretically possible to process an endless number of ideas over an infinite subjective time even with a finite energy reserve since the lowest energy needed for a computation (a thought) lowers as the universe's ambient temperature falls. The discovery of an accelerating expansion driven by a positive cosmological constant, however, suggests that the universe will not continue to cool indefinitely and that distant regions will become causally disconnected, which would prevent the indefinite survival Dyson envisioned.[509]

7.3 The Big Rip. The Big Rip is a hypothetical cosmological model in physical cosmology that describes the ultimate fate of the universe. According to this model, at some point in the future, dark energy's gravitational pull will gradually tear apart the universe's matter, including stars, galaxies, atoms, and subatomic particles, to the point where particle distances will infinitely increase.[510] The Big Rip is a fictitious, cataclysmic conclusion to the cosmos in which space's rapid expansion overwhelms all other forces and ultimately destroys all structures.[511] This scenario would tear apart solar systems, planets, galactic clusters, galaxies, and eventually atoms and subatomic particles. The process results in the violent and total breakdown of all matter if dark energy or "a form sometimes called phantom energy" becomes stronger over time.[512] The hypothesis is predicated on the notion that dark energy, which propels the expansion of the universe, may get more potent with time. The first consequence would be the disintegration of galaxy clusters. Individual galaxies and then solar systems would be torn apart by the expansion's strength. Planets would be torn off their stars in the last seconds, then the stars themselves.[513] Atoms would eventually lose their ability to stick together, and even subatomic particles would break apart.

[505]https://bigthink.com/guest-thinkers/the-big-freeze/, accessed on 04 Feb 2026
[506] George Dyson (June 2022). "Freeman John Dyson. 15 December 1923—28 February 2020". Biographical Memoirs of Fellows of the Royal Society. 73: 197–226. doi:10.1098/RSBM.2021.0050. ISSN 0080-4606, accessed on 30 Nov 2025
[507] Dyson, F. J. (1979). Time without end: Physics and biology in an open universe. Reviews of Modern Physics, 51(3), 447–460. https://doi.org/10.1103/RevModPhys.51.447, accessed on 30 Nov 2025
[508] Dyson, Freeman J. (1979). Disturbing the universe. New York: Harper & Row. ISBN 0-06-011108-9. OCLC 4956480, accessed on 30 Nov 2025
[509] Dyson, Freeman J. (1979-07-01). "Time without end: Physics and biology in an open universe". Reviews of Modern Physics. 51 (3): 447–460. doi:10.1103/RevModPhys.51.447, accessed on 30 Nov 2025
[510] https://www.space.com/universe-the-big-rip-can-we-stop-it, accessed on 30 Nov 2025
[511] Ellis, George F. R.; Maartens, Roy & MacCallum, Malcolm A. H. (2012). Relativistic Cosmology. Cambridge, UK: Cambridge University Press. pp. 146–147. ISBN 978-0-52138-115-4, accessed on 04 Feb 2026
[512] Vikhlinin, A.; Kravtsov, A. V.; Burenin, R. A.; et al. (2009). "Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints". The Astrophysical Journal. 692 (2): 1060–1074, doi:10.1088/0004-637X/692/2/1060, accessed on 04 Feb 2026
[513] Caldwell, Robert R.; Kamionkowski, Marc; Weinberg, Nevin N. (2003). "Phantom Energy and Cosmic Doomsday". Physical Review Letters. 91 (7) 071301. arXiv:astro-ph/0302506. Bibcode:2003PhRvL.91g1301C. doi:10.1103/PhysRevLett.91.071301, accessed on 28 Nov 2025

Figure 16: Nobody can stop the Big Rip?514

Let's imagine a time in the future when the cosmos breaks apart much too quickly. The clusters appear first; their galaxies being drawn apart. The galaxies then vanish. The planets and stellar systems will follow. Atoms themselves come next. In the end, space-time is shattered, making the cosmos uninhabitable. But the Big Rip is a possible future. This seems frightening and almost unthinkable, but what is most shocking is that there is some data that appears to be directly pointing to our demise. The present Hubble constant describes a rate of acceleration of the cosmos that is great enough to increase the distance between local objects like galaxies, which are bound together by gravity, but not huge enough to destroy them.[515] All material things in the universe, beginning with galaxies, would ultimately disintegrate into unbound elementary particles, radiation, and other forms in a limited amount of time if the Hubble constant were to steadily climb to infinity.[516] The cosmos ends as what is essentially a singularity when the energy density, scale factor, and expansion rate all reach infinite values. A more abrupt huge rip may happen in the unique scenario of phantom dark energy, which is thought to have negative kinetic energy and would cause a faster rate of acceleration than other cosmological constants predict.[517] Since the Big Rip depends on dark energy acting in a certain manner, it is a theoretical possibility rather than the most probable result. According to available evidence, dark energy is approximately constant, and this model does not anticipate a Big Rip.[518] According to Universe Today, some experimental evidence has suggested options that would increase the likelihood of this scenario, although the precise form of dark energy is yet unknown.[519]

7.4 The Big Crunch. The Big Crunch is a speculative scenario for the universe's ultimate fate in which the universe's expansion eventually reverses and re-collapses, bringing the cosmic scale factor to zero. This could lead to a reformation of the universe beginning with another Big Bang.[520] However, the overwhelming weight of available data suggests that this theory is incorrect. Rather, astronomical data indicate that the universe's expansion is speeding rather than being halted by gravity, indicating a significantly higher probability of a Big Chill or Big Rip.[521],[522] However, some physicists have suggested that a dark energy fluctuation may lead to a "Big Crunch-style" catastrophe. A symmetric perspective on the universe's ultimate destiny is the Big Crunch theory. Similar to how the Big Bang hypothesis began as a cosmic expansion, this theory postulates that the universe's average density will be sufficient to halt its expansion and cause it to start shrinking. The outcome is unclear; a straightforward

514 https://www.space.com/universe-the-big-rip-can-we-stop-it, accessed on 04 Feb 2026

515 Livio, M. (2011). "Lost in translation: Mystery of the missing text solved". Nature. 479 (7372): 171–173. Bibcode:2011Natur. 479..171L. doi:10.1038/479171a, accessed on 04 Feb 2026

516 Allen, S. W.; Rapetti, D. A.; Schmidt, R. W.; Ebeling, H.; Morris, R. G.; Fabian, A. C. (2008). "Improved constraints on dark energy from Chandra X-ray observations of the largest relaxed galaxy clusters". Monthly Notices of the Royal Astronomical Society. 383 (3): 879. arXiv:0706.0033. Bibcode:2008MNRAS. 383..879A. doi:10.1111/j.1365-2966.2007. 12610.x, accessed on 04 Feb 2026

517 Ludwick, Kevin J. (2017-09-14). "The viability of phantom dark energy: A review". Modern Physics Letters A. 32 (28): 1730025. arXiv:1708.06981. doi:10.1142/S0217732317300257. ISSN 0217-7323, accessed on 04 Feb 2026

518 https://www.americanscientist.org/article/tearing-apart-the-universe, accessed on 04 Feb 2026

519 https://www.universetoday.com/articles/what-is-the-big-rip, accessed on 04 Feb 2026

520 Wang, Yun; Kratochvil, Jan Michael; Linde, Andrei; Shmakova, Marina (2004). "Current observational constraints on cosmic doomsday". Journal of Cosmology and Astroparticle Physics. 2004 (12): 006, doi:10.1088/1475-7516/2004/12/006, accessed on 04 Feb 2026

521 Falk, Dan (22 June 2020). "This Cosmologist Knows How It's All Going to End". Quanta Magazine,

522 Perlmutter, Saul (April 2003). "Supernovas, Dark Energy, and the Accelerating Universe". Physics Today. 56 (4): 53–60. Bibcode:2003PhT....56d..53P. doi:10.1063/1.1580050, accessed on 04 Feb 2026

estimate would see the universe's matter and spacetime collapse into a dimensionless singularity, returning to the Big Bang, but at these scales, unidentified quantum phenomena must be taken into account.[523]

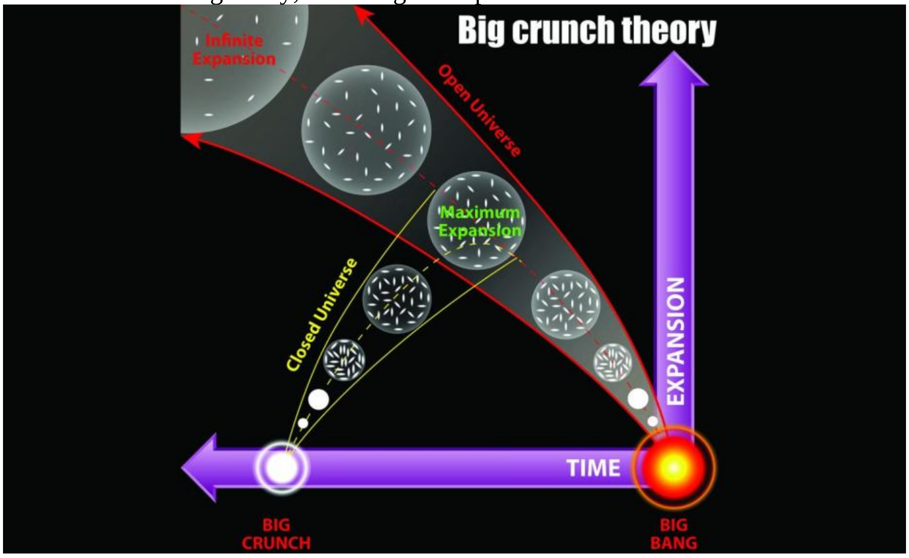


Figure 17: Universe may end with Big Crunch524

Since measurements have only been available for a brief period of time and might reverse in the future, recent information indicates that this scenario is improbable but has not been ruled out.[525] In this case, the Big Bang may happen just after a previous universe's Big Crunch. If this occurs often, a cyclic model—also referred to as an oscillatory universe—is produced.[526] The universe may thus be made up of an endless series of finite universes, each of which ends with a Big Crunch that doubles as the Big Bang of the subsequent universe. The second law of thermodynamics is incompatible with the cyclic cosmos because entropy would accumulate from oscillation to oscillation and ultimately lead to the universe's heat death.[527] The cosmos is not closed, according to current evidence. Cosmologists have given up on the oscillating universe model as a result of this. The cyclic model adopts a somewhat similar concept, but it avoids heat death because of an expansion of the branes that reduces the entropy that was created in the preceding cycle.[528]

7.5 The Big Bounce. A cosmological scenario for the beginning of the known universe is the Big Bounce theory.[529] It was initially proposed as a stage of the oscillatory universe or cyclic model explanation of the Big Bang, in which the first cosmic event was caused by the collapse of an earlier universe.[530],[531] Inflation theory arose as a solution to the horizon issue, which had arisen from improvements in observations exposing the large-scale structure of the cosmos, and it began to lose significant attention in the early 1980s.[532] A scientific theory about the origin of the known universe is called the Big Bounce. It comes from the explanation of the Big Bang known as the oscillatory universe or cyclic recurrence, in which the collapse of an earlier universe caused the initial cosmic event.[533] One interpretation of the Big Bang hypothesis of cosmology holds that the cosmos was initially endlessly dense.

[523]Rovelli, Carlo (2008). "Quantum gravity". Scholarpedia. 3 (5): 7117, doi:10.4249/scholarpedia.7117, accessed on 04 Feb 2026

[524]https://www.earth.com/news/universe-will-end-in-a-big-crunch-scientists-predict-when-cosmological-constant/, accessed on 04 Feb 2026

[525] Wang, Yun; Kratochvil, Jan Michael; Linde, Andrei; Shmakova, Marina (2004). "Current observational constraints on cosmic doomsday". Journal of Cosmology and Astro-Particle Physics. 2004 (12): 006, doi:10.1088/1475-7516/2004/12/006, accessed on 04 Feb 2026

[526] Singh, Parampreet (September–December 2014). "Loop Quantum Cosmology and the Fate of Cosmological Singularities" (PDF). Bull. Astr. Soc. India. 42 (3–4): 121. arXiv:1509.09182, accessed on 04 Feb 2026

[527] Dyer, Alan (2007-07-24). Insiders: Space. Simon & Schuster Books for Young Readers. pp. 40–41. ISBN 978-1-4169-3860-6, accessed on 04 Feb 2026

[528] Tolman, R. C. (1987) [1934]. Relativity, Thermodynamics, and Cosmology. New York: Dover. ISBN 978-0-486-65383-9. LCCN 34032023, accessed on 04 Feb 2026

[529] Abelev, B.; Adam, J.; Adamová, D.; Aggarwal, M. M.; Aglieri Rinella, G.; Agnello, M.; Agostinelli, A.; Agrawal, N.; Ahammed, Z.; Ahmad, N.; Ahmed, I.; Ahn, S. U.; Ahn, S. A.; Aimo, I.; Aiola, S. (2014-11-10). "Beauty production in pp collisions at s=2.76 TeV measured via semi-electronic decays" (PDF). Physics Letters B. 738: 97–108. doi:10.1016/j.physletb.2014.09.026, accessed on 04 Feb 2026

[530] Novello, M.; Bergliaffa, S. E. Perez (2008-07-01). "Bouncing cosmologies". Physics Reports. 463 (4): 127–213. arXiv:0802.1634. Bibcode:2008PhR...463.127N. doi:10.1016/j.physrep.2008.04.006, accessed on 04 Feb 2026

[531] Finelli, Fabio; Brandenberger, Robert (2002-05-15). "Generation of a scale-invariant spectrum of adiabatic fluctuations in cosmological models with a contracting phase". Physical Review D. 65 (10) 103522. arXiv:hep-th/0112249. Bibcode:2002PhRvD.65j3522F. doi:10.1103/PhysRevD.65.103522, accessed on 04 Feb 2026

[532] https://web.archive.org/web/20240917034537/https://press.discovery.com/us/sci/programs/how-universe-works/,

[533] Ashtekar, Abhay; Pawlowski, Tomasz; Singh, Parampreet (2 October 2006). "Quantum nature of the big bang: Improved dynamics". Physical Review D. 74 (8) 084003. arXiv:gr-qc/0607039. Bibcode:2006PhRvD.74h4003A. doi:10.1103/PhysRevD.74.084003, accessed on 04 Feb 2026

This description seems to contradict other more well recognized theories, particularly the uncertainty principle of quantum physics.[534]

Figure 18: Big Bang vs. Big Bounce?535

Consequently, a different interpretation of the Big Bang idea has emerged as a result of quantum mechanics, namely that the universe tunneled into existence and had a limited density compatible with quantum mechanics before growing in a way controlled by classical physics.[536] Also, if the universe is closed, this theory would predict that once this universe collapses it will spawn another universe in an event similar to the Big Bang after a universal singularity is reached or a repulsive quantum force causes re-expansion. In simple terms, this theory states that the universe will continuously repeat the cycle of a Big Bang, followed by a Big Crunch. Inflation was found to be inevitably eternal, creating an infinity of different universes with typically different properties, suggesting that the properties of the observable universe are a matter of chance.[537] The horizon issue was thought to have a potential predictive and falsifiable answer in the form of an alternate idea that featured a Big Bounce. As of 2022, the investigation was still ongoing.[538,539,540]

---

534 Werner Heisenberg (1989), Encounters with Einstein and Other Essays on People, Places and Particles, Princeton University Press, p. 53, accessed on 04 Feb 2026

535 https://medium.com/predict/the-big-bang-vs-big-bounce-theory-of-the-universe-8a64cbb80e3c, accessed on 04 Feb 2026

536 Halliwell, J. J. (1991). "Quantum cosmology and the creation of the universe". Scientific American. 265 (6): 76, 85. Bibcode:1991SciAm.265f..28H. doi:10.1038/scientificamerican1291-76, accessed on 04 Feb 2026

537 McKee, Maggie (25 September 2014). "Ingenious: Paul J. Steinhardt – The Princeton physicist on what's wrong with inflation theory and his view of the Big Bang". Nautilus. No. 17. NautilusThink Inc. Archived from the original on 23 January 2017, accessed on 04 Feb 2026

538 Ijjas, Anna; Steinhardt, Paul J. (January 10, 2022). "Entropy, black holes, and the new cyclic universe". Physics Letters B. 824 136823. arXiv:2108.07101. Bibcode:2022PhLB.82436823I. doi:10.1016/j.physletb.2021.136823, accessed on 04 Feb 2026

539 Lehners, Jean-Luc; Steinhardt, Paul J. (2013). "Planck 2013 results support the cyclic universe". Physical Review D. 87 (12)123533. arXiv:1304.3122. Bibcode:2013PhRvD.87l3533L. doi:10.1103/PhysRevD.87.123533. ISSN 1550-7998, accessed on 04 Feb 2026

540 Brandenberger, Robert; Peter, Patrick (2017). "Bouncing Cosmologies: Progress and Problems". Foundations of Physics. 47 (6): 797–850. arXiv:1603.05834. Bibcode:2017FoPh...47.797B. doi:10.1007/s10701-016-0057-0, accessed on 04 Feb 2026

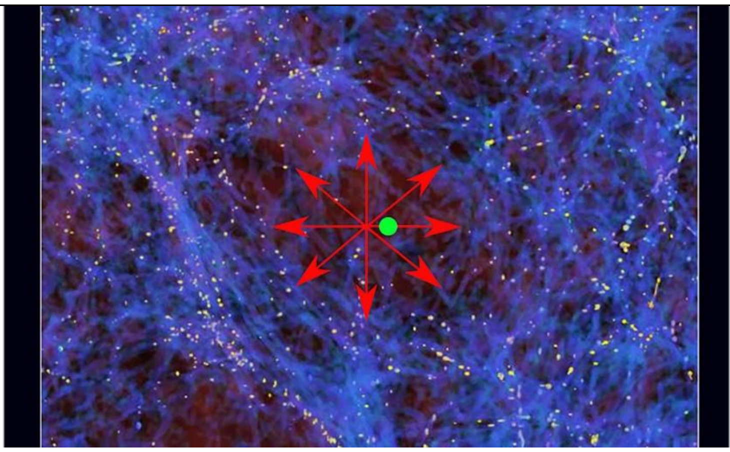


Figure 19: Some scientists believe our Milky Way galaxy may be inside a giant 'cosmic void' in the universe541

7.6 The Big Slurp. According to this idea, the cosmos is now in a false vacuum and might suddenly turn into a genuine vacuum. Understanding the Higgs field, which pervades the cosmos, is essential to comprehending the false vacuum collapse hypothesis.[542] Its potential determines its strength, much as an electromagnetic field does. As long as the cosmos is at its lowest energy state, a genuine vacuum exists; otherwise, the false vacuum idea is meaningless. The vacuum may tunnel into a lower-energy level, however, if it is not at its lowest energy state or is a fake vacuum.[543],[544] We refer to this as vacuum decay.[545] The cosmos might be drastically changed by this: in some cases, even the different physical constants could have different values, which would have a significant impact on the fundamentals of matter, energy, and spacetime. Another possibility is that all buildings will be demolished at once, with no notice.[546] However, because galaxies that are more than 4,200 megaparsecs (13 billion light-years) apart are moving away from one another faster than the speed of light, and the Big Slurp itself cannot expand faster than the speed of light, only a portion of the universe would be destroyed by the Big Slurp, leaving the majority of the universe unaffected.[547] The visible universe is presently around 46 billion light years in all directions from Earth, to put this into perspective.[548] It is believed that the universe is at least that big.

According to one of the most recent findings, our universe may be in grave danger if the Higgs boson really exists at the right mass. "If you use all the physics that we know now and you do what you think is a straightforward calculation, it's bad news," says a theoretical physicist at the Fermi National Accelerator Laboratory in Batavia, Illinois, at least. For many years, a lot of physicists have been debating the notion that our universe is on the verge of instability. In particular, physicists Frank Wilczek and Michael Turner proposed this unpleasant scenario in a 1982 paper published in Nature: "without warning, a bubble of true vacuum could nucleate somewhere in the universe and move outwards at the speed of light, and before we realized what swept by us our protons would decay away."[549] In many tens of billions of years, the universe would undergo a catastrophic event as a "bubble" formed by quantum fluctuations from some sort of alternate universe with a lower-energy state might appear in ours, expanding and ultimately destroying us, according to (Lykken) calculations (which depend on the mass of the Higgs being correct or off by a tiny margin of one percent).[550] He added, "At the speed of light, to boot." More intriguingly, his calculations raised a philosophical challenge that eventually led him to question if the condition of our universe has a role in the creation of stars, galaxies, planets, matter, and ultimately life.

541 https://www.earth.com/news/milky-way-galaxy-may-be-inside-a-giant-void-in-the-universe-explains-hubble-tension/, accessed on 04 Feb 2026

542 Goulette, Marc (15 August 2012). "What should we know about the Higgs particle?" (blog). Atlas Experiment / CERN. Archived from the original on 13 January 2022, accessed on 04 Feb 2026

543 Stone, M. (1977). "Semiclassical methods for unstable states". Phys. Lett. B. 67 (2): 186–188. Bibcode:1977PhLB...67.186S. doi:10.1016/0370-2693(77)90099-5, accessed on 06 Dec 2025

544 Coleman, S. (1977). "Fate of the false vacuum: Semiclassical theory". Physical Review D. 15 (10): 2929–36. Bibcode:1977PhRvD.15.2929C. doi:10.1103/physrevd.15.2929, accessed on 04 Feb 2026

545 Cabibbo, N.; Maiani, L.; Parisi, G.; Petronzio, R. (1979). 'Bounds on the Fermions and Higgs Boson Masses in Grand Unified Theories', accessed on 06 Dec 2025

546 Hawking, Stephen W. & Moss, I. G. (1982). "Supercooled phase transitions in the very early universe". Physics Letters B. 110 (1): 35–38. Bibcode:1982PhLB.110...35H. doi:10.1016/0370-2693(82)90946-7, accessed on 04 Feb 2026

547 Cain, Fraser; Today, Universe. "How are galaxies moving away faster than light?". phys.org. Retrieved 2023-06-15, accessed on 04 Feb 2026

548 Crane, Leah (29 June 2024). de Lange, Catherine (ed.). "How big is the universe, really?". New Scientist. p. 31,

549 Dimensionless constants, cosmology and other dark matters (with M. Tegmark, A. Aguirre, and M. Rees) astro-ph/0511774 Phys Rev D 73 N2 (2006) [MIT-CTP-3710, accessed on 04 Feb 2026

550 Hawking radiation of charged blackhole through gauge and gravitational anomaly (with Satoshi Iso, Hiroshi Umetsu) hep-th/0602146 (2006) [MIT-CTP-3714], accessed on 04 Feb 2026

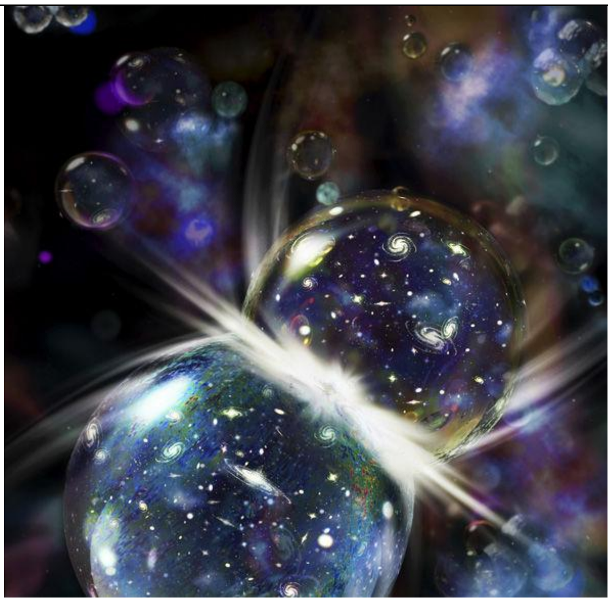


Figure 20: The ways that earth and the universe could die551

7.7 The Cosmic Uncertainty. Cosmic variance and basic physical laws are the two primary causes of the inherent limitations on our capacity to fully understand the cosmos. The statistical uncertainty that results from our limited ability to witness a single universe—a single sample with random fluctuations—is known as cosmic variance. There is a universal limit on how accurately some pairs of physical attributes (such energy and time) may be understood simultaneously due to fundamental laws like Heisenberg's uncertainty principle.[552] In reality, cosmologists are only ever able to investigate our single universe. As a result, we can only see one realization of things like the cosmic microwave background or the dispersion of galaxies on a wide scale.[553] Random quantum fluctuations in the early cosmos developed into the large-scale features we see today. This implies that the cosmos as a whole is not always represented by our observable area.[554] Measurements of cosmic attributes, such the density of matter or the features of early universe fluctuations, are always subject to statistical error due to cosmic variation. For instance, a survey area of the sky may coincidentally have a little greater or lower galaxy density.[555] Every scenario that has been discussed so far is predicated on a basic version of the dark energy equation of state. But as the name suggests, not much is currently understood about the mechanics of dark energy.[556] In the early stages of the Big Bang, the universe experienced an episode dominated by a distinct kind of dark energy if the idea of inflation is correct. However, inflation terminated, suggesting a more complicated equation of state than those thought to apply to modern dark energy. It's likely that the dark energy equation of state may shift once again, leading to an occurrence with unpredictable and hard-to-parameterize effects.[557] The possibilities of dark energy and dark matter's function in the cosmos are unclear as their nature is still mysterious, if not speculative.

[551] https://www.express.co.uk/news/science/658752/End-of-the-world-The-ways-that-Earth-and-the-universe-could-DIE, accessed on 04 Feb 2026

[552] https://physics.stackexchange.com/questions/414746/heisenbergs-uncertainty-principle-and-shape-of-universe, accessed on 04 Feb 2026

[553] https://www.researchgate.net/publication/384156316_The_Cosmic_Energy-Time_Uncertainty, accessed on 06 Feb 2026

[554] https://www.worldscientific.com/doi/10.1142/9789814355124_0014?srsltid=Afm-BOoq1VXz9zrwaWZmWvpg3YGHW6Y_ljCeOVTE8GHnIZMG-yPp4UQT3, accessed on 06 Feb 2026

[555] https://phys.org/news/2021-01-uncertainties-cosmic-expansion.html#google_vignette, accessed on 06 Feb 2025

[556] Perlmutter, S.; Turner, M.; White, M. (1999). "Constraining Dark Energy with Type Ia Supernovae and Large-Scale Structure". Physical Review Letters. 83 (4): 670–673, doi:10.1103/PhysRevLett.83.670, accessed on 06 Feb 2026

[557] Falk, Dan (2023-10-05). "The 'least crazy' idea: Early dark energy could solve a cosmological conundrum". Astronomy Magazine, accessed on 06 Feb 2026

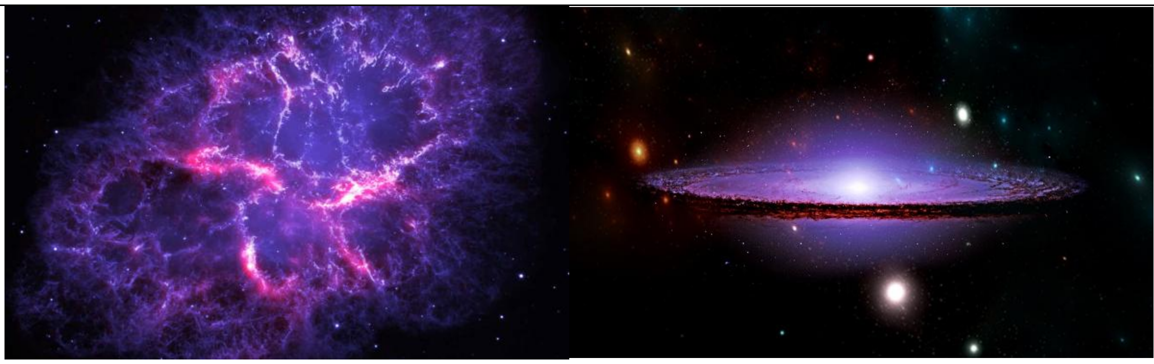


Figure 21: Cosmic uncertainty, there may antimatter worlds and Higgs particle 558,559

We will be able to make more hypotheses and verify our final line of defense in the event that the universe ends as we delve more into cosmology, bio-information theory, and the many theories of consciousness. One issue is a positive cosmic constant. Galaxies and other cosmic phenomena gradually "grow" apart due to an accelerating cosmos. There would be no way to access the energy stored in these locations. Unless, of course, the dark house—dark energy—is introduced. The nature of dark energy and its effects are mostly unknown. Thus, this is still unsettled. Going below the physical threshold required to sustain conscious awareness is the biggest barrier to attaining Dyson's perpetual intelligence. Even while organic lifeforms are undoubtedly inefficient at converting energy into computations (and awareness), it's possible that inorganic lifeforms or even other chemical architectures may produce consciousness with far less energy.

[558] https://www.newscientist.com/round-up/cosmic-uncertainty/, accessed on 06 Feb 2026
[559] https://www.livescience.com/strange-higgs-stole-antimatter-from-universe.html, accessed on 06 Feb 2026

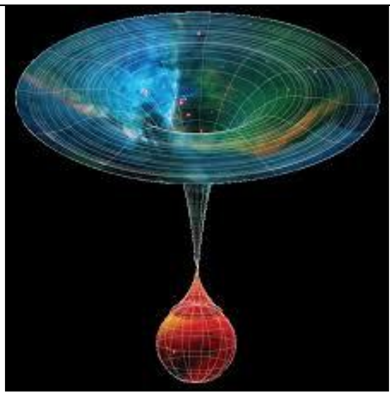


Figure 22: False vacuum decay and Big Slurp Theory560

The concept of a rigidly preset or programmed, clockwork cosmos was fatally undermined by the uncertainty principle. A quantum system's future motion can only be explained in terms of probabilities rather than absolutes as we are unable to determine its exact beginning state. Chaos theory shows that complex systems are very sensitive to beginning circumstances, even in classical, macroscopic systems, making long-term predictions—like weather forecasting or earth fast early warning—virtually unachievable. Some contemporary philosophers and physicists, such as Carlo Rovelli, investigate the nature of time itself, questioning our intuitive beliefs that the past and future are separate entities and proposing a nuanced connection between entropy, cause, and effect. In the end, physics has shifted from seeking perfect certainty to accepting reality's probabilistic character and realizing that the future is inevitably unpredictable and full of possibilities. Indeed, there are many dark rules in the environment we live in, and they undermine predicting ability by escalating mistakes and making measured quantities very susceptible to erratic outside influences.[561] Under unpredictable circumstances, there is no one solution in a dynamic environment. "Any approach to scientific inference that seeks to legitimize an answer in response to complex uncertainty is, for me, a totalitarian parody of a would-be rational learning process," as mathematician A. F. M. Smith has succinctly put it.'[562] Physics is an essential field that advances technology and transforms society in addition to expanding our knowledge of the cosmos. Physics has greatly influenced human civilization, from the discovery of basic principles to the creation of technologies that have revolutionized contemporary living. Furthermore, it tackles some of the most important issues about truth, existence, and the nature of the cosmos, inspiring interest and investigation to this day.[563]

[560] https://www.ctc.cam.ac.uk/activities/field-theory/Francesco_Muia_slides.pdf, accessed on 06 Feb 2026

[561] Robert B. Laughlin, A Different Universe. Reinventing Physics from the Bottom Down, Basic Books, A Member of the Perseus Book Group, Cambridge MA, 2006: 210, accessed on 06 Feb 2026

[562] A. F. M. Smith, "Present position and potential developments: Some personal views of Bayesian statistics," Journal of the Royal Statistical Association 147 (3): 245-259, accessed on 06 Feb 2026

[563]https://www.theguardian.com/books/2023/apr/22/physicist-michio-kaku-we-could-unravel-the-secrets-of-the-universe#:~:text=,https://www.theguardian.com/books/2023/apr/22/physicist-michio-kaku-we-could-unravel-the-secrets-of-the-universe#:~:text=, accessed on 06 Feb 2026

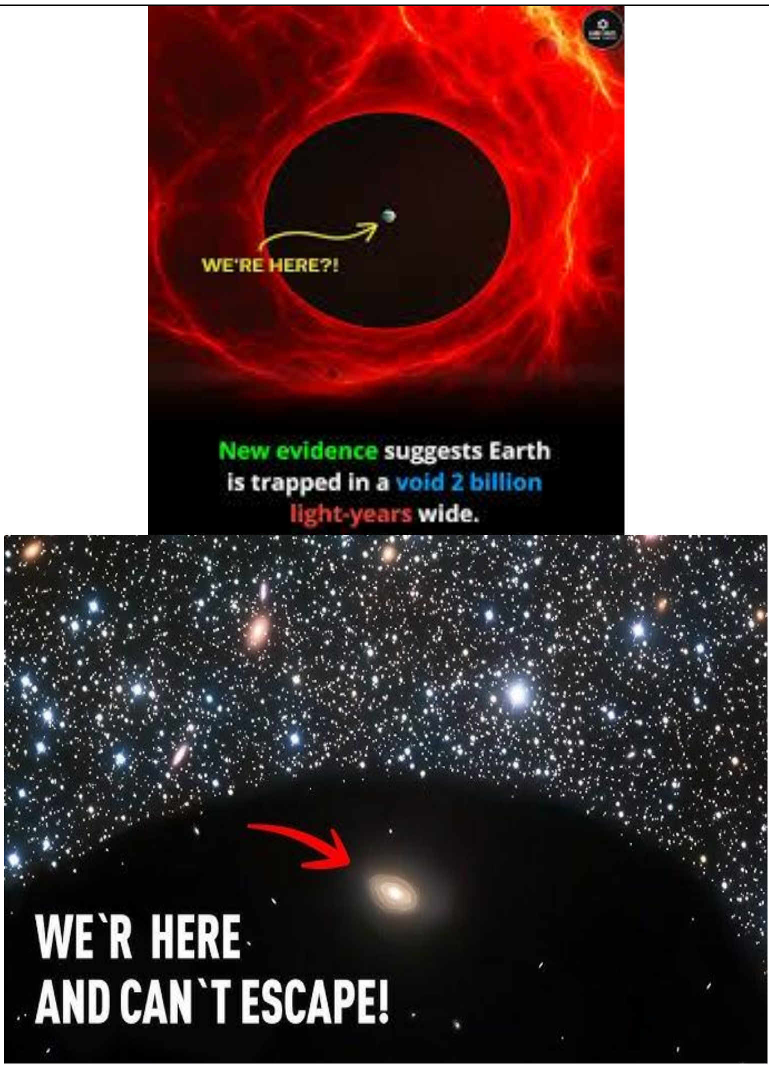


Figure 23: The Earth is trapped in a big void and can't escape[564,565]

The clockwork cosmos was blown apart by quantum theory and the general theory of relativity, which unequivocally demonstrated that human knowledge was, at best, inadequate and would likely stay that way indefinitely. Some locations in the cosmos, like black holes, might never provide any information at all. Chaos theory also illustrated how chaos is often located at the core of natural and social systems and highlighted our inherent limitations in understanding, predicting, and regulating the world around us. On the other hand, the "4D-balloon" concept, or more precisely the 3D-hypersurface of a 4D-hypersphere, is a cosmological model and pedagogical analogy used to explain the expansion and topology of the universe. It posits that our 3D universe is the "skin" of a 4D object, similar to how 2D beings would inhabit the 2D surface of a 3D balloon. However, a 3D torus, are also possible, making the 4D-balloon one of several potential, non-proven models. For decades cosmological models have loved to use a 4D hypersphere to model the universe because it perfectly seems to fit the observational data. Inappropriately, a 4D hypersphere has four spatial dimensions and our universe is limited to three, so although mathematically that model fits, intuitively for obvious reasons it's wrong since our universe has only three spatial dimensions. This model shows how a three-dimensional object can appear to be a 4-dimensional hypersphere, or hyper-shell to be more exact. This model gives an actual and very simple intuitive shape to the universe, is perfectly homogenous & isotropic, reproduces the Hubble parameter, and requires no weird dimensions.[566] Our Universe may have been 3D surface of 4D balloon. One of the goals of the science of cosmology is to discover whether our Universe is open, closed or critical.[567]

8. Conclusion

As per the most accepted theory of physics and highest number of scientists believe that, approximately 13.8 billion years ago, space and time began in a rapid, hot expansion. As the universe cooled, elementary particles formed. Around 380,000 years later, the universe cooled enough for stable atoms, mostly Hydrogen and Helium to form and releasing the 'after-

[564]https://www.facebook.com/wesaveblueplanet/posts/-astronomers-propose-we-could-live-inside-a-giant-cosmic-voida-recent-study-pres/837116198958138/, accessed on 11 Feb 2026
[565] https://www.youtube.com/watch?v=olJbmfwHm24, accessed on 11 Feb 2026
[566] https://www.facebook.com/groups/AstroCosmoNews/posts/3306929142945969/, accessed on 11 Feb 2026
[567] https://sites.ualberta.ca/~pogosyan/teaching/ASTRO_122/lect30/lecture30.html, accessed on 11 Feb 2026

glow' known as the CMB. Over billions of years, gravity caused gas to aggregate, forming the first stars and galaxies. Roughly 5 to 6 billion years ago, the universe began accelerating in its expansion, driven by a mysterious force called dark energy, which makes up about 68% of the universe. The JWST has discovered mature, massive, and highly evolved spiral galaxies existing only 200-300 million years after the Big Bang. Standard theories predicted that galaxies of this size and complexity should not have had enough time to form that early. JWST has revealed an overabundance of heavy elements like oxygen, nitrogen, and carbon in the very early universe, suggesting stellar evolution happened much faster than previously thought. There is a major conflict in science are measurements of the universe's expansion rate or Hubble constant using the early universe or CMB disagree with measurements based on modern, closer, and faster-moving objects. Some scientists suggest the universe is expanding via a 3D hypersurface of a 4D bubble, which might explain the early appearance of galaxies without requiring exotic, undiscovered dark matter particles. The concept that the Big Bang is just one of many, with different physical laws operating in other regions, is a major, though as-yet unproven, and may be alternative concept.

Our comprehension of the universe's genesis is one of physics' greatest accomplishments. Scientists have been able to investigate the underlying nature of space and time, the development of stars and galaxies, and the beginnings of the universe by using physical rules. Today, physics has made it possible for people to go beyond Earth. The concepts of classical mechanics, thermodynamics, and relativity are crucial to space exploration, which has resulted in trips to the moon, Mars, and beyond. New perspectives on planetary systems and the possibility of extraterrestrial life have been made possible by the finding of exoplanets, or planets outside of our solar system. As they look for answers to puzzles like dark matter, dark energy, and the ultimate destiny of the universe, physicists are pushing the limits of human knowledge with instruments like the Hubble Space Telescope and the soon-to-be James Webb Space Telescope. The current awareness that uncertainty is a basic, intrinsic element of nature, particularly at the quantum level, reveals a change from the classical idea of a completely predictable "clockwork" cosmos. The exact forecasting of future occurrences is inherently limited by this underlying uncertainty. Today, scientists have competing theories about how the universe ends. As per the Big Freeze theory, the universe expands forever, growing colder and more diffuse until no energy can interact. On the other hand, the Big Rip theory explain that dark energy's repulsive force increases, tearing apart galaxies, stars, and eventually atoms. At the same time the Big Crunch theory said, gravity overcomes expansion, causing the universe to collapse back into a single point.

Scientists, as enlightened philosophers believed that human reason and observation could reveal the ultimate truths of the cosmos. This new worldview has had a significant impact on art, literature, philosophy, science, and social interactions, even if we may not always be aware of it. By 'weighing' the universe—for instance, by calculating the respective contributions of matter, radiation, dark matter, and dark energy to the critical density—one might choose amongst the competing hypotheses described in this study. More specifically, data on distant supernovas, galaxy clustering, and cosmic microwave background anisotropies are used to assess difficult situations. These days, the 21st century starts with a modest acceptance of ambiguity. The book From Certainty to Uncertainty tracks the emergence and decline of the deterministic cosmos and demonstrates the changing effects that these seemingly unrelated fields now have on one another. Advanced artificial intelligence (AI), interplanetary travel, brainnets, or mind-internets, quantum computing, and the unveiling of cosmic secrets are the main topics of Michio Kaku's most recent universe predictions. Though climate change is still a direct concern, he predicts possible world ends like the Big Crunch and emphasizes how future technology might alleviate major issues, allow for human expansion like Mars colonization, and even allow direct experience sharing via brain interfaces. Thus, we may draw lessons from the past and make predictions about how we will govern our lives going forward. Physics will continue to be essential as we tackle the problems of the future, from space exploration to energy sustainability and beyond. Today, the narrative of universe is shifting from a smooth, uniform, slow-evolution model to a far more volatile, rapid-evolution, and potentially '4D-balloon' model based on new, early-galaxy data. We may expect that in the near future, we will be able to see the future of the cosmos and our existence on it more clearly.

NATURE KNOWS: SYSTEMIC INTELLIGENCE, SELF-ORGANIZATION, AND THE REORIENTATION OF SCIENTIFIC PARADIGMS

Kervalishvili P.
Georgian Academy of Natural Sciences (www.gans.ge)
https://doi.org/10.5281/zenodo.18840588

Abstract

Modern technological civilization increasingly frames its relationship with the natural world in terms of mastery and control. In the contemporary era of accelerating technological advancement, humanity increasingly positions itself as master and architect of the natural world. Yet beneath this narrative of control lies a more profound recognition: natural systems embody a form of systemic intelligence forged through 3.8 billion years of evolutionary experimentation. This paper develops the conceptual thesis encapsulated in the phrase "Nature Knows." It argues that nature's knowledge is not mystical but systemic—embedded in processes of self-organization, circularity, resilience, and adaptive optimization across scales from quantum structures to ecosystems. By examining biological networks, thermodynamic principles, evolutionary optimization, and biomimetic innovation, the paper proposes a reorientation of scientific paradigms—from domination toward dialogue, from linear extraction toward circular integration, and from reductionist fragmentation toward systemic coherence. Such a transition is not anti-technological; rather, it calls for a deeper science aligned with the structural grammar of natural intelligence. The future of sustainable civilization depends on our capacity to recognize, interpret, and cooperate with the knowledge already encoded within natural systems.

Keywords: systemic intelligence, self-organization, biomimicry, resilience, circular economy, complexity theory, sustainability paradigm

1. Introduction: Beyond the Illusion of Conquest

Technological modernity often describes progress in the vocabulary of domination—rivers are "tamed," genes "edited," and planetary resources "extracted." Yet scientific advances in ecology and complexity reveal that humanity remains embedded within a networked biosphere governed by thermodynamic and evolutionary constraints (Capra 1996; Meadows 2008).

The proposition "Nature Knows" expresses the recognition that natural systems encode adaptive solutions refined by evolutionary selection. Evolution operates as a distributed optimization process without central command (Kauffman 1993). What appears as design is emergent order arising from local interactions governed by physical law and selective retention.

The question, therefore, is not whether humanity can manipulate nature, but whether it understands the systemic consequences of doing so. (Prigogine 1984; Ramsden, Kervalishvili 2008; Odum 1994; Holling 1973; Benyus 1997).

2. Evolution as Long-Term Optimization

Life on Earth represents approximately 3.8 billion years of iterative experimentation. Natural selection preserves configurations that maintain dynamic stability within changing boundary conditions (Maynard Smith 1982).

Biological structures exhibit remarkable efficiency:

- Fractal vascular systems minimize energy expenditure (West et al. 1997).
- Photosynthesis approaches near-quantum efficiency in energy transfer (Blankenship 2014).
- Ecosystem nutrient cycles close material loops, reducing waste accumulation (Odum 1994).

Unlike human engineering, which frequently externalizes entropy in the form of pollution, biological systems internalize and redistribute material outputs.

From a thermodynamic perspective, living systems maintain order by dissipating energy gradients (Schrödinger 1944; Prigogine and Stengers 1984, Kervalishvili, Michailidis 2012). Nature's "knowledge" thus reflects compliance with physical constraints while maximizing persistence.

3. Ecosystems as Distributed Intelligence Networks

Forests exemplify emergent coordination. Mycorrhizal networks enable nutrient transfer and stress signaling among trees, enhancing collective resilience (Simard et al. 2014). No central controller directs this exchange. Order arises through decentralized feedback loops. Fig.1).

This phenomenon corresponds to principles of self-organization in complex systems (Nicolis and Prigogine 1977). Resilience theory further demonstrates that ecosystems maintain functionality through adaptive cycles of growth, collapse, and renewal (Holling 1973).

Key properties of ecosystem intelligence include:

- Redundancy and modularity
- Nonlinear feedback mechanisms
- Adaptive self-repair
- Circular nutrient flows

Such characteristics contrast with linear industrial systems, which often lack redundancy and exhibit systemic fragility (Meadows 2008).

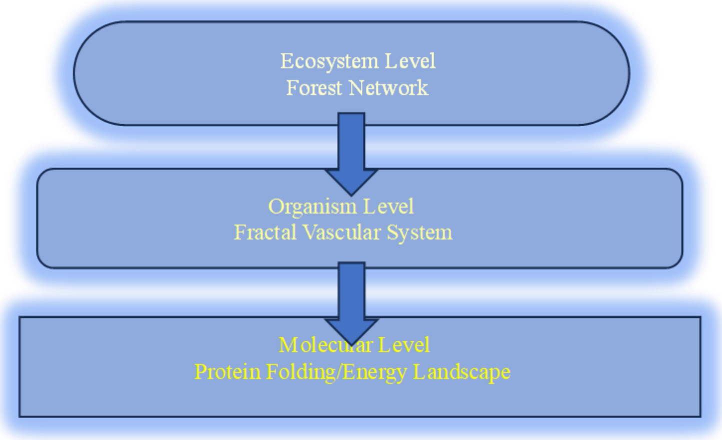


Figure 1. Multi-scale representation of systemic intelligence in natural systems. From molecular self-organization to ecosystem-level nutrient exchange, order emerges through decentralized interactions governed by thermodynamic constraints and evolutionary optimization.

4. Scale Resonance: Patterns Across Micro and Macro Realms

Nature exhibits recurring mathematical structures across scales:

- Fractal geometry in river basins and vascular networks (Mandelbrot 1982).
- Spiral dynamics in galaxies and biological growth.
- Symmetry-breaking phenomena in phase transitions (Anderson 1972).

Self-organization emerges when energy flows through matter under constraint. Whether in galactic clustering or protein folding, systems evolve toward locally stable configurations (Prigogine and Stengers 1984).

The recurrence of scale-invariant structures suggests that nature operates through universal organizing principles grounded in symmetry, conservation, and minimization of free energy.

5. Linear Economies and Thermodynamic Limits

Industrial civilization largely follows a linear throughput model:

Extract → Produce → Consume → Dispose

This model ignores the second law of thermodynamics, exporting entropy into ecological systems. By contrast, ecosystems operate cyclically, reintegrating matter into regenerative loops (Odum 1994).

The circular economy concept attempts to emulate biological metabolism (Ellen MacArthur Foundation 2013). However, circularity is not merely ethical—it is structurally necessary for long-term viability within finite planetary boundaries (Rockström et al. 2009).

Nature's intelligence eliminates the concept of waste. Human sustainability efforts must converge toward similar closure of material cycles. (Fig.2).

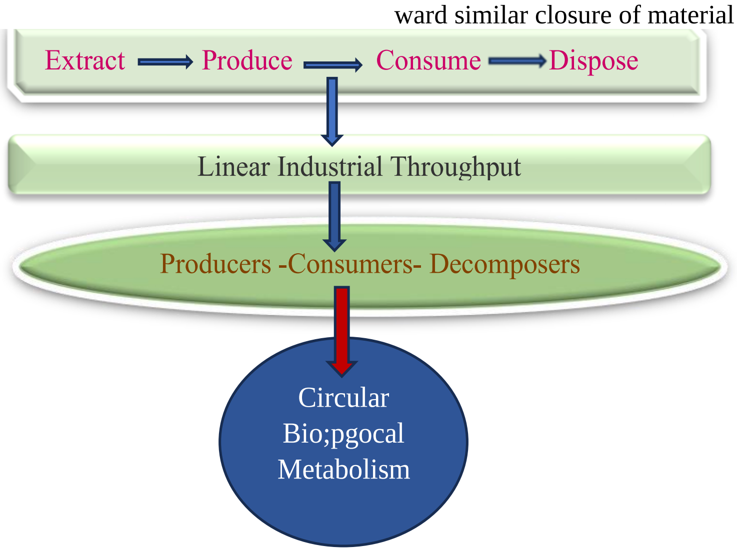


Figure 2 Comparison between linear industrial throughput systems and circular biological metabolism. Biological systems internalize material outputs into regenerative cycles, enhancing long-term stability.

6. Biomimicry as Epistemological Shift

Biomimicry formalizes the practice of consulting nature as mentor rather than resource (Benyus 1997). Engineering innovations inspired by biological systems include:

- Gecko-inspired dry adhesion systems.

- Termite mound–based passive building ventilation.
- Slime mold–inspired transport network optimization.

Such approaches illustrate that technological advancement need not oppose ecological coherence. Instead, it can extend evolutionary principles into human design domains.

Biomimicry represents a shift:

From extraction → to emulation
From control → to partnership
From fragmentation → to systems integration

7. Humility in the Anthropocene

Humanity's technological capacity—nuclear physics, synthetic biology, artificial intelligence—has reached planetary impact scale. Yet climate instability and biodiversity loss reveal limits of reductionist intervention (Rockström et al. 2009).

Complex systems exhibit nonlinear thresholds and tipping points (Scheffer 2009). Interventions based solely on short-term optimization risk destabilizing larger systemic equilibria.

Recognizing that "Nature Knows" implies epistemic humility: knowledge of parts does not guarantee mastery of wholes.

8. Toward a Dialogue-Based Scientific Paradigm

A reoriented paradigm integrates:

1. Systems thinking (Capra 1996)
2. Thermodynamic realism (Schrödinger 1944)
3. Evolutionary optimization (Kauffman 1993)
4. Resilience frameworks (Holling 1973)

Science must evolve from mechanistic fragmentation toward relational integration. (Fig,3).

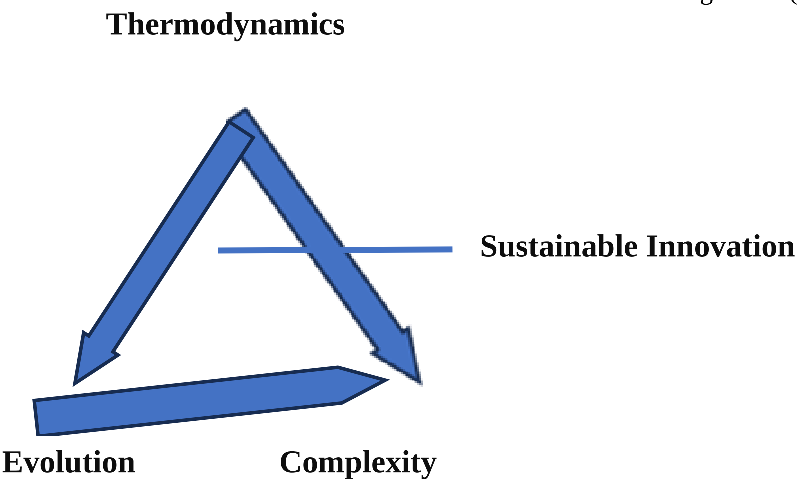


Figure 3. Conceptual framework for a dialogue-based scientific paradigm integrating thermodynamic realism, evolutionary optimization, and complexity theory to guide sustainable technological development.

Human intelligence is not separate from nature—it is an emergent property of natural evolution. Therefore, the future of innovation lies not in transcending nature, but in deepening alignment with its organizing principles.

9. Conclusion

"Nature Knows" encapsulates a scientific recognition: natural systems encode structural intelligence shaped by physical law and evolutionary refinement. Sustainable civilization depends on interpreting and cooperating with these principles rather than attempting unilateral domination.

Listening to nature is not romanticism. It is advanced systems science.

References:

1. Anderson PW (1972) More is different. Science 177:393–396.
2. Benyus JM (1997) Biomimicry: Innovation inspired by nature. HarperCollins, New York.
3. Blankenship RE (2014) Molecular mechanisms of photosynthesis. Wiley, Oxford.
4. Capra F (1996) The web of life. Anchor Books, New York.
5. Ellen MacArthur Foundation (2013) Towards the circular economy.
6. Holling CS (1973) Resilience and stability of ecological systems. Annu Rev Ecol Syst 4:1–23.
7. Kauffman SA (1993) The origins of order. Oxford University Press, Oxford.
8. Kervalishvili P. Michailidis S. (2012), Philosophy and synergy of information. NATO Science series, IOS press, v.93.
9. Mandelbrot BB (1982) The fractal geometry of nature. Freeman, New York.
10. Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press.
11. Meadows DH (2008) Thinking in systems. Chelsea Green Publishing.
12. Nicolis G, Prigogine I (1977) Self-organization in nonequilibrium systems.
13. Wiley Odum HT (1994) Ecological and general systems. University Press of Colorado.
14. Prigogine I, Stengers I. (1984) Order out of chaos. Bantam Books.
15. Ramsden J. Kervalishvili P. (2008) Complexity and Security. NATO Science Series. V.37 IOS Press.
16. Rockström J et al (2009) A safe operating space for humanity. Nature 461:472–475.
17. Scheffer M (2009) Critical transitions in nature and society. Princeton University Press.

18. Schrödinger E (1944) What is life? Cambridge University Press.
19. Simard SW et al (2014) Mycorrhizal networks and forest resilience. New Phytol 201:1096–1104.
20. West GB, Brown JH, Enquist BJ (1997) A general model for allometric scaling laws. Science 276:122–126.

TECHNICAL SCIENCES

INTEGRATION OF AN AMMONIA SENSOR INTO VIDEO ENDOSCOPY FOR EARLY DETECTION OF HELICOBACTER PYLORI INFECTION

Sabirov R.
PhD candidate in Medical Technology Engineering
Azerbaijan State Oil and Industry University, Azerbaijan, Baku.
ORCID: 0009-0008-5436-6047
https://doi.org/10.5281/zenodo.18840590

Abstract

The early diagnosis of gastrointestinal pathologies is of great importance in clinical practice. Currently, endoscopic diagnosis is carried out using visual and invasive biopsy techniques. These methods have limitations in detecting early functional and metabolic changes. Particularly in Helicobacter pylori-associated gastritis, metabolic abnormalities develop before morphological changes. As a result, false-negative results can occur during the diagnosis of H. pylori infection.

This study proposes a novel endoscopic detection procedure using an ammonia (NH_3) sensor into endoscopic systems for the early and non-invasive detection of H. pylori infection. A scientific justification of the sensor selection is presented based on the clinical literature. A design of the distal tip of the endoscope, which facilitates real-time measurement of the concentration of ammonia, is also presented. It is ensured that the integration of the sensor does not affect the optical and mechanical functionality of the endoscope. The measured ammonia levels were normalized and classified into clinical risk categories, and the raw sensor data was converted into a digital diagnostic indicator. The proposed framework of methodology increases the level of objectivity of diagnostics, minimizes unnecessary random biopsies, and offers a scientific basis for the development of artificial intelligence-based multimodal endoscopic diagnostic systems.

Keywords: Ammonia sensor, video endoscopy, H. pylori infection.

Introduction

The early detection of gastrointestinal pathologies is a problem of vital importance in clinical practice, and one of the most common methods used for their detection is endoscopy. However, the traditional methods of endoscopy are based on the evaluation of the morphological characteristics of the mucosa using white light endoscopy (WLE). Besides these methods, advanced methods like Narrow Band Imaging (NBI) and Flexible Spectral Imaging Color Enhancement (FICE) are used for the precise evaluation of vascular structures [1, p. S16-20].

Although research has proven that the combined application of these imaging techniques can enhance the diagnostic accuracy of early-stage gastric cancer, the methods currently employed are based on subjective visual interpretation and often need verification through random or selective biopsy samples [2].

Early-stage pathological changes are not clearly visible and may lead to false negative results. The pathogenesis of gastrointestinal pathologies at an early stage, particularly in the context of Helicobacter pylori-associated gastritis, is closely related to complex metabolic and biochemical changes occurring in the mucosa. As studies have shown, H. pylori infection alters the composition and metabolic activity of the gastric microbiota, leading to the formation of specific metabolic profiles that maintain chronic inflammation. Although thermal and metabolic imbalances are observed at the tissue level at this stage, structural and morphological changes generally do not reach a clinically significant level [3].

At this stage, one of the major limitations associated with endoscopic diagnosis is related to the fact that functional and metabolic changes develop prior to structural changes. Current imaging technology used in endoscopic procedures cannot directly detect increased levels of metabolic activity, biochemical changes, and imbalances in functions. As indicated in the literature, despite increased levels of metabolic activity at the tissue level in early stages of inflammatory and precancerous conditions, these conditions are not clearly visible in images at the macroscopic level and depend on subjective interpretation by the operator [4].

To summarize, pathological processes such as high-grade inflammation, dysplasia, and early neoplastic transformation are characterized by functional and metabolic abnormalities. However, there are certain limitations of the present endoscopic systems in the detection of functional abnormalities in real time. Hence, the process of diagnosis involves biopsy procedures, which increase the cost of the procedure, time taken for the procedure, and discomfort for the patients.

The purpose of the study

In modern endoscopic devices, the basis of the diagnostic process is the interpretation of images captured by the CCD and CMOS-based visual sensor. While the approach is successful in the detection of changes in the structure and morphology of the mucosa, the approach is lacking in the assessment of functional and metabolic disorders that occur in the early stages of inflammatory and infectious diseases. Specifically, in the early stages of gastritis associated with Helicobacter pylori infection, an increase in metabolic activity occurs before structural changes, there is an increase in metabolic activity before the development of morphological changes, and the diagnostic process based only on visual sensor data can lead to false negative results.

This situation highlights the need for additional sensor technologies that complement visual sensors in endoscopic platforms and can evaluate functional biomarkers in real time.

The goal of the present research is to design and scientifically verify a conceptual diagnostic method based on an ammonia (NH_3) sensor integrated into endoscopic devices for early detection of Helicobacter pylori infection. The study presents the selection of an ammonia sensor, a constructive design solution for its integration into the distal end of the endoscope, and a normalization model for NH_3 concentration based on normative ranges of clinical values. The suggested method is intended to create a methodological and technological basis for the use of functional biomarkers in endoscopic devices.

Modern Endoscopic Diagnostic Systems

Standard white light video endoscopy (WLE) is the most commonly used method in clinical practice. This method enables real-time imaging of the gastrointestinal mucosa using matrix sensors. Figure 1 shows the CCD (Charge-Coupled Device) elements of a typical video endoscope used in clinical practice.

Figure 1. Matrix sensors of video endoscope.

Ulcers, bleeding, and neoplastic lesions are very effective in the detection of macroscopic pathology. However, the diagnostic potential of WLE is basically restricted to the visual assessment of surface morphology and lacks sufficient sensitivity to identify early inflammatory, structural, and functional alterations. In particular, in the early phase of Helicobacter pylori infection and in the initial phase of low-grade dysplasia, the mucosal alterations are not visually specific, which may cause difficulties in the diagnosis and result in a potential false-negative result [5]. This limitation directly impacts to the problem of false negative results and leads to a greater reliance on random or protocol biopsy sampling, which is known to be prone to sampling error. Because biopsies are usually taken from visually suspicious regions, functionally active but visually normal regions may be overlooked. Clinical data suggest that a substantial number of precancerous gastric lesions and early inflammatory changes are overlooked during routine WLE evaluations, especially in patients with mild or patchy mucosal disease [6, p. 141-150]. Thus, WLE alone is not adequate for the early detection and risk stratification of gastric pathologies.

Narrow Band Imaging (NBI), Flexible Spectral Imaging Color Enhancement (FICE), and i-Scan spectral endoscopy use specific wavelengths of light to enhance the visualization of tissue and vascular structures. These techniques enable more accurate visualization of mucosal microarchitecture by taking into account the optical absorption of hemoglobin. Although spectral endoscopy is effective for the detection of dysplasia and neoplastic lesions, it does not offer any direct functional and biochemical information and is also operator-dependent [7].

Matrix Sensors (CCD/CMOS)

The technological basis of modern endoscopic imaging systems is semiconductor image sensors of the CCD (Charge-Coupled Device) and CMOS (Complementary Metal-Oxide-Semiconductor) type [8]. Figure 2 illustrates a macroscopic view of the active working surface and internal structural layer of a CCD image sensor.

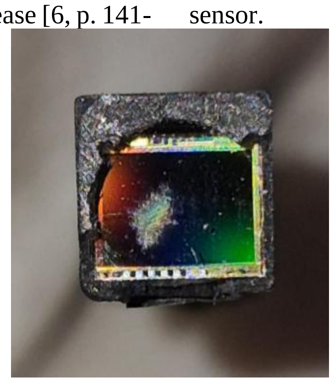


Figure 2. Appearance of the active semiconductor layer of the CCD image sensor

CCD (Charge-Coupled Device) and CMOS (Complementary Metal Oxide Semiconductor) sensors are used in endoscopic systems for the optical imaging of the structural and morphological characteristics of the mucosa in endoscopic systems. Although CCD sensors are characterized by a high signal-to-noise ratio high homogeneity of the images, their high power consumption is a disadvantage, and the sequential nature of the readout process can be a limitation for real-time processing. On the other hand, the advantages of the use of CMOS technology, including faster data transfer, low power consumption and compact system design because they have integrated readout electronics at each pixel level. For this reason, the vast majority of modern video endoscopes use CMOS-based image sensors [8].

However, the fundamental limitation of both sensor technologies is that they operate on the basis of optical information only. Both sensors are incapable of detecting biochemical and metabolic changes within the tissue, and the early stages of inflammatory processes and disorders of function, related to Helicobacter pylori, often lack visual specificity.

Therefore, though CCD and CMOS-based imaging systems have proven to be effective for detecting structural changes, there is a limitation with regard to obtaining functional and metabolic information. This creates a scientific and technological rationale for the integration of other functional sensors, particularly those with the capability of detecting real-time metabolic biomarkers like ammonia.

Thus, the use of ammonia sensors in endoscopic devices appears to hold promise for the early non-invasive detection of Helicobacter pylori infection. The increase in the local concentration of ammonia in the gastric lumen, which is caused by the high urease activity of the bacteria, can act as an indicator of the presence of Helicobacter pylori [9]. Detecting metabolic changes in the early stages may enable earlier diagnosis of infection, allowing diagnostic decisions to be made at an earlier stage.

In addition, determination of the levels of ammonia not only indicates the presence or absence of pathogens, but it also offers supplementary information regarding the functional condition of the tissue and metabolic responses to pathological processes. Obtaining this functional information in real time can enhance the informative value of endoscopic examination and partially eliminate the subjectivity of visual assessment. Early detection of functional conditions reduces the need for random tissue sampling, minimizes the overall procedure time, and reduces patient stress from an invasive procedure. All the above-mentioned advantages allow us to evaluate the use of ammonia sensors in endoscopic diagnostic systems as a viable and scientifically justified approach. Therefore, inclusion of ammonia sensors into endoscopic systems is a promising approach for early, accurate, and functional diagnosis and solves one of the main problem areas of the research.

Methods

Various clinical studies have shown that ammonia concentrations in gastric juice are high and that this is harmful in cases of H. pylori infection. It has been found that the concentration of ammonia in patients infected with H. pylori is significantly higher than that in patients who are not infected (5.58 ± 2.69 vs. 2.00 ± 1.49 μmol/L). These indicators have been statistically correlated with H. pylori density, gastritis activity, and gastric pH, demonstrating the effectiveness of the ammonia measurement device [10, p. 205].

Similarly, other studies have also shown that ammonia concentration in gastric juice is significantly increased in H. pylori-positive patients and that this indicator can be used as a diagnostic biomarker for infection. This confirms the pathophysiological importance of ammonia production.

The selection of the ammonia sensor in endoscopic studies was based on scientific criteria grounded in clinical pathophysiology and technical measurement requirements. As shown in the literature, during Helicobacter pylori infection, local ammonia production occurs in the gastric mucosa as a result of the bacterium's high urease activity, and this condition is observed with significantly higher NH_3 concentrations compared to uninfected cases. Clinical and experimental studies have shown that ammonia levels in the gastric environment increase during infection and can vary between approximately 10-100 ppm; this makes ammonia concentration an important functional and biochemical diagnostic marker [10, p. 205].

For this reason, an electrochemical ammonia sensor that is able to measure ammonia concentration at low levels was chosen for the study. The capability of the ammonia sensor to detect ammonia increases at the 1 ppm level enables the early detection of the increase in ammonia levels due to bacterial metabolism.

Sensor design

This section will present the proposed design of the distal tip for the sensor application in the integration of the sensor with the endoscopic system, together with the scientific basis of the sensor application. The major objective of the proposed sensor integration with the endoscope system is to develop an endoscope system that can perform real-time, minimally invasive, and non-invasive measurements of the local concentration of the ammonia present in the gastric mucosa, while the imaging capabilities of the endoscope system will not be compromised. Figure 3 shows the proposed design of the distal tip for the integration of the sensor with the endoscope system.

The main aim of the suggested integration of sensors is based on the development of an endoscopic diagnostic system that facilitates the real-time, minimally invasive, and non-invasive measurement of the local concentration of ammonia (NH_3) in the gastric mucosa. The acceptance of NH_3 as a metabolic biomarker of H. pylori infection is the scientific basis for the suggested sensor-based system [11, p. 51-57].

Hence, the increase in the level of ammonia in the gastric environment is detected due to the high urease activity of H. pylori, thereby providing important diagnostic information for the early detection of H. pylori infection.

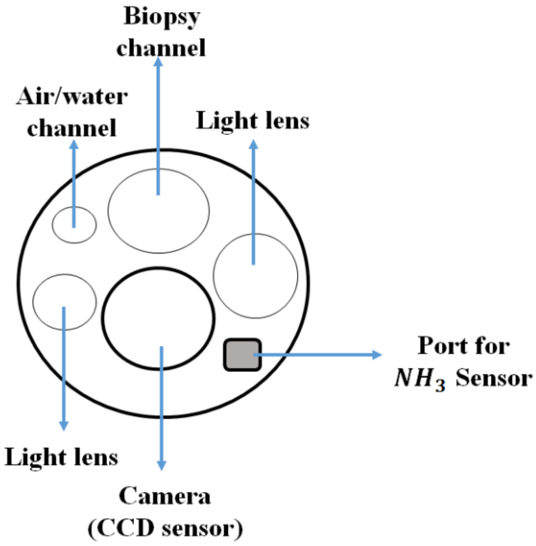


Figure 3. Conceptual diagram of a modified video endoscope distal tip for multi-mode sensor integration.

In the proposed design, the integration of the NH_3 sensor into the distal end of the endoscope, with minimal structural changes made to the structure of the biopsy channel light fibers and air/water channel, does not have any adverse effect on the optical, mechanical, and clinical functionality of the device. This ensures that the sensor can be used in real-time in a clinical environment and provides a methodological framework that can be used to evaluate quantitatively the results obtained from the sensor. Therefore, for the NH_3 concentration to be clinically meaningful, it is considered necessary to normalize the sensor outputs according to physiological ranges. According to clinical and experimental studies, ammonia concentrations in a healthy stomach environment are generally <10 ppm and range from 20-100 ppm during Helicobacter pylori infection [9]. Therefore, signals obtained from the NH_3 sensor were converted to ppm units and processed according to the normalization model presented below.

$$(NH_3)norm = \frac{NH_3 - C_{min}}{C_{max} - C_{min}} \quad (1)$$

Where,

NH_3 - measured ammonia concentration,

C_{max} and C_{min} - are the minimum and maximum physiological values based on medical literature.

Two different physiological concentration ranges were taken into account for normalising NH_3 concentration measurements. In the first scenario, a range of 10 to 100 ppm was chosen to represent a wide range of clinical concentration levels. In the second scenario, a range of 5 to 90 ppm was chosen to demonstrate the impact of low concentration changes.

In this study, it was accepted that the minimum and maximum physiological limits to be used in normalising the results of measurement from ammonia (NH_3) sensor would be 10 ppm and 100 ppm, respectively. These values were based on the results of clinical and experimental studies concerning basal levels of ammonia in a healthy stomach environment and high metabolic activity due to Helicobacter pylori infection [10, p. 205].

Table 1

Normalized NH_3 concentration and corresponding risk category based on reference range of 10 to 100 ppm.

NH_3 ppm	NH_3 *norm*	Risk Category
15	0.05	Low
20	0.11	Low
30	0.22	Suspicious
40	0.33	Suspicious
45	0.38	Suspicious
50	0.44	Suspicious
65	0.61	High
70	0.66	High
90	0.88	High

In an alternate normalisation scenario, this study set the minimum and maximum limits for ammonia concentration at 5 ppm and 90 ppm, respectively. This is with the aim of increasing sensitivity to small changes in ammonia levels and improving the assessment of early functional and metabolic changes in the gastric environment.

Table 2

Normalized NH_3 concentration and corresponding risk category based on reference range of 5 to 90 ppm.

NH3 ppm	NH3 norm	Risk Category
15	0.12	Low
20	0.18	Low
30	0.3	Suspicious
40	0.4	Suspicious
45	0.47	Suspicious
50	0.53	High
65	0.7	High
70	0.76	High
90	1	High

Figure 4 shows a comparison chart for the normalised indices calculated using two different physiological normalisation ranges ($C_{min} = 10$ $C_{max} = 100$ $C_{min} = 5$ $C_{max} = 90$) for the measurement of the ammonia (NH_3).

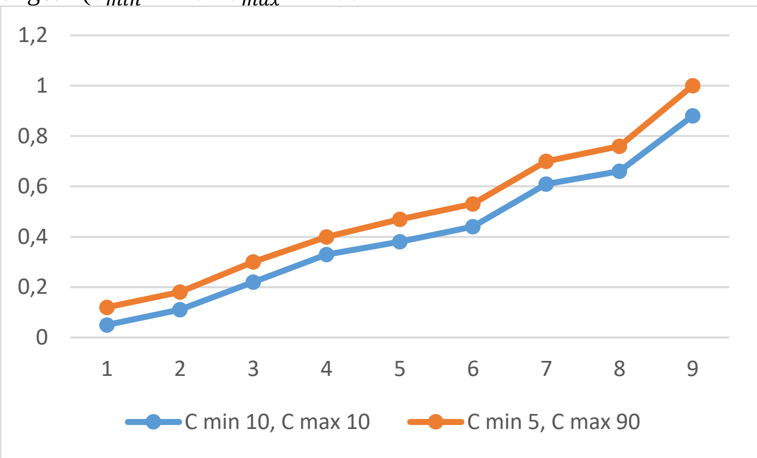


Figure 4. Comparative analysis of NH_3 measurements normalised at different minimum and maximum limits.

The ammonia concentration measured was normalized by a standard min-max scaling method to fall within the physiological range of [0,1] units, which is a common practice in biomedical signal processing and machine learning-based analysis.

Normalized NH_3 indicators are divided into clinical categories as follows:

- When the normalized ammonia level $(NH_3)norm < 0.2$, the probability of Helicobacter pylori infection is considered low.
- For, $0.2 \leq (NH_3)norm < 0.5$, the result is classified as suspicious, and additional confirmatory diagnostic procedures are recommended.
- When $(NH_3)norm \geq 0.5$, a high probability of H. pylori infection is assumed.

The above method of normalization makes it easy for the interpretation of raw values obtained from the NH_3 sensor and allows for the development of a single index that can be used for comparison in various patients and measurement conditions. In addition, the division of the normalized NH_3 index into clinical categories allows for additional information that can be used in the physician's process of making decisions in real time while carrying out an examination. The advantage of using this method is that it eliminates the possibility of missing out on H. pylori infection, especially in cases that are difficult to detect and where biopsy is not feasible.

Conclusion

The proposed methodological framework enables NH_3 measurements to be used not only as a quantitative indicator but also as a digital diagnostic indicator. The normalized sensor data is combined with multimodal endoscopic data such as standard white light images, NBI images, thermal images, or pH sensor data in the next step of the process, which can be used as suitable input data for complex analysis. This approach provides functional information input to the system to overcome the limitations of decision-making based solely on visual information in endoscopic diagnosis.

In the next phase of research, it is planned that these normalized NH_3 indicators will be integrated into artificial intelligence-based analysis. Specifically, machine learning and deep learning models that consider the temporal and spatial variation of sensor data will be applied to automatically detect H. pylori infection, assess the risk level, and recommend the most suitable areas for biopsy. This provides a scientific basis for developing smart diagnostic systems that can enhance the objectivity of endoscopic diagnosis and reduce random biopsies while allowing physician control.

References:

1. Muto M. et al. Narrow band imaging: a new diagnostic approach to visualize angiogenesis in superficial neoplasia //Clinical Gastroenterology and Hepatology. – 2005. – Т. 3. – №. 7. – С. S16-S20. https://doi.org/10.1016/S1542-3565(05)00262-4

2. Liu Z. et al. Diagnostic performance of white light endoscopy, narrow band imaging, and flexible spectral imaging color enhancement for early gastric cancer: a systematic review and meta-analysis //Quantitative Imaging in Medicine and Surgery. – 2025. – T. 15. – №. 6. – C. 5534. https://doi.org/10.21037/qims-24-482

3. Zang H. et al. Metabolic alterations in patients with Helicobacter pylori-related gastritis: The H. pylori-gut microbiota-metabolism axis in progression of the chronic inflammation in the gastric mucosa //Helicobacter. – 2023. – T. 28. – №. 4. – C. e12984. https://doi.org/10.1111/hel.12984

4. Gopakumar H. et al. Role of advanced gastrointestinal endoscopy in the comprehensive management of neuroendocrine neoplasms //Cancers. – 2023. – T. 15. – №. 16. – C. 4175. https://doi.org/10.3390/cancers15164175

5. Li H. et al. Advanced endoscopic methods in gastrointestinal diseases: a systematic review //Quantitative Imaging in Medicine and Surgery. – 2019. – T. 9. – №. 5. – C. 905. https://doi.org/10.21037/qims.2019.05.16

6. Ikenoyama Y. et al. Detecting early gastric cancer: Comparison between the diagnostic ability of convolutional neural networks and endoscopists //Digestive Endoscopy. – 2021. – T. 33. – №. 1. – C. 141-150. https://doi.org/10.1111/den.13688

7. Bidani K. et al. Optical Imaging Technologies and Clinical Applications in Gastrointestinal Endoscopy //Diagnostics. – 2025. – T. 15. – №. 20. – C. 2625. https://doi.org/10.3390/diagnostics15202625

8. Boese A. et al. Endoscopic imaging technology today //Diagnostics. – 2022. – T. 12. – №. 5. – C. 1262. https://doi.org/10.3390/diagnostics12051262

9. Popa C., Petrus M., Bratu A. M. Ammonia measurement in human breath of subjects with Helicobacter pylori using photoacoustic spectroscopy //Journal of biophotonics. – 2024. – T. 17. – №. 6. – C. e202400074. https://doi.org/10.1002/jbio.202400074

10. Lee O. J., Lee E. J., Kim H. J. Correlations among gastric juice pH and ammonia, Helicobacter pylori infection and gastric mucosal histology //The Korean journal of internal medicine. – 2004. – T. 19. – №. 4. – C. 205. https://doi.org/10.3904/kjim.2004.19.4.205

11. Graham D. Y., Miftahussurur M. Helicobacter pylori urease for diagnosis of Helicobacter pylori infection: A mini review //Journal of advanced research. – 2018. – T. 13. – C. 51-57. https://doi.org/10.1016/j.jare.2018.01.006

№175/2026

Norwegian Journal of development of the International Science

ISSN 3453-9875

It was established in November 2016 with support from the Norwegian Academy of Science.

DESCRIPTION

The Norwegian Journal of Development of the International Science, established in 2016 with support from the Norwegian Academy of Science, offers a platform for global researchers, academics, and students. Publishing 24 times annually, it covers current science topics. Submitting your article provides exposure to a diverse audience, enhancing your career. Our commitment to research and SEO ensures timely visibility. Submit to the Norwegian Journal of Development of the International Science today and join esteemed authors.

1000 copies

Norwegian Journal of development of the International Science

Iduns gate 4A, 0178, Oslo, Norway

email: publish@nor-ijournal.com

site: https://nor-ijournal.com

nor-ijournal.com